Probing Detector for Image Local Frequency Analysis

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

IN COMPUTER SCIENCE

UNIVERSITY OF REGINA

By

Bingyang Liu

Regina, Saskatchewan

June 2017

© Copyright 2017: B. Liu
Bingyang Liu, candidate for the degree of Master of Science in Computer Science, has presented a thesis titled, *Probing Detector for Image Local Frequency Analysis*, in an oral examination held on May 15, 2017. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner: Dr. Remus Floricel, Department of Mathematics and Statistics

Supervisor: Dr. Xue-Dong Yang, Department of Computer Science

Committee Member: Dr. David Gerhard, Department of Computer Science

Committee Member: Dr. Lisa Fan, Department of Computer Science

Chair of Defense: Dr. Allan Herman, Department of Mathematics and Statistics
Abstract

The Fourier transform is a century-old mathematical theory. Among many of its applications, it has been the foundation of signal processing since the middle of 20th century. Specifically, it has been hugely successful to perform voice recognition using frequency spectrum. However, its application in 2D image analysis is limited. This observation motivated the research presented in this thesis.

Most natural images contain non-stationary signals that require local image analysis techniques. The primary approach is the Short Time Fourier Transform (STFT), most notably the Gabor transform. A critical analysis presented in this thesis reveals that Gabor transform produces a spectrum with frequency resolution depending on the window width. In addition, the discrete sampling positions in the frequency domain are also varying due to the change of the window width. A novel technique, called probing detector, is proposed in this thesis to overcome these problems. The probing detector produces a spectrum with the highest possible frequency resolution and consistent discrete sampling positions in the frequency domain.
The key component of the new technique is a novel de-convolution algorithm that performs the computation directly in the domain in which convolution occurs. This provides an alternative and effective approach to solving de-convolution problems. A set of preliminary experiments was conducted to evaluate the accuracy of the Fourier coefficients reconstructed by the probing detector and to investigate its properties.
Acknowledgement

I would like to express my great appreciation to my supervisor Dr. Xue Dong Yang for his creative perspective and guidance that helped me complete my research. His patience and insight lead me to discover things on my own rather than step by step instructions to complete the research. During my graduate study, I learned how to think and solve problems when I meet new problems from my own perspective.

Along my way through the research, I was also guided by the more experienced friend Dr. Hong Xu Yang, who is a PhD candidate at TU/e & Philips. He helped me understand some background problems of my research in the beginning.

I would also like to thank my family for encouraging me to continue my education in Canada, especially my mother who gave me courage when I was in trouble. Last, the financial support of my supervisor Dr. Xue Dong Yang helped me get through my research life and kept focused on my research.
# Table of Contents

Abstract i

Acknowledgements iii

Table of Contents iv

List of Figures vi

List of Tables viii

Chapter 1 Introduction 1

  1.1 Motivation 1

  1.2 Problem of Local Frequency Analysis 5

  1.3 Contribution of the Thesis 8

  1.4 Structure of the thesis 11

Chapter 2 Related Previous Research 12

  2.1 Fundamentals of the Fourier Transform 12

  2.2 Applications of Fourier Transform 16

  2.3 Local Frequency Analysis 20

  2.4 Gabor Transform and Gabor Filters 26

Chapter 3 Probing and Convolution 29

  3.1 Early Idea of Local Image Probing 29

  3.2 Window Support Extension 30

  3.3 Convolution Theorem 35

  3.4 Previous De-Convolution Research 38

  3.5 De-Convolution Problem in Frequency Domain 41
**List of Figures**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Stationary and non-stationary signals</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Fourier transform of non-stationary signal</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Stationary image and its Fourier Domain</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>Non-stationary image and its Fourier Domain</td>
<td>4</td>
</tr>
<tr>
<td>1.5</td>
<td>Gaussian envelop</td>
<td>6</td>
</tr>
<tr>
<td>1.6</td>
<td>Region of interest and its Fourier transform</td>
<td>6</td>
</tr>
<tr>
<td>1.7</td>
<td>Local frequency distribution and its IFT</td>
<td>10</td>
</tr>
<tr>
<td>2.1</td>
<td>Continuous signal and its Fourier transform</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>Low-Pass filtering of an image</td>
<td>19</td>
</tr>
<tr>
<td>2.3</td>
<td>Periodic signal and its Fourier components</td>
<td>21</td>
</tr>
<tr>
<td>2.4</td>
<td>Non-periodic signal and its Fourier components</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>Relationship between Probing detector and Gabor transform</td>
<td>28</td>
</tr>
<tr>
<td>3.1</td>
<td>Gaussian window</td>
<td>31</td>
</tr>
<tr>
<td>3.2</td>
<td>Gaussian window with different width</td>
<td>32</td>
</tr>
<tr>
<td>3.3</td>
<td>Gaussian envelope / Probing</td>
<td>33</td>
</tr>
<tr>
<td>3.4</td>
<td>Interference introduced by the extended window support</td>
<td>34</td>
</tr>
<tr>
<td>3.5</td>
<td>Discrete convolution computation</td>
<td>36</td>
</tr>
<tr>
<td>3.6</td>
<td>Input image and result image after Laplace filtering</td>
<td>37</td>
</tr>
<tr>
<td>3.7</td>
<td>Blurred image and its deblurring results</td>
<td>40</td>
</tr>
<tr>
<td>3.8</td>
<td>Original spiral CT image and blind de-convolution image</td>
<td>41</td>
</tr>
<tr>
<td>3.9</td>
<td>Probing detector and its local frequency spectrum</td>
<td>43</td>
</tr>
</tbody>
</table>
4.1 Linear motion blurring long x direction 46
4.2 Visualization of the convolution process in 2-D 48
4.3 The convolution process in 1-D 49
4.4 Reconstruction frequency point and remain frequency spectrum 50
4.5 Illustration of the de-convolution process with a 2-D image 54
5.1 A testing example 61
5.2 The plot of Table 5.1 63
5.3 An example without clear periodical pattern 63
5.4 The plot of Table 5-2 65
5.5 Gaussian envelop with three different widths 69
5.6 Three regions with strong periodical patterns 72
5.7 A region with texture in more diversified directions 72
5.8 An office bookshelf 75
List of Tables

4.1 Peak values in the frequency spectrum 54

5.1 Mean-square-error of reconstructed images 62

5.2 Mean-square-errors for the example shown in Figure 5-3 64
Chapter 1

INTRODUCTION

Fourier analysis is one of the classic subjects in mathematics in the past century. Methods based on Fourier techniques have been widely used in science and engineering. Particularly, Fourier spectrum analysis is the fundamental theory and most frequently used tool in digital signal processing (DSP). For example, Fourier Transform is commonly used in speech recognition [2] by producing an alternative speech signal representation in the frequency domain, allowing more effective speech recognition through spectrum analysis. In astronomy, Fourier transform is also a basic technique for image restoration, such as removing specific sinusoidal interference [18].

1.1 Motivation

Fourier transform techniques have been successfully applied to 1-D signal processing, e.g. speech signal [1, 2]. One of the most widely used speech signal processing methods is Fast Fourier Transform (FFT). The formal tools for digital signal processing emerged in the middle of the 20th century with abilities to manipulate signals [2]. The essential part of traditional signal processing is to look at the signals decomposed in terms of sinusoidal basis functions with different frequencies.
For a stationary signal, each component of the signal appears throughout entire spatial domain (Figure 1-1 (a)), while the non-stationary signal, some components only appear in specific period (Figure 1-1 (b)).

![Figure 1-1](image)

(a) (b)

Figure 1-1 (a) A stationary signal exist in the entire spatial domain. (b) A non-stationary signal appears only in some specific time intervals.

The Fourier transform can decompose a signal into sinusoidal components with different frequencies. For stationary signals, all components sinusoidal waves are like signal in Figure 1-1 (a). For non-stationary signal varying along time, Fourier transform, when applied to the entire time domain, not only picks up the signal of Figure 1-1 (b), but also produces many interference noises as shown in Figure 1-2. To avoid such interference noises, windowed Fourier transform (known as Gabor transform [19] which will be reviewed in more detail in Section 2) was introduced.
Figure 1.2. (a) Fourier transform of Figure 1.1 (b).

Figure 1-2(b) The detail noises in (a)

In 2-D digital signal processing, namely image processing, Fourier analysis is also the core technique. The stationary signal can be presented as a regular pattern image, the texture of which occurs periodically, for example, Figure 1-3 (a) and its Fourier transform shown in (b). On the other hand, a non-stationary signal is a pattern that only appears in a sub-region in an image, for example, Figure 1-4 (a) and its Fourier transform in (b). In fact, most
natural images can be considered as a non-stationary signal as the content varying along the spatial coordinate.

Figure 1-3. Stationary image and its Fourier Domain

Figure 1-4. Non-stationary image and its Fourier Domain
1.2 Problem of Local Frequency Analysis

The Fourier transformation of a non-stationary pattern, such as the basket texture shown in the middle of bookshelf background in Figure 1-4, is mixed with the Fourier transform of the bookshelf. The standard Fourier transform presents the global frequency spectrum of an image without any reference to local regions in the spatial domain. In other words, every sub-region has contributions to the frequencies related to its pattern, and every frequency receives contributions from one or more sub-regions. There are many advantages to do image analysis and processing in frequency domain v.s. spatial domain. However, in many situations, analysis localized in a spatial region is required, especially in the non-stationary situation. There is a need to separate contributions from different sub-regions. The traditional approaches include the Short-Time Fourier Transform (STFT) [20] and the 2-D Gabor transform [21]. Please note that the Gabor transform is a special case of STFT in which a Gaussian window is used.

Many applications will benefit from analysis of Fourier spectrum of local image regions. For example, in Content-Based Image Retrieval (CBIR), if we are concerned with regions of specific interested texture, the standard global Fourier transform cannot satisfy the requirements. As mentioned above, STFT and Gabor transform can be used for local region analysis. In past decades, the Gabor transform has been developed into Gabor filters for texture analysis. In [3], the Gabor wavelet features are used for texture analysis in
image retrieval. However, the Gabor filter can only extract a specific range of frequency information in an entire spatial domain. For local region analysis, what we need is the entire frequency spectrum in the specially designated region of interest in the spatial domain.

![Gaussian envelop](image1-5.png)

Figure 1-5. Gaussian envelop

![Region of interest and its Fourier transform](image1-6.png)

Figure 1-6. Region of interest and its Fourier transform
For example, suppose we are interested in the central regular texture region. By applying the Gaussian envelope (shown in Figure 1-5) to the image Figure 1-4 (a), we obtain a region of interest Figure 1-6 (a). The Figure 1-6 (b) shows the Fourier transform of the image Figure 1-6 (a), i.e. the Fourier transform of the region of interest. Essentially, the Gaussian envelope ignores the contents outside the region of interest. However, the Gaussian envelope itself introduces additional spectral energy.

Precisely, let

- the Figure 1-4 (a) be the function \( f(x, y) \)
- the Gaussian envelope be the function \( g(x, y) \)
- the Fourier Transform of \( f(x, y) \) be \( F(u, v) \)
- the Fourier Transform of \( g(x, y) \) be \( G(u, v) \)

then the Figure 1-6 (a) is the product of \( f(x, y) \) and \( g(x, y) \), i.e.

\[
h(x, y) = f(x, y) \cdot g(x, y)
\]

Let the Fourier Transform of \( h(x, y) \) be \( H(u, v) \). According to the Convolution Theorem [22],

\[
H(u, v) = F(u, v) \otimes G(u, v)
\]

where ‘\( \otimes \)’ is the convolution operator.
Because the contributions of the region of interest and the Gaussian envelope are mixed together through convolution in the frequency domain, it is desirable to remove the contribution from the Gaussian envelope, such that a “clean” Fourier spectrum from the contents in the region of interest can be obtained for frequency analysis. For general situations, e.g. Figure 1-4 (b), it may be difficult to separate the contributions from the complex background from the global Fourier transform in order to obtain the contributions only from the central texture region. However, the Gaussian envelope is a simple regular shape, therefore it is feasible to remove its “contamination” from convolution result, e.g. Figure 1-6 (b).

1.3 Contributions of the Thesis

The result of local frequency analysis mentioned above, convoluted with Gaussian envelope, and cannot illustrate the original frequency spectrum distribution. In this thesis, our research focused on how to remove the Gaussian envelope interference to obtain an estimated local frequency spectrum.

Convolution is a frequently used mathematic operation in image processing, such as image blurring, edge detecting, segmentation, etc. In image processing, convolution is easy to be operated with a filter. However, if an image is convoluted with a filter, it is difficult to do the inverse process, i.e. removing the impact due to the filter. Here, we would like to define
the Gaussian envelope as a Point Spread Function (PSF) [23]. A PSF describe the response caused by the image system to the point source, the local frequency distribution.

The process of removing the PSF response can be considered as de-convolution operation. In image restoration, there are two types of situations related to convolution. One is the convolution process happened in the spatial domain, for example, blurring and filtering. The other is when the convolution process happened in the frequency domain, such as the problem described in Section 1.2. Several algorithms have been devised recently to estimate PSF using machine learning techniques and to remove its response in the spatial domain [15]. Those techniques for image restoration can, however, not be applied to the complex number in the frequency domain. In this thesis, we propose a novel de-convolution technique in the frequency domain. Our primary contribution is the design and implementation of the algorithm to de-convolute the Gaussian envelope interference from the frequency spectrum of a partial region in the spatial domain. The Gaussian envelope interference can be removed from the local frequency distribution, iteratively, one at a time, in an approximate order of magnitude. This method was inspired by the traditional motion de-blurring algorithm.

For local region analysis, we are only concerned with the region of interest. In previous research, STFT and Gabor feature have been widely used for analysis of local regions.
However, these methods have limitations in local frequency analysis. Our method can be considered as an extension of Gabor transform. This method can use the simple shape of Gaussian envelope to emphasize the contents of a local region and at the same time, obtain a full and consistent physical frequency spectrum (a new concept that will be described in detail in Chapter 3). By removing the Gaussian envelope contribution in the mixed frequency spectrum, an estimated local frequency distribution can be obtained.

Inverse Fourier Transform (IFT) has been applied to the estimated local frequency distribution. Figure 1-7 (a) shows the estimated local frequency distribution for Figure 1-6 (b) and the Figure 1-7 (b) is the image generated by IFT.

![Figure 1-7 Local frequency distribution and its IFT](image)
This experimental result demonstrates clearly that the main texture information of the local region of interest has been accurately captured (detailed accuracy analysis will be presented in Chapter 5).

1.4 Structure of the Thesis

In the first chapter, we have briefly presented the one of most frequently used techniques in image processing, the Fourier transform, as well as some of its limitations. The motivation for the use of a Gaussian envelope is detailed and its interference in Fourier domain. Chapter 2 will go into greater details with a deep review of the Fourier transform and signal analysis, put focus on the pros and cons of the Gabor transform, as they are the most relevant to this topic of research. In Chapter 3, the general solution of de-convolution algorithm will be reviewed. Previous research of de-convolution and their advantages, as well as disadvantages, will be discussed. In Chapter 4, inspired by traditional motion de-blurring algorithm, a novel de-convolution algorithm in the Fourier domain is proposed. Implementation details will be covered as well. In Chapter 5, the experimental image results and its properties will be discussed in detail. Chapter 6 will conclude with a summary of the algorithm and future direction along this algorithm.
Chapter 2
Related Previous Research

In mathematics, the Fourier transform is an operation that can transform one function to another format of expression. The fundamentals of the Fourier transform and spectrum analysis can be found in textbooks of signal/image processing. In this chapter, we will review the necessary background information of standard Fourier theorem based on the book of Digital Image Processing [4].

2.1 Fundamentals of the Fourier Transform

For any continuous and integrable function \( f(x) \) represented along the axis \( x \), the Fourier Transform can be defined as:

\[
\mathcal{F}\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} \, dx
\]  

(2-1)

where \( j = \sqrt{-1} \) and \( x \) and \( u \) are presumed to be infinite. Given a Fourier transform \( F(u) \), we can obtain the original function \( f(x) \) by Inverse Fourier Transform

\[
\mathcal{F}^{-1}\{F(u)\} = f(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} \, du
\]  

(2-2)

In the two equations above, Fourier transforms are mappings between functions in two different domains, namely \( x \) (usually referred to as time or spatial domain) and \( u \) (as frequency domain). The essential purpose of Fourier transform in signal/image processing is that the analysis of a signal can be conducted in either the time domain or the frequency
domain, or combined. However, such an analysis can be more effective and/or efficient in one domain than the other. For example, in speech recognition, the voice signal \( f(x) \) varies along the time axis \( x \). Fourier transform operation transforms signal representation from the time domain to the frequency domain, such that the energy distribution in the frequency spectrum will facilitate effective voice recognition.

![Continuous signal](a)

![Fourier transform of signal (a)](b)

Figure 2-1 (a) Continuous signal. (b) Fourier transform of signal (a).

Figure 2-1 gives an example of a one-dimensional non-periodic continuous signal (a) and its Fourier transform (b). It should be noted that, the Fourier transform generates a new function, \( F(u) \), in complex space. The original function \( f(x) \) can be either in real space or complex space. A complex number can be represented as \( R + jI \) where \( j = \sqrt{-1} \), \( R \) is
the real part and \( jI \) is the imaginary part. In complex plane, the two axes are real axis and imaginary axis respectively and any complex number can be represented as a point \((R, I)\) in this complex plane. This representation can also be written in polar coordinate in terms of the distance from the original to the point and the angle between the positive real axis and the complex number vector.

\[
r(r\cos\theta + jsin\theta)
\]
while \( r = \sqrt{R^2 + I^2} \) and \( tan\theta = I/R \). The above trigonometric form of complex number can be equivalently represented in complex exponential form (called Euler’s formula):

\[
e^{j\theta} = cos\theta + jsin\theta
\]
When Euler’s Formula is used for Fourier transform, it is clear that signals are decomposed into different frequency sinusoidal waves and the polar coordinate form provides explicitly the phase angle and magnitude information of the sinusoidal functions.

For a 2-D signal such as an image, the spatial domain contains two variables. If we consider the function \( f(x, y) \) to be continuous and integrable, Fourier transform and inverse Fourier transform are defined similarly by the following equations:

\[
\mathcal{F}\{f(x, y)\} = F(u, v) = \int_{-\infty}^{\infty} f(x, y)e^{-j2\pi(ux+vy)} \, dx \, dy
\]
(2-3)

\[
\mathcal{F}^{-1}\{F(u, v)\} = f(x, y) = \int_{-\infty}^{\infty} F(u, v)e^{j2\pi(ux+vy)} \, du \, dv
\]
(2-4)
For modern digital computers, signals are represented in discrete forms based on some sampling frequency [6]. Therefore, we need the Fourier transforms in discrete form as the following:

\[
F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)e^{-j2\pi(ux/M+vy/N)}
\]

\[
f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v)e^{j2\pi(ux/M+vy/N)}
\]

Fourier transform maps a function in the spatial domain into a complex function in the frequency domain. As the Euler’s Formula presented above, a sinusoidal wave can be converted into polar coordinates to remark the magnitude and phase angle information. Therefore, in the frequency domain, the frequency spectrum contains two types of information, i.e. magnitude and phase angle. Since the magnitude directly shows the energy of the signal at a specific frequency, it has been considered as a more important feature of Fourier transform by most researchers. However, the phase angle is a much less studied aspect of Fourier transform, but could contain other valuable information (e.g. [24]).

Similarly, for 2-D image processing, the Fourier transform can decompose an image into complex exponential functions of different frequencies. After the transformation, the representation of an image is mapped from the spatial domain to the frequency domain. Each point in the frequency domain represents the energy at a specific frequency in the
image signal and the Fourier spectrum describes the energy distribution at all frequencies for the image, which are therefore important information in image analysis [4, 5]. In the frequency domain, as we can see from Figure 1-2 and 1-3, the magnitudes of the Fourier coefficients in the low frequency range are much bigger than that in the high frequency range. Therefore, the low frequency region contains major energy of the image in the frequency spectrum [5]. On the other hand, although the high frequency region contains less energy, that, however, corresponds to fine detail of the image contents, such as edges or boundaries of objects. Therefore, it could be much more important than the low frequency region for image analysis, such as texture extraction. Fourier transform is significant in signal and image analysis because many problems are difficult, and sometimes even not feasible, to be solved in the time or spatial domain, but could be effectively and/or efficiently solved in the frequency domain [25]. Fourier transform has other useful properties, such as separability, translation, rotation, etc., which will not be discussed here.

2.2 Applications of Fourier Transform

Fourier transform has been used in a wide range of areas, such as voice recognition, identification of characteristic mechanical vibration frequency, image restoration and enhancement, texture classification, etc. [8]. Discrete Fourier Transform (DFT) is a direct way to perform Fourier transform because digital images are discrete samples of
continuous functions [7]. In practice, however, DFT is not commonly considered because of its high computational complexity. Instead, the FFT ([8, 9]) is used in most applications.

Voice signal analysis is a typical example of 1-D signal processing, and Fourier transform has naturally been applied to speech analysis. The pronunciation of each tone has a unique pattern in the frequency spectrum. The length of a voice signal could be very long and sometimes even may not have an ending in certain real-time applications. But, DFT and FFT can only be applied to finite sampled signals. So, in voice analysis, researchers have employed Hanning window to obtain sampled signals within a specified time interval. Hanning window is designed for random and unknown signals with an objective to reduce frequency leakage [2]. When Fourier transform is applied along with Hanning window, the frequency distribution allows clear extraction of the main feature components while depressing irrelevant noise. The article [1] has given an example of speech analysis for vowel O. Since Fourier transform based techniques can effectively distinguish the feature components of different tones, it has been widely used in Automatic Speech Recognition (ASR), an increasingly important application in this digital era.

In image processing, certain operations are much easier to be done in the frequency domain than the spatial domain. For example, the periodic interference noise generated during image acquisition in [10], like sinusoidal wave, is very difficult to be removed in the spatial
domain. However, this type of noises are distinctive points in the frequency domain, thus can be easily removed [6].

Most image processing systems have a preprocessing stage to reduce noise generated in image acquisition or transmission stages. For example, before edge detection or image segmentation, we would like to reduce randomly distributed sharp noise. This is typically done by low-pass filtering. This can be illustrated by the sequence of images shown in Figure 2-2 below. Given the noisy image shown in Figure 2-2 (a), its Fourier transform is given in (b). If we keep the lower frequency part, i.e. the coefficients within the ring, and ignore the other information (as shown in (c)), we can obtain a blurred image (d) by inverse Fourier transform.

It can be noticed in the low-pass filtered image that although the random noise is reduced, the detailed image contents, such as edges, are also blurred. In addition, many ghost shadow rings are generated. There are several different low-pass filters, e.g. Gaussian bandwidth filter, which performs much better, particularly creating much less ghost shadow rings.
Figure 2-2. Low-Pass filtering of an image. (a) Original Image. (b) Fourier Spectrum of the image. (c) Low-Pass within a radius of 40 pixels. (d) Inverse Fourier Transform.

In medical imaging, Fourier-based techniques allow practitioners to analyze Free Induction Decay (FID) signals in restoration of Magnetic Resonance Images [8]. The work of [12] uses Fourier transform to obtain frequency spectrums and applies Gabor filters to
perform Shape-From-Texture analysis. In image compression, DCT is a fundamental component in the international standards developed by the Joint Photographic Expert Group (JPEG) and the Moving Picture Experts Group (MPEG) [11]. In JPEG, an image is partitioned into 8X8 blocks without overlapping. DCT is then applied to each block to obtain a new representation in the frequency domain. Because human vision is less sensitive to distortion in fine detailed contents that correspond to energy in the high frequency range. Therefore, high compression ratio can be achieved by reducing data in part and this operation can be efficiently performed in the frequency domain.

2.3 Local Frequency Analysis

In Fourier transform, the signals are multiplied by sinusoidal functions with frequency varied from 0 (DC component) to the highest frequency value, then do integration over all the times. The result of the integration shows the energy of the specific frequency component, which represent magnitude.

For stationary signals, the frequency components exist through the entire duration of the signal. For example, Figure 2-3 (a) is a periodic signal in which all frequency components exist along the entire time intervals. Fourier transform can obtain the main frequency response of the stationary signal (e.g. as shown in Figure 2-3 (b)) [6].
Contrary to the stationary signal, the frequency components of a non-stationary signal may appear only in portions of the entire period. Frequency Modulation (FM) signal, widely used in radio communication, is a good example of non-stationary signal in which the frequency of the signal varies along the time [26]. Figure 2-4 (a) shows a simple non-stationary signal, in which each one of the four time intervals contains a simple sinusoidal
signal but with different frequencies from one interval to the next. Its Fourier transform is shown in Figure 2-4 (b). It can be noticed that the main frequency components of it are same as that in Figure 2-3.

The above examples demonstrate that Fourier transform can find whether a certain frequency component exists or not in both stationary and non-stationary cases. However,
when it comes to non-stationary signals, especially if a user would be interested in what frequency components occur at a designated time period, the standard Fourier transform is not capable of answering this question.

STFT is a classical method for dealing with non-stationary signal. In order to perform local analysis, the time variable is introduced into the frequency analysis. STFT assumed that portions of the non-stationary signal are stationary [6]. In other words, each main frequency component in the Fourier transform must be stationary in some subinterval of the time. A window function needs to be applied to obtain a signal on a certain time interval. The window should be narrow enough such that the part of the signal within the window can be considered to be stationary. Therefore, a signal can be divided into several small segments and the Fourier transform will be applied to each segment. The result will be a representation containing not only the frequency information, but also the associated time intervals, therefore suitable for time-frequency analysis.

The window shape and width of STFT are usually fixed when scanning through a signal. However, the width of the window may vary in order to cover different scope of intervals. The change of the window width will cause inconsistent physical frequency spectrums. In order to understand this issue clearly, let us review the definition of forward and inverse 1-D DFTs:
\[ F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x)e^{-j2\pi ux/M} \]  

(2-7)

\[ f(x) = \sum_{u=0}^{M-1} F(u)e^{j2\pi ux/M} \]  

(2-8)

When applying a window to a digital signal, the number of samples is the first factor it affects. The number of samples of the signal, i.e. the dimension of the discrete data, will be decreased by the window. Therefore, the value of \( M \) is decreased, and consequently, the range of \( u \) is also decreased. In other words, the frequency resolution will be indirectly affected. For example, suppose a 1-D array of length 1,024 stores the discrete samples of a signal for a certain period of time and a window of size 32 is used. When DFT is applied to the entire period, the resolution of frequency is 512. But the resolution of frequency from the samples within the window will only be 16. In the extreme case that the window contains only two digital samples, the frequency domain will show the DC component and the highest frequency component.

In addition to the resolution issue in the frequency domain, the discrete sampling positions in the frequency domain are also varying due to the changes of the window width. For example, if the window width \( M \) equals to 5, the discrete frequency values will take the following values:
If the window width $M$ is changed to 7, the discrete frequency values will be changed to the following values:

<table>
<thead>
<tr>
<th>$u = 0$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2\pi ux/M = 2\pi x * 0/7$</td>
<td>$2\pi x * 1/7$</td>
<td>$2\pi x * 2/7$</td>
<td>$2\pi x * 3/7$</td>
<td>$2\pi x * 4/7$</td>
</tr>
</tbody>
</table>

This implies that the energy at a particular frequency in the first case cannot simply be mapped to the energy(s) at one or more frequencies in the second case, and vice versa. There might (or perhaps should) be a relationship between the two. It is, however to our best knowledge, not known at the moment. We will refer this issue as the inconsistent frequency problem related to the window width variation.

The window function is clearly helpful so that STFT provides an approach for local region frequency analysis [13]. In STFT, the window will slide over the entire duration of signal. However, in this research, our concern is primarily in designated local regions of interest, rather than the entire image. The basic idea of STFT is the use of window function for local region analysis that can be introduced to our analysis. As we discussed before, the problem of varying frequency resolution associated with STFT will cause inconsistency between the frequency components obtained from different window widths. In time-frequency analysis, in order to find the accurate time position of a non-stationary signal
component, it is desirable to use a narrow window. But this will cause the frequency resolution to be low correspondently. It will also cause more overlapping which will be explained in Chapter 5. On the other hand, if a wider window is used in order to obtain a better frequency resolution, we may not be able to tell at which time intervals a certain frequency component occurs, and at the same time, the condition of stationary could be violated [6]. Therefore, the choice of the window size is application dependent and varies.

2.4 Gabor Transform and Gabor Filters

The Gaussian function is a popular choice of the window. In this case, STFT is called Gabor transform. Fourier transform. Commonly, Gabor transform is used as Gabor filter that applies the Gaussian window to the sinusoidal basis and then be applied to the image [27]. The following equation shows Gabor filters in 2-D.

\[ g(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} \]  
\[ f(x, y) = g(x, y)e^{-j2\pi(ux+vy)} \]

For digital signals, Gabor filters are considered as a tool for edge detection. For a specific frequency of the basis, Gabor filter will obtain a designated frequency component for the entire spatial domain. However, for a local region, we want to analyze all frequency components, therefore Gabor transform is more suitable. However, a problem associated with Gabor transform is that the frequency resolution is determined by the window width.

In this research, the objective is to obtain a full and consistent physical frequency spectrum
for a local region. We propose a new method, called probing detector, to generate such a spectrum. To understand the difference and relationship between the proposed probing detector and Gabor filter, let us examine the spatial-frequency space as illustrated in Figure 2-3. Each Gabor filter is based on a single basis frequency. After it is slid across the entire spatial domain, we obtain information along a horizontal line in the spatial-frequency space. On the other hand, our probing detector will obtain information at a specific local region, but cover the entire frequency domain, i.e. a vertical line in the spatial-frequency space.

If we slide the probing detector across the entire spatial domain, the sum of all the frequency spectrums obtained from different spatial locations will be a spectrum covering the entire frequency domain and each frequency component will contain contributions from entire spatial space, thus equivalent to the standard Fourier transform. One might wonder what information would be obtained if we vary the basis frequency from low to high for Gabor filter. We hypothesize that the total energy in the result image obtained by Gabor filter with a basis frequency $u$ will be proportional to the total energy from a ring of radius $u$ in the standard Fourier Transform spectrum.

Furthermore, what is the relationship between the proposed probing detector and Gabor transform? First, both of them will cover the entire frequency domain. But, the difference
is the following. The probing detector will produce a spectrum that has the same frequency resolution as that by the standard Fourier transform. On the other hand, Gabor transform produces a spectrum with much lower frequency resolution depending on the window size. Furthermore, if the window size varies, the frequency resolution varies accordingly. In summary, the objective of the probing detector is to produce a spectrum with the highest frequency resolution and consistent discrete sampling positions in the frequency domain, in contrast to Gabor transform that produces a spectrum with low frequency resolution and inconsistent discrete sampling positions in the frequency domain.

![Figure 2-5. Relationship between Probing detector and Gabor transform](image)

The new probing detector proposed by this research is introduced in detail in Chapters 3 and 4, and the experimental results are presented in Chapter 5.
Chapter 3

Probing and Convolution

The common choice of the window function of STFT is a Gaussian-like function, and Gabor transform is a special case of STFT. As discussed in Chapter 2, Gabor transformation suffers from its limitation in frequency resolution and consistency. To overcome these limitations, we propose an extension to the window support function.

3.1 Early Idea of Local Image Probing

Our early idea of local image probing is quite simple. Given an image, if we modify the values of pixels in a local region of interest, the difference between the global Fourier transforms of the modified image and the original image should reveal the specific information contained in that region. One simple modification method is to amplify values of the pixels in a local region. If we amplify the values by a constant value for pixels only within the region, sharp discontinuities are clearly introduced around the boundary of the local region in the image, therefore, generating much interference noise in the frequency spectrum.

To reduce the interference noise, the Gaussian function came to our mind naturally. Let the modification function be:
\[ h(x, y) = 1 + g(x, y) \]  

(3-1)

where \( g(x, y) \) is the Gaussian function defined in Eq. (2-7). If \( f(x, y) \) is an image under consideration, the modified image will be:

\[
f'(x, y) = f(x, y)h(x, y)
\]

\[
= f(x, y)(1 + g(x, y))
\]

\[
= f(x, y) + f(x, y)g(x, y)
\]

(3-2)

The difference of the global Fourier transforms of the modified image and the original image is:

\[
D(u, v) = \mathcal{F}\{f'(x, y)\} - \mathcal{F}\{f(x, y)\}
\]

\[
= \mathcal{F}\{f(x, y) + f(x, y)g(x, y)\} - \mathcal{F}\{f(x, y)\}
\]

\[
= \mathcal{F}\{f(x, y)\} + \mathcal{F}\{f(x, y)g(x, y)\} - \mathcal{F}\{f(x, y)\}
\]

\[
= \mathcal{F}\{f(x, y)g(x, y)\}
\]

(3-3)

We realized that this probing idea is very similar to Gabor transform. However, is there any difference between the two? If they are different, what is the difference? This motivated us to find such difference and formally re-define our probing technique.

### 3.2 Window Support Extension

In order to obtain a frequency spectrum at a full resolution and with consistent discrete sampling positions for a region of interest, but free of interference from information outside, we can select the width of the Gaussian window according to the size of the local...
region of interest and position the Gaussian window over on the local region, but extend the support to the entire time domain.

Specifically, assume a discrete 1-D signal has N samples. That is, the time domain is [0, N-1]. Figure 3-1 illustrates the concept of window support extension. Let \( \sigma \) to be the standard deviation of the Gaussian function. It is commonly accepted that the value of Gaussian function beyond the range \([-3\sigma, 3\sigma]\) is small enough, such that it can be considered as zero. Therefore, in practice, the support of a Gaussian window is usually set in the range \([-k\sigma, k\sigma]\) where \( k \) is a suitably selected constant. In Gabor transformation, the window support will be \([a, b]\) showed in Figure 3-1. Now, we extend the window support to the entire time range \([0, N-1]\).

![Figure 3-1. Gaussian Window](image)

It is not difficult to see that the frequency spectrum produced under the extended window support will have a full resolution with consistent sampling positions in the frequency
domain. First, the Fourier transform is applied to entire domain \([0, N-1]\), the frequency spectrum will be at the full resolution. Second, if the width of Gaussian function changes (i.e. \(\sigma\) varies as shown in Figure 3-2), the frequency resolution and the discrete sampling positions in the frequency domain remain to be consistent:

\[
\begin{array}{cccccccc}
    u & 0 & 1 & 2 & 3 & \ldots & N-2 & N-1 \\
    2\pi xu/N = & 2\pi x0/N & 2\pi x1/N & 2\pi x2/N & 2\pi x3/N & \ldots & 2\pi x(N-2)/N & 2\pi x(N-1)/N \\
\end{array}
\]

Figure 3-2. A Gaussian window with a different width.

A 2-Dimensional Gaussian function is illustrated in Figure 3-3. The part within the range \([-k\sigma, k\sigma]\), for a suitably selected constant \(k\), in both vertical and horizontal directions is referred as the Gaussian envelop. For an image to be analyzed, the center of the Gaussian envelope will be placed at the middle of the region of interest, while other parts are set to zero to avoid interference from information outside of the region of interest. As long as it extends to a common range, the frequency inconsistent problem is resolved. Theoretically, the extension could even go beyond the original image size.
Specifically, let $f(x, y)$ be the image to be analyzed, and $g(x, y, x_0, y_0)$ be the 2-Dimensional Gaussian function centered at the pixel $(x_0, y_0)$:

$$g(x, y, x_0, y_0) = \frac{1}{\sqrt{2\pi \sigma}} e^{\left(\frac{-(x-x_0)^2+(y-y_0)^2}{2\sigma^2}\right)}$$  (3-4)

By varying $(x_0, y_0)$, the Gaussian envelope can be shifted around the image. The standard Fourier transform will be applied at each stop of the Gaussian envelope:

$$F(u, v, x_0, y_0) = \mathcal{F}\{ f(x, y)g(x, y, x_0, y_0) \}$$  (3-5)

We refer this process as probing, and the Gaussian envelope as the probing detector.

Although the above extended window support solves the low frequency resolution and the inconsistent frequency sampling problem associated with previous methods, the shape of the window, unfortunately, introduces severe interference in the frequency spectrum. Figure 3-4 (a) gives an example of local frequency distribution. We assume that it is the ideal local region frequency spectrum that we want to obtain. In the probing process, it
will be interfered by the frequency spectrum of Gaussian envelope. The resulting frequency spectrum will be like that shown in (b).

![Figure 3-4. Interference introduced by the extended window support](image)

Therefore, when the Gaussian envelope is applied to a local region of an image in the spatial domain, the frequency spectrum of the local region is “contaminated”. Fortunately, there is a relationship between the spatial domain and frequency domain. The probing detector is applied to an image in the spatial domain through multiplication, i.e., \( f(x, y)g(x, y, x_0, y_0) \). According to the Convolution Theorem (to be discussed in detail in the next section), its Fourier transform in the frequency domain will be the convolution between the Fourier transform of \( f(x, y) \) and the Fourier transform of \( g(x, y, x_0, y_0) \) in the frequency domain. In order to restore the local frequency spectrum, the focus of this research is to remove the interference from the Gaussian envelop through de-convolution.
3.3 Convolution Theorem

Convolution is the mathematical approach of combining two signals into a third one. In mathematics, convolution is basically an integral operation and it is essential in image processing for filtering operation, such as image smoothing, sharpening, edge detection, etc. The inverse process of convolution, called de-convolution, is sometimes also useful in image processing, e.g. image de-blurring.

Formally, convolution is defined as [5]:

\[
h(x) = f(x) \otimes g(x) = \int f(t)g(x - t)dt
\]

(3-6)

where \( g(x) \) is usually referred as the kernel function. In image and signal processing, convolution is a common operation used for image filtering [16]. The image can be blurred, sharpened or noise reduced by using appropriate filters, such as Gaussian filter, Laplace filter, etc. In digital signal and image processing, the data are in discrete form; therefore, convolution in discrete form is used in computation:

\[
h(x) = \sum_{i=-\infty}^{\infty} f(i) \cdot g(x - i)
\]

(3-7)

Let us take image sharpening as an example. Assume that \( f(x, y) \) is the original image and \( g(x,y) \) is a sharpening kernel (e.g. the Laplace operator). The sharpening process can be modeled as:

\[
h(x, y) = f(x, y) \otimes g(x, y)
\]

(3-8)
The convolution kernel is usually smaller than the image size. Computationally, the kernel is shifted across the image and convoluted with every pixel. The essential operation of convolution is translation of the kernel function and superimposition of the kernel over a sub-region of the image (as shown in Figure 3-5). The value of an element in the convolution result can also be considered as a weighted average of the sub-region where the kernel provides the weighting coefficients. Figure 3-6 shows an example image of sharpening.

![Figure 3-5. Discrete convolution computation](image)
Let \( f(x) \) and \( g(x) \) be two functions in the spatial domain, and \( F(u) \) and \( G(u) \) the Fourier transforms of them in the frequency domain respectively. The Convolution Theorem [5] states that:

\[
\begin{align*}
 f(x) \otimes g(x) & \leftrightarrow F(u)G(u) \\
 f(x)g(x) & \leftrightarrow F(u) \otimes G(u)
\end{align*}
\]

That is, the product of two functions in the spatial domain corresponds to the convolution of between the Fourier transforms of the two functions, respectively, in the frequency domain, and vice versa. This implies that convolution computation in one domain can be achieved equivalently in the other domain. This property has been utilized in image processing frequently for improving computational efficiency.
3.4 Previous De-Convolution Research

In image processing, much research has been done on de-convolution in past decades. In [14, 15], the well-known blind de-convolution methods have been reviewed, such as simulated annealing (SA), nonnegativity and support constraints recursive inverse filtering (NASRIF). The convolution kernel is often called Point Spread Function (PSF) in de-convolution [28]. Since every pixel in new image generated by a convolution process is a weighted average of the neighboring pixels of the original image, de-convolution is difficult to achieve in the spatial domain. The traditional strategy is to perform de-convolution in the frequency domain based on the Convolution Theorem. Because a convolution in the spatial domain corresponds to a product in the frequency domain, de-convolution in the spatial domain could be performed equivalently in the frequency domain through a division operation.

Let us take the image restoration from a blurred image as an example. Let the original image be \( f(x, y) \) and the observed degraded image be \( d(x, y) \). Assume also that the blurring kernel \( g(x, y) \) is known. According to the Convolution Theorem, we have the following relationships:

\[
\begin{align*}
    d(x, y) &= f(x, y) \otimes g(x, y) \quad (3-10) \\
    D(u, v) &= F(u, v)G(u, v) \quad (3-11)
\end{align*}
\]
The goal is to find a $H(u, v)$, the inverse filter of $G(u, v)$, such that the original scene can be reconstructed from the observed degraded image:

$$D'(u, v) = \frac{F(u, v)G(u, v)}{H(u, v)}$$  \hspace{1cm} (3-12)

It appears to be quite simple that, by just letting $H(u, v) = G(u, v)$, then $D'(u, v) = F(u, v)$, and inverse Fourier transform applied to $D'(u, v)$ will produce the original image $f(x, y)$. In practice, however, it is difficult to obtain the optimal recovered image from this equation due primarily to two reasons: (1) the noise effect, measured as $N(u, v)$, in the observed image; and (2) small values in $H(u, v)$ used in division could generate very large result values that could, and indeed will, amplify noise significantly. The Wiener de-convolution algorithm [6] included a noise term in the model:

$$D(u, v) = F(u, v)G(u, v) + N(u, v)$$  \hspace{1cm} (3-13)

and introduced an improved inverse filter:

$$H(u, v) = \frac{1}{G(u, v)} \left( \frac{G(u, v)^2}{G(u, v)^2 + \frac{N(u, v)}{S(u, v)}} \right)$$  \hspace{1cm} (3-14)

Where $S(u,v)$ is the mean power spectral density of the original signal. In this expression, the noise effect is modeled as the noise-to-signal ratio. Figure 3-7 illustrates a corrupted image (a) due to motion blurring, with additional noise, the recovered image (b) by inverse filtering, and the recovered image (c) by Wiener filter.
In practice, the PSF may not be known in most situations. Several algorithms have attempted to estimate PSF known as blind de-convolution. The articles [14, 15] provide a comprehensive overview of the principles of blind de-convolution algorithms. Estimating PSF is an empirical method. For the linear degradation model, it is possible to estimate some likely PSFs then choose the best one.

Richardson-Lucy algorithm is an iterative algorithm, widely used in blind de-convolution process. Maximum a posteriori (MAP) is a static method based on Bayes Theorem, and has generated ideal PSF estimations. From the previous research, the recovered image based on blind de-convolution can be quite sharp. In medical imaging, de-convolution
algorithms have been widely applied to enhancement of CT images. Figure 3-8 shows an original spiral CT image (a) and the reconstructed image (b) using de-convolution.

![Figure 3-8. Original spiral CT image and blind de-convolution image. [15]](image)

In this research, the PSF is a Gaussian function, thus it is known. Therefore, our algorithm will not be based on ideas in the blind de-convolution process, thus the details of blind de-convolution methods will not be reviewed here.

### 3.5 De-Convolution Problem in Frequency Domain

The previous research on de-convolution deals primarily with the situations where the convolution process happens in the spatial domain, and de-convolution operation is performed in the frequency domain by simple divisions if an inverse kernel is known or can be estimated. The problem in this research is the opposite, i.e. the convolution process happens in the frequency domain due to the product between the image \( f(x, y) \) and the Gaussian envelope \( g(x, y) \) in the spatial domain. Suppose, based the principle of previous
research, which we try to achieve de-convolution in the frequency domain equivalently through division in the spatial domain:

\[
\frac{f(x, y)g(x, y)}{g(x, y)} = f(x, y)
\]  

(3-15)

then, we go back to the original image. Keep in mind that the Gaussian envelope was used to selected local regions and should not be canceled in local image analysis. That means we must do de-convolution in the frequency domain directly with a novel approach. The problem is further complicated by the fact that the convolution is between \(F(u, v)\) and \(G(u, v)\), both of which are complex valued functions.

Figure 3-9 shows an example of the problem to be solved in this research. Image (a) is the original image \(f(x, y)\); (b) is a window image \(g(x, y)\) with the probing detector located at a low-right position; (c) is the product of (a) and (b); and (d) is the Fourier transform of (c).
Figure 3-9. (a) the original image; (b) probing detector for a local region; (c) probing detector applied to the image; and (d) the frequency spectrum of the image (c).

As seen from the Figure 3-9 (d), the spectrum has been severely affected by the Gaussian envelope through the convolution process in the frequency domain. The energy distribution in the frequency domain for the region of interest has been badly “contaminated”. This research focuses on developing a reconstruction algorithm that will extract coefficients of $F(u, v)$ directly in the frequency domain, iteratively in the order of magnitudes.

Before moving onto the next chapter to present our new algorithm, it is worth to examine if the traditional Gabor transform would also suffer from the problem of “contamination”.
The probing detector is basically a revised version of the window function of Gabor transform. The difference is that, the window support of Gabor transform is limited to the Gaussian envelope, while the probing detector extends the support to the full spatial dimension. Nevertheless, the results from Gabor transform should also be the convolution result between the local image contents within the window support and the Gaussian envelope. To make this statement more clearly, it can be compared with the situation where a simple rectangular function (with the same width of a Gaussian envelope) is used as the window function. When Fourier transform is performed within this window, the result will be based only on the local image contents within the window support. Although the “contamination” problem exists in Gabor transform, it is, however, much lower than that generated from the probing detector due to the extended window support. On the other hand, Gabor transform suffers from the limitations of varying and low frequency resolutions, and inconsistent frequency sampling positions. This motivated us to solve the problem introduced in this research.
Chapter 4

De-Convolution in the Frequency Domain

For de-convolution in the spatial domain, there may not exist a model for direct analytical solution in the spatial domain. However, the Convolution Theorem converts this problem into a division problem in the frequency domain. Many existing de-convolution approaches utilize also Machine Learning algorithms [15, 28, 29]. However, the convolution problem introduced by the probing detector is in the frequency domain, further complicated by the complex valued functions in the frequency domain. Therefore, the problem that we attempt to solve is a challenging and novel problem.

4.1 An Inspiration to Our New Approach

In searching for a solution to the problem of de-convolution in the frequency domain, an existing algorithm for solving a different problem came to our mind almost “accidently” that inspired us to develop the novel approach proposed in this thesis.

In image restoration, motion blurring is a common problem. It happens when taking photos with a shaking camera during long time exposure. The blurring problem caused by uniform linear motion can be solved by the inverse filtering algorithms as reviewed in the previous chapter [6]. Interestingly, there existed also a numerical solution to perform motion de-
blurring directly in the spatial domain [4]. The principle of this algorithm is based on the following convolution model. Let $T$ stand for duration of the exposure, we have

$$g(x, y) = \int_0^T f[x - x_0(t), y - y_0(t)] dt$$

(4-1)

where $g(x, y)$ is the blurred image. Every pixel in the blurred image is the integration of several pixels of the original image. Without loss of generality, let us assume the motion is along the $x$ direction.

$$f \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot$$

\[\begin{array}{c}
\downarrow \\
g \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\
\end{array}\]

The idea is to consider differentiation along $x$ direction:

$$g'(x) = f(x) - f(x - a)$$

(4-3)

where $a$ is the motion length.

$$f(x) = g'(x) + f(x - a)$$

(4-4)

This equation may be expressed in the form

$$f(x) = \sum_{k=0}^{m} g'(x - ka) + \varphi(x - ma)$$

(4-5)

Where $\varphi(x)$ is the estimation of the first corresponding pixel located between 0 and $a$. If $\varphi(x)$ can be estimated, then the estimation of $f(x)$ can be calculated. The essential thought
of the algorithm is that if we can obtain pixel values for one column, it will be simple to obtain the pixel values of the next column in the image. Repeating this process will lead to a de-blurred image. Inspired by this thought of de-convolution process in the spatial domain where the convolution occurred, we developed an algorithm to de-convolute the Gaussian interference in the frequency domain caused by the probing detector.

4.2 Analysis of the Convolution Problem due to Probing Detector

With the extension of Gaussian window support, the convolution interference in the frequency domain has a profound impact to the spectrum. Following the idea of the algorithm reviewed in the previous section, it is helpful to examine in more detail the basic mathematical convolution process in the frequency domain.

According to the Convolution Theorem, the multiplication between the image, \( f(x, y) \), and the Gaussian envelope, \( g(x, y) \), in the spatial domain, corresponds to the convolution between the Fourier transform of the image, \( F(u, v) \), and the Fourier transform of the Gaussian envelope, \( G(u, v) \), in the frequency domain. This convolution process can be visualized in Figure 4-2. The discrete Fourier transform of the image is shown on the left where each impulse represents a Fourier coefficient \( F(u, v) \). The Fourier transform of the Gaussian envelope, \( G(u, v) \), is shown in the middle. The convolution result, \( H(u, v) \), is
illustrated on the right in which a copy of \( G(u, v) \) is placed at each \((u, v)\) location and modulated by \( F(u, v) \).

This convolution process can be visualized more clearly in 1-D situation as shown in Figure 4-3. Figure 4-3 (a) is the Fourier transform of Gaussian envelope \( G(u) \); (b) is the Fourier transform of a signal \( F(u) \), and (c) is the convolution, \( H(u) \), between \( F(u) \) and \( G(u) \).

Please note that the convolution kernel \( G(u) \) is known (still a Gaussian function). Given an \( H(u) \), the objective is to reconstruct \( F(u) \), for each \( u \). Our strategy is to estimate \( F(u) \)'s in the order of their magnitudes. Initially, we set \( F(u) \) to empty. We search \( H(u) \) to find the location \( u \) where the magnitude of \( H(u) \) is the largest. We can consider it as an estimate of \( F(u) \) at this point. So, we store the estimated value into \( F(u) \) at the position \( u \) (as shown in Figure 4-4 (a)). For the purpose of de-convolution, we use the peak value as a scalar to modulate \( G(u) \). By subtracting the scaled convolution kernel from the Figure 4-3 (c), the remaining \( H(u) \) is illustrated in Figure 3-4 (b).

Figure 4-2. Visualization of the convolution process in 2-D
Figure 4-3. The convolution process in 1-D
By repeating the above process with the next largest value in H(u), the next F(u) will be found. The iteration can continue till no value in H(u) is greater than a pre-specified magnitude.

It can be observed from Figure 4-3 (c) that, the Gaussian functions overlaps generally in the convolution result. The degree of overlapping is related to the width of the Gaussian function. The wider the width, the more of overlapping the convolution result has. It
appears that this overlapping effect may affect the accuracy of the peak points being found, in terms of both position and the estimated magnitude. This is particularly true for smaller peaks that can be buried by the tails of larger neighbors. On the other hand, however, as the larger components are successively removed, smaller peaks emerge clearly. This issue will be discussed in more detail in the next chapter. As we can see from Figure 4-3 (c), the second or third peak value is far more less than the first peak value.

Figure 4-5 illustrates the newly proposed iteratively de-convolution process by a 2-D image example. Figure 4-5 (a) is the original image and (b) is a probing detector. Figure 4-5 (c) is the Fourier transform of the product of (a) and (b), thus the “contaminated” spectrum. The first peak is, without any surprise, found at the original \((u, v) = (0, 0)\), corresponding to the DC term \(F(0, 0)\). Figure 4-5 (d) shows the first reconstructed point in \(F(u, v)\). By removing this term from the spectrum, the remaining spectrum is shown in (e). It should be pointed out that (e) is an intensity enhanced image; otherwise it would be too dark.

After ten iterations, the reconstructed Fourier transform \(F(u, v)\) is shown in Figure 4-5 (f), and the remaining spectrum in (g) (again, with an enhanced intensity for illustration). Figures (h) and (i) show the results after twenty iterations respectively.
Figure 4-5. Illustration of the de-convolution process with a 2D image.

<table>
<thead>
<tr>
<th>Frequency Spectrum</th>
<th>Peak Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Frequency Spectrum</td>
<td>0.00143099</td>
</tr>
<tr>
<td>After 1 iteration</td>
<td>0.000129201</td>
</tr>
<tr>
<td>After 2 iterations</td>
<td>0.000129196</td>
</tr>
<tr>
<td>After 10 iterations</td>
<td>0.000102169</td>
</tr>
<tr>
<td>After 20 iterations</td>
<td>0.0000783225</td>
</tr>
</tbody>
</table>

Table 4-1: Peak values in the frequency spectrum

Table 4-1 shows the magnitudes of the peak points after various numbers of iterations.

From this table, we can see that the magnitude of the first peak is significantly larger than the remaining peaks; therefore, it can be picked up reliably. Overall, the peak points observed in (c) are reasonably separated, thus overlapping problem is not a big concern.
It is interesting to see that the second and the third peaks have nearly identical magnitudes. They are actually symmetrical points: one at (39, -37) and the other at (-37, 39). These two positions are far away from each other such that Gaussian function will not affect each other. That means the overlap of the convolution will not affect much for the value estimation. Those points in the spectrum have symmetry property that could improve search time for finding peak points.

4.3 The De-Convolution Algorithm

Based on our further empirical studies, we observed that: (1) the frequency spectrum of local image regions contains mostly sparse peak points; and (2) the gaps between nearby peaks are sufficiently large. It suggests that the above relatively simple estimation method for finding peak points in this preliminary investigation of a new de-convolution algorithm is very promising. This algorithm is outlined below.

De-Convolution Algorithm:

Input: image $f(x, y)$, probing location $(x_0, y_0)$, $\sigma$ (the width of Gaussian envelope)

Output: reconstructed Fourier transform $F(u, v)$ for the local region of interest.

1. Initialize $F(u, v)$;

2. Generate the probing function $g(x, y)$ based on $(x_0, y_0)$ and $\sigma$;

3. Compute the product image $h(x, y) = f(x, y)g(x, y)$;

4. Compute the Fourier transform: $H(u, v) = \mathcal{F}[h(x, y)]$;
5. Compute the Fourier transform: \( G(u, v) = \mathcal{F}(g(x, y)) \);

6. Loop:
   a. Search the peak point in \( H(u, v) \);
   b. Scale \( G(u, v) \) by the magnitude of \( H(u, v) \);
   c. Remove the scaled \( G(u, v) \) from \( H(u, v) \);
   d. Store the peak point into the result \( F(u, v) \);

7. Until the magnitude of the peak in the remaining \( H(u, v) \) is smaller than a pre-specified \( \varepsilon \).

8. Return \( F(u, v) \).

The width of the Gaussian envelope is required by this algorithm and it can be a user specified parameter in applications. The choice of this value will have impact to the result. The properties and proper selection of this parameter is discussed in the next chapter.
In order to study the properties of the newly proposed probing detector and evaluate its performance, a set of experiments was conducted and the preliminary results are presented in this chapter.

5.1 Quality of the Reconstructed Fourier Coefficients

In the newly proposed de-convolution algorithm, the coefficients of $F(u, v)$ are estimated successively in the order of their magnitudes. However, what is the accuracy of the estimated coefficients? It is necessary to evaluate the quality of them. Since the true values of $F(u, v)$ are not available, our evaluation is indirectly done by applying inverse Fourier transform to the reconstructed $F(u, v)$, and measure the mean-square-error between the original image and the reconstructed image.

Figure 5-1 shows an example that contains strong periodical patterns. Figure 5-1 (a) is the original color image, and (b) is a gray-scale version where the red square box indicates the region of interest. Figure 5-1 (c) shows the probing detector with the Gaussian envelope centered at the region of interest. The product of (b) and (c) is given in (d), and the Fourier transform of (d) is shown in (e).
The new de-convolution algorithm is applied to (e) iteratively to reconstruct the coefficients of $F(u, v)$. After ten iterations, ten coefficients are estimated (as shown in (f)). Please note that image (f) displays only the positions of the ten coefficients, and their intensity are not proportional to the magnitudes of the coefficients. If intensity were used to display their magnitudes, smaller coefficients would be very dark even after linear or non-linear enhancement. By applying inverse Fourier transform to (f), a reconstructed image is obtained (as shown in (g)). Please note that the content in the region of interest has been duplicated to the entire image, but our comparison should only be limited to the region of interest. It is not difficult for a viewer to see that there are differences in the region of interest between the reconstructed and the original images.

If the iteration continues, we can obtain 20, 30, 40, 50, 60, 70, 80, 90, and 100 coefficients in total successively. Figures 5-1 (h) to (k) show the reconstructed images corresponding to 20, 30, 40, and 50 coefficients respectively. The viewer can observe that the difference between the reconstructed and the original images becomes smaller and smaller. Indeed, the mean-square-errors for the content within the region of interest between the original and each of the reconstructed images are calculated and presented in Table 5-1 and depicted in Figure 5-2.
(a) The original color image
(b) A gray-scale version
(c) The probing detector
(d) The product of (b) and (c)
(e) The “contaminated” spectrum

(f) Reconstructed $F(u, v)$ after 10 iterations

(g) Reconstructed image from (f)

(h) Reconstructed image with 20 coefficients
Figure 5-1. A testing example
<table>
<thead>
<tr>
<th>Estimated coefficients</th>
<th>Mean-square-error</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.6202</td>
</tr>
<tr>
<td>20</td>
<td>0.99147</td>
</tr>
<tr>
<td>30</td>
<td>0.88599</td>
</tr>
<tr>
<td>40</td>
<td>0.84762</td>
</tr>
<tr>
<td>50</td>
<td>0.830662</td>
</tr>
<tr>
<td>60</td>
<td>0.80602</td>
</tr>
<tr>
<td>70</td>
<td>0.81754</td>
</tr>
<tr>
<td>80</td>
<td>0.79169</td>
</tr>
<tr>
<td>90</td>
<td>0.768513</td>
</tr>
<tr>
<td>100</td>
<td>0.74653</td>
</tr>
<tr>
<td>110</td>
<td>0.728768</td>
</tr>
<tr>
<td>150</td>
<td>0.666302</td>
</tr>
<tr>
<td>200</td>
<td>0.608994</td>
</tr>
</tbody>
</table>

Table 5-1. Mean-square-error of reconstructed region of interest with different number of estimated coefficients
The second example is a region without a clear periodical pattern. The original image is shown in Figure 5-3 with the region of interest highlighted by the red box. The same experiments as the above were conducted on this region. Table 5-2 shows the mean-square-error and Figure 5-3 and Figure 5-4 is the plot of Table 5-2.
<table>
<thead>
<tr>
<th>Estimated coefficients</th>
<th>Mean-square-error</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.44587</td>
</tr>
<tr>
<td>20</td>
<td>1.23286</td>
</tr>
<tr>
<td>30</td>
<td>1.07875</td>
</tr>
<tr>
<td>40</td>
<td>0.967601</td>
</tr>
<tr>
<td>50</td>
<td>0.874708</td>
</tr>
<tr>
<td>60</td>
<td>0.793742</td>
</tr>
<tr>
<td>70</td>
<td>0.735185</td>
</tr>
<tr>
<td>80</td>
<td>0.671765</td>
</tr>
<tr>
<td>90</td>
<td>0.641158</td>
</tr>
<tr>
<td>100</td>
<td>0.611848</td>
</tr>
<tr>
<td>110</td>
<td>0.580257</td>
</tr>
<tr>
<td>150</td>
<td>0.501445</td>
</tr>
<tr>
<td>200</td>
<td>0.448993</td>
</tr>
</tbody>
</table>

Table 5-2. Mean-square-errors for the example shown in Figure 5-3
Based on these two experiments, we observe the following properties.

(a) The content in the region of interest can be reconstructed from the estimated coefficients with varying visual quality and numerical accuracy.

(b) As the number of estimated coefficients increases, both the visual quality and the numerical accuracy increase.

(c) The algorithm converges after 50 iterations. In other words, further estimated coefficients will only make very small improvement in both the visual quality and numerical accuracy.
The mean-square-error per pixel with 50 estimated coefficients is less than 0.9 with respect to the gray-scale range [0, 255]. That means the average numerical error per pixel is less that 0.1%.

5.2 Effect of the Width of Gaussian Envelope

As it has been briefly discussed in Chapter 2, the width of Gaussian envelope is an important issue in image local analysis. In the spatial domain, the width should not be larger than the length of the underlying non-stationary signal. On the other hand, the width should not be too small in order to generate a meaningful spectrum. Now, we examine the effect of varying width of the Gaussian envelope in the frequency domain, specifically its impact to the convolution problem.

The scaling property of Fourier transform [30] indicates that:

\[ \mathcal{F}\{g(cx)\} = \frac{1}{|c|} G \left( \frac{u}{c} \right) \]  

(5-1)

where \( c \) is a scaling factor. In other words, linear scaling in the spatial domain is reflected in an inverse scaling in the frequency domain.

Figures 5-5 (a), (b) and (c) show respectively the probing detector with three different widths, 9, 15 and 21, for the Gaussian envelope. The Fourier transforms of these three probing detectors are shown in Figures 5-4 (d), (e) and (f) respectively. It can be clearly
seen that a narrower Gaussian envelope in the spatial domain generates a wider convolution kernel in the frequency domain, and vice versa. Furthermore, it can be easily observed that a wider convolution kernel causes more overlapping in the spectrum.

It appears that heavier overlapping due to convolution might cause difficulty to the deconvolution process. Our experiments suggest that this is not a concern. The reason is that our algorithm estimates the coefficients of $F(u, v)$ in the order of their magnitude. Once a coefficient is obtained, its contribution is removed from the spectrum, thus the spectrum becomes “cleaner”. It should also be noted that the very first peak coefficient, $F(0, 0)$ (i.e. DC term) that is significantly larger than remaining coefficients, does not require any imagination.
(a) Width = 9
(b) Width = 15
(c) Width = 21
(d) Fourier transform of (a)
In conclusion, we recommend that the choice of the width for the Gaussian envelope should primarily be determined by the size of the region of interest in the spatial domain.

5.3 Potential Applications

Even though the coefficients estimated by the proposed algorithm can reconstruct images with very high accuracy, image encoding/decoding is not a primary objective of this thesis. The main objective of this research is the development of a direct de-convolution algorithm performed in the frequency domain. Our preliminary experimental results have validated the new algorithm.
Now a frequency spectrum for a local image region is available with unprecedented clarity and consistent physical frequency in the highest possible resolution, what are potential applications of it? It is well known that Fourier spectrum has played a fundamental role in voice recognition and is hugely successful. But, why it is much less useful, to our best knowledge, in 2D image analysis so far? Would the newly proposed probing detector lay a foundation for such applications? This question may be too big to be answered by this thesis. However, we would like to test the probing detector with several different types of local image contents and make some preliminary observations.

Figure 5-6 illustrates three examples in which strong periodical patterns exist. In the left image of Figure 5-6 (a) a region (highlighted in red box) containing texture mainly in the vertical direction is selected; (b) selects a region containing texture in both diagonal directions; and (c) selects also a region with texture similar to (b), but in bigger size. Their estimated frequency spectrums are shown on the right respectively. As expected, all three spectrums exhibit strong clustering of energy in the spectrum. In the cases of (b) and (c), the two clusters exhibit a similar structure.
Figure 5-6. Three regions with strong periodical patterns

Figure 5-7 shows an example containing also relatively clear texture, but in slightly diversified directions (as highlighted in red box in (a). Its frequency spectrum (shown in (b)) exhibits some clustering, however, less concentrated.

Figure 5-7. A region with texture in more diversified directions
Figure 5-8 is an example of natural photo (an office bookshelf). In the cases of (a), (b) and (c), the probing detector is placed in a region containing magazines, but at different locations respectively. The frequency spectrum in each case is shown on the right. It is clear that all of them exhibit strong clustering either in vertical or horizontal directions. The region selected in case (d) is on the plain wall. However, due to the illumination, the intensity varies slowly and smoothly across the region. Its frequency spectrum is more limited in central area. In other word, the energy is more concentrated in low frequency area. Finally, in the case (e), the region selected is also on the plain wall, but the intensity is almost uniform. Its frequency spectrum is almost totally dark, except a few values near the center.
Different image contents produce different frequency spectrums. From the few examples shown above, we may suggest that the structure of clustered distributions in the frequency spectrum can potentially form a new feature space as a foundation for investigation of new effective and efficient techniques for image analysis and pattern recognition.
Chapter 6
Summary and Future Directions

6.1 Summary

Fourier transform is a century-old mathematical theory. Among many of its applications, it has been the foundation of signal processing since the middle of 20th century. In this modern information technology era, it still plays a pivotal role in many digital technologies, such encoding/decoding of audio signals, and still/motion pictures, and signal and image analysis, etc. Specifically, it has been hugely successful to perform voice recognition using frequency spectrum produced by Fourier transform. However, its application in 2D image analysis is much limited in comparison with that for 1D signal processing. This observation motivated the research presented in this thesis.

Most natural images contain non-stationary signals that require local image analysis techniques. Local image analysis has been an active research subject in image processing for many decades. The primary approach is the windowed Fourier transform (or called STFT), most notably the Gabor transform. Gabor transform produces a spectrum with frequency resolution depending on the window size. If the window size varies, the frequency resolution varies accordingly. In addition, the discrete sampling positions in the frequency domain are also varying due to the changes of the window width. We refer this
issue as the inconsistent frequency problem related to the window width variation. A novel technique, called probing detector proposed in this thesis that overcomes these problems.

The main contributions of this thesis are the following.

a) A critical analysis is presented on the varying resolution and frequency inconsistency issues related to the existing STFT techniques, the well-known Gabor transform in particular.

b) A novel technique - probing detector - is proposed that produces a spectrum with the highest frequency resolution and consistent discrete sampling positions in the frequency domain.

c) For convolution problems occur in the spatial domain, the existing de-convolution techniques performs computation in its dual domain - the frequency domain. This thesis addresses the convolution problem occurring in the frequency domain which cannot be solved by the traditional technique. A novel de-convolution algorithm is developed that performs the computation directly in the domain in which convolution occurs. This provides an alternative and effective approach to solving de-convolution problems.

d) A concept of frequency-spatial space (Figure 2-5) is introduced. Local image transformations can be represented as points or lines in this space. It allows
visualization of relationships between different local transformations, as well as relationships between local and global transformations.

e) A set of preliminary experiments was conducted that demonstrates that the newly proposed probing detector can produce estimated Fourier coefficients for a local image region progressively in the order of their magnitudes. The local image reconstructed from the estimated Fourier coefficients has a mean-square-error less than 0.3%.

f) The estimated Fourier coefficients for local image regions have an unprecedented clarity. Several testing results show that strong clustering can be clearly observed in their frequency spectrum.

6.2 Future Research
Local image descriptors [31] have been an active research subject in recent years. Majority of them are feature vectors constructed from the gradients (the first derivative) and/or the zero-crossing (the second derivative) computed in the spatial domain. It is desirable that such descriptors are translation, rotation and scaling invariant. Among these three properties, scaling invariant appears to be the most challenging one. Multi-resolution image pyramid has been a common approach to deal with the scaling problem. A limitation of this approach is that the scales are only available at a limited number of discrete levels; and computation often involves a search through the pyramid either from top to bottom or
from bottom to top. A descriptor that captures the geometric and/or texture information and be invariant in arbitrary and continuous scales, to our best knowledge, remains to be discovered.

The Fourier coefficients generated by the probing detector for a local image region contains full information about the content of the local image region. Therefore, it can be a basis for constructing a general purpose local image descriptor, in contrast to many existing descriptors that are heuristic approaches based on selected features. More importantly, the coefficients in low frequency range provide a coarse description of the image content; and the coefficients in higher frequency ranges provide incremental refinement information of the image content in a progressive or adaptive manner. The Fourier spectrum generated from the probing detector offer a uniquely new and promising opportunity to construct a local image descriptor that is truly translation, rotation and scaling invariant. Successful development of such a new local image descriptor will have numerous applications, for example, texture classification, pattern recognition, image segmentation, contents based image retrieval, etc.

The new de-convolution algorithm presented in this thesis provides an alternative, possibly more effective, approach to solving image restoration problems that can be modeled by a convolution process.
The new de-convolution algorithm is a numerical solution. Further refinement may improve the accuracy of the estimated Fourier coefficients in terms of magnitude and peak position. In addition, computation of Fourier spectrum for local image regions using the probing detector is currently time consuming. Implementation of probing detector on GPU can improve its performance, which is particularly beneficial for real-time applications.
References


