Robust and Efficient Corner Detection for Computer Vision

Mingzhe Wang

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Computer Science and Engineering

Faculty of Engineering

The University of New South Wales

January 2022

Thesis Title and Abstract

Thesis Title

Robust and Efficient Corner Detection for Computer Vision

Thesis Abstract

Corner detection is a fundamental computer vision problem that has been widely studied in image retrieval, object tracking, motion estimation, visual localization and 3D reconstruction. The accuracy and repeatability of corner detection are important for image matching and retrieval, while the detection accuracy and localisation ability are vital for visual localisation, motion estimation and 3D reconstruction. For real-time computer vision tasks, it is important to achieve fast corner detection with high detection accuracy and repeatability.

This thesis addresses the problem of corner detection and localisation with high accuracy and repeatability. In order to improve the detection performance of the current state-of-the-art in corner detection, an improved shearlet transform and a novel complex shearlet transform are proposed to overcome the problems of traditional shearlets and achieve better localisation of distributed discontinuities, particularly with the ability to extract phase information from geometrical features. Moreover, a multi-directional structure tensor and a multi-scale corner measurement function are proposed to make full use of the structural information from the improved shearlets for detection, and a new rotary phase congruence tensor is proposed to utilize all amplitude and phase information of the complex shearlets for detection. As a result, two new corner detectors are proposed. Experimental results demonstrate that their localisation ability and detection accuracy are superior to current detectors, and their repeatability is generally higher than current corner detectors as well as recent deep learning based interest point detectors. Therefore, they have great potential for applications in computer vision tasks.

To meet the needs of real-time computer vision tasks, especially real-time portable tasks, a new type of filter that can enhance corners and suppress edges as well as noise simultaneously is proposed to simplify the detection architecture and improve its parallel computing performance, and a novel corner detector with high computational efficiency is proposed. The corresponding field programmable gate array (FPGA) design is provided as well. Experimental results show that, with very low computational cost and simple architecture, the proposed detector can achieve similar detection accuracy and repeatability of current corner detectors while being potentially useful for real-time computer vision applications.

Originality, Copyright and Authenticity Statements

ORIGINALITY STATEMENT

☑ I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

COPYRIGHT STATEMENT

☑ I hereby grant the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).

For any substantial portions of copyright material used in this thesis, written permission for use has been obtained, or the copyright material is removed from the final public version of the thesis.

AUTHENTICITY STATEMENT

☑ I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.

Inclusion of Publications Statement

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in the candidate's thesis in lieu of a Chapter provided:

- The candidate contributed **greater than 50%** of the content in the publication and are the "primary author", i.e. they were responsible primarily for the planning, execution and preparation of the work for publication.
- The candidate has obtained approval to include the publication in their thesis in lieu of a Chapter from their Supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis.

☑ The candidate has declared that **their thesis has publications - either published or submitted for publication - incorporated into it in lieu of a Chapter/s. Details of these publications are provided below.**.

Publication Details #1

Full Title:	Corner Detection Based on Shearlet Transform and Multi-directional Structure Tensor
Authors:	Mingzhe Wang, Weichuan Zhang, Changming Sun, Arcot Sowmya
Journal or Book Name:	Pattern Recognition
Volume/Page Numbers:	Volume 103, Pages 107299: 1-15
Date Accepted/Published:	July 2020
Status:	published
The Candidate's Contribution to the Work:	Primary author: methodology design, algorithm design, programming, carrying out experiments, and writing paper.
Location of the work in the thesis and/or how the work is incorporated in the thesis:	The work is incorporated in the thesis in Chapter 3.

Publication Details #2

Full Title:	Complex Shearlets and Rotary Phase Congruence Tensor for Corner Detection
Authors:	Mingzhe Wang, Changming Sun, Arcot Sowmya
Journal or Book Name:	Pattern Recognition
Volume/Page Numbers:	
Date Accepted/Published:	
Status:	submitted
The Candidate's Contribution to the Work:	Primary author: methodology design, algorithm design, programming, carrying out experiments, and writing paper.
Location of the work in the thesis and/or how the work is incorporated in the thesis:	The work is incorporated in the thesis in Chapter 4.

Publication Details #3

Full Title:	Efficient Corner Detection Based on Corner Enhancement Filters
Authors:	Mingzhe Wang, Changming Sun, Arcot Sowmya
Journal or Book Name:	Digital Signal Processing
Volume/Page Numbers:	Pages 103364: 1-13
Date Accepted/Published:	Accepted, Dec. 2021
Status:	accepted
The Candidate's Contribution to the Work:	Primary author: methodology design, algorithm design, programming, carrying out experiments, and writing paper.
Location of the work in the thesis and/or how the work is incorporated in the thesis:	The work is incorporated in the thesis in Chapter 5.

Candidate's Declaration

I confirm that where I have used a publication in lieu of a chapter, the listed publication(s) above meet(s) the requirements to be included in the thesis. I also declare that I have complied with the Thesis Examination Procedure.

Abstract

Corner detection is a fundamental computer vision problem that has been widely studied in image retrieval, object tracking, motion estimation, visual localization and 3D reconstruction. The accuracy and repeatability of corner detection are important for image matching and retrieval, while the detection accuracy and localisation ability are vital for visual localisation, motion estimation and 3D reconstruction. For real-time computer vision tasks, it is important to achieve fast corner detection with high detection accuracy and repeatability.

This thesis addresses the problem of corner detection and localisation with high accuracy and repeatability. In order to improve the detection performance of the current state-of-the-art in corner detection, an improved shearlet transform and a novel complex shearlet transform are proposed to overcome the problems of traditional shearlets and achieve better localisation of distributed discontinuities, particularly with the ability to extract phase information from geometrical features. Moreover, a multi-directional structure tensor and a multi-scale corner measurement function are proposed to make full use of the structural information from the improved shearlets for detection, and a new rotary phase congruence tensor is proposed to utilize all amplitude and phase information of the complex shearlets for detection. As a result, two new corner detectors are proposed. Experimental results demonstrate that their localisation ability and detection accuracy are superior to current detec-

tors, and their repeatability is generally higher than current corner detectors as well as recent deep learning based interest point detectors. Therefore, they have great potential for applications in computer vision tasks.

To meet the needs of real-time computer vision tasks, especially real-time portable tasks, a new type of filter that can enhance corners and suppress edges as well as noise simultaneously is proposed to simplify the detection architecture and improve its parallel computing performance, and a novel corner detector with high computational efficiency is proposed. The corresponding field programmable gate array (FPGA) design is provided as well. Experimental results show that, with very low computational cost and simple architecture, the proposed detector can achieve similar detection accuracy and repeatability of current corner detectors while being potentially useful for real-time computer vision applications.

Acknowledgment

During my Ph.D. study and research, I have received great inspiration and support from many wonderful individuals. On the occasion of the completion of the thesis, I would like to express my sincere thanks to UNSW, CSIRO Data61, my supervisors, my parents, and my friends. Thanks and appreciation to all of them for being part of this journey.

First of all, I would like to thank UNSW and CSIRO Data61. Their excellent experimental and office environment, as well as excellent support for scientific research and daily life, help me engage my research smoothly and enrich my academic skills. These allow me to concentrate on my research and study, and provide me convenience for my living. I would like to express my appreciation for the China Scholarship Council scholarship (No. 201704910811), UNSW Tuition Fee Scholarship, and CSIRO Data61 Top-up Scholarship. Their financial support allows me to have no worries in my living and scientific research during my Ph.D., and let me have a pleasant four years living in Australia.

From selecting the research topic to collecting resources, from algorithm designing to programming, from writing manuscripts to revisions of papers, I have received a lot of help from my supervisors. I would like to thank Dr. Changming Sun and Prof. Arcot Sowmya for their instructive and selfless help. At every stage of

my research, their careful guidance helps me solve many difficulties and trains my abilities in independent research and thesis writing. I would also like to thank my lab mates and colleagues at CSIRO Data61, Dr. Dadong Wang, Dr. Weichuan Zhang, Guoqing Wang, Qiangguo Jin, Liu Yang, and Qiyu Liao. Their advice, support, and friendship are a great help to my living and research.

Thanks to my parents for giving me love, care, and understanding in my life. They give me the courage to challenge difficulties and the motivation to improve myself. Their support and help allow me to devote myself to my research work and make my life more colorful. I would also like to thank my friend at UNSW, their friendship gives happiness to me and makes me not feel lonely in a foreign country.

In the days to come, I will work hard to improve my professional quality and knowledge skills, and contribute my strength to the development of computer science.

Mingzhe Wang
January 19, 2022

Publications

List of Publications

- M. Wang, W. Zhang, C. Sun, and A. Sowmya, "Corner Detection Based on Shearlet Transform and Multi-directional Structure Tensor", *Pattern Recognition*, vol. 103, pp. 107299:1-15, 2020.
- M. Wang, C. Sun, A. Sowmya, "Complex Shearlets and Rotary Phase Congruence Tensor for Corner Detection", *Pattern Recognition*, under review.
- M. Wang, C. Sun, A. Sowmya, "Efficient Corner Detection Based on Corner Enhancement Filters", accepted by *Digital Signal Processing*, 2022, in press.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The detection and matching of local features in two or more images is a long-standing and fundamental problem in computer vision [1,2], and its solutions have been widely used in image retrieval [3], object matching [4], image stitching [5], parallel tracking and mapping for augmented reality [6, 7], object tracking [8], motion estimation [9], simultaneous localisation and mapping [10, 11], robot navigation [12, 13], image and object classification [14, 15] and stereo vision [16, 17]. Naturally, it has attracted tremendous attention in computer vision.

Local features in images may be roughly divided into two categories: local sparse features and local dense features. Local sparse features fall into the framework of detection first and then description. In the detection stage, keypoints such as corners or blob-like structures are detected, and then local sparse descriptors are extracted from the image patches around the keypoints. Local dense features forego the detection stage and descriptors are extracted directly from images using deep learning approaches. Unlike local sparse features, which are focussed on low-level image structures including corners and blobs, local dense features arise from close

attention to locally unique descriptors in feature maps generated by deep learning.

Local sparse feature detection is memory-efficient, and local features can be matched efficiently by using the nearest neighbours approach [18], which is important for applications such as image retrieval, object tracking and robot navigation. Moreover, keypoints especially corners can be accurately localised in images. This is vital for image stitching, visual localization, 3D reconstruction and motion estimation. In comparison, local dense feature extraction is memory-consuming, as is the matching process. Suffering from limited generalisation ability, dense feature extraction cannot reach the performance of local sparse feature extraction [19–21], however under some extreme imaging conditions such as severe illumination changes and extreme appearance variations or in weakly textured scenes, dense features outperform local sparse features in matching accuracy at the cost of higher computations and memory consumption. It has also been demonstrated that keypoint detection is a significant weakness and bottleneck in the overall performance of the sparse feature extraction. These are the motivations for the research in this thesis.

1.1 Background and Existing Methods

Corner detection as an area of study was pioneered by Moravec [22], who found an obvious property in the grayscale patterns of images. As a significant and stable local structure pattern in images, corners can provide an alternative choice for keypoints of local sparse features [23]. Corner detection has continued to attract great attention in computer vision and has been widely used as a critical pre-processing step in many image processing and computer vision tasks. The key performance indicators for corner detection are its detection accuracy, localisation ability, computational

efficiency and detection repeatability under image scaling, rotation, noise corruption, JPEG compression, illumination variations and viewpoint changes, amongst others. The importance of these indicators varies with the nature of the computer vision applications. For instance, detection repeatability is more important for image matching and retrieval, while detection accuracy and localisation ability are vital for visual localisation, motion estimation and 3D reconstruction. For real-time computer vision tasks, especially real-time portable tasks, it is important to achieve stable and accurate corner detection with a very low computational cost and simple architecture.

Currently, there are four types of corner detectors including derivative-based, contour-based, template-based and multi-scale analysis based corner detectors. Early work on corner detection focussed on image gradients. The classical approach was that of the Harris corner detector [24], in which a two-directional structure tensor was constructed with image derivatives, and then a 'corner score' was derived per image pixel from the structure tensor and compared with a specific threshold to detect corners. Subsequent research focussed mainly on two aspects. The first was the extraction of more stable derivatives from images, where the Hessian matrix [25], Hilbert transform [26], wavelets and Gabor wavelets [27–30] were considered. The second was performance improvement of the 'corner score' function [31–35]. Non-maximum suppression (NMS) [36] and adaptive parameterised functions [37,38] were used. Meanwhile, template-based methods have also been presented, which perform corner detection by matching parameterised templates with local image areas. In order to cover different types of corners such as T-, Y- and X-shaped, various templates were constructed with Gaussian kernels [39–42], curvature characteristics [43], parametric models [44], combined models [45], and colour models [46]. Representative detectors amongst template-based methods are SUSAN [47–52] and FAST [53, 54].

In order to significantly improve computational efficiency, only one circular template was adopted in the SUSAN detector and corner detection was performed by matching the circular template with local areas of images. In FAST [53], instead of the circular template, a ring template was used for the match and a decision tree was adopted to improve detection performance and avoid unnecessary comparisons [54], and the computational efficiency has been improved further. Due to its excellent computational efficiency, the FAST algorithm has been widely used in many machine vision tasks.

Derivative-based corner detectors generally perform corner detection by comparing a 'corner score' with a specific threshold. The 'corner score' is obtained by analysing a two-directional structure tensor constructed with the first or second-order derivatives of images. While template-based methods such as SUSAN and FAST detected corners by matching a specific template with local areas of images, they are also simple and computationally efficient. However, since derivative-based and template-based methods are both sensitive to noise, affine transformations and illumination variations amongst other factors, the detection repeatability of the two types of corner detectors is insufficient for many image matching and retrieval applications. Moreover, the parameterised corner templates cannot handle all types of corners that have different orientations and subtended angles, and the two-directional derivatives are not sufficient to accurately distinguish between corners and edges [55]. For example, Noble [56] analysed the principle behind the Harris corner detector and pointed out that the Harris detector can only detect 'L' shaped corners well. Therefore, the detection accuracy of derivative-based and template-based methods is often not high enough for visual localisation, motion estimation and 3D reconstruction tasks.

In order to improve corner detection accuracy, contour-based methods have also been

presented. These methods detect corners by analysing the shape characteristics of edges in images. Earlier works detected corners by using fixed-scale curvature calculations or polynomial fitting on chain codes generated by the contour maps of images [57–63]. Then, the curvature scale space method was adopted to achieve multi-scale local curvature estimation [64–66] through various scale space theories [67–69] and curvature estimation functions [70–72]. The chord-to-point distance [73] has also been used to improve the robustness to noise and local variations [74, 75]. In general, they are more accurate than the template-based and derivative-based corner detectors, and some of them also enable corner detection at different scales, which is important for visual localisation, motion estimation and 3D reconstruction applications. However, their performance greatly depends on the qualities of the extracted contours, which are relatively sensitive to noise, affine transformations, viewpoint changes and illumination variations, and their detection repeatability is insufficient for image matching and retrieval applications. Moreover, they are time-consuming and cannot meet computational efficiency requirements of real-time computer vision tasks. Recent works [76–81] on contour-based corner detection have attempted to extract more accurate and stable contour maps from images by using multi-scale and multi-directional analysis tools instead of the conventional edge extractors such as Canny [82] and LSD [83], or perform more accurate contour analysis using recent contour analysis methods such as the directional curvature field [81] and circumferential anchors [77].

Due to their promising detection performance, multi-scale analysis based detectors have attracted more attention recently. Early works namely scale-space detectors [84–89] extended the classical derivative-based detectors to multi-scale space to realise scale invariant or multi-scale corner detection. For example, Mikolajczyk and Schmid successfully embedded Laplacian of Gaussian into Harris and Hessian de-

tectors and developed two scale invariant corner detectors called Harris-Laplace [84] and Hessian-Laplace [85]. Subsequent works focussed on improvement of corner detection accuracy and repeatability by merging the multi-scale decomposition results for corner detection. Wavelets, lifting wavelets, and dual-tree complex wavelets [90] have been used for corner detection [87–89, 91]. Since two-dimensional wavelets have a limited capability for capturing directional information, Gabor wavelets [92], Log-Gabor wavelets [93] and anisotropic Gaussian derivative filters [94] were used for considering more directional information in corner detection. Natural images contain intrinsic geometrical structures that are important for corner detection. However, while wavelets are good at isolating the discontinuities at edge points, they cannot handle equally well distributed singularities such as those along curves [95, 96]. As a result, contourlets [97], ridgelets [98], curvelets [99] and shearlets [100] have been utlised to capture intrinsic geometrical structures and applied to corner detection [101–109].

In recent years, deep learning has been extended for interest point detection [110–124]. Unlike the corner detectors which focus on low-level image structures, recent deep learning-related interest point detectors pay close attention to locally unique descriptors in feature maps generated by deep learning. It is well known that corner detection belongs to the first stage of sparse local features in a framework of first detection and then description. In the detection stage, a set of interest points such as corners or blobs is obtained, and then local descriptors are extracted from the image patches around the interest points. On the contrary, recent deep learning-related interest point detectors follow the framework of first description and then detection [110–122]. They first generate a group of feature maps with deep learning, and then the interest points are detected from the maps. D2-Net [113] directly marked the central points of the local maximum descriptors as interest points, and

SuperPoint [112] performed interest point detection by using a decoder branch. Reinforced SP [111] retrained SuperPoint via reinforcement learning to improve performance. In general, interest points detected by deep learning are less accurate but have high repeatability which is important for image matching and retrieval tasks, while low-level image structures such as corners or blobs are more accurate and memory-efficient, and their accurate localisation ability is very useful for visual localisation, motion estimation and 3D reconstruction [125]. Consequently, the accurate detection of corners with high repeatability remains a core issue of computer vision.

1.2 Inadequacies of Existing Methods

Multi-scale analysis based detectors can provide promising detection performance, and their detection accuracy and repeatability, under image scaling, rotation, noise corruption, JPEG compression, illumination variations and viewpoint changes, are superior to other types of corner detectors as well as the recent deep learning-related interest point detectors. Their computational efficiency is also remarkably higher than that of contour-based detectors. However, there are still some weaknesses that limit extensive application. First, multi-scale analysis based detectors achieve superior detection accuracy and repeatability through the use of multi-scale information, however the introduction of information from rough scales and downsampling operations commonly used in multi-scale decomposition tools may dampen the corner localisation ability. Therefore their corner localisation ability may be lower than that of derivative-based and template-based detectors, which will degrade their application in visual localisation, motion estimation and 3D reconstruction. Second, compared with the Harris and FAST detectors, multi-scale analysis based detectors

require more computational time to deal with multi-scale information, therefore their computational efficiency is lower than that of the derivative-based and template-based detectors, which limits their applications in real-time computer vision tasks. Finally, although their detection accuracy and repeatability have been proven to be superior to other types of corner detectors as well as the recent deep learning-related interest point detectors, their detection accuracy and repeatability are also insufficient for accurate computer vision tasks. Therefore, improving their detection accuracy, repeatability and localisation ability simultaneously while reducing computational complexity remains a challenging problem. In addition, for real-time computer vision tasks especially real-time portable tasks, it is important to achieve stable and accurate corner detection with very low computational cost and simple architecture. Although the Harris and FAST detectors are simple and computationally efficient, their low detection accuracy and repeatability degrade the application level and scope of real-time computer vision tasks. Therefore, for real-time computer vision tasks especially real-time portable tasks, the design of a simple and computationally efficient corner detector with sufficient detection accuracy, repeatability and adaptability remains important for real-time corner detection.

There are two factors that seriously affect the performance of multi-scale analysis based corner detectors. One is the multi-scale analysis tool used, the other is the use of multi-scale and multi-directional information for corner detection. Up till now, Laplacian of Gaussian, wavelets, lifting wavelets, Gabor wavelets, Log-Gabor wavelets, anisotropic Gaussian derivative filters, contourlets and shearlet transform have been used for image decomposition. Among them, shearlets can capture intrinsic geometrical structures as well as anisotropic features from natural images, and also have the best performance due to some unique and excellent properties [100, 109]. However, there are still some weaknesses in the current shearlets.

First, shearlet filters are all real functions so they can only capture the amplitude information and neglect the phase information completely. Second, shearlets have a narrow bandwidth in the Fourier domain, making the corresponding spatial filters have a large envelope. In areas around edges, this leads to structural information being extended at the end of the edges, i.e., the problem of edge extension. Third, the fast decay of shearlets in the Fourier domain will result in a strong side-lobe effect in the spatial domain, which leads to large side-lobes around the locations where the grayscale values drastically change, thereby resulting in bilateral margin responses for prominent edges. Finally, the shearlet decomposition at different scales has different numbers of directions, and one direction of a coarse scale corresponds to two directions of the adjacent finer scale. This makes it inconvenient to merge multi-directional information from different scales for image analysis.

In addition to the multi-scale analysis tool used, another factor that seriously affects the performance of corner detectors is the use of multi-scale and multi-directional information for corner detection. Current multi-scale analysis-based corner detectors utilise the amplitude information of multi-scale decomposition results in three ways. The first is to select amplitude coefficients in the orthogonal direction of the maximum response at the fixed scale for corner detection [93, 108, 109]. In the second, coefficients in all directions at all scales are weighted and summed into vertical and horizontal directions to establish a two-directional structure tensor for corner detection [84,85,93]. In the third, a multi-directional structure tensor is constructed by using all directional information at the finest scale for corner detection [94].

Since the neglect of most directional information in the first approach will tarnish the distinction between corners and other pixels and make detectors sensitive to local variations and additive noise, such detectors have a limited detection accuracy and

repeatability. In the second method, the consideration of all directional information at all scales will greatly improve corner detection accuracy and reduce the effects of noise. However, the fixed direction of the structure tensor is sensitive to image rotation. Moreover, the introduction of directional information from rough scales will seriously reduce corner localisation accuracy. The multi-directional structure tensor in the third method is less sensitive to image rotation and additive noise, but only the directional information at the finest scale is used for detecting candidate corners, which makes this stage sensitive to image scaling and JPEG compression. Moreover, it needs more computation time to deal with the multi-directional structure tensor, so such detectors have the highest computational complexity of all multi-scale analysis based methods.

1.3 Contributions

Based on the detailed analysis of current multi-scale analysis based corner detectors, this thesis addresses the challenging problem of improving corner detection accuracy, repeatability and localisation ability simultaneously while reducing computational complexity as much as possible. It also addresses the urgent need to provide a simple and computationally efficient corner detector with sufficient detection accuracy, repeatability and adaptability to achieve real-time detection. The main contributions are summarised as follows:

1. Two novel multi-scale and multi-directional analysis tools, namely improved shearlet transform and complex shearlet transform, are proposed, which can capture both amplitude and phase information, effectively mitigate the weaknesses of traditional shearlets and extract clear and accurate multi-scale and

multi-directional structural information from images. The detailed structural information provided by the two proposed shearlet transforms can improve the corner detection and localisation accuracy, and the use of multi-directional and multi-scale information especially phase information can greatly improve the robustness of detection and the ability to distinguish corners from other pixels. Thus, they are potentially useful in the computer vision field.

2. Two corner detectors with two novel structure tensors that makes full use of multi-directional and multi-scale structural information are proposed, which can provide better ability to distinguish corners from other pixels with high robustness. Experiments show that the two proposed corner detectors present improvements to current state-of-the-art in corner detection as well as the recent deep learning interest point detectors on most key indicators including detection accuracy, localisation accuracy and repeatability under image scaling, rotation, noise corruption, JPEG compression, illumination changes and viewpoint changes.

3. A novel efficient corner detector with a simple architecture and high parallel computing characteristics is presented, which provides similar detection accuracy and repeatability as multi-scale analysis based detectors and has a good adaptability to achieve real-time detection. It is potentially useful as an efficient corner detector for computer vision applications, especially for portable real-time tasks.

4. The ideas of enhancing corners and suppressing edges at the same time by directly filtering raw images, and the rotary phase congruence structure tensor for properly merging all the amplitude and phase information into a 2×2 structure tensor with high stability for corner detection are novel and different

from the existing corner detectors, offering a new lead for corner detection research.

The remainder of this thesis is organised as follows:

1. Related works are briefly reviewed in Chapter 2. In this chapter, the literature review of the recent corner detector algorithms is given, which contains the classification of existing algorithms, the development trajectory of each type of detectors, as well as the analysis of advantages and disadvantages of each type of algorithms. Then the possible research directions for further improving the performance of corner detection are pointed out.

2. In Chapter 3, a novel corner detector based on improved shearlet transform and multi-directional structure tensor is proposed. To improve corner detection accuracy, repeatability and localisation ability simultaneously, and overcome the weaknesses of the traditional shearlets, a new shearlet transform has been proposed, which can provide better ability to extract clear and accurate multi-scale and multi-directional structural information for detection, with more flexibility for multi-scale decomposition. To make full use of the multi-scale and multi-directional information of the proposed shearlets for corner detection, a novel multi-directional structure tensor is constructed to detect candidate corners, and a multi-scale corner measurement function is proposed to refine candidate corners. Experimental results demonstrate that the proposed corner detection method outperforms the current corner detectors in corner detection and localisation accuracies, as well as repeatability under image blurring, viewpoint changes, affine transformations, illumination changes, and JPEG compression.

3. In Chapter 4, a corner detector based on complex shearlet transform and rotary phase congruence structure tensor is proposed. This work attempts to improve the detection accuracy, repeatability and localisation ability further while reducing computational complexity as much as possible. A novel shearlet transform namely complex shearlet transform is proposed. Different from traditional shearlets, complex shearlets have a greater ability to localise distributed discontinuities and capture the phase information of geometrical features. Moreover, a type of phase congruence function is introduced and its tolerance to noise and ability of corner localisation have been improved further. Then for corner detection, a novel rotary phase congruence structure tensor is proposed to properly merge the amplitude and phase information of all scales into a 2×2 structure tensor. With less sensitivity to image scaling, rotation, blurring and noise, the rotary phase congruence structure tensor can make full use of all amplitude and phase information of complex shearlets to detect corners. Experimental results demonstrate that the proposed method presents improvements to the current state-of-the-art in corner detection on corner localisation ability, detection accuracy and repeatability under image scaling, rotation, JPEG compression, noise corruption, illumination changes and viewpoint changes with the lowest computational cost of all multi-scale analysis based detectors. It has great potential to be applied in many computer vision tasks.

4. Chapter 5 addresses the problem of providing a simple and computationally efficient corner detector with sufficient detection accuracy, repeatability and high adaptability to achieve real-time detection. In order to simplify the corner detection architecture and improve its parallel computing performance, a new type of filter is proposed that can enhance corners and suppress edges as well

as noise simultaneously. Then a novel corner detector with high computational efficiency is proposed, and the corresponding hardware design on FPGA is also presented. Experimental results show that, with very low computational cost and simple architecture, the proposed detector can achieve or even exceed the detection accuracy of multi-scale analysis based detectors, and its repeatability is similar to multi-scale analysis based detectors and clearly higher than that of other types of corner detectors. It is potentially useful as an efficient corner detector for computer vision applications, especially for portable real-time tasks.

5. Finally, in Chapter 6, a summary of the researches and contributions made by this thesis is provided. The novelties of the proposed methods are presented and analyzed. Based on the discussion of the proposed methods in this thesis, further potential works and applications are discussed.

Chapter 2

Literature Review

As a stable sparse feature of images, corners contain important structural information and play a key and irreplaceable role in computer vision and image processing with applications in many fields. Currently, there is no strict mathematical definition of corners. It is generally agreed that a corner is a pixel where

1. The grayscale changes drastically in multiple directions,
2. A local maximum of curvature on the contour map is attained, or
3. Multiple edge curves intersect.

Therefore, according to different principles, current corner detection algorithms can be roughly divided into four categories: derivative-based methods, contour-based methods, template-based methods, and multi-scale analysis based methods. These four types of methods have their own advantages and disadvantages, with different degrees of defects in practical applications.

In this chapter, a literature review of recent corner detector algorithms is presented, which contains the classification of existing algorithms, as well as an analysis of the advantages and disadvantages of each type of algorithms, then possible research directions for further improving the performance of corner detection are pointed out.

2.1 Derivative-based Corner Detectors

The derivative-based corner detection algorithms mainly detect corners by checking for grayscale changes within a local image area. As early as 1977, Moravec [22, 23] found obvious differences in the grayscale change patterns between edges, corners and homogeneous regions in images. The grayscale changes in different directions in homogeneous regions are very small, while there is only one large grayscale change perpendicular to the ridge around an edge. For corners, there are large grayscale changes in all directions of gradients. Based on this observation, a well-known corner detection algorithm, namely the Harris corner detector [24], was developed by Harris and Stephens. The Harris corner detector [24] pioneered the research on derivative-based corner detection, and it has been widely used in many computer vision tasks.

The Harris corner detector [24] detects corners by obtaining a 'corner score' from a two-directional structure tensor formed by image gradients. It has achieved great success, formed the framework of derivative-based detectors, and is widely adopted in many other detectors. The Harris corner detector [24] obtains grayscale information by moving a detection patch over the input image and detecting corners by an analysis of local intensity changes using a two-directional structure tensor function. In an input image I, after placing the detection patch over an area around (x, y)

and shifting it by (u, v), the intensity change $E(u, v)$ of the response of the detection patch is defined as

$$E(u,v) = \sum_{x,y} g_\sigma \left[I(x+u, y+v) - I(x,y)\right]^2 \tag{2.1}$$

where g_σ is a Gaussian filter with scale factor σ. To reduce the number of computations, $I(x+u, y+v)$ can be simplified by the Taylor function to

$$I(x+u, y+v) = I(x,y) + I_x u + I_y v + O(u^2, v^2) \tag{2.2}$$

where I_x and I_y are the two directional partial derivatives of the input image. Then Equation (2.1) can be rewritten as

$$\begin{aligned} E(u,v) &= \sum_{x,y} g_\sigma [I(x,y) + I_x u + I_y v + O(u^2, v^2) - I(x,y)]^2 \\ &\approx \sum_{x,y} g_\sigma \left[I_x u + I_y v\right]^2 \\ &= [u, v]\, M \begin{bmatrix} u \\ v \end{bmatrix} \end{aligned} \tag{2.3}$$

where M is the structure tensor

$$M = \sum_{x,y} g_\sigma(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{2.4}$$

As shown in Fig. 2.1, each pixel (x, y) of the image can be classified by the eigenvalues λ_1 and λ_2 of M, and there are three cases that can be summarized as follows:

1. If $\lambda_1 \gg \lambda_2$ or $\lambda_2 \gg \lambda_1$, pixel (x, y) can be considered as a point on an edge.
2. If λ_1 and λ_2 are small values, pixel (x, y) can be considered as a point on a flat region.

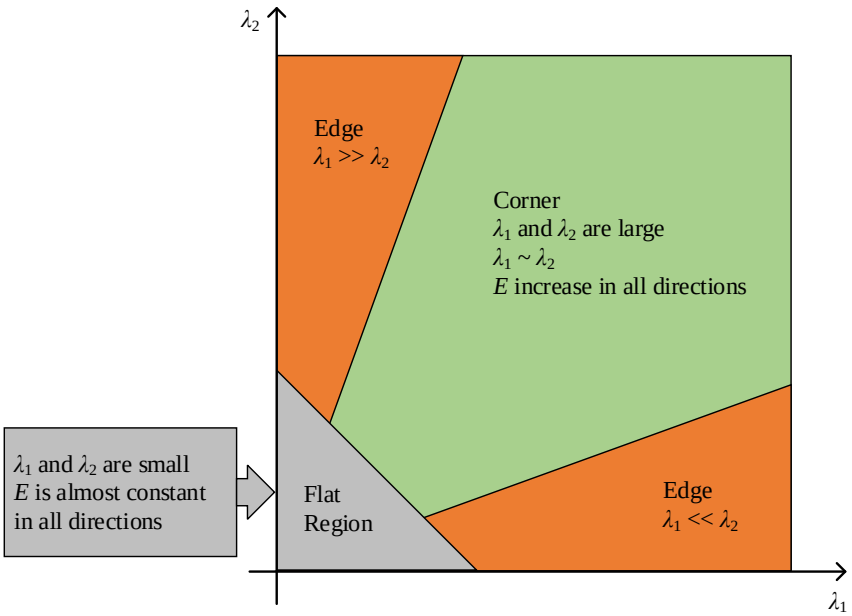


Figure 2.1: The classification of pixels by the eigenvalues λ_1 and λ_2 of M [24].

3. If λ_1 and λ_2 are both large values, pixel (x, y) can be considered as a corner.

Due to the high computational complexity of calculating eigenvalues, in order to reduce computing resources in practice, Harris and Stephens [24] proposed an approximate function, called the corner response function R, defined as

$$R = \det(M) - k_c(\text{trace}(M))^2 \tag{2.5}$$

where k_c is a constant and is usually set between 0.04 and 0.06, $\det(\cdot)$ and $\text{trace}(\cdot)$ are the determinant and trace of a matrix respectively. The value of R is the 'corner score'. When the 'corner score' is a large positive value, pixel (x, y) can be considered as a corner.

Based on the structure tensor and 'corner score' from the Harris corner detector [24], many researchers have proposed improved algorithms to achieve better performance. In the Harris corner detector [24], the response function cannot show sufficient de-

tection performance and robustness. The selection of the empirical constant in the response function affects the detection accuracy, and the detection procedure is more time-consuming than template-based corner detection algorithms that is discussed in Section 2.3. Therefore, many improved algorithms for the analysis functions of gradient and corner response functions have been proposed. Shi and Tomasi [31] constructed a new response function established directly through the smaller eigenvalue of the structure tensor, which can increase the number of detected corners and improve the detection performance. In other work [37, 38], two adaptive parameterised functions were proposed for replacing the suppression constant k_c to improve adaptability. Ryu et al. [32] proposed a new response function with a local weight function to improve the robustness to local variations. They also showed that the Harris corner detector [24] can be extended to have better localisation performance by applying the log-log scale to gradient vectors of image pixels [33]. Ando proposed an approach with image field categorisation to construct a covariance matrix of the gradient vector in a small window and applying canonical correlation analysis for each pixel to detect corners [34]. Mainali et al. [36] reduced the computational complexity by using integral images and an improved efficient non-maximum suppression (NMS). Bongjoep et al. [35] used an adaptive structure tensor for different regions to detect corners, which can improve the detection and localisation accuracy of the Harris corner detection [24].

However, the aforementioned algorithms still depend on an analysis of the two-directional derivatives of images. From the point of view of filtering, the envelope of the derivative filter is too small and sharp, only contains two directions and much detailed high frequency information is lost. Therefore, the acquired image structural information is relatively insufficient and lacks details, which is very sensitive to noise, affine transformation and local variations. Also, even though the derivatives

in two directions can estimate the grayscale changes in any direction, there are still large errors in practice especially in complex image regions. Therefore, robustness to rotation of these algorithms is insufficient. To address these shortcomings, many derivative-based corner detectors have been proposed based on more advanced filters to obtain structural information for detection. Förstner [29] proposed a least-square function to achieve better localisation accuracy than the Harris detector [24]. Moreover, other methods such as the Hessian matrix [25] and the Hilbert transform [26] have also been used to derive a 'corner score' for corner detection. Lindeberg [30] proposed an adaptive scale selection algorithm based on the maximum normalised gradient to extract image feature information.

Wavelets have shown strong performance on image decomposition and feature extraction in the field of image processing. They have been used to obtain the local grayscale changes of images and have been embedded into the Harris algorithm framework [24] to achieve corner detection [27, 126]. Furthermore, Gabor wavelets were used to obtain structural information to achieve better performance [28]. Multi-resolution methods such as Log-Gabor wavelet [127], dual-tree complex wavelet [90], contourlets [97], ridgelets [98], curvelets [99] and shearlets [100] have been introduced to capture the intrinsic geometrical structure and widely used in corner detection algorithms to improve performance on noise suppression, image feature extraction and corner localisation accuracy. These methods involve multi-scale detection and will be discussed in greater detail in Section 2.4 .

The aforementioned methods for derivative-based corner detection were proposed within the Harris algorithm framework, with more effective algorithms for the extraction of local structural information and detection of corners. Generally speaking, these algorithms perform corner detection by extraction of two-directional struc-

tural information in images to establish the structure tensor and response functions through the autocorrelation matrix. The advantage is that the simple process and fewer directions have low computational complexity, with acceptable robustness to affine transformations such as noise and illumination changes when the filter is selected appropriately. However, the shortcomings are also very obvious. These algorithms use fixed scale filters to decompose images in two directions, thus a finer scale filter can provide good localisation accuracy and detection performance, however it is very sensitive to noise and local variations, while a coarser scale filter can provide better robustness, but its detection performance and localisation accuracy are poor. Furthermore, the use of only two-directional structural information in fixed directions can only detect 'L' type corners, and cannot distinguish corners from other pixels well [56]. In consequence, the two-directional structure tensor loses a lot of useful information available in images, and cannot provide sufficient detection performance, localisation accuracy and repeatability under affine transformations and noise.

2.2 Contour-based Corner Detection

In the contour map of an image, corners always have larger curvature values than other pixels. Therefore, singularities with large curvatures on edges can be used for corner detection [57]. Contour-based corner detection algorithms mainly detect corners by analysing the characteristics of contour shapes in images. Generally, these methods first extract the edge map from images, then form closed or non-closed contours from the edges and finally detect corners by finding the maximum curvature of pixels on the contours or sharp changes in gradient directions, or by using polygonal approximation.

In early algorithms, images were segmented first to extract the Freeman chain code from the segmented image contour, then the average change of gradient and angle of each contour pixel [58], or the average change of slopes [59] within the chain code were calculated to detect corners. Beus and Tiu [60] improved the detection algorithm by calculating the average local curvature and determining outliers. Cooper et al. [61] used the pixel coordinates of the chain code to estimate the curvature and detected corners through the local maxima. Hsin-Teng [62] used polygons to approximate the shape for the chain code, and Arrebola and Sandoval [63] used a resolution coupling method to process the curvature represented by the chain code to improve the detection performance.

Compared with the Harris corner detection algorithm [24], the aforementioned algorithms need to analyse the chain code formed by contours pixel by pixel and they have much higher computational complexity. However, these algorithms do not show better performance, due mainly to the reason that the analysis of curvature depends heavily on the quality of the image contours. In general, to achieve good localisation and detection performance, the window of curvature calculation should be small. However, the small detection window cannot smooth low quality parts in the contour map. Thus the breaks, offsets and jagged edges in the contour map seriously affect detection performance, and a large number of pseudo corner responses will be generated. To solve this problem, Mokhtarian and Suomela applied the curvature scale space (CSS) [64] method to corner detection to propose the well-known CSS corner detection algorithm [65]. In this algorithm, the Canny edge detector [82] is used to extract the contour map of an image. For a pixel u in the contour map, different scale Gaussian kernels $g(u, \sigma)$ are used to smooth the edges

to obtain a multi-scale local curvature estimation $K(u,\sigma)$, which is defined as

$$K(u,\sigma) = \frac{\dot{X}(u,\sigma)\ddot{Y}(u,\sigma) - \ddot{X}(u,\sigma)\dot{Y}(u,\sigma)}{[\dot{X}(u,\sigma)^2 + \dot{Y}(u,\sigma)^2]^{\frac{3}{2}}} \tag{2.6}$$

where

$$\begin{aligned} \dot{X}(u,\sigma) &= x(u) \otimes \dot{g}(u,\sigma) \\ \ddot{X}(u,\sigma) &= x(u) \otimes \ddot{g}(u,\sigma) \\ \dot{Y}(u,\sigma) &= y(u) \otimes \dot{g}(u,\sigma) \\ \ddot{Y}(u,\sigma) &= y(u) \otimes \ddot{g}(u,\sigma) \end{aligned} \tag{2.7}$$

The $\dot{g}(u,\sigma)$ and $\ddot{g}(u,\sigma)$ are the first and second order derivatives of the Gaussian kernel respectively, and $x(u)$ and $y(u)$ are the path lengths along the contour map in x and y directions. Pixels with maximum local curvature at the coarsest scale are identified as candidate corners. Then the candidate corners are tracked from the coarsest scale to the finest scale to make the positions of candidate corners more accurate. CSS has become the standard technology for describing contour shapes of images in MPEG-7.

Although the CSS detector [65] has achieved great success, it still suffers from three main shortcomings [67, 68, 70–72]:

1. The scales of Gaussian filters used for detection are fixed.
2. The detection results are sensitive to the noise on edges.
3. The selection of the corner threshold has a great impact on performance.

In order to solve these problems, Mokhtarian and Mohanna [66] used Gaussian kernels of different scales adapted to the contour lengths to better extract the curvature

of pixels and suppress noise. Ray and Pandyanp [69] obtained an adaptive scale to smooth contours by calculating the variance of curvatures. Zhong and Liao [67] used the direct curvature scale space to model the corners and extracted the invariant geometric characteristics to reduce the computational complexity and improve noise robustness. Zhang et al. [70] proposed a multi-scale curvature product algorithm to improve detection performance, which can better suppress false corners. Gao et al. [68] proposed a detection algorithm through the saliency-scale space based on maxima modulus of wavelet transform at different scales. Masod [71] used recursive optimization to remove false corners. He and Yung [72] proposed an adaptive curvature threshold function based on a dynamic support interval to reduce the impact of threshold on performance.

However, these methods obtain candidate corners in the coarsest scale, making the contour map too smooth and causing these algorithms to miss a lot of corners. Also, the curvature of pixels still need to be calculated within a small window, which is greatly influenced by the quality of the contour map. To address these shortcomings, Awrangjeb and Lu [74] used the affine length instead of arc length for curvature estimation and proposed an improved CSS corner detector as an alternative to calculation on curvature. The CPDA detector [75] was proposed by applying a chord-to-point distance accumulation method [73] to estimate curvatures and achieve corner detection. In these algorithms, the curvature is estimated indirectly, which can reduce dependence on the quality of contours.

In recent years, researchers have realized that the uniqueness of grayscale patterns of corners is very effective for detection. Compared with curvature estimation, analysis of the grayscale pattern is more independent of the quality of the contour map and can provide more robust detection results. Therefore, gradients have replaced curva-

ture estimation in some contour-based methods. Zhang et al. [76] used the gradient distribution to construct a gradient correlation matrix for each pixel on contours and obtained the eigenvalues of the matrix to detect corners, which is very similar to the 'corner score' in the Harris corner detector [24]. Rimon and Robot [77] used the change of gradient direction on edge pixels to detect corners. Zhang et al. [78] detected corners by finding pixels with large grayscale changes obtained by Gabor filters in all directions on the contour map. The ACJ corner detector [79] detects corners through statistical modelling of normalised image gradients. Zhang and Shui [80] proposed a detector that uses the principal direction of the anisotropic directional derivatives to detect corners. Kim estimated the curvature of each pixel in the whole image through the gradients to remove the step of obtaining the contour map to improve computational efficiency [81].

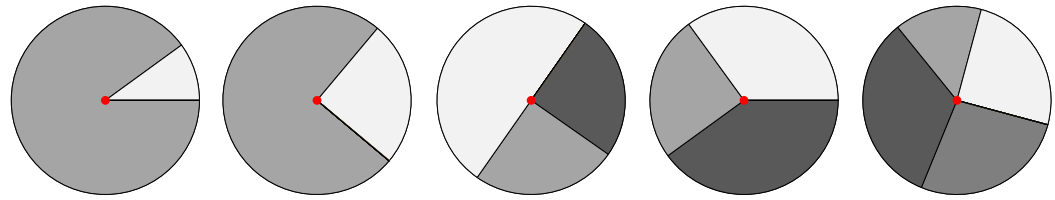
In summary, the contour-based detectors detect corners by analysing the characteristics of pixels on the contour map obtained from images. The state-of-the-art contour-based detectors can provide stable detection results through indirect curvature estimation [74, 75] or grayscale analysis [76–81]. However, their results depend heavily on the prior segmentation and boundary tracking, with the highest computational complexity among all the corner detection methods. In state-of-the-art methods [75–80], edge detection methods such as Canny [82] or LSD [83] are generally used to obtain edge contours. However, neither of these two methods can provide sufficient quality contour maps for corner detection, especially in the LSD method where contours are decomposed into many straight lines, greatly reducing the detection performance. Also, image discretisation and local variations will further increase the number of low-quality parts in the contour map such as breaks, offsets and jagged edges. This will greatly affect the subsequent analysis of the contour map, thereby greatly reducing the performance and robustness of corner

detection. Although detection on the coarsest scale can slightly overcome this problem, it also has a great impact on localisation accuracy. If a more complex and better edge detection algorithm is used, the originally high computational complexity will be further increased, making the detector impracticable.

In addition, when tracking candidate corners on multiple scales [65–72] or accumulating point-to-chord distances on contours [74, 75], a large number of repeated operations on a single pixel significantly increases the computational complexity [128]. Furthermore, multi-scale information is only used to improve the localisation accuracy of candidate corners instead of to improve overall detection performance. For methods based on grayscale analysis [76–80], the fixed detection scale of gradients greatly influences detection performance and localisation accuracy, and two-directional gradients cannot provide sufficient structural information for detection. Also, compared with the derivative-based detector, computation of the chain code from the contour map not only increases the computational complexity, but also makes the algorithms susceptible to the quality of contours.

2.3 Template-based Corner Detection

The basic approach of template-based corner detection algorithms is to use different types of parameterised templates to match corners from images. These methods do not perform analysis on first-order or second-order information from images or estimate curvatures. Therefore, template-based corner detection algorithms can provide good computational efficiency among all the types of corner detectors, and many of them can also meet the real-time requirement in some computer vision tasks.



(a) Wedge-shaped (b) L-shaped (c) T-shaped (d) Y-shaped (e) X-shaped

Figure 2.2: Different types of corner models.

Early research on template-based corner detection focussed on building and combining different types of corner models shown in Fig. 2.2. Guiducci [39] established a wedge-shaped corner model based on representation of the Gaussian kernel, and the amplitude characteristics of the representation were analyzed. Rangarajan et al. [40] established an L-shaped corner model based on the Gaussian kernel, and optimized the peak signal-to-noise ratio and localisation error of the model based on Canny edge detection [82]. Singh and Shinier [45] proposed a method that combines different types of corner templates to detect corners. Rohr [41] extended the wedge-shaped corner template [39] to T-shaped, Y-shaped and X-shaped corner templates. On this basis, Derrich and Giraudon [42] used the Gaussian kernel to characterise and analyse the wedge-shaped and Y-shaped corner models to achieve better localisation accuracy. Blaszk and Derrich [44] proposed a parametric model to describe high-order information of corners to realise detection and classification of corners. Furthermore, Parida et al. [129] proposed a model-based corner detection, classification and reconstruction method by using high-order information. Based on contour analysis, Rosin [43] analysed the curvature characteristics of wedge-shaped corners to establish corner models. Ruzon and Tomasi [46] extended the framework of template-based corner detection to colour images.

Based on the aforementioned models of corners, to meet the requirement of real-time computing, Smith and Brady [47] proposed the well-known SUSAN corner detection

algorithm that combined all the corner types. The principle of SUSAN is to compare the centre pixel with its neighbourhood and count the number of pixels that have a similar grayscale value as the centre pixel, which is defined as

$$c(\vec{r}, \vec{r_0}) = \begin{cases} 1, & |I(\vec{r}) - I(\vec{r_0})| \leq t \\ 0, & |I(\vec{r}) - I(\vec{r_0})| > t \end{cases} \tag{2.8}$$

where $\vec{r_0}$ represents the position of centre pixel (x_0, y_0) and $\vec{r}$ represents the position of another pixel in the neighborhood. $I(\vec{r})$ is the grayscale value of $\vec{r}$, and t is a constant threshold used to evaluate the similarity. Then, the count of the SUSAN template is defined as

$$N_c(x_0, y_0) = \sum_{\vec{r} \neq \vec{r_0}} c(\vec{r}, \vec{r_0}) \tag{2.9}$$

If $N_c(x_0, y_0)$ is larger than a given threshold, pixel (x_0, y_0) will be considered as a corner. Obviously, different from the detection method using first-order or second-order partial derivatives of images or the contour map of images, the SUSAN detector only compares the local zero-order information from images. Thus, this method can avoid complex analysis and processing, and provide extremely high computational efficiency. However, its detection performance and localisation accuracy are poor and it cannot detect X-shaped corners at all. Based on the SUSAN detector [47], Trajkovic and Hedley [48] futher reduced the computational complexity of corner detection. However, as with SUSAN detector, detection performance is unstable with many false corner responses in the diagonal direction edges. Bae et al. [49] combined the idea of SUSAN with a derivative-based algorithm, using two crossover operators to filter the image separately and detected corners according to the grayscale change of pixels. Cazorla and Escolano [50] extended the SUSAN template to parameterised templates with multiple angles to match corners of different structures, and achieved corner classification based on a Bayesian model of parameter estimation. Lepetit

and Fu [51] extended the SUSAN algorithm for 3D images. Lan and Zhang [52] proposed a double-circle template to replace the circular template in the SUSAN detector [47] to improve detection performance and robustness.

Although the SUSAN detector and its improved versions have extremely high computational efficiency, their performance is poor. Especially for complex scenes and noisy images, the measuring function of similarity causes a large number of pseudo corner responses which greatly reduces detection accuracy. In light of SUSAN, Rosten et al. [53] proposed an algorithm that uses a ring template to replace the circular template in SUSAN [47] and counts the number of pixels in the template with a large enough grayscale difference from the centre pixel, instead of measuring the similarity. The detection performance is further improved through the use of a machine learning algorithm, namely decision tree, to optimize the search space of corners and improve robustness of the detection template [54]. This algorithm is the well-known FAST corner detection algorithm. Compared with the SUSAN detector [47], the detection performance of the FAST detector is significantly improved, and its robustness to different image types is also significantly higher. Also due to the introduction of a relatively small template and decision trees [53, 54], the FAST detector is more computationally efficient.

The FAST detector has been widely used in many computer vision tasks due to its excellent computational efficiency and acceptable detection performance. Based on the FAST detector, Xiu et al. [130] proposed a new template with a filled circle and outer ring, and used an adaptive threshold function to improve detection accuracy and repeatability. Combined with the 'corner score' in the Harris corner detector [24], Rublee et al. [131] extended the FAST corner detector to multi-scale for feature description. Karim and Nasser [132] further improved the computational

efficiency by applying Haar wavelet to reduce the feature space. Through adaptive contrast enhancement, Kadhim and Araheemah [133] achieved better repeatability on illumination change. Wen et al. [134] extended the FAST algorithm to event cameras.

Template-based corner detection methods represented by the FAST detector [53,54] have been applied to many computer vision tasks due to higher computational efficiency among all types of corner detectors. However, these methods lack prior knowledge about grayscale change information and detect corners directly through comparison of grayscale values. Thus the detection performance of these detectors is not high, especially for complex scenes and noisy images. Meanwhile, without image smoothing, their detection performance is unstable and repeatability under affine transformations, local variations and noise is low. Furthermore, since the response function of these methods counts the number of different or similar pixels, there will be a large number of repeated responses around one corner, and non-maximum suppression does not handle it well. In practical applications, parameterised corner templates cannot handle complex image scenes well. Therefore, compared with state-of-the-art corner detectors, the template-based corner detection methods cannot achieve efficient detection performance and repeatability, and they only have an advantage in terms of computational complexity.

2.4 Multi-scale Analysis Based Corner Detection

In early research on corner detection, most algorithms detected corners at a fixed scale. Soon, researchers found that due to image characteristics as well as the influence from illumination, noise and local variations, detection algorithms based on a

single scale cannot extract all corners well. Meanwhile, whether in derivative-based methods, template-based methods or contour-based methods, detection performance is greatly affected by the size of the detection window. A large detection window, e.g., large-size filter, large-scale curvature estimation or large-size template, can greatly increase the robustness of detection results. However, it will still miss many corners, reduce detection performance and significantly lower the localisation accuracy. On the contrary, a smaller detection window can detect more corners and achieve more precise localisation accuracy, but the repeatability to local variations and noise will be greatly reduced. Therefore, researchers have gradually realized the importance of multi-scale information for corner detection. In recent years, with the development of multi-scale geometric analysis technology, local grayscale changes and geometric structure of images can be described more finely with better robustness. In consequence, these methods have begun to be widely used in corner detection algorithms to improve noise suppression capabilities, extraction performance of feature information and corner localisation accuracy.

Early work on multi-scale analysis based detectors, namely scale-space detectors, focussed on extracting corners from scale space using classical detectors. For example, Mikolajczyk and Schmid proposed the well-known Harris-Laplace detector [84], in which the Harris corner detector [24] is applied with scale space theory [30], and scale adaptive selection is performed to obtain salient corners on each scale by using the Laplacian operator. Similarly, the Hessian-Laplace corner detector [85] and multi-scale Plessey detector [86] were proposed, in which the Hessian matrix and the Plessey operator were extended into multi-scale. Furthermore, some multi-resolution image analysis tools such as wavelets were used to obtain more precise and detailed information. For example, pyramid decomposition in wavelet analysis was used for corner detection at multiple scales [87–89]. However, these methods

repeatedly apply the single-scale corner detection algorithm to different scales, although repeatability under scale change is improved, these methods cause a large number of repeated detections for one corner at different scales, and the shortcoming of the fixed detection window still exists.

In recent years, multi-scale analysis algorithms have attempted to take full advantage of multi-scale structural information to improve corner detection performance and robustness. Recently, due to their promising detection performance, these detectors have attracted more attention. Since wavelets have limited ability for directional decomposition, lifting wavelets [91], dual-tree complex wavelets [90], Gabor wavelets [92], and Log-Gabor wavelets [93] were used for obtaining structural information from images. Gao et al. [93] proposed a corner detection method, namely LGWTSMM, in which the average Log-Gabor wavelet coefficients from multi-scale in all directions were weighted and summed to establish a two-directional structure tensor as defined in Eq. (2.4) for corner detection. This method outperforms the other wavelet-based methods. However, the Gabor wavelet kernel is highly non-orthogonal and can cause redundancy in the decomposition coefficients [96], which will decrease the distinction between corners and other pixels. Moreover, the averaging process of the weighted summation of structural information in all directions will lead to the loss of some useful structural information and increase corner localisation error. Zhang and Sun [94] proposed a corner detector based on anisotropic Gaussian derivative filters. Based on eight directional derivative coefficients at the finest scale, an 8×8 structure tensor was constructed for candidate corner detection,

which is defined as

$$M = \begin{bmatrix} \sum_{i=-\frac{u}{2}}^{\frac{u}{2}} \sum_{j=-\frac{v}{2}}^{\frac{v}{2}} \nabla^2_{\sigma,\rho,1} & \cdots & \sum_{i=-\frac{u}{2}}^{\frac{u}{2}} \sum_{j=-\frac{v}{2}}^{\frac{v}{2}} \nabla_{\sigma,\rho,1}\nabla_{\sigma,\rho,8} \\ \vdots & \ddots & \vdots \\ \sum_{i=-\frac{u}{2}}^{\frac{u}{2}} \sum_{j=-\frac{v}{2}}^{\frac{v}{2}} \nabla_{\sigma,\rho,8}\nabla_{\sigma,\rho,1} & \cdots & \sum_{i=-\frac{u}{2}}^{\frac{u}{2}} \sum_{j=-\frac{v}{2}}^{\frac{v}{2}} \nabla^2_{\sigma,\rho,8} \end{bmatrix} \tag{2.10}$$

where $\nabla^2_{\sigma,\rho,k} = \nabla^2_{\sigma,\rho,k} I(x+i, y+j)$ is the directional derivative coefficients along direction k, σ and ρ are the sigma and anisotropic factor of the Gaussian kernel. The structure tensor is computed within a $(u+1, v+1)$ area of the center pixel (x, y). The coefficients of other scales were used to pick up any falsely detected corners from the candidate corners. The multi-directional structure tensor is less sensitive to image rotation and noise, but more computation time is needed to deal with the multi-directional structure tensor, therefore this detector is more time-consuming.

Natural images contain intrinsic geometrical structures that are important for corner detection. The aforementioned wavelets are good at isolating discontinuities at edge points, however they cannot handle equally well distributed singularities such as those along curves [95, 96]. As a result, contourlets [97], ridgelets [98], curvelets [99] and shearlets [100] have been introduced to capture the intrinsic geometrical structures, and have been successfully embedded in some existing corner detection frameworks [101–109] such as Harris [24] and SUSAN [47] corner detectors. Among these methods, shearlets provide the best performance due to the optimal and sparse representation of geometrical structures, high noise robustness and relatively simple generation functions [100]. The details and theoretical derivations are provided and discussed in Section 3.1. Some detectors based on shearlets have been introduced to achieve better performance and robustness. Duan et al. [106]

employed a 3D version of the traditional shearlet transform with a dual-tree structure to capture structural information from 3D magnetic resonance images, then the structure tensor in the Harris corner detector was used to analyse the structural information for image fusion. Malafronte et al. [107] constructed a spatio-temporal interest point detector based on the shearlet transform, in which the shearlet coefficients in all directions at each scale were summed. The products of the resultant values at the two finest scales were compared with a threshold for interest point detection. Duval-Poo et al. [135] proposed another interest point detection method based on shearlets. They developed a measure function by summing the shearlet coefficients in all directions at all scales for interest point detection, which has been proven to be relatively accurate in detection and highly robust to noise and scale variations. Then they proposed an edge and corner detection method with a modified shearlet transform [108,109]. They first calculated the mean of all scale shearlet coefficients in each direction, and then the sum of all directional means weighted by a sinusoidal function was considered for corner detection. Benefitting from the modified shearlets, this method yields better performance in terms of detection accuracy and robustness to noise compared with the traditional shearlet-based methods.

In summary, current multi-scale analysis based algorithms focus on merging information from multiple scales, which can address the shortcomings of the fixed detection window in single-scale detectors, and can improve detection performance with higher robustness. Multi-scale analysis based detectors contain two main parts: multi-scale decomposition of images and merging decomposition results for corner detection. For multi-scale decomposition, the shearlet has been adopted and achieved good performance [106–109,135]. However, the traditional shearlet transform suffers from some inadequacies in corner detection including the following:

1. The shearlet transform suffers from a strong side-lobe effect. The local structural information at the locations where the grayscale values drastically change, such as along edges or at corners, has multiple peaks.

2. Since a corner is not simply formed by straight lines, and there might be some intensity changes in several directions in its background, the structural information in two adjacent directions is not completely independent. However, in the shearlet transform, there is no overlap between the shearlets of adjacent directions in the Fourier domain, making the structural information of adjacent directions independent.

3. The mother function of shearlets has limited support in the Fourier domain, consequently the corresponding spatial filter has a larger envelope. This results in the problem of edge extension which causes erroneous detections in the case of the extensions of two edges intersecting with each other.

4. The spatial filter with a larger envelope depresses the high-frequency components in the image, making the detailed information less clear, and

5. In shearlets, the orientation variable is associated with the scale index. For different scales, the number of directions is different and is determined by the scale index. This is a great obstacle for subsequent corner detection when using multi-scale directional information.

Another key factor affecting the performance of corner detectors is how to use multi-scale decomposition results to detect corners. In current methods, only the amplitude information is considered, and another important structural information, namely phase information, is neglected. In methods [91,93,132], only the amplitude coefficients of one scale are utilised for corner detection. This limits their detec-

tion accuracy and repeatability under affine transformations, viewpoint changes and noise. Also, the averaging process of the weighted summation of structural information in all directions will lead to the loss of some useful structural information, and the introduction of amplitude information from coarse scales seriously reduces corner localisation accuracy. In other work [108,109], the simple usage of the amplitude coefficients at all scales by simply summing them up and comparing with a given threshold cannot accurately distinguish corners from edges, and merging multi-scale directional information by calculating their means, especially merging information from coarse scales, will greatly increase corner localisation errors. Elsewhere [94], only the amplitude coefficients at the finest scale are used for detecting candidate corners, making the results of this stage sensitive to image affine transformations and local variations, while the amplitude coefficients of other scales are used to eliminate false candidate detections. This means that information from these scales only helps reduce the false detection rate rather than improve the overall detection accuracy.

In summary, current multi-scale analysis based methods yield the best performance among all existing detectors. However, these methods do not make full use of multi-scale and multi-direction structural information from images, which influences their detection accuracy, localisation accuracy and repeatability. Multi-scale decomposition tools cannot provide sufficient structural information for detection, which causes the loss of detailed information in images and falsely detected corners. Moreover, the processing of multi-scale information and removal of false candidate corners need more computation time. Therefore, such detectors have a higher computational complexity than template-based detectors and derivative-based detectors.

2.5 Deep Learning Based Interest Point Detector

In recent years, deep learning has been extended for interest point detection. Due to there is no large database of corner labelled images that can be used for training a neural network, unlike the corner detectors which focus on low-level image structures, recent deep learning-related interest point detectors pay close attention to locally unique descriptors in feature maps generated by deep learning. Most of these algorithms cannot detect corners, but provide features similar to blob features such as DoG [136] and SIFT [137], so they are mentioned as interest point detectors in this thesis.

Early work of deep learning based interest point detectors focussed on replacing handcrafted feature detection and description pipelines with neural networks [114–116]. However, since their limited generalization abilities, these methods cannot reach the accuracy of their replaced handcrafted features [19–21, 138]. Recently, other works [110, 113, 117–122] attempted to generate and select locally unique descriptors in feature maps, and focussed on optimization of feature matching scores with metric learning. They used image patches which are pre-cropped by handcrafted features such as DoG [136] and SIFT [137] for learning and training a patch similarity network [139]. Some methods replaced the handcrafted feature detection for training by synthetic data and perform homographic adaptation for better domain adaptation performance [111, 112], while some others totally discarded interest point detection in their algorithms [123, 124]. However, the optimized accuracy in matching does not perform improvement on performance in high-level vision tasks [21, 111]. For example, the SuperPoint method [112] which is superior to SIFT [137] on matching, cannot reach the performance of SIFT [137] for estimating an essential matrix for an image pair. The LIFT detector [110] cannot provide bet-

ter reconstructions than SIFT [137] in a structure-from-motion pipeline [21]. The reason may be concluded that although the feature detection based on higher-level information can provide better performance for patches matching, which is benefited from optimization of the description of features, these methods inherently provide less accuracy for keypoints localisation and detection [113]. Meanwhile, compared to sparse features, the deep learning approaches are much less computational efficient due to the need to densely extract descriptors and generate a group of feature maps for a whole image.

The focusses of these deep learning algorithms are not on feature detection, but mainly on generating and selecting locally unique descriptors in feature maps and optimizing feature matching processing. Therefore, they are very different from sparse corner detection. It is well known that corner detection belongs to the first stage of sparse local features in a framework of detection first and then description. In the detection stage, a set of interest points such as corners or blobs is obtained, and then local descriptors are extracted from the image patches around the interest points. On the contrary, recent deep learning-related interest point detectors follow the framework of description first and then detection. They first generate a group of feature maps with deep learning, and then the interest points are detected from the maps. D2-Net [113] directly marked the central points of the local maximum descriptors as interest points, and SuperPoint [112] performed interest point detection by using a decoder branch. Reinforced SP [111] retrained SuperPoint via reinforcement learning to improve performance. In general, interest points detected by deep learning are less accurate but have high repeatability which is important for image matching and retrieval tasks, while low-level image structures such as corners or blobs are more accurate and memory-efficient, and their accurate localisation ability is very useful for visual localisation, motion estimation and 3D reconstruction [125].

Consequently, the accurate detection of corners with high repeatability remains a core issue of computer vision.

2.6 Discussion

Corner detection algorithms have undergone considerable development in recent years, especially multi-scale analysis methods that have achieved good detection performance and robustness. However, there are still some problems to be addressed. First of all, the existing corner detection algorithms only use the information of images in the horizontal and vertical directions to form a structure tensor [91,93,101–105,132], or even just the information in a single direction to form a single-directional measure function [93,108,109]. Many multi-resolution analysis tools such as wavelets can only provide structural information in two directions. However, corners are the special pixels in two-dimensional images, and their characteristics are anisotropic. The discarding of information in other directions reduces the ability of algorithms to distinguish corners from other pixels and lowers the detection performance and repeatability.

Meanwhile, in the existing methods including multi-scale analysis based methods, multi-scale information has not been fully utilised for corner detection. Many existing methods detect corners based on a fixed scale or perform repeated detection on multiple scales [94, 101–105, 132]. This greatly affects detection performance, localization accuracy and repeatability. Although some methods average multi-scale information for better performance [93, 108, 109, 135], simple average processing in multi-scale will reduce the localisation accuracy of the algorithms, and make them ignore some small regions with few grayscale changes in images. In some reported methods [65–72, 94], structural information of multiple scales is used to eliminate

false candidate detections. This means that information from these scales only helps reduce the false detection rate rather than improve overall detection accuracy.

From simple image gradients to shearlets, the tools used to extract image structural information for corner detection are more advanced and can provide better performance. The latest multi-resolution analysis tool, namely the shearlet transform, can provide multi-directional and multi-scale information with high robustness and has shown effectiveness in corner detection. However, shearlets are designed for sparse representation rather than feature detection, therefore, the decomposition results for images suffer from some drawbacks, and the insufficient information greatly influences detection performance. In addition, existing decomposition tools cannot extract phase information from images. However, compared with amplitude information, phase information can achieve better localisation accuracy for geometric features, and provide more robustness information, especially for illumination change [140, 141]. Therefore, introducing phase information into corner detection could be beneficial to detection performance, localization accuracy and repeatability.

In terms of computational complexity, although derivative-based and template-based methods have high computational efficiency, neither can provide sufficient detection performance and robustness. Especially in template-based methods, too many analysis and smoothing steps are omitted for computational efficiency. Contour-based methods need to obtain the contour map and analyse the chain code pixel by pixel, so they generally have the highest computational complexity. Although multi-scale analysis based algorithms have higher detection performance, their computational efficiency is relatively poor due to the processing of multi-scale information. If more multi-scale and multi-directional information is used to improve detection perfor-

mance, the computational complexity will also increase. Therefore, an appropriate balance in the tradeoff between computational complexity and performance is needed.

This thesis addresses these issues in designing corner detection algorithms. In order to improve detection accuracy, repeatability and localisation accuracy of the current state-of-the-art in corner detection, an improved shearlet transform and a novel complex shearlet transform, which can overcome the problems of bilateral margin responses, edge extension and lack of phase information in traditional shearlets, are proposed to better localise distributed discontinuities, especially to extract phase information from geometrical features. Moreover, a multi-directional structure tensor and a multi-scale corner measurement function are proposed to make full use of the structural information from the improved shearlets for corner detection, and a new rotary phase congruence tensor is proposed to utilise amplitude and phase information at all scales and directions of the complex shearlets for corner detection. As a result, two new corner detectors are proposed. Experimental results demonstrate that their localisation ability and detection accuracy are superior to current detectors, and their repeatability is generally higher than current corner detectors as well as recent deep learning interest point detectors. The proposed detectors have great potential for applications in the aforementioned computer vision tasks.

To meet the needs of real-time computer vision tasks, especially real-time portable tasks, a new type of filter that can enhance corners and suppress edges as well as noise simultaneously is proposed to simplify the corner detection architecture and improve its parallel computing performance. Then a novel corner detector with high computational efficiency is also proposed, and the corresponding FPGA design is provided as well. Experimental results show that, with very low computational

cost and simple architecture, the proposed detector can achieve similar detection accuracy and repeatability of current corner detectors and be potentially useful for real-time computer vision applications.

Chapter 3

Corner Detection Based on Improved Shearlet Transform and Multi-directional Structure Tensor*

As discussed in Chapter 2, current methods for corner detection do not make full use of multi-directional and multi-scale structural information, and the traditional shearlet transform suffers from some drawbacks which greatly influence corner detection performance. In this chapter, all these inadequacies are focussed and solved. An improved shearlet transform with a flexible number of directions and reasonable support is presented, which can accurately obtain multi-scale and multi-directional

* This chapter has been published: M. Wang, W. Zhang, C. Sun, and A. Sowmya, "Corner Detection Based on Shearlet Transform and Multi-directional Structure Tensor," *Pattern Recognition*, vol. 103, pp. 107299:1-15, 2020.

structural information from images. In order to make full use of the multi-scale and multi-directional information, a multi-directional structure tensor is constructed for corner detection based on the shearlet coefficients, and a multi-scale corner measurement function is proposed to reduce the influence of noise and suppress edges. The proposed corner detection method is compared with other representative corner detectors. The results obtained show that the proposed corner detection method yields some improvements on the current state-of-the-art in corner detection in terms of the detection accuracy, corner localization accuracy and robustness to affine transformations, illumination changes, noise, image blurring and viewpoint changes.

The main contributions of this chapter are summarized as follows:

1. A novel multi-scale and multi-directional analysis approach, namely improved shearlet transform, is proposed to extract accurate multi-scale and multi-directional structural information from images, and effectually mitigate the weaknesses of the traditional shearlets.

2. A multi-directional structure tensor and a multi-scale corner measurement function are proposed to take full advantage of the multi-scale and multi-directional structural information, which can effectively improve detection accuracy and depress false detection.

3. As a result, the proposed detector yields improvements to the current state-of-the-art in corner detection in terms of detection accuracy, localization accuracy and repeatability under image scaling, rotation, noise corruption, JPEG compression, illumination and viewpoint changes.

The remainder of this chapter is organised as follows. The traditional shearlet transform is briefly reviewed in Section 3.1. In Section 3.2, the inadequacies of the

traditional shearlet transform for corner detection are discussed and addressed, and an improved shearlet transform is developed. A novel corner detection with multi-directional structure tensor and a multi-scale measurement algorithm is proposed in Section 3.3. The experimental framework, datasets and extensive experimental results are given in Section 3.4. Finally, the concluding remarks are provided in Section 3.5.

3.1 Shearlet Transform

In recent years, several multi-directional image analysis methods such as ridgelets [98], curvelets [99], contourlets [97] and shearlets [100] have been introduced to capture the intrinsic geometrical structures. Compared to the other methods, shearlets can efficiently represent anisotropic features, and also have the best performance due to some unique and excellent properties [100] including the following:

1. Shearlets are generated by a single or a small set of functions.
2. Shearlets can provide optimal and sparse representation for anisotropic functions.
3. Shearlets have compact support in the Fourier domain.
4. Shearlets have fast algorithmic implementations.
5. Shearlets can provide a unified treatment for continuous and digital signals.
6. Shearlet transform is associated with classical approximation spaces.

For these reasons, shearlets are chosen in this work as a reference framework for corner detection to overcome the limitations of current state-of-the-art corner detection

methods that do not make full use of the multi-directional structural information from images. In Section 3.1.1, the theory behind the shearlet transform is reviewed, followed by the introduction of a discretization scheme for shearlets in Section 3.1.2, and the computational complexity is analyzed in Section 3.1.3.

3.1.1 Construction of Shearlet Transform

A shearlet is generated by a mother shearlet ψ in the space of square-integrable functions of 2D real space $\mathbb{R}^2$ (i.e., $L^2(\mathbb{R}^2)$) through dilation, shearing and translation, in the following manner [100]:

$$\psi_{a,s,t}(x) = a^{-\frac{3}{4}}\psi(A_a^{-1}S_s^{-1}(x-t)) \tag{3.1}$$

The shearlet depends on three parameters, the scaling matrix A_a, the shear matrix S_s and the translation parameter $t \in \mathbb{R}^2$. A_a and S_s are defined by

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, \quad S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \tag{3.2}$$

where a is the scale variable in the positive real space $\mathbb{R}^+$ which controls the scale of the shearlet, and s is the orientation variable in the real space $\mathbb{R}$ which determines the orientation of the shearlet. The representation of the mother shearlet in the Fourier domain $\widehat{\psi}$ can be factorised as

$$\widehat{\psi}(\omega_1, \omega_2) = \widehat{\psi}_1(\omega_1)\widehat{\psi}_2\left(\frac{\omega_2}{\omega_1}\right) \tag{3.3}$$

where $\widehat{\psi}_1$ is the Meyer wavelet and $\widehat{\psi}_2$ is a non-zero square-integrable function with $\text{supp}\widehat{\psi}_2 \subseteq [-1, 1]$. Here, $\text{supp}\widehat{\psi}_2$ denotes the support of function $\widehat{\psi}_2$, and $\widehat{\psi}_2$ is given

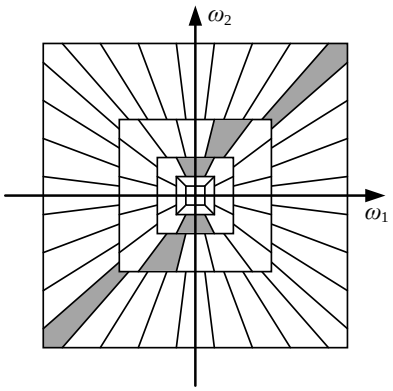


Figure 3.1: A shearlet transform system in the frequency domain with 3 shearlets of different $a,\ s,\ t$ shown in gray.

by

$$\widehat{\psi}_2(\omega) = \begin{cases} \sqrt{D(1+\omega)}, & \omega \leq 0 \\ \sqrt{D(1-\omega)}, & \text{otherwise} \end{cases} \tag{3.4}$$

where $D(x)$ is expressed as

$$D(x) = \begin{cases} 0, & x < 0 \\ 35x^4 - 84x^5 + 70x^6 - 20x^7, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases} \tag{3.5}$$

Thus, the shearlet in the frequency domain can be written as

$$\widehat{\psi}_{a,s,t}(\omega_1, \omega_2) = a^{\frac{3}{4}} \widehat{\psi}_1(a\omega_1) \widehat{\psi}_2(\frac{\omega_2 + s\omega_1}{\sqrt{a}\omega_1}) e^{-2\pi i(\omega_1, \omega_2)\cdot t} \tag{3.6}$$

$\widehat{\psi}_{a,s,t}$ with different values of $a,\ s$ and t has a different support in the frequency domain, and a shearlet system which is formed by many $\widehat{\psi}_{a,s,t}$ is a Parseval frame for $L^2(\mathbb{R}^2)$ and has compact support in the entire frequency domain. The tiling of the frequency plane of the shearlet system is shown in Fig. 3.1. The gray parts of the shearlet system in Fig. 3.1 are three shearlets in three different directions with three scales.

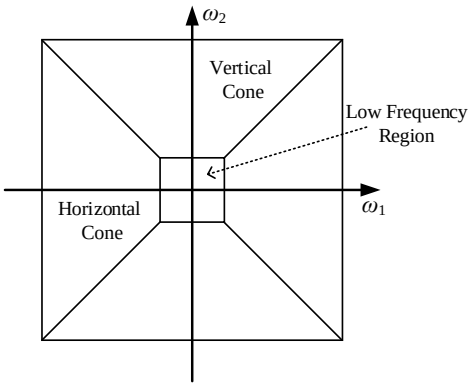


Figure 3.2: Cone-adapted decomposition of a shearlet system in the frequency domain.

For the shearlet transform defined in Eq. (3.6), controlling the shearlet orientation is complex during computation. For ease of calculation, the concept of cone-adapted shearlets was introduced [100], whose construction is based on a partition of the Fourier domain into two cones and a square low frequency region shown in Fig. 3.2. The horizontal cone and vertical cone $\mathcal{C}_c,\ c \in \{h, v\}$ can be defined as

$$\begin{aligned} \mathcal{C}_h &= \left\{(\omega_1, \omega_2) \in \mathbb{R}^2 : \left|\frac{\omega_2}{\omega_1}\right| \leq 1, |\omega_1| > \frac{1}{2}\right\} \\ \mathcal{C}_v &= \left\{(\omega_1, \omega_2) \in \mathbb{R}^2 : \left|\frac{\omega_1}{\omega_2}\right| \leq 1, |\omega_2| > \frac{1}{2}\right\} \end{aligned} \tag{3.7}$$

The central square region in Fig. 3.2 is the low frequency part, which is defined as

$$\left\{(\omega_1, \omega_2) \in \mathbb{R}^2 : |\omega_1| < 1, |\omega_2| < 1\right\} \tag{3.8}$$

For a signal $f \in L^2(\mathbb{R}^2)$, the shearlet transform $\mathcal{SH}(f)$ is defined by

$$\mathcal{SH}^c(f)(a, s, t) = \langle f, \psi^c_{a,s,t}(x)\rangle \tag{3.9}$$

where $\langle f, \psi^c_{a,s,t}(x)\rangle$ is the scalar product in $f \in L^2(\mathbb{R}^2)$ and $c \in \{h, v\}$ is the parameter for the selection of two cones.

3.1.2 Discretisation of Shearlets

In order to deal with discrete signals such as images, the shearlet system should be discretised. There are many different discretisation schemes available [142–145]. In this work, the fast finite shearlet transform (FFST) [142, 143] is adopted, which is constructed in the frequency domain and can provide more computation efficiency. For a digital image I with size $N_1 \times N_2$, the scaling, shearing and translation parameters a, s and t are discretised as

$$\begin{aligned} a_j &= 2^{-2j}, && j = 0, \ldots, j_0 - 1 \\ s^h_{j,k} &= k \times 2^{-j}, && -2^j \leq k \leq 2^j - 1 \\ s^v_{j,k} &= k \times 2^{-j}, && -2^j + 1 \leq k \leq 2^j \\ t_m &= (\tfrac{m_1}{N_1}, \tfrac{m_2}{N_2}), && m \in I \end{aligned} \tag{3.10}$$

where $I = \{(m_1, m_2) \colon m_1 = 0, \ldots, N_1 - 1, m_2 = 0, \ldots, N_2 - 1\}$ and j_0 and k are the number of considered scales and directions respectively. It is worth noting that the number of directions in each scale is 2^{j+2}, and the number of directions in a coarse scale is half the number in the adjacent finer scale (see Fig. 3.1). Therefore, the shearlet in the frequency domain for horizontal cone can be written as

$$\widehat{\psi}^h_{a_j, s_{j,k}, t_m}(x) = 2^{\frac{3j}{4}} \widehat{\psi}_1(2^{-j}\omega_1) \widehat{\psi}_2(2^{\frac{j}{2}} \frac{\omega_2}{\omega_1} - k) e^{\frac{2\pi i (m_1\omega_1 + m_2\omega_2)}{N}} \mathcal{XC}_h(\omega_1, \omega_2) \tag{3.11}$$

where $\mathcal{XC}_c, c \in \{h, v\}$ is a Boolean operation to select cones. Similarly, the shearlet for vertical cone only needs to swap the positions of ω_1 and ω_2. The discrete shearlet transform of an image I is now defined as

$$\mathcal{SH}^c(I)(j, k, m) = \langle I, \psi^c_{a_j, s_{j,k}, t_m}(x) \rangle \tag{3.12}$$

Based on the Plancherel formula, the discrete shearlet transform can be efficiently computed by applying the 2D fast Fourier transform (fft2) and its inverse (ifft2).

Thus, the discrete shearlet transform can be rewritten as

$$\mathcal{SH}(I)(j,k,m) = \begin{cases} \text{ifft2}\Big(\widehat{\psi}(4^{-j}\omega_1, 4^{-j}k\omega_1 + 2^{-j}\omega_2) \cdot \widehat{I}(\omega_1,\omega_2)\Big), & \text{for } h \text{ cone} \\ \text{ifft2}\Big(\widehat{\psi}(4^{-j}\omega_2, 4^{-j}k\omega_2 + 2^{-j}\omega_1) \cdot \widehat{I}(\omega_1,\omega_2)\Big), & \text{for } v \text{ cone} \end{cases} \tag{3.13}$$

where $\widehat{I}(\omega_1,\omega_2)$ is fft2(I). As discussed earlier in this subsection, shearlets have compact support in the entire frequency domain. Therefore, shearlets can be used to obtain multi-scale and multi-directional local structural information for corner detection from images without any loss. For pixels from the flat areas of images, $|\mathcal{SH}(I)(j,k,m)|$ is very small in all directions. For edge points, $|\mathcal{SH}(I)(j,k,m)|$ only reaches its peak in the orthogonal direction of an edge. As for corners, $|\mathcal{SH}(I)(j,k,m)|$ has multiple extreme values. This property can be used to accurately distinguish and detect corners. In addition, the effect of noise, which is not strong enough to observably change the structure characteristics of $|\mathcal{SH}(I)(j,k,m)|$ in all coarse scales, can be reduced effectively by verifying the detected candidate corners based on the shearlet coefficients at coarse scales.

3.1.3 Computational Complexity

For applications in high-level tasks, high computation speed for corner detection is very important. For an image I of size $N \times N$, the computational complexity of the shearlet transform defined in Equation (3.13) is $\mathcal{O}(XN^2 + 2XN^2\log N)$, where X is the number of shearlets used. Although X is proportional to the exponential of the number of considered scales j_0, for high resolution images, the localization of corners is inaccurate at large scales. In other words, j_0 should not be very large. In this case, the computational complexity of the shearlet transform is $\mathcal{O}(N^2\log N)$,

which is the same as the 2D fast Fourier transform.

3.2 Improved Shearlets for Corner Detection

The inadequacies of the traditional shearlet transforms for corner detection are discussed in Section 3.2.1. In Section 3.2.2, these inadequacies are focussed and addressed, and an improved shearlet transform is proposed to accurately extract the local structural information at multiple scales and in multiple directions.

3.2.1 Inadequacies of Shearlets for Corner Detection

As discussed in Section 3.1, the mother shearlet $\psi \in L^2(\mathbb{R}^2)$ is constructed by using $\widehat{\psi}_1$ and $\widehat{\psi}_2$. The usual choice for $\widehat{\psi}_1$ is the Meyer wavelet as shown in Fig. 3.3(a), and $\widehat{\psi}_2$ is a non-zero square-integrable function which is defined by Eq. (3.4). The generation of the 2D shearlet can be described simply by rotating the one-dimensional wavelet $\widehat{\psi}_1$ in a two-dimensional coordinate system. The rotation allows the wavelet to have limited support in its orthogonal direction and make the wavelet anisotropic. The angle of rotation is determined by supp$\widehat{\psi}_2$.

The frequency tiling of the traditional shearlet transform with 3 scales in the Fourier domain are shown in Fig. 3.3(b), and one of the shearlets in the Fourier and spatial domain is shown in Fig. 3.3(c) and Fig. 3.3(d) respectively. In Fig. 3.4, the absolute values of shearlet coefficients in each direction from an image are added in 3 scales to measure the ability of the shearlet transform to obtain local structural information [108]. It is worth noting that all the resultant images in Fig. 3.4 are complement images for convenience of viewing. As shown in Fig. 3.3(a), the

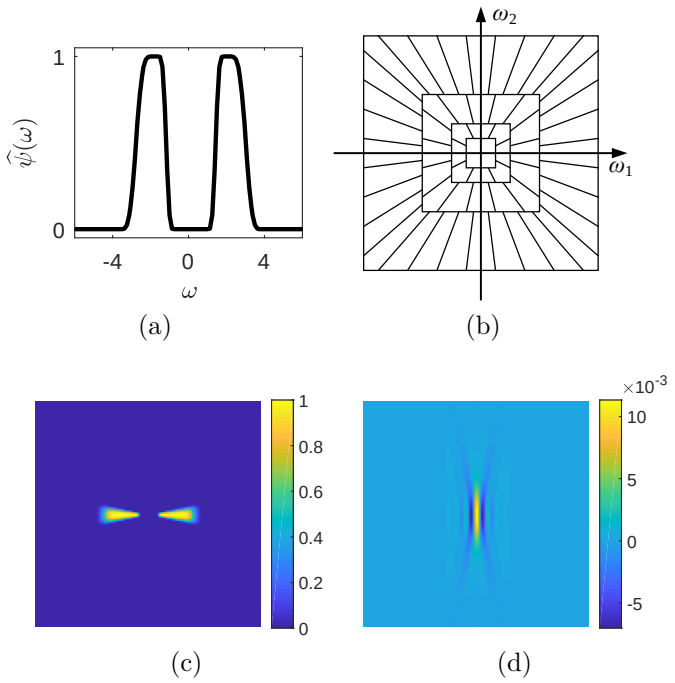
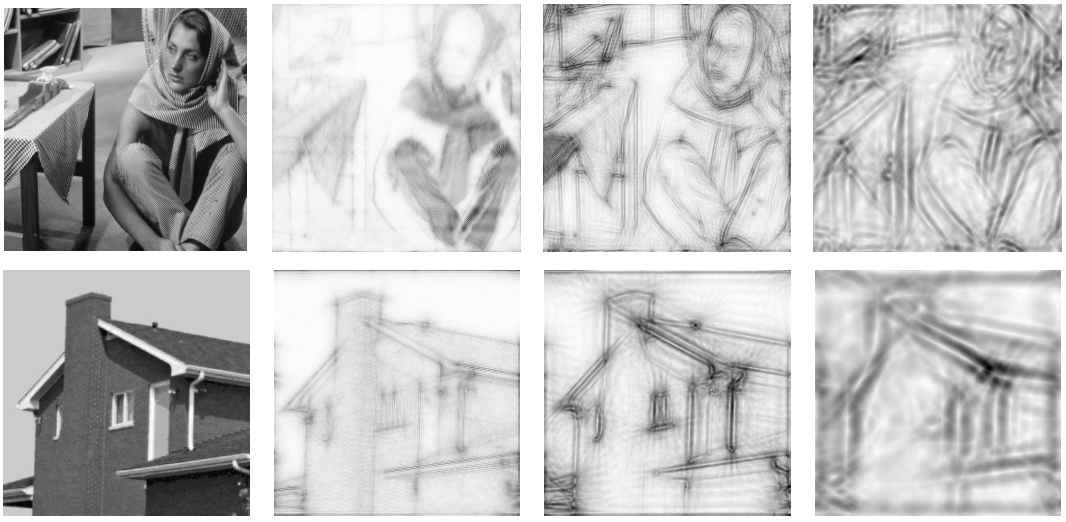


Figure 3.3: The traditional shearlet transform. (a) Function envelope of $\widehat{\psi}_1$, i.e., Meyer wavelet mother function, (b) frequency tiling for 3 scales in the Fourier domain, (c) shearlet for $j = 1$ and $k = 0$ in the Fourier domain, (d) the corresponding version in the spatial domain.

Meyer wavelet is a band-limited function where the speed of decay is very fast in the Fourier domain. This will lead to a strong side-lobe effect in the spatial domain (see Fig. 3.3(d)), causing multiple responses at locations where the grayscale values drastically change, such as along edges and at corners. The structural information obtained by the traditional shearlet transform in Fig. 3.4 also confirms this. Therefore, duplicate detections may occur for a corner.

As can be seen from Fig. 3.3(b), in the traditional shearlet transform, there is no overlap between the shearlets of adjacent directions. The advantage of this property is that the shearlets can provide optimally sparse representations for anisotropic functions, and this is very advantageous for image compression and sparse representation. However, the non-overlapping shearlets cannot provide sufficient structural



(a) (b) (c) (d)

Figure 3.4: Original images and resultant images of local structural information obtained by the traditional shearlets with 3 scales. (a) Original images, (b) resultant images of local structural information with $j = 2$, (c) resultant images of local structural information with $j = 1$, (d) resultant images of local structural information with $j = 0$.

information to detect corners. In other words, because a corner is not simply formed by straight lines and there might be intensity changes in several directions in its background, the structural information in two adjacent directions is not completely independent. Therefore, the relevance of structural information in all adjacent directions should be considered.

Since $\widehat{\psi}_2$ is a non-zero square-integrable function with $\text{supp}\widehat{\psi}_2 \subseteq [-1, 1]$, the mother function of shearlets has limited support in the Fourier domain, and consequently the corresponding spatial filter has a larger envelope. In areas around edges, this leads to the structural information being extended at the end of the edges, namely the problem of edge extension, which causes erroneous detections in the case of extensions of two edges intersecting with each other. Furthermore, a filter with limited bandwidth in the Fourier domain and a larger envelope in the spatial do-

main depresses the high frequency components, making the detailed information extracted from images less clear. It can be observed that the structural information in Fig. 3.4 is unclear and has multiple peaks with the problem of edge extension. These problems will greatly affect the subsequent corner detection.

It can be seen from Eq. (3.10) and Fig. 3.3(b) that the number of directions in scale j is 2^{j+2}. In other words, the orientation variable s is associated with the scale index j. Thus, the shearlet decomposition at different scales has different number of directions, and one direction of a coarse scale corresponds to two directions of the adjacent finer scale (see Fig. 3.3(b)). At the coarse scales, shearlets cannot provide sufficiently clear local structural information, and fewer directions can effectively reduce the number of shearlet coefficients. This is very important for image compression and sparse representation. However, for corner detection, a unified detection framework that can be easily applied to multiple scales is very important. Keeping the same number of directions in the shearlet coefficients of different scales is very beneficial to make full use of multi-scale information for corner detection and reduce the difficulty of algorithm design.

3.2.2 Improved Shearlets for Corner Detection

In order to overcome the problem of multiple peaks resulting from the side-lobe effect, in the improved shearlets, an odd wavelet proposed by Mallat and Zhong [146] with an analytical expression in the Fourier domain is selected to replace the Meyer wavelet, which is defined as

$$\widehat{\psi}_1(\omega) = i\omega \left(\frac{\sin(\frac{\omega}{4})}{\frac{\omega}{4}} \right)^{2n+2} \tag{3.14}$$

As shown in Fig. 3.5(a), the decay speed of this wavelet is much slower than the

Meyer wavelet. Accordingly, the side-lobe effect in the spatial domain can be depressed.

To overcome the problem of edge extension and make the structural information of adjacent directions interrelated, $\widehat{\psi}_2$ is redefined as

$$\widehat{\psi}_2(\omega) = \begin{cases} \frac{1}{b}\sqrt{D(1+\omega/b)}, & \omega \leq 0 \\ \frac{1}{b}\sqrt{D(1-\omega/b)}, & \omega > 0 \end{cases} \tag{3.15}$$

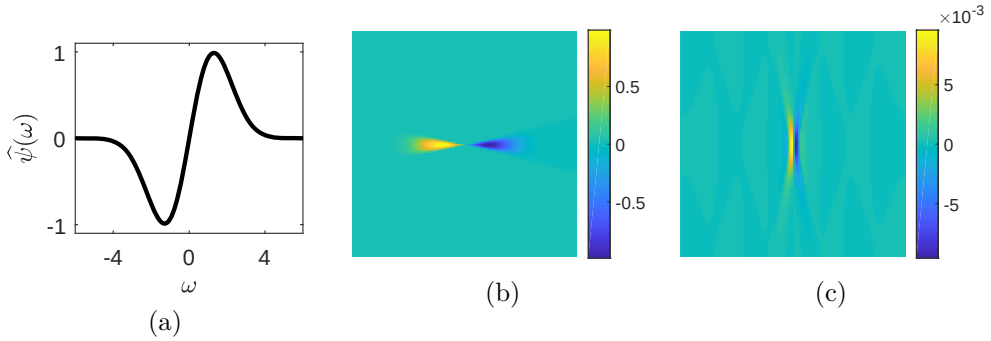
where $b \in \mathbb{R}^+$ and $D(x)$ is the auxiliary function defined in Eq. (3.5). It is easy to see from Eq. (3.4) and Eq. (3.15) that the new $\widehat{\psi}_2$ has a flexible support $\text{supp}\widehat{\psi}_2 \subseteq [-b, b]$ in the Fourier domain, which can be easily adjusted by the value of b. Therefore, the shearlet in Eq. (3.6) can be rewritten as

$$\widehat{\psi}_{a,s,t}(\omega_1,\omega_2) = \begin{cases} \dfrac{i\omega_1}{ba^{\frac{4}{7}}}\sqrt{D(1+\dfrac{\omega_2+s\omega_1}{\sqrt{a}b\omega_1})}\left(\dfrac{\sin(\frac{a\omega_1}{4})}{\frac{a\omega_1}{4}}\right)^{2n+2} e^{-2\pi i\langle(\omega_1,\omega_2),t\rangle}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \dfrac{\omega_2+s\omega_1}{\sqrt{a}\omega_1} \leq 0 \\ \dfrac{i\omega_1}{ba^{\frac{4}{7}}}\sqrt{D(1-\dfrac{\omega_2+s\omega_1}{\sqrt{a}b\omega_1})}\left(\dfrac{\sin(\frac{a\omega_1}{4})}{\frac{a\omega_1}{4}}\right)^{2n+2} e^{-2\pi i\langle(\omega_1,\omega_2),t\rangle}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \dfrac{\omega_2+s\omega_1}{\sqrt{a}\omega_1} > 0 \end{cases} \tag{3.16}$$

In order to make the improved shearlet transform provide a uniform number of directions for different scales, the parameter s in Eq. (3.10) is re-defined as

$$\begin{aligned} s^h_{j,k} &= k2^{-j} \times \frac{2^{j+2}}{K} = \frac{4k}{K}, \qquad -2^p \leq k \leq 2^p - 1 \\ s^v_{j,k} &= k2^{-j} \times \frac{2^{j+2}}{K} = \frac{4k}{K}, \qquad -2^p + 1 \leq k \leq 2^p \end{aligned} \tag{3.17}$$

where p is a nonnegative integer and $K = 2^{p+2}$ which is the number of directions for all scales. It is easy to see from Eq. (3.17) that for the continuous orientation variable



(a) (b) (c)

Figure 3.5: The modified shearlet transform constructed by using the odd wavelet defined in Eq. 3.14 as $\widehat{\psi}_1$ instead of the Meyer wavelet. (a) Function envelope of the odd wavelet $\widehat{\psi}_1$, (b) shearlet of the modified shearlet transform for $j = 1$ and $k = 0$ in the Fourier domain, (c) the corresponding version in the spatial domain.

s, the number of sampling points can be adjusted by K which is independent with scale index j. However, for the coarse scales such as $j = 0$, too many directional divisions will lead to an excessive overlap of shearlets in adjacent directions, so the support of $\widehat{\psi}_2$ should be adjusted in accordance with K and scale index j. Therefore, $\widehat{\psi}_2$ is finally rewritten as

$$\widehat{\psi}_2(\omega) = \begin{cases} \frac{K}{2^{j+2}b}\sqrt{D(1+\frac{K\omega}{2^{j+2}b})}, & \omega \leq 0 \\ \frac{K}{2^{j+2}b}\sqrt{D(1-\frac{K\omega}{2^{j+2}b})}, & \omega > 0 \end{cases} \tag{3.18}$$

In this way, the support of each shearlet at all scales has an appropriate size for corner detection, and every directional shearlet at each scale corresponds to one

(a) (b) (c) (d)

Figure 3.6: Original images and resultant images of local structural information provided by the modified shearlet transform with 3 scales. (a) Original images, (b) resultant images of obtained structural information with $j = 2$, (c) resultant images of obtained structural information with $j = 1$, (d) resultant images of obtained structural information with $j = 0$.

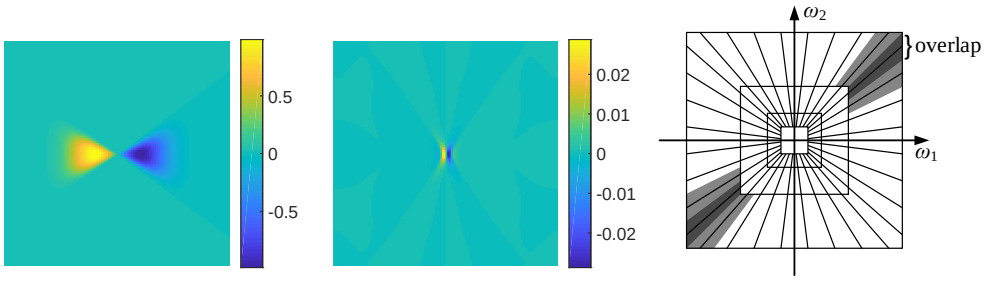
directional shearlet in other scales. Therefore, Eq. (3.11) can be rewritten as

$$\widehat{\psi}^{h}_{j,k,m}(\omega_1,\omega_2) = 2^{\frac{-3j}{2}}\widehat{\psi}_1(4^{-j}\omega_1)\widehat{\psi}_2(2^j\frac{K\omega_2+4k\omega_1}{K\omega_1})e^{\frac{-2\pi i(m_1\omega_1+m_2\omega_2)}{N}}\mathcal{XC}_h(\omega_1,\omega_2)$$

$$= \begin{cases} 2^{\frac{-9j-4}{2}}\dfrac{iK\omega_1}{b}\sqrt{D(1+\dfrac{K\omega_2+4k\omega_1}{4b\omega_1})}\left(\dfrac{\sin(4^{-j-1}\omega_1)}{4^{-j-1}\omega_1}\right)^{2n+2} \\ \qquad e^{-2\pi i(\frac{m_1\omega_1}{N_1}+\frac{m_2\omega_2}{N_2})}\mathcal{XC}_h(\omega_1,\omega_2), \quad \dfrac{K\omega_2+4k\omega_1}{K\omega_1} \le 0 \\ 2^{\frac{-9j-4}{2}}\dfrac{iK\omega_1}{b}\sqrt{D(1-\dfrac{K\omega_2+4k\omega_1}{4b\omega_1})}\left(\dfrac{\sin(4^{-j-1}\omega_1)}{4^{-j-1}\omega_1}\right)^{2n+2} \\ \qquad e^{-2\pi i(\frac{m_1\omega_1}{N_1}+\frac{m_2\omega_2}{N_2})}\mathcal{XC}_h(\omega_1,\omega_2), \quad \dfrac{K\omega_2+4k\omega_1}{K\omega_1} > 0 \end{cases} \tag{3.19}$$

Similarly, the shearlet for the vertical cone only needs to swap the positions of ω_1 and ω_2.

In contrast with the traditional shearlet, it can be seen from Fig. 3.3(d) and Fig. 3.5(c)



(a) (b) (c)

Figure 3.7: The improved shearlet transform with $b = 3$. (a) An improved shearlet for $j = 1$ and $k = 0$ in the Fourier domain, (b) the corresponding version in the spatial domain, (c) diagram of directions with 3 scales and 16 directions in the Fourier domain.

(a) (b) (c) (d)

Figure 3.8: Original images and resultant images of local structural information obtained by the improved shearlets with 3 scales. (a) Original images, (b) resultant images of obtained structural information with $j = 2$, (c) resultant images of obtained structural information with $j = 1$, (d) resultant images of obtained structural information with $j = 0$.

that, by replacing the Meyer wavelet with Eq. (3.14), the side-lobe effect in the spatial domain is effectively suppressed. As a consequence, it can be seen from Fig. 3.6 that compared with Fig. 3.4, the problem of multi-peak responses is overcome and the obtained structural information at each scale is clearer.

Moreover, according to Eq. (3.18), by redefining $\widehat{\psi}_2$, the envelope size can be controlled by changing the value of b. It can be seen from Fig. 3.7(a) that the envelope of the improved shearlet in the Fourier domain is significantly enlarged, which can extract more high frequency information in each direction. Accordingly its envelope in the spatial domain has been significantly reduced and the extension of structural information can be effectively depressed. Moreover, an adjustable and larger bandwidth of shearlet in the Fourier domain means that shearlets in adjacent directions will have an adjustable overlap area, and the relevance of structural information in all adjacent directions can be considered. It can be seen from Fig. 3.7(c) that the envelope of the improved shearlet is b times larger than that of the traditional shearlet, and the overlap size between the adjacent shearlets is $b-1$. As can be observed from Fig. 3.8, by redefining $\widehat{\psi}_2$, the problem of extension of structural information has been effectively alleviated, and the obtained structural information at each scale is much clearer compared with Fig. 3.5 and Fig. 3.3.

Furthermore, according to Eq. (3.17) and Eq. (3.19), the number of directions K is independent of the scale index j, and the number of directions in each scale can be adjusted and kept consistent with other scales (see Fig. 3.7(c)). Obviously, the same number of directions has a benefit for subsequent corner detection by using multi-scale directional information.

3.3 Multiple Directional Corner Detection and Measurement Algorithm

As discussed in Section 2.4, current multi-scale analysis based corner detectors usually utilise the results of multi-scale decomposition for corner detection in two ways.

One is to establish a two-directional structure tensor based on the weighted summation of the structural coefficients in all directions, and the second is to select the perpendicular direction of the maximum response in multi-scale decomposition or the weighted summation of all directional means of all scale shearlet coefficients. Obviously, these corner detectors do not make full use of the multi-scale and multi-directional structural information. The two-directional structure tensor cannot effectively distinguish corners from pixels with complex backgrounds or on curves with large curvatures. Moreover, the merging or selection process will lead to loss of some useful structural information and increase the corner localisation error.

In this thesis, a novel multi-directional structure tensor with multi-directional shearlet coefficients is constructed for corner detection, and a multi-scale corner measurement function is developed to remove false candidate corners by using multi-scale directional information. The proposed multi-directional structure tensor obtains structural information by moving a detection patch in the input image, and detects corners by finely mining the local intensity changes in multiple directions and scales. For an image I, after placing the multi-directional detection patch over a circular area around (x, y) and shifting it by Δt, the intensity change $E(\Delta t)$ of the response of the detection patch can be defined as

$$E(\Delta t) = \frac{1}{K(w+1)^2} \sum_{x=-\frac{w}{2}}^{\frac{w}{2}} \sum_{y=-\frac{w}{2}}^{\frac{w}{2}} \sum_{k=1}^{K} c(x,y) \left[I(x + \Delta t \cos\theta_k, y + \Delta t \sin\theta_k) - I(x,y) \right]^2 \tag{3.20}$$

with a square patch having width $w+1$ centred at (x, y). $c(x, y)$ is a circular mask where the values within the circle are 1 and 0 outside, K is the number of directions and θ_k is the angle between the horizontal axis and the k-th direction. In Eq. (3.20),

the term $I(x+\Delta t\cos\theta_k, y+\Delta t\sin\theta_k)$ can be simplified by the Taylor function to

$$I(x+\Delta t\cos\theta_k, y+\Delta t\sin\theta_k) = I(x,y) + \Delta t I_x\cos\theta_k + \Delta t I_y\sin\theta_k + O(\Delta t^2\cos\theta_k^2, \Delta t^2\sin\theta_k^2) \tag{3.21}$$

where $I_x\cos\theta_k$ and $I_y\sin\theta_k$ are the directional derivatives along direction k. Then Eq. (3.20) can be rewritten as

$$\begin{aligned} E(\Delta t) =& \frac{1}{K(w+1)^2}\sum_{x=-\frac{w}{2}}^{\frac{w}{2}}\sum_{y=-\frac{w}{2}}^{\frac{w}{2}}\sum_{k=1}^{K} c(x,y)[I(x,y) + \Delta t I_x\cos\theta_k + \Delta t I_y\sin\theta_k + \\ & O(\Delta t^2\cos\theta_k^2, \Delta t^2\sin\theta_k^2) - I(x,y)]^2 \\ =& \frac{1}{K(w+1)^2}\sum_{x=-\frac{w}{2}}^{\frac{w}{2}}\sum_{y=-\frac{w}{2}}^{\frac{w}{2}}\sum_{k=1}^{K} c(x,y)[\Delta t I_x\cos\theta_k + \Delta t I_y\sin\theta_k + \\ & O(\Delta t^2\cos\theta_k^2, \Delta t^2\sin\theta_k^2)]^2 \\ \approx& \frac{1}{K(w+1)^2}\sum_{x=-\frac{w}{2}}^{\frac{w}{2}}\sum_{y=-\frac{w}{2}}^{\frac{w}{2}}\sum_{k=1}^{K} c(x,y)[\Delta t(I_x\cos\theta_k + I_y\sin\theta_k)]^2 \end{aligned} \tag{3.22}$$

It is worth noting that

$$I_x\cos\theta_k + I_y\sin\theta_k \simeq \mathcal{SH}(I)(j,k,m) \tag{3.23}$$

Therefore, Eq. (3.22) can be rewritten as

$$\begin{aligned} E(\Delta t) \approx & \frac{1}{K(w+1)^2} \sum_{x=-\frac{w}{2}}^{\frac{w}{2}} \sum_{y=-\frac{w}{2}}^{\frac{w}{2}} c(x,y) \\ & \left([\mathcal{SH}(I)(j,1,m), \mathcal{SH}(I)(j,2,m), \ldots, \mathcal{SH}(I)(j,K,m)] \begin{bmatrix} \Delta t \\ \Delta t \\ \vdots \\ \Delta t \end{bmatrix} \right)^2 \\ = & \frac{1}{K(w+1)^2} [\Delta t, \Delta t, \ldots, \Delta t] M \begin{bmatrix} \Delta t \\ \Delta t \\ \vdots \\ \Delta t \end{bmatrix} \end{aligned} \tag{3.24}$$

where M is defined as

$$M = \sum_{x=-\frac{w}{2}}^{\frac{w}{2}} \sum_{y=-\frac{w}{2}}^{\frac{w}{2}} c(x,y) \begin{bmatrix} \mathcal{SH}_{1,j}^2(I) & \mathcal{SH}_{1,j}(I)\mathcal{SH}_{2,j}(I) & \ldots & \mathcal{SH}_{1,j}(I)\mathcal{SH}_{K,j}(I) \\ \mathcal{SH}_{2,j}(I)\mathcal{SH}_{1,j}(I) & \mathcal{SH}_{2,j}(I)^2 & \ldots & \mathcal{SH}_{2,j}(I)\mathcal{SH}_{K,j}(I) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{SH}_{K,j}(I)\mathcal{SH}_{1,j}(I) & \mathcal{SH}_{K,j}(I)\mathcal{SH}_{2,j}(I) & \ldots & \mathcal{SH}_{K,j}(I)^2 \end{bmatrix} \tag{3.25}$$

where $\mathcal{SH}_{k,j}(I) = \mathcal{SH}(I)(j,k,m)$. M is a structure tensor with K directions at scale j which is a symmetric $K \times K$ matrix. Since the magnitudes of the eigenvalues of M can represent the intensity changes of a pixel in K directions, the K eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_K$ of M can be used to form a new corner detection function to distinguish corners from other pixels. For each image pixel (x, y), the corner response

function R is defined by

$$R = \frac{\prod_{k=1}^{K} \lambda_k}{\sum_{k=1}^{K} \lambda_k + q} \tag{3.26}$$

where q is the minimum float number in a specific programming system, e.g., $q = 2.2204 \times 10^{-16}$ in MATLAB, which is used to avoid zero denominator. For each image pixel, if the value of the corner response function R is larger than threshold T_1, this pixel can be considered as a candidate corner.

For an edge pixel, the structural information only reaches its peak in the orthogonal direction of the edge. In other words, the structure tensor M of the edge pixel has one large eigenvalue while the other eigenvalues are small. In contrast, a corner will have two or more large eigenvalues. This property can be used to better distinguish corners from edge pixels. Meanwhile, the effect of noise, which is not strong enough to observably change the structure characteristics of eigenvalues in all coarse scales, can be reduced effectively by verifying the detected candidate corners at all scales. In the proposed method, a corner measurement function V is proposed to remove false detections from candidate corners, which is defined as

$$V = \frac{\sum_{k=1}^{K} \lambda_k}{\max(\lambda_k)} \tag{3.27}$$

For a candidate corner, if the value of the corner measurement function V is larger than threshold T_2 at all scales, the candidate corner can be confirmed as a corner.

The proposed corner detection algorithm is summarised in Algorithm 3.1.

Algorithm 3.1 Proposed Shearlet-based Multi-directional Corner Detection Algorithm.

1: Use the improved shearlet transform to obtain local structural information from images at 3 scales in K directions.

2: For each pixel, compute the multi-directional structure tensor (see Eq. (3.25)) at the finest scale.

3: Obtain the eigenvalues of the structure tensor for all pixels and compute the corner responses based on Eq. (3.26).

4: If the response of a pixel is a local maximum within a 7×7 area and is larger than T_1, mark this pixel as a candidate corner.

5: Check all candidate corners based on Eq. (3.27). For each candidate corner, if the corner measure is larger than T_2 at all scales, then the candidate corner is confirmed as a corner.

3.4 Experimental Results and Performance Analysis

Experimental assessments on the proposed corner detector were performed. The detection performance, robustness to affine transformations, illumination changes, noise, image blurring, viewpoint changes and JPEG compression, and computational cost were evaluated and compared with representative corner detectors such as Harris [24], Hessian-Laplace [85], FAST [53], ORB [131], CF [81], ANDD [80], ACJ [79], LGWTSMM [93], SMCD [108] and Zhang and Sun's method [94]. Representative interest point detectors such as DoG [136], SIFT [137], SURF [147], deep learning based method LIFT [110] and shearlet blob detector (SBD) [135] were also used for robustness evaluation.

3.4.1 Evaluation Metrics and Datasets

The evaluation metrics and datasets for detection performance, localisation accuracy and repeatability for corner detectors are introduced in Sections 3.4.1.1 and 3.4.1.2.

3.4.1.1 Metrics and Datasets for Detection Precision and Localisation Accuracy

In detection performance evaluation, the key performance indicators including detection precision and corner localisation accuracy were considered. For each corner (x_i, y_i) of the ground truth, if a corner $(\widetilde{x}_i, \widetilde{y}_i)$ can be detected within a distance of 4 pixels, the two corners are regarded as a matched pair. Otherwise, the corner (x_i, y_i) is counted as a missed corner. Similarly, for each detected corner $(\widetilde{x}_i, \widetilde{y}_i)$, if a corner (x_i, y_i) can be found in the ground truth within a distance of 4 pixels, the two corners are regarded as a matched pair. Otherwise, the corner $(\widetilde{x}_i, \widetilde{y}_i)$ is counted as a false corner. For all the match pairs, the localisation error L is defined as the average distance of the matched pairs, which is calculated by

$$L = \frac{1}{P}\sum_{i=1}^{P}\sqrt{(\widetilde{x}_i - x_i)^2 + (\widetilde{y}_i - y_i)^2} \tag{3.28}$$

where P is the number of the matched pairs. Let N_m denote the number of missed corners, and N_f be the number of false corners. Suppose there are N_p labeled corners, then *F-score* is defined as [2]

$$F\text{-}score = \frac{2 \times precision \times recall}{precision + recall} \tag{3.29}$$

where *precision* is given by

$$precision = \frac{N_p - N_m}{N_p - N_m + N_f} \tag{3.30}$$

and *recall* is defined as

$$recall = \frac{N_p - N_m}{N_p} \tag{3.31}$$

The *precision*, *F-score* and localisation error L were used as the metrics of detection accuracy.

For detection accuracy and corner localisation accuracy evaluation, a total of 20 different images with ground truth including 7 real images and 13 synthetic images were collected from different sources [39, 79–81, 93, 94, 148, 149], parts of which are shown in Fig. 3.9.

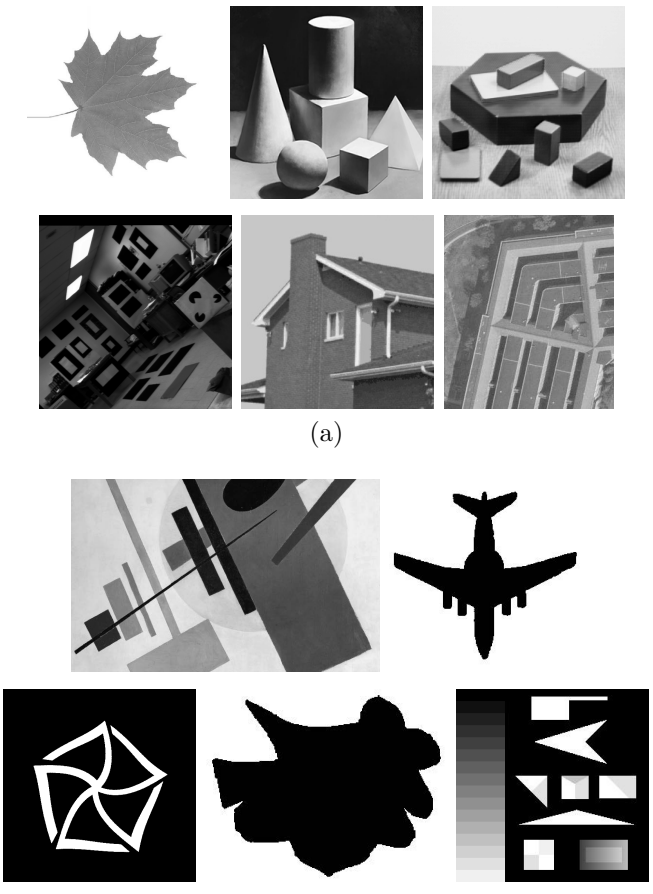


Figure 3.9: Sample images (a) from the 7 real images and (b) from the 13 synthetic images.

3.4.1.2 Metrics and Datasets for Repeatability

In evaluating the robustness to affine transformations, illumination changes, noise and JPEG compression, the repeatability score RS on a given image sequence is regarded as the metric, which is defined as [108]

$$RS_i = \frac{|CR_{1i}|}{\min(|C_1|, |C_i|)} \tag{3.32}$$

where C_i is the set of detected corners in image I_i, and CR_{1i} is the set containing all the corner correspondences between image I_1 and the transformed image I_i. To calculate CR_{1i}, the regions representing the corners are mapped from one image to the other using homography. Then their overlap with a corner region from the other image by more than 60% was checked. Since repeatability score RS can reflect the potential of all interest point detectors to be extended into local feature descriptors, representative interest point detectors such as DoG [136], SIFT [137], SURF [147], deep learning based method LIFT [110] and shearlet blob detector (SBD) [135] were also used to compare robustness evaluation.

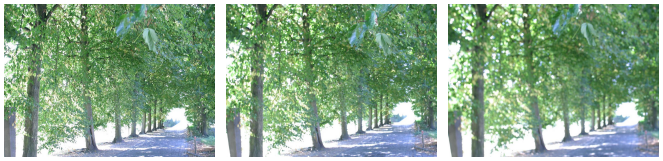
In evaluating the robustness to affine transformations, illumination changes, noise and JPEG compression, three datasets were considered. Some of the images are shown in Fig. 3.10 and Fig. 3.11. Dataset I, namely Oxford Affine Dataset [94, 108, 135], consists of 36 real images forming 6 image sequences including two sequences with blur acquired by varying the camera focus, one sequence with illumination changes obtained by varying the camera aperture, two sequences with viewpoint changes captured by varying the viewpoint of a fronto-parallel view camera from 20° to 60° with an increment of 10°, and one sequence with JPEG compression produced by setting the quality factor as 40, 20, 10, 5, and 2 respectively.

Dataset II partly used elsewhere [80, 94] contains 1,800 images transformed from 50

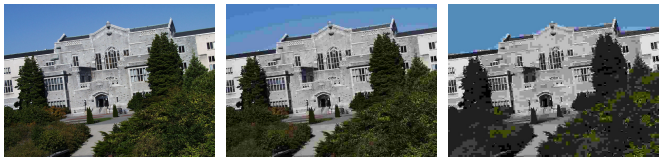
(a) Bikes

(b) Graffiti

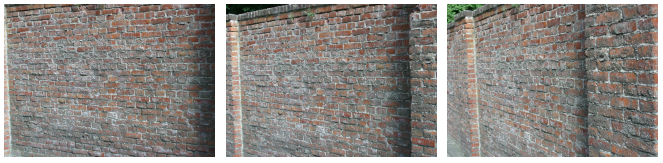
(c) Leuven



(d) Tree



(e) Ubc



(f) Wall

Figure 3.10: Sample frames of Dataset I produced by image blurring, illumination changes, viewpoint changes and JPEG compressions. (a) 'Bikes' (blur), (b) 'Graffiti' (viewpoint change), (c) 'Leuven' (illumination change), (d) 'Tree' (blur), (e) 'Ubc' (JPEG compression), (f) 'Wall' (viewpoint change).

Figure 3.11: Sample images of Dataset II.

natural images, which are combined into six different styles of image sequences. The blur sequence and illumination change sequence were obtained by using Gaussian blur and Gamma correction to deal with the 50 natural images respectively, the standard deviation of Gaussian blur varies from 1 to 3 with an increment of 0.5 and the Gamma changes from 1.5 to 2.3 with an increment of 0.2. Each natural image was transformed with rotation and scale transforms, with the rotation angle varying from 15° to 165° with an increment of 15° and the scale factor changing from 1.2 to 2 with an increment of 0.2, producing the rotation and scaling sequences respectively. The viewpoint change sequence was acquired by using homography transformation to simulate the viewpoint varying from 20° to 60° with an increment of 10°.

Dataset III contains three types of image sequences formed by 840 images which were generated from Datasets I and II by adding three types of noise in the following manner. For the per image sequence, the first frame was corrupted by Gaussian noise with zero mean, and the standard deviation was set to 7, 9, 11, 13, and 15 respectively, forming 56 image sequences. Similarly, another 112 image sequences

were generated, 56 sequences of which were corrupted by salt and pepper noise with density varying from 0.03 to 0.15 with an increment of 0.03. The others were corrupted by multiplicative speckle noise with normalised standard deviation varying from 0.12 to 0.2 with an increment of 0.02.

3.4.2 Setting of Parameters

Among the 20 test images with ground truth shown in Fig. 3.9, two real images and four synthetic images were selected and used for setting the parameters of the proposed and compared corner detectors. For each method, its threshold was set to the value at which the detector achieves the maximum average *F-score*.

For threshold T_1 of the proposed corner detector, step 5 in Algorithm 3.1 was bypassed, and the candidate corners determined by the first four steps with $b = 3.4$, $K = 16$, and T_1 varying from 0 to 10^{25} were regarded as the resultant corners. The average *recall* on the six images is summarised in Fig. 3.12(a). As can be seen, most of the marked corners can be correctly detected until T_1 reaches 10^{14}. Therefore, in subsequent experiments, T_1 was set to $10^{13.5}$. For threshold T_2, all steps in Algorithm 3.1 were used, with $T_1 = 10^{13.5}$ and T_2 varying from 1 to 1.3 with an increment of 0.01. The corresponding average *precision* and *F-score* on the six images are summarised in Fig. 3.12(b) and Fig. 3.12(c) respectively. From the two figures, it can be seen that as T_2 increases, the false corners have been effectively picked out from the candidate corners, and the average *precision* and *F-score* increase as well. When T_2 is beyond 1.11, beside the false corners, some marked corners have also been picked out, so the average *precision* continues to increase and the *F-score* begins to decline. Therefore, in the detection performance evaluation experiments, T_2 is set to 1.11. Note that when T_2 approaches 1.3, most

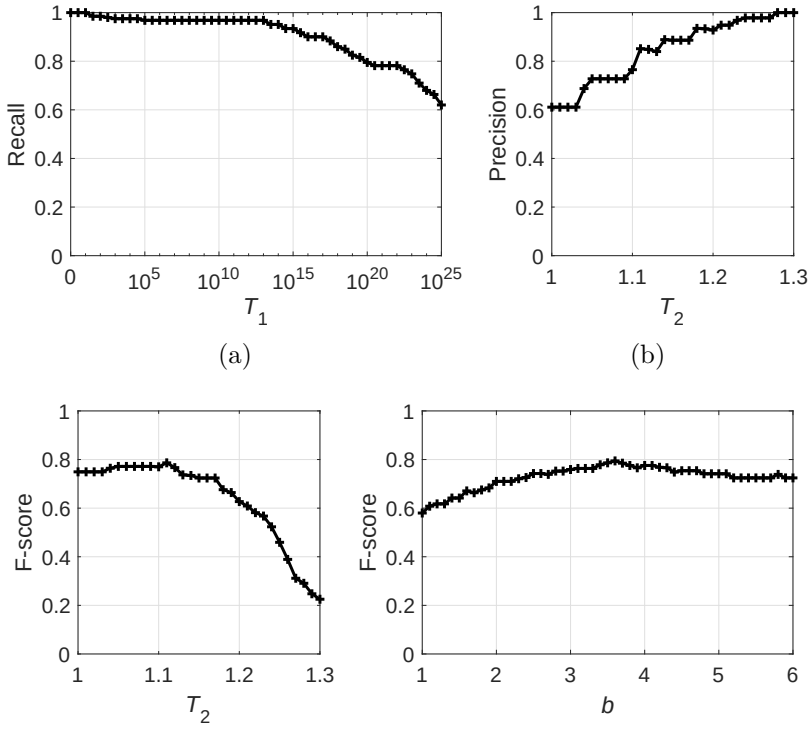


Figure 3.12: Setting of parameters. (a) *Recall* curve on threshold T_1, (b) *Precision* curve on threshold T_2, (c) *F-score* curve on threshold T_2, (d) *F-score* curve on parameter b.

of the false corners have been eliminated from the candidate corners, and the corresponding *precision* approaches 1. This demonstrates that in the proposed detector, by adjusting parameter T_2, the number of false and missed corners can be tuned for different applications. With $T_1 = 10^{13.5}$, $T_2 = 1.11$ and b varying from 1 to 6 with an increment of 0.1, the average *F-score* of the proposed detector on the six images is given in Fig. 3.12(d). As can be seen from this figure, for a smaller value of b, the extensions of edges will induce many false detection corners, which results in a low *F-score*. As b increases, the extensions of edges have been effectively depressed and the *F-score* increases as well. When b is beyond 3.6, beside the extensions of edges, the differences between the responses on edges and corners are been depressed, and

F-score begins to decline. Therefore, in the subsequent experiments b is set to 3.6.

The threshold parameter for each compared corner detector was also finetuned to make its average *F-score* on the six images to be the best, and the corresponding threshold value was determined and fixed in the subsequent detection performance evaluation experiments. However, the aim of the robustness evaluation experiments was to verify the tolerance of each method to affine transformations, illumination changes, noise and JPEG compression, rather than detection accuracy. For the sake of fairness, it is important to ensure that all detectors achieve a similar number of corner detections for the first frame of each image sequence. Therefore, in robustness evaluation experiments, for each image sequence, the thresholds of all the compared methods and the proposed detector were adjusted to obtain a similar number of corners detected on the first frame.

3.4.3 Evaluation of Detection Performance

To verify the detection precision and corner localisation accuracy of the proposed corner detector, *precision*, *F-score* and localisation error L defined in Eq. (3.28), Eq. (3.29) and Eq. (3.30) were used as the metrics. Except the six images used in Section 3.4.2, the remaining 14 images with ground truth including five real images and nine synthetic images were considered. Representative corner detectors such as Harris [24], FAST [53], CF [81], ANDD [80], ACJ [79], LGWTSMM [93], SMCD [108] and Zhang and Sun's method [94] were used for comparison.

Detection results on two typical images are shown in Fig. 3.13 and Fig. 3.14, and the average and maximum *precision*, *F-score*, and localisation error on all the 14 images are summarised in Table 3.1. From the results, several observations can be

made:

1. The average *precision*, average *F-score*, and maximum *F-score* of the proposed detector are evidently higher than those of other detectors. This indicates that for most images, the proposed method outperforms current corner detectors in detection accuracy. Since ANDD [80] and CF [81] methods detect corners based on edge contours, for the second synthetic image in Fig. 3.9(b) which has distinct edge contours, they could achieve a good performance. Therefore only the maximum *precision* of the two methods is slightly higher than that of the proposed method.

2. The corner localisation accuracy of the proposed method is relatively higher than that of the other methods. This is because the proposed method performs corner detection on the finest scale and there is no down-sampling operation in the improved shearlet transform. The relatively high localisation error of

Table 3.1: Detection performance and localization accuracies.

Detectors	*Precision* (%)		*F-score* (%)		Localization error
	Avg.	Max.	**Avg.**	Max.	
Harris [24]	60.59	85.45	63.91	80.99	1.538
FAST [53]	53.93	86.49	61.56	75.81	1.644
CF [81]	74.55	96.77	64.20	77.23	1.526
ANDD [80]	76.37	**96.88**	73.48	93.33	1.446
ACJ [79]	66.87	87.10	67.86	87.72	2.173
LGWTSMM [93]	73.88	92.00	71.12	84.40	1.944
SMCD [108]	70.73	84.91	66.44	80.36	2.071
Zhang and Sun [94]	83.10	94.44	74.10	84.40	1.748
Proposed (K=4)	85.86	93.94	76.02	90.60	1.427
Proposed (K=8)	86.48	94.59	77.37	95.89	1.433
Proposed (K=16)	**88.43**	96.30	**78.08**	**97.30**	**1.414**

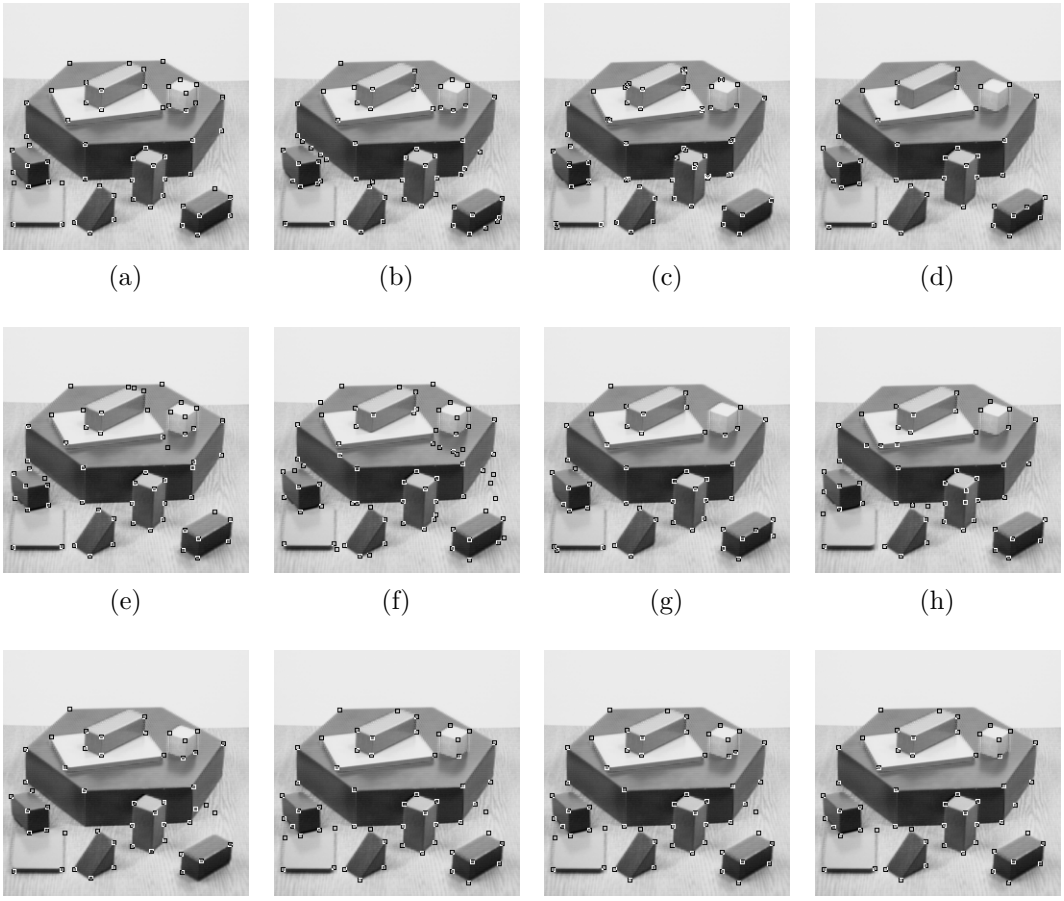


Figure 3.13: Detection results of the proposed method and the compared corner detectors on the image 'Block'. (a) Ground truth, (b) Harris, (c) FAST, (d) CF, (e) ANDD, (f) ACJ, (g) LGWTSMM, (h) SMCD, (i) Zhang and Sun's method, (j) proposed method (K=4), (k) proposed method (K=8), (l) proposed method (K=16).

ACJ detector [79] is the result of the down-sampling operation in the line segment detector. Since merging information from coarse scales in the multi-scale analysis based corner detectors such as LGWTSMM [93] and SMCD [108] will increase localisation errors, their localisation errors are higher than the other methods except for ACJ [79].

3. Compared with current multi-scale analysis based corner detectors especially

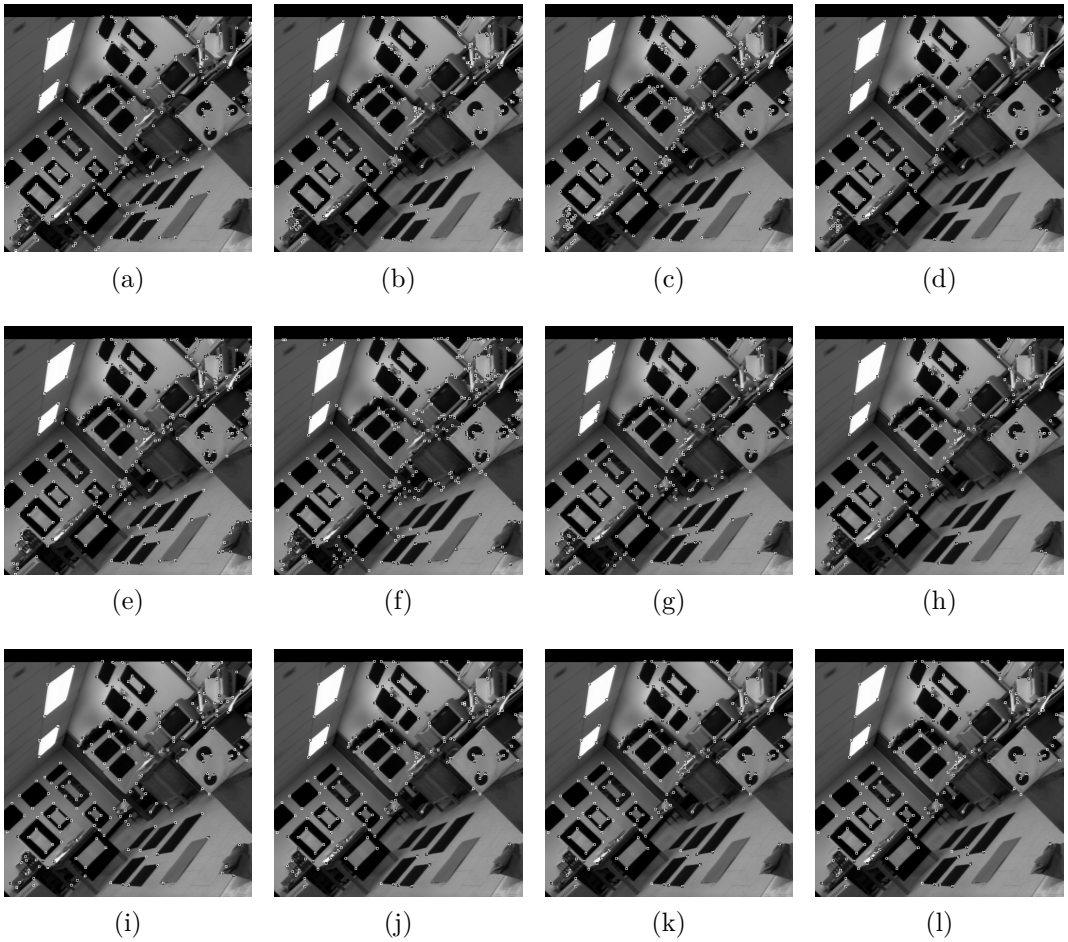


Figure 3.14: Detection results of the proposed method and the compared corner detectors on the image 'Lab'. (a) Ground truth, (b) Harris, (c) FAST, (d) CF, (e) ANDD, (f) ACJ, (g) LGWTSMM, (h) SMCD, (i) Zhang and Sun's method, (j) proposed method (K=4), (k) proposed method (K=8), (l) proposed method (K=16).

SMCD, all of the listed indicators of the proposed method are remarkably higher. This demonstrates that the proposed method outperforms current multi-scale analysis based corner detectors.

4. From the results of the proposed detector with different K, it can be concluded that using structural information from more directions is useful to improve the detection accuracy. The slight difference in the localisation errors of the

proposed detector for different K may be the result of the change in the number of detected corners.

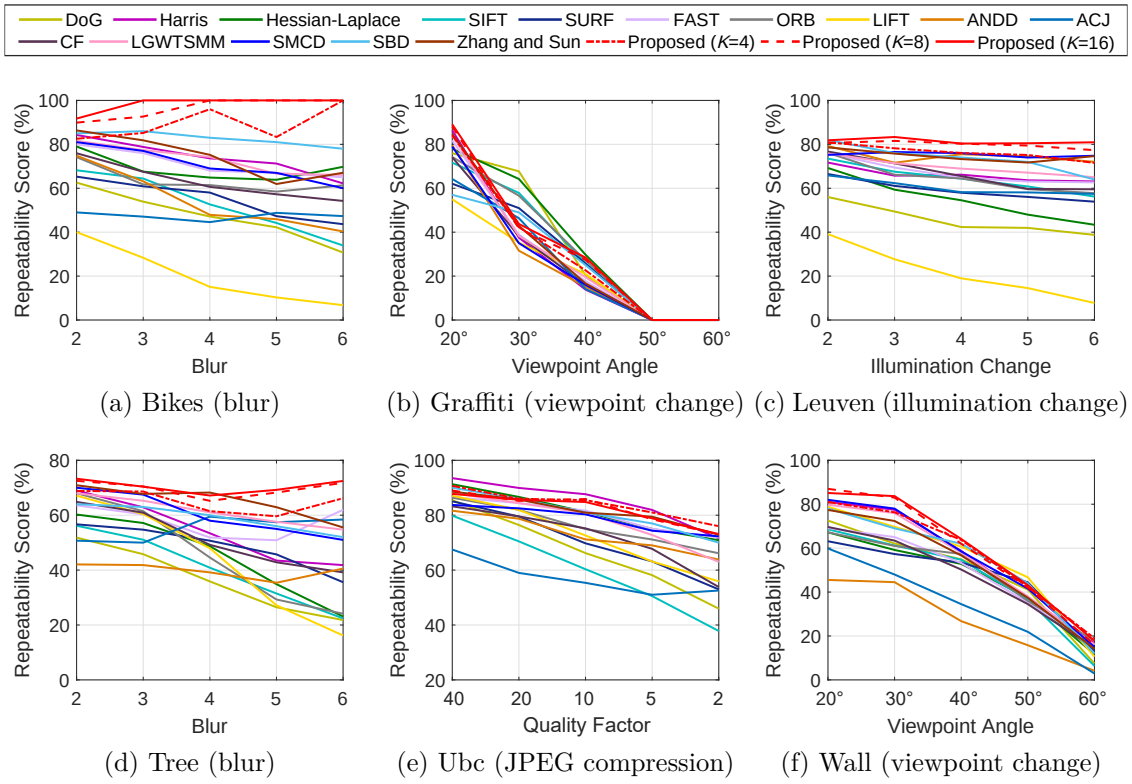
In order to verify the performance of the proposed corner detector further, a cross-validation procedure was also performed. The 20 images with ground truth were randomly divided into three folds. Folds 1 and 2 consist of two real images and four synthetic images, and fold 3 comprises three real images and five synthetic images. One fold was used for setting parameters, and the remaining folds were utilised for corner detection. The average *F-score* on the test images in each fold and the average value of the results from three folds are summarised in Table 3.2. It can be observed from these results that the proposed corner detector achieves a high detection accuracy close to that of the above experiment on all folds.

Table 3.2: Detection performance of the proposed method with cross-validation.

	Fold 1	Fold 2	Fold 3	Average result
Avg. *F-score*	78.15%	77.91%	78.46%	78.17%

3.4.4 Evaluation of Repeatability

The robustness of the proposed method to viewpoint changes, affine transformations, illumination changes and JPEG compression is evaluated in this subsection. The two datasets shown in Fig. 3.10 and Fig. 3.11 were considered and the repeatability score RS defined in Eq. (3.32) was utilised as the metric. A comparison with representative corner detectors including Harris [24], Hessian-Laplace [85], FAST [53], ORB [131], CF [81], ANDD [80], ACJ [79], LGWTSMM [93], SMCD [108] and Zhang and Sun's method [94] was performed. Considering that repeatability can reflect the potential for all interest point detectors to be extended into local feature descrip-



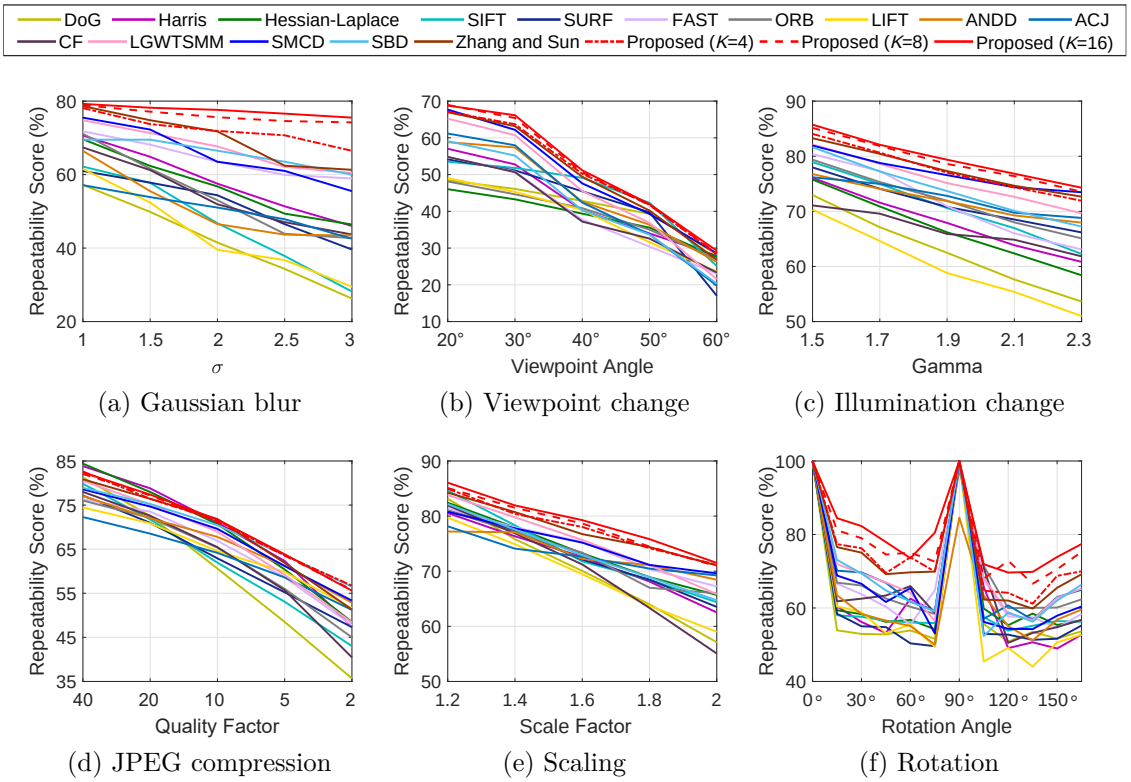
(a) Bikes (blur) (b) Graffiti (viewpoint change) (c) Leuven (illumination change)

(d) Tree (blur) (e) Ubc (JPEG compression) (f) Wall (viewpoint change)

Figure 3.15: Repeatability scores of the proposed method and the compared detectors on Dataset I. (a) 'Bikes' (blur), (b) 'Graffiti' (viewpoint change), (c) 'Leuven' (illumination change), (d) 'Tree' (blur), (e) 'Ubc' (JPEG compression), (f) 'Wall' (viewpoint change).

tors, in order to comprehensively evaluate the potential of the proposed detector, the representative interest point detectors such as DoG [136], SIFT [137], SURF [147], deep learning based method LIFT [110] and shearlet blob detector (SBD) [135] were compared as well.

The experimental results on the six sequences in Dataset I are summarised in Fig. 3.15. For Dataset II, it should be noted that each type of image sequence was divided into 50 subsequences according to the source images. For each type of image sequence, the average repeatability score on 50 subsequences was determined and considered as the result. The corresponding results are summarised in Fig. 3.16. As can be seen from Fig. 3.15(a),(c),(d) and Fig. 3.16(a),(c),(e),(f), in terms of ro-



(a) Gaussian blur (b) Viewpoint change (c) Illumination change

(d) JPEG compression (e) Scaling (f) Rotation

Figure 3.16: Repeatability scores of the proposed method and the compared detectors on Dataset II. (a) Blur, (b) viewpoint change, (c) illumination change, (d) JPEG compression, (e) scaling, (f) rotation.

bustness to image blurring, scaling, rotation and illumination change, the proposed method with $K = 8$ and 16 outperforms all compared methods, and its performance with $K = 4$ is similar to that of Zhang and Sun's method [94] but higher than that of the other compared methods.

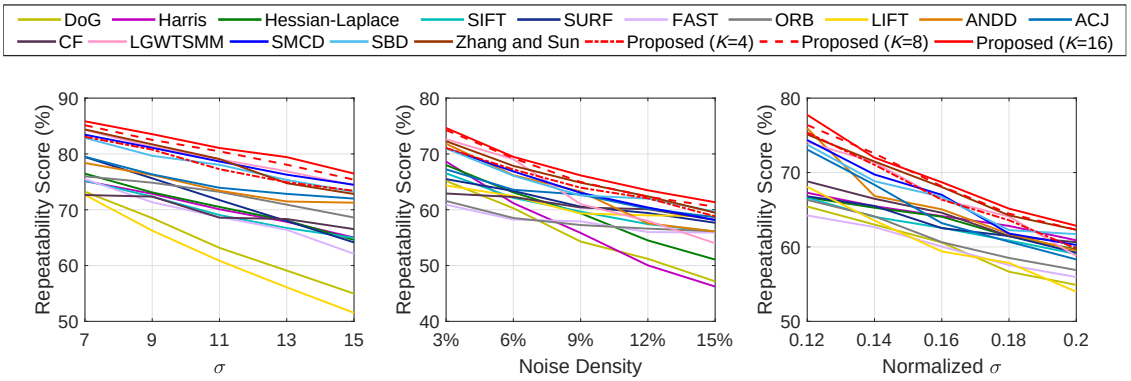
For JPEG compression, as can be observed from Fig. 3.15(e) and Fig. 3.16(d), the repeatability scores of all methods tend to drop as the quality factor decreases, among them the downward trend of the proposed method is slower than that of the other methods. It is also clear that for JPEG compression with larger quality factors, the decreases in the repeatability score of the derivative-based methods such as Harris [24] and Hessian-Laplace [85] are lower than those of the other methods,

which may be due to the fact that the two-directional structure tensor constructed from derivatives is less sensitive to the loss of very high frequency information in images. These observations demonstrate that the robustness of the proposed method to JPEG compression is generally higher than that of the other methods, and only for JPEG compression with larger quality factors, the derivative-based methods outperform the proposed method. As for viewpoint change, it can be seen from Fig. 3.16(b) that the average repeatability score of the proposed method is higher than that of the compared methods. This indicates that for most images, the proposed method yields a higher robustness to viewpoint change in comparison with the other methods. It should be noted from Fig. 3.15(b),(f) that this type of robustness on a specific image is not only associated with the detectors, but also strongly dependent on the characteristics of the image.

In summary, the proposed method outperforms current corner detectors and representative interest point detectors in terms of robustness to image blurring, viewpoint change, affine transformations, illumination changes and parts of JPEG compression. Only for JPEG compression with larger quality factors, the robustness of the proposed method is slightly lower than that of the derivative-based corner detectors but higher than that of the other methods.

3.4.5 Evaluation of Noise Robustness

In this subsection, experiments were conducted to verify the robustness of the proposed method in the presence of image noise. The three typical kinds of image noise, namely Gaussian noise, salt and pepper noise and multiplicative speckle noise, contained in Dataset III were considered in this experiment. Each kind of noisy sequence consists of 56 subsequences, and the average repeatability score was obtained and is



(a) Gaussian noise (b) Salt and pepper noise (c) Multiplicative speckle noise

Figure 3.17: Repeatability scores of the proposed method and the compared detectors on image sequences corrupted by noise. (a) Gaussian noise, (b) salt and pepper noise, (c) multiplicative speckle noise.

summarised in Fig. 3.17.

From the results, several observations may be made:

1. Multi-scale analysis based corner detectors such as SMCD [108], LGWTSMM [93], SBD [135] and the proposed method can provide higher tolerances to the three typical kinds of image noise in comparison with the other detectors.

2. The proposed detector with $K = 8$ and 16 outperforms the other methods.

3. From the results of the proposed method with different K, it can be concluded that using structural information from more directions is helpful in reducing the effects of noise.

3.4.6 Evaluation of Computational Cost

The computational cost of the proposed corner detector was evaluated and compared with the aforementioned corner detectors. The experiments were run on a personal

computer with Intel Core i5-6500 CPU and 4 GB memory, using MATLAB 2018a on Microsoft Windows 10 Enterprise, and the 20 images in Fig. 3.11 were considered.

Table 3.3: Computational cost of the eleven corner detectors.

Detectors	Computation time (s)
Harris [24]	0.045
Hessian-Laplace [85]	0.263
FAST [53]	0.019
ORB [131]	0.144
CF [81]	0.278
ANDD [80]	6.515
ACJ [79]	3.617
LGWTSMM [93]	0.622
SMCD [108]	0.591
Zhang and Sun [94]	0.942
Proposed ($K = 16$)	1.891
Proposed ($K = 8$)	0.997
Proposed ($K = 4$)	0.785

Each detector performed corner detection on each image ten times. The average computation time on the test images was obtained and listed in Table 3.3. It can be seen from Table 3.3 that among all the detectors, the computational efficiency of the template-based method FAST [53] is highest and the contour-based detectors such as ANDD [80] and ACJ [79] have the highest computational cost. As for the derivative-based detectors, the Harris [24] and Hessian-Laplace [85] detectors have higher computational efficiencies. Since more computation time is required to deal with the multi-directional structure tensor, the proposed method and Zhang and Sun's method [94] have relatively higher computational costs than LGWTSMM [93] and SMCD [108]. In Zhang and Sun's method [94], the number of directions was fixed to 8, therefore its computational cost is slightly lower than that of the proposed

detector with $K = 8$, and higher than that of the proposed detector with $K = 4$.

For the proposed method, considering the experimental results in the Sections 3.4.3, 3.4.4 and 3.4.5, it can be concluded that using structural information from more directions is helpful in improving the detection accuracy and robustness, but the computational cost has increased as well. An appropriate detector could be selected from the proposed detector with $K = 4,\ 8$, or 16 depending on the requirements of detection performance and computational efficiency in practical applications. It also can be observed that from $K = 4$ to $K = 8$, the performance improvement is significant, and the average computational time increases by about 20%. However, compared with $K = 8$, the improvement on performance of $K = 16$ is not significant but the average computational time increases by about 100%. Therefore, it can be conducted that $K = 8$ can provide relatively balanced performance and computational complexity when using multi-directional information for corner detectors.

3.5 Remarks

In this chapter an improved shearlet transform was proposed, and a novel corner detection method that can take advantage of multi-scale and multi-directional structural information was developed based on the improved shearlet transform, multi-directional structure tensor and multi-scale corner measurement function. The improved shearlet transform can provide an accurate and highly robust multi-directional and multi-scale decomposition of images for corner detection, and effectively overcome the inadequacies of the traditional shearlet transform such as edge extensions and multiple peak responses. Compared with the traditional shearlet

transform, the improved shearlet transform can extract clearer and more accurate local structural information with a flexible number of directions from images. Thus it is potentially useful as a multi-scale decomposition tool in image analysis.

Based on multi-directional and multi-scale information obtained by the improved shearlet transform from images, a novel multi-directional structure tensor was proposed to effectively improve detection accuracy and robustness. In addition, in order to further reduce the number of false detections and improve detection performance, a multi-scale measurement function was proposed by using multi-scale structural information to further refine the corners. Experimental results demonstrate that the proposed corner detection method outperforms current corner detectors in corner detection and localisation accuracies. It was also shown that its robustness to image blurring, viewpoint changes, affine transformations, illumination changes and parts of JPEG compression is generally superior to not only current corner detectors but also representative interest point detectors. Thus, it has great potential for extension as a local feature descriptor and applications in object tracking, image matching and many other tasks. However, the computational cost of the proposed method is higher than that of derivative-based methods and most multi-scale analysis based methods. Addressing this problem can greatly increase practicability of the potential applications. In Chapter 4, a corner detector with a more efficient way to use multi-scale and multi-directional structural information is proposed. This detector attempts to improve the detection accuracy, repeatability and localisation ability further while reducing computational complexity as much as possible.

Chapter 4

Corner Detection Based on Complex Shearlet Transform and Rotary Phase Congruence Structure Tensor*

Current multi-scale analysis based corner detectors do not make full use of multi-scale and multi-directional structural information from images. They do not provide sufficient structural information for detection and discard phase information, only using amplitude information for detection. Their computational complexity is also high. In order to address these shortcomings, a corner detection method was proposed in Chapter 3 that was based on a multi-directional structure tensor and an

* This chapter has been submitted and is under review: M. Wang, C. Sun, and A. Sowmya, "Complex Shearlets and Rotary Phase Congruence Tensor for Corner Detection," *Pattern Recognition*, 2021.

improved shearlet transform. Experimental results showed that the proposed detector can provide better detector performance, localisation accuracy and repeatability. However, only the amplitude coefficients at the finest scale were used for detecting candidate corners, making the results of structure tensor sensitive to image affine transformations and local variations, while the amplitude coefficients of other scales were used to eliminate false candidate detections. This means that information from these scales only helps reduce the false detection rate rather than improve the overall detection accuracy. Also, the improved shearlet cannot obtain phase information from images, therefore this information cannot be utilised by the detector. Moreover, the processing of the multi-directional structure tensor and removal of false candidate corners are computationally expensive, therefore the proposed detector has the highest computational complexity among all multi-scale analysis based methods.

In this chapter, these weaknesses of current multi-scale analysis based methods are addressed, and a novel type of shearlet transform, namely complex shearlet transform, is proposed by redefining the generation functions and the discretisation process for shearlets. Contrary to traditional shearlets, complex shearlets have a greater ability to localise distributed discontinuities, in particular the ability to capture the phase information of geometrical features. The weaknesses of traditional shearlets such as the problem of edge extension and bilateral margin responses for prominent edges are effectively suppressed. Moreover, in light of the phase congruence model [140, 141] and the idea of image denoising based on statistical modelling of wavelet coefficients [150], a new type of phase congruence function is introduced, and its tolerances to noise and ability of corner localisation have been improved by screening and normalising the amplitude information. Based on the phase congruence function, a novel rotary phase congruence structure tensor is proposed to properly merge all multi-scale and multi-directional information into a two-directional

structure tensor for corner detection. It is well known that phase congruence is less sensitive to scaling and blurring than amplitude derivatives [140], and the rotary phase congruence structure tensor is invariant to rotation in theory. Therefore, the rotary phase congruence structure tensor, with less sensitivity to image scaling, rotation, blurring and noise, can make full use of all the multi-scale and multi-directional information for corner detection.

The main contributions of this work can be summarised as follows:

1. A novel multi-scale and multi-directional analysis approach, namely the complex shearlet transform, is proposed for image processing and analysis, that can capture both amplitude and phase information, and effectually suppress bilateral margin responses and edge extensions of traditional shearlets. Moreover, the complex shearlet decomposition at all scales has the same number of directions, and each direction of any scale is consistent with one direction of the other scales, which provides benefits for subsequent image analysis by the combination of multi-scale directional information.

2. A phase congruence function is introduced to combine phase and amplitude information of geometric features for corner detection, which can provide benefits for detection performance and robustness, and its tolerances to noise and ability of corner localisation have been improved by screening and normalising the amplitude information.

3. A novel rotary phase congruence structure tensor is proposed that can properly merge all multi-scale and multi-directional information into a two-directional structure tensor. It is less sensitive to image scaling, rotation, blurring and noise.

4. Benefitting from the complex shearlet transform and rotary phase congruence structure tensor, the proposed corner detection method yields improvements on the current state-of-the-art in corner detection in terms of detection accuracy, corner localisation accuracy and robustness to affine transformations, illumination changes, noise and JPEG compression.

This chapter is organised as follows: complex shearlets are proposed and analysed in Section 4.1, followed by the phase congruence function in Section 4.2. In Section 4.3, the rotary phase congruence structure tensor and corner detector are proposed. Experimental results and discussions are in Section 4.4, and Section 4.5 contains overall remarks.

4.1 Complex Shearlet Transform

In Section 3.1, the traditional shearlet transform was briefly introduced, and the inadequacies of traditional shearlets for corner detection were discussed. In this section, a novel complex shearlet transform is proposed, which can better localise distributed discontinuities and especially with the ability to extract phase information from geometrical features, while also overcoming the shortcomings of the traditional shearlet transform.

According to Eq. (3.3), the mother shearlet $\psi \in L^2\mathbb{R}^2$ is constructed by using $\widehat{\psi}_1$ and $\widehat{\psi}_2$. In the proposed complex shearlet transform, the two generation functions $\widehat{\psi}_1^c$ and $\widehat{\psi}_2^c$ are defined as follows:

Definition 4.1 *The generation function $\widehat{\psi}_1^c(\omega)$ of the proposed complex shearlet*

transform is defined as

$$\widehat{\psi}_1^c(\omega) = \begin{cases} i\omega \left(\frac{\sin(\frac{\omega}{4})}{\frac{\omega}{4}} \right)^{2l+2}, & \omega > 0 \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

where l is a constant.

Definition 4.2 *The generation function $\widehat{\psi}_2^c(\omega)$ is defined as*

$$\widehat{\psi}_2^c(\omega) = \begin{cases} \frac{1}{B}\sqrt{D(1-\omega/B)}, & \omega > 0 \\ \frac{1}{B}\sqrt{D(1+\omega/B)}, & \text{otherwise} \end{cases} \tag{4.2}$$

where $B \in \mathbb{R}^+$ and $D(x)$ is an auxiliary function which is expressed as Eq. (3.5).

Definition 4.3 *In the proposed complex shearlet transform, the scaling factor a and shearing factor s are discretised as*

$$\begin{aligned} a_j &= 2^{-2j}, && j = 0, \ldots, j_0 - 1 \\ s_{j,k}^h &= k \times 2^{-j} \times \tfrac{2^{j+2}}{K} = \tfrac{4k}{K}, && -2^p \le k \le 2^p - 1 \\ s_{j,k}^v &= k \times 2^{-j} \times \tfrac{2^{j+2}}{K} = \tfrac{4k}{K}, && -2^p + 1 \le k \le 2^p \end{aligned} \tag{4.3}$$

where j_0 denotes the number of considered scales, p is a nonnegative integer and K is $2^{(p+2)}$.

Definition 4.4 *The parameter B in Eq. (4.2) is defined as*

$$B = \frac{2^{(j+2)}b}{K}, \qquad j = 0, \ldots, j_0 - 1 \tag{4.4}$$

with $b \in \mathbb{R}^+$, so Eq. (4.2) can be rewritten as

$$\widehat{\psi}_2^c(\omega) = \begin{cases} \frac{K}{2^{j+2}b}\sqrt{D\left(1 - \frac{K\omega}{2^{j+2}b}\right)}, & \omega > 0 \\ \frac{K}{2^{j+2}b}\sqrt{D\left(1 + \frac{K\omega}{2^{j+2}b}\right)}, & \text{otherwise} \end{cases} \tag{4.5}$$

Remark 4.1 *The shearlets constructed by $\widehat{\psi}_1^c$ and $\widehat{\psi}_2^c$ are a set of complex filters, while traditional shearlets are a set of real filters in 2D space.*

According to Eq. (3.13), in the spatial domain, the shearlet transform can be visualised as a given image I being convolved with a set of filters. Without loss of generality, in the subsequent illustrations the listed remarks are demonstrated in the case of the h cone which is defined in Eq. (3.7), and the filters $\psi_{j,k}(x,y)$ are expressed as

$$\begin{aligned}\psi_{j,k}(x,y) &= \text{ifft2}\left(\widehat{\psi}\left(4^{-j}\omega_1, 4^{-j}k\omega_1 + 2^{-j}\omega_2\right)\right)\\ &= \text{ifft2}\left(\widehat{\psi}_1\left(4^{-j}\omega_1\right)\widehat{\psi}_2\left(2^j\frac{\omega_2}{\omega_1}+k\right)\right)\\ &= \text{ifft2}\left(\widehat{\psi}_1\left(4^{-j}\omega_1\right)\widehat{\psi}_2\left(\frac{2^j}{\omega_1}\left(\omega_2+\frac{\omega_1}{2^j}k\right)\right)\right)\end{aligned} \tag{4.6}$$

According to the scaling and translation properties of Fourier transform,

$$\begin{aligned}&\text{ifft}\left(\widehat{\psi}_2\left(\frac{2^j}{\omega_1}\left(\omega_2+\frac{\omega_1}{2^j}k\right)\right)\right)\\ =\ &\text{ifft}\left(\widehat{\psi}_2\left(\frac{2^j}{\omega_1}\omega_2\right)\right)e^{-i\frac{k\omega_1}{2^j}2\pi y}\\ =\ &\left|\frac{\omega_1}{2^j}\right|\psi_2\left(\frac{\omega_1 y}{2^j}\right)e^{-i\frac{k\omega_1}{2^j}2\pi y}\end{aligned} \tag{4.7}$$

Thus, Eq. (4.6) can be rewritten as

$$\psi_{j,k}(x,y) = \text{ifft}\left(\widehat{\psi}_1\left(4^{-j}\omega_1\right)\left|\frac{\omega_1}{2^j}\right|\psi_2\left(\frac{\omega_1 y}{2^j}\right)e^{-i\frac{k\omega_1}{2^j}2\pi y}\right) \tag{4.8}$$

where $\psi_2(y)$ is $\text{ifft}(\widehat{\psi}_2(\omega))$. Let $\widehat{R}(\omega_1, y)$ be $\widehat{\psi}_1(4^{-j}\omega_1)\left|\frac{\omega_1}{2^j}\right|\psi_2(\frac{\omega_1 y}{2^j})$, then

$$\psi_{j,k}(x,y) = \text{ifft}\left(\widehat{R}\left(\omega_1, y\right)e^{-i\frac{k\omega_1}{2^j}2\pi y}\right) \tag{4.9}$$

For the traditional shearlets, $\widehat{\psi}_1$ and $\widehat{\psi}_2$ are even symmetric real functions, therefore, $\psi_2(y)$ is also an even symmetric real function. Obviously, $\widehat{R}(\omega)$ is an even symmetric

real function as well. Hence, $\widehat{R}(\omega_1, y)e^{-i\frac{k\omega_1}{2^j}2\pi y}$ is conjugate symmetric, so $\psi_{j,k}(x, y)$ is a set of real functions. In the proposed shearlet transform, $\widehat{\psi}_2^c(\omega)$ is an even symmetric real function, and $\widehat{\psi}_1^c(\omega)$ is an asymmetric imaginary function. Therefore, the proposed shearlets $\psi_{j,k}^c(x, y)$ can be deduced as

$$\begin{aligned}\psi_{j,k}^c(x,y) &= \text{ifft2}\left(\widehat{\psi}_1^c\left(4^{-j}\omega_1\right)\widehat{\psi}_2^c\left(2^j\frac{\omega_2}{\omega_1}+k\right)\right)\\ &= \text{ifft2}\left(\widehat{\psi}_1^c\left(4^{-j}\omega_1\right)\frac{K}{2^{j+2}b}\widehat{\psi}_2\left(\frac{K}{4b\omega_1}\left(\omega_2+\frac{\omega_1}{2^j}k\right)\right)\right)\end{aligned} \tag{4.10}$$

Similarly,

$$\psi_{j,k}^c(x,y) = \text{ifft}\left(\widehat{\psi}_1^c\left(4^{-j}\omega_1\right)\frac{|\omega_1|}{2^j}\psi_2\left(\frac{4b\omega_1}{K}y\right)e^{-i\frac{k\omega_1}{2^j}2\pi y}\right) \tag{4.11}$$

Let $\widehat{R}^c(\omega_1, y)$ be $\widehat{\psi}_1^c\left(4^{-j}\omega_1\right)\frac{|\omega_1|}{2^j}\psi_2\left(\frac{4b\omega_1}{K}y\right)$, then

$$\psi_{j,k}^c(x,y) = \text{ifft}\left(\widehat{R}^c\left(\omega_1, y\right)e^{-i\frac{ky}{2^j}2\pi\omega_1}\right) \tag{4.12}$$

Since $\psi_2(y)$ is an even symmetric real function and $\widehat{\psi}_1^c(\omega)$ is an asymmetric imaginary function, $\widehat{R}^c(\omega)$ is an asymmetric purely imaginary function. Therefore, $\psi_{j,k}^c(x, y)$ is a set of complex functions. The generation function $\widehat{\psi}_1$, frequency tiling for 3 scales and shearlet in the Fourier and spatial domains for $j = 1$ and $k = 1$ of the traditional shearlets are shown in Fig. 4.1, and the corresponding terms for the complex shearlets are shown in Fig. 4.2. The transform results of the traditional and proposed complex shearlets, illustrated with the sum of shearlet coefficients over all directions at scale j, are shown in Fig. 4.3 and Fig. 4.4 respectively. As can be seen from Fig. 4.1 and Fig. 4.2, the proposed shearlets are a set of complex filters, while the traditional shearlets are a set of real filters. Moreover, it can be seen that in Fig. 4.3(b) and (c), for each edge especially a prominent edge, bilateral margin responses and edge extensions to their both ends can be observed, whereas the

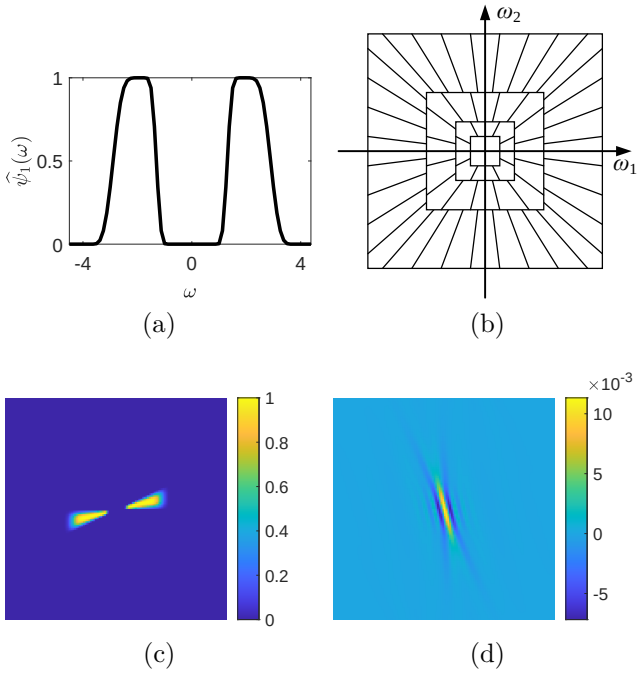


Figure 4.1: Traditional shearlets. (a) Meyer wavelet mother function, i.e., $\widehat{\psi}_1$, (b) frequency tiling for 3 scales, (c) shearlet for $j = 1$ and $k = 1$ in the Fourier domain, (d) the corresponding version in the spatial domain.

bilateral margin responses and edge extensions are greatly depressed in Fig. 4.4(b) and (c). This results in that Fig. 4.4(b) and (c) are clearer than Fig. 4.3(b) and (c). Furthermore, compared Fig. 4.3(a) and (c) with Fig. 4.4(a) and (c), more details of images can be observed. For example, the details in the girl's face can be hardly seen in Fig. 4.3(a) but can be clearly observed in Fig. 4.4(a). This indicates that either the real part or the imaginary part of the proposed shearlets can extract clearer details from images than the traditional shearlet transform.

Remark 4.2 *Any filter $\psi^c_{j,k}(x, y)$ of the complex shearlets can be interpreted as the sum of a pair of complex functions, where one is conjugate symmetric and the other is conjugate antisymmetric. Their phases in the Fourier domain differ by $\frac{\pi}{2}$, and their amplitudes in the Fourier domain are equal.*

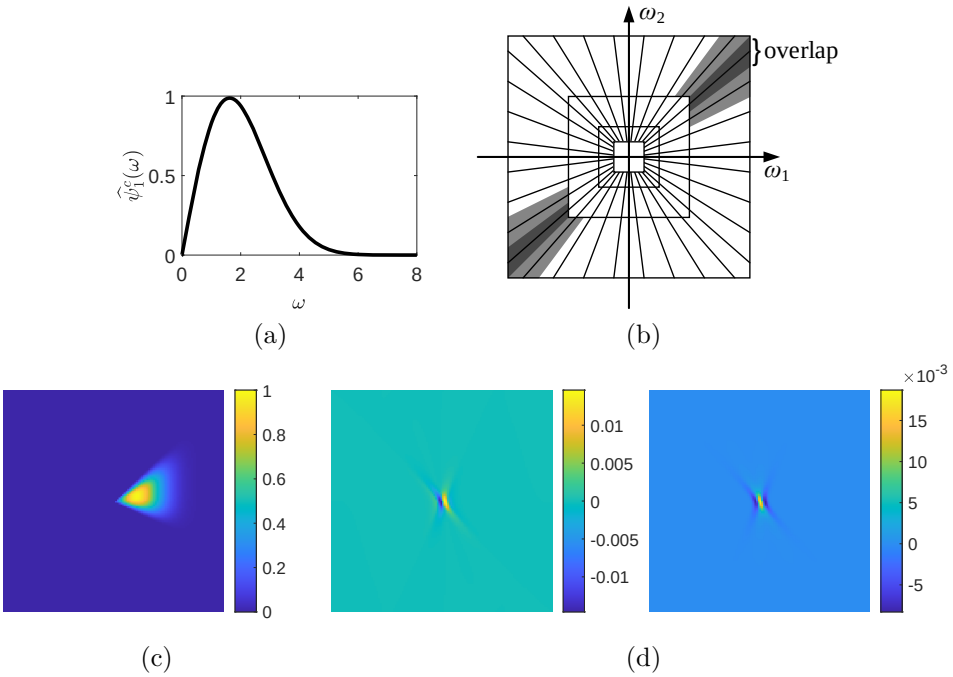
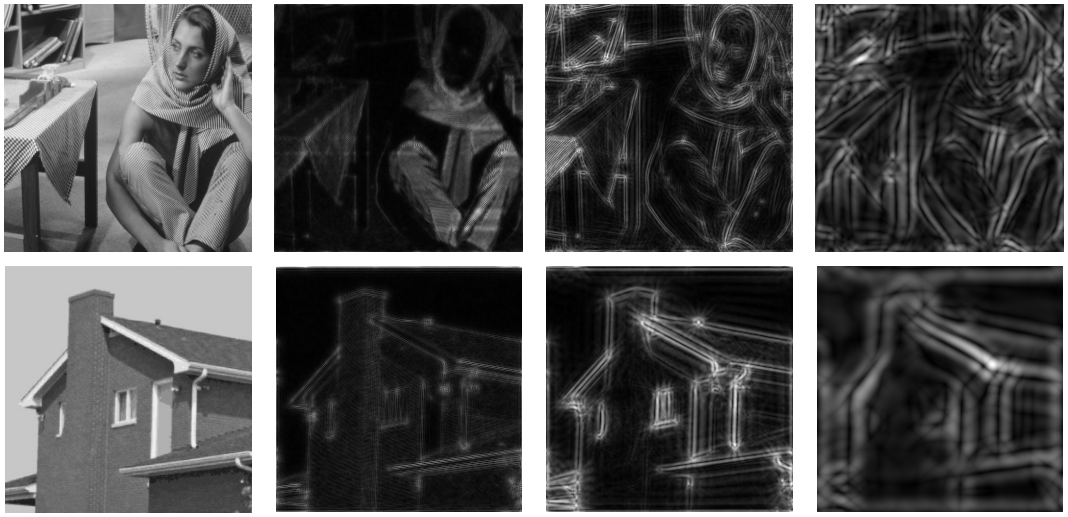


Figure 4.2: Complex shearlets. (a) Generation function $\widehat{\psi}_1^c$, (b) frequency tiling for 3 scales, (c) shearlet for $j = 1$ and $k = 1$ in the Fourier domain, (d) the corresponding version in the spatial domain, left: real part, right: imaginary part.

According to Eq. (4.12), $\psi_{j,k}^c(x, y)$ can be written as

$$\begin{aligned}
\psi_{j,k}^c(x,y) &= \text{ifft}\left(\widehat{R}^c(\omega_1, y)e^{-i\frac{ky}{2^j}2\pi\omega_1}\right) \\
&= \text{ifft}\left(\widehat{R}^c\left(\omega_1, y\right)\cos\left(\frac{ky}{2^j}2\pi\omega_1\right) - i\widehat{R}^c\left(\omega_1, y\right)\sin\left(\frac{ky}{2^j}2\pi\omega_1\right)\right) \\
&= \text{ifft}\left(\widehat{R}^c\left(\omega_1, y\right)\cos\left(\frac{ky}{2^j}2\pi\omega_1\right) - i\widehat{R}^c\left(\omega_1, y\right)\cos\left(\frac{ky}{2^j}2\pi\omega_1 + \frac{\pi}{2}\right)\right) \\
&= \psi_{j,k}^{co}(x,y) + \psi_{j,k}^{ce}(x,y)
\end{aligned} \tag{4.13}$$

with $\psi_{j,k}^{co}(x, y) = \text{ifft}\left(\widehat{R}^c\left(\omega_1, y\right)\cos\left(\frac{ky}{2^j}2\pi\omega_1\right)\right)$ and $\psi_{j,k}^{ce}(x, y) = \text{ifft}\left(-i\widehat{R}^c\left(\omega_1, y\right) \times \sin\left(\frac{ky}{2^j}2\pi\omega_1\right)\right)$. Obviously, the phase difference between $\psi_{j,k}^{co}(x, y)$ and $\psi_{j,k}^{ce}(x, y)$ is $\frac{\pi}{2}$ and their amplitudes are equal. Note that $\widehat{R}^c(\omega)$ is a purely imaginary function, $\widehat{R}^c\left(\omega_1, y\right)\cos\left(\frac{ky}{2^j}2\pi\omega_1\right)$ is a purely imaginary function, and $-i\widehat{R}^c\left(\omega_1, y\right)\sin\left(\frac{ky}{2^j}2\pi\omega_1\right)$ is a real function. Therefore, $\psi_{j,k}^{co}(x, y)$ is a conjugate antisymmetric function and



(a) (b) (c) (d)

Figure 4.3: The transform results of the traditional shearlets. (a) Original images; the sum of shearlet coefficients over all directions at scale (b) $j = 2$, (c) $j = 1$ and (d) $j = 0$.

$\psi^{ce}_{j,k}(x, y)$ is a conjugate symmetric function.

Remark 4.3 *In the traditional shearlet transform, shearlets in the Fourier domain (Fig. 4.1(c)) have a narrow bandwidth and their spatial versions (Fig. 4.1(d)) have a strong side-lobe effect, while the complex shearlets (Fig. 4.2(c)) have a relatively wide bandwidth in the Fourier domain, and the side-lobe effect in their spatial versions (Fig. 4.2(d)) has been significantly reduced.*

In the traditional shearlets, $\widehat{\psi}_1(\omega)$ was defined as

$$\widehat{\psi}_1(\omega) = \sqrt{g^2(2\omega) + g^2(\omega)} \tag{4.14}$$

(a) (b) (c) (d)

Figure 4.4: The transform results of the proposed complex shearlets. (a) Original images; the sum of shearlet coefficients in real part (top) and imaginary part (bottom) over all directions at scale (b) $j = 2$, (c) $j = 1$ and (d) $j = 0$.

where $g(\omega)$ was given by

$$g(\omega) = \begin{cases} \sin\left(\frac{\pi}{2} v(|\omega| - 1)\right), & 1 \leq |\omega| \leq 2 \\ \cos\left(\frac{\pi}{2} v(|\omega| - 1)\right), & 2 < |\omega| \leq 4 \\ 0, & \text{otherwise} \end{cases} \tag{4.15}$$

with

$$v(x) = \begin{cases} 0, & x < 0 \\ 2x^2, & 0 \le x \le \frac{1}{2} \\ 1 - 2(1-x)^2, & \frac{1}{2} < x \le 1 \\ 1, & x > 1 \end{cases} \tag{4.16}$$

As shown in Fig. 4.1(a), $\widehat{\psi}_1(\omega)$ is a band-limited function ($[-4, -\frac{1}{2}] \cup [\frac{1}{2}, 4]$) with a steep descending boundary. As a result, there is a strong side-lobe effect in the spatial domain (see Fig. 4.1(d)). For the complex shearlets, $\widehat{\psi}_1^c(\omega)$ defined in Eq. (4.1) has a relatively wide bandwidth, and the bandwidth of $\widehat{\psi}_2^c(\omega)$ is B times the bandwidth of $\widehat{\psi}_2(\omega)$. So the complex shearlets have a relatively wider bandwidth than that of traditional shearlets. Moreover, $\widehat{\psi}_1^c(\omega)$ declines much slower than $\widehat{\psi}_1(\omega)$ at their boundary, and the side-lobe effect can be significantly reduced.

Remark 4.4 *For the traditional shearlet transform, the total number of directions at scale j varies with j, and the direction k of scale j corresponds to the direction $2k$ plus one half of $2k-1$ and $2k+1$ directions of scale $j+1$, while complex shearlet decomposition has the same number of directions at all scales and the direction k of any scale is consistent with the direction k of other scales.*

According to Eq. (3.10), in the traditional shearlet transform, the shearing parameter s is discretised as $s_{j,k}^h = k \times 2^{-j},\ -2^j \le k \le 2^j - 1$ and $s_{j,k}^v = k \times 2^{-j},\ -2^j + 1 \le k \le 2^j$. This means that in scale j, the number of directions is 2^{j+2}, and the direction k of scale j corresponds to the direction $2k$ plus one half of $2k-1$ and $2k+1$ directions of scale $j+1$ (as shown in Fig. 4.1(b)). For the complex shearlets, according to Definition 4.3, the shearing parameter s is discretised as $s_{j,k}^h = k \times 2^{-j} \times \frac{2^{j+2}}{K} = \frac{4k}{K},\ -2^p \le k \le 2^p - 1$ and $s_{j,k}^v = k \times 2^{-j} \times \frac{2^{j+2}}{K} = \frac{4k}{K},\ -2^p + 1 \le k \le 2^p$. In other

words, the number of directions in all scales is $K = 2^{(p+2)}$ and the direction k of any scale is consistent with the direction k of other scales (as shown in Fig. 4.2(b)).

Based on these remarks, the advantages of the complex shearlets over traditional shearlets can be summarised as follows. First, in light of Remark 4.1, the complex shearlet transform can capture both amplitude and phase information, while the traditional shearlet transform only provides amplitude information. Secondly, according to the phase congruence model [140,141], the conjugate antisymmetric filter in Remark 4.2 will respond maximally to step edges, and the conjugate symmetric filter will respond maximally to lines and roof shapes. Therefore, complex shearlets will produce maximal phase congruence responses at the pixels with drastic gray-scale value changes, e.g., at corners or along edges. Thirdly, on the basis of Remark 4.3, in the traditional shearlets, the side-lobe effect will cause large side-lobes around the locations where gray-scale values change drastically. Thus, bilateral margin responses can be found for prominent edges (as shown in Fig. 4.3). In contrast, in the complex shearlets, the side-lobe effect has been significantly reduced, therefore, there is only a strong single response for each prominent edge in the results of the complex shearlet transform (as shown in Fig. 4.4). Moreover, the narrow bandwidth of the traditional shearlet enables the corresponding filter in the spatial domain to have a large envelope. This leads to the structural information being extended at the end of edges in areas around the edges, which is named as the problem of edge extension (see Fig. 4.3). Whereas the complex shearlet has a relatively wide bandwidth, and its envelope in the spatial domain is significantly reduced. Accordingly, the extensions of edges are also effectually reduced (see Fig. 4.4). In addition, complex shearlets with wide bandwidths can extract more high frequency details, making the structural information obtained at all scales much clearer than the traditional shearlet. Finally, according to Remark 4.4, complex shearlet decomposition has the

same number of directions at all scales and the direction k of any scale is consistent with the direction k of other scales, which provides benefits for subsequent image analysis via combination of the multi-scale directional information.

4.2 Phase Congruence Model

According to the phase congruence model [140,141] and Remark 4.2, complex shearlets will produce maximal phase congruence responses at the locations where gray-scale values change drastically, such as along edges or at corners. For an image $f(x, y)$, the multi-scale and multi-directional coefficients $csh_{j,k}(x, y)$ can be determined by using the complex shearlet transform. Based on the phase congruence model, the phase congruence function at location (x, y) along direction k, can be defined as

$$PC_k(x, y) = \frac{\sum_j A_{j,k}(x, y) \cos\left(\phi_{j,k}(x, y) - \bar{\phi}_k(x, y)\right)}{\sum_j A_{j,k}(x, y)} \tag{4.17}$$

where $A_{j,k}(x, y)$ and $\phi_{j,k}(x, y)$ represent the amplitude and local phase of the complex shearlet coefficients $csh_{j,k}(x, y)$ at position (x, y) respectively. $\bar{\phi}_k(x, y)$ denotes the average phase along direction k which can be obtained by

$$\bar{\phi}_k(x, y) = \frac{\sum_{j=0}^{j_0-1} \phi_{j,k}(x, y)}{j_0} \tag{4.18}$$

where j_0 denotes the number of considered scales. At position (x, y) where the intensity changes drastically, the local phases $\phi_{j,k}(x, y)$ tend to be congruent, therefore, $\phi_{j,k}(x, y)$ approximately equals $\bar{\phi}_k(x, y)$, and $PC_k(x, y)$ approaches 1. For flat regions in images, there is no coherence of phases and $PC_k(x, y)$ falls to a minimum of 0.

Remark 4.5 *The phase congruence function $PC_k(x, y)$ is invariant to linear variations in image illumination.*

For an image $I(x, y)$, its transformed image with a linear variation in illumination is given by

$$I^{(1)}(x, y) = \xi I(x, y) \tag{4.19}$$

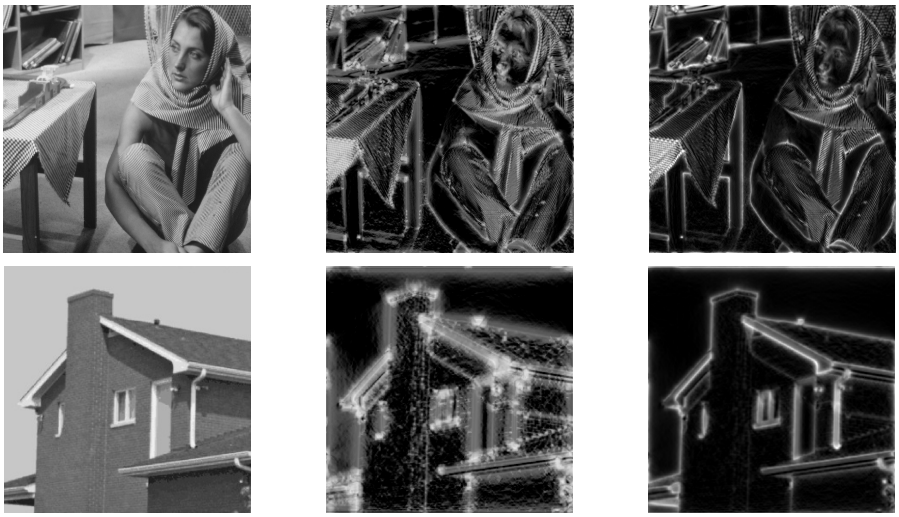
where $\xi \in \mathbb{R}^+$ reflects the linear variation in illumination. According to Remark 4.2,

$$\begin{aligned} csh^{(1)}_{j,k}(x, y) &= I^{(1)}(x, y) * \psi^c_{j,k}(x, y) \\ &= \xi I(x, y) * \psi^c_{j,k}(x, y) \\ &= \xi \times csh_{j,k}(x, y) \end{aligned} \tag{4.20}$$

where $*$ denotes the convolution operation. Substituting Eq. (4.20) into Eq. (4.17),

$$PC^{(1)}_k(x, y) = PC_k(x, y) \tag{4.21}$$

It is well known that phase congruence is less affected by the filter bandwidth, which means that its positioning accuracy is less affected by the filter envelope [140, 141]. However, for natural images, the amplitude of a frequency component generally decreases with the increase of frequency [151]. Therefore, the phase congruence function $PC_k(x, y)$ mainly depends on low-frequency components. Compared with high-frequency components, the phase change of low-frequency components with spatial position is slower, which will affect the localisation accuracy. For improvement in corner localisation accuracy and noise robustness of the phase congruence function, and according to the statistical characteristics of wavelet coefficients in



(a) Original images (b) $PC_k(x, y)$ (c) $NPC_k(x, y)$

Figure 4.5: Original images and phase congruence images obtained by complex shearlet transform with 3 scales and 8 directions. (a) Original images, (b) phase congruence images of $PC_k(x, y)$, (c) phase congruence images of $NPC_k(x, y)$.

detail components, where noise is usually scattered on all wavelet coefficients while the useful signal is concentrated on some large wavelet coefficients [141, 150], the phase congruence function is redefined as

$$NPC_k(x,y) = \begin{cases} \dfrac{\sum\limits_j A^1_{j,k}(x,y)\cos\left(\phi_{j,k}(x,y) - \bar{\phi}_k(x,y)\right)}{\sum\limits_j A^1_{j,k}(x,y)}, & \sum\limits_j A^1_{j,k}(x,y) \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{4.22}$$

and

$$A^1_{j,k}(x,y) = \begin{cases} \dfrac{A_{j,k}(x,y)}{\max(A_{j,k})}, & A_{j,k}(x,y) \geq \varepsilon_T \\ 0, & \text{otherwise} \end{cases} \tag{4.23}$$

where ε_T $(= 0.1)$ is used to reduce the effect of noise. According to Eq. (4.22) and Eq. (4.23), the contribution of amplitude to $NPC_k(x, y)$ is screened and normalised

with the maximum coefficient of the same scale, so the effect of the low-frequency components on localisation accuracy and the effect of noise have been effectively reduced. The images shown in Fig. 4.5(a) were decomposed with the complex shearlet transform into 3 scales and 8 directions. The phase congruence value for each direction was determined using Eq. (4.17) and Eq. (4.22) and added over all directions to form phase congruence images, which are shown in Fig. 4.5(b) and (c). Clearly the phase congruence images of $NPC_k(x, y)$ are much clearer than those of $PC_k(x, y)$.

4.3 Rotary Phase Congruence Structure Tensor and Proposed Corner Detector

In this section, a novel rotary phase congruence structure tensor for corner detection is proposed by using the multi-directional phase congruence information. For a pixel (m, n) in image I, a square patch of size $P \times P$ is placed over the image area centered at (m, n) and is rotated until its x axis coincides with the direction $k_1 = \underset{k}{\operatorname{argmax}}\,(NPC_k(m, n))$. After translating the rotated square patch by $(\Delta x, \Delta y)$, the corresponding intensity change $E(\Delta x, \Delta y)$ is given by

$$E(\Delta x, \Delta y) = \frac{1}{P^2} \sum_{u=-\frac{P-1}{2}}^{\frac{P-1}{2}} \sum_{v=-\frac{P-1}{2}}^{\frac{P-1}{2}} [I(m + u + \Delta x, n + v + \Delta y) - I(m + u, n + v)]^2 \tag{4.24}$$

Taylor function can be used to simplify $I(m + u + \Delta x, n + v + \Delta y)$ to

$$I(m + u + \Delta x, n + v + \Delta y) = I(m + u, n + v) + \Delta x I_x + \Delta y I_y + O\left(\Delta x^2, \Delta y^2\right) \tag{4.25}$$

where I_x is the derivative in the direction k_1 and I_y denotes the derivative perpendicular to the direction k_1. Then, Eq. (4.24) can be rewritten as

$$E(\Delta x, \Delta y) \approx \frac{1}{P^2} \sum_{u=-\frac{P-1}{2}}^{\frac{P-1}{2}} \sum_{v=-\frac{P-1}{2}}^{\frac{P-1}{2}} [\Delta x I_x + \Delta y I_y]^2 \tag{4.26}$$

According to Section 4.2, the phase congruence information defined in Eq. (4.22) can reflect image grayscale changes. Moreover, in image $I(x, y)$, the direction $k_1 = \underset{k}{\operatorname{argmax}} (NPC_k(m, n))$ coincides with the x axis. Therefore, $NPC_{k_1}(m, n)$ and $NPC_{k_1+\frac{\pi}{2}}(m, n)$ can be used instead of I_x and I_y, which are calculated by projecting the phase congruence information of all directions on the k_1 and $k_1 + \frac{\pi}{2}$ directions:

$$\begin{aligned} E(\Delta x, \Delta y) \approx \frac{1}{P^2} & \sum_{u=-\frac{P-1}{2}}^{\frac{P-1}{2}} \sum_{v=-\frac{P-1}{2}}^{\frac{P-1}{2}} \sum_{k=0}^{K-1} \\ & [\Delta x NPC_k(m, n) \cos\theta_k + \Delta y NPC_k(m, n) \sin\theta_k]^2 \quad (4.27) \\ = \frac{1}{P^2} & \sum_{u=-\frac{P-1}{2}}^{\frac{P-1}{2}} \sum_{v=-\frac{P-1}{2}}^{\frac{P-1}{2}} [\Delta x, \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \end{aligned}$$

where $\theta_k = \frac{\pi}{K}(k - k_1)$ and the rotary phase congruence structure tensor M is defined as

$$M = \begin{bmatrix} \sum_{k=0}^{K-1} NPC_k^2(m, n) \cos^2\theta_k & \frac{1}{2} \sum_{k=0}^{K-1} NPC_k^2(m, n) \sin 2\theta_k \\ \frac{1}{2} \sum_{k=0}^{K-1} NPC_k^2(m, n) \sin 2\theta_k & \sum_{k=0}^{K-1} NPC_k^2(m, n) \sin^2\theta_k \end{bmatrix} \tag{4.28}$$

Since the magnitudes of the eigenvalues of M can represent the intensity changes near a pixel, the eigenvalues λ_1 and λ_2 of M can be used to form a corner detection function to distinguish corners from other pixels. For each image pixel (m, n), based on the eigenvalues λ_1 and λ_2 of M, a corner response function R is defined by

$$R = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2 + 1} \tag{4.29}$$

For pixels in flat areas or on an edge, none or only the phase congruence of the direction perpendicular to the edge is larger, and both eigenvalues or one of them are small, causing a relatively small value for R. For corners, the phase congruences of two or more directions are larger, so both eigenvalues of M are larger, forming a relatively larger R. Thus, corners can be detected by comparing R with an appropriate threshold.

Remark 4.6 *The rotary phase congruence structure tensor defined in Eq. (4.28) is invariant to image translation and rotation.*

For pixel (m, n), any image translation has no effect on its coefficients $NPC_k(m, n)$. So, the phase congruence structure tensor M is invariant to image translation. In addition, the direction of the rotary phase congruence structure tensor is normalized by rotating the structure tensor until its x axis coincides with direction $k_1 = \underset{k}{\operatorname{argmax}}\left(NPC_k(m, n)\right)$. Thus, it is invariant to image rotation. Moreover, Eq. (4.29) can be directly calculated by

$$\begin{aligned} R &= \frac{\det(M)}{\operatorname{trace}(M) + 1} \\ &= \frac{\sum_{k=0}^{K-1} NPC_k^2(m, n)\cos^2\theta_k \sum_{k=0}^{K-1} NPC_k^2(m, n)\sin^2\theta_k - \frac{1}{4}\left(\sum_{k=0}^{K-1} NPC_k^2(m, n)\sin 2\theta_k\right)^2}{1 + \sum_{k=0}^{K-1} NPC_k^2(m, n)} \end{aligned} \tag{4.30}$$

where $\det(M)$ and $\operatorname{trace}(M)$ represent the determinant value and trace of M respectively. Compared with the multi-directional structure tensor in other methods [94, 152], in which the matrix eigenvalue decomposition process is necessary for each pixel, the proposed method (detailed in Algorithm 4.1) has high computational efficiency.

Algorithm 4.1 Proposed Corner Detection Algorithm Based on Complex Shearlet Transform and Rotary Phase Congruence Structure Tensor.

1: Decompose the input image with the complex shearlet transform into 3 scales and 8 directions.

2: Calculate the phase congruence function $NPC_k(x, y)$ with Eq. (4.22).

3: For each pixel, compute R based on Eq. (4.30).

4: Pixels with local maximum of R larger than a threshold are regarded as corners.

4.4 Performance Evaluation and Experimental Analysis

In this section, the performance of the proposed method is evaluated by experiments. It is well known that corner detection and localisation accuracies, computational efficiency and repeatability under various image transformations are the main factors that determine the application potential of corner detectors. Performance was evaluated from these four perspectives. A comparison of the proposed corner detector with representative corner detection methods including FAST [53], Harris [24], Hessian-Laplace [85], ANDD [80], ACJ [79], CF [81], ORB [131], SMCD [108], LGWTSMM [93], Zhang and Sun's method [94], the proposed method in Chapter 3, and the recent deep learning interest point detection methods D2-Net [113] and Reinforced SP [111] was performed.

4.4.1 Corner Detection and Localisation Accuracies

In this subsection, the corner detection and localisation accuracies of the proposed corner detector were verified. A total of 20 different images with ground truths

including 7 real images and 13 synthetic images introduced in Section 3.4.1.1 was adopted. To measure detection accuracy, *F-score* and *precision* were used as metrics, and the corner localisation error L was utilized to evaluate the corner localisation accuracy.

Table 4.1: Detection performance and localisation accuracies.

Detectors	*Precision* (%)		*F-score* (%)		L
	Avg.	Max.	**Avg.**	Max.	
Harris [24]	60.59	85.45	63.91	80.99	1.538
FAST [53]	53.93	86.49	61.56	75.81	1.644
CF [81]	74.55	96.77	64.20	77.23	1.526
ANDD [80]	76.37	96.88	73.48	93.33	1.446
ACJ [79]	66.87	87.10	67.86	87.72	2.173
LGWTSMM [92]	73.88	92.00	71.12	84.40	1.944
SMCD [108]	70.73	84.91	66.44	80.36	2.071
Zhang and Sun [94]	83.10	94.44	74.10	84.40	1.748
Method in Chapter 3	86.48	94.59	77.37	95.89	1.433
Proposed	**88.76**	**97.33**	**78.73**	**97.53**	**1.297**

For each method, its threshold was set to the value at which the detector achieves the maximum average *F-score* on 6 images randomly selected from the 20 images. The remaining 14 images were used for evaluation, and the average and maximum *precision*, *F-score*, and localisation error L are shown in Table 4.1. As can be seen from the results, the multi-scale analysis based methods perform better among all compared detectors. Compared with other detectors, all performance measures of the proposed method are significantly better. The reason is that the multi-scale structural information merged with the phase congruence function is fully utilised to improve the detection performance, which can avoid the detection performance degradation and the impact on localisation accuracy caused by the fixed detection scale in other algorithms. Further, for the traditional shearlets in SMCD [108], bi-

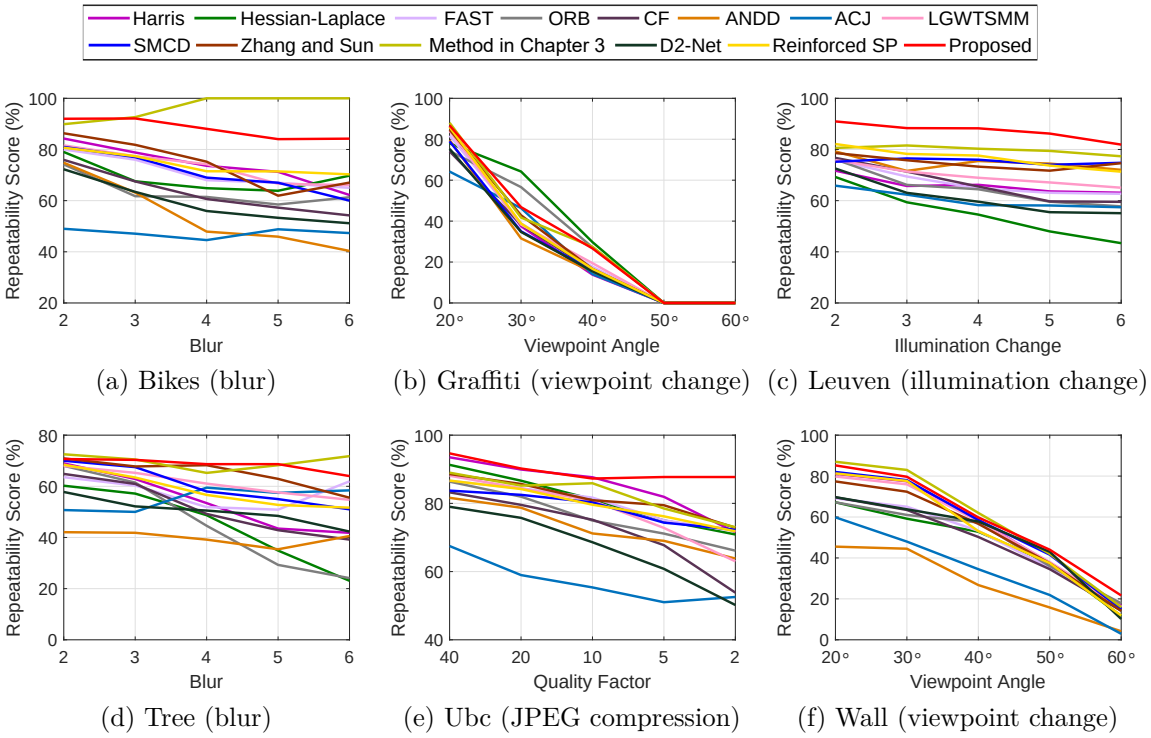
lateral margin responses and edge extensions will cause interference for positioning edges and details. In the proposed complex shearlet transform, according to Remark 4.3, the bilateral margin responses and edge extensions have been greatly reduced. As a result, the edges or details can be localized more accurately. Compared with the improved shearlets in the method in Chapter 3, according to Remark 4.2, one of the proposed complex shearlet can be regarded as the sum of a pair of complex functions, where one is conjugate symmetric and the other is conjugate antisymmetric. Considering that the conjugate antisymmetric filter will respond maximally to step edges, and the conjugate symmetric filter will respond maximally to lines and roof shapes [141], the proposed complex shearlet transform can respond to all types of discontinuities. So, the proposed complex shearlet transform can better localise distributed discontinuities, especially with the ability to extract the phase information of geometrical features for better corner detection performance. Since the phase information and the proposed phase congruence model can better locate geometric features rather than the amplitude information in images, and with higher robustness to local variations especially for illumination changes, the proposed method performs better than existing methods in terms of *precision* and *F-score*, and the localisation accuracy L is further improved by screening and normalising the amplitude information. Therefore, the proposed method can achieve the best localisation error L amongst all the compared methods.

4.4.2 Evaluation of Repeatability

The main objective of this subsection is to verify repeatability under image blurring, scaling, rotation, noise corruption, JPEG compression, viewpoint changes and illumination changes. Three datasets (partly shown in Fig. 3.10 and Fig. 3.11) in-

troduced in Section 3.4.1.1 were used for evaluation, and the repeatability score RS defined in Eq. (3.32) used as the evaluation metric for each image sequence. For the sake of fairness, it is important to ensure that all methods yield a similar number of detected corners in the first frame of each image sequence. Therefore, in this section, for each image sequence, the thresholds of all corner detectors were finetuned to achieve a similar number of detected corners in the first frame.

Experimental results on Dataset I are presented in Fig. 4.6. Each type of image sequence in Dataset II and Dataset III consists of 50 or 56 subsequences, and the average repeatability score is obtained and summarised in Fig. 4.7 and Fig. 4.8. From these results, several observations may be made. First, according to Remark 4.5, the phase congruence function is invariant to linear variations in illumination, therefore, the proposed detector has very strong repeatability under illumination changes (Figs. 4.6(c) and 4.7(c)). Compared with Fig. 4.6(c), in which illumination has been changed naturally, the downward slope in Fig. 4.7(c) may be the result of using gamma correction to simulate illumination changes, which is not exactly linear. Second, it is well known that phase congruence is less sensitive to scaling and blurring than amplitude derivatives [140, 141]. Moreover, the rotary phase congruence structure tensor is rotationally normalised, therefore, the proposed detector performs better than compared detectors in terms of repeatability under image scaling and rotation. The changes in the curves of Fig. 4.7(f) may be the result of resampling and relocation errors in rotating discrete images. In terms of repeatability under image blurring, the proposed detector, the method in Chapter 3 and Zhang and Sun's detector [94] outperform the other methods. Also, the reason that the repeatability of the method in Chapter 3 is superior to the proposed detector could be due to the usage of a multi-directional structure tensor, which can obtain more useful information from blurred images. Third, for JPEG compressed images, the



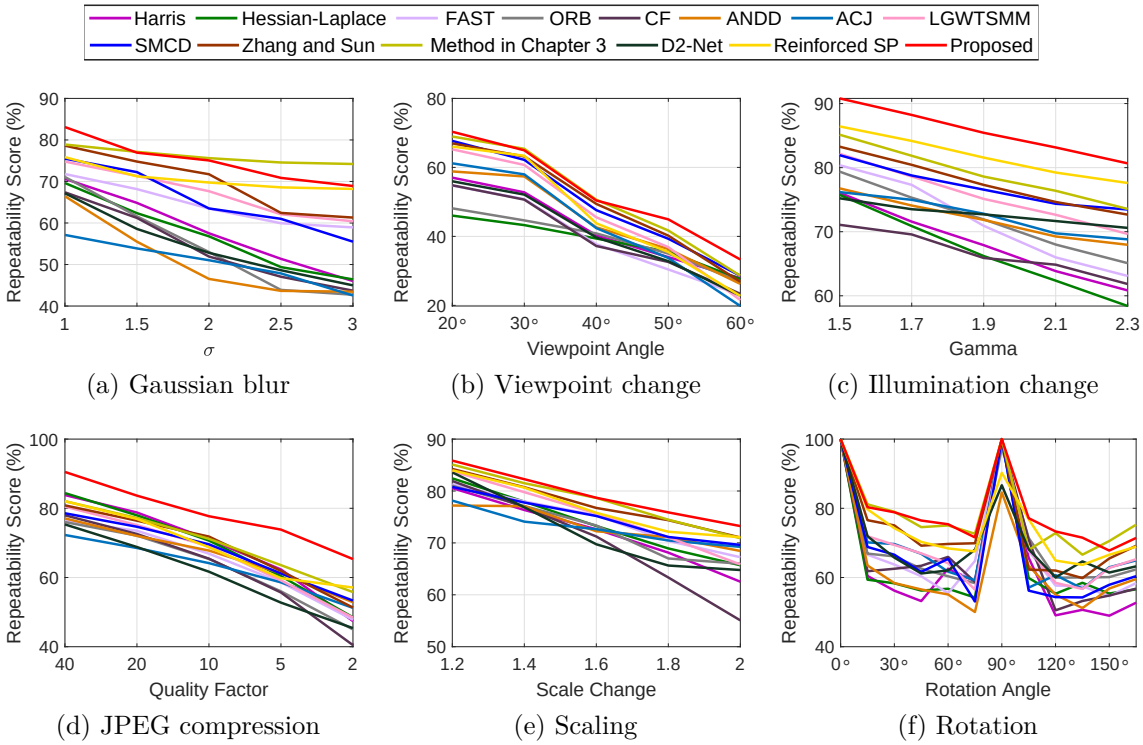
(a) Bikes (blur) (b) Graffiti (viewpoint change) (c) Leuven (illumination change)

(d) Tree (blur) (e) Ubc (JPEG compression) (f) Wall (viewpoint change)

Figure 4.6: Repeatability scores of the proposed method and the compared detectors on Dataset I. (a) 'Bikes' (blurring), (b) 'Graffiti' (viewpoint changing), (c) 'Leuven' (illumination changing), (d) 'Tree' (blurring), (e) 'Ubc' (JPEG compression), (f) 'Wall' (viewpoint changing).

proposed detector performs better than the other detectors. Fourth, for viewpoint change, in Fig. 4.7(b), the proposed detector yields higher average repeatability score especially for large viewpoint angles, which demonstrates that the proposed detector performs better than other detectors in term of repeatability under viewpoint changes for most images. It is also shown in Fig. 4.6(b) and (f) that repeatability under viewpoint changes greatly depends on the characteristics of a specific image.

It may be concluded that, compared with other methods as well as the two recent deep learning interest point detectors, the proposed detector yields the highest repeatability under image scaling, rotation, illumination changes and JPEG compression. In addition, for most images, it yields higher repeatability under image

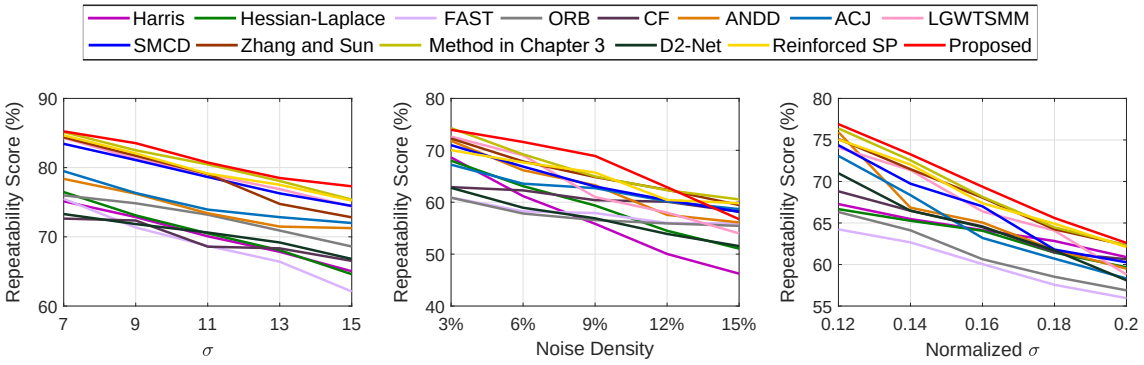


(a) Gaussian blur (b) Viewpoint change (c) Illumination change

(d) JPEG compression (e) Scaling (f) Rotation

Figure 4.7: Repeatability scores of the proposed method and the compared detectors on Dataset II. (a) Blurring, (b) viewpoint changing, (c) illumination changing, (d) JPEG compression, (e) scaling, (f) rotation.

viewpoint changes especially large viewpoint changes. Only for repeatability under image blurring, none but the method in Chapter 3 is superior to the proposed detector. From Fig. 4.8, it may also be concluded that the proposed detector achieves higher repeatabilities under Gaussian noise, salt and pepper noise with low-density and multiplicative speckle noise than current corner detectors, and only for salt and pepper noise with high-density, its repeatability is lower than that of some multi-scale analysis based corner detectors and the deep learning algorithm Reinforced SP [111].



(a) Gaussian noise (b) Salt and pepper noise (c) Multiplicative speckle noise

Figure 4.8: Repeatability scores of the proposed method and the compared detectors on Dataset III. (a) Gaussian noise, (b) salt and pepper noise, (c) multiplicative speckle noise.

4.4.3 Computational Cost

In this subsection, the computational cost of the proposed detector is tested using MATLAB 2018a on a personal computer with Intel Core i5-6500 CPU, 4 GB memory and Microsoft Windows 10 Enterprise. The images in Section 3.4.1.1 were used. Each detector performed corner detection on each image ten times. The average computation time on these images was obtained and listed in Table 4.2. From these results, it may be observed that the template-based method FAST [53] has the lowest computational cost, while the derivative-based detectors such as Harris [24] and Hessian-Laplace [85] have a relatively lower computational cost and the contour-based detectors such as ANDD [80] and ACJ [79] have the highest computational cost. Another contour-based detector CF [81] does not perform contour extraction and the tolerance to geometric transformations was not fully considered in its curvature analysis, therefore, the computational cost is relatively low. Multi-scale analysis based corner detectors such as LGWTSMM [93], SMCD [108], the method in Chapter 3, Zhang and Sun's method [94] and the proposed detector are evidently faster than the contour-based detectors in general, and slower than the template

Table 4.2: Computational cost of the evaluated corner detectors.

Detectors	Computation time (s)
Harris [24]	0.045
Hessian-Laplace [85]	0.263
FAST [53]	0.019
ORB [131]	0.144
CF [81]	0.278
ANDD [80]	6.515
ACJ [79]	3.617
LGWTSMM [93]	0.622
SMCD [108]	0.591
Zhang and Sun [94]	0.942
Method in Chapter 3	0.997
Proposed	0.515

and derivative-based detectors. Among them, the proposed detector has the lowest computational cost, while the high costs of the method in Chapter 3 and Zhang and Sun's detector [94] mainly result from the processing of their multi-directional tensors.

4.5 Remarks

In this chapter, a novel shearlet transform, namely the complex shearlet transform, was proposed, and a novel corner detection method was developed based on the complex shearlet transform and a rotary phase congruence structure tensor. The advantages of the complex shearlets over traditional shearlets include the following:

1. They can capture both amplitude and phase information from images.

2. Both their real parts and imaginary parts can extract much clearer structural information from images, and the bilateral margin responses and edge extensions in traditional shearlets are effectually suppressed.

3. The complex shearlet decomposition at all scales has the same number of directions and each direction of any scale is consistent with one direction of other scales, which provides benefits for subsequent image analysis by a combination of multi-scale directional information.

Therefore, the complex shearlet transform is potentially useful as a multi-scale decomposition approach in image processing and analysis.

Based on the proposed complex shearlet transform, a phase congruence function and rotary structure tensor for corner detection were proposed. Compared with current multi-scale analysis based corner detectors, the rotary phase congruence structure tensor can properly merge all the multi-scale and multi-directional information into a two-directional structure tensor and is less sensitive to image rotation. Moreover, the proposed phase congruence function is invariant to linear variations in image illumination and less sensitive to scaling, blurring and noise. In addition, the computation process for the complex shearlet transform using the fast Fourier transform and the introduction of the rotary phase congruence structure tensor can effectively reduce the computation cost. Experimental results demonstrate that the proposed corner detector outperforms current corner detectors in corner detection accuracy, localisation accuracy and robustness to illumination changes, JPEG compression, scaling, rotation and viewpoint changes especially large viewpoint changes for most images. Only for robustness to blurring, none but the method in Chapter 3 is superior to the proposed detector. It was also shown that the computational cost of the proposed method is lower than that of current multi-scale analysis based corner

detectors. Therefore, it has great potential as a local feature descriptor and for applications such as object tracking and image matching. However, for some real-time computer vision tasks especially real-time portable tasks, the proposed method may not meet the requirement of the low computational cost and simple architecture. In Chapter 5, a simple and computationally efficient corner detector with sufficient detection accuracy, repeatability and high adaptability to achieve real-time detection is provided, which is potentially useful for real-time computer vision applications.

Chapter 5

Efficient Corner Detection Based on Corner Enhancement Filters*

For real-time computer vision tasks such as on-line visual localisation, motion estimation and 3D reconstruction, it is crucial to quickly and accurately detect corners with high repeatability. Especially for real-time portable applications, it is important to realise corner detection in hardware, e.g., field programmable gate array (FPGA) device. The popular classical Harris corner detector [24] detects corners with a 2×2 structure tensor formed by the horizontal and vertical derivatives of images. Since the two-directional derivatives are sensitive to noise and not sufficient to distinguish between corners and edges, the detection accuracy and repeatability are unsatisfactory. Another popular classical corner detector, namely the FAST corner detector [53], performs corner detection by matching a parameterised corner

* This chapter has been accepted by a journal and is in press: M. Wang, C. Sun, and A. Sowmya, "Efficient Corner Detection Based on Corner Enhancement Filters," *Digital Signal Processing*, 2022.

template with local areas of images. Since the parameterised corner template is sensitive to local variations and cannot handle many complex scenes well, its detection accuracy is not high enough for many computer vision applications. The advantage of the two detectors is that they are computationally cost-efficient. Although their detection accuracy and repeatability are insufficient, they are still the preferred corner detectors in many real-time applications [13, 153–155] for this reason.

To improve corner detection accuracy, contour-based detectors have been proposed. Their detection accuracy is generally high in comparison with derivative-based and template-based detectors. However, their computational cost is much higher than that of the other two types of methods, which limits their applications. Recently, due to their promising detection accuracy and repeatability, multi-scale analysis based detectors have attracted more attention. In Chapter 3 and Chapter 4, two multi-scale analysis based detectors were proposed on the basis of making full use of multi-scale and multi-directional structural information. Their detection accuracy, repeatability and localisation ability are superior to many existing detectors. The method proposed in Chapter 4 has also greatly improved computational efficiency, which is the best among all multi-scale analysis based detectors. However, they are too complex to be implemented in hardware, are time-consuming and not suitable for real-time computer vision tasks.

For real-time computer vision tasks, it is important to achieve fast and accurate corner detection with high repeatability, especially for real-time portable tasks. This chapter proposes a novel efficient and accurate corner detector with a simple architecture, high parallel computing characteristics and adaptability to achieve real-time detection. Experimental results on benchmark datasets show that the proposed detector can achieve or even exceed the detection performance of the multi-scale anal-

ysis based detectors. Moreover, benefitting from the simple architecture and high parallel computing characteristic, it has a good adaptability to achieve real-time detection. Consequently, it has great advantages for application in various real-time tasks including portable real-time applications.

The main contributions of this work can be summarised as follows:

1. A novel efficient corner detector with a simple architecture and high parallel computing characteristics is presented, which provides detection accuracy and repeatability that are similar to multi-scale analysis based detectors and has a good adaptability to achieve real-time detection.

2. The idea of enhancing corners and suppressing edges at the same time by directly filtering raw images is novel and different from existing corner detectors, including derivative-based, template-based, contour-based and multi-scale analysis based detectors, offering a new lead for corner detection research.

3. A FPGA architecture design of the proposed corner detector with parallel architecture and pipeline design is presented to meet the requirements of portable real-time computer vision tasks.

This chapter is organised as follows: in Section 5.1, a corner enhancement filter is defined and its performance on corner enhancement and edge suppression is discussed. Then, the proposed corner detector is described as well. Its FPGA architecture design is presented in Section 5.2. Experimental results are presented and discussed in Section 5.3, and the chapter ends with some remarks in Section 5.4.

5.1 Efficient Corner Enhancement and Detection

Current corner detectors, including contour-based and multi-scale analysis based detectors, follow a two-phase framework. The first phase is to enhance useful information such as contours or multi-scale and multi-directional structural information, and the second phase is to detect corners by mining the resultant information of the first phase with a structure tensor or with contour analysis. Accordingly, they generally have a relatively complicated architecture, high computational cost and low parallel computing performance, especially in the second phase. Template-based and derivative-based corner detectors are simple and efficient, however, their detection accuracy and repeatability are insufficient for most computer vision tasks. Filtering is the basic means of signal and image processing, and the related convolution computation has an inherent parallel computing characteristic. In this section, a new type of filter, namely corner enhancement filter, is proposed, that can enhance corners and suppress edges as well as noise simultaneously. Its performance on corner enhancement and edge suppression is also discussed. Based on the definition and analysis results, the proposed fast corner detector is then introduced.

5.1.1 Corner Enhancement Filter

In the spatial domain, a corner enhancement filter is defined as

$$l(x,y) = \begin{cases} 0, & \text{if } \ x = x_0 \ \text{ or } \ y = y_0 \\ \text{sign}((x-x_0)(y-y_0)) \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}}, & \text{otherwise} \end{cases} \tag{5.1}$$

where (x_0, y_0) is the centre point of the filter, and $\text{sign}(\cdot)$ denotes the sign function.

For any point (x, y), its symmetry point (x', y') with respect to line $x = x_0$ can be denoted as $(x' = 2x_0 - x, y' = y)$, so

$$
\begin{aligned}
l(x', y') &= \begin{cases} 0, & \text{if } x = x_0 \text{ or } y = y_0 \\ \text{sign}((x' - x_0)(y' - y_0)) \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x'-x_0)^2+(y'-y_0)^2}{2\sigma^2}}, & \text{otherwise} \end{cases} \\
&= \begin{cases} 0, & \text{if } 2x_0 - x = x_0 \text{ or } y = y_0 \\ \text{sign}((2x_0 - x - x_0)(y - y_0)) \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(2x_0-x-x_0)^2+(y-y_0)^2}{2\sigma^2}}, & \text{otherwise} \end{cases} \\
&= \begin{cases} 0, & \text{if } x = x_0 \text{ or } y = y_0 \\ -\text{sign}((x - x_0)(y - y_0)) \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}}, & \text{otherwise} \end{cases} \\
&= -l(x, y)
\end{aligned}
\tag{5.2}
$$

Similarly for any point (x, y), its symmetry point (x'', y'') with respect to line $y = y_0$ can be denoted as $(x'' = x, y'' = 2y_0 - y)$, and it can be easily deduced that

$$
\begin{aligned}
l(x'', y'') &= \begin{cases} 0, & \text{if } x = x_0 \text{ or } y = y_0 \\ -\text{sign}((x - x_0)(y - y_0)) \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}}, & \text{otherwise} \end{cases} \\
&= -l(x, y)
\end{aligned}
\tag{5.3}
$$

In other words, the corner enhancement filter is anti-symmetric on either line $x = x_0$ or line $y = y_0$. $l(x, y)$ with $\sigma = 3$ is shown in Fig. 5.1, and it can be seen from it that the absolute values of $l(x, y)$ are centrally symmetric and decay along the radial direction in accordance with a Gaussian function. The decay speed can be controlled with σ. Moreover, $l(x, y)$ is anti-symmetric on the horizontal or vertical central axis.

Clearly, in smooth regions of images, pixel values are approximately equal. Since the filter function is anti-symmetric on the horizontal and vertical central axes, its

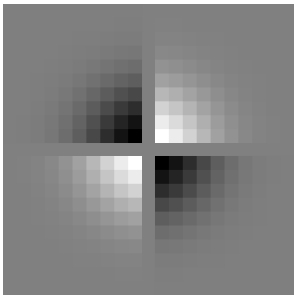
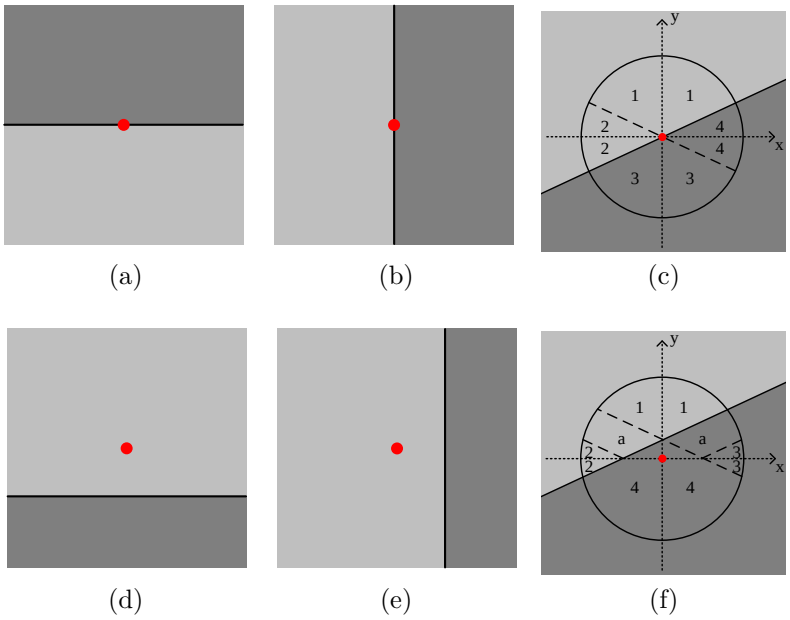


Figure 5.1: Proposed corner enhancement filter $l(x, y)$ with $\sigma = 3$.

response approaches zero in smooth regions. As for edges, there are six typical cases including pixels located exactly on a horizontal edge, a vertical edge or a slope edge, and close to a horizontal edge, a vertical edge or a slope edge. For pixels located exactly on a horizontal edge (shown in Fig. 5.2(a)) or close to a horizontal edge (shown in Fig. 5.2(d)), the values of their circumambient pixels are approximately symmetric on the vertical central axis, so the response of the filter approaches zero.

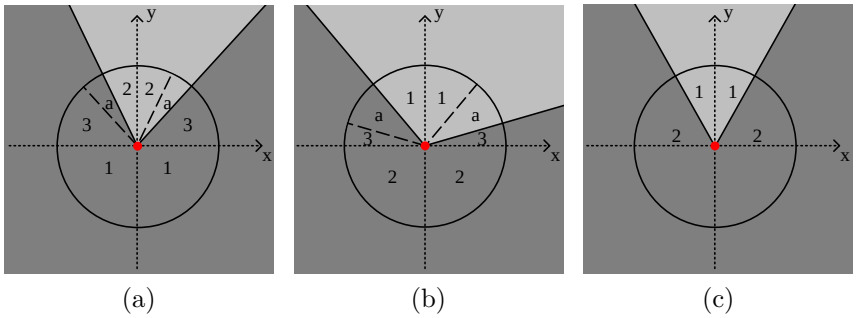


(a) (b) (c)

(d) (e) (f)

Figure 5.2: Six typical cases for edges. (a) Pixels located exactly on a horizontal edge, (b) on a vertical edge, (c) on a slope edges, (d) pixels close to a horizontal edge, (e) to a vertical edge, (f) to a slope edge.

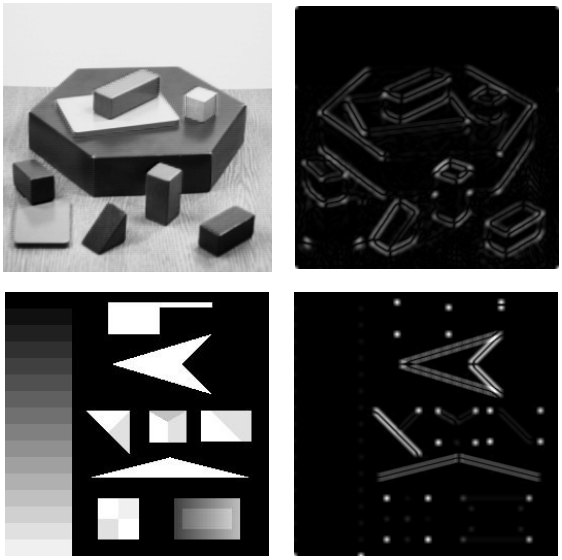
Similarly, in cases of pixels located exactly on a vertical edge (shown in Fig. 5.2(b)) and close to a vertical edge (shown in Fig. 5.2(e)), the values of their circumambient pixels are approximately symmetric on the horizontal central axis, and the response of the filter approaches zero as well. As for the case of pixels located exactly on a slope edge (shown in Fig. 5.2(c)), regions labelled with the same number are symmetric on the horizontal or vertical central axis in pairs, and the response of the filter approaches zero. Only in the case of pixels close to a slope edge as shown in Fig. 5.2(f), regions labelled with the same number are symmetric on the horizontal or vertical central axis in pairs, but the regions labelled with character 'a' are not symmetric. Therefore, the absolute values of the filter responses are generally greater than those of smooth areas and edges.



(a) (b) (c)

Figure 5.3: Three typical cases for corners. (a) Unsymmetrical acute corner, (b) unsymmetrical obtuse corner, (c) symmetrical corner.

For corners, there are two typical cases including unsymmetrical corners and symmetrical corners on the horizontal or vertical central axis. In the case of unsymmetrical corners as shown in Fig. 5.3(a) and (b), regions labelled with the same number are symmetric on the horizontal or vertical central axis in pairs, while regions labeled with character 'a' are not symmetric. Therefore, the absolute values of the filter responses at the vertices of corners are generally greater than those of smooth areas and edges. For symmetrical corners on the horizontal or vertical central axis



(a) (b)

Figure 5.4: Filtered results of two images. (a) Input images, (b) filtered results.

as shown in Fig. 5.3(c), regions labelled with the same number are symmetric on the horizontal or vertical central axis, and the response of filters at the vertices of this type of corner approaches zero.

Filtered results on two classical images frequently used in corner detection research are shown in Fig. 5.4. From the results, the following observations may be made. First, in smooth regions and regions with low textures, the responses of the filter approach zero. This demonstrates that the proposed filter can effectively suppress the background and low textured areas in images. Second, for many corners, the responses of the filter are higher than in other areas. The reason that some corners are not enhanced is that the corner is approximately symmetric on the horizontal or vertical central axis. This indicates that this filter has great ability on unsymmetrical corner enhancement. Third, at points located exactly on an edge, the responses of the filter approach zero. However, for points close to a slope edge, the absolute values of the filter responses are generally greater than that of smooth areas and edges.

Therefore, in the resultant images, the horizontal and vertical edges are effectively suppressed, and there are two bright lines located at two sides of a black slope edge line. This indicates that the proposed filter can effectively suppress various edges in images, but points located on two sides of slope edges have been enhanced as well. These issues will be dealt with later in Section 5.1.2.

Filtered results on two images corrupted by zero mean white Gaussian noise, multiplicative speckle noise and salt and pepper noise are shown in Fig. 5.5. As can be seen from the results, noise has been suppressed to a certain extent, and compared

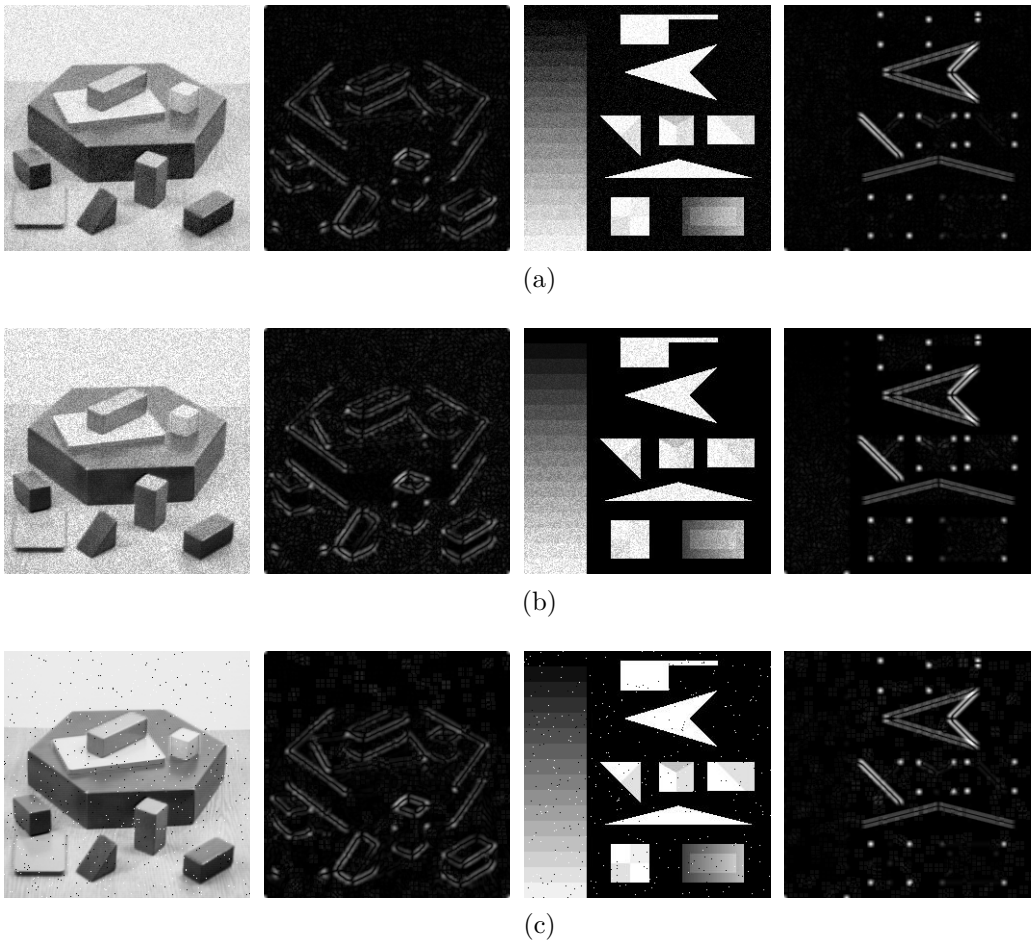


Figure 5.5: Filtered results of two images corrupted by three typical types of noise. (a) Gaussian noise, (b) multiplicative speckle noise, (c) salt and pepper noise.

with enhanced corners, the residual noise information is too weak to disturb subsequent corner detection. It may be concluded that the proposed filter has the ability of unsymmetrical corner enhancement and edge suppression. It also has ability to handle background variation, low texture area and noise suppression.

5.1.2 Proposed Corner Detector

As discussed in Section 5.1.1, the proposed filter has the ability for edge suppression and corner enhancement. However, there are two drawbacks that must be considered first. One is the issue of no enhancement on symmetrical corners, and the second is the enhancement of points close to a slope edge. For symmetrical corners with respect to the corner enhancement filter defined in Eq. (5.1), if the filter is rotated by 45°, as shown in Fig. 5.6, these corners are no longer symmetrical for the rotated filter, and the size of unsymmetrical regions labelled with character 'a' reaches a maximum. This means that the corners missed by the original filter can be enhanced further by a rotated filter. In light of this, the problem of no enhancement on symmetrical corners can be solved in the following manner: the input image is filtered by the proposed filter and the rotated version, and the maximum of the absolute values of the two filter responses on a pixel is selected to form the resultant

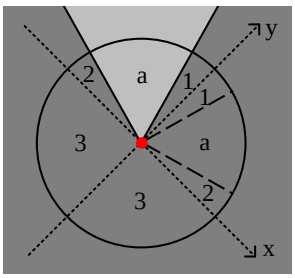


Figure 5.6: Enhancement of a symmetrical corner with respect to the corner enhancement filter defined in Eq. (5.1), with the corner enhancement filter rotated by 45°.

image.

Note that the proposed filter can effectively suppress various edges in images. The two bright lines located at two sides of some black edge lines are the result of enhancement of points close to the edges. Therefore, to overcome the second problem, the binary edge image of the input image is obtained by using the Canny edge detector and multiplied pixel by pixel with the resultant image from the two filters. By this means, the two bright lines located at two sides of a black edge line can be suppressed effectively. Some resultant images of these two steps are shown in Fig. 5.7. Note that these images are shown in colour according to the filter response

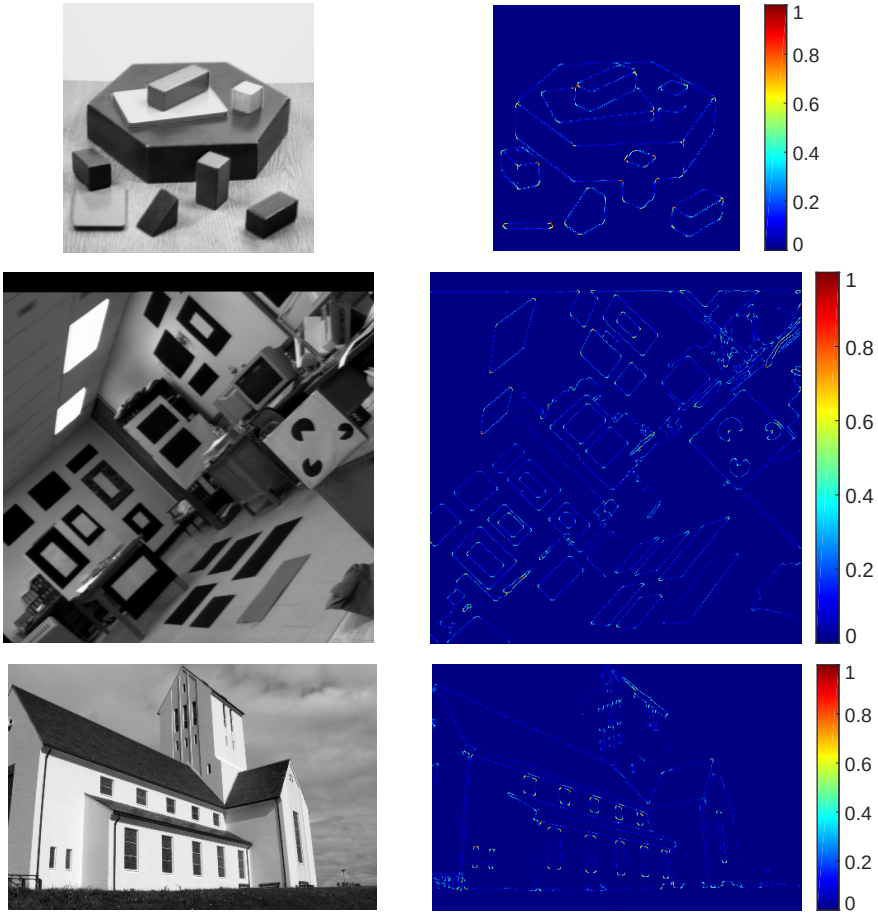


Figure 5.7: Some resultant images of the first two steps of Algorithm 5.1.

intensities of the pixels for viewing convenience. It can be seen that the corners have been enhanced remarkably, while the low textures have been suppressed completely. Although there is some residual edge information resulting from edge location error, in general this information is weaker than corners and can easily be distinguished from corners by an appropriate threshold.

As shown in Algorithm 5.1, the proposed corner detector consists of three steps. The objective of step 1 is to enhance corners, as well as suppress edges and smooth areas in an image. In this step, the image is filtered by the proposed filter and the rotated version, and the maximum of the absolute values of the two filter responses on a pixel is selected to form the resultant image. The purpose of step 2 is to suppress the enhancement of points close to a slope edge. In this step, the resultant image from step 1 is multiplied pixel by pixel with the binary edge image obtained by using the Canny edge detector. Then corners are detected with an appropriate threshold in step 3.

Algorithm 5.1 Fast Corner Detection Algorithm Based on Corner Enhancement Filters.

1: The input image is filtered by the proposed filter and the 45° rotated version, and the maximum of the absolute values of the two filter's responses on a pixel is selected to form the output image.

2: The binary edge image from the Canny edge detector is multiplied with the output image from step 1 pixel by pixel.

3: If the value of a pixel in the resultant image is the local maximum and larger than threshold T, the pixel is regarded as a corner.

5.2 Design of FPGA Architecture for Proposed Detector

The architecture design of the proposed detector is shown in Fig. 5.8, which consists of three modules including corner enhancement, edge detection and non-maximum suppression & thresholding. The first two modules deal with an input image in parallel. Their results are multiplied and pass through the non-maximum suppression & thresholding module, and then the corner map of the image can be finally obtained.

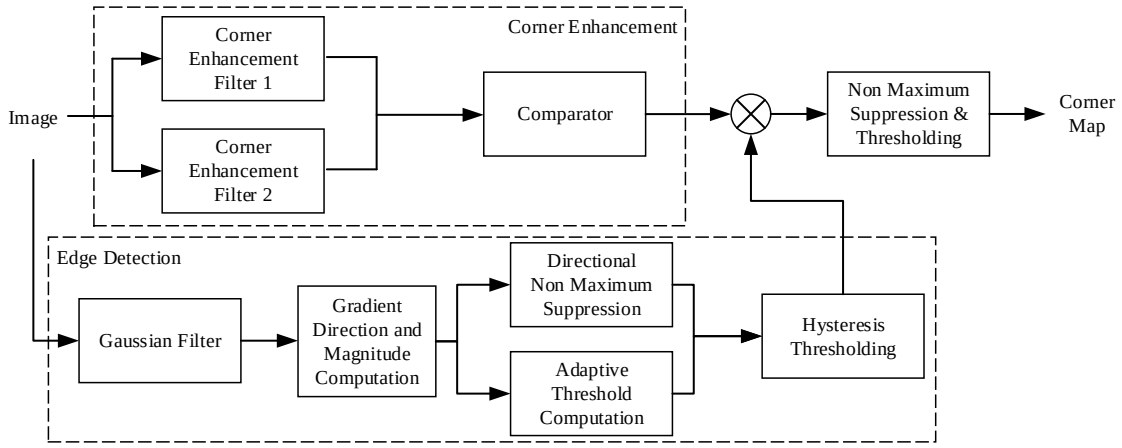


Figure 5.8: The architecture of the proposed detector.

5.2.1 Sliding Window

It is well known that for FPGA, images are input by rows in a linear manner. However, to compute convolution, non-maximum suppression and hysteresis threshold, computations within a sliding window are required. If the computations are performed by loading the corresponding window after the input image is ready, there would be a long delay and huge storage resource requirements. To synchronise the

computations with input of an image, a pipeline architecture of a sliding window is constructed, which can create an $M \times N$ window in the image, and can slide by one pixel to next window per clock cycle. Meanwhile, the required computations can be carried out at the same time as the input of the image with a few delays.

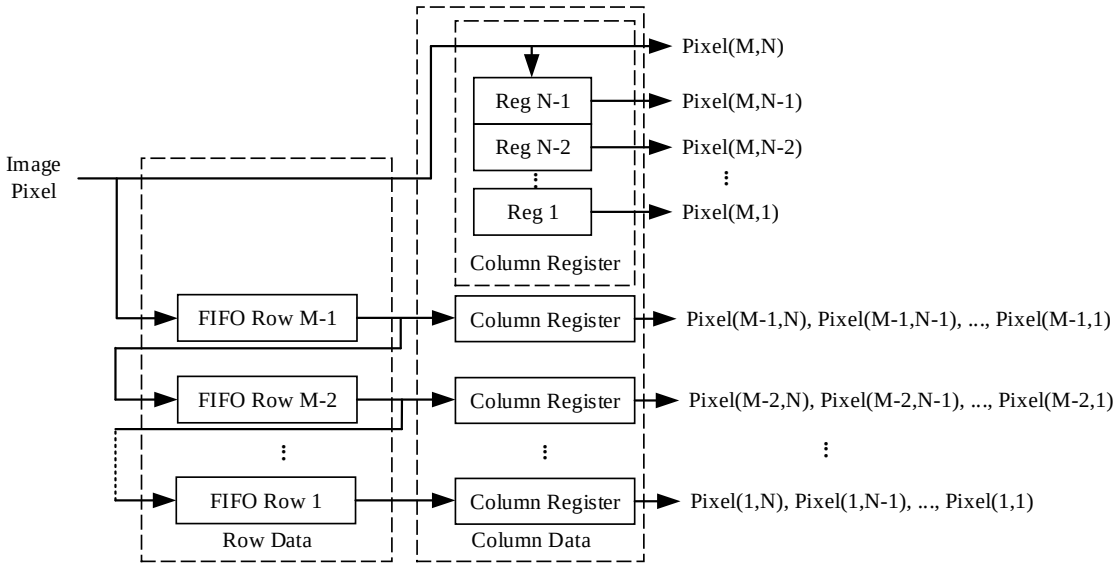


Figure 5.9: Pipeline architecture of sliding window.

The pipeline architecture of sliding window is shown in Fig. 5.9, which consists of $M-1$ cascaded FIFOs to store the image data in a cascading shift way. As the $((M+1)/2)$th FIFO is filled, the row data of the first sliding window is ready. When the first pixel of the image reaches the $((N+1)/2)$th register of the row, the first sliding window is established. Each register outputs the pixel data of the corresponding position at the same time and the required calculation can be completed through the corresponding arithmetic unit. Then, the window moves with the input of data in each clock cycle. In this way, only with a few delays, the required computations such as convolution can be achieved in the same time with the input of image data. The arithmetic unit only needs to multiply each output pixel of the sliding window module with its corresponding weight, and then to sum

them up. In this work, a 5×5 Gaussian filter is used for smoothing, two 3×3 directional gradient operators are adopted for gradient calculation and two corner enhancement filters with a size of 9×9 are used for corner enhancement.

5.2.2 Submodules of Edge Detection

After filtering the input image, the direction and magnitude of the gradient for each pixel are required to compute the directional non-maximum suppression result and the adaptive threshold. Then, the edge map of the input image can be obtained by hysteresis thresholding. The submodules designed for these calculations are presented in this subsection.

5.2.2.1 Magnitude and Direction of Gradient

After the calculation of gradients, the arctangent and square root functions are required to compute magnitude and direction for each pixel. However, as these functions are resource- and time-consuming in FPGA, a challenge is to find a fast way to simply perform the gradient magnitude and direction calculations in the pipeline of the edge detection module. To calculate the gradient magnitude, a square root approximation [156] with a low resource requirement is adopted, in which the magnitude can be approximately calculated by

$$M = \max((0.875a + 0.5b), a) \tag{5.4}$$

where

$$\begin{cases} a = \max(|f_x(x,y)|, |f_y(x,y)|) \\ b = \min(|f_x(x,y)|, |f_y(x,y)|) \end{cases} \tag{5.5}$$

In this way, the square root can be calculated simply by some compare, shift and add operations, which can be easily performed by bit operations. The direction of gradient can be estimated by comparing the $|f_y(x,y)|$ with $|f_x(x,y)|\tan 22.5°$ and $|f_x(x,y)|\tan 67.5°$. Here, $\tan 22.5°$ and $\tan 67.5°$ are approximated as

$$\begin{cases} \tan 22.5° = 0.4142 \approx \dfrac{1}{2} - \dfrac{1}{16} - \dfrac{1}{64} - \dfrac{1}{128} \\ \tan 67.5° = 2.4142 = 2 + \tan 22.5° \end{cases} \tag{5.6}$$

Thus, the multiplication of $|f_x(x,y)|$ with $\tan 22.5°$ or $\tan 67.5°$ can be efficiently calculated by shifting and adding operations. Then, the direction can be determined with a Look Up Table (LUT) as listed in Table 5.1, and the corresponding FPGA design is shown in Fig. 5.10.

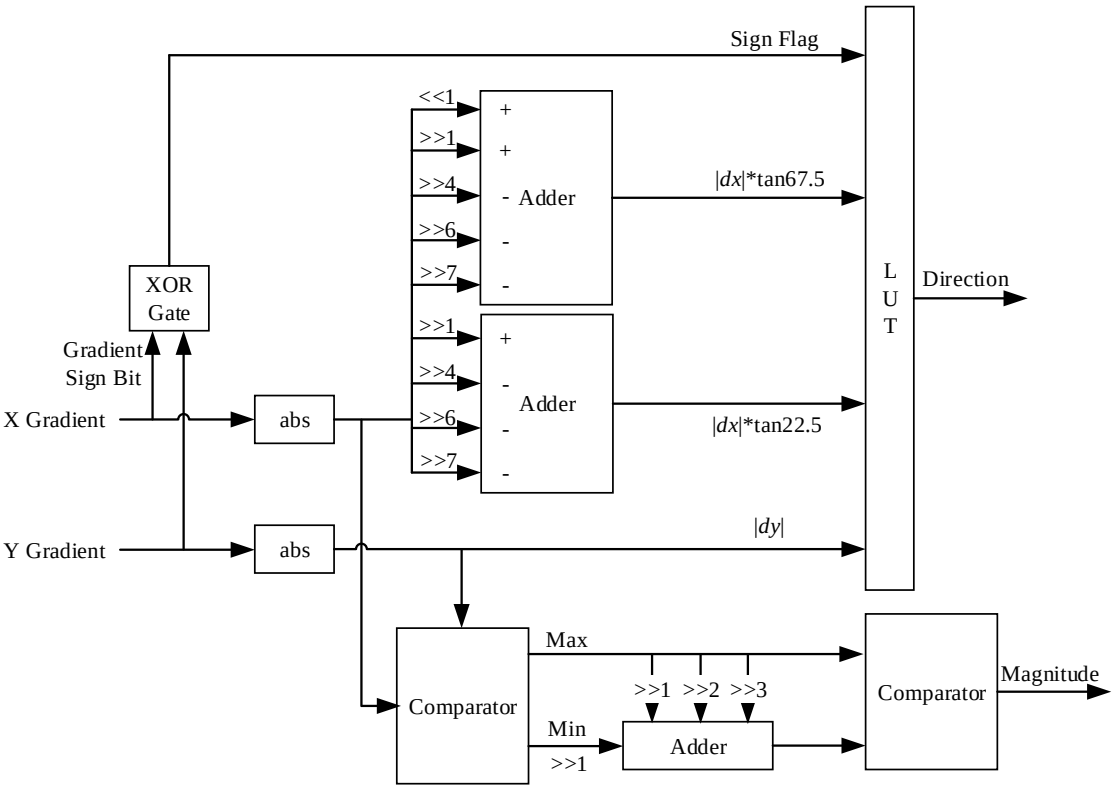


Figure 5.10: Gradient direction and magnitude computations.

Table 5.1: Direction decision in LUT.

Range of $f_x(x,y)$ and $f_y(x,y)$	Sign bit	Direction
$\lvert f_y(x,y)\rvert < \lvert f_x(x,y)\rvert \tan 22.5°$		$0°$
$\lvert f_y(x,y)\rvert > \lvert f_x(x,y)\rvert \tan 67.5°$		$90°$
$\lvert f_x(x,y)\rvert \tan 22.5° < \lvert f_y(x,y)\rvert < \lvert f_x(x,y)\rvert \tan 67.5°$	$(1,1)$ or $(0,0)$	$45°$
$\lvert f_x(x,y)\rvert \tan 22.5° < \lvert f_y(x,y)\rvert < \lvert f_x(x,y)\rvert \tan 67.5°$	$(1,0)$ or $(0,1)$	$-45°$

5.2.2.2 Adaptive Threshold Decision

In the adaptive threshold submodule (see Fig. 5.11), the threshold is determined by calculating the histogram of gradient magnitude. Each gradient magnitude is quantized to six bits (0-63) by a shifter and matched with the addresses of 64 registers. Then, the registers with address greater or equal to the magnitude will increase by 1. As the input of last magnitude has finished, the address of the register whose value just exceeds the product of the image size and a pregiven percentage is rebounded to its original range by a shifter. The result is output as the high threshold, and multiplied by a pregiven ratio to form the low threshold.

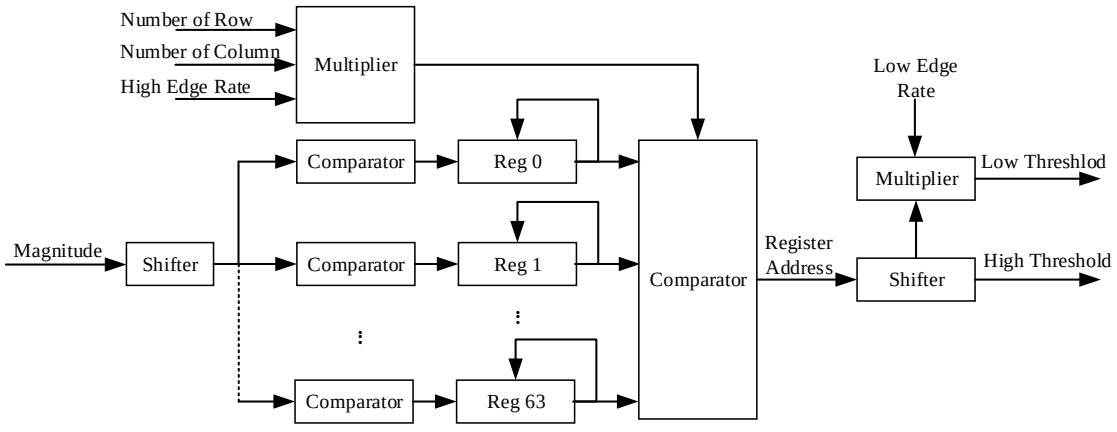


Figure 5.11: Adaptive threshold.

5.2.2.3 Directional Non-Maximum Suppression

Directional non-maximum suppression shown in Fig. 5.12 is executed in parallel with the adaptive threshold submodule. Note that the thresholds can only be output after the last magnitude processing is completed. Therefore, this submodule outputs its results to a buffer. The sliding window with 3×3 size is constructed as in Fig. 5.9. Based on the direction, the selector outputs the amplitudes of two pixels marked as a and b, which correspond to the specified direction in the sliding window. If the magnitude of this pixel is greater than a and b, the magnitude will be output, otherwise the output will be 0. The calculation of this submodule is carried out at the same time as the output of the magnitude and direction calculation submodule, which only contains an input delay caused by preparing a few rows of data for FIFOs.

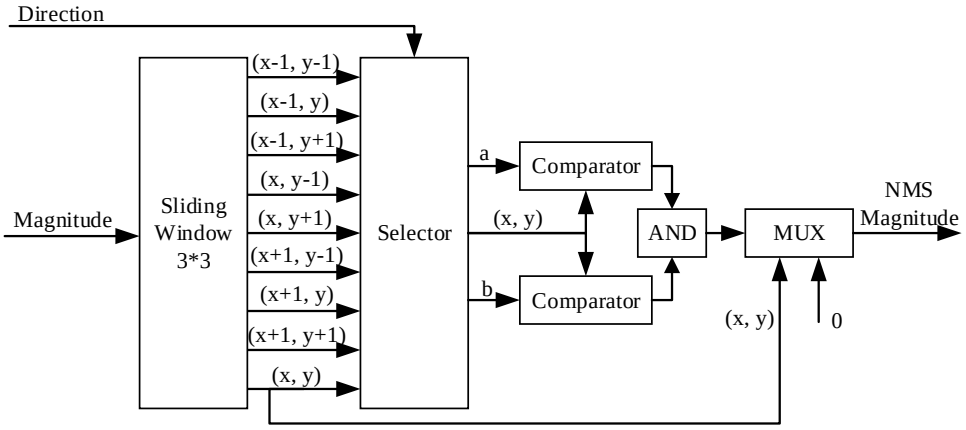


Figure 5.12: Directional non-maximum suppression.

5.2.2.4 Hysteresis Thresholding

In the hysteresis thresholding submodule shown in Fig. 5.13, the magnitude from the buffer is compared with the high threshold. If it is greater than the high threshold, the strong edge flag E_H is marked as 1, otherwise marked as 0. Meanwhile, the

magnitude is also compared with the low threshold. The corresponding weak edge flag E_L is stored at a FIFO, and then its connection with E_H in its 3×3 neighbour sliding window is checked by two logical gates. Finally, $E_H = 1$ or $E_L = 1$ connected with one or more $E_H = 1$ will be output as 1, otherwise output as 0 to form the final edge map.

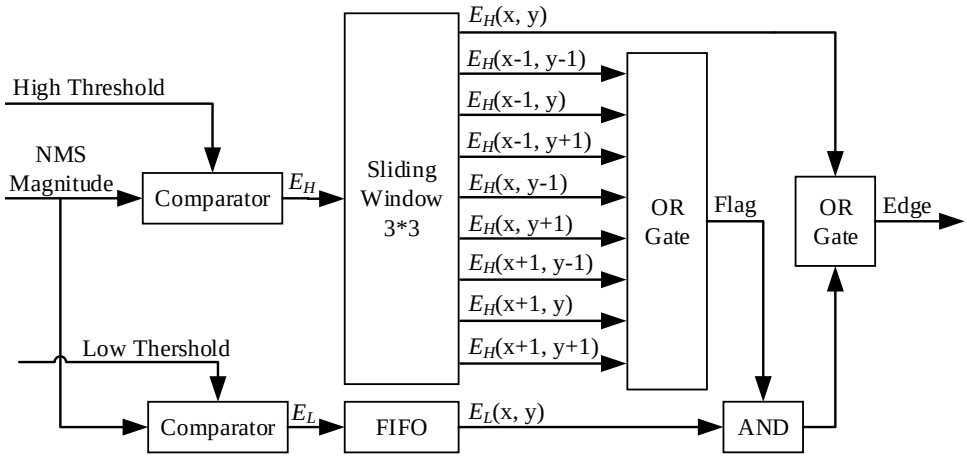


Figure 5.13: Hysteresis thresholding.

5.2.3 Detection of Corners

The module of non-maximum suppression and thresholding for corner detection is shown in Fig. 5.14. The 7×7 sliding window is constructed as in Fig. 5.9. If a pixel

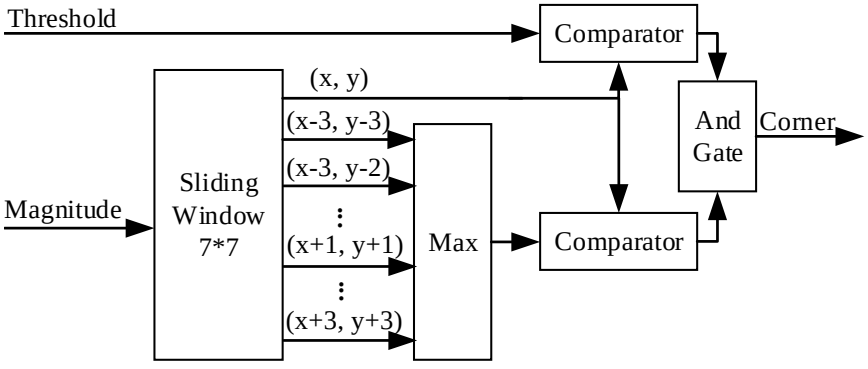


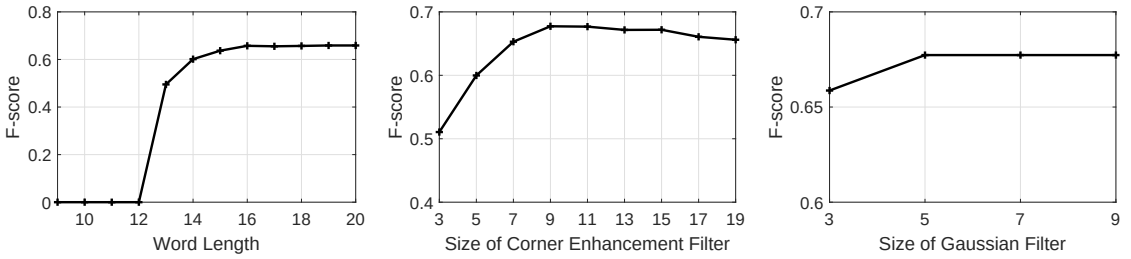
Figure 5.14: Non-maximum suppression and thresholding for corner detection.

has the largest magnitude in the sliding window and is greater than the threshold, the corresponding output of this pixel is marked as 1, otherwise it is marked as 0. The corner map of the input image can be obtained by the results of this module.

5.2.4 Relevant Parameters

In this design, the relevant parameters are the word length, sizes of the corner enhancement filter and Gaussian filter in edge detection, which affect its resource consumption and detection performance. For example, too short a word length will produce a large calculation error, and too long a word length will increase resource consumption and computation time. Also, the sizes of the corner enhancement filter and Gaussian filter determine the resource consumption and time delay of the cascaded FIFOs in Fig. 5.9. To set these parameters, a dataset with ground truth including 7 real images and 13 synthetic images introduced in Section 3.4.1.1 of Chapter 3 were considered, and some of them are shown in Fig. 3.9. The first four synthetic and two real images in this figure were used for setting parameters, and the *F-score* defined in Eq. (3.29) was used as the metric.

The experiments were performed strictly following the fixed point limitations and operations in this design by using MATLAB 2018a. The average *F-scores* on the six images with word length varying from 9 to 20, 7×7 corner enhancement filter and 7×7 Gaussian filter are summarized in Fig. 5.15(a). It can be seen that as the word length increases from 9 to 16, the detection accuracy increases due to the improvement in calculation precision. However, when the word length exceeds 16, the contribution of the improvement of calculation precision to the detection accuracy is minimal. Therefore, the word length was set as 16. The average *F-scores* on the six images with 16 bit word length, 7×7 Gaussian filter, and the size of the corner



(a) (b) (c)

Figure 5.15: Module parameter setting. *F-score* curves on (a) number of bits, (b) size of corner enhancement filter, (c) size of Gaussian filter.

enhancement filter varying from 3 to 19 are summarized in Fig. 5.15(b). It can be seen from Fig. 5.15(b) that as the size of the corner enhancement filter increases from 3 to 9, more relevant surrounding pixels are considered and the *F-score* increases as well, and the slow decline of detection accuracy from 11 to 19 indicates that the introduction of too distant pixels may be unfavourable to detection accuracy. Therefore, the size of the corner enhancement filter was set to 9×9. Similarly, the average *F-scores* on the six images with 16 bit word length, 9×9 corner enhancement filter, and the size of the Gaussian filter varying from 3 to 7 are summarized in Fig. 5.15(c). Based on this figure, the size of the Gaussian filter is set to 5×5.

The proposed detector with 16 bit word length, 9×9 corner enhancement filter and 5×5 Gaussian filter, which is regarded as the fixed point base case, was simulated using Xilinx Vivado software based on Xilinx's Virtex-5 XC5VLX155T FPGA. The simulation results for a 256×256 image are listed in Table 5.2. As can be seen, the overall processing time is 1.33 ms, which is sufficient for real-time applications. From the aforementioned designs, the reduction of the computation time comes from three aspects: (1) In Fig. 5.8, the corner enhancement module processes the input image in parallel with the edge detection module. Furthermore, in the corner enhancement module, the two corner enhancement filters deal with the image in parallel, and

in the edge detection module, the directional non-maximum suppression submodule processes the image in parallel with the adaptive threshold computation submodule. It is well known that for two parallel modules, only the computation time of the slower one is considered in the computation time for the entire system. (2) Generally, the computation time of a system equals to the product of the data size and the individual datum processing time. But for the pipeline architecture in FPGA, its computation time is the required time when all the data were passed through the pipeline architecture. So, its total computation time is the sum of the data size and the individual datum processing time as well as the delay time in the pipeline architecture. (3) The design of the pipeline architecture for sliding windows in Fig. 5.9 minimizes the required delays for convolution, non-maximum suppression and hysteresis threshold computations as far as possible.

Table 5.2: Resource occupancy and execution time (for 256×256 image).

Used Slices	Used Slice LUTs	Frequency	Computation Time
7728/24320	15912/97280	100 MHz	1.33 ms

5.3 Performance Evaluation and Experimental Analysis

In this section, the performance of the proposed method is evaluated on benchmark datasets. It is well known that detection accuracy, computational cost and repeatability are the main factors that determine the application potential of corner detectors. Therefore, the performance of the proposed corner detector was evaluated on these three factors. A comparison of this corner detector with other representative corner detectors including Harris [24], FAST [53], CF [81], ANDD [80], ACJ [79],

SMCD [108], Zhang and Sun's method [94] and the method in Chapter 3, as well as the current deep learning-based interest point detectors D2-Net [113] and Reinforced SP [111] were performed. In order to comprehensively verify the performance of the proposed detector, the results of its fixed point version are provided as well. In this case, the detector used the fixed point base as described in Section 5.2. The proposed method was simulated and the evaluation experiments were conducted by using MATLAB 2018a on a personal computer with Intel Core i5-6500 CPU and 4 GB memory, and the operation system was Microsoft Windows 10 Enterprise.

5.3.1 Corner Detection Accuracy and Computational Cost

In this subsection, corner detection accuracy and computational cost were verified. A total of 20 different images with ground truth including 7 real images and 13 synthetic images introduced in Section 3.4.1.1 of Chapter 3 were used. The corner detection accuracy includes two aspects: corner detection and localisation accuracies. For detection accuracy, *F-score* and *precision* were used as the metrics, and the corner localisation error L was utilised to evaluate corner localisation accuracy.

In this experiment, four synthetic and two real images in Fig. 3.9 were utilised to set the threshold parameter. The threshold of each detector was set to the value at which the detector achieves the maximum average *F-score* on the six images. The remaining fourteen images are used for evaluation. The average and maximum *precision*, *F-score* and average localisation error L on all the 14 images are listed in Table 5.3, and average computation times are listed in Table 5.4. Note that the average computation time of each detector listed in Table 5.4 is calculated with the average detection time of 10 repeated detections per image for all images. From the results, the following conclusions may be drawn. First, Harris, FAST and the

proposed detectors are fast corner detectors, and their computational costs are remarkably less than that of the other detectors. The proposed detector has a similar detection speed to the FAST detector, but it has half the computational cost compared with the Harris detector. Its corner localisation accuracy is slightly lower compared with the Harris detector, but higher than that of the FAST detector. Furthermore, the proposed detector outperforms the Harris and FAST detectors in terms of average and maximum *precision* and *F-score*. Considering all these factors, the proposed detector is more advantageous than the two fast detectors for real-time computer vision tasks. Second, the average *precision* of the proposed detector is higher in comparison with all compared detectors except for the method in Chapter 3, and the average *F-score* is higher than that of some current detectors including the contour-based detectors CF [81] and ACJ [79], and the multi-scale analysis based detector SMCD [108]. This indicates that the proposed detector can achieve or even exceed the detection accuracy of some current corner detectors with very low computational cost. Hence it is potentially useful as a fast corner detector

Table 5.3: Detection performance and localisation accuracy.

Detectors	*Precision* (%)		*F-score* (%)		Localization error
	Avg.	Max.	**Avg.**	Max.	
Harris [24]	60.59	85.45	63.91	80.99	1.538
FAST [53]	53.93	86.49	61.56	75.81	1.644
CF [81]	74.55	96.77	64.20	77.23	1.526
ANDD [80]	76.37	96.88	73.48	93.33	1.446
ACJ [79]	66.87	87.10	67.86	87.72	2.173
SMCD [108]	70.73	84.91	66.44	80.36	2.071
Zhang and Sun [94]	83.10	94.44	74.10	84.40	1.748
Method in Chapter 3	86.48	94.59	77.37	95.89	1.433
Proposed	83.93	92.86	68.34	84.62	1.618
Proposed (Fixed Point)	83.85	90.91	67.73	87.67	1.646

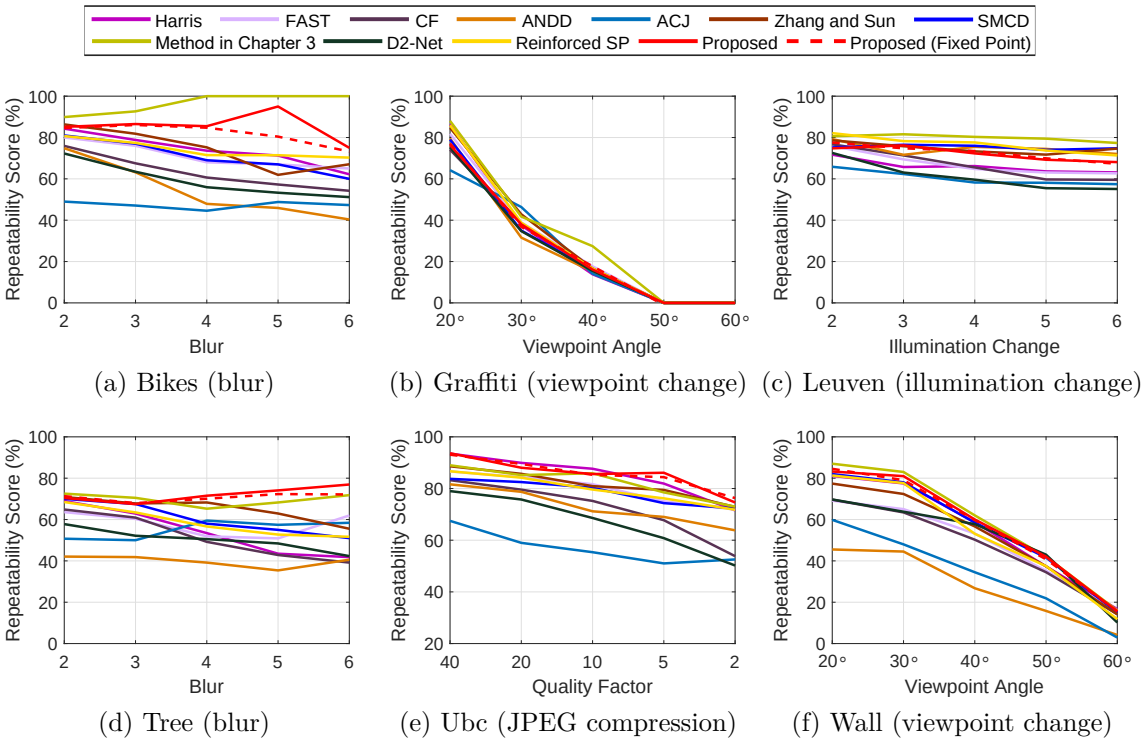
Table 5.4: Computational cost of the evaluated corner detectors.

Detectors	Computation time (s)
Harris [24]	0.045
FAST [53]	0.019
CF [81]	0.278
ANDD [80]	6.515
ACJ [79]	3.617
SMCD [108]	0.591
Zhang and Sun [94]	0.942
Method in Chapter 3	0.997
Proposed	0.022
Proposed (Fixed Point)	0.024

for computer vision applications. Third, the loss of computation precision in the fixed point base case generally leads to a slight drop in the corner detection and localization accuracies. The increase of maximum *F-score* shows that, for a specific image, this loss is not always harmful to the detection accuracy especially the *recall* rate. Note that the computation time of the fixed point version in Table 5.4 includes the time consumption for converting floating-point operations to fixed-point.

5.3.2 Evaluation of Repeatability

The objective is to verify the repeatability of the proposed detector under image viewpoint change, scaling, rotation, blurring, JPEG compression, noise corruption and illumination change. Three datasets used to evaluate repeatability in Section 3.4.1.2 were considered. The repeatability score RS was used as the metric. Note that it is necessary to ensure that all detectors acquire a similar number of detected interest points in the first frame of each image sequence in evaluating re-



(a) Bikes (blur) (b) Graffiti (viewpoint change) (c) Leuven (illumination change)

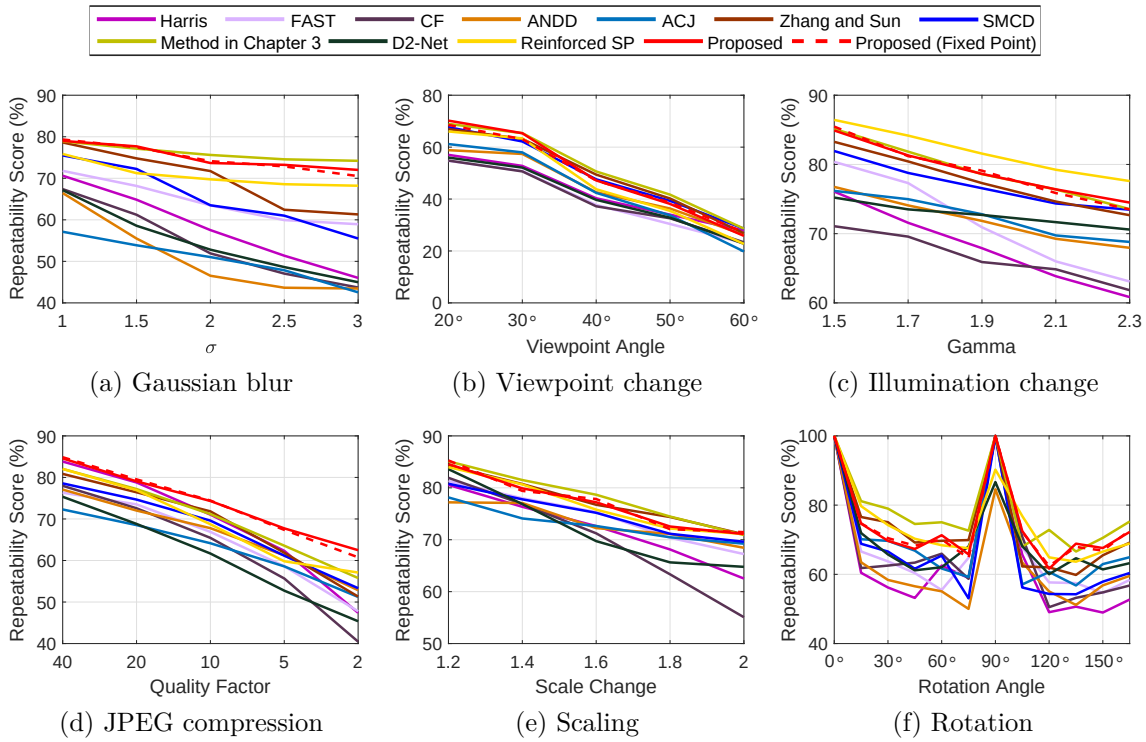
(d) Tree (blur) (e) Ubc (JPEG compression) (f) Wall (viewpoint change)

Figure 5.16: Repeatability scores of the proposed detector and the compared detectors on Dataset I. (a) 'Bikes', (b) 'Graffiti', (c) 'Leuven', (d) 'Tree', (e) 'Ubc', (f) 'Wall'.

peatability. Thus, for each image sequence in this experiment, the thresholds of all detectors were finetuned to ensure that all detectors acquire a similar number of detected interest points from the first frame.

Repeatability scores on the sequences in Dataset I are shown in Fig. 5.16. In Datasets II and III, there are 50 or 56 subsequences in each style of sequences, so the average repeatability scores on the 50 or 56 subsequences were determined and summarised in Fig. 5.17 and Fig. 5.18. As can be seen from these figures:

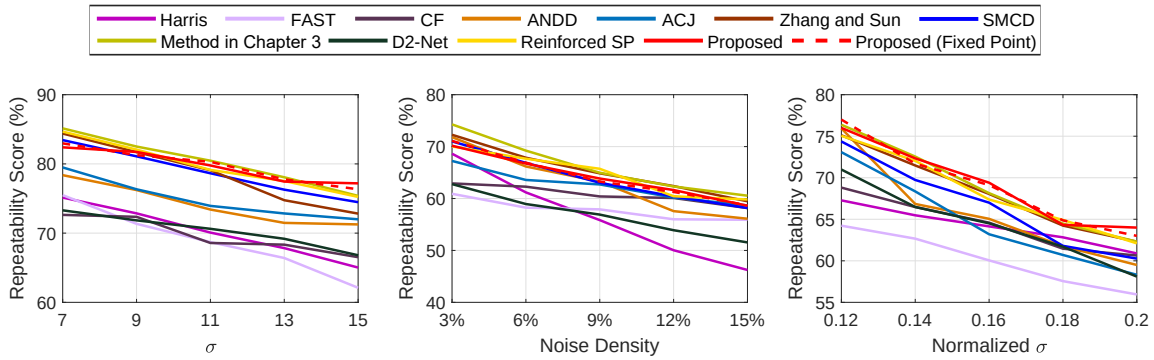
1. Detection repeatability under image blurring for the proposed detector is only lower than the method in Chapter 3, but higher than the other corner detectors and current deep learning-based interest point detectors (see Fig. 5.16(a), (d)



(a) Gaussian blur (b) Viewpoint change (c) Illumination change

(d) JPEG compression (e) Scaling (f) Rotation

Figure 5.17: Repeatability scores of the proposed detector and the compared detectors on Dataset II. (a) Image blurring, (b) image viewpoint change, (c) illumination change, (d) JPEG compression, (e) image scaling, (f) image rotation.



(a) Gaussian noise (b) Salt and pepper noise (c) Multiplicative speckle noise

Figure 5.18: Repeatability scores of the proposed detector and the compared detectors on Dataset III. (a) Gaussian noise, (b) salt and pepper noise, (c) multiplicative speckle noise.

and Fig. 5.17(a)).

2. Repeatability under viewpoint change for the proposed detector is slightly

lower than multi-scale analysis based detectors, but higher than the other detectors (see Fig. 5.16(b), (f) and Fig. 5.17(b)).

3. For JPEG compression, the repeatability of the proposed detector is generally higher than the other methods (see Fig. 5.16(e) and Fig. 5.17(d)).

4. For most images under illumination change, the deep learning-based interest point detector Reinforced SP [111] achieves the highest repeatability, and the proposed detector obtains the second-highest repeatability (see Fig. 5.17(c) for illumination). For the specified image sequence 'Leuven', its repeatability is lower than Reinforced SP [111] and a little short of the repeatability of the multi-scale analysis based detectors (see Fig. 5.16(c)).

5. Repeatability under image scaling, rotation and salt and pepper noise is similar to that of Reinforced SP [111] and two multi-scale analysis based detectors including Zhang and Sun's method [94] and the method in Chapter 3, but higher than the other detectors (see Fig. 5.17(e), (f) and Fig. 5.18(b)).

6. For images with Gaussian and multiplicative speckle noise, repeatability of the proposed detector is similar to that of the method in Chapter 3 and generally higher than other detectors.

It can be concluded that, with a very low computational cost and simple architecture, the repeatability of the proposed detector is similar to some of the multi-scale analysis based detectors and the machine learning-based interest point detector Reinforced SP [111], and clearly higher than other types of corner detectors as well as the machine learning-based interest point detector D2-Net [113]. Moreover, the repeatability of the fixed point version is similar to that of the floating-point detector.

5.4 Remarks

In this chapter, a novel fast corner detector was proposed, based on the novel idea of corner enhancement filtering. Experimental results demonstrate that the proposed detector has a similar detection speed to the FAST detector, and twice the detection speed of the Harris detector. It was also shown that the proposed corner detector performs notably better than existing fast corner detectors such as Harris and FAST in terms of detection accuracy, tolerances to noise and robustness to affine transformations, illumination changes, noise and JPEG compression. Therefore, it is more advantageous for real-time computer vision tasks. Moreover, with a very simple architecture and low computational cost, the proposed detector can achieve or even exceed the detection accuracy of some current corner detectors. Its repeatability is similar to some of the multi-scale analysis based detectors and the deep learning-based interest point detector Reinforced SP [111], and clearly higher than that of other types of corner detectors as well as the deep learning-based interest point detector D2-Net [113]. Therefore, it is potentially useful as a fast corner detector for many computer vision applications.

Chapter 6

Conclusions and Future Work

Corner detection is a fundamental computer vision problem that has been widely studied in many computer vision tasks. Current corner detectors cannot provide sufficient detection accuracy, repeatability and localisation ability. Based on a detailed analysis of current multi-scale analysis based corner detectors, this thesis addresses the challenging problem of improving corner detection accuracy, repeatability and localisation ability simultaneously while also reducing computational complexity.

6.1 Thesis Summary

In order to improve corner detection accuracy, repeatability and localisation ability simultaneously in images, and overcome the weaknesses of edge extension, bilateral margin responses for prominent edges and different numbers of directions at different scales in the traditional shearlets, in Chapter 3, a new shearlet transform has been proposed, which has a greater ability to localise distributed discontinuities, can

effectually mitigate the weaknesses of the traditional shearlets, and can provide more flexible direction divisions. To make full use of the multi-scale and multi-directional information of the proposed shearlets for corner detection, in Chapter 3, a novel multi-directional structure tensor is constructed at the finest scale to detect candidate corners, and a multi-scale corner measurement function is proposed to remove false candidate corners. Experimental results demonstrate that the proposed corner detection method outperforms current corner detectors in corner detection and localisation accuracies. It is also shown that its detection repeatability under image blurring, viewpoint changes, affine transformations, illumination changes and parts of JPEG compression is generally superior to not only current corner detectors but also representative interest point detectors.

While the proposed corner detector in Chapter 3 outperforms current corner detectors in terms of corner detection accuracy, repeatability as well as localisation ability, its computational cost is higher than that of derivative-based methods and most multi-scale analysis based detectors. Chapter 4 attempts to improve detection accuracy, repeatability and localisation ability further while reducing computational complexity as much as possible. In Chapter 4, a novel shearlet transform, namely complex shearlet transform, is proposed by redefining the generation functions and the discretisation process. Different from traditional shearlets, complex shearlets have greater ability to localise distributed discontinuities and can capture phase information of geometrical features. Moreover, in light of the phase congruence model [140, 141] and statistical characteristics of wavelet coefficients in detail components [150], a type of phase congruence function is introduced, and its tolerance to noise and ability of corner localisation are improved by screening and normalising the amplitude information. Then, for corner detection, a novel rotary phase congruence structure tensor is proposed to properly merge the amplitude and phase

information of all scales into a 2×2 structure tensor. With less sensitivity to image scaling, rotation, blurring and noise, the rotary phase congruence structure tensor can make full use of all amplitude and phase information of the complex shearlets to detect corners. In addition, fast and direct processing of the rotary phase congruence structure tensor is presented, which can greatly reduce the computational cost. Experimental results demonstrate that the proposed method in Chapter 4 presents improvements to the current state-of-the-art in corner detection on corner localisation ability, detection accuracy and repeatability under image scaling, rotation, JPEG compression, noise corruption, illumination changes and viewpoint changes with the lowest computational cost in all multi-scale analysis based detectors.

Finally, in Chapter 5, the thesis also addresses the issue of providing a simple and computationally efficient corner detector with sufficient detection accuracy, repeatability and high adaptability to achieve real-time detection. In order to simplify the corner detection architecture and improve its parallel computing performance, a new type of filter is proposed that can enhance corners and suppress edges as well as noise simultaneously. Then a novel efficient corner detector is proposed, which can easily achieve real-time detection with high performance. The corresponding FPGA design is also presented. Experimental results show that, with very low computational cost and simple architecture, the proposed detector can achieve or even exceed the detection accuracy of multi-scale analysis based detectors, and its repeatability is similar to that of multi-scale analysis based detectors and clearly higher than other types of corner detectors.

6.2 Contributions

Based on detailed analysis of current multi-scale analysis based corner detectors, this thesis addresses the challenging problem of improving corner detection accuracy, repeatability and localisation ability simultaneously while reducing computational complexity as much as possible. It also addresses the urgent need to provide a simple and computationally efficient corner detector with sufficient detection accuracy, repeatability and high adaptability to achieve real-time detection. The main contributions are summarised as follows:

1. Two novel multi-scale and multi-directional analysis tools, namely improved shearlet transform and complex shearlet transform, are proposed, which can capture both amplitude and phase information, effectively mitigate the weaknesses of traditional shearlets and extract clear and accurate multi-scale and multi-directional structural information from images. The detailed structural information provided by the two proposed shearlet transforms can improve the corner detection and localisation accuracy, and the use of multi-directional and multi-scale information especially phase information can greatly improve the robustness of detection and the ability to distinguish corners from other pixels. Thus, they are potentially useful in corner detection. Moreover, because the proposed two shearlets have a greater ability to obtain detailed and high frequency information, and can effectually mitigate the weaknesses of the traditional shearlets, they are potentially useful as a multi-scale decomposition and analysis approach for many tasks in the field of computer vision such as detection and analysis of edges and ridges, image compression, image denoising, interpolating and watermarking.

2. Two corner detectors with two novel structure tensors that makes full use of multi-directional and multi-scale structural information are proposed, which can provide better ability to distinguish corners from other pixels with high robustness. Experiments show that the two proposed corner detectors present improvements to current state-of-the-art in corner detection as well as the recent deep learning interest point detectors on most key indicators including detection accuracy, localisation accuracy and repeatability under image scaling, rotation, noise corruption, JPEG compression, illumination changes and viewpoint changes. Moreover, the detector proposed in Chapter 4 yields improvement in computational efficiency, and the proper usage of phase information may be helpful in improvements to the detection accuracy and localisation ability. Thus, the two proposed detectors have great potential to be applied as robust and accurate feature detectors in many computer vision tasks such as image retrieval, object tracking, motion estimation, and visual localization.

3. A novel efficient corner detector with a simple architecture and high parallel computing characteristics is presented, which provides similar detection accuracy and repeatability as multi-scale analysis based detectors and has a good adaptability to achieve real-time detection. It is potentially useful as an efficient corner detector for computer vision applications, especially for portable real-time tasks.

4. The ideas of enhancing corners and suppressing edges at the same time by directly filtering raw images, and the rotary phase congruence structure tensor for properly merging all the amplitude and phase information into a 2×2 structure tensor with high stability for corner detection are novel and different from the existing corner detectors, offering a new lead for corner detection research.

6.3 Limitations and Future Work

This thesis addresses the challenging problem of improving corner detection accuracy, repeatability and localisation ability simultaneously while reducing computational complexity as much as possible. However, there are still some existing limitations. The proposed two shearlet transforms decay too fast on coarse scales, the two proposed multi-scale analysis based corner detectors are too complex for real-time detection, and the proposed efficient detector still has room for performance improvement. Meanwhile, the proposed detectors have great potential for applications in many computer vision tasks such as image retrieval, object tracking, motion estimation, visual localization and 3D reconstruction. However, the potential extensions for a feature descriptor and further applications have not been presented in this thesis. Therefore, potential future research arising from this thesis is summarised as follows.

First of all, the proposed improved shearlet transform and complex shearlet transform have greatly improved the ability to obtain structural information and locate geometric features from images compared with the traditional shearlet transform. However, the obtained detailed structural information decays too fast as the scale becomes coarser, which restricts the choice of available scales. Moreover, compared to shearing, rotation is more natural in image processing. The use of rotation instead of shearing to construct a filter bank of shearlets may provide better performance. Therefore, further improving the performance of complex shearlet transform is one of the potential research works in future. The application of the shearlet transform in computer vision tasks can also be further explored, such as detection and analysis of edges and ridges, image compression, image denoising, interpolating and watermarking.

The two proposed multi-scale analysis based corner detection algorithms yield some improvements to the current state-of-the-art in corner detection on localisation ability, detection accuracy and repeatability. They have great potential for extension as feature descriptors and be applied in many computer vision tasks. It is also shown that the computational efficiency of the detector proposed in Chapter 4 is better than that of the current multi-scale analysis based corner detectors. However, for some real-time computer vision tasks especially real-time tasks on portable devices, the proposed two methods may not meet the requirements with regard to low computational cost, parallel computing, and simple architecture. Thus, how to further simplify the algorithm architecture and to improve their computational efficiency and parallel computing characteristics so as to make them suitable for hardware implementation such as on FPGA will be focused in future work. Furthermore, the extension of the proposed detectors to local feature descriptors could be addressed, and other possible applications of corner detectors could also be investigated.

The proposed fast corner detector can provide extremely high computational efficiency. However, the detection performance and repeatability still leave room for improvement. Extending the proposed corner enhancement filter to multiple directions and scales may be beneficial to detection performance, and decomposing the corner enhancement filter into a combination of multiple filters with a better ability to obtain detailed information is also one of the possible research directions. In future work, how to further improve its detection accuracy, localisation accuracy and repeatability could be addressed. Some potential applications in real-time computer vision and image processing tasks could be further investigated.

Last but not the least, for the application of corner detection, feature descriptors are necessary. In this regard, deep learning has achieved some advantages in recent

years. However, contrary to the idea of sparse features, most existing deep learning algorithms follow the framework of first description and then detection, they do not provide sufficient memory efficiency. Therefore, it is possible to consider combining corner detection with deep learning to propose an architecture that is different from existing deep learning interest point detection algorithms. The new framework could first perform detection and then description to provide sufficient detection accuracy and memory efficiency, while improving the repeatability and robustness of the feature matching.

References

[1] P. Loncomilla, J. R. del Solar, and L. Martínez, "Object recognition using local invariant features for robotic applications: A survey," *Pattern Recognition*, vol. 60, pp. 499–514, 2016.

[2] Y. Li, S. Wang, Q. Tian, and X. Ding, "A survey of recent advances in visual feature detection," *Neurocomputing*, vol. 149, pp. 736–751, Feb. 2015.

[3] M. Kumar, P. Chhabra, and N. K. Garg, "An efficient content based image retrieval system using BayesNet and K-NN," *Multimedia Tools and Applications*, vol. 77, no. 16, pp. 21 557–21 570, 2018.

[4] J. Huang, G. Zhou, X. Zhou, and R. Zhang, "A new FPGA architecture of FAST and BRIEF algorithm for on-board corner detection and matching," *Sensors*, vol. 18, no. 4, 2018.

[5] M. Zhu, W. Wang, B. Liu, and J. Huang, "A fast image stitching algorithm via multiple-constraint corner matching," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–6, 2013.

[6] K. Dohi, Y. Yorita, Y. Shibata, and K. Oguri, "Pattern compression of FAST corner detection for efficient hardware implementation," in *2011 21st Inter-*

national Conference on Field Programmable Logic and Applications, 2011, pp. 478–481.

[7] G. Klein and D. Murray, "Parallel tracking and mapping for small AR workspaces," in *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007, pp. 225–234.

[8] S. Gauglitz, T. Höllerer, and M. Turk, "Evaluation of interest point detectors and feature descriptors for visual tracking," *International Journal of Computer Vision*, vol. 94, no. 3, p. 335, 2011.

[9] F. Ferreira, G. Veruggio, M. Caccia, and G. Bruzzone, "ROV vision-based motion estimation: a comparison study," *IFAC Proceedings Volumes*, vol. 45, no. 22, pp. 96–101, 2012.

[10] A. Gil, O. M. Mozos, M. Ballesta, and O. Reinoso, "A comparative evaluation of interest point detectors and local descriptors for visual SLAM," *Machine Vision and Applications*, vol. 21, no. 6, pp. 905–920, 2010.

[11] C. A. V. Hernández and F. A. P. Ortiz, "A corner detector algorithm for feature extraction in simultaneous localization and mapping." *Journal of Engineering Science & Technology Review*, vol. 12, no. 3, 2019.

[12] A. Schmidt, M. Kraft, and A. Kasiński, "An evaluation of image feature detectors and descriptors for robot navigation," in *International Conference on Computer Vision and Graphics*, Warsaw, Poland, Sept. 2010, pp. 251–259.

[13] M. Wu, S. Lam, and T. Srikanthan, "A framework for fast and robust visual odometry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3433–3448, 2017.

[14] J. Zhou, J. Yan, T. Wei, K. Wu, X. Chen, and S. Hu, "Sharp corner/edge recognition in domestic environments using RGB-D camera systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 10, pp. 987–991, 2015.

[15] C. Jian, X. Xiang, and M. Zhang, "Mobile terminal gesture recognition based on improved FAST corner detection," *IET Image Processing*, vol. 13, no. 6, pp. 991–997, 2019.

[16] L. Puglia, M. Vigliar, and G. Raiconi, "Real-time low-power FPGA architecture for stereo vision," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, no. 11, pp. 1307–1311, 2017.

[17] W. Prawira, E. Nasrullah, S. R. Sulistiyanti, and F. X. A. Setyawan, "The detection of 3D object using a method of a Harris corner detector and Lucas-Kanade tracker based on stereo image," in *2017 International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2017, pp. 163–166.

[18] M. Muja and D. G. Lowe, "Scalable nearest neighbor algorithms for high dimensional data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2227–2240, 2014.

[19] T. Sattler, W. Maddern, C. Toft, A. Torii, L. Hammarstrand, E. Stenborg, D. Safari, M. Okutomi, M. Pollefeys, J. Sivic, F. Kahl, and T. Pajdla, "Benchmarking 6DOF outdoor visual localization in changing conditions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[20] T. Sattler, Q. Zhou, M. Pollefeys, and L. Leal-Taixe, "Understanding the limitations of CNN-based absolute camera pose regression," in *Proceedings*

of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.

[21] J. L. Schonberger, H. Hardmeier, T. Sattler, and M. Pollefeys, "Comparative evaluation of hand-crafted and learned local features," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[22] H. P. Moravec, "Towards automatic visual obstacle avoidance," in *International Conference on Artificial Intelligence*, Massachusetts, 1977, p. 584.

[23] H. P. Moravec, "Visual mapping by a robot rover," in *Proceedings of the 6th International Joint Conference on Artificial Intelligence*, Tokyo, Japan, 1979, pp. 598–600.

[24] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. of the Fourth ALVEY Vision Conference (ACV'88)*, Manchester, UK, 1988, pp. 147–151.

[25] P. R. Beaudet, "Rotationally invariant image operators," in *Proc. of the 4th International Joint Conference on Pattern Recognition (IJCPR'78)*, Kyoto, Japan, Nov. 1978, pp. 579–583.

[26] K. Kohlmann, "Corner detection in natural images based on the 2-D Hilbert transform," *Signal Processing*, vol. 48, no. 3, pp. 225–234, Feb. 1996.

[27] C.-H. Chen, J.-S. Lee, and Y.-N. Sun, "Wavelet transformation for gray-level corner detection," *Pattern Recognition*, vol. 28, no. 6, pp. 853–861, 1995.

[28] A. Quddus and M. M. Fahmy, "Corner detection using Gabor-type filtering," in *Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS'98)*, vol. 4, Monterey, CA, USA, May 1998, pp. 150–153.

[29] W. Förstner, "A framework for low level feature extraction," in *Proc. of the 3rd European Conference on Computer Vision (ECCV'94)*, Secaucus, NJ, USA, 1994, pp. 383–394.

[30] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, Nov. 1998.

[31] J. Shi and C. Tomasi, "Good features to track," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, Seattle, WA, USA, Jun. 1994, pp. 593–600.

[32] J.-B. Ryu, C.-G. Lee, and H.-H. Park, "Formula for Harris corner detector," *Electronics Letters*, vol. 47, pp. 180–181, 02 2011.

[33] J.-B. Ryu and H.-H. Park, "Log-log scaled Harris corner detector," *Electronics Letters*, vol. 46, pp. 1602–1604, 12 2010.

[34] S. Ando, "Image field categorization and edge/corner detection from gradient covariance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 2, pp. 179–190, 2000.

[35] B. Kim, J. Choi, Y. Park, and K. Sohn, "Robust corner detection based on image structure," *Circuits, Systems, and Signal Processing*, vol. 31, 08 2012.

[36] P. Mainali, Q. Yang, G. Lafruit, L. Van Gool, and R. Lauwereins, "Robust low complexity corner detector," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 4, pp. 435–445, 2011.

[37] H. Wang and M. Brady, "Real-time corner detection algorithm for motion estimation," *Image and Vision Computing*, vol. 13, no. 9, pp. 695–703, Nov. 1995.

[38] Z. Zheng, H. Wang, and E. Khwang Teoh, "Analysis of gray level corner detection," *Pattern Recognition Letters*, vol. 20, no. 2, pp. 149–162, 1999.

[39] A. Guiducci, "Corner characterization by differential geometry techniques," *Pattern Recognition Letters*, vol. 8, no. 5, pp. 311–318, 1988.

[40] K. Rangarajan, M. Shah, and D. Van Brackle, "Optimal corner detector," *Computer Vision, Graphics, and Image Processing*, vol. 48, no. 2, pp. 230–245, 1989.

[41] K. Rohr, "Recognizing corners by fitting parametric models," *International Journal of Computer Vision*, vol. 9, no. 3, pp. 213–230, 1992.

[42] R. Deriche and G. Giraudon, "A computational approach for corner and vertex detection," *International Journal of Computer Vision*, vol. 10, no. 2, pp. 101–124, 1993.

[43] P. Rosin, "Measuring corner properties," *Computer Vision and Image Understanding*, vol. 73, pp. 291–307, Feb. 1999.

[44] R. Deriche and T. Blaszka, "Recovering and characterizing image features using an efficient model based approach," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1993, pp. 530–535.

[45] A. Singh and M. Shneier, "Grey level corner detection: A generalization and a robust real time implementation," *Computer Vision, Graphics, and Image Processing*, vol. 51, no. 1, pp. 54–69, 1990.

[46] M. Ruzon and C. Tomasi, "Edge, junction, and corner detection using color distributions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1281–1295, Nov. 2001.

[47] S. M. Smith and J. M. Brady, "SUSAN–a new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45–78, May 1997.

[48] M. Trajkovic and M. Hedley, "Fast corner detection," *Image and Vision Computing*, vol. 16, pp. 75–87, Feb. 1998.

[49] S. C. Bae, I. S. Kweon, and C. D. Yoo, "COP: a new corner detector," *Pattern Recognition Letters*, vol. 23, no. 11, pp. 1349–1360, 2002.

[50] M. Cazorla and F. Escolano, "Two Bayesian methods for junction classification," *IEEE Transactions on Image Processing*, vol. 12, pp. 317–327, Mar. 2003.

[51] V. Lepetit and P. Fua, "Keypoint recognition using randomized trees," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1465–1479, Sept. 2006.

[52] J. Lan and M. Zhang, "Fast and robust corner detector based on double-circle mask," *Optical Engineering*, vol. 49, no. 12, pp. 1–8, 2010.

[53] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *10th IEEE International Conference on Computer Vision (ICCV'05)*, vol. 2, Beijing, Oct. 2005, pp. 1508–1515.

[54] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

[55] J. Wang and W. Zhang, "A survey of corner detection methods," in *2nd International Conference on Electrical Engineering and Automation (ICEEA 2018)*, vol. 139, Jan. 2018, pp. 214–219.

[56] J. A. Noble, "Finding corners," *Image and Vision Computing*, vol. 6, no. 2, pp. 121–128, 1988.

[57] F. Attneave, "Some informational aspects of visual perception," *Psychological Review*, vol. 61, no. 3, pp. 183–193, 1954.

[58] A. Rosenfeld and J. Weszka, "An improved method of angle detection on digital curves," *IEEE Transactions on Computers*, vol. C-24, no. 9, pp. 940–941, Sept. 1975.

[59] H. Freeman and L. S. Davis, "A corner-finding algorithm for chain-coded curves," *IEEE Transactions on Computers*, vol. C-26, no. 3, pp. 297–303, Mar. 1977.

[60] H. Lynn Beus and S. S. Tiu, "An improved corner detection algorithm based on chain-coded plane curves," *Pattern Recognition*, vol. 20, no. 3, pp. 291–296, 1987.

[61] J. Cooper, S. Venkatesh, and L. Kitchen, "Early jump-out corner detectors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 823–828, Aug. 1993.

[62] H.-T. Sheu and W.-C. Hu, "A rotationally invariant two-phase scheme for corner detection," *Pattern Recognition*, vol. 29, no. 5, pp. 819–828, 1996.

[63] F. Arrebola and F. Sandoval, "Corner detection and curve segmentation by multiresolution chain-code linking," *Pattern Recognition*, vol. 38, no. 10, pp. 1596–1614, 2005.

[64] F. Mokhtarian and A. K. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 1, pp. 34–43, Jan. 1986.

[65] F. Mokhtarian and R. Suomela, "Robust image corner detection through curvature scale space," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1376–1381, Dec. 1998.

[66] F. Mokhtarian, "Enhancing the curvature scale space corner detector," in *Proc. Scandinavian Conf. on Image Analysis*, Bergen, Norway, 2001, pp. 145–152.

[67] B. Zhong and W. Liao, "Direct curvature scale space: Theory and corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 508–512, Jan. 2007.

[68] X. Gao, F. Sattar, A. Quddus, and R. Venkateswarlu, "Multiscale contour corner detection based on local natural scale and wavelet transform," *Image and Vision Computing*, vol. 25, pp. 890–898, June 2007.

[69] B. K. Ray and R. Pandyan, "ACORD—an adaptive corner detector for planar curves," *Pattern Recognition*, vol. 36, no. 3, pp. 703–708, 2003.

[70] X. Zhang, M. Lei, D. Yang, Y. Wang, and L. Ma, "Multi-scale curvature product for robust image corner detection in curvature scale space," *Pattern Recognition Letters*, vol. 28, no. 5, pp. 545–554, 2007.

[71] A. Masood, "Optimized polygonal approximation by dominant point deletion," *Pattern Recognition*, vol. 41, no. 1, pp. 227–239, 2008.

[72] X. He and N. Yung, "Corner detector based on global and local curvature properties," *Optical Engineering*, vol. 47, pp. 1–12, May 2008.

[73] J. H. Han and T. Poston, "Chord-to-point distance accumulation and planar curvature: a new approach to discrete curvature," *Pattern Recognition Letters*, vol. 22, no. 10, pp. 1133–1144, 2001.

[74] M. Awrangjeb and G. Lu, "An improved curvature scale-space corner detector and a robust corner matching approach for transformed image identification," *IEEE Transactions on Image Processing*, vol. 17, no. 12, pp. 2425–2441, Dec. 2008.

[75] M. Awrangjeb and G. Lu, "Robust image corner detection based on the chord-to-point distance accumulation technique," *IEEE Transactions on Multimedia*, vol. 10, pp. 1059–1072, Oct. 2008.

[76] X. Zhang, H. Wang, A. W. Smith, X. Ling, B. C. Lovell, and D. Yang, "Corner detection based on gradient correlation matrices of planar curves," *Pattern Recognition*, vol. 43, no. 4, pp. 1207–1223, 2010.

[77] R. Elias and R. Laganiere, "JUDOCA: Junction detection operator based on circumferential anchors," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2109–2118, Apr. 2012.

[78] W.-C. Zhang, F.-P. Wang, L. Zhu, and Z.-F. Zhou, "Corner detection using Gabor filters," *IET Image Processing*, vol. 8, pp. 639–646, Nov. 2014.

[79] G.-S. Xia, J. Delon, and Y. Gousseau, "Accurate junction detection and characterization in natural images," *International Journal of Computer Vision*, vol. 106, no. 1, pp. 31–56, Jan. 2014.

[80] W.-C. Zhang and P.-L. Shui, "Contour-based corner detection via angle difference of principal directions of anisotropic Gaussian directional derivatives," *Pattern Recognition*, vol. 48, no. 9, pp. 2785–2797, 2015.

[81] S. Kim, "Robust corner detection by image-based direct curvature field estimation for mobile robot navigation," *International Journal of Advanced Robotic Systems*, vol. 9, no. 5, pp. 187:1–12, 2012.

[82] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[83] R. Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 722–732, Apr. 2010.

[84] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in *Proc. of the Eighth IEEE International Conference on Computer Vision (ICCV'01)*, vol. 1, Vancouver, Canada, July 2001, pp. 525–531.

[85] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[86] X. Gao, W. Zhang, F. Sattar, R. Venkateswarlu, and E. Sung, "Scale-space based corner detection of gray level images using Plessey operator," in *2005 5th International Conference on Information Communications Signal Processing*, 2005, pp. 683–687.

[87] M. Liu, C. Wu, and Y. Zhang, "Motion vehicle tracking based on multi-resolution optical flow and multi-scale Harris corner detection," in *2007 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2007, pp. 2032–2036.

[88] M. Liu, C. Wu, and Y. Zhang, "Multi-resolution optical flow tracking algorithm based on multi-scale Harris corner points feature," in *2008 Chinese Control and Decision Conference*, 2008, pp. 5287–5291.

[89] W. Xu, X. Huang, and W. Zhang, "A multi-scale visual salient feature points extraction method based on Gabor wavelets," in *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2009, pp. 1205–1208.

[90] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[91] H. Zhao and Y. Zhao, "Image corner location algorithm based on sub-blocking with dynamic threshold and lifting wavelet transform," in *3rd International Conference on System Science, Engineering Design and Manufacturing Informatization*, vol. 2, Oct. 2012, pp. 60–63.

[92] J. Ren, N. Chang, and W. Zhang, "A contour-based multi-scale vision corner feature recognition using Gabor filters," in *Advances in Brain Inspired Cognitive Systems*, Guangzhou, China, July 2019, pp. 433–442.

[93] X. Gao, F. Sattar, and R. Venkateswarlu, "Multiscale corner detection of gray level images based on Log-Gabor wavelet transform," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 7, pp. 868–875, Jul. 2007.

[94] W. Zhang and C. Sun, "Corner detection using multi-directional structure tensor with multiple scales," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 438–459, 2020.

[95] D. D. Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1610–1620, Jun. 2006.

[96] B. Manjunath, G. Haley, W.-Y. Ma, and S. Newsam, "Multiband techniques for texture classification and segmentation," in *Handbook of Image and Video Processing (Second Edition)*. Burlington: Academic Press, 2005, pp. 455–470.

[97] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[98] E. J. Candès and D. L. Donoho, "Ridgelets: A key to higher-dimensional intermittency?" *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, vol. 357, no. 1760, pp. 2495–2509, Sept. 1999.

[99] E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities," *Communications on Pure and Applied Mathematics*, vol. 57, no. 2, pp. 219–266, Feb. 2004.

[100] D. Labate, W. Q. Lim, G. Kutyniok, and G. Weiss, "Sparse multidimensional representation using shearlets," in *Proc. of SPIE*, vol. 5914, San Diego, California, USA, Sept. 2005, pp. 254–262.

[101] X.-M. Long and J. Zhou, "Research of image preprocessing and corner detection in 3D measurement based on Ridgelet transform," in *2009 IEEE International Conference on Granular Computing*, 2009, pp. 430–433.

[102] Y.-B. Li and J.-J. Li, "Harris corner detection algorithm based on improved contourlet transform," *Procedia Engineering*, vol. 15, pp. 2239–2243, 2011.

[103] S. R. Saydam, I. A. El rube', and A. A. Shoukry, "Contourlet based interest points detector," in *2008 20th IEEE International Conference on Tools with Artificial Intelligence*, vol. 2, 2008, pp. 509–513.

[104] J. Xiang, "A new corner detecting method based on contourlet transfrom," in *MIPPR 2009: Automatic Target Recognition and Image Analysis*, vol. 7495, Yichang, China, Oct. 2009, pp. 707–710.

[105] J. Fauqueur, N. Kingsbury, and R. Anderson, "Multiscale keypoint detection using the dual-tree complex wavelet transform," in *2006 International Conference on Image Processing*, 2006, pp. 1625–1628.

[106] C. Duan, S. Wang, Q. Huang, T. Wen, C. Zhu, and Y. Xu, "Feature level MRI fusion based on 3D dual tree compactly supported shearlet transform," *Journal of Visual Communication and Image Representation*, vol. 60, pp. 319–327, 2019.

[107] D. Malafronte, F. Odone, and E. De Vito, "Detecting spatio-temporally interest points using the shearlet transform," in *8th Iberian Conference on Pattern Recognition and Image Analysis*, Faro, Portugal, June 2017, pp. 501–510.

[108] M. A. Duval-Poo, F. Odone, and E. D. Vito, "Edges and corners with shearlets," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3768–3780, Nov. 2015.

[109] M. A. Duval-Poo, F. Odone, and E. De Vito, "Enhancing signal discontinuities with shearlets: An application to corner detection," in *18th International Conference on Image Analysis and Processing*, Genova, Italy, Sept. 2015, pp. 108–118.

[110] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua, "LIFT: Learned invariant feature transform," in *European Conference on Computer Vision (ECCV'16)*, Amsterdam, Netherlands, Oct. 2016, pp. 467–483.

[111] A. Bhowmik, S. Gumhold, C. Rother, and E. Brachmann, "Reinforced feature points: Optimizing feature detection and description for a high-level task," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 4947–4956.

[112] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superpoint: Self-supervised interest point detection and description," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR)*, Jun. 2018, pp. 337–349.

[113] M. Dusmanu, I. Rocco, T. Pajdla, M. Pollefeys, J. Sivic, A. Torii, and T. Sattler, "D2-Net: A trainable CNN for joint description and detection of local features," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 8084–8093.

[114] A. Kendall, M. Grimes, and R. Cipolla, "PoseNet: A convolutional network for real-time 6-DOF camera relocalization," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015.

[115] B. Ummenhofer, H. Zhou, J. Uhrig, N. Mayer, E. Ilg, A. Dosovitskiy, and T. Brox, "DeMoN: Depth and motion network for learning monocular stereo," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[116] V. Balntas, S. Li, and V. Prisacariu, "RelocNet: Continuous metric learning relocalisation using neural nets," in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

[117] Y. Tian, B. Fan, and F. Wu, "L2-Net: Deep learning of discriminative patch descriptor in Euclidean space," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[118] X. Han, T. Leung, Y. Jia, R. Sukthankar, and A. C. Berg, "MatchNet: Unifying feature and metric learning for patch-based matching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[119] V. Balntas, E. Riba, D. Ponsa, and K. Mikolajczyk, "Learning local feature descriptors with triplets and shallow convolutional neural networks." in *British Machine Vision Conference (BMVC)*, 2016.

[120] Y. Verdie, K. Yi, P. Fua, and V. Lepetit, "TILDE: A temporally invariant learned detector," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[121] A. Mishchuk, D. Mishkin, F. Radenovic, and J. Matas, "Working hard to know your neighbor's margins: Local descriptor learning loss," in *Proceedings of the 31th International Conference on Neural Information Processing Systems (NeurIPS)*, 2017, pp. 4829–4840.

[122] N. Savinov, A. Seki, L. Ladicky, T. Sattler, and M. Pollefeys, "Quad-Networks: Unsupervised learning to rank for interest point detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[123] C. B. Choy, J. Y. Gwak, S. Savarese, and M. Chandraker, "Universal correspondence network," in *Proceedings of the 30th International Conference on Neural Information Processing Systems (NeurIPS)*, 2016, p. 2414–2422.

[124] E. Simo-Serra, E. Trulls, L. Ferraz, I. Kokkinos, P. Fua, and F. Moreno-Noguer, "Discriminative learning of deep convolutional feature point descriptors," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015.

[125] J. L. Schönberger, M. Pollefeys, A. Geiger, and T. Sattler, "Semantic visual localization," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2018, pp. 6896–6906.

[126] N. Sebe and M. S. Lew, "Comparing salient point detectors," *Pattern Recognition Letters*, vol. 24, no. 1, pp. 89–96, 2003.

[127] D. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A*, vol. 4, no. 12, pp. 2379–2394, 1987.

[128] M. Awrangjeb, G. Lu, and C. Fraser, "Performance comparisons of contour-based corner detectors," *IEEE Transactions on Image Processing*, vol. 21, pp. 4167–4179, May 2012.

[129] L. Parida, D. Geiger, and R. Hummel, "Junctions: Detection, classification, and reconstruction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, pp. 687–698, 1998.

[130] Y. Xing, D. Zhang, J. Zhao, M. Sun, and W. Jia, "Robust fast corner detector based on filled circle and outer ring mask," *IET Image Processing*, vol. 10, no. 4, pp. 314–324, 2016.

[131] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *2011 International Conference on Computer Vision*, Barcelona, Spain, Nov. 2011, pp. 2564–2571.

[132] A. A. Karim and E. F. Nasser, "Improvement of corner detection algorithms (Harris, FAST and SUSAN) based on reduction of features space and complexity time," *Engineering and Technology Journal*, vol. 35, pp. 112–118, Dec. 2017.

[133] H. A. Kadhim and W. A. Araheemah, "A method to improve corner detectors (Harris, Shi-Tomasi & FAST) using adaptive contrast enhancement filter," *Periodicals of Engineering and Natural Sciences*, vol. 8, pp. 508–515, Mar. 2020.

[134] S. Wen, W. An, H. Li, and H. Wei, "Event-based improved FAST corner feature detection algorithm," in *2019 Chinese Control And Decision Conference (CCDC)*, 2019, pp. 4956–4961.

[135] M. A. Duval-Poo, N. Noceti, F. Odone, and E. D. Vito, "Scale invariant and noise robust interest points with shearlets," *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2853–2867, Jun. 2017.

[136] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. of the Seventh IEEE International Conference on Computer Vision (ICCV'99)*, vol. 2, Kerkyra, Greece, Sept. 1999, pp. 1150–1157.

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

[138] K. M. Yi, E. Trulls, Y. Ono, V. Lepetit, M. Salzmann, and P. Fua, "Learning to find good correspondences," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[139] M. Brown, G. Hua, and S. Winder, "Discriminative learning of local image descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 43–57, 2011.

[140] R. Owens, "Feature-free images," *Pattern Recognition Letters*, vol. 15, no. 1, pp. 35–44, 1994.

[141] P. Kovesi, "Phase congruency detects corners and edges," in *International Conference on Digital Image Computing: Techniques and Applications*, vol. 2003, Sydney, Dec. 2003.

[142] S. Häuser and G. Steidl. (2012, Feb.) Fast finite shearlet transform. [Online]. Available: https://arxiv.org/abs/1202.1773/

[143] S. Häuser and G. Steidl, "Convex multiclass segmentation with shearlet regularization," *International Journal of Computer Mathematics*, vol. 90, no. 1, pp. 62–81, Jan. 2013.

[144] G. Kutyniok, W. Q. Lim, and R. Reisenhofer, "ShearLab 3D: Faithful digital shearlet transforms based on compactly supported shearlets," *ACM Transactions on Mathematical Software*, vol. 42, no. 1, pp. 5:1–42, Feb. 2016.

[145] G. Kutyniok, M. Shahram, and X. Zhuang, "ShearLab: A rational design of a digital parabolic scaling algorithm," *SIAM Journal on Imaging Sciences*, vol. 5, no. 4, pp. 1291–1332, 2012.

[146] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, Jul. 1992.

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

[148] S. W. Teng, R. M. N. Sadat, and G. Lu, "Effective and efficient contour-based corner detectors," *Pattern Recognition*, vol. 48, no. 7, pp. 2185–2197, 2015.

[149] B. Kerautret, J.-O. Lachaud, and B. Naegel, "Curvature based corner detector for discrete, noisy and multi-scale contours," *International Journal of Shape Modeling*, vol. 14, no. 02, pp. 127–145, 2008.

[150] M. K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 300–303, 1999.

[151] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A*, vol. 4, no. 12, pp. 2379–2394, Dec. 1987.

[152] M. Wang, W. Zhang, C. Sun, and A. Sowmya, "Corner detection based on shearlet transform and multi-directional structure tensor," *Pattern Recognition*, vol. 103, pp. 107 299:1–15, 2020.

[153] S. Lam, T. C. Lim, M. Wu, B. Cao, and B. A. Jasani, "Area-time efficient FAST corner detector using data-path transposition," *IEEE Transactions on Circuits and Systems*, vol. 65, no. 9, pp. 1224–1228, 2018.

[154] Y. Zhao, H. Xing, S. Guo, Y. Wang, J. Cui, Y. Ma, Y. Liu, X. Liu, J. Feng, and Y. Li, "A novel noncontact detection method of surgeon's operation for a master-slave endovascular surgery robot," *Medical and Biological Engineering and Computing*, no. 58, pp. 871–885, Feb. 2020.

[155] Y. H. Lee, T. C. Chen, H. C. Liang, and J. X. Liao, "Algorithm and architecture design of FAST-C image corner detection engine," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 1–12, 2020.

[156] D. D. Gajski, *Principles of Digital Design*. USA: Prentice-Hall, Inc., 1996.