# Local Phase Coherence Measurement For Image Analysis and Processing

by

Rania Khairy Mohammed Hassen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2013

# Abstract

The ability of humans to perceive significant pattern and structure in an image is something which humans take for granted. We can recognize objects and patterns independent of changes in image contrast and illumination. In the past decades, it has been widely recognized in both biology and computer vision that phase contains critical information in characterizing the structures in images.

Despite the importance of local phase information and its significant success in many computer vision and image processing applications, the coherence behavior of local phases at scale-space is not well understood. This thesis concentrates on developing an invariant image representation method based on local phase information. In particular, considerable effort is devoted to study the coherence relationship between local phases at different scales in the vicinity of image features and to develop robust methods to measure the strength of this relationship. A computational framework that computes local phase coherence (LPC) intensity with arbitrary selections in the number of coefficients, scales, as well as the scale ratios between them has been developed. Particularly, we formulate local phase prediction as an optimization problem, where the objective function computes the closeness between true local phase and the predicted phase by LPC. The proposed framework not only facilitates flexible and reliable computation of LPC, but also broadens the potentials of LPC in many applications.

We demonstrate the potentials of LPC in a number of image processing applications. Firstly, we have developed a novel sharpness assessment algorithm, identified as LPC-Sharpness Index (LPC-SI) , without referencing the original image. LPC-SI is tested using four subject-rated publicly-available image databases, which demonstrates competitive performance when compared with state-of-the-art algorithms. Secondly, a new fusion quality assessment algorithm has been developed to objectively assess the performance of existing fusion algorithms. Validations over our subject-rated multi-exposure multi-focus image database show good correlations between subjective ranking score and the proposed image fusion quality index. Thirdly, the invariant properties of LPC measure have been employed to solve image registration problem where inconsistency in intensity or contrast patterns are the major challenges. LPC map has been utilized to estimate image plane transformation by maximizing weighted mutual information objective function over a range of possible transformations. Finally, the disruption of phase coherence due to blurring process is employed in a multi-focus image fusion algorithm. The algorithm utilizes two activity measures, LPC as a sharpness activity measure along with local energy as a contrast activity measure. We show that combining these two activity measures result in notable performance improvement in achieving both maximal contrast and maximal sharpness simultaneously at each spatial location.

# Acknowledgments

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed in the completion of this work.

First and foremost, to Dr. Magdy Salama for his insights and unfailing support as my main dissertation adviser. My utmost gratitude to Dr. Zhou Wang for his gentle guidance and steadfast encouragement that allowed me to have the knowledge required to contribute in our field. In each of them I have also greatly valued their enormous breadth of knowledge that spanned many disciplines.

I would like to acknowledge my Examination Committee for their extensive time, effort, and expertise that each Committee member devoted towards this work.

I must also thanks my colleagues in Image and Video Communication (IVC) research group in the Department of Electrical and Computer Engineering at the University of Waterloo for sharing valuable insights and uncountable number of useful technical discussions.

I would also like to acknowledge the administrative stuff in the Department of Electrical and Computer Engineering for the enjoyable working environment that they contribute too.

Last but not the least, my father Dr. Khairy Hassen, and my husband Dr. Ahmed Shaheen for always being there. My two lovely boys Marwan and Yaseen for their patience and tolerance while Mama busy. My mother soul for her inspiration. And the one above all of us, the omnipresent God, for answering my prayers for giving me the strength to plod on despite my constitution wanting to give up and throw in the towel, thank you so much.

# Dedication

*TO MY PARENTS WHO GAVE ME WINGS TO FLY, AND TO MY HUSBAND WHO TEACHES ME HOW TO FLY…*

# Contents

**List of Tables** viii

**List of Figures** xi

**List of Abbreviations** xii

**1 Introduction** 1

1.1 Motivation . . . 1

1.2 Problem Statement . . . 2

1.3 Contributions . . . 2

1.4 Thesis Overview . . . 3

**2 Literature Review** 5

2.1 Importance of Phase Information . . . 5

2.2 Phase in Estimating Binocular Disparity and Motion . . . 7

2.3 Phase in Image Quality Assessment . . . 9

2.4 Phase in Image Registration . . . 10

**3 Local Phase Coherence: Theory and Computation** 13

3.1 Defining Phase Coherence . . . 13

3.2 Calculating Local Phase Coherence . . . 14

3.3 Local Phase Coherence Measure . . . 17

3.4 Calculating Local Phase Coherence in Arbitrary Scales . . . 19

3.5 Efficient LPC Evaluation Approach . . . 21

3.6 Conclusions . . . 24

**4 Image Sharpness Assessment 27**
4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
4.2 Previous Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
4.3 LPC-Based Sharpness Measure . . . . . . . . . . . . . . . . . . . . 29
4.4 Validation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
4.5 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42

**5 Image Fusion 43**
5.1 Blind Objective Quality Assessment of Image Fusion . . . . . . . . 43
5.1.1 Previous Work . . . . . . . . . . . . . . . . . . . . . . . . 44
5.1.2 Quality Assessment Method . . . . . . . . . . . . . . . . . 44
5.1.3 Validation . . . . . . . . . . . . . . . . . . . . . . . . . . 48
5.2 Multi-focus Image Fusion . . . . . . . . . . . . . . . . . . . . . . 55
5.2.1 Multiresolution Fusion Scheme . . . . . . . . . . . . . . . . 56
5.2.2 Experimental Results . . . . . . . . . . . . . . . . . . . . . 59
5.3 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 62

**6 Multi-sensor Image Registration 65**
6.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 65
6.2 Proposed Registration Algorithm . . . . . . . . . . . . . . . . . . . 66
6.3 Experimental Results . . . . . . . . . . . . . . . . . . . . . . . . . 69
6.4 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 72

**7 Conclusion and Future Work 73**
7.1 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 73
7.2 Future Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 74

**References 76**

# List of Tables

3.1 Weight Solutions For Three Successive Scales . . . . . . . . . . . . . . 25

3.2 Weight Solutions For Four Successive Scales . . . . . . . . . . . . . 25

3.3 Weight Solutions for Five Successive Scales . . . . . . . . . . . . . . 26

4.1 Performance Evaluation Over Four Databases . . . . . . . . . . . . 34

4.2 Direct and Weighted Average Performance Over Four Databases . . 35

4.3 Statistical Significance Analysis Based On Quality Prediction Residuals On LIVE [1], TID2008 [2], CSIQ [3] and IVC [4] Databases . . 37

4.4 Runtime Comparison of Sharpness Measures For Images of 1024 × 1024 Resolution . . . . . . . . . . . . . . . . . . . . . . . . . . 41

5.1 Fusion Rules For Multi-resolution Fusion Algorithms . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50

5.2 Performance Evaluation Using 14 Image Sets And 6 Fusion Algorithms . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51

5.3 Comparison Between Seven Objective Fusion Quality Measures Over 14 Image Dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 54

6.1 RMS Error Comparison of MI and WMI Methods . . . . . . . . . . 69

# List of Figures

2.1 When phase information of image $A$ is combined with magnitude information of image $B$, it is phase information that prevails in the reconstructed image $R2$. It is the other way around for $R1$. . . . . . 6

3.1 A one dimensional signal and its corresponding amplitude and phase scalogram. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15

3.2 Local phase coherence structure near (a) a sharp step edge and (b) a blurred step edge, where the equal-phase contours in the scale space are shown. The phase congruency relationship corresponds to the central vertical contours at position $x_0$ in both (a) and (b), but does not describe the phase structures in the vicinity of $x_0$. (c): positions of discrete sampling across three consecutive dyadic scales in the scale space. . . . . . . . . . . . . . . . . . . . . . . . . . . . 16

3.3 Local phase coherence map using dyadic scales in the scale space plan of (a)a 1D signal (b) a phantom image and (c) a natural image. 18

3.4 Local phase coherence map using arbitrary scale ($d = 1/2$) in the scale space plan of (a)a 1D signal (b) a phantom image and (c) a natural image. . . . . . . . . . . . . . . . . . . . . . . . . . . . 24

4.1 (a) Out-of-focus blur (b) Motion-blur (c) Blur due to limited depth of field when objects are at different distances. . . . . . . . . . . . . 28

4.2 Original (a) and Gaussian blurred "Monarch" images ((b) and (c)) at two blur levels, together with their corresponding LPC maps (d)-(f). 31

4.3 Scatter plots between subjective and objective scores (after nonlinear mapping) of six blur metrics over four blur image databases. Top to bottom rows: BLIINDS-II [5, 6], BRISQUE [7, 8], $S_3$ [9], CPBD [10], JNBM [11],and the proposed LPC-SI measure; Left to right columns: LIVE, TID2008, CSIQ and IVC blur databases. . . . . . . . . . . 36

4.4 (a) Sample images and their corresponding (b) CPBD [10], (c) H-metric [12], (d) $S_3$ [9], and (e) LPC-SI sharpness maps, where brighter indicates higher sharpness. . . . . . . . . . . . . . . . . . . . . . . . 38

4.5 Relationship between LPC-SI and Gaussian blur standard deviation. 39

4.6 Example of original and noise corrupted images. . . . . . . . . . . . 40

4.7 Relationship between LPC-SI and noise standard deviation $\sigma$. . . . 40

4.8 Computational Runtime Comparison between LPC-SI and Existing NR-Sharpness Measures with Image Size. . . . . . . . . . . . . . . 42

5.1 Source images (top) and subject-rank fusion results (bottom). . . . 45

5.2 Sample images with multi-exposure and multi-focus effects. . . . . . 49

5.3 Dendrogram of fusion measures. . . . . . . . . . . . . . . . . . . . . . 55

5.4 Synthetic images with different sharpness and contrast in corresponding objects. . . . . . . . . . . . . . . . . . . . . . . . . . . 56

5.5 Diagram of the proposed image fusion algorithm. . . . . . . . . . . 57

5.6 (a) Local phase coherence map and (b) Local energy map of Fig. 5.4(a). 58

5.7 Fusion results for the images in Fig. 5.4 using (a) local energy activity measure only and (b) the proposed fusion method. . . . . . . . . . . 59

5.8 (a),(b): Synthetic noisy images; (c): Fused image using the proposed method. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

5.9 LPC Measurment with Different Noise Levels. . . . . . . . . . . . . 61

5.10 (a),(b): Two images of radioleria acquired by optical microscope; (c): Fused Image using the proposed algorithm. . . . . . . . . . . . 61

5.11 Fusion of natural images using the proposed method: (a) Original image, (b) blurred image 1, (c) blurred image 2, and (d) the fused image. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63

5.12 Fusion of natural images using the proposed method: (a) low contrast image, (b) blurred image, (c) fused image. . . . . . . . . . . . 64

6.1 Overall architecture of the proposed registration algorithm. . . . . . 66

6.2 (a) Reference image,(b) Target image translated and rotated around the image center, (c) Registered image using WMI, (d) The phase coherence map of both images and the corresponding joint histograms before and after registration using both MI and WMI, along with the proposed weighting function. . . . . . . . . . . . . . . . . . . . . 68

6.3 (a) Noisy reference image; (b) Noisy target image; (c) Registered image; (d),(e) Comparison of RMS estimation error in the $x$-direction and $y$-direction against different noise levels. . . . . . . . . . . . . 70

6.4 Registration results using the proposed method in the case of simulated missing data in brain images (a) Original reference image; (b) reference image with missing data; (c) target image; (d) Registered image. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 71

6.5 Registration results using the proposed method in the case of real missing data in brain images (a) Original reference image; (b) target image; (d) Registered image. . . . . . . . . . . . . . . . . . . . . . . . . 72

# List of Abbreviations

| | |
|---|---|
| AVG | Average |
| BLIINDS-II | BLind Image Integrity Notator using DCT Statistics-II |
| BRISQUE | Blind/Referenceless Image Spatial QUality Evaluator |
| CM | Choose Maximum |
| CMCV | Choose Maximum with Consistency Verification |
| COP | COntrast Pyramid |
| CPBD | Cumulative Probability Blur Detection |
| CSF | Contrast Sensitivity Function |
| CSIQ | Categorical Image Quality database |
| CW-SSIM | Complex Wavelet-Structural SIMilarity |
| DCT | Discrete Cosine Transform |
| DMOS | Difference of MOS |
| DOF | Depth-Of-Field |
| DWT | Discrete Wavelet Transform |
| FFT | Fast Fourier Transform |
| FQI | Fusion Quality Index |
| FR | Full Reference |
| GRW | Generalized Random Walk |
| HVS | Human Visual System |
| IQA | Image Quality Assessment |
| JNB | Just Noticeable Blur |

| | |
|---|---|
| KRCC | Kendall Rank-order Correlation Coefficient |
| LAP | LAplacian Pyramid |
| LPC | Local Phase Coherence |
| LPC-SI | Local Phase Coherence-Sharpness Index |
| MI | Mutual Information |
| MOS | Mean Opinion Score |
| MSE | Mean Squared Error |
| NR | No Reference |
| PLCC | Pearson Linear Correlation Coefficient |
| PSNR | Percentage Signal to Noise Ratio |
| RAP | RAtio Pyramid |
| RMSE | Root Mean Squared Error |
| SRCC | Spearman Rank-order Correlation Coefficient |
| SSIM | Structural SIMilarity |
| TID2008 | Tampere Image Database 2008 |
| TV | Total Variation |
| UPGMA | Unweighted Pair Group Method with Arithmetic Mean |
| VQEG | Video Quality Expert Group |
| WA | Weighted Average |
| WMI | Weighted Mutual Information |

# Chapter 1

# Introduction

## 1.1 Motivation

The human visual system (HVS) is able to reliably identify salient image features under widely varying conditions. Our interpretation of image structure is not greatly affected by changes in image contrast/illumination or by blur distortion. Over the past three decades, researchers from many communities have converged to a view that phase information brings more insights about image structure [13]. Some effort has been devoted to investigate the phase structure in natural images for purposes of improving the performance of image processing [14] and computer vision systems [15, 16, 17], and also for furthering our understanding of biological visual systems [18, 19, 20]. More recently, a quantitative measure for the alignment of the local phases of all Fourier harmonics, known as *Phase Congruency*, has been developed [21]. Phase congruency reflects the behavior of an image in the frequency domain and provides localization of significant features based on phase information. However, little attention has been paid to the relationship between local phases at different scales at the vicinity of image features until the work done by Wang and Simoncelli [22]. In this work they showed that when an image is blurred through convolution with a symmetric linear filter, the phase information in the *global* Fourier transform domain does not change at all, but blur will disrupt the *local* phases pattern. This work elevated what is known as the local phase coherence theory.

As many research work agreed that local phase is crucial in the perception of images, interest in developing novel image representation measures based on local phases has been growing. Inspired in part by the biological importance of local phase information and motivated by the work in [22], this thesis concerned with the study of the coherence relationship between local phases of wavelet coefficients in the scale-space domain. In particular, we present a measure of local phase coherence strength to describe sharp structural information of an image. This measure can automatically detect blur image features and quntitaivly estimate the amount of blur distortion. In the case of a distortion-free image, the proposed measure creates a feature map that could be potentially used as a low-level image representation which is invariant to image illumination, contrast changes, and to some extent robust to noise. Such a map would provide an absolute feature detection that could be universally applied to any image without any prior knowledge about image structure or characteristics.

## 1.2 Problem Statement

The objectives of this thesis are to develop the concept of local phase coherence so as to provide reliable feature detectors that can quantitatively measure the strength of local phase coherence regardless of image illumination and contrast, and to extend the novel phase coherence measurements to real-world image processing applications.

## 1.3 Contributions

The main contributions of this thesis are summarized as follow:

- The basis idea behind the computation of phase coherence has been extended to any number of scales with fractional scale ratios. The scale ratio does not have to be the same between successive scales. This is a useful feature in practical applications because real world signals often contain mixtures of many distinctive features, and thus local measurement at coarse scales often suffers from interference from nearby features. If the scale ratios can be fractional (preferably less than 2), then we will be able to carry out closer scale-space analysis of local features and avoid interference from nearby features. This leads to the development of a new measure of phase coherence strength based on the extended fractional scale computation. This new measure provides a very localized feature map for a given image allowing features to be detected and quantified based on coherence intensity at each location close to the feature.

- An efficient LPC evaluation method is proposed that provides a computationally low cost algorithm for the assessment of phase coherence at every location in an image. A weighted-phase prediction model is developed for $N$ scales evenly spaced in either linear or logarithmic scale.

- A new interpretation of image sharpness has been introduced. A no-reference sharpness estimator is developed upon the notion of LPC that can evaluate the perceived image sharpness without a reference image. The proposed algorithm does not require traditional edge width, local gradient, or energy fall-off computations, which are often limited by the complication of the large variations between real-world images in terms of information content and complexity.

- A new approach to image fusion quality assessment has been proposed, where LPC sharpness measure of fused image and LPC feature maps of source images are incorporated to build a robust fusion metric that well predicts human visual preference of fusion results.
- LPC measure has been utilized as sharpness activity measure in a simple yet efficient multi-focus fusion approach. This approach is different from existing methods, which share a common assumption that high local energy or contrast is a direct indication of local sharpness. Our algorithm shows that combining two activity measures, LPC for sharpness and local energy for contrast, results in notable performance improvement in achieving both maximal contrast and maximal sharpness simultaneously at each spatial location.
- The invariance characteristics of LPC measure to image illumination and contrast has been employed in a multi-sensor image registration approach. LPC feature map has been used as an intermediate image representation to eliminate contrast and/or illumination variations between multi-sensor images. The feature maps are utilized to estimate image plane transformation by maximizing an objective function based on weighted mutual information over a range of possible transformations.

## 1.4 Thesis Overview

Chapter 2 reviews the major machine vision and image processing approaches that have used phase information as a mechanism for localizing/representing features in natural images. Firstly the importance of phase information in representing signals and images is defined. Then a number of vision applications based on local phases, such as motion estimation, disparity measure, image registration and quality assessment, is presented and discussed.

Chapter 3 sets out to thoroughly present the theoretical aspects and computational approach of local phase coherence. It starts by defining the meaning of phase coherence in 1D and 2D signals and describing the prior computation of phase prediction. Subsequently, the calculation of the proposed LPC measure is presented. Finally a novel flexible framework which allows the computation of LPC in arbitrary fractional scales is introduced.

Chapter 4 presents an approach to evaluate the sharpness of a given image without having reference image. Existing blur/sharpness evaluation algorithms are mostly based on edge width, local gradient, or energy reduction of global/local high frequency content. In this chapter, we understand sharpness prediction from a different perspective, where sharpness is identified as strong LPC near distinctive image features evaluated in the complex wavelet transform domain. A new sharpness assessment algorithm without referencing the original image is presented and tested using four subject-rated publicly-available image databases, which demonstrates competitive performance when compared with state-of-the-art algorithms.

Chapter 5 describes an approach to assess the performance of image fusion algorithms. A blind objective fusion quality metric is proposed and validated over our multi-exposure multi-focus subject-rated image database. The proposed fusion metric shows very good correlation with human scores and outperforms seven existing fusion metrics. Further, a novel fusion algorithm is built to integrate multi-focus images in a multi-resolution framework. The algorithm employs LPC measurement as sharpness activity measure together beside local energy as contrast activity measure. We show that combining these two activity measures results in notable performance improvement in achieving both maximal contrast and maximal sharpness simultaneously at each spatial location.

Chapter 6 uses LPC as an underlying representations to common features in multi-sensor images. This application utilizes changes in intensity/contrast invariant property in LPC measure to align images acquired using different imaging devices. Further, an objective function based on weighted mutual information is presented,tested on both synthetic and medical images and evaluated based on registration accuracy.

Finally, Chapter 7 concludes this work and discusses the areas that could be developed further in future work.

# Chapter 2

# Literature Review

## 2.1 Importance of Phase Information

In the Fourier representations of signals and images, phases have been long realized to be more important than magnitudes in the reconstruction of signals structure. In the signal processing literature, Oppenheim, Lim, et al. are in the first group of people who emphasized this phenomenon [13, 23]. They used numerical experiments to illustrate the similarity between a signal and its phase-preserved reconstruction. Given two images $A$ and $B$, one can compute their 2D Fourier transforms $F(A)$ and $F(B)$. Then two image reconstructions are conducted. The first reconstructed image denoted by $R1$ is the inverse Fourier transform of the combination of the magnitudes of $F(A)$ and the phases of $F(B)$. The other reconstructed image $R2$ is the inverse Fourier transform of the combination of the magnitudes of $F(B)$ and the phases of $F(A)$. It was found that in nearly all cases, reconstruction $R1$ looks similar to image $B$, while reconstruction $R2$ looks similar to image $A$. In other words, the structures seen in a reconstructed image clearly corresponds to those in the image from which phases are preserved. A natural interpretation of this phenomenon is that in image/signal reconstruction, important features of a image/signal are contained in phases and that phase information is more important than magnitude information. A demonstration of this phenomena is repeated here in Fig. 2.1. Since then the importance of phase in image/signal representation has received considerable attention.

A widely accepted concept in computational neuroscience is that the biological visual systems attempt to encode a visual image in an economical form. This encouraged researchers to believe that there may exist in the human visual system detectors specialized to respond to distinctive features such as edges and lines. Tolhurst [24] and Kulikowski et al. [25] suggested that visual detectors may take advantage of the local symmetry of these features. Psychophysical experiments reveal that the visual system is sensitive to two-dimensional spatial phase [26, 27] and that localized phase may play an important role in vision. Moreover, findings in physiology indicate that the visual system encodes information by means of a

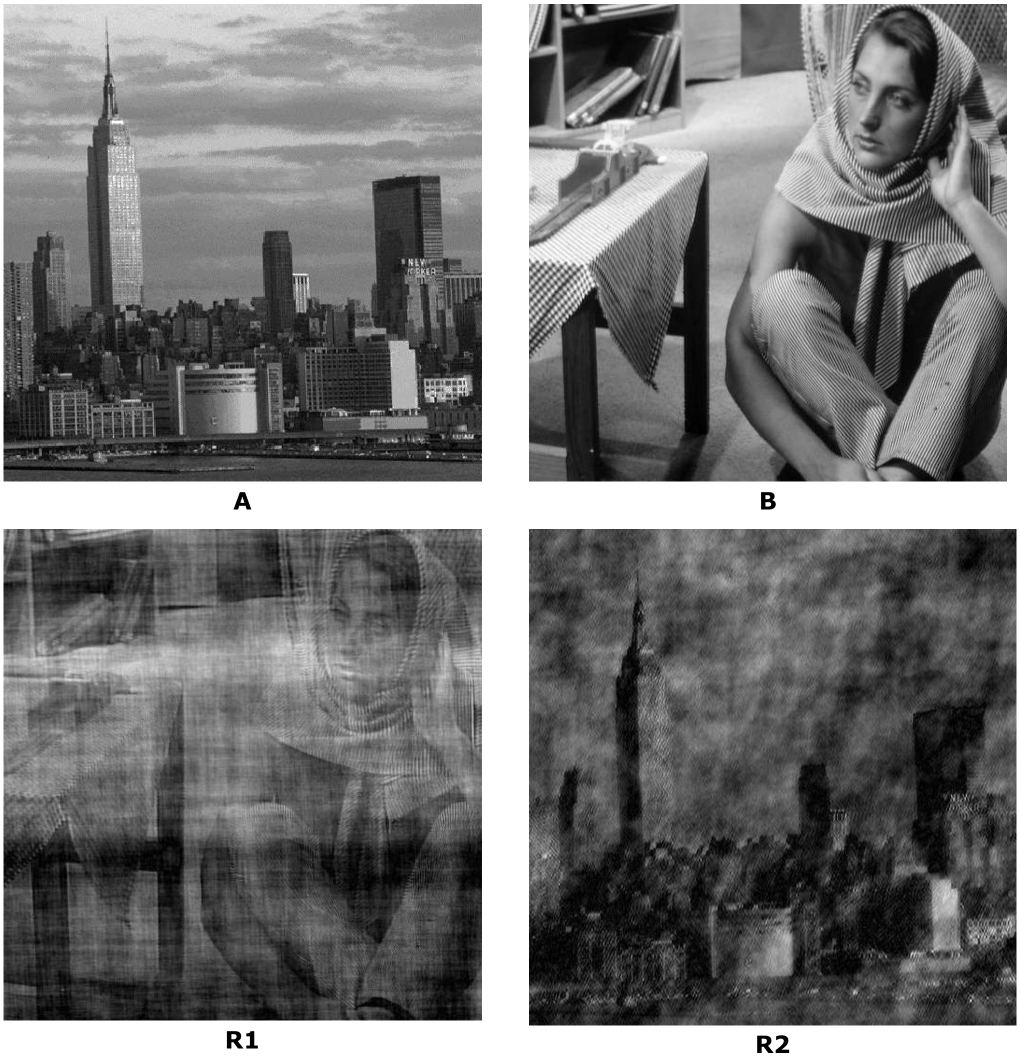


Figure 2.1: When phase information of image $A$ is combined with magnitude information of image $B$, it is phase information that prevails in the reconstructed image $R2$. It is the other way around for $R1$.

scheme which, to a reasonable approximation, is similar to representing a signal using bases of multi-scale bandpass oriented filters [28, 29]. Additional evidences suggest that there may be visual detectors that have both even- and odd-symmetric receptive fields [30].

In Fourier space, images of natural scenes diverge from each other primarily in terms of their phase and not their amplitude spectra. Phase is particularly important for edge and line features, since these features require an alignment of the phase of different spatial frequency components. Several models of edge and line detection have been developed. One of the significant works is done by Morrone and Burr [16]. In their work they proposed a feature detection model that operates on the bases of phase information, known in the literature as "*Local Energy Model*". The model makes use of a common property of lines and edges by considering the local Fourier representation of them. This property is the symmetry of lines and edges in the phase spectrum of Fourier space. A line is an even-symmetric function while an edge is an odd-symmetric function. Thus when the edge or line is chosen as the origin, the Fourier phase spectrum is constant: zero at all frequencies for a line and $\pi/2$ for an edge. For any origin, at the point of isolated line or edge, the arguments (also called arrival phases) of all Fourier harmonics will be identical. Conversely, the value of the average arrival phase at any point determines the nature of the feature. That is to say, values near zero correspond to a line, and values near $\pi/2$ correspond to an edge. That definition of lines and edges had been successfully used to predict "Mach Bands", the illusory light and dark stripes seen where luminance gradients meet plateaux [15]. Based on this definition, Kovesi [21] developed a method to measure the significance of the "congruency of arrival phase" via a dimensionless quantity in wavelet domain.

From a statistical point of view, manipulation in phase spectrum do not affect first- and second-order statistics but affect higher-order image structures such as edges and corners [31]. In a series of elegant articles, Thomson and colleagues explored the properties of phase spectra of natural images within a statistical framework [32, 33, 34, 35]. They extend the commonly used first- and second-order statistics analysis by computing higher-order image statistics attempting to find whether regularities in the phase spectra of images are reflected in their higher-order statistics. In [33] Thomson showed that whitened natural scenes have a strictly positive kurtosis, whereas phase randomized versions of the same images have positive and negative kurtosis values very close to zero. Recently, Wichmann et al. [36] investigated the effect of phase manipulations on the processing of natural images by adding random phase noise, and manipulating image's contrast. They demonstrated that phase noise is more disruptive than simple contrast reduction because it changes visual features such as local edges, thereby degrading object boundaries.

The use of phase information in computer vision and image processing applications is inspired by experiments and results described above. In the following sections we will briefly review various computer vision and image processing applications utilizing phase information.

## 2.2 Phase in Estimating Binocular Disparity and Motion

The localization of objects in the real world from a stereo presentation has been one of the major goals of computer vision researchers. We begin by examining binocular disparity techniques. The word binocular comes from two Latin roots which means "two eyes". Binocular viewing of a scene creates two slightly different images of the scene in the two eyes due to the eyes' different positions (parallax) on the head. These differences, referred to as binocular disparity, provide information that the brain can use to calculate depth in the visual scene, providing a major means of depth perception. In the last two decades, the problem of computing binocular disparity has been formulated in terms of phase matching, that is, determining the shift required so that the phases of the left and right signals become equal [37, 38, 39]

Phase correlation based on Fourier shift property was used to compute binocular disparity in [40]. Fourier shift theory states that a shift in the coordinate frames of two functions is transformed in the Fourier domain as linear phase differences. The assumption is made that the left and right images are simply shifted versions of each other, and that the measurement of this shift is equivalent to determining the stereo parallax. Unfortunately, this assumption is not precisely valid. Horn [41] and Jenkin et al. [42] have shown that interocular differences in lighting and the effects of the perspective imaging of binocular systems pose difficult problems to correlation algorithms that fundamentally assume that the left and right eyes see simply shifted versions of the same structure. In a recognized state-of-the-art disparity estimation work, Jenkin and Jepson [42] presented a new technique for measuring disparity as the local phase difference between bandpass versions of the two images. They presented a new definition of disparity that is tied to the interocular phase difference in bandpass versions of the monocular images. Jenkin and Jepson's disparity measurement technique [42] bears similarity to other techniques ([37, 39, 40]). All of them expressed the task of measuring binocular disparity as the problem of computing the phase difference between two signals. However, Jenkin and Jepson [42] computed the phase difference locally without using Fourier-based schemes. This technique results in disparity detectors capable of producing dense, robust responses and can be tuned to deal with interocular differences such as orientation, contrast, and nonzero disparity gradients.

Estimating object's motion in a visual scene is often referred to as 2D image velocity or *optical flow*. The goal of optical flow estimation is to compute an approximation to the motion field from time-varying image intensity. Fleet and Jepson in [43] defined component image velocity in terms of the gradient of phase output from velocity-tuned linear filters. A sequence of image varying in time is first represented with a family of constant-phase spatiotemporal velocity-tuned linear filters, each of which is tuned to a narrow range of orientation, speed, scale, and has a local spatiotemporal support. The local phase gradient is then computed from the output of each filter to obtain estimates of image velocity. Comparing phase-based motion estimation approaches to amplitude-based ones, it was found that phase information yields accurate and robust estimation of image velocity. In particular, local phase information was shown to be more robust than amplitude under variations in lighting conditions, surface orientation, as well as changes in local orientation, wavelength, and speed due to geometric deformation in space-time. In a review evaluation of motion estimation techniques, Barron et al. [44] compared nine different motion estimation techniques of various approaches, namely differential, matching, energy-based, and phase-based ones. They have tested these algorithms on several standard image sequences synthetic and realistic. For synthetic image sequences, the ground truth motion fields are known, and for realistic sequences, the desired motion fields are not known and only qualitative comparisons can be made. They found that, the phase-based technique of Fleet and Jepson [43] and differential technique of Lucas and Kanade [45] produced the most accurate results overall.

Since phase information has shown many desirable properties, the robustness of local phase information for measuring image velocity and binocular disparity has been a concern. A thorough theoretical analysis of phase stability has been done in [17, 46]. In these works, Fleet et al. studied the stability of phase with respect to geometric deformations, and its linearity as a function of spatial position. They showed that the phase stability and linearity depend on the form of the filters and their bandwidths. For example, for any filter type, as the bandwidth increases, the extent of the phase stability increases, while the spatial extent over which phase is expected to be linear decreases. They further showed that phase information can become unstable in the vicinity of phase singularities. Phase singularities were described as the neighborhoods of which the phase behavior is extremely sensitive to input scale perturbations, small changes in spatial domain, and small amounts of noise. These phase singularities are a source of significant errors for phase-difference and phase-gradient techniques, as well as gradient-based techniques, zero-crossing techniques, and phase-correlation techniques. The detection of unstable phase regions is essential to the reliable performance of phase-matching techniques and it has been used to improve performance, measurement accuracy, and robustness of binocular disparity works done by Jenkin and Jepson [42], and Sanger [37].

## 2.3 Phase in Image Quality Assessment

Image quality is a characteristic of an image that measures the perceived image degradation. Images can be degraded in numerous ways during acquisition, transmission, storage and reconstruction. Identifying and quantifying the type of degradation factor and its impact on the image is very useful in many applications such as improving the acquisition system for producing better image quality. Subjective image quality measure has been considered in many applications because the quality of the image is dependent on the image's consumers. In subjective image quality measure, human viewers are asked to provide the ultimate judgment of the image quality. The major disadvantages of this type of quality measurements are slowness, and dependence of viewing conditions and vision capabilities. Therefore, objective image quality measures that can avoid these drawbacks are desired for the sake of reliable computer vision applications without human involvement.

In recent years, there has been an increasing need to develop objective image quality assessment measure that can provide an automatic quantitative indication of the quality of an image. Such measures find numerous image processing and pattern recognition applications. The most widely used objective image quality metric is Mean Squared Error (MSE) . MSE is simple to compute, has a clear physical meaning and is an excellent metric in the context of optimization problems [47]. However, in [48] the authors have shown that MSE performs poorly in image quality assessment and pattern recognition tasks. In that work, a new spatial domain image quality assessment method called Structural SIMilarity (SSIM) index was proposed. The main idea behind SSIM is that the human visual system is highly adapted to extract image structure from visual scenes. Therefore image quality measure should be able to retain structural information. SSIM performs remarkably well across a wide variety of image quality measures. The main drawback of SSIM is that it is highly sensitive to non-structural distortion such as translation, rotation and scaling of images. Non-structured image distortions are typically caused by the movement of the acquisition devices, rather than the changes of image structures. To handle such situations, a modified version of SSIM called Complex Wavelet-SSIM (CW-SSIM) index was developed [49]. The contribution of CW-SSIM index is by considering a new component based on the idea that the structural information of local image features is mainly contained in the relative phase patterns of the wavelet coefficients. Moreover, a constant phase shift of all wavelet coefficients will not change the structure of local image feature. Recently, it has been shown that CW-SSIM generalizes well to a wide variety of image similarity tasks [50].

## 2.4 Phase in Image Registration

Image registration is one of the most common widely encountered image analysis problems. Image registration refers to the process of geometrically align images acquired of the same scene using different sensors. The main challenging problem in multimodal registration is that images are acquired using different sensors, have different intensity mappings for the same content, making them difficult to compare directly. One possible solution to this problem is to build image representations that allow for direct comparison independent of the underlying image intensity and based only on image structural characteristics. Utilizing phase information of multimodal images would be a perfect match to solve such problems. Liu et al. [51] developed a multimodal registration technique based on the gradient of phase information. First local frequency representation of the image data is obtained by filtering the image with a Gabor filter tuned to a certain frequency/orientation and then the gradient of the phase of the filtered image is computed. To align the local frequency representation of two images, namely target and source images, they developed a statistically matching criterion based on the minimization of the integral squared error or $L_2$ error between Gaussian model of the residual (refers to the difference between the local frequency representations of the transformed source and target images) and the true density function of the residual. They tested their proposed algorithm on 2D MR T1- and T2-weighted scans and on 3D CT and MR data. A related approach was proposed by Wong et al. in [52], where they used the phase-congruency model proposed by Kovesi [21] in the detection of control points for the purpose of rigid registration between satellite and aerial remotely sensed images. All the above mentioned algorithms remove different intensity values between images to be registered, by extracting common structural features based on local phase information, but they do not use local phase information directly. Mellor et al. [53] utilized local phase directly as the representation basis for non-rigid registration of multimodal medical images. Local phase has been estimated using monogenic signal [54], which is a multi-dimensional extension to analytic signal. In this method, the displacement field is computed iteratively. At each iteration, the joint distribution between the source image and the target image phases is estimated. Then at each pixel, mutual information forces are computed assuming a partial volume interpolation scheme [55]. Finally the displacement field is estimated by minimizing the sum of all mutual information force vectors computed at every scale. The deformation field is then smoothed and applied to the source image at the end of the current iteration and the process is repeated.

Phase correlation [40] is another common and popular technique used for aligning two images based on phase information. The phase correlation method is based on the well-known Fourier shift property. Specifically, a shift in the coordinate frame of two functions results in a linear phase difference in the Fourier transform of the two functions. Given a pair of two-dimensional functions, $f(x, y)$ and $g(x, y)$, which differ only by displacement $(a, b)$, i.e.,

$$g(x, y) = f(x - a, y - b) \tag{2.1}$$

their corresponding elements of the Fourier transform denoted by $G$ and $F$ are related by:

$$G(w_x, w_y) = F(w_x, w_y)\, e^{-j(w_x\, a + w_y\, b)} \tag{2.2}$$

where $(w_x, w_y)$ are the Fourier domain coordinates. In other words, the two images have the same Fourier magnitude but a phase difference directly related to their displacement. In order to compute the translation parameters $a$ and $b$, the normalized cross-power spectrum between $F$ and $G$ is used to obtain the phase correlation matrix as follows:

$$Q(w_x, w_y) = \frac{G(w_x, w_y)F(w_x, w_y)^*}{|F(w_x, w_y)F(w_x, w_y)^*|} = e^{-j(w_x\, a + w_y\, b)} \tag{2.3}$$

where $F^*$ is the complex conjugate of $F$, the Shift theorem guarantees that the phase of the cross-power spectrum is equivalent to the phase difference between the images. Furthermore, taking the inverse Fourier transform of the cross-power spectrum $Q(w_x, w_y)$, results in an impulse function in the spatial domain. The resulting image is approximately zero everywhere except at the location of displacement, which is needed to optimally register two images. This can be described as follow:

$$q(x, y) = \delta(x - a, y - b). \tag{2.4}$$

The basic idea behind phase correlation registration is to determine the location of the peak of the inverse Fourier transform of the cross-power spectrum. This property of using the phase information for correlation is sometimes referred to as a *whitening* of an image. The main advantages of phase correlation technique is its robustness to noise and computational efficiency because fast Fourier transform is much faster than computing the cross-correlation between large images. Moreover, it is invariant to changes in brightness, which makes the phase correlation measure relatively scene independent. On the other hand, the drawbacks of phase correlation appears in case of significant amount of noise spread across all frequencies, in such a case the location of the peak will be inaccurate because phase difference at each frequency is corrupted. However, other transformation parameters such as rotation and scale cannot be computed in a straight forward method. An extension of the phase correlation technique proposed by De Castro and Morandi in [56], where they computed the phase of the cross-power spectrum as a function of the rotation angle $\phi$ and used the polar coordinate to simplify the cross-power spectrum computation equation as follow:

$$G(r,\theta;\phi) = \frac{F_1(r,\theta)F_2^*(r,\theta-\phi)}{|F_1(r,\theta)F_2^*(r,\theta-\phi)|} \tag{2.5}$$

The main contribution in this work is that they separate the computation of rotation and translation by first determining the angle $\phi$ which makes the inverse Fourier transform of the cross-power spectrum the closest approximation to an impulse, and then compute translation factor in term of the impulse location. Even though their method succeeded in computing correct rotation and translation values, it is computationally costly because of the difficulty in testing for each $\phi$. Another intersting work related to phase correlation in estimating registration parameters was done by Reddy and Chatterji in [57]. They extended phase correlation technique to cover translation, rotation and scaling. In their work, Fourier scaling properties and Fourier rotational properties are used to find scale and rotational movement. They demonstrated that scaling and rotational movement can be reduced to translational displacement using logarithmic scale and polar coordinates respectively. The registration is performed in the Fourier log-polar plane where rotation angle and scaling factor are estimated first. Then the image is transformed using the estimated rotation and scale parameters to finally compute translation using the ordinary search for peak location in the cross-power spectrum. The main advantage of this method is that it computes the motion parameters in fixed time for fixed size of image irrespective of the amount of translation, rotation and scale. A major drawback in phase correlation technique is the difficulty of handling non-integer displacement between images to be registered. Foroosh et al. [58] proposed a phase correlation approach to compute subpixel accuracy registration by means of analytic expression of phase correlation on down sampled images. The main drawback of their method is its limitation to translations only.

# Chapter 3

# Local Phase Coherence: Theory and Computation

## 3.1 Defining Phase Coherence

In order to define phase coherence, it is important to obtain spatially localized frequency information of a signal/images. Wavelets provides perhaps the best way of doing this because in part it may be viewed as biological resemblance to simple cortical cells in HVS, and in the other part it can provide local frequency information in terms of amplitude and phase values that can help in image local analysis. The use of Wavelet Transform for frequency analysis of a signal was developed by Morlet et al. [59]. The basic idea behind Wavelet analysis is that one uses a bank of filters to analyze a signal. The filters are all created from scaling and shifting of one wave shape. Each filter is designed to pick up one band of frequency in the signal being analyzed. We are interested in calculating local phase information of a signal. So in order to preserve phase information, linear-phase filters are desired. This means symmetric/anti-symmetric wavelets are more appropriate in this case than orthogonal wavelets.

Analysis of a signal is done by convolving the given signal with a wavelet filter bank. As described in [22], let $f(x)$ denote a signal localized near the position $x_0$, where $f(x) = f_0(x - x_0)$, a general expression of the wavelet transform can be written as:

$$\begin{aligned} F(s,p) &= \int_{-\infty}^{\infty} f(x) w_{s,p}^{*}(x) dx \\ &= \left[ f(x) * \frac{1}{\sqrt{s}} g\left(\frac{x}{s}\right) e^{\jmath \omega_c x/s} \right]_{x=p} , \end{aligned} \tag{3.1}$$

where $s \in R^{+}$ is the scale factor, $p \in R$ is the translation factor, and the family of wavelets are derived from a scaled and shifted mother wavelet $w(x) = g(x)e^{j\omega_c x}$ by:

$$\begin{aligned} w_{s,p}(x) &= \frac{1}{\sqrt{s}}\, w\left(\frac{x-p}{s}\right) \\ &= \frac{1}{\sqrt{s}}\, g\left(\frac{x-p}{s}\right) e^{j\omega_c (x-p)/s}\,, \end{aligned} \tag{3.2}$$

where $\omega_c$ is the center frequency of the modulated band-pass filter, and $g(x)$ is a slowly varying, non-negative and symmetric envelop function. For example, when $g(x)$ has the shape of a Gaussian, $w_{s,p}(x)$ constitutes a family of Gabor wavelets - sine and cosine each modulated by a Gaussian wave. However, the general derivation below also applies to other options of $g(x)$ as well. Using two filters in quadrature, we can calculate the amplitude and phase of the responses of each quadrature pair filters for a particular frequency at given spatial location as

$$\begin{aligned} A_{s,p}(x) &= \sqrt{\left(f(x) * w^e_{s,p}(x)\right)^2 + \left(f(x) * w^o_{s,p}(x)\right)^2} \\ \Phi_{s,p}(x) &= \tan^{-1} \frac{f(x) * w^o_{s,p}(x)}{f(x) * w^e_{s,p}(x)}\,, \end{aligned} \tag{3.3}$$

where $w^e_{s,p}(x)$ and $w^o_{s,p}(x)$ can be thought as the even and odd wavelets at a scale $s$ and a translation $p$. The results of convolution with wavelet filter bank can be graphically displayed with a scalogram in Fig. 3.1. Each row in the scalogram represents the result of convolving the signal with quadrature pair wavelets at a certain scale. The vertical axis of the scalogram is the frequency scale, with the lowest frequency at the top. Each column of the scalogram can be considered as the local Fourier spectrum for each point. Note that the phase values are mapped from $0 - 2\pi$ to $0 - 1$ and displayed as gray levels. The black/white discontinuities in the phase scalogram represent the wrap-around of phase values. The phase scalogram is of particular interest. The level sets local phases at different scales form straight lines which all converge to the location of the signal feature (in this case step feature). This relationship between local phases in scale-space domain is referred to as Local Phase Coherence (LPC) [22]. Phase congruency [21] is considered as a subset of LPC as we will mathematically describe in the following section. To the best of our knowledge, local phase has been used in great body of research work in image processing and computer vision literature. However studying the relationship between local phases at different scale nearby image features and quantifying the strength of coherence have not been thoroughly investigated.

## 3.2 Calculating Local Phase Coherence

Mathematically, we can describe local phase relationship of a particular scale at any given location as the following. Back to Eq. 3.1 and using the convolution theorem, and the shifting and scaling properties of the Fourier transform, we can

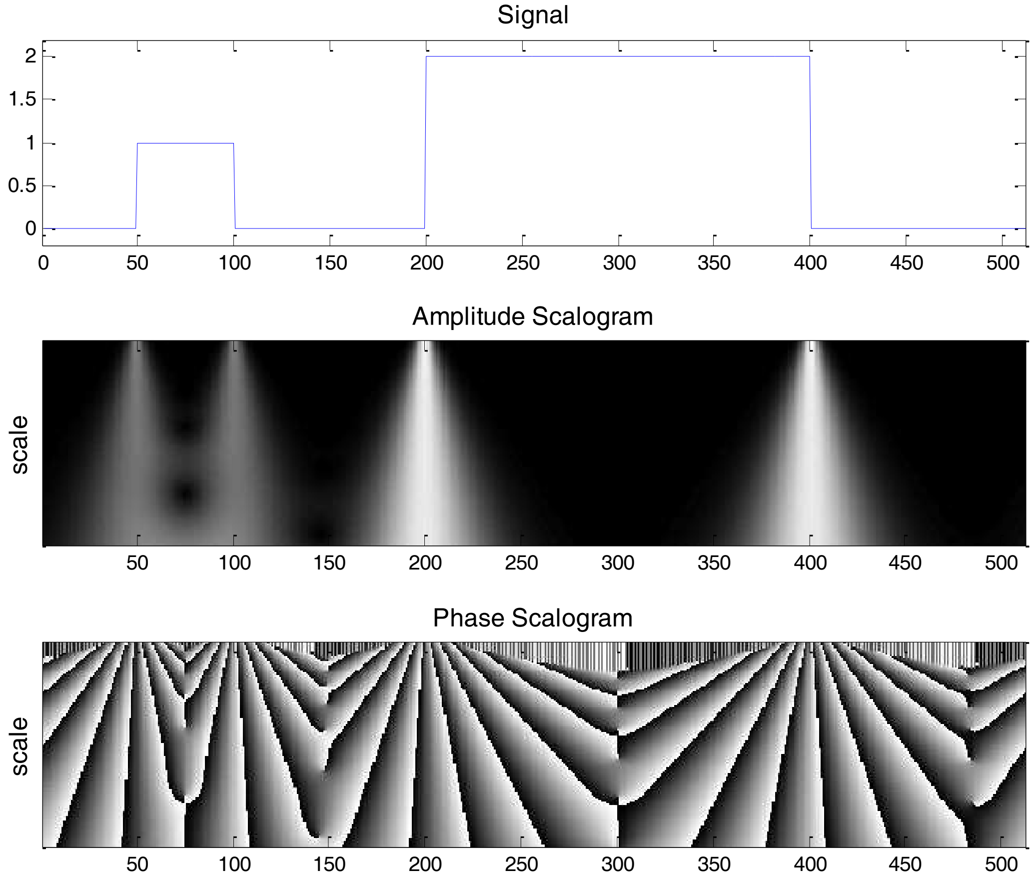


Figure 3.1: A one dimensional signal and its corresponding amplitude and phase scalogram.

derive [22]:

$$\begin{aligned} F(s,p) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\sqrt{s}G(s\,\omega-\omega_c)e^{j\omega p}d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} F_0(\omega)\sqrt{s}G(s\,\omega-\omega_c)e^{j\omega(p-x_0)}d\omega \\ &= \frac{1}{2\pi\sqrt{s}}\int_{-\infty}^{\infty} F_0\left(\frac{\omega}{s}\right)G(\omega-\omega_c)\,e^{j\omega(p-x_0)/s}d\omega\,, \end{aligned} \tag{3.4}$$

where $F(\omega)$, $F_0(\omega)$ and $G(\omega)$ are the Fourier transforms of $f(x)$, $f_0(x)$ and $g(x)$, respectively. The phase of $F(s,p)$ depends on the nature of $F_0(\omega)$. If $F_0(\omega)$ is scale invariant, meaning that:

$$F_0(\omega/s) = K(s)F_0(\omega)\,, \tag{3.5}$$

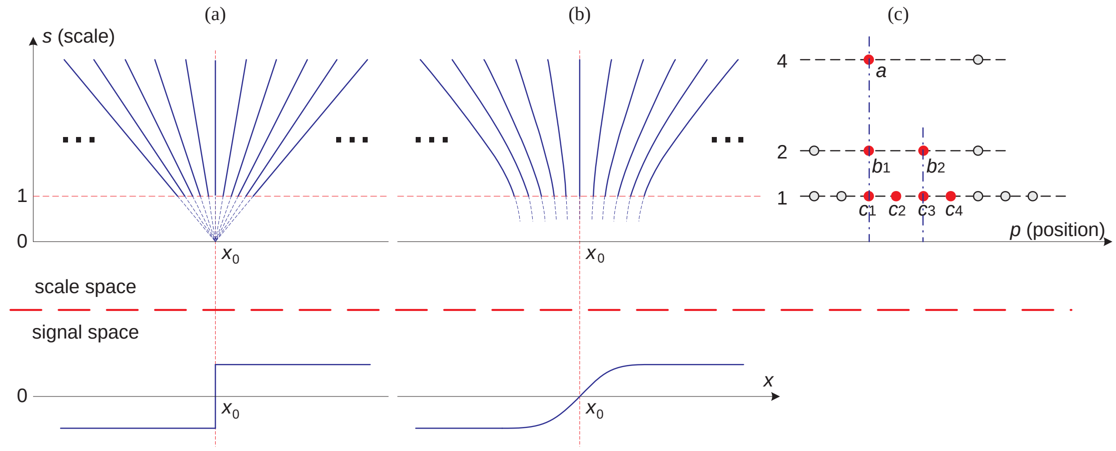


Figure 3.2: Local phase coherence structure near (a) a sharp step edge and (b) a blurred step edge, where the equal-phase contours in the scale space are shown. The phase congruency relationship corresponds to the central vertical contours at position $x_0$ in both (a) and (b), but does not describe the phase structures in the vicinity of $x_0$. (c): positions of discrete sampling across three consecutive dyadic scales in the scale space.

where $K(s)$ is a real function of only $s$, but independent of $\omega$, then:

$$F(s,p) = \frac{K(s)}{2\pi\sqrt{s}} \int_{-\infty}^{\infty} F_0(\omega)\, G(\omega - \omega_c)\, e^{j\omega(p-x_0)/s} d\omega \,, \tag{3.6}$$

This suggests a predictable structure of $F(s,p)$ from $F(1,p)$ given by

$$F(s,p) = \frac{K(s)}{K(1)\sqrt{s}} F(1, x_0 + \frac{p - x_0}{s}) \,. \tag{3.7}$$

Since both $K(s)$ and $s$ are real, we obtain the following phase relationship of $F(s,p)$ [22]:

$$\Phi(F(s,p)) = \Phi(F(1, x_0 + \frac{p - x_0}{s})) \,. \tag{3.8}$$

This result mathematically validates the local phase relationship across scale and space, where equal phase contours in the $(s,p)$ plane form straight lines that converge exactly at the location of the feature $x_0$, as schematically illustrated in Fig. 3.2(a). These straight lines are defined by $x_0 + (p - x_0)/s = D$, where $D$ is a constant. The phase congruency relationship constitutes a subset of LPC by predicting the phases in the same way for the center vertical line ($p = x_0$) only. Note that the derivation of the LPC phase pattern is based on the assumption that $f_0$ is a scale invariant signal, which turns out to be true for distinctive sharp features such as an isolated impulse or a step edge in a 1D signal, or an edge, a line, or an isolated impulse in a 2D image. It is a common view that edges and lines are particularly rich sources of image information, and that organization of these features could provide the basis for an efficient description of an image. This gives another property that distinguishes the LPC relationship with phase congruency, which does not require the scale-invariance feature and holds for other types of features also (e.g., a blurred edge). This is illustrated in Fig. 3.2(b), where blur does not change the central vertical contour ($p = x_0$), but distorts the general LPC patten away from the center.

## 3.3 Local Phase Coherence Measure

In [60], the behavior of local phases at adjacent scales was first studied. The importance of local phases is demonstrated in synthesizing 3D texture images. A coarse-to-fine scale statistical phase prediction (also called relative phase) is developed to capture the linearly varying behavior of local phases with distance from features. Wang and Simoncelli [22] examined the local phase structures at distinctive features in more depth. Similar to [60], a coarse-to-fine scale phase prediction is developed to demonstrate the effect of blur on disrupting local phases coherence. One common limitation in [60, 22] is that coarse-to-fine scale can only be computed with limited number of coefficients and fixed integer scale factors. In subsequent sections, new methods have been developed to overcome this limitation. It is worth noting that, non of the previous works attempt to develop an absolute measure of LPC strength for image features. The importance of such a measure seems intuitively sensible from the perspective of visual function. In particular, the accurate localization of image features is critical to a variety of visual applications, including stereopsis and motion estimation. The basic idea behind LPC strength measure is that the phase of a wavelet coefficient can be predicted using the phases of its neighboring coefficients in the scale-space when LPC relationship is satisfied at a spatial location [22]. The prediction accuracy (i.e., the closeness between $\Phi(c_1)$ and $\hat{\Phi}(c_1)$) can then be used as a local measure of the strength of the LPC relationship of a particular scale and at a given location. An example is shown in Fig. 3.2(c), where the finest scale coefficients $c_i$ for $i = 1, 2, 3, 4$ can be predicted from their coarser scale neighbors $a$, $b_1$ and $b_2$. As an example defined in [22],

$$\hat{\Phi}(c_1) = -2\Phi(a) + 3\Phi(b_1)\,, \tag{3.9}$$

where $\hat{\Phi}(c_1)$ denotes the prediction of the true phase $\Phi(c_1)$.

In our work [61, 62, 63], a method was proposed to compute a spatial LPC map in the complex wavelet transform domain based on the complex version of the steerable pyramid decomposition [64]. Following a 3-scale, multi-orientation steerable pyramid transform, the phases of the finest scale coefficients are predicted using their corresponding parent and grandparent coefficients using an LPC-based phase predictor (such as (3.9)). At each spatial location, an LPC strength measure is then defined as

$$P_i = \frac{\sum_j |c_{ij}| \cos\Big(\Phi\left(\{c_{ij}\}\right) - \widehat{\Phi}\left(\{c_{ij}\}\right)\Big)}{\sum_j |c_{ij}|}\,, \tag{3.10}$$

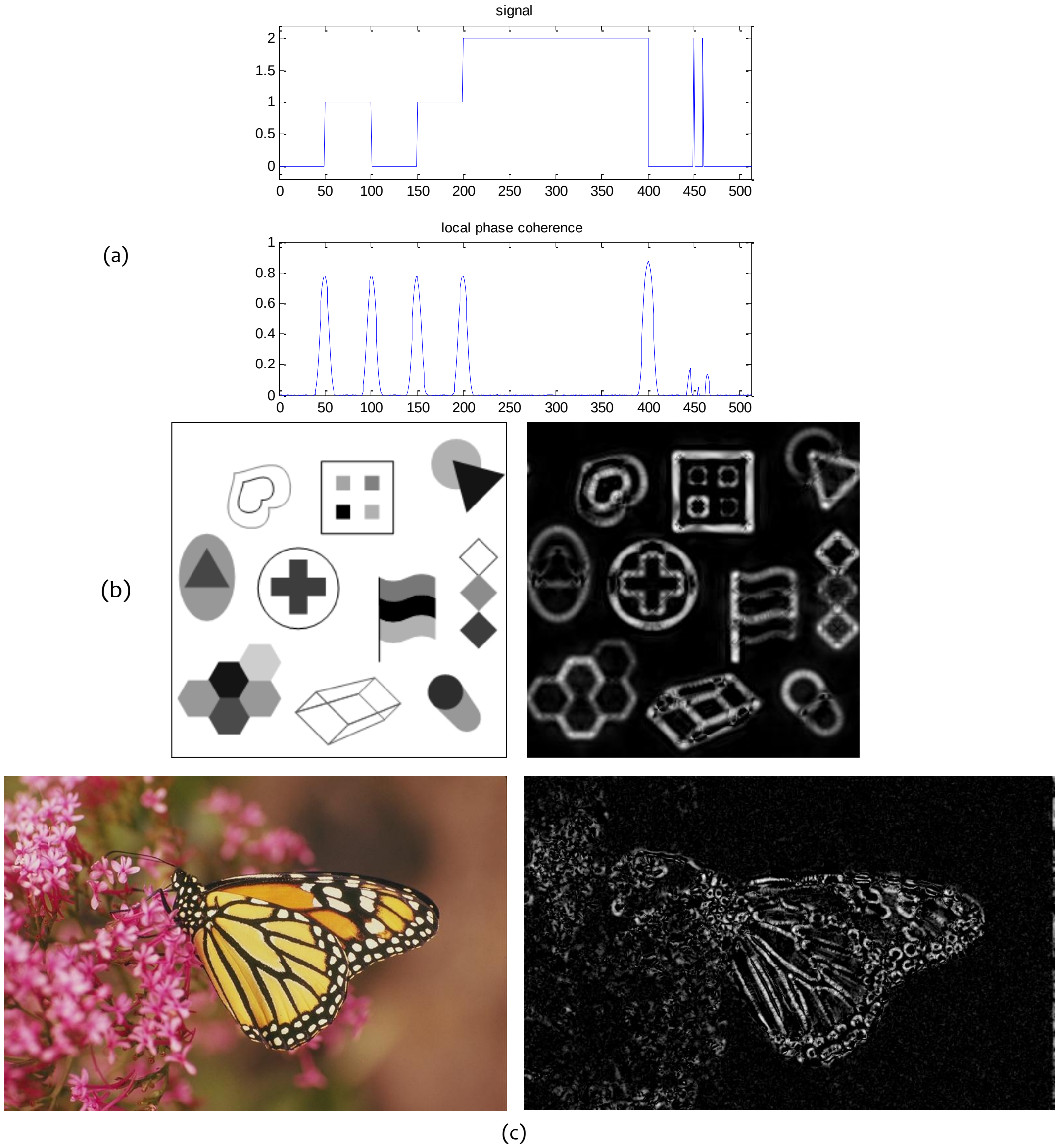


Figure 3.3: Local phase coherence map using dyadic scales in the scale space plan of (a)a 1D signal (b) a phantom image and (c) a natural image.

where $\Phi\left(\{c_{ij}\}\right)$ and $\widehat{\Phi}\left(\{c_{ij}\}\right)$ are the true and predicted phases of the $i$-th coefficient in the $j$-th orientation, respectively. This measure achieves the maximal value of unity when the phase prediction (and thus LPC relationship) is perfect. This is expected to occur in the vicinity of distinctive sharp image features. The measure is weighted by the magnitudes of the coefficients over orientations, so that the orientations that contain more energy are given higher weights. Figure 3.3 shows examples of LPC maps of 1D signal, a phantom image, and a natural image. It can be observed that high LPC values are achieved at sharp features, regardless of feature type (edges or lines), contrast or luminance strength, and the polarity of foreground-background intensities. Such a LPC measure had been employed in image registration [61], fusion [62], and sharpness assessment [63] applications as we will describe in the following chapters.

## 3.4 Calculating Local Phase Coherence in Arbitrary Scales

The computation of LPC described in [22] has strong limitations. Specifically, the computation is applicable to three spatially co-locating complex wavelet coefficients spread in three consecutive dyadic scales only, where the widths of the wavelet filters expand by factors of 1, 2, and 4, respectively, from the finest to the coarsest scales. In practice, this limitation on one hand restricts the accuracy of phase measurement hence feature detection precision and on the other hand limits its application, as such a large expansion of the coarsest scale filter is often problematic because nearby image features may interfere with the responses of these coefficients (but may or may not affect the finer scale coefficients). As a result, the LPC relationship is often corrupted even if the local image feature is sharp. Consequently, it is highly desirable to develop more flexible (preferably more space- and scale-localized) LPC computation methods.

To develop a new method to compute LPC, we would need to closely examine the phase patterns at sharp features. Like in [22], we analyze the LPC relationship in 1D and the results can be directly extended to 2D. In 1D, the most common sharp features are impulses and ideal step edges. In the case of an impulse $f_0(x) = A\delta(x)$, where $A$ is a non-zero constant and $\delta(x)$ is the impulse function, the Fourier transform of $f_0(x)$ is $F_0(\omega) = A$ for all $\omega$, and thus $F_0(\omega/s) = A$ and $K(s)$ in (3.5) equals unity. Substitute this into (3.6), we have

$$\begin{aligned} F(s,p) &= \frac{1}{2\pi\sqrt{s}} \int_{-\infty}^{\infty} A\, G(\omega-\omega_c)\, e^{j\omega(p-x_0)/s} d\omega \\ &= \frac{A}{\sqrt{s}}\, g\left(\frac{p-x_0}{s}\right) e^{j\omega_c \frac{p-x_0}{s}} \;, \end{aligned} \tag{3.11}$$

The derivation above is based on the fact that the RHS of the first row constitutes the inverse Fourier transform of $\frac{A}{\sqrt{s}}G(\omega-\omega_c)$ evaluated at $(p-x_0)/s$. The phase of $F(s,p)$ is then

$$\Phi(F(s,p)) = \begin{cases} \frac{\omega_c(p-x_0)}{s} & \text{when } A > 0 \\ \frac{\omega_c(p-x_0)}{s} + \pi & \text{when } A < 0 \end{cases} . \tag{3.12}$$

In the case of an ideal step edge, $f_0(x) = B\left[u(x) - \frac{1}{2}\right]$, where $B$ is a non-zero constant and $u(x)$ is the step function. The Fourier transform of $f_0(x)$ is $F_0(\omega) = \frac{B}{j\omega}$

and $F_0\left(\frac{\omega}{s}\right) = \frac{Bs}{\jmath\omega} = sF_0(\omega)$. Thus $K(s) = s$ in (3.5). Substitute this into (3.6), we have

$$F(s,p) = \frac{B\sqrt{s}}{2\pi}\int_{-\infty}^{\infty}\frac{G(\omega-\omega_c)}{\jmath\omega}\,e^{j\omega(p-x_0)/s}d\omega\,, \tag{3.13}$$

where the RHS constitutes the inverse Fourier transform of $\frac{B\sqrt{s}G(\omega-\omega_c)}{\jmath\omega}$ evaluated at $(p-x_0)/s$. Based on the integration property of Fourier transform, we obtain

$$\begin{aligned} F(s,p) &= B\sqrt{s}\int_{-\infty}^{\frac{p-x_0}{s}} g(x)\,e^{j\omega_c x}dx \\ &= \frac{B\sqrt{s}}{\jmath\omega_c}\left[g(x)e^{j\omega_c x}\Big|_{-\infty}^{\frac{p-x_0}{s}} - \int_{-\infty}^{\frac{p-x_0}{s}} g'(x)\,e^{j\omega_c x}dx\right]. \end{aligned} \tag{3.14}$$

Since $g(x)$ is a slowly varying function localized near $x=0$, we have $g(-\infty)=0$ and $g'(x)\approx 0$. Therefore,

$$F(s,p) \approx \frac{B\sqrt{s}}{\jmath\omega_c}\,g\left(\frac{p-x_0}{s}\right)e^{j\omega_c\frac{p-x_0}{s}}\,. \tag{3.15}$$

The phase is then computed as

$$\Phi(F(s,p)) \approx \begin{cases} \frac{\omega_c(p-x_0)}{s} - \frac{\pi}{2} & \text{when } B>0 \\ \frac{\omega_c(p-x_0)}{s} + \frac{\pi}{2} & \text{when } B<0 \end{cases}. \tag{3.16}$$

Combining (3.12) and (3.16), we obtain a general LPC relationship given by

$$\Phi(F(s,p)) \approx \hat{\Phi}(s,p) \equiv \frac{\omega_c(p-x_0)}{s} + \frac{k\pi}{2}\,, \tag{3.17}$$

where $k$ is an integer depending on the nature of the sharp feature. Specifically, it equals 0, 2, $-1$, and 1 for positive impulse, negative impulse, ideal ascending step edge, and ideal descending step edge, respectively.

The general relationship in (3.17) provides a model of the phase pattern in the vicinity of a sharp image feature. Given the phases of $N$ sample coefficients arbitrarily extracted from $F(s,p)$ near $x_0$ and their corresponding predictions

$$\mathbf{\Phi} = \begin{bmatrix} \Phi[F(s_1,p_1)] \\ \Phi[F(s_2,p_2)] \\ \vdots \\ \Phi[F(s_N,p_N)] \end{bmatrix}, \quad \hat{\mathbf{\Phi}}_{x_0,k} = \begin{bmatrix} \hat{\Phi}(s_1,p_1) \\ \hat{\Phi}(s_2,p_2) \\ \vdots \\ \hat{\Phi}(s_N,p_N) \end{bmatrix}, \tag{3.18}$$

we can quantify the strength of LPC by assessing the closeness between the true phases of these coefficients and their optimal model predictions:

$$S_{LPC} = \max_{\{x_0,k\}} S(\mathbf{\Phi}, \hat{\mathbf{\Phi}}_{x_0,k})\,, \tag{3.19}$$

where $S(.,.)$ is a similarity measure between the true and predicted phase samples given in vector form as in (3.18).

The most straightforward way to define the similarity function $S$ is to make it monotonically decreasing with the squared error between the vectors. For example, we may define

$$S(\mathbf{\Phi}, \hat{\mathbf{\Phi}}_{x_0,k}) = \frac{1}{1 + \alpha \, \|\mathbf{\Phi} - \hat{\mathbf{\Phi}}_{x_0,k}\|^2} \,, \tag{3.20}$$

where $\alpha$ is a positive constant which controls the speed of decaying in the similarity measure as a function of the mean squared error between the true and predicted phases. This similarity function is bounded between 0 and 1. As a result, solving the optimization problem in (3.19) is equivalent to solving a least square optimization problem. For each given value of $k$ (for $k = -1, 0, 1, 2$), the optimal value $x_0$ in (3.19) can be found in closed-form, and the overall optimization problem can then be resolved by picking the value $k$ and its corresponding $x_0$ that lead to the minimal squared error. Although this method is conceptually simple and computationally efficient, the solutions turn out to be erroneous and unreliable in our experiment. The reason is because phase variables are not normal scalar quantities but are angular variables that have the $2\pi$ wrap-around effect (for example, an error of $2\pi$ is equivalent to 0 error for an angular variable but is significant in the linear scale).

To overcome the wrap-around effect as well as the ambiguity between different types of features (that correspond to different $k$ values), we used a different method in [65], which corresponds to defining a similarity function as

$$S(\mathbf{\Phi}, \hat{\mathbf{\Phi}}_{x_0,k}) = \frac{1}{N} \sum_{i=1}^{N} \cos\{4\Phi[F(s_i, p_i)] - 4\hat{\Phi}(s_i, p_i)\} \,. \tag{3.21}$$

This similarity function is bounded between $-1$ and 1. Notice that the factor 4 here makes the last term in (3.17) a multiplier of $2\pi$. This factor, when combined with the use of the cosine function, eliminates both the wrap-around effect and the ambiguity between different features. Although this similarity definition is conceptually elegant, it makes the optimization problem in (3.19) difficult to solve, and in general, no closed-form solution can be found. In practice, we have to resort to numerical optimization tools, which often lead to extremely high computational cost (as in [65]), especially when this approach is applied to 2D images, where the optimization procedure needs to be carried out at every location in the image.

## 3.5 Efficient LPC Evaluation Approach

Given the difficulties in the LPC evaluation methods described above, our focus below will be on practical algorithms that lead to efficient assessment of LPC. It is worth noting that the general formulation laid out in (3.19) allows for arbitrary selections of samples (in terms of both the number and positions of the samples) of

$F(s, p)$ in the scale-space. In practice, this is unlikely and unnecessary. Motivated by the method used in [22], we may pick a set of samples at specific positions and scales in the scale-space, so that the LPC computation can be largely simplified. In particular, if we can find a set of samples $\hat{\mathbf{\Phi}}$, together with a corresponding set of weights $\mathbf{w} = [w_1, w_2, \cdots, w_N]^T$, such that

$$\mathbf{w}^T \, \hat{\mathbf{\Phi}} = 0 \, , \tag{3.22}$$

then we may define a simple measure of LPC strength by

$$S_{LPC} = \cos(\mathbf{w}^T \, \mathbf{\Phi}) \, . \tag{3.23}$$

The value of this LPC measure is bounded between $-1$ and 1, and the maximal value is achieved when $\mathbf{w}^T \, \mathbf{\Phi} = 0$, which is consistent with the phase relationship defined in (3.22).

To provide an example (that may lead to a practically useful LPC measure), let us assume that we extract $N$ samples in the scale-space that are aligned at the same position $p$ (which may not be aligned with the feature position $x_0$) but at different scales. Substitute (3.17) into (3.22), we obtain

$$\omega_c(p - x_0) \left( \sum_{i=1}^{N} \frac{w_i}{s_i} \right) + \frac{k\pi}{2} \left( \sum_{i=1}^{N} w_i \right) = 0 \, . \tag{3.24}$$

In order for this to be true for all possible values of $\omega_c$, $p$, $x_0$ and $k$, we would need the following simultaneous equations to be true

$$\begin{cases} \sum_{i=1}^{N} w_i = 0 \\ \sum_{i=1}^{N} (w_i / s_i) = 0 \end{cases} \, . \tag{3.25}$$

Without loss of generality, we assume $w_1 = 1$. This results in $N - 1$ unknowns ($w_2$, $w_3$, $\cdots$, $w_N$) with two equations. In the case that $N = 3$, the solutions are unique and are given by

$$\begin{cases} w_1 = 1 \\ w_2 = \frac{s_2(s_3 - s_1)}{s_1(s_2 - s_3)} \\ w_3 = \frac{s_3(s_2 - s_1)}{s_1(s_3 - s_2)} \end{cases} \tag{3.26}$$

When $N > 3$, we can solve for a least square weight energy solution under the constraints of (3.25). Using a Lagrange multiplier approach, we define the following energy function

$$E = \sum_{i=1}^{N} w_i^2 + \lambda_1 \left( \sum_{i=1}^{N} w_i \right) + \lambda_2 \left( \sum_{i=1}^{N} \frac{w_i}{s_i} \right) \, , \tag{3.27}$$

where $\lambda_1$ and $\lambda_2$ are the Lagrange multipliers. Taking derivative of $E$ with respect to $\lambda_1$, $\lambda_2$ and $w_i$ for $i = 1, 2, \cdots, N$ and setting them to zero, we obtain a linear system of equations

$$\mathbf{A}\mathbf{w}^* = \mathbf{b} \, , \tag{3.28}$$

where

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} 1 & 0 & \cdots & 0 & 1/2 & 1/2s_2 \\ 0 & 1 & \cdots & 0 & 1/2 & 1/2s_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1/2 & 1/2s_N \\ 1 & 1 & \cdots & 1 & 0 & 0 \\ 1/s_2 & 1/s_3 & \cdots & 1/s_N & 0 & 0 \end{bmatrix} \\ \mathbf{w}^* &= \begin{bmatrix} w_2 & w_3 & \cdots & w_N & \lambda_1 & \lambda_2 \end{bmatrix}^T \\ \mathbf{b} &= \begin{bmatrix} 0 & 0 & \cdots & 0 & -1 & -1/s_1 \end{bmatrix}^T . \end{aligned} \tag{3.29}$$

We can then solve for the weights by

$$\mathbf{w}^* = \mathbf{A}^{-1}\mathbf{b} \, . \tag{3.30}$$

In practice, one would likely to choose $s_1$ to be the finest scale ($s_1 = 1$) for maximal localization, and choose the other $s_i$ values to be evenly spaced in either linear or logarithm scale. For example, in $N = 3$ case,

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1+d \\ 1+2d \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 \\ r \\ r^2 \end{bmatrix} . \tag{3.31}$$

The weight solutions of these two cases are computed as

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2(1+d) \\ 1+2d \end{bmatrix} \text{ and } \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -(1+r) \\ r \end{bmatrix} , \tag{3.32}$$

respectively. The solutions of the two cases for specific selections of $d$ and $r$ values are given in Table 3.1. Interestingly, the previous LPC computation (3.9) becomes a special case that can be directly derived from the row for $r = 2$ in Table 3.1. In the cases of $N = 4$ and $N = 5$, the least square weight energy solutions for specific values of $d$ and $r$ are shown in Tables 3.2 and 3.3, respectively.

Figure 3.4 depict the resulted LPC map using three scales with scale factor $d = 1/2$ for same images in Fig. 3.3. It is worth noting that in case of natural image, features are usually smoothed or blurred in someway due to sampling and noise, this reduces the spread of frequencies present in the signal, which in turn results in a poorly localized phase coherence. A very simple thresholding method has been employed to address this issue in the new approach. Comparing Fig. 3.3 and Fig. 3.4, we can conclude that this new approach of calculating local phase coherence provides much better localization and also offers other possibilities in the use of local phase coherence in image processing applications as we will see in the following chapters.

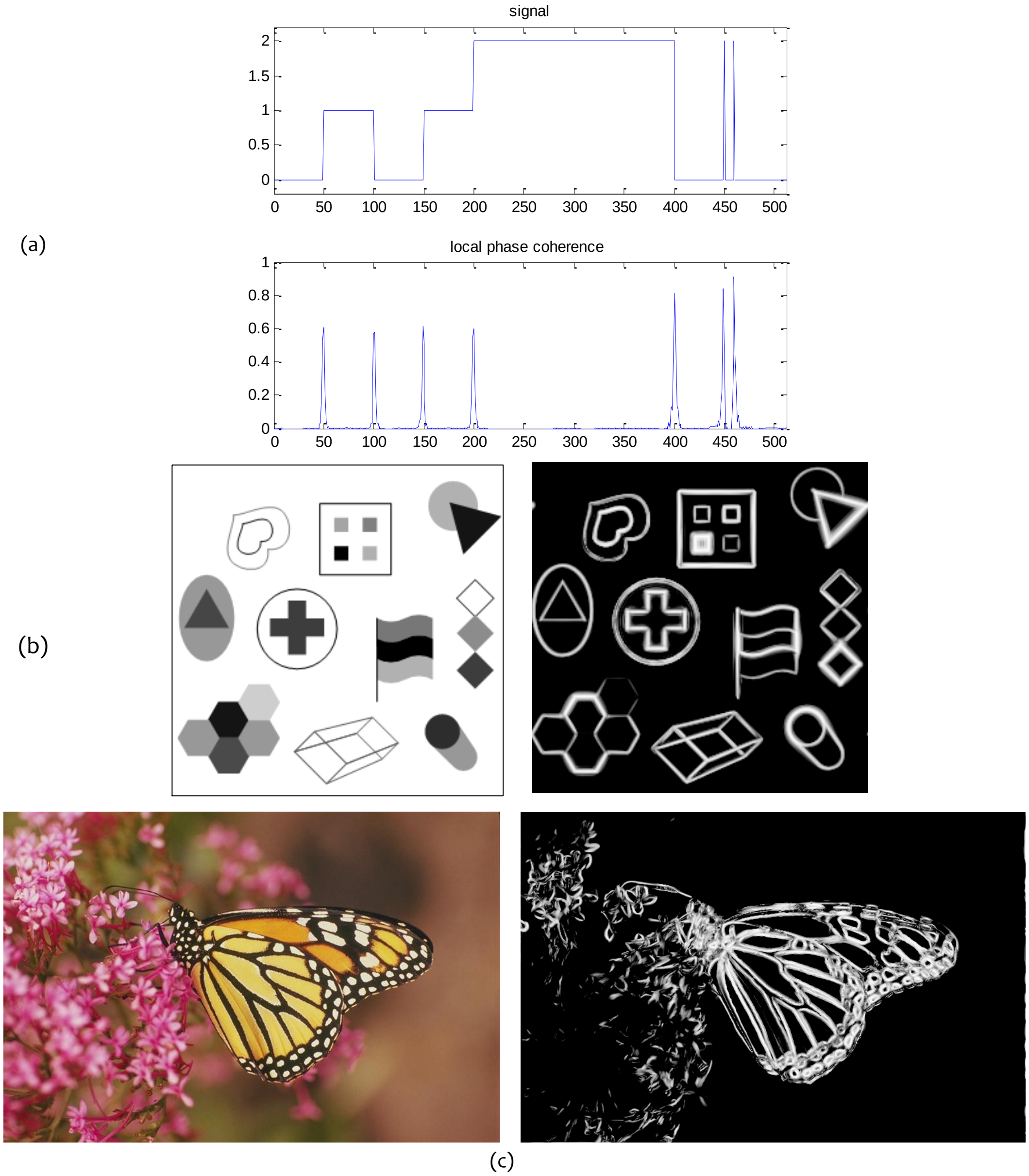


Figure 3.4: Local phase coherence map using arbitrary scale ($d = 1/2$) in the scale space plan of (a)a 1D signal (b) a phantom image and (c) a natural image.

## 3.6 Conclusions

The purpose of this chapter is to extend the theory and methodology of local phase coherence, so that it can be converted to more practical techniques that can be applied to various signal processing applications for the analysis of signals and the detection of features. The major contribution of the current work as opposed to

Table 3.1: Weight Solutions For Three Successive Scales

| | $s_1$ | $s_2$ | $s_3$ | $w_1$ | $w_2$ | $w_3$ |
|---|---|---|---|---|---|---|
| | 1 | $1+d$ | $1+2d$ | 1 | $-2(1+d)$ | $1+2d$ |
| $d=1/4$ | 1 | $5/4$ | $3/2$ | 1 | $-5/2$ | $3/2$ |
| $d=1/2$ | 1 | $3/2$ | 2 | 1 | $-3$ | 2 |
| $d=1$ | 1 | 2 | 3 | 1 | $-4$ | 3 |
| $d=3/2$ | 1 | $5/2$ | 4 | 1 | $-5$ | 4 |
| $d=2$ | 1 | 3 | 5 | 1 | $-6$ | 5 |
| | 1 | $r$ | $r^2$ | 1 | $-(1+r)$ | $r$ |
| $r=5/4$ | 1 | $5/4$ | $25/16$ | 1 | $-9/4$ | $5/4$ |
| $r=\sqrt{2}$ | 1 | $\sqrt{2}$ | 2 | 1 | $-1-\sqrt{2}$ | $\sqrt{2}$ |
| $r=\sqrt{3}$ | 1 | $\sqrt{3}$ | 3 | 1 | $-1-\sqrt{3}$ | $\sqrt{3}$ |
| $r=2$ | 1 | 2 | 4 | 1 | $-3$ | 2 |
| $r=\sqrt{5}$ | 1 | $\sqrt{5}$ | 5 | 1 | $-1-\sqrt{5}$ | $\sqrt{5}$ |
| $r=3$ | 1 | 3 | 9 | 1 | $-4$ | 3 |

Table 3.2: Weight Solutions For Four Successive Scales

| | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $w_1$ | $w_2$ | $w_3$ | $w_4$ |
|---|---|---|---|---|---|---|---|---|
| $d=1/3$ | 1 | $4/3$ | $5/3$ | 2 | 1 | -1.9474 | -0.1316 | 1.0789 |
| $d=1/2$ | 1 | $3/2$ | 2 | $5/2$ | 1 | -2.2347 | -0.0408 | 1.2755 |
| $d=2/3$ | 1 | $5/3$ | $7/3$ | 3 | 1 | -2.5166 | 0.0464 | 1.4702 |
| $d=1$ | 1 | 2 | 3 | 4 | 1 | -3.0714 | 0.2143 | 1.8571 |
| $d=2$ | 1 | 3 | 5 | 7 | 1 | -4.7089 | 0.6962 | 3.0127 |
| $r=2^{1/3}$ | 1 | $2^{1/3}$ | $2^{2/3}$ | 2 | 1 | -1.5962 | -0.2401 | 0.8363 |
| $r=3^{1/3}$ | 1 | $3^{1/3}$ | $3^{2/3}$ | 3 | 1 | -1.7828 | -0.1683 | 0.9511 |
| $r=4^{1/3}$ | 1 | $4^{1/3}$ | $4^{2/3}$ | 4 | 1 | -1.9320 | -0.1084 | 1.0404 |
| $r=6^{1/3}$ | 1 | $6^{1/3}$ | $6^{2/3}$ | 6 | 1 | -2.1686 | -0.0097 | 1.1784 |
| $r=2$ | 1 | 2 | 4 | 8 | 1 | -2.3571 | 0.0714 | 1.2857 |

existing LPC computation is to formulate the problem using an optimization framework. Several technical issues have been studied in order to overcome a series of problems encountered in formulating the optimization problem and in finding the optimal solutions. The resulting LPC computation exhibits significantly broadened flexibilities such that it can be computed with arbitrary grouping of neighboring complex wavelet coefficients spread at any fractional scale ratios between successive scales. These flexibilities make our approach desirable in many potential applications, especially in the cases when multiple features exist and are close to each other, when only partial information of local phases is available, and/or when significant noise exists in the signal. In the following chapters, we will demonstrate how the methodology developed in this work can be applied to practical image processing applications so as to better exploit the advantages of LPC.

Table 3.3: Weight Solutions for Five Successive Scales

| | $s_1$ | $s_2$ | $s_3$ | $s_4$ | $s_5$ | $w_1$ | $w_2$ | $w_3$ | $w_4$ | $w_5$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $d = 1/4$ | 1 | 5/4 | 3/2 | 7/4 | 2 | 1 | -1.4477 | -0.4827 | 0.2067 | 0.7237 |
| $d = 1/2$ | 1 | 3/2 | 2 | 5/2 | 3 | 1 | -1.8458 | -0.4581 | 0.3744 | 0.9295 |
| $d = 3/4$ | 1 | 7/4 | 5/2 | 13/4 | 4 | 1 | -2.2252 | -0.4350 | 0.5289 | 1.1314 |
| $d = 1$ | 1 | 2 | 3 | 4 | 5 | 1 | -2.5957 | -0.4137 | 0.6774 | 1.3320 |
| $d = 3/2$ | 1 | 5/2 | 4 | 11/2 | 7 | 1 | -3.3237 | -0.3745 | 0.9661 | 1.7321 |
| $r = 2^{1/4}$ | 1 | $2^{1/4}$ | $\sqrt{2}$ | $2^{3/4}$ | 2 | 1 | -1.1937 | -0.4932 | 0.0958 | 0.5911 |
| $r = 3^{1/4}$ | 1 | $3^{1/4}$ | $\sqrt{3}$ | $3^{3/4}$ | 3 | 1 | -1.3271 | -0.4818 | 0.1604 | 0.6484 |
| $r = \sqrt{2}$ | 1 | $\sqrt{2}$ | 2 | $2\sqrt{2}$ | 4 | 1 | -1.4314 | -0.4698 | 0.2102 | 0.6910 |
| $r = 6^{1/4}$ | 1 | $6^{1/4}$ | $\sqrt{6}$ | $6^{3/4}$ | 6 | 1 | -1.5930 | -0.4466 | 0.2858 | 0.7538 |
| $r = 8^{1/4}$ | 1 | $8^{1/4}$ | $2\sqrt{2}$ | $8^{3/4}$ | 8 | 1 | -1.7185 | -0.4255 | 0.3434 | 0.8006 |

# Chapter 4

# Image Sharpness Assessment

## 4.1 Introduction

Sharpness is one of the most important factors in photographic image quality. It is also a useful indicator of the amount of details the imaging system can reproduce. It can be described as how clear details of objects in a photograph are visible. Blur is closely associated with sharpness, because it is the most common type of distortion that impairs visual sharpness. In practice, blur may occur during image acquisition, e.g., due to atmospheric turbulence, camera motion, or out-of-focus of the optical system. It may also be a side effect of certain image processing operations such as compression and de-noising. Figure 4.1 shows examples of blur types. Amazingly, the HVS can effortlessly point out blur and sharp regions in an image without seeing the original image. Yet, this task remains quite challenging for a computer system. With the increasing number of cameras in mobile devices, automatically assisting and quantifying the sharpness of captured images becomes more important.

In a large number of applications, the reference image is unavailable. Therefore, no-reference (NR) IQA that does not require any access to the reference image is highly desirable [66]. Applications of sharpness assessment measures may include, auto-enhancement algorithms to sharpen images in a spatially adaptive fashion[67]. It is also useful in computer vision tasks such as detecting shadows which often have softer edges, adjusting parameters of recognition/tracking methods [68], and identifying in-focus and out-of-focus areas of an image [69]. When capturing a scene, it is often difficult for a user to determine whether a photo is well focused on a small mobile screen, so a real-time method for estimating sharpness directly can be useful to provide feedback to users in case a re-capture is required. In this chapter, we present a novel NR image sharpness/blur estimator from a different and complementary perspective. The idea is originated from the notion of local phase coherence theory described in Chapter 3.

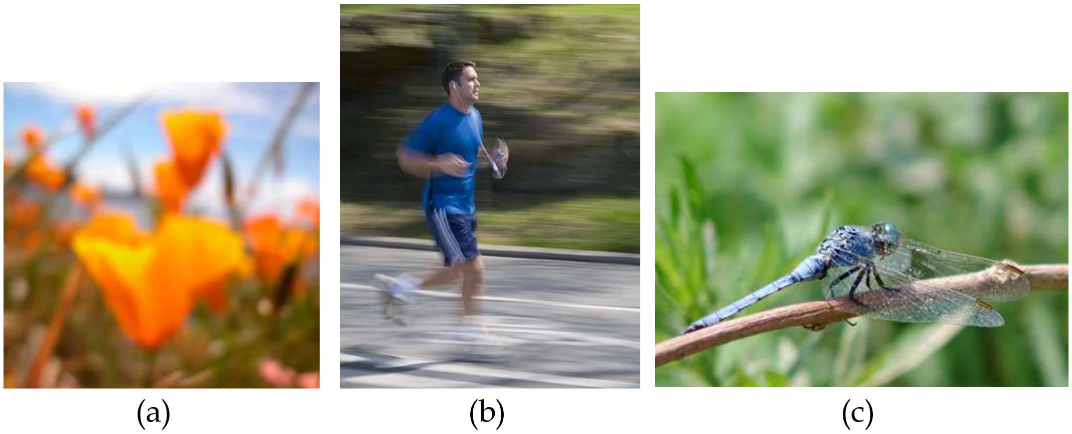

(a) (b) (c)

Figure 4.1: (a) Out-of-focus blur (b) Motion-blur (c) Blur due to limited depth of field when objects are at different distances.

## 4.2 Previous Work

In the literature, several types of computational models have been proposed to account for the perception of blur. The first type of models examine the power spectrum of an image in global/local frequency transformations such as the Fourier transform [70]. This is motivated by the fact that blur is often created through low pass filtering, which smooths images and reduces their high frequency energy. As a result, the power spectrum of a blurred image falls faster than the original sharp natural image, whose power spectrum falloff often obeys a power law [70]. Following the framework, a sharpness/blur measure may be implemented by evaluating high-frequency energy [71], by computing an HVS-weighted summation of normalized power spectrum [72], by counting near-zero high frequency discrete cosine transform (DCT) coefficients [73], or by calculating the ratio between high and low frequency energy [74].

The second class of models focus on edges and associate blur with edge width. Edge detection algorithms are first employed to find edges and edge widths are estimated along either horizontal/vertical [75] or local gradient direction [76], followed by a blur measure computed as the average edge width over all edges detected [75]. In [11], the edge detection and edge width approaches were incorporated with a novel concept of just noticeable blur (JNB), which is a perceptual model indicating the probability of detecting a blur distortion by human eyes. It was shown that the JNB value is a function of local contrast and psycho-physical experiments were carried out to calibrate the model [77]. The JNB method was further refined in a cumulative probability of blur detection (CPBD) measure [10], which is based on a probabilistic framework on the sensitivity of human blur perception at different contrasts. The CPBD measure achieves state-of-the-art performance when tested with multiple image databases [10].

The third category of blur prediction models are based on statistics on the distribution of pixel intensities or transform coefficients. Pixel intensity statistical methods are based on the hypothesis that sharper images have larger variance or higher entropy in their intensity values [78, 79]. Kurtosis of 2D DCT coefficient distribution has also been found to be a useful measure to characterize image sharpness [80]. In [80], such a kurtosis measure is combined with an edge profile based method, leading to an effective image sharpness measure. Total variation (TV) between adjacent pixel values in an image region has been employed as a measure of local sharpness in a probabilistic framework [81].

The fourth type of models employ local gradient measures based on the observation that the strength of local gradient is sensitive to image blur. In [12], singular value decomposition is applied to groups of gradient vectors computed within local patches. The two resulting singular values provide useful measures of gradient strength along the dominant direction and its perpendicular direction, respectively. A sharpness measure $H$ is then defined by making use of the dominant singular value as well as prior knowledge about the noise variance. In [82], a normalized form that accounts for the relative gradient strength between the dominant and its perpendicular directions was proposed. This results in an improved sharpness measure $Q$ that was employed for automatic parameter selection of image denoising algorithms.

All aforementioned four types of blur/sharpness models are well-motivated and are shown to be effective in capturing certain aspects about the impact of blur on perceived sharpness, and they have achieved notable success when tested using a number of subject-rated databases. On the other hand, these methods are often limited by the complication of the large variations between real-world images in terms of information content and complexity. For example, it was pointed out that individual images show significant variability in their amplitude spectra both in shape and in the speed of falloff [83]. For another example, the overall gradient strength of an image not only depends on the degree of blur, but is also largely affected by the amount of sharp detail presented in the original source image. A useful idea is to combine the merits of different blur/sharpness models. For example, the recently proposed $S_3$ algorithm [9] combines TV-based spatial sharpness measure and block-wise power spectral falloff features and achieves significantly improved performance.

## 4.3 LPC-Based Sharpness Measure

Given an input image whose sharpness is to be evaluated, wavelet analysis is done by convolving the input image with log-Gabor filters. Log-Gabor filters are chosen based on the suggestion of Field [20]. These filters allow arbitrarily large bandwidth filters to be constructed while maintaining a zero DC component in the even-symmetric filter. The algorithm starts by first passing the input image through a series of $N$-scale $M$-orientation log-Gabor filters without any subsequent down-sampling process. This results in $MN$ "sub-bands" and there are $MN$ complex coefficients at each spatial location across all orientations and all scales. Let $c_{ijk}$ be the complex coefficient at the $i$-th scale, the $j$-th orientation and the $k$-th spatial location. Then at the $j$-th orientation and the $k$-th location, we can compute the LPC strength by

$$\begin{aligned} S_{LPC}^{\{j,k\}} &= \cos(\mathbf{w}^T \mathbf{\Phi}_j) = \cos\left(\sum_{i=1}^{N} w_i \Phi\{c_{ijk}\}\right) \\ &= \cos\left(\Phi\left\{\prod_{i=1}^{N} c_{ijk}^{w_i}\right\}\right) \\ &= \frac{\Re\left\{\prod_{i=1}^{N} c_{ijk}^{w_i}\right\}}{\left|\left\{\prod_{i=1}^{N} c_{ijk}^{w_i}\right\}\right|}, \end{aligned} \tag{4.1}$$

where $\Re\{\cdot\}$ denotes the real part of a complex number. This LPC strength measure is combined at each spatial location $k$ by a weighted average across all orientations, where the weights are determined by the magnitude of the first (finest) scale coefficient $c_{1jk}$ , so that the orientations that contain more energy are given higher importance:

$$S_{LPC}^{\{k\}} = \frac{\sum_{j=1}^{M} |c_{1jk}|\, S_{LPC}^{\{j,k\}}}{\sum_{j=1}^{M} |c_{1jk}| + C}, \tag{4.2}$$

where a constant $C$ is added to avoid instability when the magnitudes of the coefficients are close to zero.

The collection of $S_{LPC}^{\{k\}}$ at all locations constitutes a spatial LPC map. An example LPC map of a natural image of Fig. 4.2(a) is shown in Fig. 4.2(d). It can be observed that the local LPC strength measure responds strongly to sharp image structures around the sharp foreground region but weakly to the background out-of-focus regions. When the image is blurred as in Fig. 4.2(b), the strength of local LPC is reduced, reflecting the observation presented in Fig. 3.2, where it shows blur weakens the LPC relationship in the scale-space. When the image is severely blurred as in Fig. 4.2(c), the LPC relationship is completely disrupted, as shown in Fig. 4.2(f).

In order to provide an overall evaluation about the sharpness of the test image, we need to pool the LPC map into a single sharpness index. An effect in subjective sharpness assessment is that humans tend to make their judgment based on the sharpest region in the image [63]. For example, Fig. 4.4(a) is typically rated as a sharp image regardless of the out-of-focus background. This suggests that in pooling the LPC map, a mechanism is necessary to put more emphasis on the sharpest regions in the image. Here we propose a weighted averaging method based on ranked LPC values: Let $S_{LPC}^{\{k\}}$ for $k = 1, 2, ..., K$ be a collection of LPC values extracted from the LPC map, and let $S_{LPC}^{\{(k)\}}$ for $k = 1, 2, ..., K$ denote the sorted

LPC strength values such that $S_{LPC}^{\{(1)\}} \geq S_{LPC}^{\{(2)\}} \geq ... \geq S_{LPC}^{\{(K)\}}$. Then the overall LPC-based sharpness index (LPC-SI) is defined as

$$S_{LPC} = \frac{\sum_{k=1}^{K} u_k \, S_{LPC}^{\{(k)\}}}{\sum_{k=1}^{K} u_k} , \tag{4.3}$$

where $u_k$ is the weight assigned to the $k$-th ranked spatial LPC value and is computed as an exponentially decaying function given by

$$u_k = \exp\left[ - \left( \frac{k-1}{K-1} \right) \Big/ \beta_k \right] , \tag{4.4}$$

which gives a weight 1 to the highest LPC value and the decaying speed of the weights is controlled by the parameter $\beta_k$. In all the experimental results reported in the next section, the parameters are set as $M = 8$, $N = 3$, $s_1 = 1$, $s_2 = 3/2$, $s_3 = 2$, $C = 2$, and $\beta_k = 1e - 4$, respectively. These parameters are set empirically, but are found to be insensitive to the overall performance in our test with various blur image databases.

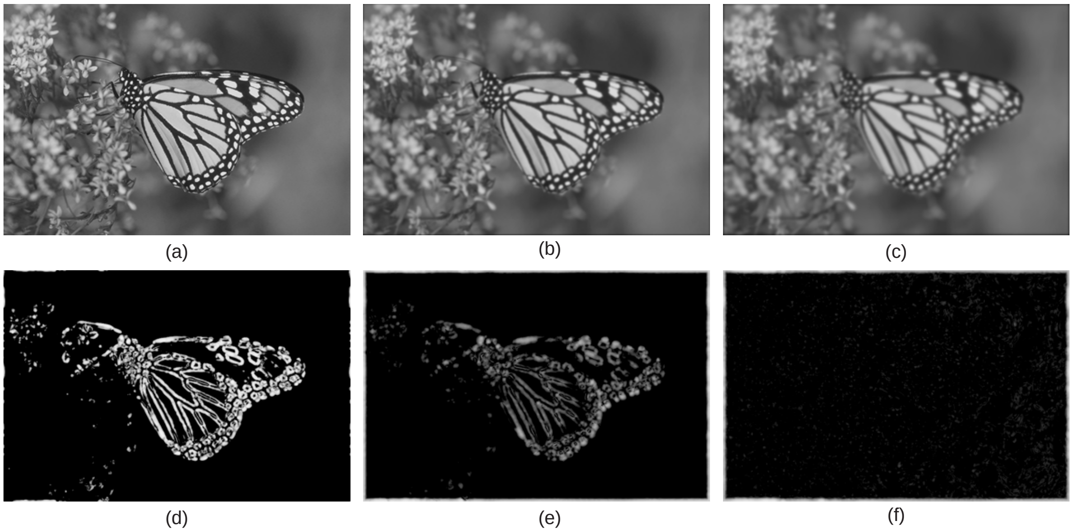


Figure 4.2: Original (a) and Gaussian blurred ”Monarch” images ((b) and (c)) at two blur levels, together with their corresponding LPC maps (d)-(f).

## 4.4 Validation

We test the performance of the proposed LPC-SI sharpness measure using four blur data sets obtained from four independent publicly available image databases:

- The LIVE database [1] was developed at The University of Texas at Austin, where the blur data set contains 174 images including 145 blurred and 29 reference images. All images were rated by 20-25 subjects on a continuous linear scale divided into 5 regions, namely *Bad*, *Poor*, *Fair*, *Good*, and *Excellent*. For each image, the mean opinion score (MOS) and the standard deviation between subjective scores were recorded.

- The Tampere Image Database 2008 (TID2008) [2] includes 100 Gaussian-blurred images obtained from 25 original images. Each image was evaluated by subjects from 0 to 9. The final MOS is obtained by averaging evaluated scores for a given image. Observers from three different countries (Italy, Finland and Ukraine) have performed two types of tests. The first test was conducted by direct evaluation of distorted images and the second by relative comparison between the quality of image pairs.

- The Categorical Image Quality (CSIQ) database [3] was developed at Oklahoma State University, where the blur data set contains 150 Gaussian-blurred images created from 30 original images at four to five distortion levels. The images are subjectively rated base on a linear displacement of the images across four calibrated LCD monitors placed side by side with equal viewing distance to the observer. Ratings are reported in the form of Difference of MOS (DMOS) between the original and blurred images.

- The IVC database [4] was developed at Ecole Polytechnique de l'Universite de Nantes. Four reference images have been distorted with 5 levels of Gaussian blur with a total of 20 blurred images. 15 subjects were asked to evaluate the distortion with respect to the original image on a five point scale. Subjective scores were reported in the form of MOS.

We compare the proposed LPC-SI method against seven existing NR sharpness measures, which include CPBD [10], JNBM [11], the H-metric [12], the Q-metric [82], $S_3$ [9], BRISQUE [7, 8], and BLIINDS-II [5, 6]. We have also included in our experiments two FR measures, PSNR and SSIM [48], which provide useful comparisons on the relative performance against the most widely used IQA measures. It should be noted that BRISQUE and BLIINDS-II are both general-purpose NR IQA techniques that are not designed for assessing blur/sharpness only.

Four criteria are employed for performance evaluation by comparing subjective and objective quality measures. Some of the criteria were included in previous tests carried out by the video quality experts group (VQEG) [84]. Other criteria were adopted in previous publications, e.g., [85]. These evaluation criteria are 1) Spearman rank-order correlation coefficient (SRCC); 2) Kendall's rank-order correlation coefficient (KRCC); 3) Pearson linear correlation coefficient (PLCC) after a nonlinear modified logistic mapping between the subjective and objective scores [86]; and 4) Root mean squared error (RMSE) between MOS values and model predicted MOS. SRCC and KRCC are both used to assess prediction monotonicity

[84]. PLCC and RMSE are employed to evaluate prediction accuracy [84]. A good objective quality measure is expected to achieve high values in PLCC, SRCC and KRCC, and low values in RMSE. In all experiments, only the blurred images in each data set have been included in the computation (i.e. reference images are excluded).

Table 4.1 summarizes the performance comparison results based on all four databases. For each evaluation criterion, we highlight the top two results obtained by NR methods with boldface. To provide an overall evaluation, Table 4.2 computes the average PLCC, SRCC and KRCC results for each objective measure over all four databases. The averages are computed using two methods. The first calculates the direct average of the correlation scores over the four databases, and the second computes a weighted average based on the sizes (or the numbers of images) of the image databases (specifically, the weight factors are 145 for LIVE, 100 for TID2008, 150 for CSIQ, and 20 for IVC databases, respectively). The results of both BRISQUE and BLIINDS-II are not included for the LIVE database because they both use LIVE images for training. Subsequently, their average results are computed over the other three databases only. From Tables 4.1 and 4.2, we observe that the proposed LPC-SI measure performs reasonably well and consistently over all four databases and in general are among the most competitive NR methods. Although the comparison is unfair, LPC-SI is often comparable to the FR PSNR, but is inferior to the FR SSIM measure. It is worth mentioning that the good performance of the proposed method is achieved without any edge detection or training processes.

Figure 4.3 shows the scatter plots of subjective versus objective quality scores after nonlinear mapping for the four databases, where each sample point represents one test image. The proposed LPC-SI method generally gives reasonable quality predictions, where the sample points tend to be clustered closer to the diagonal lines (representing perfect prediction) than the other five NR methods under comparison. Interestingly, we observe saturation effects of LPC-SI at both low and high quality ends of the scatter plots. The effect is more apparent in LIVE and CSIQ databases, where LPC-SI do not provide further distinctions when the quality of image is beyond or below the saturation levels. Similar effects are also observed in other state-of-the-art sharpness measures such as CPBD [10] and $S_3$ [9] in case of the CSIQ database. This may be due to the facts that image sharpness is not significantly affected with minor image distortions but is completely lost when the distortion is beyond certain level (and thus further distortions do not lead to further degradations of the sharpness measures). This effect also suggests that image sharpness assessment alone may not provide a complete solution to evaluating the quality of blurred images. Other approaches such as naturalness measures may be combined to improve the quality prediction performance.

Statistical significance analysis based on variance-based hypothesis testing provides additional information regarding the relative performance of different image quality models [86]. The assumption behind such analysis is that the residual difference between the subjective score and its prediction by the objective score

Table 4.1: Performance Evaluation Over Four Databases

LIVE blur database (145 images) [1]

| Measure | PLCC | SRCC | KRCC | RMSE |
|---|---|---|---|---|
| PSNR | 0.8782 | 0.8694 | 0.6920 | 13.898 |
| SSIM [48] | 0.9997 | 0.9990 | 0.9780 | 0.7416 |
| H-metric [12] | 0.7849 | 0.7147 | 0.5347 | 9.7687 |
| Q-metric [82] | 0.6971 | 0.5548 | 0.4056 | 11.302 |
| JNBM [11] | 0.8130 | 0.7821 | 0.6015 | 9.1797 |
| CPBD [10] | 0.9024 | 0.9271 | 0.7714 | 6.7943 |
| $S_3$[9] | **0.9494** | **0.9517** | **0.8157** | **4.9503** |
| BRISQUE [7, 8] | training images | | | |
| BLIINDS-II [5, 6] | training images | | | |
| LPC-SI | **0.9219** | **0.9501** | **0.7994** | **6.1092** |

TID2008 blur database (100 images) [2]

| Measure | PLCC | SRCC | KRCC | RMSE |
|---|---|---|---|---|
| PSNR | 0.8729 | 0.8702 | 0.7327 | 0.5726 |
| SSIM [48] | 0.9460 | 0.9550 | 0.8147 | 0.3803 |
| H-metric [12] | 0.5144 | 0.5106 | 0.3182 | 1.0063 |
| Q-metric [82] | 0.3074 | 0.3290 | 0.2208 | 1.1167 |
| JNBM [11] | 0.6931 | 0.6681 | 0.4947 | 0.8459 |
| CPBD [10] | 0.8237 | 0.8418 | **0.6297** | 0.6655 |
| $S_3$ [9] | **0.8541** | **0.8418** | 0.6124 | **0.6103** |
| BRISQUE [7, 8] | 0.8046 | 0.7989 | 0.6229 | 0.6968 |
| BLIINDS-II [5, 6] | 0.8260 | 0.8205 | 0.6245 | 0.6614 |
| LPC-SI | **0.8455** | **0.8431** | **0.6249** | **0.6267** |

CSIQ blur database (150 images) [3]

| Measure | PLCC | SRCC | KRCC | RMSE |
|---|---|---|---|---|
| PSNR | 0.9252 | 0.9287 | 0.7539 | 0.1087 |
| SSIM [48] | 0.9472 | 0.9605 | 0.8246 | 0.0919 |
| H-metric [12] | 0.8355 | 0.7997 | 0.6274 | 0.1575 |
| Q-metric [82] | 0.7237 | 0.6528 | 0.4860 | 0.1978 |
| JNBM [11] | 0.8061 | 0.7624 | 0.5971 | 0.1696 |
| CPBD [10] | 0.8822 | 0.8790 | 0.6905 | 0.1349 |
| $S_3$ [9] | **0.9106** | **0.9058** | **0.7290** | **0.1184** |
| BRISQUE [7, 8] | **0.9279** | **0.9032** | **0.7353** | **0.1069** |
| BLIINDS-II [5, 6] | 0.8930 | 0.8765 | 0.6783 | 0.1290 |
| LPC-SI | 0.9061 | 0.8931 | 0.7022 | 0.1212 |

IVC blur database (20 images) [4]

| Measure | PLCC | SRCC | KRCC | RMSE |
|---|---|---|---|---|
| PSNR | 0.8883 | 0.8105 | 0.6632 | 0.5243 |
| SSIM [48] | 0.9463 | 0.9353 | 0.7789 | 0.3690 |
| H-metric [12] | **0.9423** | **0.9263** | 0.7684 | **0.3822** |
| Q-metric [82] | 0.9375 | **0.9338** | **0.7789** | 0.3972 |
| JNBM [11] | 0.6983 | 0.6737 | 0.4947 | 0.8172 |
| CPBD [10] | 0.8012 | 0.7744 | 0.6105 | 0.6832 |
| $S_3$ [9] | 0.9274 | 0.8691 | 0.7090 | 0.4269 |
| BRISQUE [7, 8] | 0.8300 | 0.8239 | 0.6561 | 0.6367 |
| BLIINDS-II [5, 6] | 0.7806 | 0.5262 | 0.3979 | 0.7136 |
| LPC-SI | **0.9574** | 0.9202 | **0.7831** | **0.3295** |

Table 4.2: Direct and Weighted Average Performance Over Four Databases

| Direct Average | | | |
|---|---|---|---|
| Measure | PLCC | SRCC | KRCC |
| PSNR | 0.8911 | 0.8697 | 0.7105 |
| SSIM [48] | 0.9598 | 0.9625 | 0.8491 |
| H-metric [12] | 0.7693 | 0.7378 | 0.5622 |
| Q-metric [82] | 0.6664 | 0.6176 | 0.4728 |
| JNBM [11] | 0.7526 | 0.7216 | 0.5470 |
| CPBD [10] | 0.8524 | 0.8556 | 0.6755 |
| $S_3$ [9] | **0.9103** | **0.8921** | **0.7165** |
| BRISQUE [7, 8] | 0.8541 | 0.8420 | 0.6714 |
| BLIINDS-II [5, 6] | 0.8332 | 0.7411 | 0.5669 |
| LPC-SI | **0.9077** | **0.9016** | **0.7274** |

| Database Size-Weighted Average | | | |
|---|---|---|---|
| Measure | PLCC | SRCC | KRCC |
| PSNR | 0.8944 | 0.8882 | 0.7228 |
| SSIM [48] | 0.9652 | 0.9714 | 0.8736 |
| H-index [12] | 0.7456 | 0.7064 | 0.5273 |
| Q-index [82] | 0.6244 | 0.5541 | 0.4081 |
| JNBM [11] | 0.7761 | 0.7423 | 0.5690 |
| CPBD [10] | 0.8713 | 0.8818 | 0.7003 |
| $S_3$ [9] | **0.9113** | **0.9046** | **0.7302** |
| BRISQUE [7, 8] | 0.8749 | 0.8463 | 0.6878 |
| BLIINDS-II [5, 6] | 0.8598 | 0.8298 | 0.6376 |
| LPC-SI | **0.8995** | **0.9023** | **0.7214** |

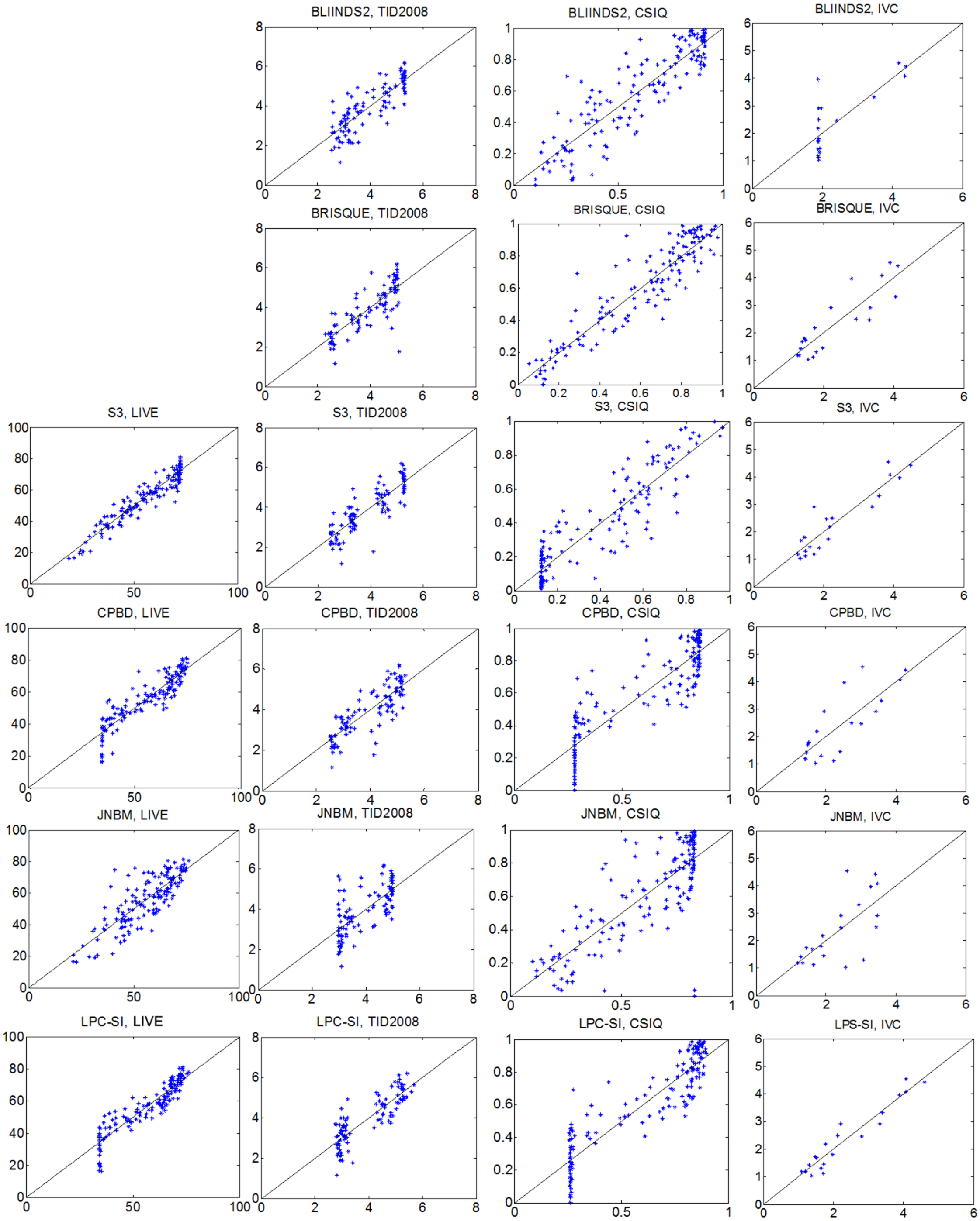


Figure 4.3: Scatter plots between subjective and objective scores (after nonlinear mapping) of six blur metrics over four blur image databases. Top to bottom rows: BLIINDS-II [5, 6], BRISQUE [7, 8], $S_3$ [9], CPBD [10], JNBM [11],and the proposed LPC-SI measure; Left to right columns: LIVE, TID2008, CSIQ and IVC blur databases.

Table 4.3: Statistical Significance Analysis Based On Quality Prediction Residuals On LIVE [1], TID2008 [2], CSIQ [3] and IVC [4] Databases

| model | PSNR | SSIM | H-Metric | Q-Metric | JNBM | CPBD | LPC-SI | $S_3$ | BRISQUE | BLIINDS-II |
|---|---|---|---|---|---|---|---|---|---|---|
| PSNR | ---- | 000- | -10- | -10- | 010- | 0-0- | 0-00 | 0-0- | x--- | x-0- |
| SSIM [48] | 111- | ---- | 111- | 111- | 1111 | 11-1 | 11-- | 11-- | x111 | x111 |
| H-Metric [12] | -01- | 000- | ---- | 1-0- | 0--1 | 0001 | 000- | 000- | x011 | x011 |
| Q-Metric [82] | -01- | 000- | 0-1- | ---- | 00-1 | 0001 | 000- | 000- | x011 | x011 |
| JNBM [11] | 101- | 0000 | 1--0 | 11-0 | ---- | 000- | 0000 | 000- | x011 | x011 |
| CPBD [10] | 1-1- | 00-0 | 1110 | 1110 | 111- | ---- | 0--0 | ---0 | x-1- | x-1- |
| LPC-SI | 1-11 | 00-- | 111- | 111- | 1111 | 1--1 | ---- | ---- | x-11 | x-11 |
| $S_3$ [9] | 1-1- | 00-- | 111- | 111- | 1111 | ---1 | ---- | ---- | x-1- | x-11 |
| BRISQUE [7, 8] | x--- | x000 | x100 | x100 | x-0- | x-0- | x-00 | x-0- | x--- | x-0- |
| BLIINDS-II [5, 6] | x-1- | x000 | x100 | x100 | x10- | x-0- | x-00 | x-00 | x-1- | x--- |

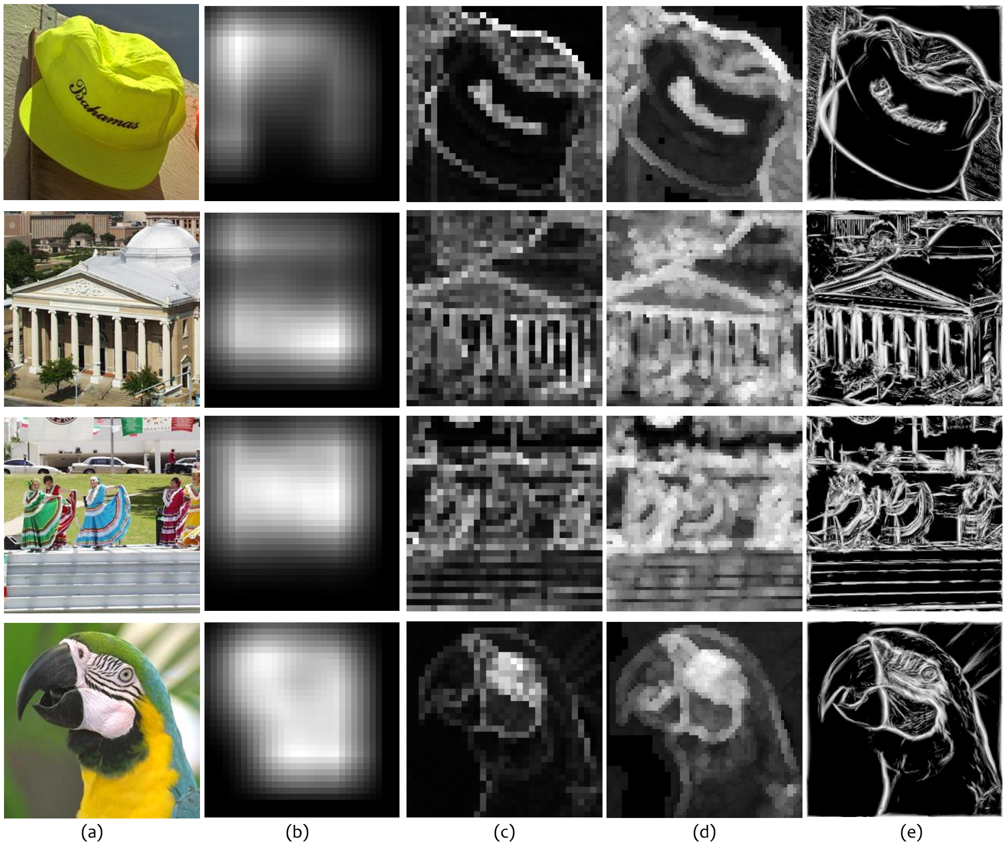


Figure 4.4: (a) Sample images and their corresponding (b) CPBD [10], (c) H-metric [12], (d) $S_3$ [9], and (e) LPC-SI sharpness maps, where brighter indicates higher sharpness.

is Gaussian distributed. In reality, this assumption is not always met perfectly, but is somewhat reasonable because with the large number of sample points, the Central Limit Theorem comes into play and the distribution of the residual difference approximates Gaussian. For a given image database, F-statistic can then be employed to compare the variances of two sets of prediction residuals by two objective methods, so as to determine whether the two sample sets come from the same distribution. As such, we can make a statistically sound judgment regarding superiority or inferiority of one objective method against another. A statistical significance analysis matrix is created and shown in Table 4.3, where each entry consists of four characters which correspond to the four blur databases in the order of LIVE, TID2008, CSIQ and IVC, respectively. A symbol "-" denotes that the two objective methods are statistically indistinguishable, "1" denotes the method of the row is statistically better than that of the column, and "0" denotes that the method of the column is better than that of the row. A symbol "x" denotes unfeasible analysis between row and column method. This is mainly in the case of BRISQUE and BLIINDS-II algorithms over trained data from the LIVE database.

It can be observed that $S_3$ and LPC-SI are statistically indistinguishable for all databases and outperform all other NR sharpness methods for all databases.

One useful feature of the proposed LPC-SI approach is that it provides an LPC map that indicates the spatial variations of local sharpness. Sample images are given in Fig. 4.4, together with their corresponding sharpness maps produced by CPBD [10], H-Metric [12], $S_3$ [9] and LPC-SI algorithms. Since the CPBD algorithm requires a large block of size $64 \times 64$ [10], to produce more localized measurement, overlapping $64 \times 64$ windows of step size 8 are used, resulting in blurry sharpness maps, as shown in Fig. 4.4. The H-metric map is computed by non-overlapping $8 \times 8$ blocks, and thus creates sharpness maps that have the same resolution but are less burry. $S_3$ employs different block sizes and overlapping factors for spectral and spatial measurement, and in general further improves the resolution of the sharpness maps. Unlike the other algorithms, LPC-SI does not employ block-based computation, and produces the most localized sharpness map. Visual comparisons of the sample images together with the sharpness maps in Fig. 4.4 suggest that all four types of sharpness maps convey useful information regarding local sharpness variations, and the H-Metric, $S_3$ and LPC-SI, to a significant extent, agree with each other, though the localization of the measurement could be very different. It is worth mentioning that these sharpness maps are computed to compare the local behavior of the competing algorithms only, some of which may not be designed to generate accurate local sharpness measurement but to predict the overall human sharpness assessment of the whole image. Moreover, without precise local sharpness assessment by human subjects (which will be our future work), it is difficult to conclude about the relative performance of these maps.

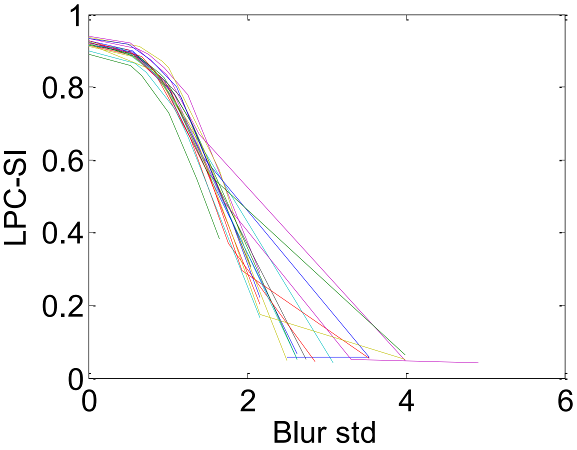


Figure 4.5: Relationship between LPC-SI and Gaussian blur standard deviation.

In Fig. 4.5, we examine the monotonicity between LPC-SI and the numerical level of blur, where Gaussian blur with standard deviation between $\sigma = 0$ and $\sigma = 4$ are tested using source images in the LIVE database. It can be seen that

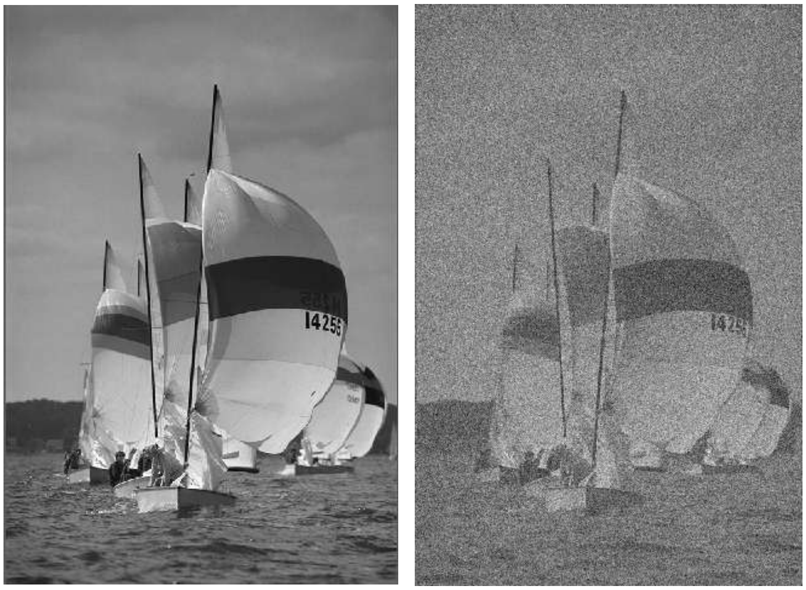


Figure 4.6: Example of original and noise corrupted images.

LPC-SI has a monotonic behavior and therefore can successfully rank-order the source image with different blur parameters. On the other hand, the sharpness measures behave differently when blurred images generated from different source images are compared. Similar behavior has also been observed when the same test is applied to the other NR measures in Tables 4.1 and 4.2. It is worth noting that the goal of a perceptual sharpness measure is to predict the perceived blur, but not the numerical factor used to create the blurred image. The perceptual blur and the numerical blur may not perfectly agree with each other, especially when the images are significantly different in texture content and complexity.

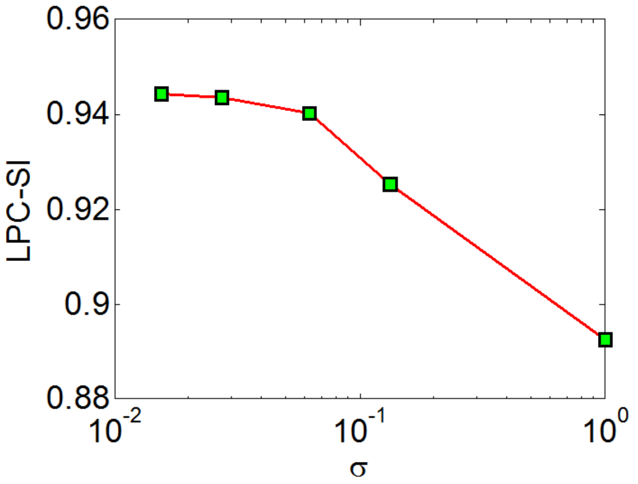


Figure 4.7: Relationship between LPC-SI and noise standard deviation $\sigma$.

Table 4.4: Runtime Comparison of Sharpness Measures For Images of 1024 × 1024 Resolution

| model | runtime (second) |
|---|---|
| H-metric [12] | 3.86 |
| JNBM [11] | 8.37 |
| CPBD [10] | 11.29 |
| $S_3$ [9] | 142.5 |
| BRISQUE [7, 8] | 1.03 |
| BLIINDS-II [5, 6] | 572.0 |
| LPC-SI | 4.37 |

Although our algorithm is not designed to work with noise, it is interesting to observe how it reacts when images are contaminated with noise. The impact of noise on perceived sharpness is a complicated issue. Adding noise may (or may not) increase the visual sharpness of flat image regions, but it could significantly affect the perceived sharpness at the sharp edges near an object, which in turn may make it more difficult for the visual system to discern detailed structures in an image. Figure 4.6 demonstrates this by showing the "Sail Boats" image that is severely corrupted by noise. In this example, the sharpness of the main objects appear to be weakened by noise. To examine how the proposed method reacts to noise, we plot LPC-SI versus noise level for the "Sail Boats" image in Fig. 4.7, where we observe that LPC-SI decreases with the increase of noise. We hypothesize that how noise affects perceptual sharpness is content-dependent. It is currently not a conclusive issue and is worth future investigations.

The majority of the computational cost of the proposed LPC-SI algorithm is in the log-Gabor filtering process, which is implemented using the fast Fourier transform (FFT) approach. In the case of $L$-scale $M$-orientation filtering, a total of one forward 2-dimensional (2D) FFT and $LM$ inverse 2D-FFTs are performed, and all other computations are linear with respect to the number of pixels $N$ in the image. Therefore, the computational complexity of the proposed algorithm is determined by the complexity of 2D-FFT and is thus in the order of $O(N \log N)$. We have also carried out further experiment to compare the runtime of seven sharpness measures applied on images with 1024 × 1024 resolution. This test was performed using a computer configured with Intel Core i7 CPU at 2.40 GHz, 8GB RAM, Windows 7 64-bit, and Matalab 7.10. The results are summarized in Table 4.4, which gives a rough estimate of the relative complexity of the algorithms because the Matlab code is not optimized for speed. The BRISQUE algorithm requires a long training process but is the fastest in the testing phase among all algorithms being compared. The slowest methods are BLIINDS2 and $S_3$, both of which involve sophisticated block-based computation. LPC-SI achieves highly competitive perceptual sharpness prediction and is among the fastest algorithms, giving it advantages in real-world applications. Figure 4.8 depicts the runtime comparison with increasing image size. LPC-SI is significantly faster than the most competitive no-sharpness assessment

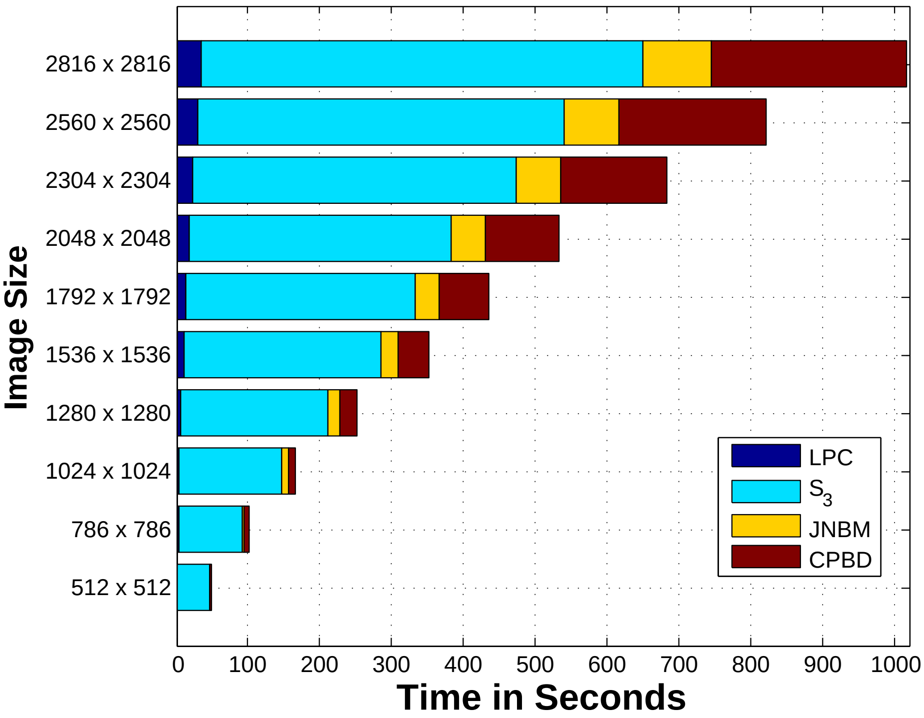


Figure 4.8: Computational Runtime Comparison between LPC-SI and Existing NR-Sharpness Measures with Image Size.

measures namely $S_3$, JNBM, and CPBD.

## 4.5 Conclusions

In this chapter, we presented an LPC-based method for the assessment of perceived image sharpness without referencing the original image. The underlying assumption is that blur affects the LPC relationship near sharp image features and the degradation of LPC strength can be employed to measure image sharpness. We develop a novel LPC-SI image sharpness measure based on the computational framework described in previous chapter, which shows competitive performance when compared with state-of-the-art algorithms. These promising results inspire us to extend the current work to other image processing applications where LPC may be employed as a new type of prior model of sharp natural images.

# Chapter 5

# Image Fusion

This chapter presents two important applications of local phase coherence measurement in image fusion problems. First, the objective quality assessment of image fusion is introduced. An overview of existing quality measures and the challenges that need to be addressed are presented. Followed by in details description of the design of the proposed fusion quality measure. Second, an simple novel image fusion algorithm is presented. The fusion algorithm development and results are subsequently described.

## 5.1 Blind Objective Quality Assessment of Image Fusion

The growing interest in image and video fusion to serve many military, surveillance, and medical applications contributed to the development of a number of objective image fusion performance metric. In these applications, it is paramount to understand the quality of an image or video sequence. In general, image fusion algorithms attempt to improve the image quality by combining features from multiple image sources into a single image for either human consumption or as an introductory stage to another task. Therefore, the performance of fusion algorithm must be justified in terms of improvement in the following tasks. Traditionally, the assessment of a fusion scheme is carried out by subjective evaluation, which is known to be slow and expensive, and most importantly cannot be embedded into automated frameworks for parameters optimization and algorithm improvement. A valuable alternative to subjective evaluation is objective image fusion quality measures that are consistent and well-correlate with human visual perception.

Various fusion algorithms presented in the literature [87] construct an “ideal” fused image and objectively evaluate the algorithm’s performance by using this image as a reference (or ground truth) for comparison with the experimental results [88, 89]. The assumption of the existence of ground truth is not always valid because the reference image is not available in many practical and real-world applications.

In this section, we present a novel blind objective fusion index inspired in part by LPC measurement and in another part by the structural similarity approach [48].

### 5.1.1 Previous Work

Some newly developed quality measures examine the information preserved in the fused image using various techniques. Qu et al. [90] developed an information-theory related measure. In his measure, the summation of the mutual information (MI) between the fused image and two input images is obtained to represent the difference in quality. Xydeas and Petrovic [91] proposed an objective edge based performance measure which evaluates the relative amount of edge information that is transferred from the input images to the fused image. Similarly, Wang and Liu [92] proposed a measure based on a multi-scale scheme implemented with a two-level Haar wavelet. The edge information in this measure is retrieved from the high and band-pass components of the wavelet decomposition. Recently, objective image fusion performance measures based on the structural similarity measure [48] have come into being. For example, Piella and Heijmans [93] proposed a new quality measure for image fusion based on the SSIM measure. The measure computes the SSIM maps between each input image and the fused image and uses it as a measurement of the loss of structural information. SSIM maps have been weighted by local salience measure. The weighting factor depends on the similarity in spatial domain between input images. Likewise, Yang et al. [94] proposed another way to use SSIM for fusion assessment. In an attempt to take into consideration human visual perception, Chen and Varshney [95] proposed an image fusion quality measure based on the HVS. The method employs the Contrast Sensitivity Function (CSF) on the entire image and then considers the local spatial information transfer by calculating the summation of squares of edge intensities on a region-by-region basis. The reader is suggested to read [96] for more details and comparisons of image fusion measures.

### 5.1.2 Quality Assessment Method

In photography, when one scene contains objects in different distance or when a camera has a limited focal length or crop factor, the camera Depth-of-Field (DOF) will be affected, causing some objects in the scene to be out-of-focus. In addition, in a scene with low or strong lighting conditions, there will be an out-of-range physical amount of light from the subject allowed to fall on the film, which cause significant information loss during capture, recorded as “black” (underexposed) or “white” (overexposed) values. These two common types of problem in photography can be solved with image fusion techniques. There is evidence that humans exploit different features in an image for different tasks [97]. In fact, perception studies indicate that the best fusion strategy is task dependent [98]. Therefore, we expect that a fusion quality measure should be task specific, and the “best” fusion algorithm

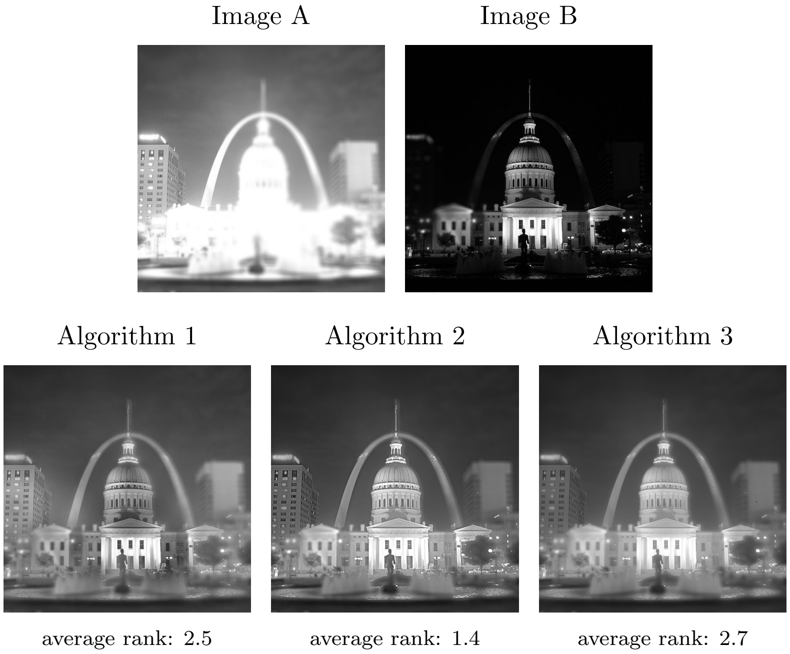


Image A

Image B

Algorithm 1

Algorithm 2

Algorithm 3

average rank: 2.5

average rank: 1.4

average rank: 2.7

Figure 5.1: Source images (top) and subject-rank fusion results (bottom).

changes from task to task. In case of multi-exposure and multi-focus fusion, the task is to fuse input images with different focal points and/or different exposure values to get one fused image with sharp features and good exposure compared to input images. A large research body in image fusion literature proposed fusion algorithms to solve these two common problems [99, 100, 101, 102, 103]. To this end, the effectiveness and efficiency of fusion algorithms need to be verified with actual human performance. Figure 5.1 shows two input images with different exposure levels and complementary locations of out-of-focus objects in the scene. Three different fusion algorithms have been used to fuse these images. The slowness and tediousness of subjective quality assessment leads to a growing needs to the development of automatic objective assessment approaches for fusion results.

Good fusion quality measures can be obtained through a thorough understanding of the HVS. Much work has investigated and modeled the HVS and it is clear that humans respond to contrast and edges in imagery. However, it is still not clear how to quantify the way humans interpret a particular image/video for a specific task. Assume we have two input images $A$ and $B$ fused to produce $F$. A robust fusion quality measure is required to account for a set of features to assess the quality of $F$ given $A$ and $B$. The question is: How should one go about choosing a set of features to assist the quality of $F$? When applying the concepts of human perception to fusion quality measure development, we found that three characteristics are extremely important, which will be discussed in the following subsections.

### Local Contrast

A simple and effective measure to quantify the local contrast of a given signal is the standard deviation. SSIM [48] employed the local standard deviation of images under comparison in a simple contrast comparison function that is conceptually consistent with the contrast-masking feature of the HVS. A good fusion measure needs to evaluate the similarity of local contrast between the input and the fused images. The contrast of the fused image should be close to the input image with higher contrast. In other words, the fusion measure algorithm should penalize the cases that the fused signal strength is significantly lower than any of the input signal's strength. Let $a$, $b$, and $f$ be local image patches of the input and fused images, respectively. The local contrast comparison function is defined as:

$$c(a,b,f)=\frac{2\sigma_f\max\left(\sigma_a,\sigma_b\right)+C_1}{\sigma_f^2+\max\left(\sigma_a^2,\sigma_b^2\right)+C_1} \tag{5.1}$$

where $\sigma_a$, $\sigma_b$, and $\sigma_f$ are the local standard deviations of local image patches, and $C_1$ is a positive stabilizing constant. The local contrast measure is applied to an image using a sliding window that runs across the image space. This results in a map that is bounded between 0 and 1, where 0 and 1 represent completely insignificant and completely significant signal strengths, respectively. To provide a single score of contrast, the map is pooled by averaging as follow:

$$Q_c=\frac{1}{N}\sum_{i=1}^{N}c(a_i,b_i,f_i) \tag{5.2}$$

where $a_i$,$b_i$ and $f_i$ are the $i$-th patches in the input images and fused image being compared, respectively, and $N$ is the number of patches.

### Sharpness

Sharpness is a very desirable property in a good fusion index. The HVS has a remarkable capability to detect image blur/sharpness without seeing the original image. Our LPC-based sharpness estimator in [104], described in Chapter 4, is incorporated to evaluate the sharpness of a fused image and eventually used as a second feature in the proposed fusion quality index. The proposed sharpness estimator has many attractive properties that nominate it as a good candidate for the current context. Firstly, it is based on local phase information obtained from band-pass filters which resembles neuronś receptive field properties in mammalian visual cortex. Secondly, no reference image is required which is appropriate for practical applications. Thirdly, it correlates well with human subject scores over different databases, and is among the most competitive no reference sharpness estimators in the literature. And finally, the LPC-SI produces a spatially varying localized sharpness map that could potentially used in many applications, as will be described shortly. Motivated by all of the above reasons the sharpness of the fused image is evaluated based on LPC.

The fused image is first passed through a series of $N$-scale $M$-orientation log-Gabor filters without any subsequent down-sampling process. This results in $MN$ complex coefficients $c_{ijk}$ at the $i$-th scale, the $j$-th orientation and the $k$-th spatial location. Then the LPC sharpness map of fused image is computed as:

$$LPCS_{map} = \frac{\sum_{j=1}^{M} |c_{1jk}| \cos\left(\Phi\left\{\prod_{i=1}^{N} c_{ijk}^{w_i}\right\}\right)}{\sum_{j=1}^{M} |c_{1jk}| + C}\,, \tag{5.3}$$

where the magnitude of the finest scale coefficient $c_{1jk}$ is used to give higher importance to orientations that contain more energy, $w_i$ are an optimal set of weights to compute accurate predicted phase compared to true phases of finest scale coefficients (more details in computing these weights are given in Chapter 3), and $C$ is a constant to avoid instability in case of small magnitude coefficients. In order to compute one score for fused image sharpness, $LPCS_{map}$ is pooled using a weighted averaging method based on ranked LPC values by

$$Q_{sh} = \frac{\sum_{k=1}^{K} u_k \, LPCS_{map}}{\sum_{k=1}^{K} u_k}\,, \tag{5.4}$$

where $u_k$ is the weight assigned to the $k$-th ranked spatial LPC value and is computed as an exponentially decaying function, as in Chapter 4. In all the experimental results reported in the next section, the parameters are set as $M = 8$, $N = 3$, $s_1 = 1$, $s_2 = 3/2$, $s_3 = 2$, $w_1 = 1$, $w_2 = -3$, $w_3 = 2$ and $C = 2$, respectively. These parameters are set empirically, but are found to be insensitive to the overall performance.

### Weighted-Structural Preservation

It is important for a fusion quality measure to ensure that structural information content in the input images has been preserved in the fused image. This is the third characteristic modeled in our proposed fusion measure. The SSIM [48] provides a useful design philosophy as well as a simple method for measuring structural fidelity between images. Given input images $A$ and $B$ and the fused image $F$, let $a$, $b$, and $f$ be local image patches of the input and fused images respectively, the structural comparison is conducted between each of the input images and the fusion results as follow

$$\begin{aligned} S_{AF} &= \frac{\sigma_{af} + C_2}{\sigma_a \sigma_f + C_2} \\ S_{BF} &= \frac{\sigma_{bf} + C_2}{\sigma_b \sigma_f + C_2}\,, \end{aligned} \tag{5.5}$$

where $\sigma_a$, $\sigma_b$, and $\sigma_f$ are the local standard deviations of local image patches, $\sigma_{af}$ and $\sigma_{bf}$are cross correlation between the two corresponding patches, respectively, and $C_2 = C_1/2$ as suggested in [48]. $S_{AF}$ and $S_{BF}$ are spatially varying quality maps which deliver information about how similar each input image is to the final fused image. Because multi-focus input images contain complementary sharp regions, the fusion index algorithm should be able to make the amount of sharp features transferred from input to fused signals. Since the LPC map indicates the location of sharp features in each input image, it is plausible to weight the structure similarity map with the LPC map. This will assign more weights to sharp features transfer from each input image to the fusion results. The impact of the blurred or less similar feature decreased accordingly. In doing so, we are able to measure more accurately the fusion performance, especially in an experimental setup where the input images are multi-focus variants of the same scene. The weighted-structural preservation constraint is defined as:

$$Q_{ws} = \frac{1}{2}\left\{ \frac{\sum_{k=1}^{K} LPCS_{map}^{\{k\}}(A) S_{AF}^{\{k\}}}{\sum_{k=1}^{K} LPCS_{map}^{\{k\}}(A)} + \frac{\sum_{k=1}^{K} LPCS_{map}^{\{k\}}(B) S_{BF}^{\{k\}}}{\sum_{k=1}^{K} LPCS_{map}^{\{k\}}(B)} \right\} , \tag{5.6}$$

where $LPCS_{map}(A)$ and $LPCS_{map}(B)$ are the LPC map defined in Eq. 5.3 for both input images $A$ and $B$, respectively. Similarly, $S_{AF}$ and $S_{BF}$ are the structural similarity maps between each input image and the fusion results. Both maps are bounded between 0 and 1, hence $Q_{ws} = 0$ represents blur dissimilar structure and $Q_{ws} = 1$ represents sharp similar structure.

#### Fusion Quality Assessment Model

The local contarst measure $Q_c$ in Eq. 5.2, the sharpness estimator $Q_{sh}$ in Eq. 5.4, and the weighted-structural preservation measure $Q_{ws}$ in Eq. 5.6 are the three building blocks of our fusion quality measure. The overall Fusion Quality Index (FQI) puts these features together as

$$FQI = [Q_c]^{\alpha} \cdot [Q_{sh}]^{\beta} \cdot [Q_{ws}]^{\gamma} , \tag{5.7}$$

where $\alpha$, $\beta$, and $\gamma$ are parameters used to adjust the relative importance of the three components. In the current implementation we set $\alpha = \beta = \gamma = 1$ for simplicity. We find that the overall performance of the fusion index is insensitive to these parameters within an order of magnitude.

### 5.1.3 Validation

Numerous algorithms have been developed for multi-exposure and/or multi-focus image fusion and evaluated on some online individual images. As an initiative, we build a database of simultaneous multi-exposure multi-focus images. The database contains 14 pair of images. Each pair has an overexposed and underexposed image which has been corrupted by simulated out-of-focus at complementary locations. The out-of-focus effect has been simulated using foveated filters with different fixation points. Examples of images are given in Fig. 5.2.

Figure 5.2: Sample images with multi-exposure and multi-focus effects.

Table 5.1: Fusion Rules For Multi-resolution Fusion Algorithms

| Fusion Rules | LAP | COP | RAP | DWT |
|---|---|---|---|---|
| Base Images | CM | CMCV | WA | CM |
| HF Images | AVG | MIN | AVG | MIN |

Six fusion algorithm were considered in this experiment with various rules to extract and combine salient features in input images, including (1) pixel average (AVG), (2)Laplacian pyramid (LAP), (3) contrast pyramid (COP), (4) ratio pyramid (RAP), (5) discrete wavelet transform (DWT), and recently proposed (6) generalized random walk (GRW) [100]. Rockingers MATLAB toolbox [105] is used as the reference for the implementation of the multiresolution methods (2), (3), and (4). Daubechies 2 wavelet is used to fuse approximation and detail coefficients at each scale before reconstruction. Typical fusion schemes aim to locally select the most salient image features. Four most popular method were used to select high frequency (HF) coefficients with the largest activity level at each pixel location in multiresolution techniques include (1) choose maximum (CM) absolute coefficient, (2) weighted average (WA), and (3) choose maximum with consistency verification (CMCV). Average (AVG) and minimum (MIN) between base images are alternatively employed to combine low frequency bands of input images. Table 5.1 summarize the rules used in each algorithm.

A well-adopted method to validate a fusion measure is comparing objective quality assessment results with subjective data. To do so, we conducted a subjective test using the image database developed by ourselves. Eighteen subjects were provided with 14 sets of fused images, each of which includes 6 images fused by the fusion algorithm mentioned above. The positions of the fused images were placed in random order in the test to avoid any positional selection bias. The two input images for each set were displayed at the same location on the screen as a reference point for the observer. The test environment was kept constant during the tests. The observers viewed the images on a 42 inch HD LCD monitor. Before taking the test, the observers were briefly instructed with an example of images. Then the observers were asked to give an opinion score between 1 and 6, where “1” stands for best perceptual quality, and “6” denotes the worst perceptual quality. The observers performed the tests individually with no time limit. The rank scores were recorded anonymously on a form given to each subject. The subjective rankings for each image is then averaged, resulting in its mean ranking score within the set.

To efficiently evaluate the performance of the proposed fusion measure, we employed two correlation. Correlation coefficients can range from $+1$ to $-1$, which represent a perfect agreement or disagreement between objective and subjective ranks.

- Spearmans rank-order correlation coefficient (SRCC) is defined as

$$SRCC = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N^2-1)}\,, \tag{5.8}$$

Table 5.2: Performance Evaluation Using 14 Image Sets And 6 Fusion Algorithms

| Image Set | KRCC | | | SRCC | | |
|---|---|---|---|---|---|---|
| | $\mu$ sbj perfrmnc | $\sigma$ sbj perfrmnc | FQI perfrmnc | $\mu$ sbj perfrmnc | $\sigma$ sbj perfrmnc | FQI perfrmnc |
| 1 | 0.7961 | 0.2455 | 0.8667 | 0.8689 | 0.2107 | 0.9429 |
| 2 | 0.8039 | 0.1499 | 0.8667 | 0.8789 | 0.1276 | 0.9429 |
| 3 | 0.7098 | 0.2460 | 0.8667 | 0.7882 | 0.2422 | 0.9429 |
| 4 | 0.8509 | 0.1993 | 0.8667 | 0.9126 | 0.1486 | 0.9429 |
| 5 | 0.8482 | 0.1301 | 0.8667 | 0.9216 | 0.0837 | 0.9429 |
| 6 | 0.8051 | 0.1103 | 0.9661 | 0.8908 | 0.0880 | 0.9856 |
| 7 | 0.8849 | 0.0983 | 0.8281 | 0.9396 | 0.0632 | 0.9276 |
| 8 | 0.7302 | 0.1992 | 0.8667 | 0.8289 | 0.1729 | 0.9429 |
| 9 | 0.8039 | 0.1771 | 0.8667 | 0.8789 | 0.1383 | 0.9429 |
| 10 | 0.7991 | 0.1849 | 0.7333 | 0.8703 | 0.1538 | 0.8286 |
| 11 | 0.8353 | 0.1108 | 1.0000 | 0.9126 | 0.0759 | 1.0000 |
| 12 | 0.6941 | 0.2615 | 1.0000 | 0.7882 | 0.2328 | 1.0000 |
| 13 | 0.7725 | 0.1811 | 1.0000 | 0.8622 | 0.1579 | 1.0000 |
| 14 | 0.7647 | 0.1455 | 0.8667 | 0.8521 | 0.1144 | 0.9429 |
| **Average** | **0.7928** | **0.1743** | **0.8901** | **0.8710** | **0.1436** | **0.9489** |

where $d_i$ is the difference between the i-th images ranks in subjective and objective evaluations.

- Kendalls rank-order correlation coefficient (KRCC) is defined as:

$$KRCC = \frac{2(N_c - N_d)}{N(N-1)}, \tag{5.9}$$

where $N_c$ and $N_d$ are the numbers of concordant and discordant pairs in the data set, respectively.

To evaluate the proposed FQI, we calculate the KRCC and SRCC values between the mean ranking scores and the objective quality measures for each image set. The results are given in Table 5.2. To provide a reference point in evaluating the performance of FQI, we compare it with the behavior of an average subject. To do this, we first compute the KRCC and SRCC values between the mean ranking scores and the ranking scores given by each individual subject for each image set. We then compute the mean and standard deviation of these KRCC and SRCC values over subjects, which are shown in Table 5.2. The average KRCC and SRCC values over all 14 image sets are given in the last row. It can be seen that for all image sets, the KRCC and SRCC values of FQI are well within the range of ±1 standard deviation from the KRCC and SRCC values of the mean over all subjects. This indicates that FQI behaves quite similarly to an average subject.

To further investigate the performance of the proposed quality measure, we compared it against the state-of-the-art fusion measures. In this experiment, seven objective fusion measures have been considered. The parameters used in all fusion measures are set as their default values. The fusion measures are as follow:

1. *Normalized Mutual Information ($Q_{MI}$) [90]*: MI for fusion performance measure is the summation of the MI between fused image $F$ and the two input images $A$ and $B$. This measures the amount of information transferred from the source images to the fused output image. The larger the MI value, the better the fused result is.

2. *Piellas Measure($Q_S$)* [93]: The fusion measure is defined based on SSIM [48]. First it computes the structural similarity between fused image $F$ and the two input images $A$ and $B$. Then a local measure of image salience used to indicate the relative importance of image $A$ compared to image $B$. The larger the $Q_S$ values, the better the fused results are.

3. *Xydeas's Measure($Q_G$)* [91]: The amount of edge information which is transferred from input images to the fused image is calculated using Sobel edge operator. The edge strength and orientation of each input image and fused image is firstly computed. A relative edge measure between input and fused image is obtained and employed to drive edge strength and orientation preservation values. The final assessment is obtained from the weighted average of the edge information preservation values. The higher the value of $Q_G$, the better the quality.

4. *Yang's Measure ($Q_Y$)* [94]: This measure again uses SSIM to evaluate the similarity between input images and fused image. It also computes SSIM between the input images and employed it as a match measure based on hard threshold. If the match measure is equal to or larger than a given threshold, then the weighted average between $SSIM(A, F)$ and $SSIM(B, F)$ is taken as the local quality, otherwise, the larger one of the two is taken. Just as the above fusion measures based on structural similarity, the larger the $Q_Y$ value, the better the fusion results.

5. *Chen-Varshney's Measure ($Q_{CV}$)* [95] : First CSF is applied to the difference between each input image and the fused image in the frequency domain. Then an edge strength map is created for each input image as well as the fused image using Sobel edge detector and divided into non-overlapped windows. A measure of salience is calculated using the summation of squares of edge intensities in each local window. Finally the quality of fused image is the weighted summation over all of the non-overlapping windows. The lower the value of $Q_{CV}$, the better the fusion results.

6. *Wang's Measure ($Q_M$)* [92] : Multiscale scheme is used to decompose input and fused images using a two-level Haar Wavelet. At each level, an edge preservation function is calculated for the high frequency coefficients in the horizontal, vertical and diagonal directions. Then the edge preservation values are weighted by the energy at each location. Finally the overall quality is computed by combining the measurements at each scale level. The higher the value of $Q_M$, the better the quality.

7. *Chen-Blum's Measure ($Q_{CB}$)* [106] : This measure is based on HVS model. The input and fused images are first filtered by a CSF, after which a local contrast map is computed for each image. Second, a contrast preservation map is generated to describe the relationship between the fused image and each input image. Finally, the preservation maps are weighted by a salience map to obtain an overall quality map. The mean of the quality map indicates the quality of the fused image.

Kendell and Spearman correlations between subjective score and objective score of all fusion measures under study and for all images in the database are given in Table 5.3. The average performance of fusion measures across images are listed on the last column. We highlight the top three results obtained by fusion measures with boldface. We observe that $Q_s$ and $Q_g$ show a good correlation with human score compared to $Q_{MI}$, $Q_{CV}$. Surprisingly $Q_y$ gives very low correlation even though it is based on structural similarity features like $Q_s$. It appears that $FQI$ outperforms all existing fusion measures.

To better understand the similarities between the fusion measures under study, a visual representation of correlations between fusion measures is created using dendrogram plot in Fig. 5.3. Kendell correlation coefficients between the eight fusion measures are firstly obtained and then averaged across images. The dendrogram transforms this correlation coefficients into distances and clusters the coefficients using the unweighted pair group method with arithmetic mean (UPGMA) algorithm. Euclidean distance is used as a distance measure for data clustering. The horizontal axis represents the fusion measures along with the MOS values referred to as leaf nodes. The vertical axis is labeled distance and refers to a distance measure between nodes. The height of the node can be thought of as the distance value between the right and left sub-branch clusters. The fusion measures are clustered based on similarity of performance. The dendogram shows how close the proposed fusion measure is to human scores. $Q_s$ and $Q_g$ are both close to $FQI$ which agree with the previous results in Table 5.3.

Table 5.3: Comparison Between Seven Objective Fusion Quality Measures Over 14 Image Dataset

| | KRCC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Image Set | $Q_{MI}$ | $Q_S$ | $Q_G$ | $Q_Y$ | $Q_{CV}$ | $Q_M$ | $Q_{CB}$ | FQI |
| 1 | 0.0667 | 0.7333 | 1.0000 | 0.6000 | 0.0667 | 0.8667 | 0.2000 | 0.8667 |
| 2 | 0.0667 | 0.4667 | 0.4667 | 0.2000 | 0.0667 | 0.3333 | 0.6000 | 0.8667 |
| 3 | 0.2000 | 0.8667 | 0.7333 | 0.0667 | 0.3333 | 0.6000 | 0.0667 | 0.8667 |
| 4 | 0.0667 | 0.7333 | 0.6000 | 0.0667 | 0.6000 | 0.4667 | 0.2000 | 0.8667 |
| 5 | 0.0667 | 0.8667 | 0.6000 | 0.0667 | 0.6000 | 0.6000 | 0.4667 | 0.8667 |
| 6 | 0.2760 | 0.8281 | 0.6901 | 0.1380 | 0.9661 | 0.2760 | 0.4140 | 0.9661 |
| 7 | 0.0000 | 0.6901 | 0.5521 | 0.1380 | 0.1380 | 0.4140 | 0.4140 | 0.8281 |
| 8 | 0.0667 | 0.6000 | 0.6000 | 0.0667 | 0.0667 | 0.6000 | 0.0667 | 0.8667 |
| 9 | 0.0667 | 0.2000 | 0.7333 | 0.0667 | 0.6000 | 0.3333 | 0.0667 | 0.8667 |
| 10 | 0.2000 | 0.7333 | 0.8667 | 0.7333 | 0.4667 | 0.7333 | 0.0667 | 0.7333 |
| 11 | 0.0667 | 0.8667 | 0.7333 | 0.4667 | 0.3333 | 0.7333 | 0.3333 | 1.0000 |
| 12 | 0.0667 | 1.0000 | 0.7333 | 0.0667 | 0.7333 | 0.3333 | 0.3333 | 1.0000 |
| 13 | 0.0667 | 0.7333 | 0.8667 | 0.2000 | 0.7333 | 0.4667 | 0.0667 | 1.0000 |
| 14 | 0.0667 | 0.8667 | 0.7333 | 0.4667 | 0.0667 | 0.7333 | 0.3333 | 0.8667 |
| Average | 0.0959 | **0.7275** | **0.7078** | 0.2388 | 0.4122 | 0.5350 | 0.2591 | **0.8901** |

| | SRCC | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Image Set | $Q_{MI}$ | $Q_S$ | $Q_G$ | $Q_Y$ | $Q_{CV}$ | $Q_M$ | $Q_{CB}$ | FQI |
| 1 | 0.0286 | 0.8857 | 1.0000 | 0.7714 | 0.1429 | 0.9429 | 0.0857 | 0.9429 |
| 2 | 0.0857 | 0.6000 | 0.6000 | 0.2571 | 0.0857 | 0.1429 | 0.6571 | 0.9429 |
| 3 | 0.3143 | 0.9429 | 0.8857 | 0.1429 | 0.5429 | 0.7143 | 0.0857 | 0.9429 |
| 4 | 0.2000 | 0.8857 | 0.6571 | 0.0857 | 0.7714 | 0.5429 | 0.2571 | 0.9429 |
| 5 | 0.0857 | 0.9429 | 0.6571 | 0.0286 | 0.7714 | 0.7714 | 0.6000 | 0.9429 |
| 6 | 0.2609 | 0.8986 | 0.8117 | 0.1739 | 0.9856 | 0.4638 | 0.5218 | 0.9856 |
| 7 | 0.0579 | 0.7827 | 0.6377 | 0.2319 | 0.4058 | 0.4638 | 0.5508 | 0.9276 |
| 8 | 0.1429 | 0.7714 | 0.7714 | 0.0286 | 0.0857 | 0.6571 | 0.1429 | 0.9429 |
| 9 | 0.0286 | 0.4286 | 0.8857 | 0.0857 | 0.7714 | 0.4857 | 0.0857 | 0.9429 |
| 10 | 0.3143 | 0.8286 | 0.9429 | 0.8286 | 0.6000 | 0.8286 | 0.0286 | 0.8286 |
| 11 | 0.0286 | 0.9429 | 0.8286 | 0.6000 | 0.4857 | 0.8286 | 0.4857 | 1.0000 |
| 12 | 0.0286 | 1.0000 | 0.8286 | 0.0286 | 0.8857 | 0.4857 | 0.4857 | 1.0000 |
| 13 | 0.0857 | 0.8286 | 0.9429 | 0.2571 | 0.8857 | 0.6000 | 0.0286 | 1.0000 |
| 14 | 0.0286 | 0.9429 | 0.8286 | 0.6000 | 0.0857 | 0.8286 | 0.4857 | 0.9429 |
| Average | 0.1207 | **0.8344** | **0.8056** | 0.2943 | 0.5361 | 0.6254 | 0.3215 | **0.9489** |

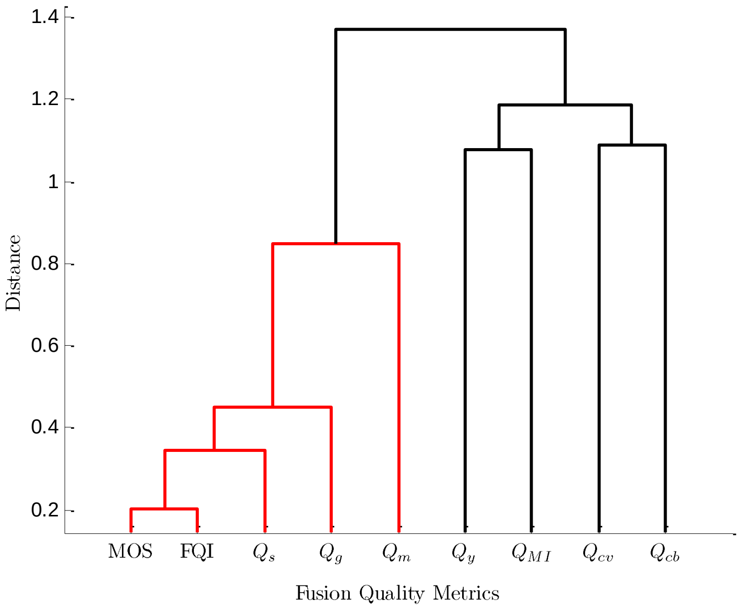


Figure 5.3: Dendrogram of fusion measures.

## 5.2 Multi-focus Image Fusion

The imaging properties of an optical system depend on the acquisition parameters such as focal length and the distances to the objects imaged. Due to limited depth-of-focus of optical lenses, it is often not possible to acquire an image that contains all relevant objects in-focus. Multi-focus fusion process can create a single image where all objects are in-focus. Having said that, it is not an easy task to distinguish out-of-focus objects. To take the advantages of LPC properties, image fusion technique employed to combine multiple multi-focus images captured using different focus levels. In case that optical out-of-focus is the major source of quality degradations, it is natural to assume that the image region that has higher local energy/contrast and higher sharpness is more active and thus more informative. In existing fusion algorithms, a common implicit assumption is that finding the high energy/contrast regions is equated with finding the high sharpness regions [53]. While this might be true in many application environments, it may not always hold, especially when the images are captured using different instrument modalities. A synthetic example is given in Fig. 5.4, which demonstrates that local energy/contrast and local sharpness measurements can be two independent events. For example, the top-right square in the first image has lower contrast than that in the second image, but apparently it has higher sharpness. This motivates us to develop a novel image fusion algorithm that uses two different activity measures to assess local energy/contrast and local sharpness, respectively. We can then create a new fused image that is maximal in both local contrast and sharpness at each spatial location. At the core of our

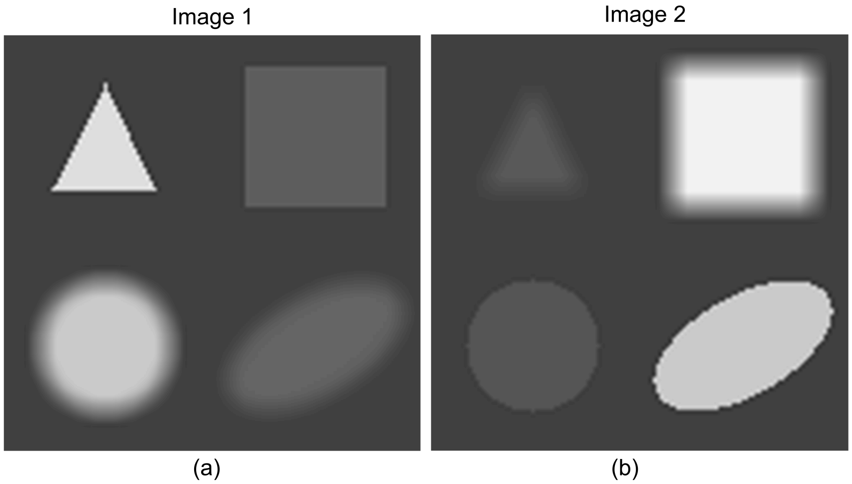


Figure 5.4: Synthetic images with different sharpness and contrast in corresponding objects.

approach is a novel complex-wavelet domain image sharpness measure based on the LPC measurement described in previous chapters.

### 5.2.1 Multiresolution Fusion Scheme

In recent years, multiresolution analysis has become a widely adopted method for image fusion techniques for various reasons: (1) real-world objects consist of structure at different scales, (2) there is a strong evidence that the HVS process retinal images in multiresolution fashion, and (3) multiresolution methods offer computational advantages and appear to be more robust. The basic idea of multiresolution image fusion is to perform multiresolution decomposition on each input image, integrate the decompositions to form a composite representation based on certain fusion rules, and then reconstruct the fused image by performing an inverse multiresolution transform.

In Fig. 5.5, we show a detailed hierarchy of the proposed fusion scheme, which consists mainly of three modules: *activity measure*, *decision rule*, and *combination rule*. Below, we give a short description of each building block.

*Activity Measure*: The purpose here is to provide an indication of informative content at each spatial location. In the literature, a body of research work computes the activity as some sort of energy calculation. Two of the most commonly used activities are CM scheme and WA scheme. In the former scheme the fused coefficient is obtained by selecting the maximum magnitude of the corresponding coefficients in all images, while in the later one the fused coefficient is calculated as a weighted average of the corresponding coefficients. Different from existing approaches, we compute two distinctive activity measures, one for local sharpness and the other for local energy/contrast. We observe that the local phase coherence

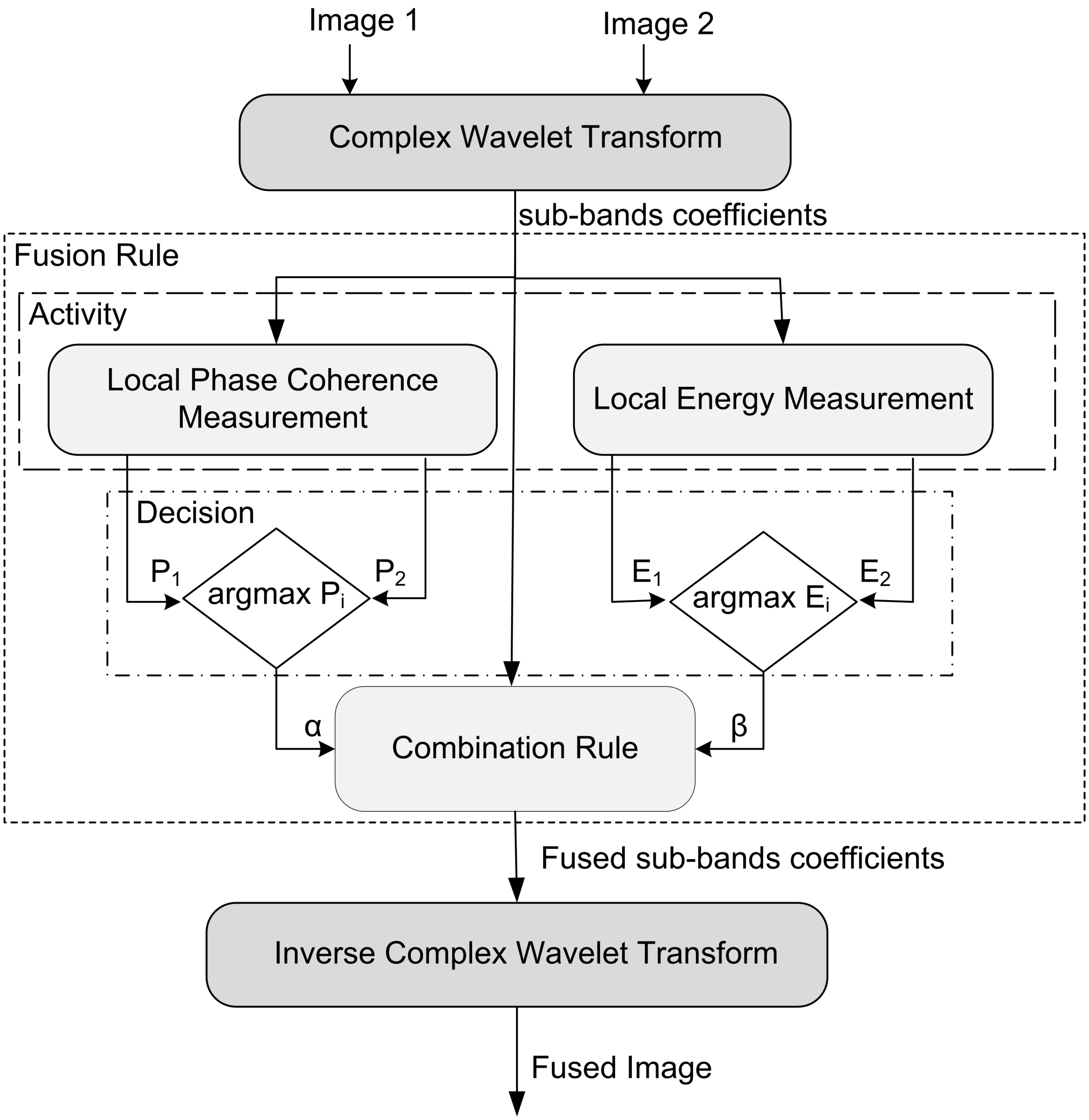


Figure 5.5: Diagram of the proposed image fusion algorithm.

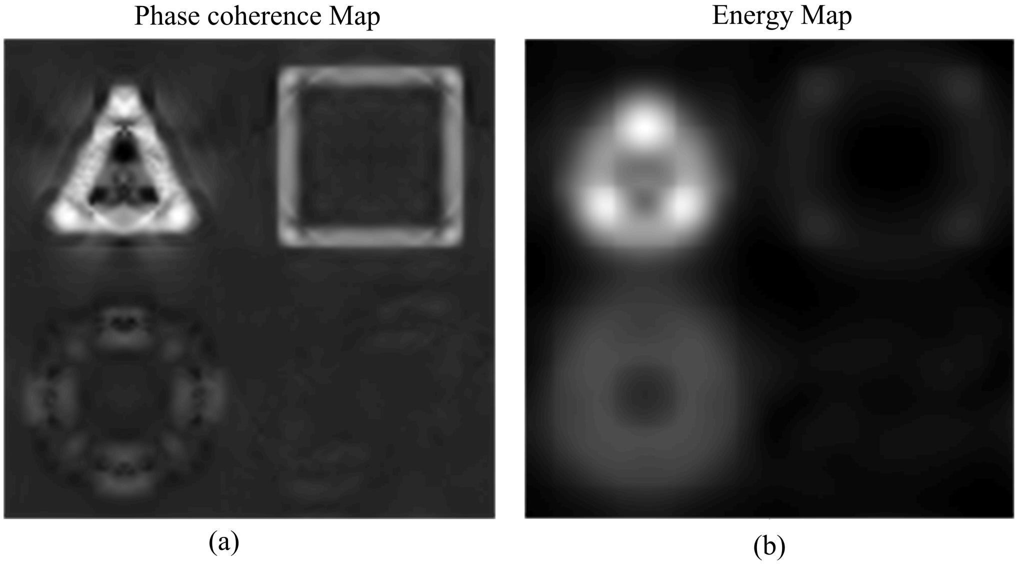


Figure 5.6: (a) Local phase coherence map and (b) Local energy map of Fig. 5.4(a).

measurement described in the previous chapters gives strong responses at sharp image features and almost ignore blurred image features. However, local energy gives strong responses to high contrast features and low responses to low contrast feature. This motivates us to use local phase coherence measurement Eq.( 3.10) as a measure of sharpness. The local energy/contrast measure is defined as:

$$E_i = W * \sum_{o} |c_i| \ , \tag{5.10}$$

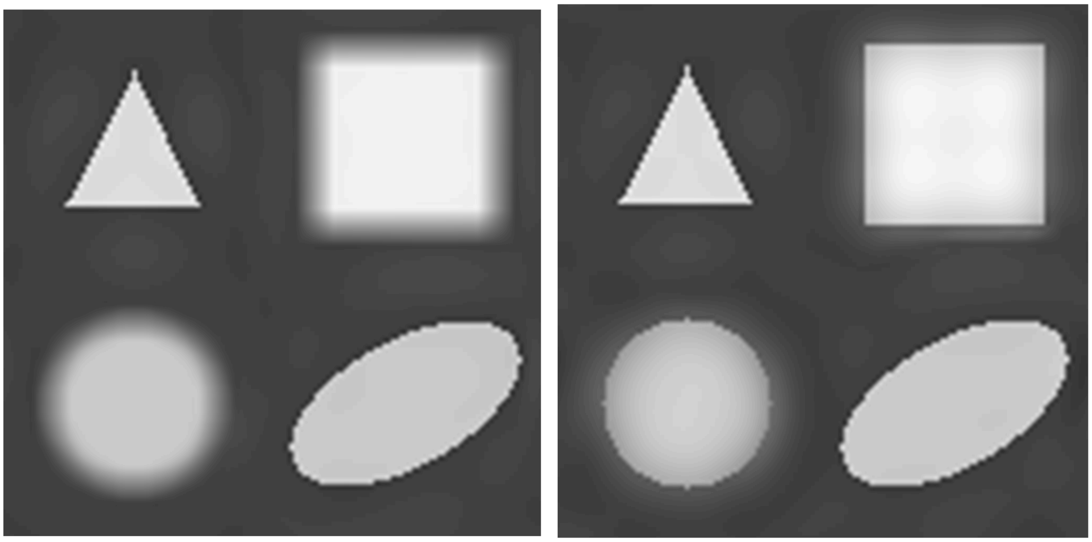
where $W$ is a smoothing filter, which is convolved with the sum of the magnitudes of wavelet coefficients to provide a smooth local energy map. Fig. 5.6 shows an example of local phase coherence map and local energy map of the test image shown in Fig. 5.4(a). It is clear that they emphasize on different types of activities in the image.

*Decision and Combination Rules*: These two components represent the core of the fusion algorithm. Their output governs the actual combination between the coefficients of the multiresolution decomposition of the source images. For each level $k$, orientation band $p$, and location $n$, the combination rule and decision rule produce a fused coefficients. The decision rule is mainly used to choose the coefficients which give high contrast and high sharpness activity measures. In other words, the decision process decides that the most salient coefficient (has maximum energy and sharpness) is the best choice for the composite coefficient and tells the combination process where they are in order to select them.

The selective combination then fuse the coefficients which have been selected by the decision rule to produce maximum energy and maximum sharpness. Two cases could occur here. In the first case, the maximum energy and maximum sharpness coefficient belongs to the same source image. In this case this coefficient is included in the fused band as it is. In the second case, the maximum energy coefficient belongs to one of the source images while the maximum sharpness coefficient belongs to another. In this situation, the combination rule will choose the coefficient with maximum sharpness and boost their energy to the level of the maximum local energy coefficient. By doing so, sharp and high contrast features from both images are combined. The decision and combination rules are given by

$$\begin{aligned} \alpha &= arg \max_i LPCM_i \,, \\ \beta &= arg \max_i E_i \,, \\ c_{fused} &= \frac{E_\beta}{E_\alpha} \cdot c_\alpha \,, \end{aligned} \tag{5.11}$$

where $LPCM_i$ and $E_i$ are the $i$-th pixel value obtained from the LPC and the local energy maps, respectively. Currently, the algorithm is tested on two source images, however the above general rule can be directly extended for fusing three or more source images.



(a) (b)

Figure 5.7: Fusion results for the images in Fig. 5.4 using (a) local energy activity measure only and (b) the proposed fusion method.

### 5.2.2 Experimental Results

The proposed fusion algorithm is demonstrated using four experiments. In the first experiment, we show the necessity of using separate activity measures for local contrast and local sharpness. Two synthetic test images are shown in Fig. 5.4, where the same four objects appear in both images at the same spatial locations. The only difference is in their contrast and sharpness. In particular, compared with those in Image 2, the top-left, top-right, bottom-left and bottom-right objects in Image 1 have higher-contrast/higher-sharpness, lower-contrast/higher-sharpness, higher-contrast/lower-sharpness, and lower-contrast/ lower-sharpness, respectively. We apply the proposed fusion algorithm to these images and compare it with the result obtained by using the local energy activity measure (as in Eq. (5.10)) only. The fused images are shown in Fig. 5.7. It appears that using only the local energy measure, the higher sharpness of the top-right object in Image 1 and bottom-left object in Image 2 cannot be incorporated. By contrast, all useful information has been appropriately fused by the proposed method.

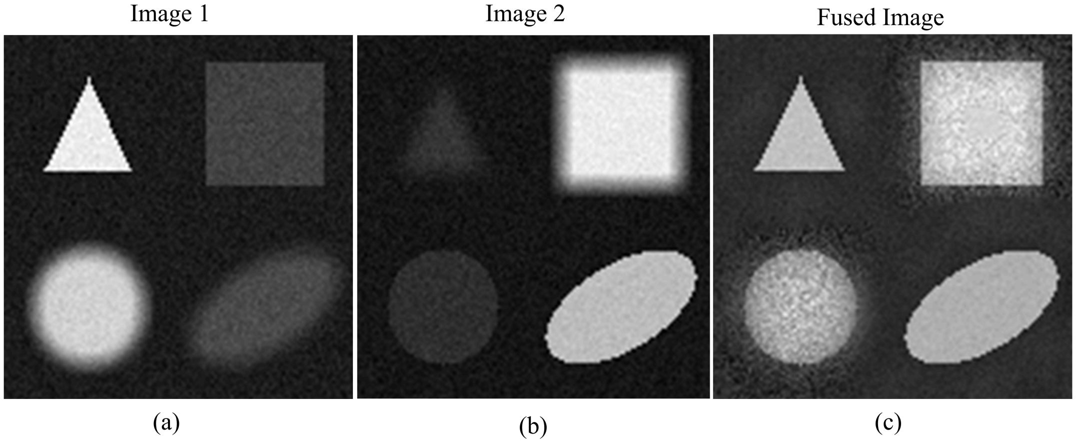


Figure 5.8: (a),(b): Synthetic noisy images; (c): Fused image using the proposed method.

The purpose of the second experiment is to test the robustness of the proposed method in the presence of noise. The same images used in the first experiment are employed, but with added independent white Gaussian noise. The results are shown in Fig. 5.8. It can be observed that the performance of the proposed algorithm does not change demonstrating the robustness of the local phase coherence and local energy measures as well as the fusion algorithm. To provide more insights about the effect of noise on LPC, Fig. 5.9 plots LPC measurmeant with different noise levels. It can be observed that LPC values slightly decay with increased noise level.

In the third experiment, we test the proposed fusion algorithm using microscopy images acquired by light microscope with varying focus settings. The result is shown in Fig. 5.10, where the microscopy images were obtained from the Vision Research lab web site at University of California, Santa Barbra [107]. It can be seen that high contrast and sharp features from both images are appropriately merged in the fused image.

The last experiment is performed on natural images obtained from Computer

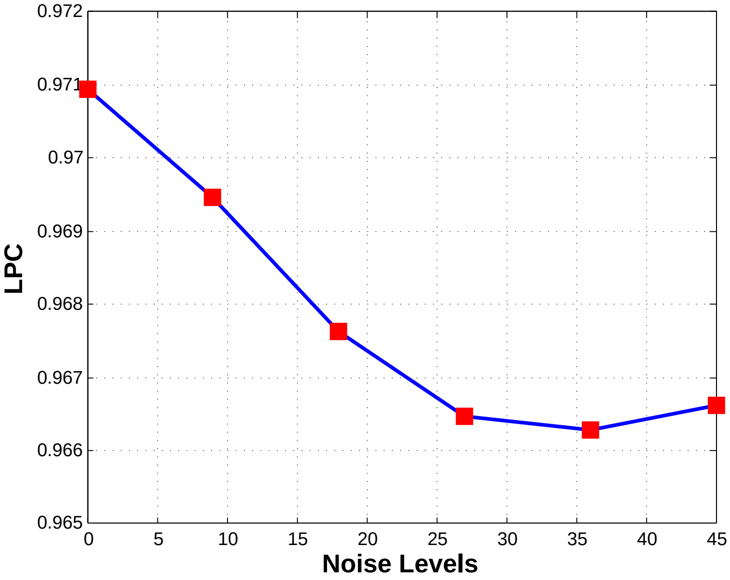


Figure 5.9: LPC Measurment with Different Noise Levels.

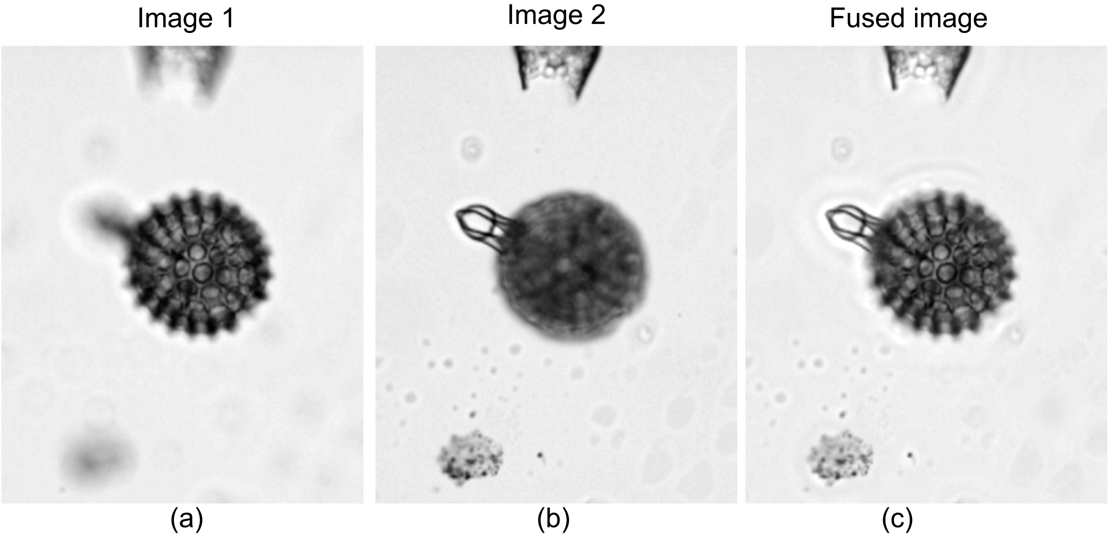
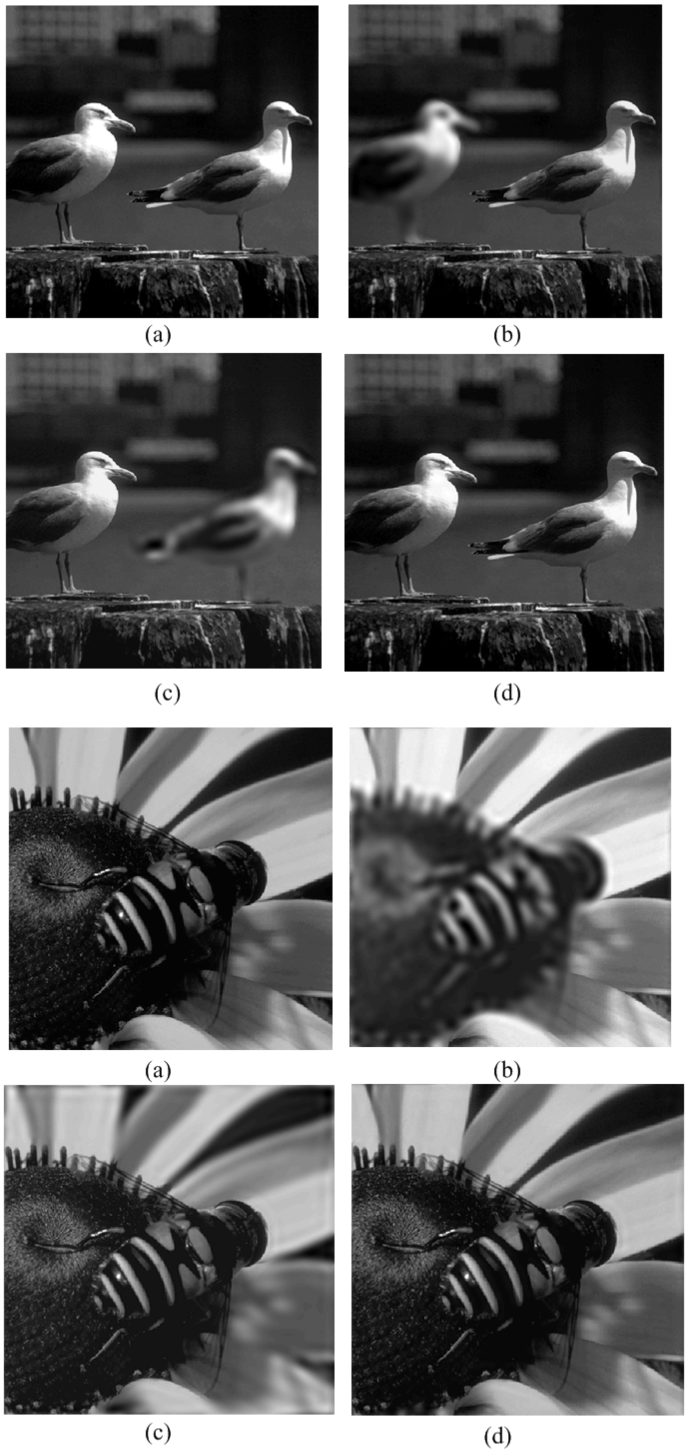


Figure 5.10: (a),(b): Two images of radioleria acquired by optical microscope; (c): Fused Image using the proposed algorithm.

Vision Group, University of Granada [108]. This experiment is divided into two parts. In the first part, we used natural images with two or more objects in the scene. Each image is manipulated to produce two input images, such that each object in the scene appears to be sharp in one image and blurred in the other image. The proposed fusion scheme is then applied to the input images in order to have one image with all objects are sharp everywhere. The results are shown in Fig. 5.11 which demonstrate the success of the proposed sharpness measure to detect blurred objects in any type of images. The second part of the experiment is conducted in another set of images with one object in the scene. The original image is blurred by Gaussian function with different standard deviations to produce one input image while the second input image is obtained from the original image by decreasing the contrast of the original image to produce a low contrast image. The blurred and the low contrast images are both fed to the proposed fusion scheme and the results are shown in Fig. 5.12. It is clear that the fused image combines the sharpness and the high contrast features from both input images.

## 5.3 Conclusions

In this chapter we proposed a novel blind objective quality assessment for image fusion. Three main characteristics have been taken into account: local contrast, sharpness, and structural preservation measure to estimate how well the important information in the source images is represented by the fused image. The structural preservation measure is designed as a weighted sum of the structural similarity maps between each source image and the fusion result. The source images LPC sharpness maps, introduced in the previous chapter, has been employed as local weights to the structural similarity maps. This is intuitively plausible because the sharp image features would be given more weights and hence results in a comparison measure that penalize both dissimilar and unsharp features in the final evaluation of the fused image. A simultaneous multi-exposure multi-focus image database has been created and subjectively rated to serve in the validation process. Validations using our subject-rated image database and comparison with state-of-the-art fusion metrics show good correlations between subjective ranking score and the proposed image fusion quality metric. Furthermore, we demonstrate the extended applications of LPC sharpness estimators to develop a fusion algorithm for multi-focus images. The fusion algorithm incorporates LPC measure as a sharpness activity measure to choose sharp image features in a multiresolution framework. Local energy is employed as high contrast activity measure. Results demonstrate that combining these two activity measures results in both maximal contrast and sharpness in the fusion results.



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

Figure 5.11: Fusion of natural images using the proposed method: (a) Original image, (b) blurred image 1, (c) blurred image 2, and (d) the fused image.

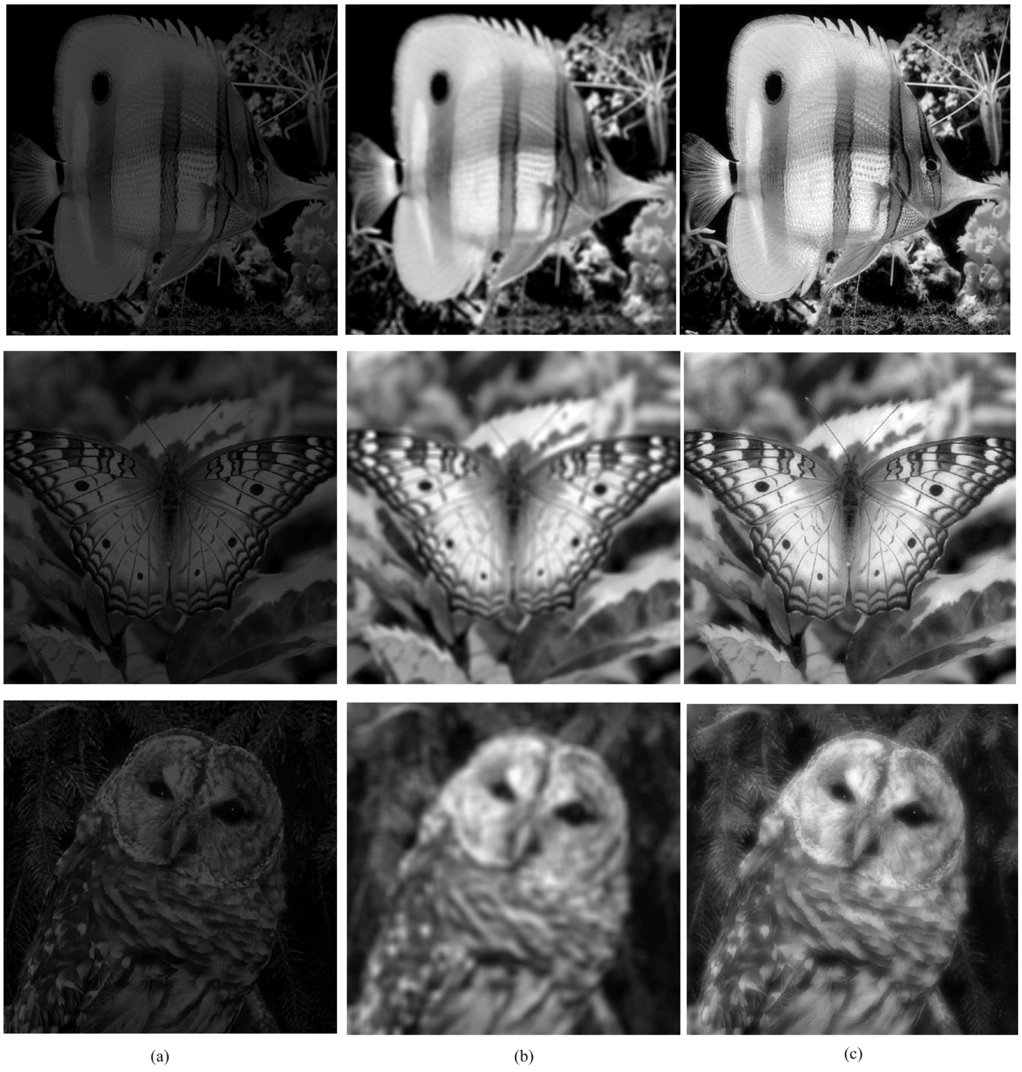


Figure 5.12: Fusion of natural images using the proposed method: (a) low contrast image, (b) blurred image, (c) fused image.

# Chapter 6

# Multi-sensor Image Registration

## 6.1 Introduction

Image registration is a fundamental task in image processing that aims to match two or more images taken at different times, different points of view, or different sensors (sometimes called multi-sensor or multi-modal). A great number of image processing and computer vision systems require image registration as an intermediate step, examples of such systems including, but not limited to, real-time target recognition, matching stereo images for 3D shape reconstruction, fusion of satellite images, and aligning medical images for diagnosis purpose.

Extraordinary advances in sensor technology and microelectronics in the past few decades have brought a strong need for processing techniques that can effectively combine information from different sensors. In order to do so, multi-sensor image registration become a crucial step in image analysis tasks in which the final information is gained from the combination of various data sources. The main goal of multi-sensor image registration is to geometrically align images of the same scene taken by different sensors or modalities. Challenging issues found in multisensor images include: (1) the relationship is unknown between intensity values of corresponding pixels; (2) Image contrast may differ in the same regions from each other; (3) Multiple intensity values in one image may map to a single intensity value in another image; (4) Structures present in one image may not appear in another image. Most of these problems, and many related variations, cannot be solved by intensity-based registration methods. However, optimal transformation between images can be found by choosing proper image features that are able to detect the salient image structure so that the points in one image can be mapped correctly to their corresponding points in another image. This type of registration is called feature-based registration. The performance of feature-based methods depends highly on the features being used. In this research, we proposed a novel feature-based registration method based on our LPC measurement. As we have shown before that local phase coherence measurement is invariant to intensity or contrast variations. In the context of multi-sensor registration, local phase coherence measurement seems

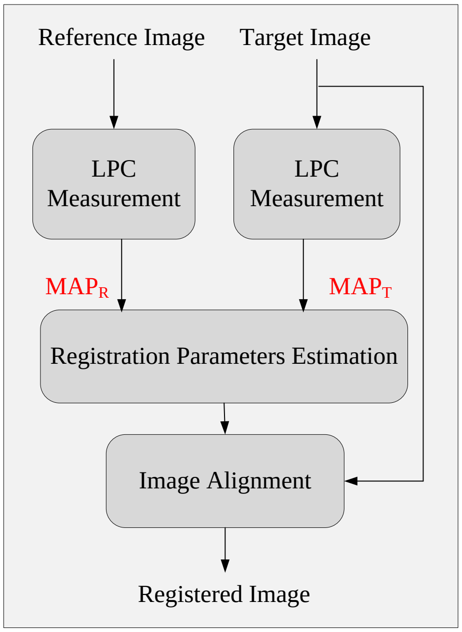


Figure 6.1: Overall architecture of the proposed registration algorithm.

apt to capture the salient common features between different modalities. This is the premise underlying the development of our multi-sensor registration algorithm to be described subsequently.

## 6.2 Proposed Registration Algorithm

The overall architecture of the proposed registration scheme is shown in Fig. 6.1. Given a reference image and a target image to be registered, the proposed algorithm starts by computing the local phase coherence map for both images using Eq.( 3.10). In multi-sensor images, the process of local phase estimation can be thought of as a conservative information reduction, ideally suited to registration problem for which an obvious structural relationship between images exists, while an intensity relationship does not. Once we compute the local phase coherence representation, we are ready to find the registration which best matches these images. Currently, a popular method for registration using image intensities is by analyzing the joint distribution of intensities at each iteration. The general effect of misalignment is to spread out the joint distribution in some way. The best registration is typically assumed to be the one which minimizes the amount of spreading. Therefore, it is mandatory to define a measure of dispersion. Mutual information is a famous information theoretic similarity measure which provides an excellent dispersion measure. Mutual information in its formal form is defined as [109]:

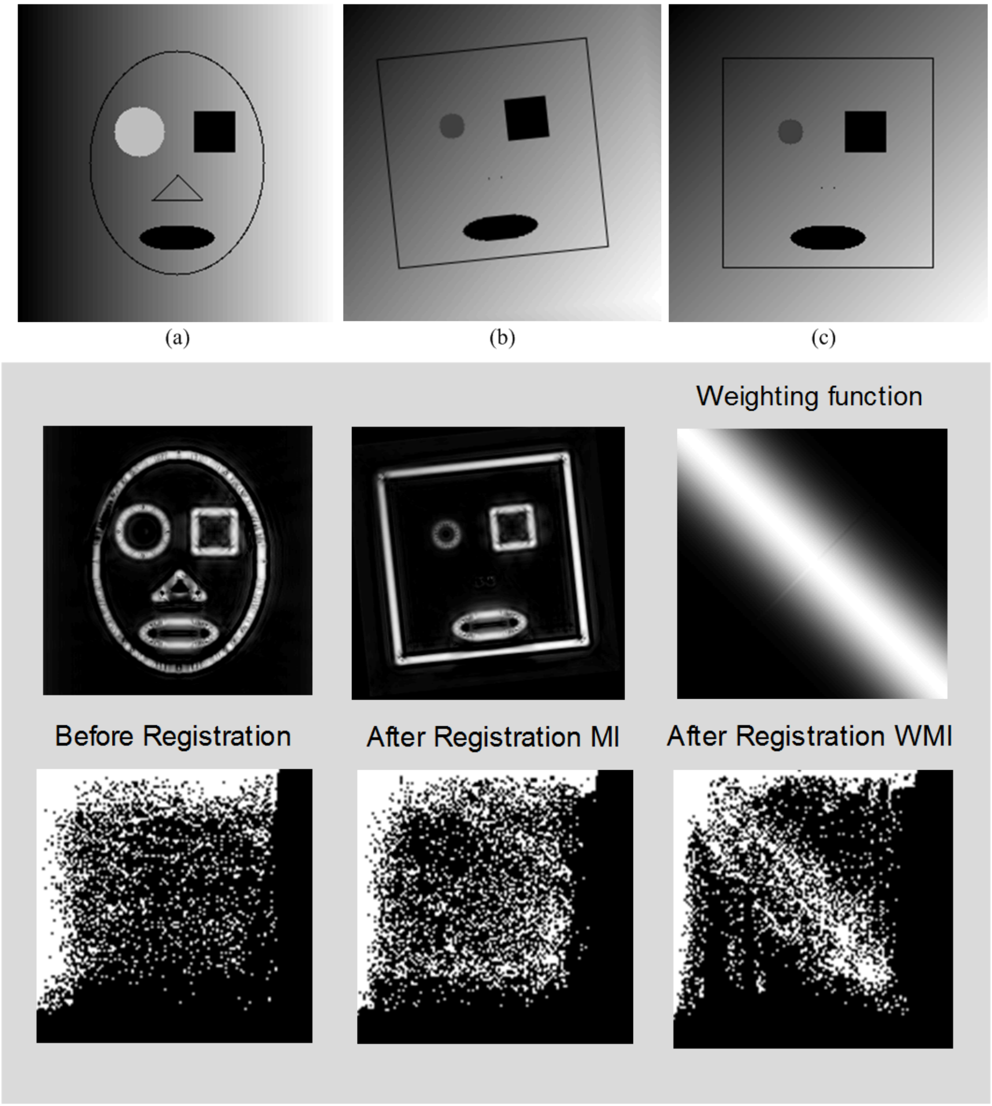
$$I(\mathbf{x}) = \sum_i \sum_j P_{ij} \log \left( \frac{P_{ij}}{P_i P_j} \right), \tag{6.1}$$

where $P_i$ and $P_j$ are the marginal distributions of the reference and target images, respectively, and $P_{ij}$ is the joint distribution. The main problem in this formal mutual information measure is that all the points in the joint histogram are treated the same way everywhere. This is not appropriate in multisensor images, specifically when some objects appear in one image but not in the other. The joint distribution in this case will show many points near one of the coordinate axes which corresponds to the image with more objects and thus mutual information will give small values since the joint distribution disperses. This motivates us to use weighted mutual information as a dispersion measure for multisensor image registration. In this research we used weighted mutual information (WMI) as a cost function computed from the local phase coherence representation of images to be registered. WMI is defined as:

$$I(\mathbf{x}) = \sum_i \sum_j w_{ij} P_{ij} \log \left( \frac{P_{ij}}{P_i P_j} \right), \tag{6.2}$$

where $P_i$ and $P_j$ are the marginal histograms of the phase coherence maps for the reference and target images, respectively, $P_{ij}$ is the joint histogram between the two maps, and $w_{ij}$ is the weight given to each entry in the joint histogram. As a special case, the conventional mutual information based approach corresponds to the case that all $w_{ij} \equiv 1$. This is problematic because the existence of partial or missing information between the two images may cause significant loss of correspondence in the joint histograms, and thus largely affect the estimation of the registration parameters.

Fig. 6.2 shows two phantom images created to demonstrate the idea behind the proposed algorithm. Two images have different intensity values and image structures except for two objects, namely the squared right eye and the ellipsoid mouth. The second image is altered by rotation and translation transforms. The joint histogram of LPC values shows clouds near both axes, which correspond to objects appear in one image but not in the other. The proposed weighting function is also shown in Fig. 6.2. In our approach, we use this weighting function that assigns larger weights along the main diagonal of the joint histogram, and decreasing weights to points far from the diagonal. By doing so, we give less weights to outliers or points in the joint histogram with no corresponding data between the images. In other words, we put more emphasis on the points where there are possible matches between the two images. In our implementation, we allow for global translation and rotation of the images and the estimated transformation parameters $\mathbf{x}$ are obtained using a nonlinear unconstrained simplex search algorithm. It appears that the proposed weighting function helps the optimizer avoid being trapped in local maxima while retaining a high degree of precision in the determination of the transformation parameters. This is illustrated in Fig. 6.2, where the joint



(a) (b) (c)

(d)

Figure 6.2: (a) Reference image,(b) Target image translated and rotated around the image center, (c) Registered image using WMI, (d) The phase coherence map of both images and the corresponding joint histograms before and after registration using both MI and WMI, along with the proposed weighting function.

Table 6.1: RMS Error Comparison of MI and WMI Methods

| Method | WMI | MI |
|---|---|---|
| | $x$ ,$y$ ,$\theta$ | $x$ ,$y$ ,$\theta$ |
| TD1 | 0.0796 ,0.0721 ,0.151 | 1.645 ,1.325 ,2.397 |
| TD2 | 0.2181 ,0.2036 ,0.152 | 5.049 ,4.474 ,3.800 |

histogram becomes more concentrated along the main diagonal using the proposed WMI approach when compared with the conventional MI-based method. Finally, the registration parameters estimated using the local phase coherence maps are employed to transform the target image to create the registered image, shown in Fig. 6.2.

## 6.3 Experimental Results

In this section, the proposed registration algorithm is evaluated with three sets of experiments using both synthetic images (Test Data 1 (TD1) shown in Fig. 6.2) and MRI brain images obtained from the National Library of Medicine Visible Human Project (Test Data 2 (TD2) shown in the first row of Fig. 6.4(a,c)) [110]. The synthetic images and the brain images are initially in perfect alignment, so that the ground truth transformation is already known. The goal of the first experiment is to show the advantage of WMI Eq. (6.2), over conventional mutual information Eq. (6.1) in the case of partial and missing data. For the synthetic images in TD1, the target image was altered by 11 sets of rigid transformations (translation and rotation), where the transformations range from -20 to 16 pixel shift along x-direction, -15 to 16 pixel shift along y-direction, and from 3 to 10 degrees of counter-clockwise rotation around the center of the image. Different transformation parameters were applied to the T1-MRI brain image in TD2, where the transformation parameters ranged from -24 to 20 pixel shift in x-direction, -30 to 18 pixel shift in y-direction, and from -10 to 10 degrees of rotation around the image center. To evaluate registration accuracy, RMS error between the ground truth transformation and the estimated registration parameters is calculated. Table 6.1 compares the average RMS results between conventional mutual information and the WMI methods. It can be seen that for both test data sets, the proposed method have significantly smaller RMS error, demonstrating the effectiveness of the WMI approach.

The purpose of the second experiment is to evaluate the robustness of the proposed algorithm in the presence of noise. Both the reference and target images are corrupted by additive independent white Gaussian noise at a wide range of noise levels. An example is given in Fig. 6.3, where it can be observed that the proposed method provides us with accurate registration result, regardless of the existence of heavy noise. The RMS performance comparison of the proposed algorithm and the conventional MI method is shown in Fig. 6.3. It can be seen that the proposed

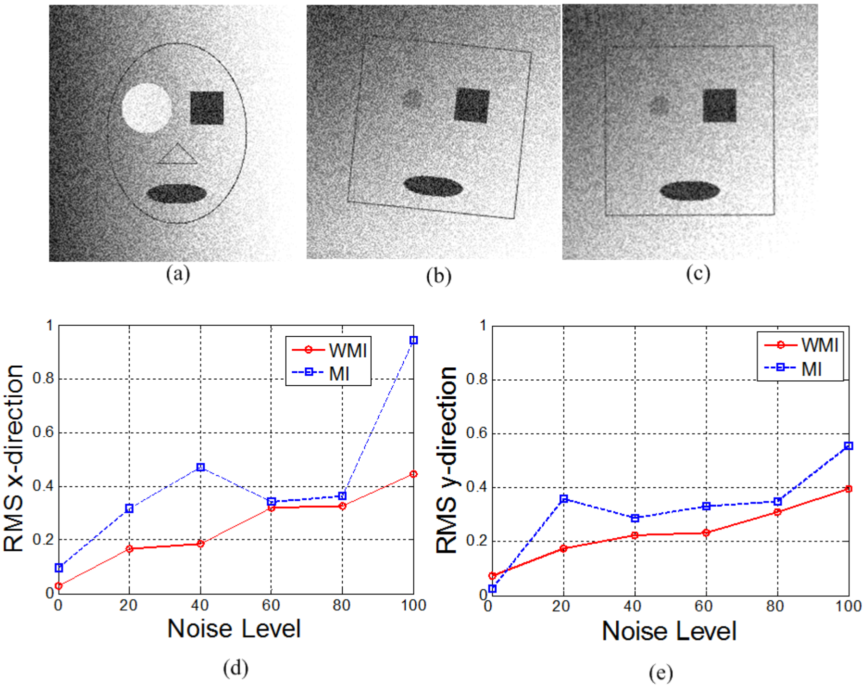


Figure 6.3: (a) Noisy reference image; (b) Noisy target image; (c) Registered image; (d),(e) Comparison of RMS estimation error in the $x$-direction and $y$-direction against different noise levels.

method performs better at almost all noise levels and the resulting RMS values are quite small with the existence of very heavy noise, demonstrating strong robustness against noise. We believe that such noise-robustness is a result of the stability of the proposed LPC measurement which creates small values for noise patterns (because of their phase randomness and relatively low magnitudes).

The third experiment is designed to show the performance of the proposed algorithm in the case of significant missing information between the images. As shown in Fig. 6.4, the missing data is simulated by cutting half of the target image, which is to be registered with a reference image. Some of the MRI images used in this experiment are obtained from BrainWeb database [111]. It can be seen that the proposed algorithm correctly re-aligns the target image into the right position, regardless of the large amount of missing data, the apparent intensity variations between different imaging modalities, as well as the significant translation and rotation.

In practice, partial and missing data are quite common, as demonstrated in Fig. 6.5, where the images are obtained from the Whole Brain Atlas website [112]. In the first row, the first image is T2-weighted MR which shows very clearly a large tumor region while in the second image which is T1-weighted MR, the tumor boundaries are not well-defined. The same situation appears in the second row of Fig. 6.5. Here the first image is T1-weighted MRI and the second image is CT scan. In this case the small dark circle region in T1-MRI corresponds to a localized

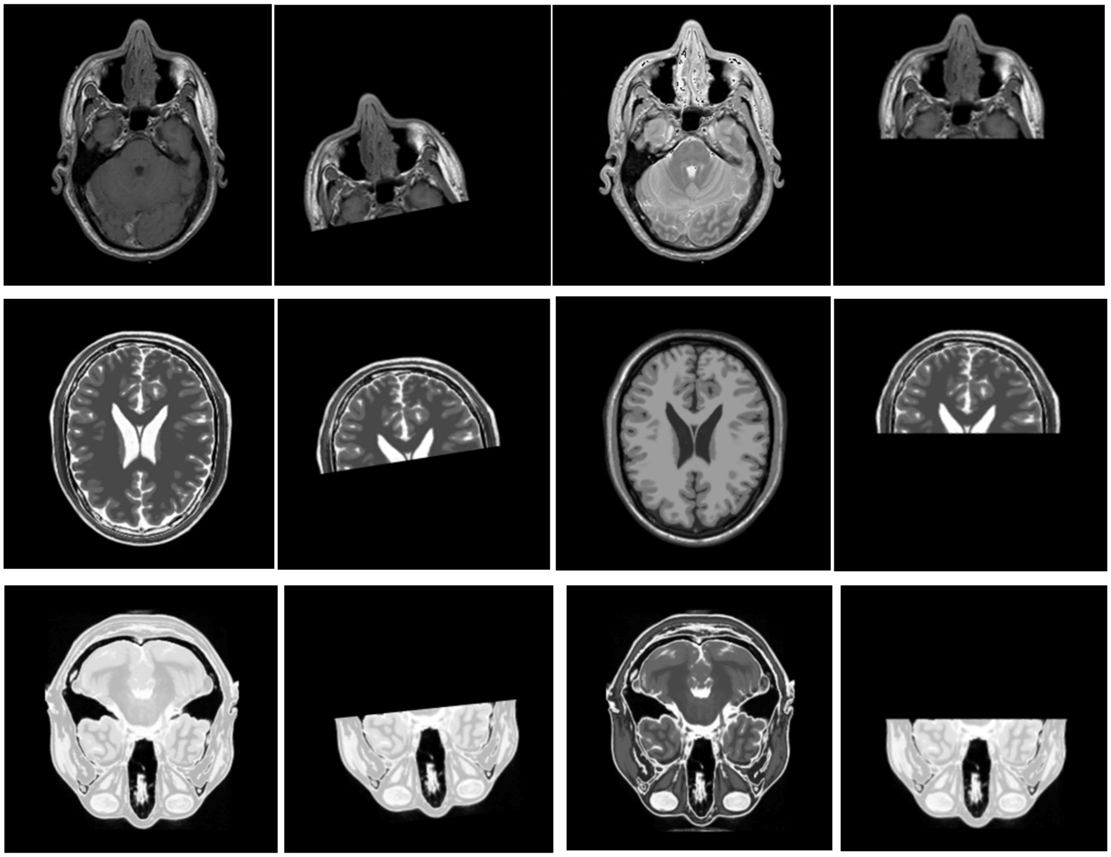

(a) (b) (c) (d)

Figure 6.4: Registration results using the proposed method in the case of simulated missing data in brain images (a) Original reference image; (b) reference image with missing data; (c) target image; (d) Registered image.

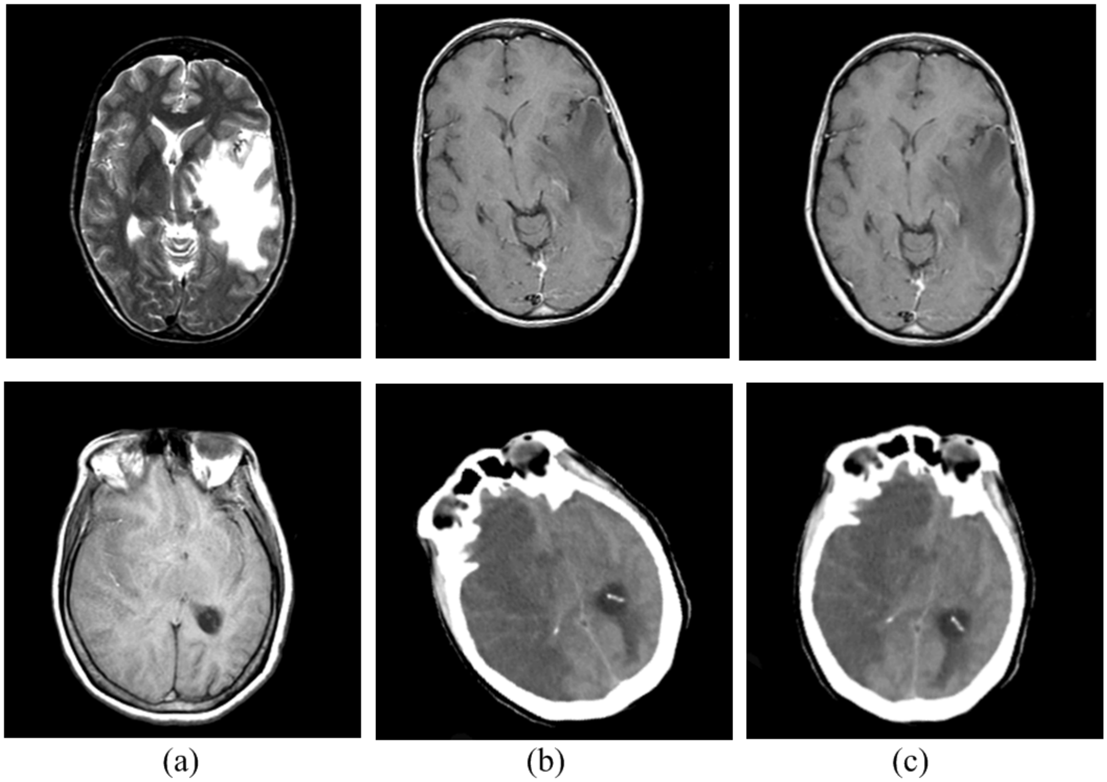


Figure 6.5: Registration results using the proposed method in the case of real missing data in brain images (a) Original reference image; (b) target image; (d) Registered image.

necrosis resulting from obstruction of the blood supply. It is obviously visible in T1-MRI but not clear in CT scan. In both cases the proposed registration algorithm is able to re-align the target image with the reference image.

## 6.4 Conclusions

In this chapter, we propose a novel method for multi-sensor image registration, which is based on two key ideas. The first is the introduction of using local phase coherence measure as a new saliency feature, which is simultaneously insensitive to luminance/contrast variations and noise contamination. The second idea is the use of a weighted mutual information-based objective function in searching the optimal registration parameters, so that more weights are given to the points that are likely to find correspondence between the images being registered. The proposed method demonstrates good performance with both synthetic and MRI brain images in the cases of significant intensity and contrast variations, noise contamination, and partial and missing data between the images.

# Chapter 7

# Conclusion and Future Work

## 7.1 Conclusion

This thesis builds upon the idea of local phase coherence and develops a series of methodologies for several image processing applications. Local phase coherence has been shown to provide useful insights in the visual perception of blur. However, not much has been done to find reliable low-level measure of phase coherence in images. By studying the behavior of local phases at different scales for different image features, this thesis has developed a low-level measure of the strength of local phase coherence. This measure helps to better understand the phase coherence phenomena and broaden the potential applications of local phase coherence in processing images. In achieving thesis overall objectives, a number of novel ideas has been developed, which are summarized as follow:

- Develop a novel framework that can measure of the strength of LPC from samples of any number of scales and with fractional scale ratios.
- Propose a computationally efficient LPC evaluation method that can be easily adopted in practical image processing applications.
- Introduce a novel no-reference sharpness estimator upon the notion of LPC.
- Propose a new blind fusion quality assessment metric incorporating LPC measurement that well correlates with human visual perception of fusion results.
- Build a simple yet efficient multi-focus fusion algorithm utilizing LPC along with energy as two separate fusion rules in order to achieve both sharp and high contrast fusion results.
- Introduce a multi-sensor image registration approach which reduce image contrast/illumination variations between multi-sensor images by employing LPC maps as underlying images representation to compute optimal transformation parameters.

## 7.2 Future Work

The research presented in this dissertation provides a foundation for research prospects of local phase coherence in a wide variety of image processing problems. Many of the ideas developed in this work can be improved and/or adopted to better interpret image properties and better process images. Potential future research directions include the following:

- One aspect that was not fully addressed was the incorporation of local phase coherence into statistical image modeling. We believe that, such a statistical model is likely to lead to substantial improvements in a variety of applications, such as image deblurring, denoising, compression, and quality assessment.
- The computation of LPC can be further extended to provide accurate phase measurements. Such a measurement can be used to detect localized image features using phase samples measured above a certain allowable scale rather than spatial samples. This would provide a deeper implication on how the visual system could “see beyond the Nyquist rate” which could explain a number of visual phenomena and is critical to a variety of visual capabilities including various forms of hyperacuity, stereopsis, and motion estimation.
- The proposed LPC sharpness estimation can be easy incorporated into iterative algorithms as a stopping criterion for different image processing applications such as deblurring and super-resolution.
- All the images used to estimate sharpness in this work were of gray-scale intensity. It is intersting to investigate the effect of color channels on the perceived sharpness and to extend the LPC for estimation sharpness of color images.
- A single fusion metric is not sufficient to account for all the requirements for the fusion applications like satellite images, surveillance, and medical imaging. In multi-focus multi-exposure imaging, the input images are of the same modality, but in medical imaging, heterogeneous images are fused. A more reliable and universal fusion metric is desirable in future research.
- The image fusion algorithm can possibly be improved in many ways. One potential extension is to optimize the fusion algorithm for better image quality by taking advantage of the proposed fusion quality metric.
- For the purpose of this research, we have primarily focused on the idea of employing LPC measurement to discount the contrast/illumination variations between different modality images for the task of registration. In the optimization process used to determine the alignment between two images, we have mainly employed deterministic optimization approaches, which might be trapped by local optima. In future work, it is worth investigating stochastic optimization approaches.

- A great body of work performs image registration as a separate preprocessing step before image fusion using different techniques for registration and fusion. In this dissertation, we demonstrated the potentials of LPC measurment to perform both registration and fusion. In future work, we can develop a general framework to perform both registration and fusion using one single technique based on LPC measurment.

- For the purpose of this research, we have primarily focused on the design of image fusion algorithm to solve multi-focus multi-exposure problem that appeared in photography images. As future work, it is worth investigating the possible extension to medical applications.

# Published Papers

R.Hassen, Z.Wang, M.Salama, “Blind Fusion Quality Assessment Based on Local Phase Coherence”, in prep, 2013.

R.Hassen, Z.Wang, M.Salama, “Image Sharpness Assessment Based on Local Phase Coherence”, *IEEE Transactions on Image Processing*, accepted, to appear in 2013.

R.Hassen, Z.Wang, M.Salama, “A Flexible Framework for Local Phase Coherence Computation”, *International Conference of Image Analysis and Recognition*, Burnaby, BC, Canada, June 22-24, 2011.

R.Hassen,Z.Wang,M.Salama, “No-Reference Image Sharpness Assessment Based On Local Phase Coherence Measurement”, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Dallas, Texas, U.S.A, March 15-19, 2010.

R.Hassen, Z.Wang, M.Salama, “Multisensor Image Registration Based-on Local Phase Coherence”, *IEEE International Conference of Image Processing*, Cairo, Egypt, November 7-11, 2009.

R.Hassen, Z.Wang, M.Salama, “Multifocus Image Fusion Using Local Phase Coherence Measurement”, *International Conference of Image Analysis and Recognition*, Halifax, Canada, July 6-8, 2009.

# References

[1] H.R. Sheikh, Z. Wang, L. Cormack, and A.C. Bovik, "LIVE image quality assessment database release 2," http://live.ece.utexas.edu/research/quality. viii, 31, 34, 37

[2] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008 - a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, no. 4, pp. 30–45, 2009. viii, 32, 34, 37

[3] E.C. Larson and D.M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 011006–011006, 2010. viii, 32, 34, 37

[4] P. Le Callet and F. Autrusseau, "Subjective quality assessment IRCCYN/IVC database 2005," 2005, http://www.irccyn.ec-nantes.fr/ivcdb/. viii, 32, 34, 37

[5] M.A. Saad, A.C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339–3352, 2012. ix, 32, 34, 35, 36, 37, 41

[6] M.A. Saad, A.C. Bovik, and C. Charrier, "DCT statistics model-based blind image quality assessment," in *IEEE International Conference on Image Processing (ICIP 2011)*. IEEE, 2011, pp. 3093–3096. ix, 32, 34, 35, 36, 37, 41

[7] A. Mittal, A. Moorthy, and A. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, to appear, 2012. ix, 32, 34, 35, 36, 37, 41

[8] A. Mittal, A. K. Moorthy, and A. C. Bovik, "Referenceless image spatial quality evaluation engine," in *45th Asilomar Conference on Signals, Systems and Computers*, November 2011. ix, 32, 34, 35, 36, 37, 41

[9] C.T. Vu, T.D. Phan, and D.M. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 934, 2012. ix, 29, 32, 33, 34, 35, 36, 37, 38, 39, 41

[10] N. Narvekar and L. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE Transactions on Image Processing*, , no. 99, pp. 1–1, 2011. ix, 28, 32, 33, 34, 35, 36, 37, 38, 39, 41

[11] R. Ferzli and L.J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 717–728, 2009. ix, 28, 32, 34, 35, 36, 37, 41

[12] X. Zhu and P. Milanfar, "A no-reference sharpness metric sensitive to blur and noise," in *1st International Workshop on Quality of Multimedia Experience (QoMEX)*, 2009. ix, 29, 32, 34, 35, 37, 38, 39, 41

[13] AV Oppenheim and JS Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981. 1, 5

[14] J. Behar, M. Porat, and YY Zeevi, "Image reconstruction from localized phase," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 736–743, 1992. 1

[15] MC Morrone, J. Ross, DC Burr, and R. Owens, "Mach bands depend on spatial phase," *Nature*, vol. 324, no. 6094, pp. 250–253, 1986. 1, 7

[16] M.C. Morrone and DC Burr, "Feature detection in human vision: A phase-dependent energy model," *Proceedings of the Royal Society of London. Series B, biological sciences*, pp. 221–245, 1988. 1, 7

[17] D.J. Fleet, A.D. Jepson, and M.R.M. Jenkin, "Phase-based disparity measurement," *CVGIP: Image understanding*, vol. 53, no. 2, pp. 198–210, 1991. 1, 9

[18] FW Campbell and JG Robson, "Application of Fourier analysis to the visibility of gratings," *The Journal of Physiology*, vol. 197, no. 3, pp. 551–566, 1968. 1

[19] D. J. Field and D. J. Tolhurst, "The structure and symmetry of simple-cell receptive field profiles in the cats visual cortex," *Proc. Roy. Soc. London B*, vol. 228, pp. 379–400, 1986. 1

[20] D.J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Am. A*, vol. 4, no. 12, pp. 2379–2394, 1987. 1, 29

[21] P. Kovesi, "Image features from phase congruency," *Videre: Journal of Computer Vision Research*, vol. 1, no. 3, pp. 1–26, 1999. 1, 7, 11, 14

[22] Z. Wang and E.P. Simoncelli, "Local phase coherence and the perception of blur," *Advances in Neural Information Processing Systems*, vol. 16, 2004. 1, 13, 14, 15, 16, 17, 19, 22

[23] A. Oppenheim, J. Lim, G. Kopec, and S. Pohlig, "Phase in speech and pictures," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'79.* IEEE, 1979, vol. 4, pp. 632–637. 5

[24] DJ Tolhurst, "On the possible existence of edge detector neurones in the human visual system," *Vision Research*, vol. 12, no. 5, pp. 797–IN1, 1972. 5

[25] JJ Kulikowski and PE King-Smith, "Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation," *Vision Research*, vol. 13, no. 8, pp. 1455–1478, 1973. 5

[26] DC Burr, "Sensitivity to spatial phase," *Vision research*, vol. 20, no. 5, pp. 391, 1980. 5

[27] T. Caelli and P. Bevan, "Visual sensitivity to two-dimensional spatial phase," *Journal of the Optical Society of America*, vol. 72, no. 10, pp. 1375, 1982. 5

[28] JG Daugmann, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision Res*, vol. 20, no. 10, pp. 847–56, 1980. 6

[29] JG Daugmann, "Complete discrete 2-d gabor transforms by neural networks for image analysis and compression," *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, vol. 36, no. 7, pp. 1169–1179, 1988. 6

[30] DA Pollen and SF Ronner, "Phase relationships between adjacent simple cells in the visual cortex," *Science*, vol. 212, no. 4501, pp. 1409–1411, 1981. 6

[31] M.O.F.B. Scholkopf, "Implicit wiener series for higher-order image analysis," in *Advances in Neural Information Processing Systems 17: Proceedings of the 2004 Conference.* MIT Press, 2005, vol. 17, p. 465. 7

[32] M. Thomson, "Visual coding and the phase structure of natural scenes," *Network: Computation in Neural Systems*, vol. 10, no. 2, pp. 123–132, 1999. 7

[33] M.G.A. Thomson, "Higher-order structure in natural scenes," *Journal of the Optical Society of America A*, vol. 16, no. 7, pp. 1549–1553, 1999. 7

[34] M.G.A. Thomson and D.H. Foster, "Role of second-and third-order statistics in the discriminability of natural images," *Journal of the Optical Society of America A*, vol. 14, no. 9, pp. 2081–2090, 1997. 7

[35] M.G.A. Thomson, D.H. Foster, and R.J. Summers, "Human sensitivity to phase perturbations in natural images: a statistical framework," *Perception*, vol. 29, pp. 1057–1069, 2000. 7

[36] F.A. Wichmann, D.I. Braun, and K.R. Gegenfurtner, “Phase noise and the classification of natural images,” *Vision Research*, vol. 46, no. 8-9, pp. 1520–1529, 2006. 7

[37] T.D. Sanger, “Stereo disparity computation using gabor filters,” *Biological Cybernetics*, vol. 59, no. 6, pp. 405–418, 1988. 8, 9

[38] M. Jenkin and A.D. Jepson, *The measurement of binocular disparity*, in Computational processes in human vision, (Z. Pylyshyn, Ed.), Ablex, New Jersey, 1988. 8

[39] J. Stokes, “Image matching with phase shift methods,” in *ISPRS Comm III. Symposium*, 1986, pp. 638–652. 8

[40] CD Kuglin and DC Hines, “The phase correlation image alignment method,” in *Proc. Int. Conf. on Cybernetics and Society*, 1975, vol. 4, pp. 163–165. 8, 11

[41] BKP Horn, “Non-correlation methods for stereo matching,” *Photogrammetric Engineering and Remote Sensing*, vol. 49, no. 4, pp. 535–536, 1983. 8

[42] M.R.M. Jenkin, A.D. Jepson, and J.K. Tsotsos, “Techniques for disparity measurement,” *CVGIP: Image Understanding*, vol. 53, no. 1, pp. 14–30, 1991. 8, 9

[43] D.J. Fleet and A.D. Jepson, “Computation of component image velocity from local phase information,” *International Journal of Computer Vision*, vol. 5, no. 1, pp. 77–104, 1990. 8, 9

[44] JL Barron, DJ Fleet, and SS Beauchemin, “Performance of optical flow techniques,” *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994. 8

[45] B.D. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in *International joint conference on artificial intelligence*, 1981, vol. 81, pp. 674–679. 9

[46] D.J. Fleet and A.D. Jepson, “Stability of phase information,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 12, pp. 1253–1268, 1993. 9

[47] Z. Wang and A. C. Bovik, “Mean squared error: love it or leave it? - a new look at signal fidelity measures,” *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009. 10

[48] Z. Wang, AC Bovik, HR Sheikh, and EP Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004. 10, 32, 34, 35, 37, 44, 46, 47, 52

[49] Z. Wang and EP Simoncelli, "Translation insensitive image similarity in complex wavelet domain," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05)*, 2005, vol. 2. 10

[50] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey, "complex wavelet structural similarity: A new image similarity index," *IEEE Transactions on Image Processing*, to appear 2009. 10

[51] J. Liu, BC Vemuri, and JL Marroquin, "Local frequency representations for robust multimodal image registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 462–469, 2002. 10

[52] A. Wong and D.A. Clausi, "Arrsi: Automatic registration of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1483, 2007. 11

[53] M. Mellor and M. Brady, "Phase mutual information as a similarity measure for registration," *Medical Image Analysis*, vol. 9, no. 4, pp. 330–343, 2005. 11, 55

[54] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, 2001. 11

[55] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997. 11

[56] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transforms," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 9, no. 5, pp. 700–703, 1987. 12

[57] BS Reddy and BN Chatterji, "An fft-based technique for translation, rotation, and scale-invariant image registration," *IEEE Transactions on Image Processing*, vol. 5, no. 8, pp. 1266–1271, 1996. 12

[58] H. Foroosh, JB Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, 2002. 12

[59] J. Morlet, G. Arens, E. Fourgeau, and D. Giard, "Wave propagation and sampling theory; part i, complex signal and scattering in multilayered media," *Geophysics*, vol. 47, no. 2, pp. 203–221, 1982. 13

[60] Javier Portilla and Eero P Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–70, 2000. 17

[61] R. Hassen, Z. Wang, and M. Salama, “Multi-sensor image registration based-on local phase coherence,” in *IEEE International Conference on Image Processing (ICIP 2009)*. IEEE, 2009, pp. 181–184. 17, 19

[62] R. Hassen, Z. Wang, and M. Salama, “Multifocus image fusion using local phase coherence measurement,” *International Conference on Image Analysis and Recognition*, pp. 54–63, 2009. 17, 19

[63] R. Hassen, Z. Wang, and M. Salama, “No-reference image sharpness assessment based on local phase coherence measurement,” in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 2010)*. IEEE, 2010, pp. 2434–2437. 17, 19, 30

[64] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, “Shiftable multi-scale transforms,” *IEEE Transaction on Information Theory*, vol. 38, no. 2, pp. 587–607, Mar 1992. 17

[65] R. Hassen, Z. Wang, and M. Salama, “A flexible framework for local phase coherence computation,” *International Conference on Image Analysis and Recognition*, 2011. 21

[66] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment (Syntheses Lectures on Image, Video and Multimedia Processing)*, Morgan-Claypool Publishers, 2006. 27

[67] D. Shaked and I. Tastl, “Sharpness measure: Towards automatic image enhancement,” in *IEEE International Conference on Image Processing, 2005. ICIP 2005.* IEEE, 2005, vol. 1, pp. I–937. 27

[68] C. Liu, W.T. Freeman, R. Szeliski, and S.B. Kang, “Noise estimation from a single image,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,*. IEEE, 2006, vol. 1, pp. 901–908. 27

[69] Y.C. Chung, J.M. Wang, R.R. Bailey, S.W. Chen, and S.L. Chang, “A non-parametric blur measure based on edge analysis for image processing applications,” in *IEEE Conference on Cybernetics and Intelligent Systems,*. IEEE, 2004, vol. 1, pp. 356–360. 27

[70] D.J. Field and N. Brady, “Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes,” *Vision Research*, vol. 37, no. 23, pp. 3367–3383, 1997. 28

[71] L. Firestoneand, K. Cook, N. Talsania, and K. Preston, “Comparison of autofocus methods for automated microscopy,” *Cytometry*, vol. 12, pp. 195–206, 1991. 28

[72] N. B. Nill and B. H. Bouzas, “Objective image quality measure derived from digital image power spectra,” *Optical Engineering*, vol. 31, no. 4, pp. 813–825, 1992. 28

[73] X. Marichal, W. Ma, and H. J. Zhang, "Blur determination in the compressed domain using dct information," in *IEEE International Conference on Image Processing (ICIP 1999)*, Oct. 1999, vol. 2, pp. 386–390. 28

[74] D. Shaked and I. Tastl, "Sharpness measure: Towards automatic image enhancement," in *IEEE International Conference on Image Processing (ICIP 2005)*, 2005, vol. 1, pp. 937–940. 28

[75] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," in *IEEE International Conference on Image Processing (ICIP 2002)*, Rochester, NY, 2002, IEEE, vol. 3, pp. 57–60. 28

[76] E. Ong, W. Lin, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, and F. Moschetti, "A no-reference quality metric for measuring image blur," in *International Symposium on Signal Processing and Its Applications*. IEEE, 2003, vol. 1, pp. 469–472. 28

[77] R. Ferzli and L. J. Karam, "Human visual system based on no-reference objective image sharpness metric," in *IEEE International Conference on Image Processing (ICIP 2006)*, 2006, pp. 2949–2952. 28

[78] S. Erasmus and K. Smith, "An automatic focusing and astigmatism correction system for the SEM and CTEM," *J. Microscopy*, vol. 127, pp. 185–199, 1982. 29

[79] N. K. Chern, N. P. A. Neow, and M. H. Ang, Jr, "Practical issues in pixel-based autofocusing for machine vision," in *IEEE International Conference on Robotics and Automation*, 2001, vol. 3, pp. 2791–2796. 29

[80] J. Caviedes and F. Oberti, "A new sharpness metric based on local kurtosis, edge and energy information," *Signal Processing: Image Communication*, vol. 19, no. 2, pp. 147–161, 2004. 29

[81] G. Blanchet, L. Moisan, and B. Rougé, "Measuring the global phase coherence of an image," in *IEEE International Conference on Image Processing (ICIP 2008)*. IEEE, 2008, pp. 1176–1179. 29

[82] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3116–3132, 2010. 29, 32, 34, 35, 37

[83] Y. Tadmor and DJ Tolhurst, "Discrimination of changes in the second-order statistics of natural and synthetic images," *Vision Research*, vol. 34, no. 4, pp. 541–554, 1994. 29

[84] "Final report from the video quality experts group on the validation of objective models of video quality assessment," March 2000, http://www.vqeg.org/. 32, 33

[85] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Transactions on Image Processing*, , no. 99, pp. 1–1, 2010. 32

[86] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3440–3451, 2006. 32, 33

[87] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," *Information Fusion*, vol. 4, no. 4, pp. 259–280, 2003. 43

[88] H. Li, BS Manjunath, and S.K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical models and image processing*, vol. 57, no. 3, pp. 235–245, 1995. 43

[89] O. Rockinger, "Image sequence fusion using a shift-invariant wavelet transform," in *International Conference on Image Processing, 1997.* IEEE, 1997, vol. 3, pp. 288–291. 43

[90] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics letters*, vol. 38, no. 7, pp. 313–315, 2002. 44, 52

[91] CS Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000. 44, 52

[92] P. Wang and B. Liu, "A novel image fusion metric based on multi-scale analysis," in *International Conference on Signal Processing, ICSP 2008.* IEEE, 2008, pp. 965–968. 44, 53

[93] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *International Conference on Image Processing, ICIP 2003.* IEEE, 2003, vol. 3, pp. 170–173. 44, 52

[94] C. Yang, J.Q. Zhang, X.R. Wang, and X. Liu, "A novel similarity based quality metric for image fusion," *Information Fusion*, vol. 9, no. 2, pp. 156–160, 2008. 44, 52

[95] H. Chen and P.K. Varshney, "A human perception inspired quality metric for image fusion based on regional information," *Information fusion*, vol. 8, no. 2, pp. 193–207, 2007. 44, 52

[96] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 94–109, 2012. 44

[97] J. Puzicha, J.M. Buhmann, Y. Rubner, and C. Tomasi, "Empirical evaluation of dissimilarity measures for color and texture," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on.* IEEE, 1999, vol. 2, pp. 1165–1172. 44

[98] A. Toet and E.M. Franken, “Perceptual evaluation of different image fusion schemes,” *Displays*, vol. 24, no. 1, pp. 25–37, 2003. 44

[99] W. Huang and Z. Jing, “Evaluation of focus measures in multi-focus image fusion,” *Pattern Recognition Letters*, vol. 28, no. 4, pp. 493–500, 2007. 45

[100] R. Shen, I. Cheng, J. Shi, and A. Basu, “Generalized random walks for fusion of multi-exposure images,” *IEEE Transactions on Image Processing,*, vol. 20, no. 12, pp. 3634–3646, 2011. 45, 50

[101] B. Yang and S. Li, “Multifocus image fusion and restoration with sparse representation,” *Instrumentation and Measurement, IEEE Transactions on*, vol. 59, no. 4, pp. 884–892, 2010. 45

[102] R. Redondo, F. Šroubek, S. Fischer, and G. Cristóbal, “Multifocus image fusion using the log-gabor transform and a multisize windows technique,” *Information Fusion*, vol. 10, no. 2, pp. 163–171, 2009. 45

[103] T. Jinno and M. Okuda, “Multiple exposure fusion for high dynamic range image acquisition,” *Image Processing, IEEE Transactions on*, vol. 21, no. 1, pp. 358–365, 2012. 45

[104] R. Hassen, Z. Wang, and M. Salama, “Image sharpness assessment based on local phase coherence,” *IEEE Transactions on Image Processing, Accepted - to appear 2013*, 2013. 46

[105] Oliver Rockinger, “Image fusion toolbox,” 1999, http://www.metapix.de/toolbox.htm. 50

[106] Yin Chen and Rick S. Blum, “A new automated quality assessment algorithm for image fusion,” *Image and Vision Computing*, vol. 27, no. 10, pp. 1421 – 1432, 2009. 53

[107] University of California at Santa Barbara, “Vision research lab,” 2000, http://vision.ece.ucsb.edu/. 60

[108] University of Granada, “Computer vision group,” 2002, http://decsai.ugr.es/cvg/CG/base.html. 62

[109] P. Viola and W.M. Wells III, “Alignment by maximization of mutual information,” *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997. 67

[110] National Institute of Health National Library of Medicine, “Visible human project,” 2003, http://www.nlm.nih.gov/research/visible/visible_human.html. 69

[111] McGill University McConnell Brain Imaging Centre (BIC) of the Montreal Neurological Institute, “Brainweb: Simulated brain database,” 1997, http://www.bic.mni.mcgill.ca/brainweb/. 70

[112] Keith A. Johnson and J. Alex Becker, "The whole brain atlas," 1999, http://www.med.harvard.edu/AANLIB/home.html. 70