Efficient Stockwell Transform with Applications

to Image Processing

by

Yanwei Wang

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Applied Mathematics

Waterloo, Ontario, Canada, 2011

© Yanwei Wang 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Multiresolution analysis (MRA) has fairly recently become important, and even essential, to image processing and signal analysis, and is thus having a growing impact on image and signal related areas. As one of the most famous family members of the MRA, the wavelet transform (WT) has demonstrated itself in numerous successful applications in various fields, and become one of the most powerful tools in the fields of image processing and signal analysis.

Due to the fact that only the scale information is supplied in WT, the applications using the wavelet transform may be limited when the absolutely-referenced frequency and phase information are required. The Stockwell transform (ST) is a recently proposed multiresolution transform that supplies the absolutely-referenced frequency and phase information. However, the ST redundantly doubles the dimension of the original data set. Because of this redundancy, use of the ST is computationally expensive and even infeasible on some large size data sets. Thus, I propose the use of the discrete orthonormal Stockwell transform (DOST), a non-redundant version of ST.

This thesis will continue to implement the theoretical research on the DOST and elaborate on some of our successful applications using the DOST. We uncover the fast calculation mechanism of the DOST using an equivalent matrix form that we discovered. We also highlight applications of the DOST in image compression and image restoration, and analyze the global and local translation properties. The local nature of the DOST suggests that it could be used in many other local applications.

Acknowledgments

Contents

List of Figures

x

List of Tables

List of Abbreviations

MRA	Multiresolution Analysis
FT	Fourier Transform
IFT	Inverse Fourier transform
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
IFFT	Inverse Fast Fourier Transform
STFT	Short Time Fourier Transform
GT	Gabor Transform
WT	Wavelets Transform
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
ST	Stockwell Transform
DST	Discrete Stockwell Transform
DOST	Discrete Orthonormal Stockwell Transform
FDOST	Fast Discrete Orthonormal Stockwell Transform
PSNR	Peak Signal to Noise Ratio
TV	Total Variation
FOV	Field of View
MRI	Magnetic Resonance Imaging

Chapter 1

Introduction

In signal and image processing, the Fourier transform (FT) [5] is commonly used to decompose a signal into its frequency components. Explicitly, the FT of a one-dimensional function, $h(t) \in \mathbf{L}^1(\mathbb{R})$, is defined [1] as

$$H(f) = \mathcal{F}\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-i2\pi ft}dt, \qquad (1.1)$$

where $i^2 = -1$.

The inverse Fourier transform (IFT) of $H(f)$ is defined as

$$h(t) = \mathcal{F}^{-1}\{H(f)\} = \int_{-\infty}^{\infty} H(f)e^{i2\pi ft}df. \qquad (1.2)$$

The FT offers the convenience to study and modify the signal in a different manner – frequency space (also known as $k-$space in some application areas, especially in medical imaging) [18, 7, 23]. But the global property of the FT – that each sample affects every Fourier coefficient (and vice versa) – makes it unfavorable in applications where local information is preferred (e.g. signal denoising and

---

[1]Other equivalent definitions are available for the pair of FT and IFT with the possible modulation factor, $1/2\pi$ or $1/\sqrt{2\pi}$, in the exponential. Additional factors in front of the integral may arise on both the forward and inverse definitions.

compression). For instance, when denoising a signal with useful information in both high frequencies and low frequencies, if the noise is only localized within a certain region, the FT would be incapable of separating the noise from the high frequency information. This issue can be explained by the following well-know Fourier uncertainty principle [4], which is derived from the Heisenberg uncertainty principle [29] from Quantum Mechanics.

To elaborate the principle, we define the term, deviation, of a function $g(x)$ as

$$\triangle_a g = \frac{\int_{-\infty}^{\infty} (\lambda - a)^2 |g(\lambda)|^2 d\lambda}{\int_{-\infty}^{\infty} |g(\lambda)|^2 d\lambda}. \tag{1.3}$$

Then the Fourier uncertainty principle states that:

Theorem 1.0.1. *(Fourier uncertainty principle)*

Suppose f is a function in $\mathbf{L}^1(\mathbb{R})$. Then

$$\triangle_a f \cdot \triangle_\alpha \hat{f} \geq \frac{1}{4}, \tag{1.4}$$

for all points $a \in \mathbb{R}$ and $\alpha \in \mathbb{R}$.

As is manifested by the uncertainty principle, due to the perfect localization of Fourier transform in frequency domain, the information in the time or space domain has been entirely smeared into all Fourier coefficient. Tiny deviations of the Fourier coefficients could cause huge deviations of the time component. That is to say, a function in real world can never be both band-limited (compact in Fourier domain) and time-limited (compact in time domain).

In order to resolve this global issue one may use the short-time Fourier transform (STFT) [31], such as Gabor transform (GT) [19], defined as

$$\mathbf{STFT}\{x(t)\} \equiv X(\tau, f) = \int_{-\infty}^{\infty} x(t)w(t - \tau)e^{-ift}\, dt, \tag{1.5}$$

2

where $w(t)$ is the predefined window function. The STFT offers the way to calculate the spectrum localized by the window function and has been demonstrated to be viable in various fields for applications [19, 1]. However, besides the absence of the perfect reconstruction algorithm in general, the manually defined window size has put another significant barrier in applications using the STFT. We use a chirp signal in Figure 1.1 (a) as an example and intuitively show how the GT decomposes the signal into temporal-frequency domain. From its filled contour plot in Figure 1.1 (b), we can see that the frequency is increasing with respect to the precession of the signal. However, the horizontal width of the substantial coefficients band, which illustrates the resolution of the corresponding frequency, remains the same for all frequency components. As well known by the sampling theorem (Chapter 2), a higher frequency requires more resolution to pursue a flexible manipulation or to avoid the aliasing phenomenon during real applications. On the other hand, it would be redundant to put excessive resolution for the low frequency component.

As will be elaborated in the next chapter, based on the multiresuolution analysis (MRA), the wavelet transform (WT) [15] has successfully overcome the shortcomings of the STFT mentioned above by applying local decomposition filters to a signal on multiple scales. Normally, the continuous wavelet transform (CWT) for a continuous-domain input $h(t) \in \mathbf{L^2}(\mathbb{R})$ is defined as the integral

$$W(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} h(t)\psi\left(\frac{t - \tau}{s}\right) dt, \qquad (1.6)$$

where $\psi(t)$, called the mother wavelet, is a continuous-domain function of both the time and the scale; τ is the translation factor and s is the scale factor. By convention, some discrete versions of wavelet are used in applications. For example, the Daubechies wavelet [14, 15] of order K is defined by the conditions that the

3

mother wavelet satisfies

$$\int x^k \psi(x)dx = 0, \quad 0 \le k \le K - 1. \tag{1.7}$$

Each specific wavelet (in terms of different K) has a number of zero moments or vanishing moments equal to half the number of coefficients, $2K$, which are normally involved in various wavelet applications.

The upsampling and downsampling algorithms [4] are available in applying the discrete wavelet transform to applications with a computational complexity of $\mathcal{O}(N)$, where N is the size of the input. However, the self-similarity constraint among the wavelet basis functions destroys the phase information, so that the coefficients will only supply locally-referenced scale information. Most of the wavelet transforms, which have the complexity of $\mathcal{O}(N)$, will end up with compact basis functions, which cause a perfect localization in time or space domain. While using these wavelets to decompose the input, the overlap in the frequency domain becomes non-avoidable. So, even though the term "scale" can be approximately interpreted as "frequency" due to its ability in adjusting the size of the basis function, there is no straightforward way to turn this scale information into proper frequency information.

In response to this restriction, the Stockwell transform (ST, sometimes called the S-transform) [39] was published in 1996. The ST is a time-frequency decomposition that offers absolutely-referenced frequency and phase information (i.e. the phase information is referenced to time $t = 0$) [17, 26, 27, 39]. Sharing the same frame of definition with other integral transforms discussed above, but with a different kernel function, the Stockwell transform of $h(t) \in \mathbf{L^1}(\mathbb{R})$ is defined as

$$S(\tau, f) = \mathcal{S}\{h(t)\} = \int_{-\infty}^{\infty} h(t)\frac{|f|}{\sqrt{2\pi}}e^{-\frac{(\tau-t)^2 f^2}{2}}e^{-i2\pi ft}dt, \tag{1.8}$$

4

where f is the frequency, and t and τ are both time variables. The ST decomposes a signal into both temporal (τ) and frequency (f) components. The Gaussian part inside the integral acts as a frequency sensitive window function, which creates a comparably narrow window for large values of f (high frequency), and a relatively wide window for small values of f (low frequency). The value of τ represents the center of the window function, and thus, by exhausting all possible values for τ, the ST coefficients cover the whole temporal axis and create the full resolutions for each designated frequency.

Moreover, considering the integral property of the Gaussian function,

$$\frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} e^{-\frac{x^2 f^2}{2}} \, dx = \frac{\sqrt{2\pi}}{|f|}, \tag{1.9}$$

the accumulation over all the Stockwell coefficients for a certain value of f will recover the corresponding Fourier coefficients,

$$\int_{-\infty}^{\infty} S(\tau, f) d\tau = H(f), \tag{1.10}$$

highlighting a special feature of the ST and its close relation with the FT.

For application convenience, the discretized Stockwell transform (DST) can be achieved from its continuous version and will consistently maintain the temporal-frequency nature of the ST (see Chapter 2 for more elaboration).

As such, Figure 1.1 (c) gives the filled contour plot of the 2-D ST coefficients of the same 1-D chirp signal in Figure 1.1 (a). When we compare Figure 1.1 (c) with Figure 1.1 (b), we can see that the ST offers more substantially non-zero coefficients (the dark blue pixel represents zero values) at higher frequency location, while the GT always supplies the same substantially non-zero coefficients for low frequency and high frequency. Again, see Chapter 2 for more theoretical details about their comparison.

The obvious shortcoming of the ST can be discovered immediately by its definition – redundancy. In (1.8), we can see that for each specified t value, the Stockwell coefficients over all the possible f values will be calculated, which requires a huge amount of calculation time and storage for transforming even a moderate size signal into its DST coefficients. For a signal of length N, the DST generates N^2 coefficients. The computational complexity to generate these coefficients is $\mathcal{O}(N^2 \log N)$ by taking advantage of the fast Fourier transform (FFT). This has become the main obstacle preventing the Stockwell transform from being applied to larger size images or higher dimension data sets.

To combat the redundancy issue of the ST and maintain its advantages, in these scenarios, a suitable non-overlapping partition strategy is applied on the time-frequency domain. Consequently, higher frequencies will have more partitions than lower frequencies. For example, the DC frequency will have fewer partitions than higher frequencies, yielding a total of N sub-regions for the whole time-frequency domain. Parameters and basis functions are defined corresponding to each sub-region, and yield the discrete orthonormal Stockwell transform (DOST) [35]. The dot-product between the input signal and DOST basis functions gives a brute force way to calculate the DOST coefficients. Compared to the ST, the DOST transform has successfully kept its multiresolution nature and the absolutely-referenced frequency and phase information by reducing the computational complexity to $\mathcal{O}(N^2)$. Still, compared to the original frequency analysis tool, the FT, which has a complexity of $\mathcal{O}(N \log N)$, the computation of the DOST is still expensive for large signals, such as audio processing and remote sensing, and higher dimensional data sets, such as medical imaging and volumetric imaging. A fast algorithm to compute the DOST coefficients based on the proposed matrix expression of the DOST

transform is presented in Chapter 4 as the Fast DOST (FDOST) [45]. Details about how the time-frequency domain is partitioned and more analysis on its time-frequency properties will be elaborated in Chapter 3.

1.1   Research Motivation and Objectives

Considering the many successful applications using the Fourier transform, the Gabor transform, and the wavelet transform, we are interested in studying another multiresolution facility, the Stockwell transform and its discrete orthonormal version, the DOST. As I continued to spend an increasing amount of time in this topic during the past three years, I focused on solidifying the theoretical integrity of the newly invented DOST and on mining more reasonable applications using the DOST, such as image compression and etc.

1.2   Thesis Organization

As a starting point of my thesis, in Chapter 2 we provide a brief review of the multiresolution analysis of various transforms, such as the Gabor transform, the wavelet transform and the Stockwell transform. Considering that the redundancy and computational complexity of the Stockwell transform are still significant, in Chapter 3, we propose a partition strategy adopted by the DOST and pursue a detailed theoretical analysis of the DOST design. Besides, the negative frequency parameters have been appropriately chosen to achieve the conjugate symmetry for a real input signal. An alternative symmetric version of the DOST is also delineated to show the freedom of defining the DOST with re-arranged parameters. By reasonably varying the parameters, which can maintain the

orthogonality, the Nyquist criterion and the fast algorithm of Chapter 4, different DOSTs can be defined to allow arbitrary windowing and interpolation over the whole time-frequency domain. Brute-force calculation of the DOST coefficients is expensive. To combat this issue, in Chapter 4, we propose a suitable matrix form of the DOST calculation, which is directly related to the Fourier coefficients. Hence, with the same computational complexity class as the Fourier transform, a fast way of computing the DOST coefficients is recovered. The rigorous proof on its complexity is available in that chapter. In Chapter 6, the global and local translation properties on the DOST are studied individually. We discuss the results that we have reached and state that, due to the Fourier uncertainty principle, the local translation detection can not be done precisely. Nevertheless, numerical experiment has convinced us that a possible approximation strategy might exist for local translation detection. We also propose a mathematical system for a possible analysis tool to benefit further researchers and applications on local translation. In Chapter 5 and 6, we present two applications using the DOST on image compression and image restoration. In Chapter 8, we state some practical forward branching related to the DOST. Various diversities of the branching will benefit either theoretical analysis or application fields, such as image processing and designs of medical imaging devices.

In the Appendix, useful Matlab codes of the fast DOST are attached. We would be especially delighted to see that more applications are born based on Dr. Stockwell's initiation of this field and based on our extension in both theoretical and practical aspects of the DOST.

1.3 Contribution of this thesis

Both the ST and the DOST are younger than other well established transforms, such as the FT, the GT and the WT. As one of the many efforts in this thesis, I have endeavored to integrate the theoretical structure of the ST and the DOST by offering in depth understanding, intuitive properties and instructive comparisons to other transforms. Proofs on essential equations and theorems, and the time-frequency analysis including comparison to the GT and the WT have become the major contributions of the first two chapters.

On various aspects (sampling theorem, spectrum analysis and more) in signal and image processing, the virtues of the DOST have been highlighted in Chapter 3, which has formed a nearly comprehensive analysis on the DOST. The matrix factorization and thus the fast DOST algorithm, with detailed proof of computational complexity and experimental comparison, are some other major contributions of this thesis. The analysis on the translation properties is entirely new in this area. The global translation property is completely developed and the local translation property is reasonably analyzed.

As another two major contributions in application aspects, the DOST is used for image compression and image restoration. As proved in sufficient details in Chapter 6 and Chapter 7, the DOST has outperformed the wavelet transform at entrance level. More advanced applications will interest new researchers to gain better results than the state of art wavelet techniques.
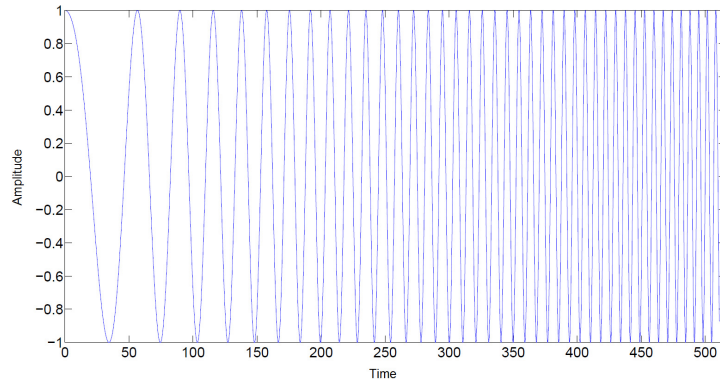
(a) A chirp with increasing frequency
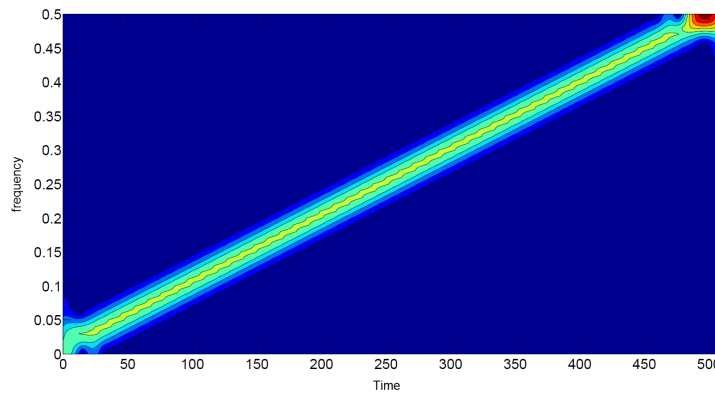
(b) Filled contour plot of the GT coefficients of (a)
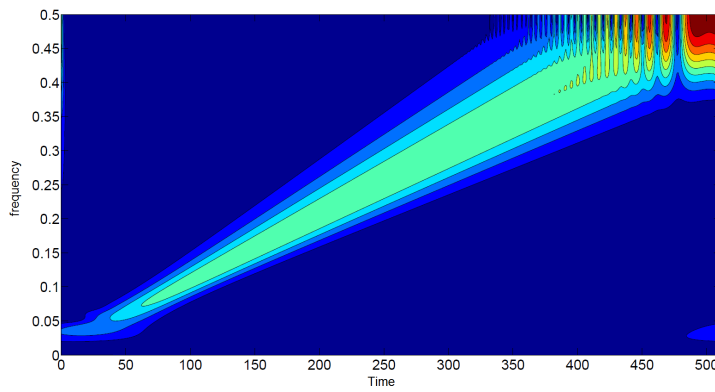
(c) Filled contour plot of the ST coefficients of (a)

Figure 1.1: A comparison between the GT and the ST.

Chapter 2

Stockwell Transform and Time-Frequency Analysis

For a given signal, the Stockwell transform (ST) [39, 27, 26, 17] gives a full time-frequency decomposition, which perfectly maintains the absolutely-referenced frequency and phase information. In this chapter, we first give a quick review of the ST and then highlight the comparison between the ST and other modern transforms, such as STFT and WT, in time-frequency analysis.

2.1 Stockwell Transform

2.1.1 1-D Continuous Stockwell Transform

For continuity, we will repeat the formal definition of the Stockwell transform (ST). The ST of a given function, $h(t) \in \mathbf{L}^1(\mathbb{R})$, is defined as [39, 27, 26, 17]

$$S(\tau, f) = \mathcal{S}\{h(t)\} = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt, \qquad (2.1)$$

11
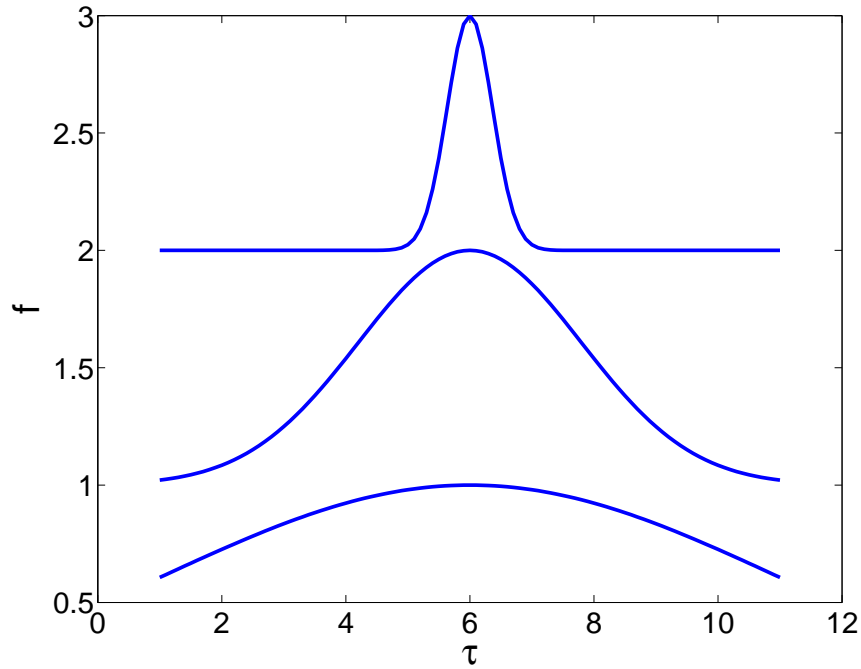
Figure 2.1: Each different f-value generates a different width of Gaussian curve, and hence a different width of kernel function and different resolution.

where f is the frequency, and t and τ are both time variables. The ST decomposes a signal into temporal (τ) and frequency (f) components. The value of τ represents the center of the window function, and thus, by picking all possible values for τ, the ST coefficients will cover the whole temporal axis and create full resolutions for each designated frequency. Different values of f adjust the sizes of the Gaussian windows over the temporal axis to realize multiresolution over different frequencies, i.e. higher resolution on higher frequencies and lower resolution on lower frequencies. Figure 2.1 illustrates different widths of Gaussian curves in resizing the kernel functions generated by different f values, and hence different resolutions for different f.

Considering the integral property of the Gaussian function,

$$\frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} e^{-\frac{x^2 f^2}{2}} \, dx = \frac{\sqrt{2\pi}}{|f|}, \tag{2.2}$$

the accumulation over all the Stockwell coefficients for a certain value of f will recover the Fourier coefficients,

$$\int_{-\infty}^{\infty} S(\tau, f) d\tau = H(f), \tag{2.3}$$

highlighting the special feature of the ST, its close relation to the FT.

Hence, the original function $h(t)$ can be recovered by calculating the inverse Fourier transform of $H(f)$,

$$h(t) = \mathcal{S}^{-1}\{S(\tau, f)\} = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} S(\tau, f) d\tau \right\} e^{i2\pi ft} df. \tag{2.4}$$

In general, the Stockwell coefficients $S(\tau, f)$ are complex, so we can write

$$S(\tau, f) = A(\tau, f) e^{i\Phi(\tau, f)}, \tag{2.5}$$

where $A(\tau, f)$ is the "amplitude S-spectrum" and $\Phi(\tau, f)$ is the "phase S-spectrum". The phase $\Phi(\tau, f)$ allows the definition of a broadband generalization of instantaneous frequency [38]. The absolutely-referenced phase information allows the comparison of phases derived from similar time series for correlation analysis [39].

2.1.2 1-D Discrete Stockwell Transform

As we will show below, taking advantage of the fast Fourier transform (FFT), there is an equivalent frequency-domain definition of the continuous Stockwell transform.

Theorem 2.1.1. *In the Fourier domain, the definition of the ST (equation (2.1))*

becomes

$$S(\tau, f) = \int_{-\infty}^{\infty} H(\alpha + f)e^{-\frac{2\pi^2\alpha^2}{f^2}}e^{i2\pi\alpha\tau}d\alpha, \qquad f \neq 0. \tag{2.6}$$

Proof. We start by substituting

$$H(\alpha + f) = \int_{-\infty}^{\infty} h(t)e^{-2\pi i(\alpha+f)t}dt, \tag{2.7}$$

into (2.6) to eventually derive (2.1).

After the substitution, (2.6) can be written as

$$\int_{-\infty}^{\infty} H(\alpha + f)e^{-\frac{2\pi^2\alpha^2}{f^2}}e^{i2\pi\alpha\tau}d\alpha$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t)e^{-2\pi i(\alpha+f)t}e^{-\frac{2\pi^2\alpha^2}{f^2}}e^{2\pi i\alpha\tau}dt d\alpha$$

$$= \int_{-\infty}^{\infty} h(t)e^{-2\pi fti}\left(\int_{-\infty}^{\infty} e^{-\frac{2\pi^2\alpha^2}{f^2}}e^{-2\pi i\alpha(t-\tau)}d\alpha\right)dt. \tag{2.8}$$

To evaluate the integral in the brackets of (2.8), we use the integral formula 2.33.1
on page 108 of [21]

$$\int e^{-(ax^2+2bx+c)}dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}\exp\left(\frac{b^2-ac}{a}\right)\mathrm{erf}\left(\sqrt{a}x + \frac{b}{\sqrt{a}}\right), \tag{2.9}$$

where

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt, \tag{2.10}$$

and

$$\mathrm{erf}(x)|_{-\infty}^{\infty} = \frac{2}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}dt = 2. \tag{2.11}$$

In our case, $a = \frac{2\pi^2}{f^2}$, $b = \pi i(t - \tau)$ and $c = 0$, so

$$\int_{-\infty}^{\infty} e^{-\frac{2\pi^2\alpha^2}{f^2}}e^{-2\pi i\alpha(t-\tau)}d\alpha$$

$$= \frac{1}{2}\frac{|f|}{\sqrt{2\pi}}\exp\left(-\frac{(t-\tau)^2 f^2}{2}\right)\mathrm{erf}\left(\sqrt{2}\frac{\pi}{f}\alpha + \sqrt{2}i(t-\tau)f\right)\Big|_{-\infty}^{\infty}$$

$$= \frac{|f|}{\sqrt{2\pi}}\exp\left(-\frac{(t-\tau)^2 f^2}{2}\right), \tag{2.12}$$

14

which supplies the Gaussian in (2.1) and completes the proof. $\qquad\square$

We may discretize (2.6) to define the discrete Stockwell transform (DST) [39]. For an input $h(m)$, $m = 0, \cdots , N - 1$, its DST can be written as

$$S(j,n) = \sum_{m=0}^{N-1} H(m+n) e^{\frac{-2\pi^2 m^2}{n^2}} e^{\frac{i2\pi m j}{N}}, \qquad j = 0, \cdots , N - 1, \qquad (2.13)$$

for $n = 1, \cdots , N - 1$, where $H(\cdot)$ is the DFT of $h(\cdot)$. For the $n = 0$ voice, define

$$S(j,0) = \frac{1}{N} \sum_{m=0}^{N-1} h(m). \qquad (2.14)$$

It has been shown [27] that

$$\frac{1}{N} \sum_{j=0}^{N-1} S(j,n) = H(n), \qquad (2.15)$$

where $H(n)$ is the discrete Fourier coefficient. Thus, the original signal can be recovered from the Stockwell coefficients as

$$h(k) = \left(\frac{1}{N}\right)^2 \sum_{j=0}^{N-1} \sum_{n=0}^{N-1} S(j,n) e^{\frac{i2\pi k}{N}}. \qquad (2.16)$$

It is not hard to see that if we also want multiple components over the frequency axis (assuming that we also want N samples for each temporal axis), via the discrete Stockwell transform, an N-tuple input signal will be decomposed into N^2 Stockwell coefficients. To get the explicit values of these N^2 coefficients, if the original discretized basis functions are involved for the dot product with the input signal, a total of $\mathcal{O}(N^3)$ operations would be required. However, taking advantage of the FFT [11], definition (2.13) offers a shortcut and calculates the Stockwell coefficients in an efficient way. More specifically, for a fixed value of j in (2.13), the DST coefficients for different n can be regarded as the inverse Fourier transform of the term $H[m + n]e^{-2\pi^2 m^2/n^2}$, so it could be done with

15

operations of order $\mathcal{O}(N \log N)$, which is identical to the computational complexity of FFT. Consequently, rather than a total of $\mathcal{O}(N^3)$, a total number of $\mathcal{O}(N^2 \log N)$ operations are sufficient to evaluate all N^2 Stockwell coefficients.

Multiresolution is a direct application of the Nyquist-Shannon sampling theorem, which is stated as following.

Theorem 2.1.2. *(Nyquist-Shannon sampling theorem)*

If a function $x(t)$ contains no frequencies higher than W hertz, it is completely determined by giving its ordinates at a series of points spaced $1/2W$ seconds apart.

In other words, lower frequency signals require fewer samples, and higher frequency signals require more samples. However, the DST includes N coefficients for each of the N frequency bands resulting in obvious redundancy in the low-frequency components according to the sampling theorem.

To achieve a reduced subset for each f, we might expect to require fewer coefficients for lower frequencies and more coefficients for higher frequencies, and thus form a key subset of all coefficients. However, if we accumulate the numbers of the coefficients in this key subset, it is the sum of an $N-$element arithmetic sequence. So, this reverse hierarchy still produces $\mathcal{O}(N^2)$ number of coefficients. Unless we have some prior of the signal (either high or low frequency dominates) or we know what specific actions (either high or low pass) need to be done for the signal, the way the ST is defined limits itself from being reconstructed by some substantially small subset of all coefficients. How to deal with the compromise between the temporal resolution and the frequency resolution becomes the design purpose of the discrete orthonormal Stockwell transform (DOST). The smart way of partitioning the time-frequency domain, and how it relates to the sampling theorem, will be explained in Chapter 3.

2.1.3   2-D Stockwell Transform

Like the FT, the ST is a separable transform over different dimensions. For a 2-D continuous-domain function $h(x', y') \in \mathbf{L}^1(\mathbb{R}^2)$, the 2-D ST with a 2-D Gaussian envelope can be analogously defined as

$$S(x, y, k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x', y') \frac{|k_x||k_y|}{2\pi} e^{-\frac{(x'-x)^2 k_x^2 + (y'-y)^2 k_y^2}{2}} e^{-i2\pi(k_x x' + k_y y')} dx' dy'.$$
(2.17)

As seen in (2.1), the Gaussian kernel changes shape with respect to spatial frequencies k_x and k_y. Due to this separability, the calculation can be pursued first over one dimension and then over another.

Integration of $S(x, y, k_x, k_y)$ over the variables x and y gives the 2-D Fourier spectrum,

$$H(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(x, y, k_x, k_y) dx dy.$$
(2.18)

Then the 2-D inverse Fourier transform can be applied to $H(k_x, k_y)$ to recover the original function.

Following the similar proof of (2.6), the ST (2.17) can also be defined as operations on the Fourier Spectrum $H(\alpha, \beta)$,

$$S(x, y, k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(\alpha + k_x, \beta + k_y) e^{-\frac{2\pi^2 \alpha^2}{k_x^2}} e^{-\frac{2\pi^2 \beta^2}{k_y^2}} e^{i2\pi(\alpha x + \beta y)} d\alpha d\beta,$$
(2.19)

for $k_x \neq 0$ and $k_y \neq 0$, where α and β are both frequency variables. In order to take advantage of the FFT calculation, the discrete 2-D Stockwell coefficients of an image $h(p, q)$, where $p = 0, \cdots, N-1$ and $q = 0, \cdots, M-1$, can be expressed explicitly as

$$
\begin{aligned}
&S(p, q, n, m) \\
&= \sum_{n'=0}^{N-1} \sum_{m'=0}^{M-1} H\left(n' + n, m' + m\right) e^{-\frac{2\pi^2 n'^2}{n^2}} e^{\frac{i2\pi n' p}{N}} e^{-\frac{2\pi^2 m'^2}{m^2}} e^{\frac{i2\pi m' q}{M}},
\end{aligned}
$$
(2.20)

for $n \neq 0$ and $m \neq 0$ (non-DC case).

For the case of $n = 0$ and $m \neq 0$, we need to use (2.13) with respect to m' indices first and then apply (2.14) for n' indices. For the case of $n \neq 0$ but $m = 0$ and $n = m = 0$, we can combine their definition similarly from (2.13) and (2.14).

It has been shown [27] that

$$\frac{1}{M} \sum_{q=0}^{M-1} \frac{1}{N} \sum_{p=0}^{N-1} S\left(p, q, n, m\right) = H\left(n, m\right), \tag{2.21}$$

where $H(n, m)$ are the discrete 2-D Fourier coefficients. Thus the original image can be reconstructed using

$$h(p, q) = \left(\frac{1}{M}\right)^2 \sum_{q'=0}^{M-1} \sum_{m=0}^{M-1} \left(\frac{1}{N}\right)^2 \sum_{p'=0}^{N-1} \sum_{n=0}^{N-1} S\left(p', q', n, m\right) e^{\frac{i2\pi p}{N}} e^{\frac{i2\pi q}{M}}. \tag{2.22}$$

In the Stockwell coefficients, each discrete point of the image has a 2-dimensional spatial-frequency representation, so the 2-D discrete Stockwell transform is a complex function of x, y, k_x and k_y. This 2-D DST offers the convenience and the freedom to manipulate data over spatial and frequency domains, but processing the 4-D set of Stockwell coefficients does tax computer resources and time; visualizing and analyzing these coefficients is a big challenge. Because of this reason, normally only relevant components of the $S(x, y, k_x, k_y)$ are computed and stored during real applications. Some strategies in dealing with 4-D data sets have been adopted successfully in various research fields [39, 32], which are described later in this section.

2.1.4   Properties of the ST

In order to maintain the scope of this thesis, we will only discuss the properties based on the 1-D case, with the exception of the rotation property of the ST. Most

of the properties for the 2-D transform can be derived analogously. And due to definition, the ST shares some similar properties of the FT. They are:

- Linearity: Assuming $h(t), g(t) \in \mathbf{L}^1(\mathbb{R})$ and a, b are arbitrarily complex numbers, the linear property holds as

$$\mathcal{S}\{ah(t) + bg(t)\} = a\mathcal{S}\{h(t)\} + b\mathcal{S}\{g(t)\}. \tag{2.23}$$

- Symmetry: The ST of a real function is a conjugate-symmetric function so that half the calculation can be saved in decomposition.

- Modulation: Shifting a function introduces into its spectrum a phase shift that is linear with frequency besides the shifting on the coefficients itself.

$$\mathcal{S}\{h(t - t_0)\} = e^{-i2\pi f t_0} S(\tau - t_0, f). \tag{2.24}$$

This alters the distribution of energy between the real and imaginary parts of the spectrum without changing the total energy.

- Scaling: Narrowing a function with a scale a will broaden its ST coefficients in the scale of $1/a$, and vice versa,

$$\mathcal{S}\{h(at)\} = S\left(a\tau, \frac{f}{a}\right). \tag{2.25}$$

- Rotation Invariance: Rotating a function rotates its ST coefficients on both spatial and frequency axes. Specifically, in Cartesian coordinate system, for a rotation operator $\mathcal{R}$,

$$\mathcal{S}\{h(\mathcal{R}(\overrightarrow{x}))\} = S\{\mathcal{R}(\overrightarrow{\tau}), \mathcal{R}(\overrightarrow{f})\}, \tag{2.26}$$

where

$$\mathcal{S}\{h(\overrightarrow{x})\} = S\{\overrightarrow{\tau}, \overrightarrow{f}\}. \tag{2.27}$$

The proofs of Linearity and Modulation are elaborated and can be found in Appendix A. Using the definition of the ST, the rest properties can be proven similarly.

2.1.5 Current Applications Using the ST

In this section, we highlight two applications of the ST using the DST, that have come to light since the ST was published [39]: one in medical image processing and the other in geophysics.

As one of the most accurate and efficient technologies in tumor study and cancer detection, MRI is becoming increasingly powerful and popular because of its non-invasive nature and increasing resolution. Today, the time it takes to acquire data has dropped significantly due to modern image processing techniques and improved technical design of the hardware itself. However, the movement of the object, either inside or outside of the field of view (FOV), is still one of the main sources of artifacts.

In Figure 2.2 (a), which shows a T2*-weighted fMRI image, the patient's coughing outside of the FOV causes obvious ghost view (the white-grey blurs inside the FOV). Ghost intensity is relatively high and overlaps the visual cortex. As located in the signal panel, Figure 2.2 (c), the high intensity peaks of the artifact can be observed. After processing with the designed 1-D ST filter, the ghost intensity magnitude is reduced to nearly baseline levels, as shown in Figure 2.2 (b). It is stated that, compared to other filter designs, the ST is a fairly powerful tool to deal with the artifacts caused by movement outside of the FOV with minimal impact on the data detected in the cortex.

In its field of origin, geophysics, the ST has also built significant applications

20
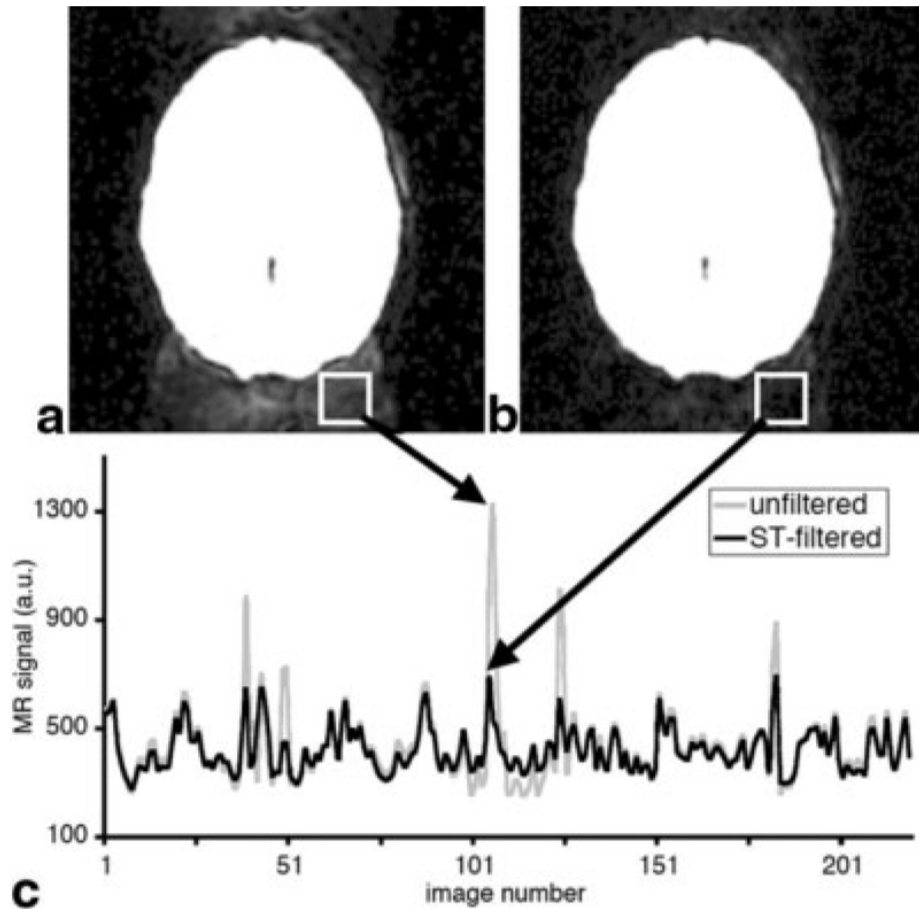
Figure 2.2: Stockwell transform (ST) filtering of fMRI data significantly reduces ghost intensity. a: T2*-weighted image collected when a subject was coughing. Ghost intensity is relatively high and overlaps the visual cortex. b: ST filtering reduces ghost intensity magnitude to the near baseline levels. c: Average time course of image intensity for image pixels inside the white boxes. ST filtering removes high frequency artifacts from the MR signal. (Used by permission of Dr. Hongmei Zhu.)

[36, 37, 32, 33]. The following is an example in image segmentation [32]. Figure 2.3 is the sample picture of the deposited Fanshawe Section in Southern Ontario. Quite
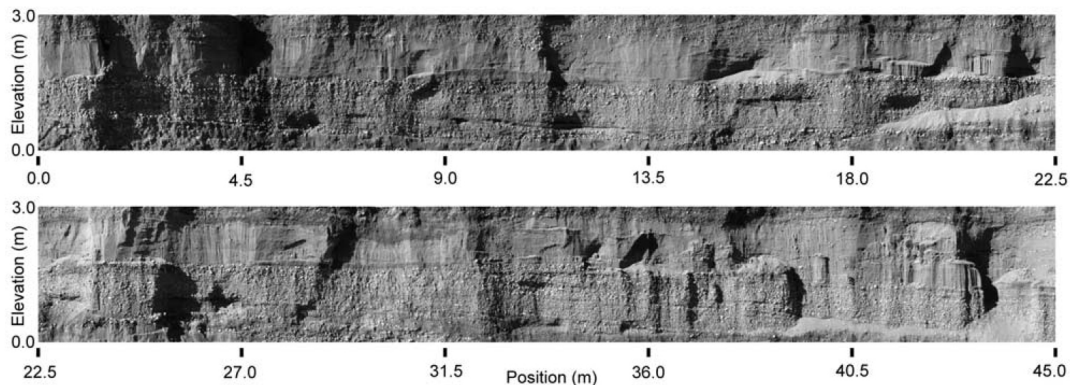
Figure 2.3: Photographic mosaic of the subsection data set for the Fanshawe section. Pixel resolution is 1.7mm. (Used by permission of Dr. Greg Oldenborger.)

different textures are visible in this area because of the year by year depositing. Researchers want to segment the sample in order to look for water or petroleum. In this application, a suitable treatment on the 4-D Stockwell coefficients is required so that the 2-D local spectrum on each pixel can be evaluated to a specific quantity for reference of the texture. Various treatments on the coefficients have been tried and compared by Dr. Oldenborger and an outstanding result is achieved in terms of the segmentation. Figure 2.4 shows the result based on their strategy.

The DST is also used in a lot of other areas. For example, in geophysics it is used for analyzing internal atmospheric wave packets [36], atmospheric studies [30], characterization of seismic signals and global sea surface temperature analysis [26]. It is used in electrical engineering [13], mechanical engineering [28], in digital signal processing [34], in the medical field in human brain mapping [2], in cardiovascular studies [40] and in studying the physiological effects of drugs [3].

22
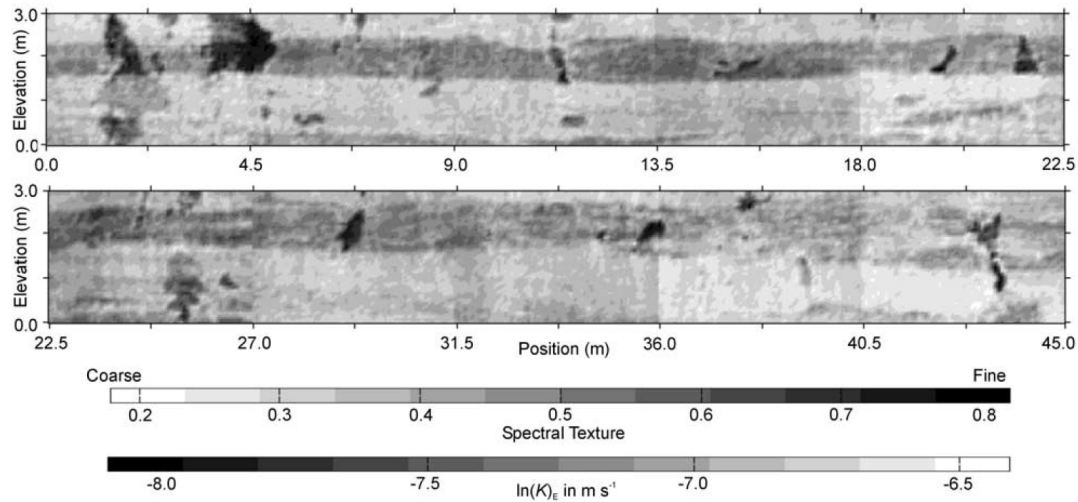
Figure 2.4: Spectral texture map and estimated log-transformed hydraulic conductivity field for the Fanshawe section. (Used by permission of Dr. Greg Oldenborger.)

2.2 Time-Frequency Analysis on ST, STFT and WT

In many applications, such as signal processing, image processing, etc., since the local information is usually required and needed to be treated, various techniques in the time-frequency analysis have been proposed and are widely used. The ST, STFT and WT are all popular transforms in terms of the time-frequency analysis. We will provide a brief review on the STFT and the WT, and then compare them in detail to the newly invented ST.

23

2.2.1 STFT vs ST

Generally, for an input function $h(t) \in \mathbf{L}^1(\mathbb{R})$, its short-time Fourier transform is defined as

$$\mathbf{STFT}\{h(t)\} \equiv X(\tau, f) = \int_{-\infty}^{\infty} w(t-\tau)e^{-i2\pi ft}h(t)dt, \qquad (2.28)$$

where $w(t-\tau)$ is the pre-selected window function and τ represents the center of the window. Adjusting the size and the center of the window allows the STFT to detect the local information from the input. However, only a few choices of window functions will yield a perfect reconstruction algorithm. Also, prior information is preferable to determine the window size in applying the STFT to real applications.

Normally, the window function is chosen to be a Gaussian window function, and thus defines the famous transform, the Gabor transform (GT). Explicitly, given a function $h(t) \in \mathbf{L}^1(\mathbb{R})$, the Gabor transform is formally defined as

$$\mathcal{G}(\tau, f) = \int_{-\infty}^{\infty} e^{-\pi(t-\tau)^2} e^{-i2\pi ft}h(t)\, dt, \qquad (2.29)$$

which offers the feasibility of recovering the original signal due to the integral properties of the Gaussian function.

In theory, the ST outperforms the STFT in two main aspects. First, the window size of the STFT is fixed for all frequency components, and thus needs to be pre-defined. As a consequence, there would be a chance that a specific frequency component will not be detected using the STFT (see this in detail in the experiments below). On the other hand, the window size of the ST is self adjusted in the sense that higher frequencies require more details and a higher temporal resolution. Second, the STFT is not usually invertible, but the ST is perfectly invertible, which makes the ST ideal for applications where reconstruction is involved.

In 1996 [39], the ST and the STFT were compared in real experiments. Based on the ST and STFT decomposition, two experiments were run to detect the short window of high frequency bursts. Their experimental setup and results are shown in Figure 2.5 and 2.6. First, to compare the performance of the ST and the STFT, a high frequency signal, a low frequency signal and a high frequency burst signal were combined to design the test signal of the experiment. In one result, both the ST and the STFT succeeded to detect the high frequency burst with noticeable non-zero coefficients at the right time interval. However, as the STFT uses a constant window width, it leads to having poorer temporal definition in the result. In the second experiment two non-overlapped but closely located high frequency bursts were added to crossed chirps signal. In the result, only the ST succeeded to detect both frequencies and to generate a clear separation between the bursts. But, as seen in the contour plot of the STFT coefficients, there were some non-zero coefficients between those two burst windows indicating that there was extra information beyond the crossed chirps over that region; however, no separation between the bursts was detected. The STFT coefficients of the bursts were compromised and the accuracy in the time axis was lost. Another time-frequency analysis facility, Wigner distribution, was also compared, but the result was not comparable to the ST and no burst was detected. ST was shown more useful than the other transforms since it indicated the bursts more clearly. This suggests its functionality in other applications.

2.2.2 WT vs ST

The Wavelet transform (WT) is a tool that cuts up data, functions or operators into different spatial-scale components, and then studies each component with a

25
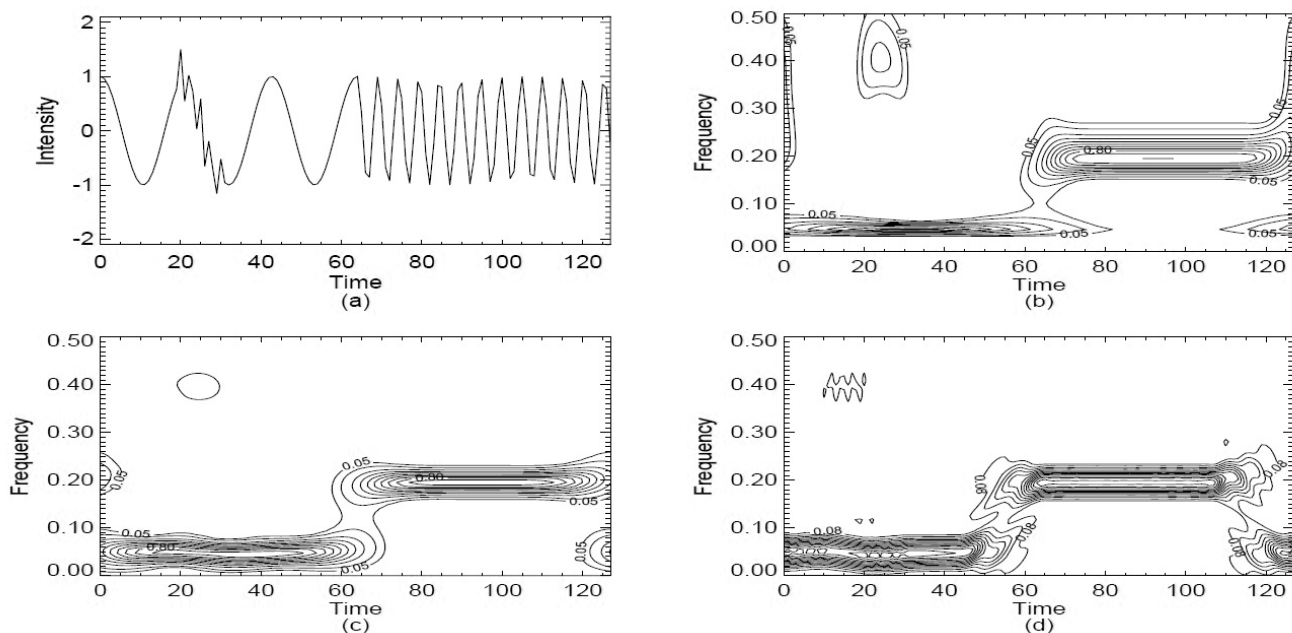
Figure 2.5: (a): A synthetic time series consisting of a low frequency signal for the first half, a middle frequency signal for the second half, and a high frequency burst at t=20. The function is $h[0 : 63] = cos(2\pi t * 6.0/128.0)$, $h[63 : 127] = cos(2\pi t * 25.0/128.0)$, $h[20 : 30] = h[20 : 30] + 0.5 * cos(2\pi t * 52.0/128.0)$. (b): The amplitude of the S transform of the time series. (c): The Short Time Fourier transform (STFT) of the time series using a fixed gaussian window of standard deviation = 8 units. (d): Same as (c) except that the window is a boxcar of length = 20 units. (Used by permission of Dr. Stockwell)

resolution matched to its scale. The continuous wavelet transform (CWT) for a continuous-domain input $h(t) \in \mathbf{L}^2(\mathbb{R})$ is defined as the integral

$$W(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} h(t)\psi\left(\frac{t - \tau}{s}\right) dt, \qquad (2.30)$$

where $\psi(t)$, called the mother wavelet, is a continuous-domain function of both the time and the scale; τ is the translation factor and s is the scale factor.

To recover the original input $h(t)$, based on the resolution of the identity

26
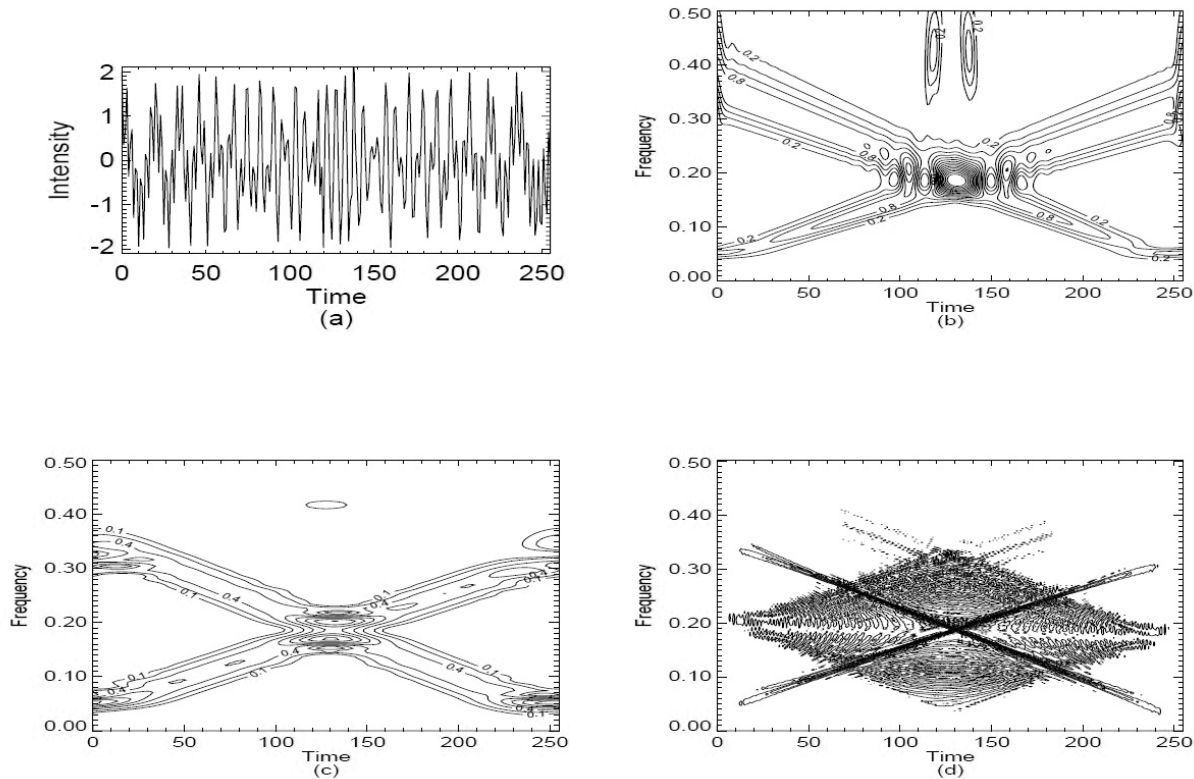
Figure 2.6: (a): A synthetic time series consisting of two cross chirps and two high frequency bursts. The time series is: $h[0 : 255] = cos(2\pi(10 + t/7) * t/256) + cos(2\pi(256/2.8 - t/6.0) * t/256)$, $h[114 : 122] = h[114 : 122] + cos(2\pi t * 0.42)$ and $h[134 : 142] = h[134 : 142] + cos(2\pi t * 0.42)$. (b): The amplitude of the S transform of the time series. (c): The amplitude of the STFT (with a Gaussian window) of the time series. (d): The amplitude of the Wigner distribution of the time series. (Used by permission of Dr. Stockwell)

formula, the inverse wavelet transform (IWT) is defined as

$$h(t) = \int_0^\infty \int_{-\infty}^\infty \frac{1}{s^2} W(\tau, s) \frac{1}{\sqrt{|s|}} \phi\left(\frac{t - \tau}{s}\right) d\tau ds, \qquad (2.31)$$

where $\phi(t)$ is the scaling function.

As one of the most important properties of the CWT, by conventionally choosing the scale factor as 2, it satisfies the conditions of the Multiresolution

27

Analysis (MRA) defined as the following,

Definition 1. *(Multiresolution Analysis)*

Let V_j, $j = \cdots, -2, -1, 0, 1, 2, \cdots$ be a sequence of subspaces of functions in $\mathbf{L}^2(\mathbb{R})$. The collection of spaces $\{V_j, \ j \in \mathbb{Z}\}$ is called a multiresolution analysis, with the scaling function ϕ, if the following conditions hold:

- *(nested) $V_j \subset V_{j+1}$, which creates an increasing subset of $\mathbf{L}^2(\mathbb{R})$.*

- *(density) $\overline{\bigcup V_j} = \mathbf{L}^2(\mathbb{R})$, which makes sure that any function in $\mathbf{L}^2(\mathbb{R})$ will belong to a V_j and hence $V_{j+1}, V_{j+2}, \cdots$, due to the nested property.*

- *(separation) $\bigcap V_j = 0$, which means the interception of all subset contains only one element, 0.*

- *(scaling) The function $f(x)$ belongs to V_0 if and only if the function $f(2^j x)$ belongs to V_j,*

- *(orthonormal basis) The function ϕ belongs to V_0 and the set $\{\phi(x-k), \ k \in \mathbb{Z}\}$ is an orthonormal basis ($\mathbf{L}^2$ inner product) of V_0.*

Figure 2.7 shows the nesting relations among the series of the of sets V_k and gives the intuition of the MRA.

In real applications, the CWT can not be used conveniently due to the requirement on continuous or infinite storage. The discrete wavelet transform (DWT) can be defined based on the multiresolution analysis. Normally, a DWT is obtained from a continuous representation by discretizing the dilation and translation parameters, s and τ. The dilation parameter is typically discretized by an exponential with base 2 and the translation parameter is chosen as integers.
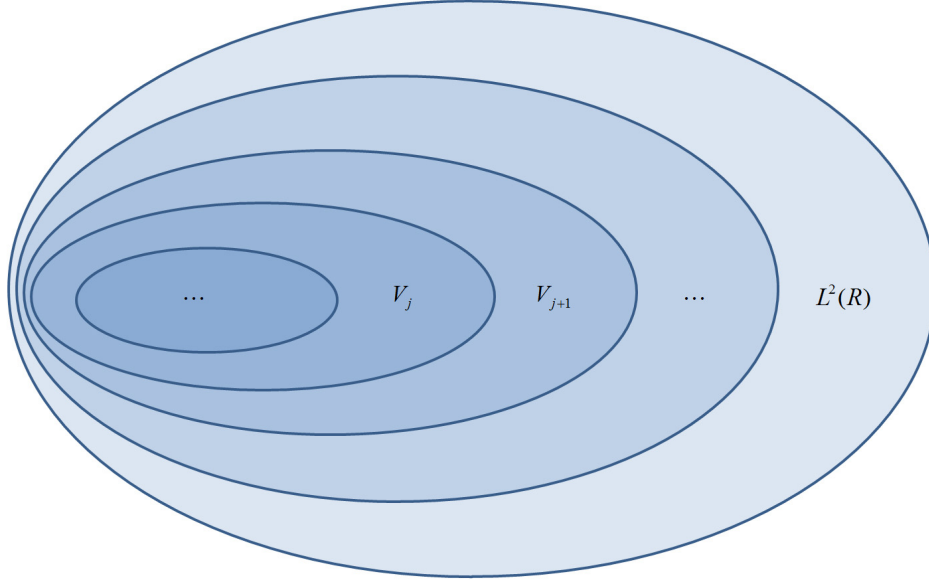
Figure 2.7: Intuition of the Mutiresolution Analysis

Explicitly, given a series coefficients $p_i, i \in \mathbb{Z}$, the scaling function for DWT can be defined as the function $\phi(x)$ that satisfies

$$\frac{1}{\sqrt{2}}\phi\left(\frac{x}{2}\right) = \sum_{k\in\mathbb{Z}} p_k \phi(x-k), \tag{2.32}$$

and the mother wavelet function $\psi(x)$ is defined as

$$\frac{1}{\sqrt{2}}\psi\left(\frac{x}{2}\right) = \sum_{k\in\mathbb{Z}} (-1)^k \overline{p}_{1-k}\phi(x-k). \tag{2.33}$$

This definition offers the basic properties of the scaling functions and the wavelet functions – self similarity – which implies the calculation advantage of the DWT and ability of using the DWT in other areas such as numerical analysis.

There are many kinds of wavelets, the Daubechies wavelet [14], in which

$$\int x^k \psi(x)dx = 0, \quad k = 0, \cdots, K-1, \tag{2.34}$$

29

(a) Mother wavelet function   (b) Scaling function
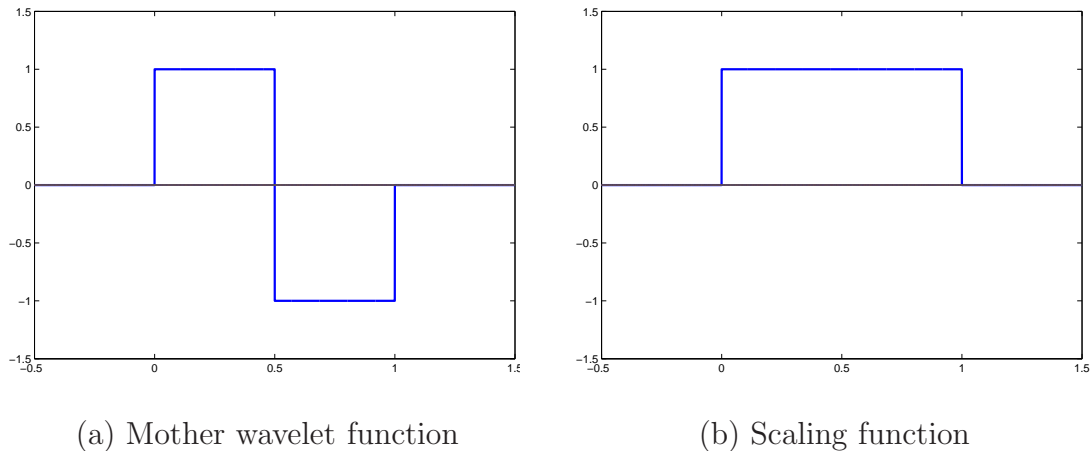
Figure 2.8: Mother wavelet function and scaling function for Haar wavelet.

the B-spline wavelet [10], the Shannon wavelet, etc., which prevail over various fields [25], such as image processing and pattern recognition. In image processing, it turns out the Daubechies wavelet has become one of the most often used wavelets [9, 15, 42]. The Daubechies wavelet forms a family of orthogonal wavelets with a finite set of non-zero coefficients. Generally, for the order-K Daubechies wavelet, there are $2K$ non-zero coefficients. For example, the Haar wavelet has two non-zero coefficients, which makes the Haar wavelet the simplest wavelet. The mother wavelet and the scaling function for the Haar wavelet are shown in Figure 2.8.

Due to the intrinsic relation, self similarity, between the scale functions and wavelet functions, the wavelet coefficients need not be calculated by the dot-product between the wavelet basis function and the input signal. Instead, especially for the Daubechies wavelet, the finite numbers of non-zero scaling coefficients, which generate the finer level basis functions from the coarser level, play an important role during the calculation of the wavelet coefficients. Super-fast implementations, known as downsampling and upsampling

operators [4, 15], can be iteratively used to generate the Daubechies wavelet coefficients within a computational complexity of $\mathcal{O}(N)$, where N is the size of input for a 1-D case. Consequently, the decomposition algorithm of the input signal $\{h^j\}$ can be diagrammatically shown as the following pyramid tree. All the "leaves" of the tree form the set of the wavelet coefficients.

Downsample :

$$\{h^j\} \quad \longrightarrow \quad \{h^{j-1}\} \quad \longrightarrow \quad \{h^{j-2}\} \quad \cdots \quad \longrightarrow \quad \{h^1\} \quad \longrightarrow \quad \{h^0\}$$

$$\searrow \qquad\qquad \searrow \qquad\qquad \searrow \qquad\qquad \searrow$$

$$\{\omega^{j-1}\} \qquad \{\omega^{j-2}\} \quad \cdots \qquad \{\omega^1\} \qquad \{\omega^0\},$$

As an example in discussing the complexity, an input $\{h^j\}$ of size N is decomposed using the basic Harr wavelet with only two non-zero coefficients. The first level decomposition in the following diagram takes $N/2 * 2 = N$ operations to generate $\{h^{j-1}\}$ and $\{w^{j-1}\}$ which each has $N/2$ elements. Thus, the second level decomposition will take $N/4 * 2 = N/2$ operations to the third level, and so on. To achieve the last level, only two operations are required since $\{h^0\}$ is scalar. So, in total, as the sum of an arithmetic progress, $N, N/2, \cdots, 1$, the total complexity to decompose $\{h^j\}$ is $2N - 1$, which is of order $\mathcal{O}(N)$.

Similarly, the reconstruction pyramid is

Upsample :

$$\{h^j\} \quad \longleftarrow \quad \{h^{j-1}\} \quad \longleftarrow \quad \{h^{j-2}\} \quad \cdots \quad \longleftarrow \quad \{h^1\} \quad \longleftarrow \quad \{h^0\}$$

$$\nwarrow \qquad\qquad \nwarrow \qquad\qquad \nwarrow \qquad\qquad \nwarrow$$

$$\{\omega^{j-1}\} \qquad \{\omega^{j-2}\} \quad \cdots \qquad \{\omega^1\} \qquad \{\omega^0\}.$$

During the decomposition and the reconstruction, multiplications only occur between the suitably chosen scaling coefficients and N wavelet coefficients, which makes the complexity of order $\mathcal{O}(N)$.

The calculation advantage of the DWT and its multiresolution ability highlights it as one of the most widely used tools in signal processing and image processing. The current computational complexity of the ST is still $\mathcal{O}(N^2 \log N)$. As we will show in Chapter 4, even for the latest discretized orthonormal version of the ST, the DOST, the optimum complexity we have achieved is still $\mathcal{O}(N \log N)$.

However, other factors have made the ST outperform the WT in the real world – frequency and phase advantage and quality advantage. First, as is well known from the Fourier theory, the translation in time of any function corresponds to a phase modulation in the Fourier spectral domain. The wavelet basis functions are self-similar due to the dilation, the translation and the linear combination of these operations. Hence, the wavelet voices have a phase modulation applied to them in the spectral domain. However, the basis functions of the ST are not translated and not self-similar to each other. Moreover, due to its direct relation to the FT (2.3), the ST successfully maintains the ability to recover the phase and frequency information from the ST coefficients without reconstructing the signal, which is known as keeping the absolutely-referenced frequency and phase information. The DOST, the compact version of the ST, has perfectly inherited this ability and highlighted its usefulness with an improved computational complexity. In particular, the absolutely-referenced frequency and phase information makes the ST and the DOST more appropriate to be applied when, for example, instantaneous frequency (IF) information is involved [39, 35]. Second, it is well known that in image processing using the DWT, block pattern often appears in the residual image: the visualization of the error between the original image and the reconstructed image. Various strategies have been attempted to solving this issue. However, as will be explained in detail in Chapter 5, even the global application of the DOST on the test image, the

residual image has a much milder block pattern.

There are discussions whether the ST is actually a member of the WT family. However, even though the ST and the WT share some similar properties, they are different kinds of transforms due to the following reasons:

- The CWT has a mother wavelet which makes all the wavelet functions dilations or translations of it. However, the Gaussian functions under different scales are not self-similar. Also, because of this reason, the ST will not be able to satisfy all conditions of the MRA. Nevertheless, the various resolutions obtained at different voices qualifies the ST to be used in the area where the MRA has conquered and entitles the ST as a quasi-MRA.

- In the CWT, the integral of the mother wavelets function is zero and the integral of the scaling function is one. However, the integral of the ST basis function, which is the multiplication between the Fourier basis and the Gaussian window, has no fixed value.

- In the CWT, only the scale information can be expected with a modulated phase information. However, in the ST, a direct relation exists between the ST coefficients and the FT coefficients, and the exact frequency and phase information can be achieved without reconstructing the images.

Overall, since the inception of the ST in 1996, its special advantages (multiresolution ability, keeping absolutely-referenced frequency and phase information) has helped the ST outperform the STFT and the WT in various fields. On the other hand, in solving the inefficiency and storage issues, either intelligent strategies may be applied under different situations to deal with the high dimensional data set, or attempts can be made to switch to the use of the

discrete orthonormal version of the ST, the DOST. This will be elaborated upon and discussed in the rest of this thesis.

Chapter 3

Discrete Orthonormal Stockwell Transform

As we already mentioned in the previous chapters, the mutiresolution decomposition of the ST is useful but redundant and computationally expensive. Starting from this chapter, we will focus on its discrete orthonormal version, the DOST, to achieve the desired efficiency and compactness.

3.1   Overview of the DOST

The DOST [35] is a pared-down version of the fully redundant ST. In the sense of multiresolution, less temporal resolution is required for a lower frequency band based on the sampling theorem. As discussed in the previous chapter, the ST has redundantly stored an equal amount of data in each low frequency band as in each high frequency band, despite the fact that the Nyquist criterion indicates that these two bands have very different sampling requirements.

The DOST manages to pursue a non-overlapping "multiresolution" partition

over the time-frequency domain. It does this by constructing a set of N orthogonal unit-length basis vectors in $\mathbb{C}^N$, each of which targets a particular region in the time-frequency domain. The regions are described by a set of parameters: ν specifies the center of each frequency band (voice), β is the width of that band, and τ specifies the location in time. Using these parameters, the kth basis vector is defined as

$$D[k]_{[\nu,\beta,\tau]} = \frac{1}{\sqrt{\beta}} \sum_{f=\nu-\beta/2}^{\nu+\beta/2-1} \exp\left(-i2\pi\frac{k}{N}f\right) \exp\left(i2\pi\frac{\tau}{\beta}f\right) \exp\left(-i\pi\tau\right), \qquad (3.1)$$

for $k = 0, \cdots, N-1$, which can be summed analytically to

$$D[k]_{[\nu,\beta,\tau]} = ie^{-i\pi\tau}\frac{e^{-i2\alpha(\nu-\beta/2-1/2)} - e^{-i2\alpha(\nu+\beta/2-1/2)}}{2\sqrt{\beta}\sin\alpha}, \qquad (3.2)$$

where $\alpha = \pi(k/N - \tau/\beta)$ can be regarded as the center of the temporal window. This partition strategy can be found in Figure 3.1. An alternative view is shown later in Figure 3.7 (a) with detailed coefficients.

To make the family of basis vectors in (3.2) orthogonal, the parameters ν, β and τ have to be chosen suitably. Letting the variable p index the frequency bands, Dr. Stockwell defines the DOST basis vectors of the positive frequency for each p on page 5 in [35] using,

- $p = 0$, (one basis vector)

  $\nu = 0$,

  $\beta = 1$,

  $\tau = 0$, $D[k]_{[\nu,\beta,\tau]} = 1$;

- $p = 1$, (one basis vector)

  $\nu = 1$,

  $\beta = 1$,

36
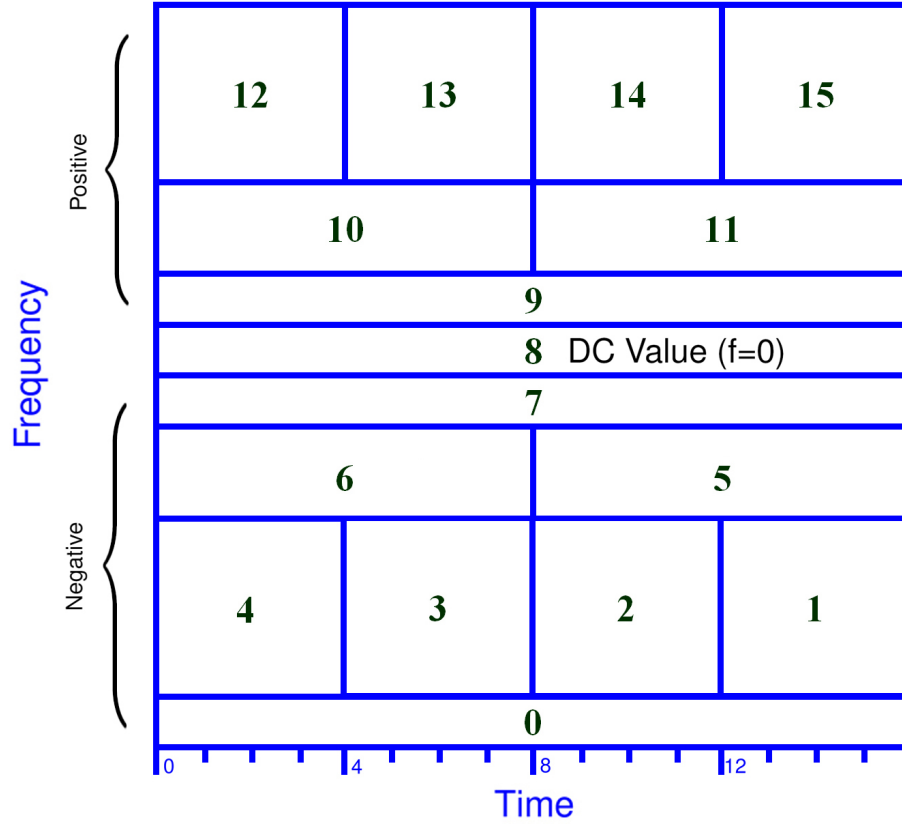
Figure 3.1: The order of the 2-D DOST coefficients into an 1-D N-vector.

$$\tau = 0,$$

$$D[k]_{[\nu,\beta,\tau]} = \exp(-i2k\pi/N);$$

- $p = 2, 3, \cdots, \log_2 N - 1$, $(2^{p-1}$ basis vectors for each frequency band)

$$\nu = 2^{(p-1)} + 2^{(p-2)},$$

$$\beta = 2^{(p-1)},$$

$$\tau = 0, \cdots, \beta - 1,$$

$$D[k]_{[\nu,\beta,\tau]} = ie^{-i\pi\tau} \frac{e^{-i2\alpha(\nu-\beta/2-1/2)} - e^{-i2\alpha(\nu+\beta/2-1/2)}}{2\sqrt{\beta}\sin\alpha}.$$

$$(3.3)$$
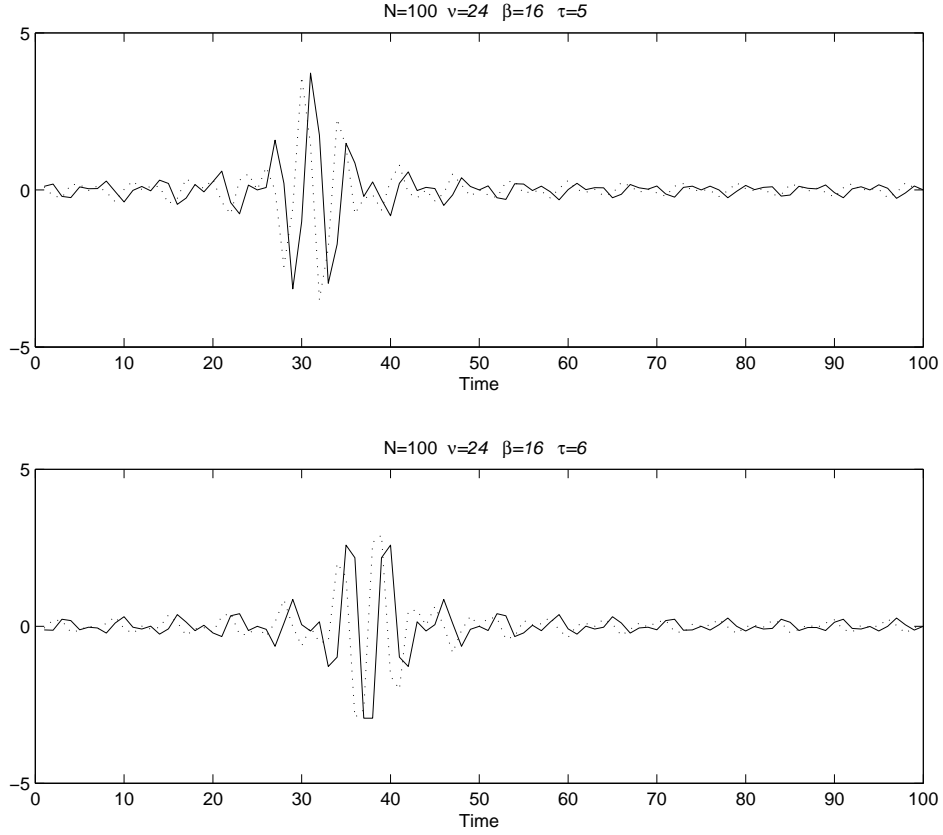
Figure 3.2: Plotting of two DOST basis functions.

Figure 3.2 shows some examples of the DOST basis functions. As is quite clear, the basis functions are not self similar.

Mathematically, we can prove that these basis vectors are orthonormal,

$$\frac{1}{N} \int_0^N D[k]_{[\nu',\beta',\tau']} D[k]^*_{[\nu,\beta,\tau]} dk = \delta_{\nu',\nu} \, \delta_{\beta',\beta} \, \delta_{\tau',\tau} \, , \tag{3.4}$$

where

$$\delta_{x,y} = \begin{cases} 1 & \text{for } x = y \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

is the Kronecker delta.

Combining these basis vectors with the basis vectors for the negative frequencies (described in the next section), we can prove that these parameter choices generate a basis of N orthogonal unit vectors, hence N DOST coefficients. For real applications, it is helpful to order these N coefficients into a 1-D vector. The ordering we use is shown in Fig. 3.1 for a signal of length 16 (see Fig. 3.7 (a) for more details). By convention, our time index (τ) traverses the time axis in the negative direction for negative frequencies. Doing so creates a symmetric correspondence between the positive- and negative-frequency coefficients in the 1-D representation. That is, for a given coefficient with index i in the 1-D DOST vector, its negative-frequency analog is at index $N - i$. This indexing convention will help later to gain symmetry of the DOST.

3.2   DOST and Sampling Theorem

As another way to address the Nyquist criterion, for a band limited signal with a maximum frequency of W Hz, we require $2W$ pieces of information to achieve a perfect reconstruction. By linear algebra, to recover this bandlimited signal, we will need W pieces of information to represent W harmonics and another W pieces to represent the corresponding phases. In this sense, this set of Fourier coefficients (W complex coefficients, or $2W$ degrees of freedom) is equivalent to the $2W$ samples, because the basis functions used in the sampling theorem are actually spanning the same subspace (band-limited signal space for $|f| \leq W$) as the Fourier decomposition and reconstruction.

For a bandpass signal staring from frequency W_L and ending at frequency W_H ($W_L \leq |f| \leq W_H$), we will need $(W_H - W_L + 1)$ harmonics and corresponding phases to reconstruct the signal. Generalized to linear algebra, $2(W_H - W_L + 1)$

pieces of information are the minimum required to reconstruct this signal.

For the sampling theorem, recall that the signal is reconstructed within the space spanned by the family of sinc functions

$$\psi_n(t) = \text{sinc} \left\{ \frac{t - nT}{T} \right\}, \tag{3.6}$$

where T is the temporal sample spacing. The Fourier spectrum of the sinc function is

$$\Psi_n(f) = \begin{cases} Te^{-i2\pi nfT} & \text{for } |f| \leq W \\ 0 & \text{for } |f| > W. \end{cases} \tag{3.7}$$

As seen from the spectrum of the sinc function, for each temporal sampling, it generates partial weights for all frequencies lower than W. Also, notice that $\Psi_n(f)$ are orthogonal for different n, and so are the corresponding sinc functions. This orthogonality offers the perfect equivalence between the Fourier-spanned signal space and the sinc-functions-spanned signal space. However, in our case of the bandpass signal, the low frequency components ($|f| < W_L$) involved in the sinc spectrum will need to be canceled out. Hence, a higher sampling rate than $2(W_H - W_L + 1)$ is required. The sampling theorem for this bandpass signal has been studied and a formal theory regarding the sampling rate is stated [12, 22]

$$\frac{2(W_H + 1)}{n} \leq f_s \leq \frac{2W_L}{n - 1}, \text{ for } n \text{ satisfying: } 1 \leq n \leq \left\lfloor \frac{W_H + 1}{W_H - W_L + 1} \right\rfloor, \tag{3.8}$$

where $\lfloor \cdot \rfloor$ rounds toward negative infinity.

The DOST basis function offers a perfect frequency response between the designed frequency values. In this sense, the design of the DOST might be a good supplement to the study of the sampling theorem (especially to the non-uniform sampling theorem). Instead of calculating the locations of the sampling points, some DOST coefficients can be calculated and used. This will form an interesting branch of study for the DOST.

40

Moreover, the number of DOST coefficients for a bandpass signal is consistent with the sampling requirement stated in (3.8). For example, in Figure 3.7 (a) the topmost band, along with its negative counterpart cover the frequency from four to seven. According to the sampling theorem, eight samples are required to recover that band. On the other hand, according to the definition of the parameter τ, we have exactly eight DOST coefficients available, corresponding to the case of $n = 2$ in (3.8).

For applications in which narrower high-frequency bands are desired, a reverse combination (wider frequency band at lower frequencies, and vice versa) can generate different types of DOST. We claim that research in designing new DOSTs for different applications is an exciting theoretical branch of research. As will be shown in later sections and chapters, flexible ways of partitioning the time-frequency domain and suitable definitions of the parameters (ν, β and τ) are possible to maintain the orthogonality, conjugate symmetry and fast calculation strategy. In section 3.7, we develop a new DOST, that keeps all the properties of the original DOST and its calculation advantages by reasonably varying the parameters according to the design diagram.

3.3 Visualization of Time-Frequency DOST Coefficients

For computation and storage convention, the DOST coefficients of a signal of size N have been stored as an N-tuple vector. However, it is important to be able to analyze the data set back into its 2-D nature. For this purpose we implement the 2-D visualization according to the order of Figure 3.1.
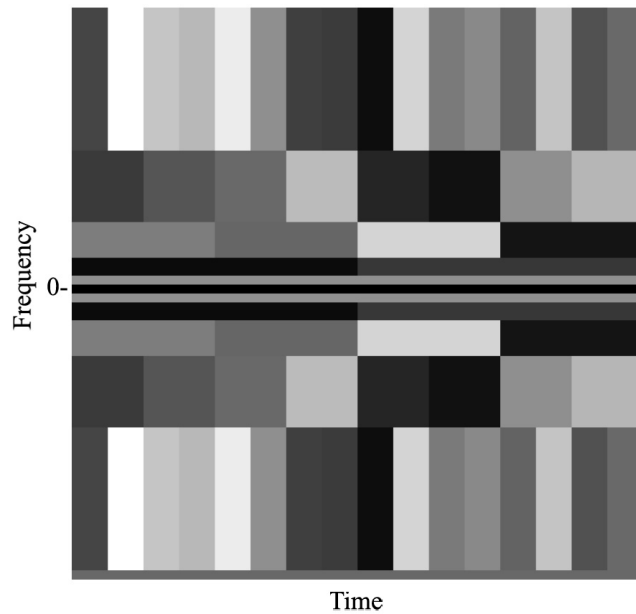
Figure 3.3: 2-D visualization of the DOST coefficients of a signal of size 64.

Figure 3.3 gives an outlook of this visualization, where the horizontal axis is consistent with the temporal index of the signal and the vertical axis is consistent with the ordered frequency bands.

3.4 Conjugate Symmetry of the DOST

If we pick the parameters (ν, β and τ) suitably, a real-valued input signal yields a set of conjugate symmetric DOST coefficients.

More explicitly, if we use the negative integers p to index the negative frequency bands, and let $q = -p$, then we can choose the parameters using:

- $q = 1$, (one basis vector)
  $\nu = -1$,

$$\beta = 1,$$

$$\tau = 0, \ D[k]_{[\nu,\beta,\tau]} = \exp(i2k\pi/N);$$

- $q = 2, 3, \cdots, \log_2 N - 1$, ($2^{q-1}$ basis vectors for each frequency band)

$$\nu = -2^{(q-1)} - 2^{(q-2)} + 1,$$

$$\beta = 2^{(q-1)},$$

$$\tau = 0, \cdots, \beta - 1,$$

$$D[k]_{[\nu,\beta,\tau]} = ie^{-i\pi\tau} \frac{e^{-i2\alpha(\nu-\beta/2-1/2)} - e^{-i2\alpha(\nu+\beta/2-1/2)}}{2\sqrt{\beta}\sin\alpha};$$

- $q = \log_2 N$, (one basis vector)

$$\nu = -2^{q-1},$$

$$\beta = 1,$$

$$\tau = 0,$$

$$D[k]_{[\nu,\beta,\tau]} = \exp(-ik\pi).$$

$$(3.9)$$

Theorem 3.4.1. *The DOST basis functions under the parameters ν, β and τ according to the rules of (3.3) and (3.9) form an orthonormal basis in $\mathbb{C}^N$. Moreover, for a real-valued input signal, the DOST coefficients are conjugate symmetric about the DC value (p=0).*

Proof. The orthogonality has been implied by the definition (3.1). We will focus on the conjugate symmetry here.

For an arbitrarily given band index $|p|$ ($p \neq 0$ and $\log_2 N$), we have two groups of basic functions: one group corresponding to the positive frequencies, and the other group corresponding to the negative frequencies. Notice here the values of β are the same and the values of τ have the same range, $0, 1, \cdots, \beta - 1$. We will distinguish the positive-frequency parameters from the

negative-frequency parameters using a superscripted positive sign or negative sign. Then

$$
\begin{aligned}
D[k]^*_{[\nu^+,\beta,\tau]} &= \left(ie^{-i\pi\tau} \frac{e^{-i2\pi(k/N-\tau/\beta)(\nu^+-\beta/2-1/2)} - e^{-i2\pi(k/N-\tau/\beta)(\nu^++\beta/2-1/2)}}{2\sqrt{\beta}\sin\pi(k/N-\tau/\beta)} \right)^* \\
&= -ie^{i\pi\tau} \frac{e^{i2\pi(k/N-\tau/\beta)(\nu^+-\beta/2-1/2)} - e^{i2\pi(k/N-\tau/\beta)(\nu^++\beta/2-1/2)}}{2\sqrt{\beta}\sin\pi(k/N-\tau/\beta)}, \quad (3.10)
\end{aligned}
$$

where $*$ denotes complex conjugation. Note that, for the corresponding negative index $-p$, we have $\nu^+ = -(\nu^- - 1)$. Thus, (3.10) can be written

$$
\begin{aligned}
D[k]^*_{[\nu^+,\beta,\tau]} &= -ie^{i\pi\tau} \frac{e^{i2\pi(k/N-\tau/\beta)(-\nu^-+1-\beta/2-1/2)} - e^{i2\pi(k/N-\tau/\beta)(-\nu^-+1+\beta/2-1/2)}}{2\sqrt{\beta}\sin\pi(k/N-\tau/\beta)} \\
&= ie^{i\pi\tau} \frac{e^{-i2\pi(k/N-\tau/\beta)(\nu^--\beta/2-1/2)} - e^{-i2\pi(k/N-\tau/\beta)(\nu^-+\beta/2-1/2)}}{2\sqrt{\beta}\sin\pi(k/N-\tau/\beta)}, \quad (3.11)
\end{aligned}
$$

where we have swapped the terms in the numerator in (3.10). Since τ is an integer, $e^{i\pi\tau}$ is always real. Thus $e^{i\pi\tau} = e^{-i\pi\tau}$. Making these substitutions, we can write (3.11) as

$$
\begin{aligned}
D[k]^*_{[\nu^+,\beta,\tau]} &= ie^{-i\pi\tau} \frac{e^{-i2\pi(k/N-\tau/\beta)(\nu^--\beta/2-1/2)} - e^{-i2\pi(k/N-\tau/\beta)(\nu^-+\beta/2-1/2)}}{2\sqrt{\beta}\sin\pi(k/N-\tau/\beta)} \\
&= D[k]_{[\nu^-,\beta,\tau]}, \quad (3.12)
\end{aligned}
$$

which means, if the same τ values have been picked, the basis vectors for the positive-frequency band p are conjugate symmetric to the corresponding basis vectors for the negative-frequency band $-p$. Hence, the corresponding DOST coefficients will exhibit conjugate symmetry when the input is real-valued. $\square$

3.5 Spectrum Analysis of the DOST

As already shown in many works on wavelet analysis [4, 14, 20], the spectrum intensity of the wavelets is not well localized, which is one explanation why the

44

scale information of the wavelet is not equivalent to the frequency information. In this section, we examine the localization of the DOST spectrum. As we have seen, the DOST basis functions are grouped according to the same frequency band with different temporal localizations. Basis functions within the same group share the same ν and β values, but have different τ values.

In Figure 3.4, we plot the amplitude of their Fourier spectrum for some selected basis functions at different bands. As we can see, each band covers one specific frequency band and never overlaps in frequency. Within the same band, the different basis vectors all have the same Fourier spectrum amplitude, which means the basis functions in that group can cover the same frequency band as we have analyzed. Only the phase plotting of these Fourier spectrums is different. In Figure 3.5, within the period of 2π, we plot some phases of the Fourier spectrum for some basis functions in the topmost band. In order to achieve the orthogonality within the same frequency band, as illustrated by the definition in (3.1), different τ values are picked for different basis functions to realize the different phases with respect to different frequency components. According to the Fourier shift theorem, these phase differences between different basis functions are equivalent to the same unit shifts in the temporal domain. Thus, the whole set of τ values forms a full coverage on the time duration. However, due to Fourier uncertainty principle, we can not expect these temporal coefficients within one specific frequency band to be non-overlapping, which creates an essential difficulty in studying the local translation property of the DOST. More discussion on this topic is in Chapter 5.

The perfect frequency localization and continuous phase resolution highlight the DOST and illustrate the particular advantages for filter design using the DOST.

45
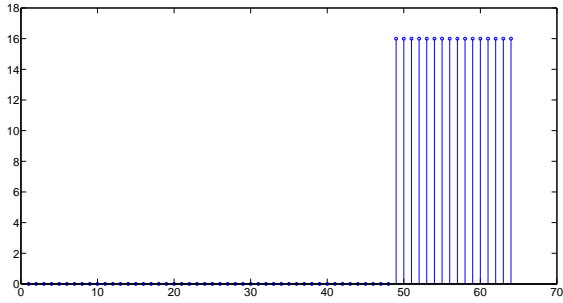
(a) $\nu = 24,\ \beta = 16,\ \tau = 0$
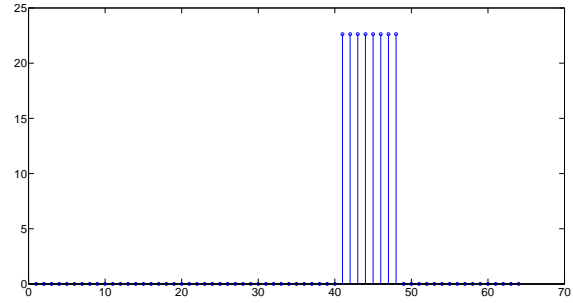
(b) $\nu = 12,\ \beta = 8,\ \tau = 0$
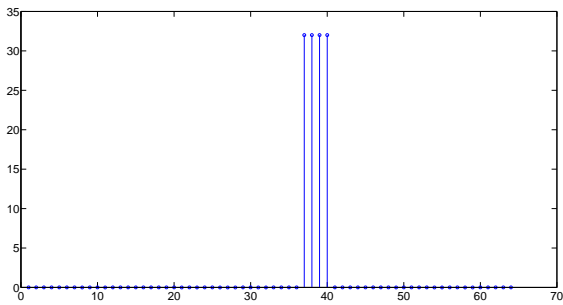
(c) $\nu = 6,\ \beta = 4,\ \tau = 0$
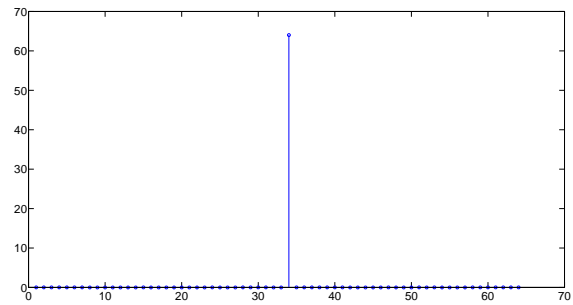
(d) $\nu = 0.5,\ \beta = 1,\ \tau = 0$

Figure 3.4: The amplitude of the Fourier spectrum of the basis functions, which covers different band of the frequency. (a) stands for a high frequency band which consist of multiple high frequencies, and while (d) stands for the lowest frequency band, DC.

3.6 Sub-band Coding and the DOST

Sub-band coding breaks a signal into a number of different frequency bands that can be independently coded. This decomposition is often the first step in data compression for audio and video signals. A high quality and low complexity sub-band filter bank is usually applied before the signal is encoded in each individual band. Effort has been made to make this algorithm faster while, at the same time, preserving fairly high quality. As one of the earliest works in this field, Dr. Burt et al. [6] proposed a Laplacian pyramid to accelerate the filter
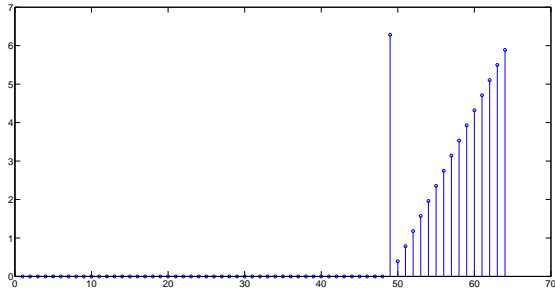
(a) $\nu = 24$, $\beta = 16$, $\tau = 15$
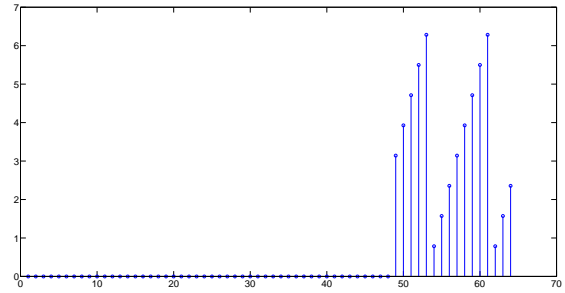
(b) $\nu = 24$, $\beta = 16$, $\tau = 14$
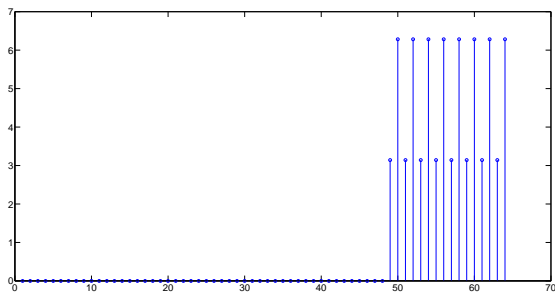
(c) $\nu = 24$, $\beta = 16$, $\tau = 8$
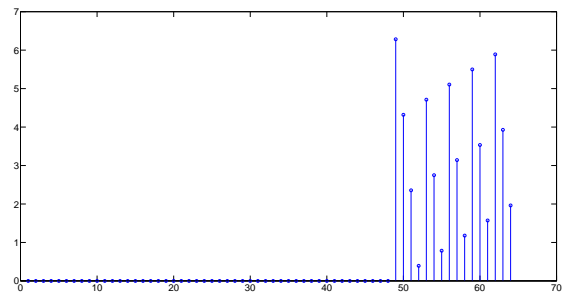
(d) $\nu = 24$, $\beta = 16$, $\tau = 5$

Figure 3.5: Some phase plotting of the Fourier spectrum of the basis functions in group Figure 3.4 (a). A complete combination of all the phases implies full temporal resolution.

bank execution. The idea was that a low pass filter blurs an image so that an alternative filter with a finite number of coefficients can be convolved equivalently to realize an approximate blur. By doing that, the computational complexity was reduced to $\mathcal{O}(N)$ comparing the general FFT, which is $\mathcal{O}(N \log N)$. However, an obvious frequency leakage happened, which could be easily explained by the Fourier uncertainty principle.

Demand of high compression quality in consumer audio applications for wireless networks continues to grow. Also, in order to accommodate the properties of human ears, a highly accurate sub-band division is always preferred. To my knowledge, to achieve a perfect sub-band decomposition in

47

frequency space, the Fourier transform is a minimum to avoid the frequency leakage caused by many other transforms, such as the wavelet transform, Laplacian pyramid, etc. Without sacrificing the computational complexity, the DOST has effectively reorganized the frequency band and, on the other hand, offered nearly localized temporal resolution to achieve more freedom. These advantages will be able to benefit sub-band coding at some level and form an interesting future research area on the DOST.

3.7 Approximation Using the DOST

Normally, in the field of image processing, signal processing, etc., people are dealing with low-frequency dominant sequences. In those sequences, the useful information is kept in the low frequencies, which makes it reasonable to drop some high frequency information to achieve a good approximation. Due to the multiresolution nature of the DOST and its time-frequency representation, approximation can be achieved by dropping or manipulating frequency- and/or time-specific DOST coefficients. To test the approximating ability of the DOST, we select one row from a test image, Barbara (which will be used in the next chapter for a compression experiment). The length of the signal is 512 and its intensity ranges from 0 to 255, stored as 8 bit integers.

We decompose the original input (the solid black curve in Figure 3.6 (a)) into the DOST domain and then accumulate, band by band with reconstruction, from low frequency band to high frequency band (small ν to large ν). Note that the input signal is real-valued, so the DOST coefficients are conjugate symmetric. When a band is included, its corresponding negative band is also included. In theory, the input signal is decomposed into 9 bands, with the lowest band, DC,

labeled as band 1, a higher band labeled as band 2 and so on. Figure 3.6 (a) shows a step by step approximation produced by picking the first two bands (only four coefficients), the first five bands (1/16 of the coefficients) or the first seven bands (1/4 of the coefficients). In Figure 3.6 (b), the difference for the accumulation up to a certain band is plotted. As we can read from the plot, by keeping only half of the DOST coefficients (band 1 to band 8), the image intensity only bears a sacrifice of 1.4 out of 256 (about 0.5%). Even for the approximation with only 1/4 of the coefficients (band 1 to band 7), on average, only 1.5% of the accuracy is lost for each entry.

3.8 Alternative Symmetric DOST

Figure 3.7 (a) shows how the parameters ν, β and τ partition the time-frequency domain and how the DOST gives conjugate symmetry when the input is real. Motivated by the Fourier Shift Theorem, we can modify the definitions of the parameters and define an alternative, fully-symmetric DOST. In Ref. [43], we previously proposed a symmetric DOST where the basis vectors were altered so that the resulting coefficients were conjugate symmetric for real-valued input. The new basis is also orthonormal. Here, then, we repeat the derivation.

Imposing the conjugate symmetry requirement on the DOST coefficients [43] gives us

$$D_{[\nu,\beta,\tau]} = \left(D_{[-\nu,\beta,\tau]}\right)^*. \tag{3.13}$$

This symmetry constraint is satisfied for all non-zero ν if we simply shift all the samples away from the zero frequency by $1/2$. We are then left with a gap from -1 to 1 containing one coefficient whose band extends from $-1/2$ to $1/2$. Instead, we split the gap into two coefficients, one with a band from 0 to 1, and the other
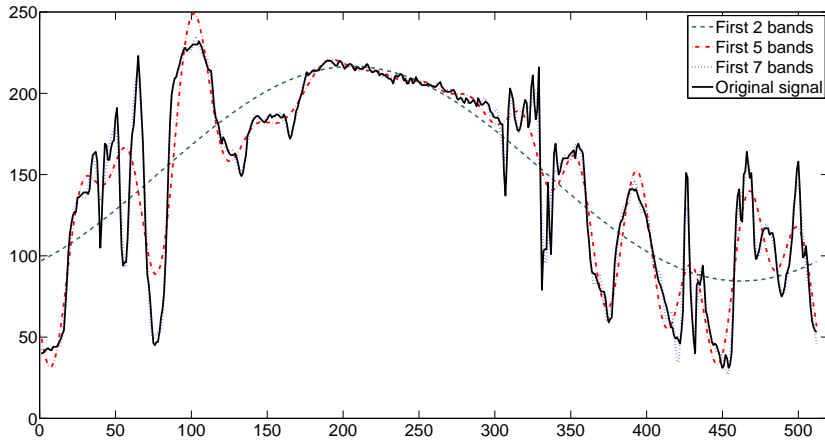
(a) Step by step approximating to the signal.
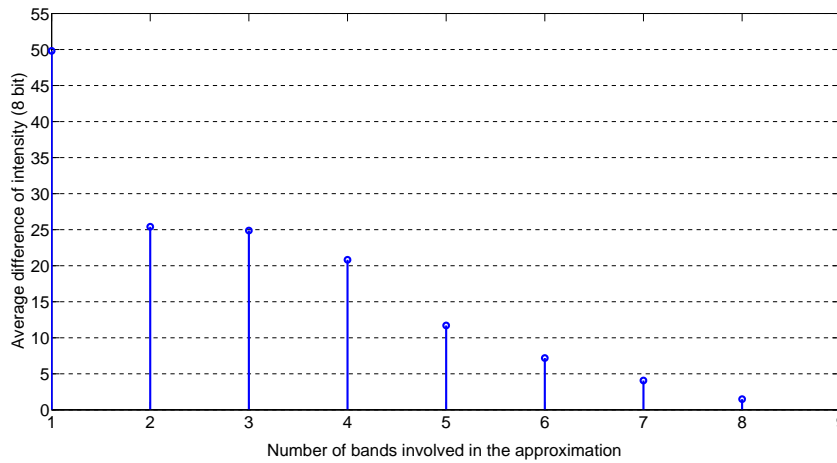
(b) Step by step averaging of the difference.

Figure 3.6: Study on the approximation ability of the DOST. Good approximation is achieved by keeping only a small part of the DOST coefficients.

with a band from -1 to 0. This alternative partition can be implemented by simply replacing ν with $(\nu + 1/2)$ in (3.2) for the positive frequencies; a partition that is mirrored for the negative frequencies. Then, the basis vectors for this symmetric
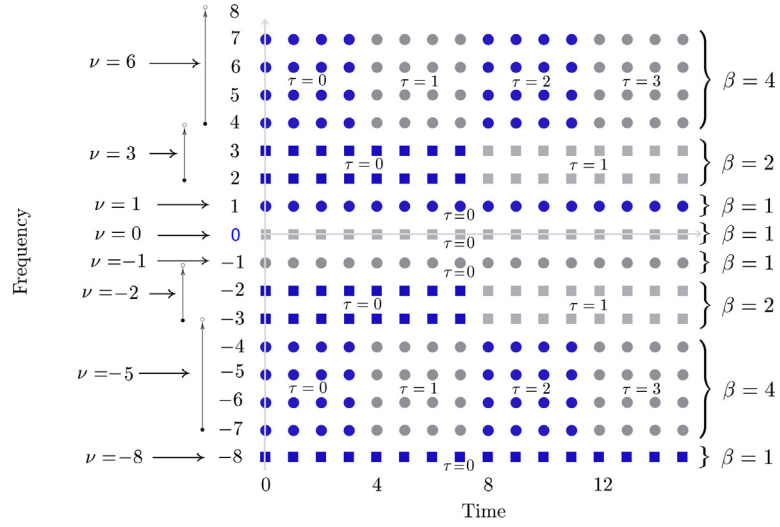
(a) DOST.
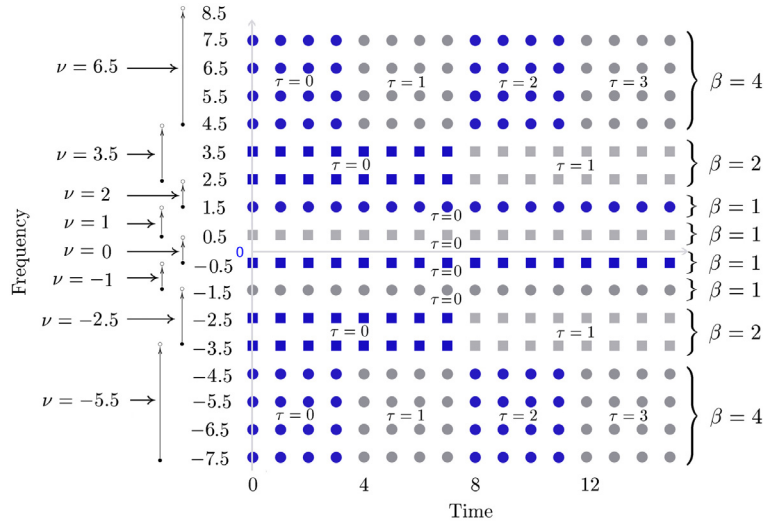
(b) Alternative Symmetric DOST.

Figure 3.7: Partition diagram of the time-frequency domain. Each rectangle block area with the same shape and color markers corresponds to one DOST coefficient. In (b), the partitions for the symmetric DOST have been shifted along the frequency axis according to the description in section 3.8 .

DOST, denoted by $\tilde{D}$, can be written

$$\tilde{D}[k]_{[\nu,\beta,\tau]} = ie^{-i\pi\tau}\frac{e^{-i2\alpha(\nu-\beta/2)} - e^{-i2\alpha(\nu+\beta/2)}}{2\sqrt{\beta}\sin\alpha}. \tag{3.14}$$

The orthogonality property still holds for this family of basis vectors. Figure 3.7 (b) shows the partition over the time-frequency domain of this symmetric DOST, and how it differs from that of the original DOST. This work was published in [43] in 2008.

3.9   2-D DOST

The 2-D ST is a separable transform, as is the 2-D DOST. Figure 3.8 gives the impression of how the DOST coefficients distribute in an ordered 2-D expression. The input is a black image, of size 1024×1024, with only one white dot at position of $(360, 90)$. For a better comparison between the coefficients, the plot of the coefficients is in log-scale.

Figure 3.9 shows the logarithm of the magnitude of the 2-D DOST coefficients for a popular example image, Lena. As we can see, the coefficients decay very quickly, which makes the DOST a powerful tool for image compression and other applications. Moreover, the DOST coefficients decay in a consistent way. As you can easily observe from the log-scale magnitude plot, there are still small "Lenas" on each corner of the plot. And even for the square and rectangular blocks inside the plot, where frequency bands with respect to different spatial axes of the image overlap, the stretched "Lenas" are still visible.

Due to the side-lobes seen in the plotting of the DOST basis function (see Figure 3.2), the 2-D DOST coefficients are non-zero almost everywhere, even for the one-dot image. This dispersion has created difficulties in using the DOST for
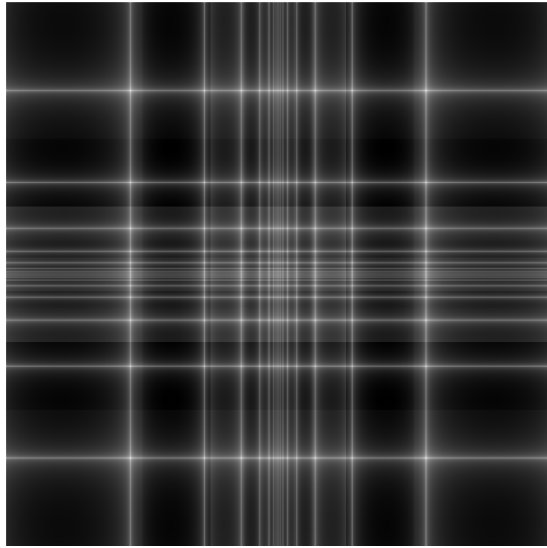
Figure 3.8: Logarithm of the DOST coefficients of an image with one white dot.
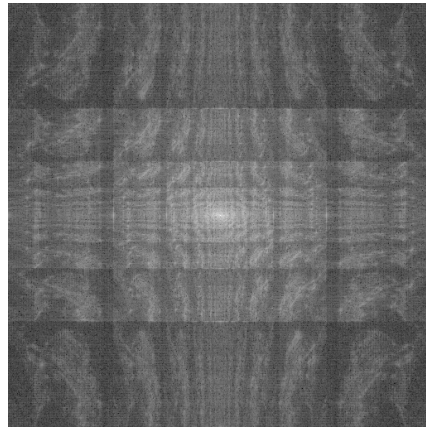
Figure 3.9: Lena and the logarithm of its DOST coefficients.

some applications. However, the extent of the temporal side-lobes is the price that must be paid for perfect frequency banding. This can be a nuisance, as will be seen in the discussion of the local translation in Chapter 5.

3.10 Current Applications Using the DOST

The DOST is fairly young. However, compared to other transforms and strategies, it has been demonstrated to be useful in some fields. The DOST has been successfully applied in signal analysis to channel instantaneous frequency analysis [35]. It has also been recently applied to image processing in image texture analysis [16], image compression [46] and image restoration [44]. The details of these applications can be found in the corresponding references and in Chapter 6 and 7 of this thesis.

Chapter 4

The Fast DOST

We stated above that the matrix-vector implementation of the DOST has computational complexity of $\mathcal{O}(N^2)$. However, the DOST can be calculated in a faster manner by taking advantage of the FFT. While this was mentioned in [35], we developed our method independently, and supply a rigorous proof of its computational complexity class here. This work has been published in the SIAM Journal on Scientific Computing (SISC) [45] in 2009.

4.1 FDOST Algorithm

Consider the inner product between $D[k]_{[\nu,\beta,\tau]}$, as shown in (3.1), and the input signal $h[k]$ (of length N). The resulting expression is the DOST coefficient, S, for the region corresponding to the choice of $[\nu, \beta, \tau]$, and can be expressed as

$$
\begin{aligned}
S_{[\nu,\beta,\tau]} &= \langle D[k]_{[\nu,\beta,\tau]}, h[k] \rangle \\
&= \frac{1}{\sqrt{\beta}} \sum_{k=0}^{N-1} \sum_{f=\nu-\beta/2}^{\nu+\beta/2-1} \exp\left(-i2\pi\frac{k}{N}f\right) \exp\left(i2\pi\frac{\tau}{\beta}f\right) \exp\left(-i\pi\tau\right) h[k]. \text{ (4.1)}
\end{aligned}
$$

55

In the above summation, the order of the sums can be switched and the common factors can be taken out. Then (4.1) becomes

$$\frac{1}{\sqrt{\beta}} \sum_{f=\nu-\beta/2}^{\nu+\beta/2-1} \exp\left(-i\pi\tau\right) \exp\left(i2\pi\frac{\tau}{\beta}f\right) \left[\sum_{k=0}^{N-1} \exp\left(-i2\pi\frac{k}{N}f\right) h[k]\right]. \qquad (4.2)$$

The part in the square brackets is $H[f]$, the discrete Fourier coefficient of our signal, evaluated at the frequency index f. Hence, we have

$$S_{[\nu,\beta,\tau]} = \frac{1}{\sqrt{\beta}} \sum_{f=\nu-\beta/2}^{\nu+\beta/2-1} \exp\left(-i\pi\tau\right) \exp\left(i2\pi\frac{\tau}{\beta}f\right) H[f], \qquad (4.3)$$

where the value of f is summed only on a certain band (depending on ν and β). Hence, this summation can be represented by the inner product between a row in a sparse matrix and the vector of the Fourier coefficients, H.

This strategy can be summarized as in Figure 4.1 (a). The block-diagonal nature of the transform matrix T offers the opportunity to calculate the DOST coefficients in a block-wise fashion. Hence, this sparse matrix allows for more efficient matrix multiplication.

The alternative symmetric DOST can be represented in a similar way (as shown in Figure 4.1(b)) by first multiplying the signal by a phase ramp. Despite the fact that the symmetric DOST corresponds to a 1/2-sample shift along the frequency axis, there is no loss of information due to resampling because the phase ramp that precedes the FFT implements the shift by the Fourier shift theorem. Note that the transform matrix is slightly different for the symmetric DOST. However, these transform matrices essentially have the same structure, and are block-diagonal in both cases.

Not only is T sparse, but each block of T has a special structure that facilitates efficient matrix multiplication. To see this, consider the top-left block, labeled T_1.

(a) DOST.

(b) Alternative Symmetric DOST

Figure 4.1: Calculation strategies of the DOST and the alternative symmetric DOST. The symmetric DOST is equivalent to the shifted version of the DOST with a different transform matrix.

In the case where $N = 16$, T_1 is

$$\frac{1}{\sqrt{\beta}} \begin{pmatrix} e^{-\pi i \tau_0} e^{2\pi i \frac{\tau_0}{\beta}(A)} & e^{-\pi i \tau_0} e^{2\pi i \frac{\tau_0}{\beta}(A+1)} & e^{-\pi i \tau_0} e^{2\pi i \frac{\tau_0}{\beta}(A+2)} & e^{-\pi i \tau_0} e^{2\pi \frac{\tau_0}{\beta}(A+3)} \\ e^{-\pi i \tau_1} e^{2\pi i \frac{\tau_1}{\beta}(A)} & e^{-\pi i \tau_1} e^{2\pi i \frac{\tau_1}{\beta}(A+1)} & e^{-\pi i \tau_1} e^{2\pi i \frac{\tau_1}{\beta}(A+2)} & e^{-\pi i \tau_1} e^{2\pi i \frac{\tau_1}{\beta}(A+3)} \\ e^{-\pi i \tau_2} e^{2\pi i \frac{\tau_2}{\beta}(A)} & e^{-\pi i \tau_2} e^{2\pi i \frac{\tau_2}{\beta}(A+1)} & e^{-\pi i \tau_2} e^{2\pi i \frac{\tau_2}{\beta}(A+2)} & e^{-\pi i \tau_2} e^{2\pi i \frac{\tau_2}{\beta}(A+3)} \\ e^{-\pi i \tau_3} e^{2\pi i \frac{\tau_3}{\beta}(A)} & e^{-\pi i \tau_3} e^{2\pi i \frac{\tau_3}{\beta}(A+1)} & e^{-\pi i \tau_3} e^{2\pi i \frac{\tau_3}{\beta}(A+2)} & e^{-\pi i \tau_3} e^{2\pi i \frac{\tau_3}{\beta}(A+3)} \end{pmatrix}.$$

where we have replaced $(\nu - \beta/2)$ with A for notational simplicity. Noting that $\tau_k = k$, if we index the rows with k and the columns with j (where $j, k = 0, \cdots, \beta - 1$),

then the (j, k) element of T_1 is

$$
\begin{aligned}
\beta^{-\frac{1}{2}} e^{-\pi i \tau_k} e^{2\pi i \frac{\tau_k}{\beta}(A+j)} &= \beta^{-\frac{1}{2}} e^{-\pi i \tau_k \left(1 - 2\frac{A}{\beta}\right)} e^{2\pi i \frac{\tau_k}{\beta} j} \\
&= \beta^{-\frac{1}{2}} e^{-\pi i k \left(1 - 2\frac{A}{\beta}\right)} e^{2\pi i \frac{k}{\beta} j}.
\end{aligned}
\tag{4.4}
$$

From (4.4), we can see that T_1 can be factored into a product of two matrices,

$$
T_1 = R_1 V_1,
\tag{4.5}
$$

where R_1 is a diagonal phase-ramp matrix with entries $r_k = \beta^{-1/2} e^{-\pi i k (1 - 2A/\beta)}$ and V_1 is the inverse Fourier matrix (of size $\beta = 4$ in our example).

Therefore, the process of multiplying by T_1 can be broken into two parts: applying V_1 which takes $\mathcal{O}(\beta \log \beta)$, and applying R_1 which takes $\mathcal{O}(\beta)$. Accumulating the operation counts over all the blocks in T (i.e. for $\beta = N/4, N/8, \cdots, 1, \cdots, N/8, N/4, 1$), the complexity to modify the Fourier coefficients to get the DOST coefficients is $\mathcal{O}(N \log N)$. A formal and detailed proof of the computational complexity of this technique will be given in next section. Since the initial FFT in Figure 4.1 (a) also has a complexity of $\mathcal{O}(N \log N)$, the total complexity for calculating the DOST coefficients is $\mathcal{O}(N \log N)$.

By studying the entries of the phase-ramp matrix in our algorithm, it turns out (taking into consideration how the parameters have been chosen) that

$$
r_k = e^{-2\pi i \frac{k}{\beta}(\beta - \nu)} = e^{-2\pi i \frac{k}{\beta}\frac{\beta}{2}},
\tag{4.6}
$$

which means the slope is $\beta/2$ in the algorithm we presented here. According to the Fourier Shift Theorem, that slope is equivalent to a shift over the input sequence before the IFFT is taken, which makes our algorithm equivalent to the one described in [16], where the shift of $-N_y/2$ is applied before the IFFT.

58

Let us now consider the operation of reconstruction, the inverse DOST. All the blocks of T are unitary matrices, so T is a unitary matrix. Hence the inverse of T is the adjoint (conjugate transpose) of T. The adjoint of T has the same structure as T, and can still be decomposed into a diagonal matrix and a Fourier matrix, and therefore applied with computational complexity $\mathcal{O}(N \log N)$. The other matrix factors shown in Figure 4.1 (a) are all trivially invertible and applied with the same computational complexity as the forward operators. Thus, the inverse DOST can also be computed in $\mathcal{O}(N \log N)$.

Moreover, during the decomposition and reconstruction, at no point does a matrix need to be explicitly stored. The FT matrices are implemented by the FFT, and the other matrices are all diagonal.

Besides the computational advantages, the matrix decomposition helps to elucidate the nature of the DOST decomposition. In the series of calculations to get the DOST, the input signal is transformed into pure frequency information first. Next, an inverse Fourier transform is applied to a narrow frequency band, yielding time-domain coefficients specific to that frequency band. Thus the final coefficients will carry both frequency and temporal information. This explanation is similar to the rationale given in [35] and [16].

Figure 4.2 plots the logarithm of the execution time for computing the FFT and FDOST. Both curves show the same growth trend, although the FDOST appears to be slower by a constant factor. As a comparison, the ideal $\mathcal{O}(N \log N)$ line is plotted as well.

Since the FDOST method is in a different computational complexity class than the brute-force DOST computation (using vector dot-products), we did not embark on a formal study to compare the execution times between the two methods. However, we include here a realistic example to give an impression of the speed
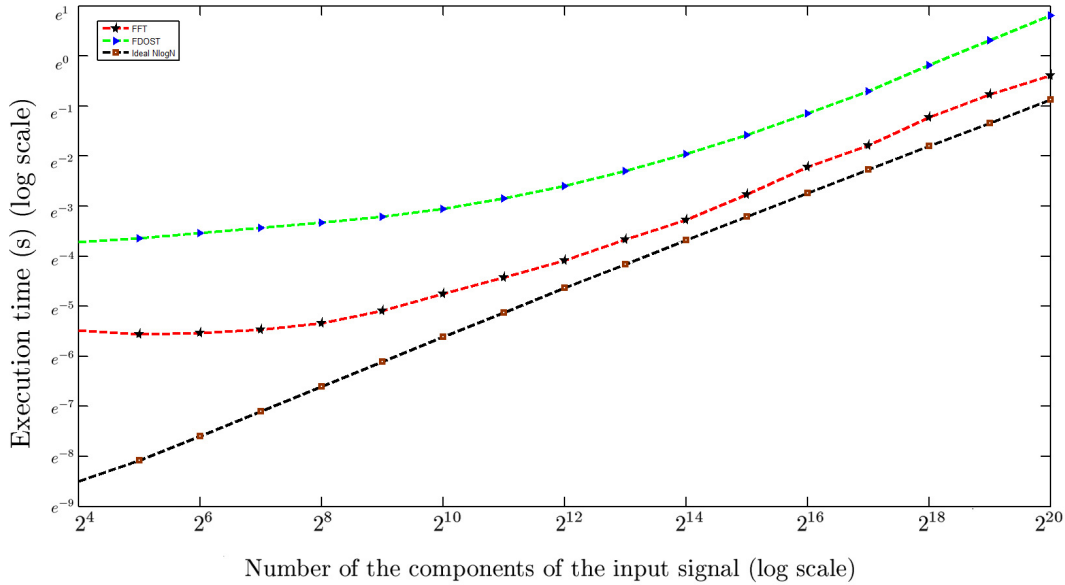
Figure 4.2: The comparison of time between the FDOST and FFT for various sizes of input signals.

difference. On a signal of length 1024, it took 2.285 seconds to compute the DOST using vector dot-products (including constructing the basis vectors), but only 0.0086 seconds using our FDOST method. It is worth noting, however, that these timings were run in Matlab. Although every effort was made to implement the two methods on a "level playing-field" (using Matlab's vectorization wherever possible), the timings ultimately depend on the particular Matlab implementation.

The alternative symmetric DOST has a slightly different transform matrix, $\widetilde{T}$, as well as a different ramp matrix (e.g. R_1 in (4.5)). However, both matrices have the same structure as their regular-DOST counterparts, so the symmetric FDOST algorithm also has complexity $\mathcal{O}(N \log N)$. Moreover, if the input signal is real-valued, the symmetry property allows one to compute only half of the coefficients.

4.2 Computational Complexity

Theorem 4.2.1. *The computational complexity of the fast DOST and fast inverse DOST algorithms, as described in section 4.1, is $\mathcal{O}(N \log N)$. The fast algorithms for the alternative symmetric DOST are also $\mathcal{O}(N \log N)$.*

Proof. Assume we have an input series, h, of size N. As well known, the computational complexity of taking the FFT on h is $\mathcal{O}(N \log N)$. Assume that the actual number of floating-point operations of the FFT (and IFFT) algorithm is $\alpha N(\log N)$.

First assume $N = 2^n$, where n is a positive integer larger than three.[1] The total accumulation of the DOST operations has been divided into two stages.

Stage 1: In this stage, we take the global FT using the FFT, i.e. the right-most matrix multiplication in Figure 4.1 (a). The operation count for this stage is

$$S_1 = \alpha N \log N. \tag{4.7}$$

Stage 2: In this stage, we perform the block-wise matrix multiplication of the Fourier coefficients (from stage 1) with T, i.e. the matrix multiplication on the left in Figure 4.1 (a).

Based on the partition strategy, in the left-most matrix of Figure 4.1 (a) we have a series of matrices of size

$$\{2^{n-2}, 2^{n-3}, \cdots, 2, 1, 1, 1, 2, \cdots, 2^{n-3}, 2^{n-2}, 1\}.$$

---

[1]In this thesis, we have focused on the dyadic length signals or images. The non-dyadic length case was mentioned in [35] by Dr. Stockwell. However, a formal theoretical structure and decent verification of this topic will be required and form one possible future work of the DOST.

Recall from (4.5) that the matrix block can be factored into a diagonal matrix (R) and a Fourier matrix (V). For a block of size 2^m, the number of floating-point operations required to perform the IFFT and diagonal matrix multiplication is

$$\alpha 2^m \log 2^m + 2^m = \alpha m 2^m \log 2 + 2^m. \tag{4.8}$$

So the total operations needed in this stage will be:

$$
\begin{aligned}
S_2 &= 2 \sum_{m=0}^{n-2} (\alpha m 2^m \log 2 + 2^m) + 2 * 2^0 \\
&= 2\alpha \log 2 \sum_{m=1}^{n-2} m 2^m + 2 \sum_{m=0}^{n-2} 2^m + 2.
\end{aligned}
\tag{4.9}
$$

Now we need to evaluate the sum of an arithmetic-geometric sequence, $m 2^m, m = 1, \cdots, n-2$. Letting

$$U = \sum_{m=1}^{n-2} m 2^m, \tag{4.10}$$

multiply by 2 on both sizes

$$2U = \sum_{m=1}^{n-2} m 2^{m+1} = \sum_{m=2}^{n-1} (m-1) 2^m. \tag{4.11}$$

Subtracting (4.10) from (4.11), we get

$$U = (n-2) 2^{n-1} - \sum_{m=2}^{n-2} 2^m - 2. \tag{4.12}$$

Using the fact $n = \log N / \log 2$,

$$
\begin{aligned}
S_2 &= 2\alpha \log 2 \left((n-2) 2^{n-1} - \sum_{m=2}^{n-2} 2^m - 2 \right) + 2 \sum_{m=0}^{n-2} 2^m + 2 \\
&= \alpha(n-2) 2^n \log 2 - \alpha 2^n \log 2 + 2^n + 4\alpha \log 2 \\
&= \alpha N \log N - (3\alpha \log 2 + 1) N + 4\alpha \log 2. \tag{4.13}
\end{aligned}
$$

Thus, the total number of floating-point operations required to calculate the DOST coefficients is

$$
\begin{aligned}
S &= S_1 + S_2 \\
&= 2\alpha N \log N - (3\alpha \log 2 + 1)N + 8\alpha \log 2 \\
&= \mathcal{O}(N \log N).
\end{aligned}
\tag{4.14}
$$

The computational complexity for the reconstruction and the alternative symmetric version can be proven in a similar fashion, which completes this proof. $\qquad\square$

The fast DOST and the separability between dimensions offer a way of using the DOST to analyze higher dimensional data sets.

Chapter 5

Global Translation and Local Translation of the DOST

The FT has a convenient representation for image translation. When an image is translated in a periodic manner (so that its contents wrap around), its Fourier coefficients are modified by the addition of a linear component to its phase, which is known as the Fourier Shift Theorem [5].

The ST is a compromise between localities in the temporal and frequency domains. In this chapter, we will initialize the study of the DOST Shift theorem. Considering the local properties of the DOST, we will attempt to achieve some local translation properties of the DOST, so that local translations can be detected and corrected based on the DOST coefficients themselves.

5.1   Global Translation

In Fourier theory, a circular shift of the input x_n corresponds to multiplying the Fourier coefficients X_k by a linear phase. Explicitly,

Corollary 5.1.1. *(Fourier Shift Theorem)*

If $\{x_n\}$ represents the input vector x then

$$\mathcal{F}(\{x_{n-m}\})_k = X_k \cdot e^{-\frac{2\pi i}{N}km}. \tag{5.1}$$

Recalling the matrix form presented in Chapter 4, the DOST coefficients can be achieved by applying a global FT first and then block-wise inverse FTs with ramp matrices. This calculation strategy offers the convenience to analyze the global translation property on the DOST.

Theorem 5.1.2. *(DOST Shift Theorem) If a one-dimensional signal is translated, then the entire DOST coefficients are equivalently translated according to the Fourier shift theorem on each frequency band.*

Proof. Denote $x = \{x_k\}$ as the original signal. Regarding the matrix order of the Fourier coefficients in the matrix form of Chapter 4, the index k to be taken as $k = N/2 - 1, N/2 - 2, \cdots, 0, \cdots, -N/2$ needs.

Denote $X_k = \mathcal{F}(\{x_n\})_k$, $k = N/2 - 1, \cdots, 0, \cdots, -N/2$ correspondingly, as its Fourier coefficients. Assume that x' is the translated version of the original signal. Without loss of generality, we will assume that the signal is translated to the right by the amount of m.

According to the Fourier shift theorem,

$$\mathcal{F}(\{x_n \cdot e^{\frac{2\pi i}{N}nm}\})_k = X_{k-m}, \tag{5.2}$$

$$\mathcal{F}(\{x_{n-m}\})_k = X_k \cdot e^{-\frac{2\pi i}{N}km}. \tag{5.3}$$

Equivalently, the shift theorem on the inverse Fourier transform states

$$\mathcal{F}^{-1}(\{X_{k-m}\})_k = x_k \cdot e^{\frac{2\pi i}{N}nm}, \tag{5.4}$$

$$\mathcal{F}^{-1}(\{X_k \cdot e^{-\frac{2\pi i}{N}km}\})_k = x_{n-m}. \tag{5.5}$$

Taking advantage of the matrix expression of the DOST developed in Chapter 4, the DOST coefficients of x, $S_{[\nu,\beta,\tau]}$, can be expressed as

$$\begin{aligned} S_{[\nu,\beta,\tau]} &= T \cdot \mathcal{F}(x) \\ &= T \cdot X. \end{aligned} \tag{5.6}$$

After plugging the translation into the input, the DOST coefficients of the translated signal, $S'_{[\nu,\beta,\tau]}$, can be expressed as

$$\begin{aligned} S'_{[\nu,\beta,\tau]} &= T \cdot \mathcal{F}(x') \\ &= T \cdot Q \cdot X, \end{aligned} \tag{5.7}$$

where Q is the diagonal phase-ramp matrix, with the diagonal components of $\exp\{-\frac{2\pi i}{N}km\}$, $k = N/2 - 1, \cdots, 0, \cdots, -N/2$. So, $Q \cdot X$ is a vector which has components $\exp\{-\frac{2\pi i}{N}km\}X_k$, $k = N/2 - 1, \cdots, 0, \cdots, -N/2$.

Recall that the transformation matrix T is a block-diagonal matrix. Therefore the components in $Q \cdot X$ can be partitioned accordingly. We first consider the positive frequency portion. Without loss of generality, only the top two blocks need to be analyzed.

Using the expression of (4.5), the top block of the transformation matrix T is

$$T_1 = R_1 V_1. \tag{5.8}$$

Denote the size of T_1 as $\beta_1 = N/4$. Let $[QX]_1$ denote the first β_1 elements of $Q \cdot X$. Denote the multiplication between V_1 and $[QX]_1$ as $[VQX]_1$ and the multiplication between V_1 and X (for $k = N/2 - 1, \cdots, N/2 - \beta_1$, or equivalently $k = 2\beta_1 - 1, \cdots, \beta_1$) as $[VX]_1$. Hence, we can write the band of shifted DOST coefficients as $R_1[VQX]_1$.

Without loss of generality, for now, the value of $m/4$ can be assumed to be integer. The non-integer case will occur in the lower frequency bands. If $m/4$ is not an integer, the interpolation among the highest frequency band will be required at the very beginning.

We rewrite $k = 3\beta_1/2 + k_1$ ($k_1 = \beta_1/2 - 1, \cdots, -\beta_1/2$) so that k_1 is centered in the voice, then rewrite the term in $[QX]_1$ as

$$
\begin{aligned}
\exp\left\{-\frac{2\pi i}{N}km\right\}X_k &= \exp\left\{-\frac{2\pi i}{4\beta_1}(\frac{3\beta_1}{2} + k_1)m\right\}X_k \\
&= \exp\left\{\pi i\frac{m}{4}\right\}\exp\left\{-\frac{2\pi i}{\beta_1}k_1\frac{m}{4}\right\}X_k \\
&= (-1)^{\frac{m}{4}}\exp\left\{-\frac{2\pi i}{\beta_1}k_1\frac{m}{4}\right\}X_k.
\end{aligned}
\tag{5.9}
$$

Recall from section 4.1 that V_1 is the inverse FT matrix. Then $[VQX]_1$ turns into the inverse Fourier transform of $[QX]_1$. Based on (5.5), the result will be the translated version of $[VX]_1$ by the amount of $m/4$ to the right with a possible minus sign. Notice that a general translation permutation matrix H commutes with R_1, or commutes with R_1 with an additional factor of -1. Indeed, for an example of size 4, when translation is odd,

$$
R_1 \cdot H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix},
$$

$$
H \cdot R_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{pmatrix},
$$

67

and when translation is even,

$$R_1 \cdot H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

$$H \cdot R_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Multiplication between R_1 and $[VX]_1$ provides the DOST coefficients of the original input on the positive highest frequency band. So, the Multiplication between R_1 and $[VQX]_1$ gives the translated DOST coefficients (an $m/4$-vector) on the same band with possible minus signs depending on the parity of $m/4$.

The next frequency band of DOST coefficients is half the size, and so are the transformation matrix and the ramp matrix involved in the calculations. So, on the next level of the matrix form, corresponding to the second highest frequency block, we have the Fourier coefficients from $k = \beta_1 - 1, \cdots, \beta_1 - \beta_2 = 2\beta_2 - 1, \cdots, \beta_2$, where $\beta_2 = N/8 = \beta_1/2$.

We rewrite $k = 3\beta_2/2 + k_2$, then $k_2 = \beta_2/2 - 1, \cdots, -\beta_2/2$. Following the same analysis and notations used above, we can rewrite the term in $[QX]_2$

$$
\begin{aligned}
\exp\left\{-\frac{2\pi i}{N}km\right\} X_k &= \exp\left\{-\frac{2\pi i}{8\beta_2}(\frac{3\beta_2}{2} + k_2)m\right\} X_k \\
&= \exp\left\{-\frac{3m\pi i}{8}\right\} \exp\left\{-\frac{2\pi i}{\beta_2}k_2\frac{m}{8}\right\} X_k, \quad (5.10)
\end{aligned}
$$

considering the assumption we made on m ($m/4$ is integer). Besides the factor from commuting the matrices, the former exponential (independent of the index

k) supplies another constant factor for this band. So if the $m/8$ is still an integer, following the analysis of the topmost frequency block, the DOST coefficients will be translated to the right by the amount of $m/8$ with possible minus signs. If not, the above calculation procedure matches the Fourier shift theorem with an additional constant factor, which is from both commuting and the index partition.

Moreover, the same analysis can go as deep as the smallest block in the center of the matrix form, in which only one DOST coefficient is involved. Over this block, no translation is required.

The complex conjugate property between the negative frequency and the positive frequency guarantees the same analysis can be done for the negative half of the frequency in the matrix form, which completes the proof. □

Theorem 5.1.2 forms the fundamental property of the DOST translation. With the amount of translation (usually called "offset" in applications), coefficients of the original signal on different bands can be manipulated in parallel to achieve new DOST coefficients. The lowest frequency coefficient in the DOST is the DC, and the conjugate symmetric second lowest frequency coefficients are actually Fourier coefficients regarding the definition of their corresponding basis functions in Chapter 3. So, due to the Fourier shift theorem, to quantify the amount of global translation is fairly straightforward by observing the phase change of the second lowest frequency. However, in the wavelet transform, the explicit relation between the coefficients before and after translation is not obvious. In Multi-Carrier Modulation using the wavelet, the algorithm needs to deal with the effect caused by the offset. Different techniques have been discussed [24], however, no satisfactory result has been achieved. Considering the multiresolution ability of the DOST and its direct relation with

respect to the offset, the DOST could be a good candidate in similar areas.

5.2   Local Translation

Let us now consider a one-dimensional signal, where there is a periodic translation over a small window inside the signal. Figure 5.1 shows an example of this behavior. We intend to build the relation between the DOST coefficients before and after the local translation, and thus detect the translation with only the DOST coefficients before and after the local translation.

As we have already seen in the previous chapters, the DOST is perfectly band-limited in the frequency domain. Due to the uncertainty principle between the Fourier domain and the time domain, the DOST can not be perfectly compact in the time domain. The side-lobes of the Gaussian window functions disperse the contribution of a single component to every DOST coefficient in a given band. Thus, we can not expect to find exact closed-form representations for temporally local phenomena. Rather, we look for approximate representations, a valid pursuit given that the Gaussian side-lobes fall off rapidly.

Let $x \in \mathbb{R}^N$ be a signal. Suppose that part of x undergoes a periodic shift, so that samples l through $l + L$ shift to the right.

Define $P \in \mathbb{R}^N$ as

$$P_i = \begin{cases} x_i & \text{for } l \leq i \leq l + L \\ 0 & \text{otherwise} \end{cases} \qquad (5.11)$$

and $B \in \mathbb{R}^N$ as

$$B = x - P. \qquad (5.12)$$

Since the shift is periodic, we can decompose P into two parts, one that shifts to the right, and the other part that effectively shifts to the left because of the

wrapping. Thus, if P is shifted by m samples $(m < L)$, then define P^1 and P^2 as

$$P_i^1 = \begin{cases} x_i & \text{for } l \leq i \leq l + L - m \\ 0 & \text{otherwise} \end{cases} \tag{5.13}$$

$$P_i^2 = \begin{cases} x_i & \text{for } l + L - m + 1 \leq i \leq l + L \\ 0 & \text{otherwise} \end{cases} \tag{5.14}$$

So, the original signal has been decomposed into three mutually exclusive parts, B, P^1 and P^2. Following the notation from the previous chapter, we use a letter D before the terms defined above to indicate the DOST transformation matrix. S is used to denote the original DOST coefficients and S' is the DOST coefficients of the locally translated signal. $T_k(\cdot)$ is used to denote the global translation to the right by k samples. Then the model can be expressed explicitly as

$$\begin{aligned} S &= DB + DP^1 + DP^2, \\ S' &= DB + DT_m(P^1) + DT_{m-M}(P^2). \end{aligned} \tag{5.15}$$

To identify a local translation, we will need to solve for m and M. Or, equivalently, we need to solve for P^1 and P^2. Using the result of Theorem 5.1.2,

$$S'_\nu = DB + \mathcal{F}^{-1}(R_{\nu,m} \cdot \mathcal{F}(DP^1)) + \mathcal{F}^{-1}(R_{\nu,m-M} \cdot \mathcal{F}(DP^2)), \tag{5.16}$$

where $R_{\nu,m} = \{e^{-\frac{2\pi i}{N}km}\}$ and $R_{\nu,m-M} = \{e^{-\frac{2\pi i}{N}k(m-M)}\} = \{e^{-\frac{2\pi i}{N}km} \cdot e^{\frac{2\pi i}{N}kM}\}$, $k = 1, \cdots, \beta_\nu$, are the diagonal phase-ramp matrices corresponding to the translation of P^1 and P^2 on level β of the DOST coefficients. When we subtract S' from S in (5.15), the effect from B cancels out with leaving only the coupling of the opposite translations. In a special case where $M = N$, which means the translation window is identical to the original length of the signal, the factor $e^{\frac{2\pi i}{N}kM}$ becomes one and makes it possible to combine $DT(P^1)$ and $DT(P^2)$, to solve for m. Otherwise,

the effects from the P^1 are coupled with the effects of P^2, which means there are not enough equations to exactly solve (5.15). In other special cases, if we do have some priori information available, we can use (5.15) to solve for the rest of the information. The total complexity remains $\mathcal{O}(N \log N)$ compared to $\mathcal{O}(MN^2)$, which is the order of the brute-force solution for finding the local translation.

As an empirical study, we used the signal in Figure 5.1 (a) and its local translated version in Figure 5.1 (b) as test signals to compare their DOST coefficients in terms of magnitudes and phases. The results shown in Figure 5.2 are consistent with our analysis. The magnitude difference and phase difference of the coefficients are mostly concentrated over the same area, which helps to locate the window approximately. There could easily be an approximation method to determine the window, and amount of translation. In particular, phase difference information in the low-frequency bands gives a general indication of the local translation. More precise information could be gleaned from the higher-frequency DOST coefficients in the same temporal location. For example, in Figure 5.2, the phase difference in the lower frequency might signal an inspection of the DOST coefficients for the region in higher frequencies. This future work will also be pursued in more detail in parallel with the theoretical model-analysis mentioned above.

5.3   Conclusion

The DOST shift theorem states the explicit relations between coefficients before and after global translation and opens the door for research of local translation. Unfortunately, due to the non-local nature of the DOST coefficients, coupling two opposite translations makes it difficult to solve the system (5.15). The
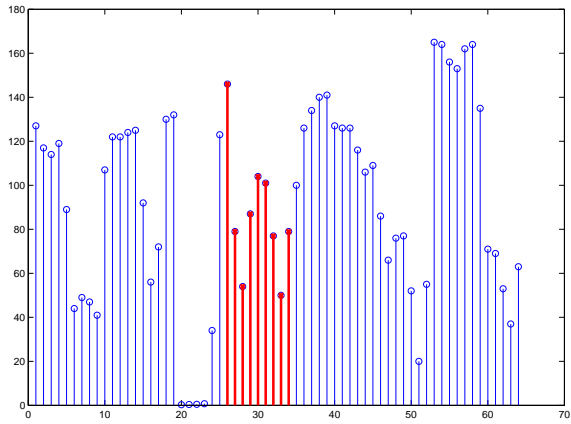
(a) Original Signal.
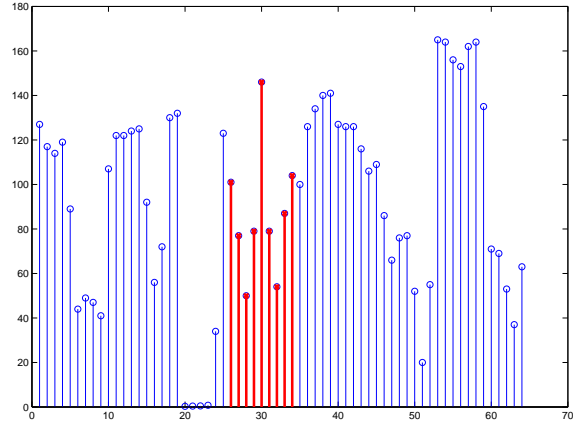
(b) Local translated Signal.

Figure 5.1: The original signal and the locally translated signal. Samples 26 to 34 (indicated by the solid filling) are periodically translated to the right by 4 samples.
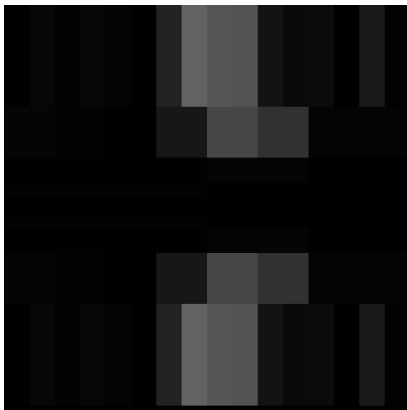
(a) Phase Difference
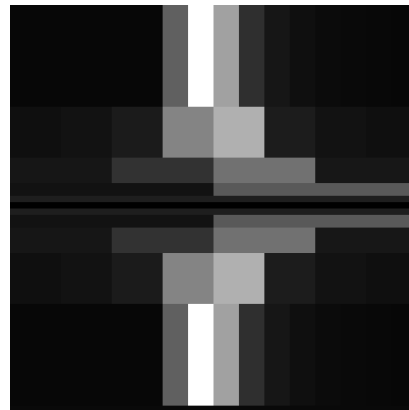
(b) Magnitude difference

Figure 5.2: The phase and magnitude difference between the DOST coefficients of the signals in Figure 5.1.

DOST does not yield a simple way to detect where the exact translation window is and how much the signal inside the window has been translated. However, numerical results give a hint that some approximation methods might be possible, which would directly benefit the multiple windows translation. Considering the computational complexity advantage using the DOST, we will leave this as possible future research.

Chapter 6

Image Compression Using the DOST

Image compression is an important step in many image-processing pipelines, allowing for smaller storage size, and faster download. Currently, JPEG image compression is one of the most prevalent image compression standards [42]. The most recent JPEG standard, called JPEG2000 [9], uses wavelets. Wavelets are currently regarded as the leading technology for image compression.

Before the use of wavelets, the Fourier transform (FT) was commonly used in image compression. The FT decomposes the image into its component frequencies, but does so globally so that each pixel affects every Fourier coefficient. Wavelets give a multiresolution decomposition in the spatial-scale domain. Even though the scale information can be approximately treated as frequency information (i.e. the fine scale information corresponds to the high frequency information and vice versa), the wavelet basis functions (e.g. the compactly supported Daubechies wavelets) are not entirely smooth. Hence, wavelet compression can be suboptimal on smooth parts of an image.

The Stockwell transform (ST) provides a continuous and infinitely differentiable kernel function and a full decomposition over the spatial-frequency domain. The orthonormal version of the Stockwell transform is the Discrete Orthonormal Stockwell transform (DOST) discussed earlier, which gives a spatial-frequency decomposition with no redundancy. In this chapter, we use an image compression experiment to demonstrate the advantages of the DOST by analyzing the peak signal to noise ratio (PSNR). We will see that a better approximation is achieved in the smooth areas of the image without sacrificing crisp edges. The result has been published in SPIE Proceedings [46] in 2009.

6.1 Methods

Our goal is to introduce the ST as a candidate tool for image compression. As an initial stab at determining the ST's capabilities, we compare it to two other transforms in a rudimentary compression methodology – simply dropping a percentage of the smallest coefficients (in modulus) and then reconstructing the images.

For our experiments, we used one of the most efficient families of wavelets, the Daubechies wavelets [14]. The Daubechies wavelets form a family of orthogonal wavelets with a small number of coefficients. Generally, the order-K Daubechies wavelet has $2K$ non-zero coefficients, which makes the Daubechies wavelets efficient for image compression [15].
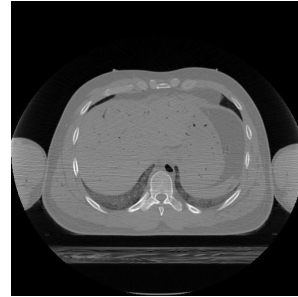
To compare the capabilities of the compression methods (DOST, FT, Daubechies), we conducted an experiment in which we applied each of the three methods to three different test images (shown in Figure 6.1) at different compression rates. The test images are all 512×512 pixels in size.

76

(a) Babara   (b) Lena   (c) CT

Figure 6.1: Original sample images.

Definition 2. *Peak Signal to Noise Ratio (PSNR)*

Peak signal to noise ratio is mostly defined via the mean squared error (MSE) which for two $m \times n$ monochrome images I and K where one of the images is considered a noisy approximation of the other is defined as:

$$\mathbf{MSE} = \frac{1}{m\,n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2.$$

The PSNR is defined as:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$
$$= 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right). \tag{6.1}$$

Tables 6.1-6.3 report the PSNR of the compressed images for our experiment. In all cases, the DOST method yields a substantially higher PSNR than the FT and Daubechies methods. In addition (though not reported in the tables), the maximum intensity errors are roughly the same between the DOST and the Daubechies methods.

Figure 6.2 compares the original Barbara and different compressed versions using DOST, FT and Daubechies-2 (compressing by 90%, in other words,

Table 6.1: PSNR for compression using 80% of coefficients.

Transform	Barbara	Lena	CT
DOST	90.27	88.87	86.69
FT	52.39	56.61	53.87
Haar	74.42	76.56	78.10
Daubechies-2	87.13	83.18	82.36
Daubechies-5	84.68	82.29	81.85
Daubechies-15	88.1665	84.59	80.22
Daubechies-38	81.9181	80.84	79.95

reconstructing using only 10% of the coefficients). As we can see, the DOST version remains sharper and keeps more detailed information (e.g. shadows behind the door, expression on the face and texture over the pants) than the wavelet version.

Figure 6.3 shows the corresponding intensity errors for different compression methods. The distribution of the non-zero elements hints at each method's strengths and weaknesses. In particular, the FT method exhibits its largest errors in the regions containing high-frequency content. The DOST method shows relatively small errors throughout. Similar observations are made over different compression rates.

To see more detailed comparison, as has been marked from the original image, four different regions with different textures are chosen and magnified. We can clearly see that the DOST compressed image has managed to maintain more original textures than the same level wavelet compression.

In Figure 6.9 and 6.10, we give another example for comparison. Similar comparison result can be achieved as above.

Table 6.2: PSNR for compression using 50% of coefficients.

Transform	Barbara	Lena	CT
DOST	55.17	53.20	52.45
FT	39.80	38.25	40.76
Haar	48.35	47.41	48.34
Daubechies-2	50.21	51.24	48.10
Daubechies-5	51.00	48.06	48.95
Daubechies-15	50.68	48.68	47.68
Daubechies-38	48.70	47.34	46.90

Table 6.3: PSNR for compression using 10% of coefficients.

Transform	Barbara	Lena	CT
DOST	34.31	33.40	33.25
FT	27.80	26.25	26.76
Haar	31.07	30.41	28.34
Daubechies-2	31.27	31.24	30.10
Daubechies-5	32.56	30.06	30.95
Daubechies-15	32.44	29.96	31.68
Daubechies-38	31.84	29.34	30.90

6.2 Conclusion and Discussion

Over the baseline comparison, the DOST is a valuable tool for image compression by giving a higher PSNR than the wavelet and the FT.

From the residual images of the experiment, Figure 6.3 and Figure 6.10, we can see that there are some block patterns in both wavelet and DOST methods.

79

(a) Original

(b) DOST compressed

(c) FT compressed

(d) Daubechies-2 compressed

Figure 6.2: Original and compressed versions of Barbara using 10% of coefficients.

However, we do find that the block pattern in the DOST residual is milder than the one in the wavelet residual. Even though there is no explicit standard to quantify the block pattern as one of the specifications of the image quality, this block does introduce artifacts. We analyze this phenomenon using the following experiment. First we set both the DOST coefficients and the wavelet coefficients of the image to zero, and then assign a random value to random positions of those

(a) FT
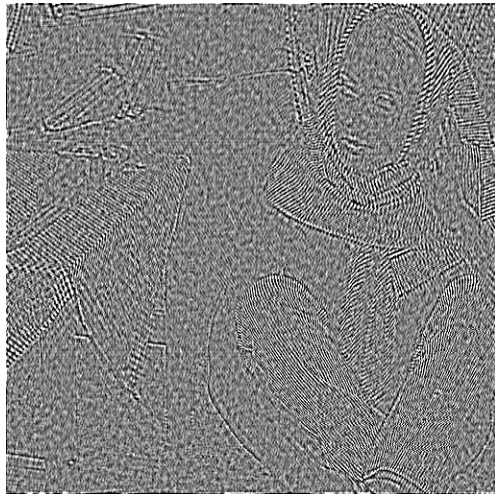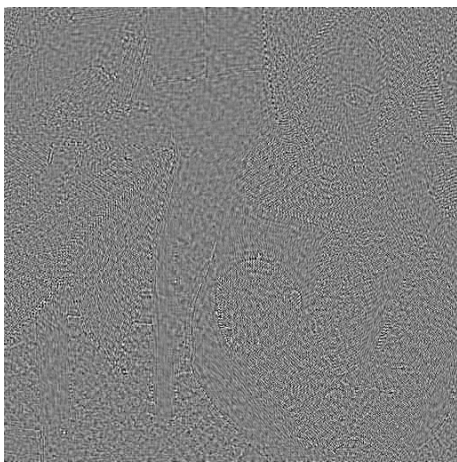
(b) DOST
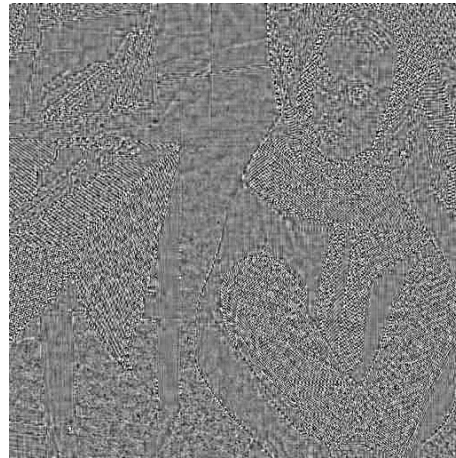
(c) Daubechies-2 wavelet

Figure 6.3:   Intensity errors for Barbara image using 10% of coefficients (see Table 6.3).
The gray level is set so that -20 maps to black and 20 maps to white.

Figure 6.4: Selected regions for detailed comparison.

two matrices. We run the corresponding reconstruction algorithms and study the reconstructed images. In the wavelet experiment, aside from the information we assigned, we see globally distributed information all over the image, which can be explained by the upsampling algorithm (2.35). Considering (2.35), we know that all the leaves of the tree are wavelet coefficients, which will contribute to the reconstruction. So, once a random coefficient is assigned, the original image would be affected over a larger area. The lower the coefficient is located in the pyramid tree, the further reaching its influence. Moreover, in real applications, the high frequency information, corresponding to the bottom positions of the tree, tends to be dropped, which consequently has a global influence on the image. On the other hand, as we can see from Figure (6.11) (b), a single DOST coefficient affects only a small region of the image on its vertical and horizontal directions, and thus introduces much milder blur to the image.

(a) Original

(b)DOST compressed

(c) Daubechies-2 wavelet compressed

Figure 6.5: Magnified comparison of Barbara's eyes.

We understand that the state-of-the-art compression methods that use the FT and WT are considerably more complex than the simplistic compression method used in our experiments. This project is simply a pilot study, and we believe it presents the motivation for more investigation which is needed to assess the degree to which the DOST can challenge the best compression methods.

(a) Original

(b)DOST compressed

(c) Daubechies-2 wavelet compressed

Figure 6.6: Magnified comparison of the hood on Barbara's right shoulder.

(a) Original

(b)DOST compressed

(c) Daubechies-2 wavelet compressed

Figure 6.7: Magnified comparison of the toy on the table.
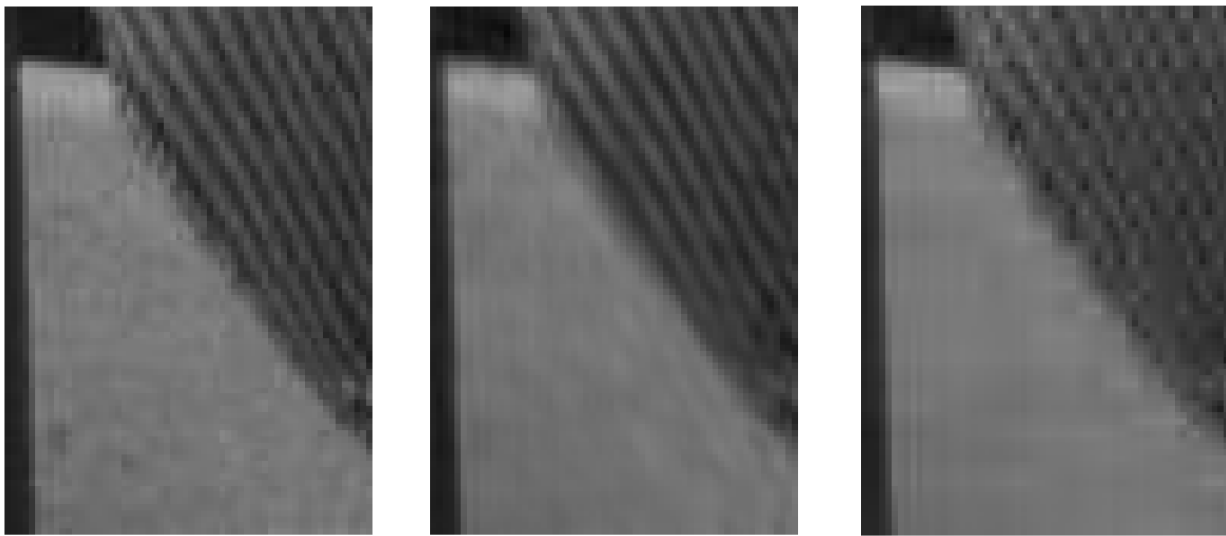
(a) Original   (b)DOST   (c) Daubechies-2 wavelet

Figure 6.8:   Magnified comparison of the table cloth.
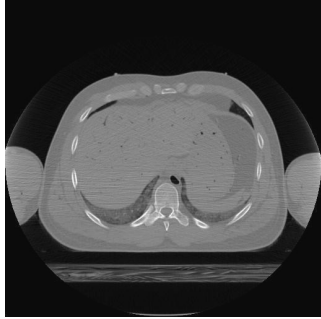
(a) Original
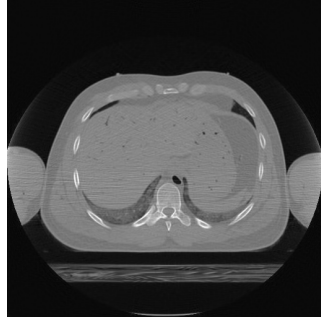
(b) DOST compressed
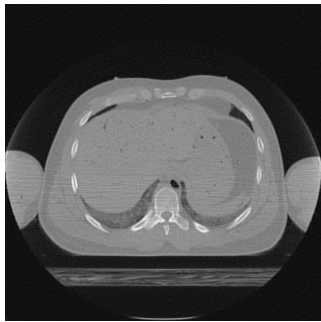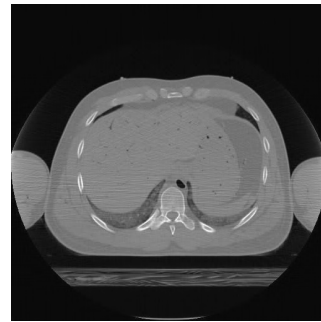
(c) FT compressed

(d) Daubechies-2 compressed

Figure 6.9: Original and compressed versions of the CT image using 10% of coefficients.

(a) FT
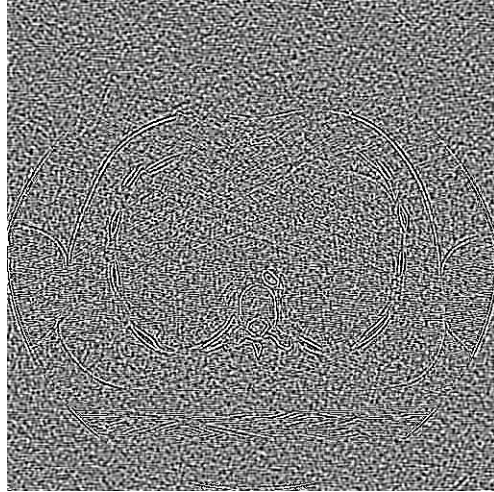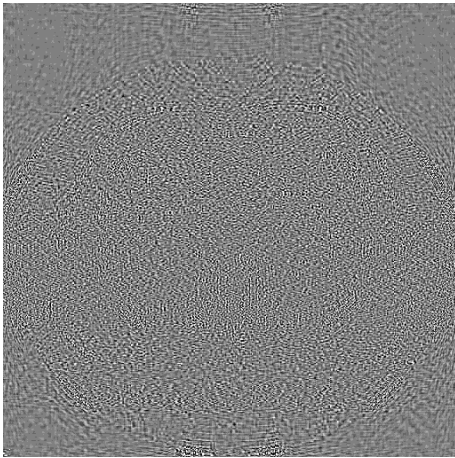
(b) DOST
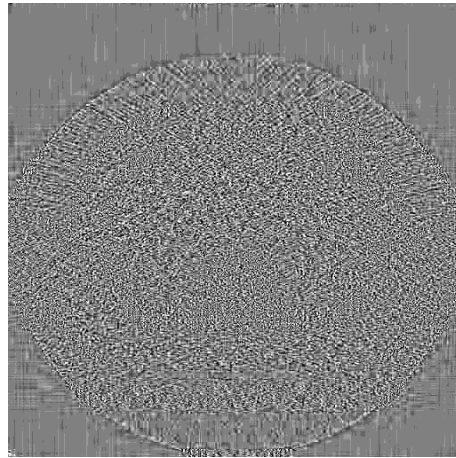
(c) Daubechies-2 wavelet

Figure 6.10:   Intensity errors for CT image using 10% of coefficients (see Table 6.3). The gray level is set so that -5 maps to black and 5 maps to white.
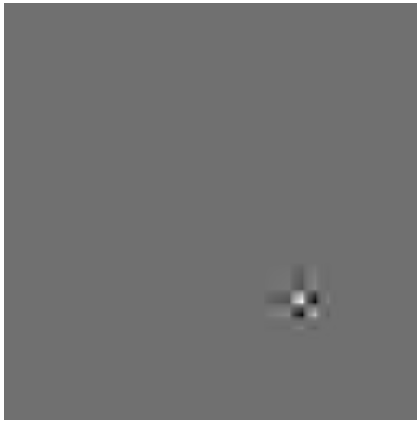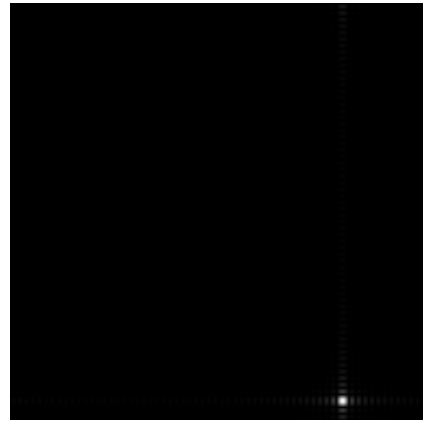
(a) Wavelet   (b) DOST

Figure 6.11: The reconstructed images from one wavelet coefficient or one DOST coefficient.

Chapter 7

Image Restoration Using the DOST

7.1 Introduction

Wavelets have been used to compress images since the 1970s. When the compressed images are electronically transferred, the wavelet coefficients are transferred. Due to the transportation, the routers and several other reasons, the coefficients could get partially lost. To restore the partially lost information we need to use a restoration algorithm so the image is acceptable to human eyes.

Even though wavelets have dominated image compression for years, other competitive methods for compression are being studied. In Ref. [46] and Chapter 5, a compression algorithm was described using the newly invented discrete orthonormal Stockwell transform (DOST) [35]. The possibility to restore a DOST-encoded image with its partially lost information (coefficients) will make the reality of DOST compression more convincing. These results have been published in ICIP [44] in 2009.

7.2   Method and Algorithm

In this section, we propose a total variation (TV) DOST based image inpainting model for the case of additive noise. We start with a standard image model,

$$u(x) = u_0(x) + n(x), \tag{7.1}$$

where $u_0(x)$ is the original noise free image and $n(x)$ the Gaussian white noise.

Assume $u_0(x)$ has the following DOST decomposition,

$$u_0(f, x) = \sum_{i,j} f_{i,j} D[x_1]_i D[x_2]_j, \tag{7.2}$$

where $f_{i,j}$ is the (i, j)th DOST coefficient (for simplicity of notation, the DOST coefficients for an image have two indices, hence form a 2-D matrix as a whole) and $x = (x_1, x_2) \in \mathbf{R}^2$ is the 2-D spatial coordinate.

Model: Define the total variation of the image as [8] [41]

$$\min_{d_{i,j}:(i,j)\in I} F(u; d) = \int_{\mathbf{R}^2} |\nabla_x u(d, x)| dx = \mathrm{TV}(u(d, x)), \tag{7.3}$$

where $u(d, x)$ is the damaged image and has the DOST decomposition:

$$u(d, x) = \sum_{i,j} d_{i,j} D[x_1]_i D[x_2]_j, \qquad i, j \in \mathbf{Z}, \tag{7.4}$$

with the constraint

$$d_{i,j} = f_{i,j}, \ (i, j) \notin \mathbf{I}, \tag{7.5}$$

where $\mathbf{I}$ is the index set of damaged or lost coefficients.

The TV minimization model has been used to get promising quality in denoising [41] and has started to benefit image restoration. In Ref. [8], Chan et al. presented a method to restore a wavelet-encoded image by minimizing the TV. In the above model, we let the first derivative of $F(u; d)$ – with respect to

the coefficients $d_{i,j}$ – equal zero and solve for d. Following similar analysis in section 2 and section 3 of Ref. [8], the minimization of TV is equivalent to solving the corresponding Euler-Lagrange Equation which is, for this case,

$$0 = \int \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) D[x_1]_i D[x_2]_j dx, \qquad (i,j) \in I. \tag{7.6}$$

The following explicit finite difference iterative algorithm has been used to numerically solve the Euler-Lagrange equation (7.6) and find the minimizer.

Algorithm:

1. Start with $n = 0$ and initial guess $d_{i,j}^{new} = d_{i,j}$. Set $d_{i,j}^{old} = 0$, and the initial error $E = ||d^{new} - d^{old}||_2$.

2. while $n < N$ or $E > \delta$, do

   - Set $d^{old} = d^{new}$,

   - Calculate d^{TV} as described after the main pseudo-code here.

   - For all (i, j), update

   $$d_{i,j}^{new} = d_{i,j}^{old} + \frac{\Delta t}{\Delta x} \lambda_{i,j}, \tag{7.7}$$

   where $\lambda_{i,j} = d_{i,j}^{TV} \chi_{i,j}$, and

   $$\chi_{i,j} = \begin{cases} 1 & \text{if } (i,j) \in \mathbf{I} \\ 0 & \text{if } (i,j) \notin \mathbf{I} \end{cases}$$

   - Compute error $E = ||d^{new} - d^{old}||_2$, and set $n = n + 1$.

Inside the algorithm above, the d^{TV}, which corresponds to the non-linear integral inside the minimization process, can be calculated using $u = \text{IDOST}(d)$,

where IDOST is the inverse DOST transform. For all (i, j), compute

$$
\begin{aligned}
\mathrm{curv}_{i,j} \;=\; & D_1^- \left(\frac{D_1^+ u_{i,j}}{\sqrt{|D_1^+ u_{i,j}|^2 + |D_2^+ u_{i,j}|^2 + \epsilon}} \right) \\
& + D_2^- \left(\frac{D_2^+ u_{i,j}}{\sqrt{|D_1^+ u_{i,j}|^2 + |D_2^+ u_{i,j}|^2 + \epsilon}} \right),
\end{aligned}
\tag{7.8}
$$

where

$$
D_1^+ u_{i,j} = u_{i+1,j} - u_{i,j},
$$

$$
D_2^+ u_{i,j} = u_{i,j+1} - u_{i,j}
$$

are the forward differences,

$$
D_1^- u_{i,j} = u_{i,j} - u_{i-1,j},
$$

$$
D_2^- u_{i,j} = u_{i,j} - u_{i,j-1}
$$

are the backward differences, and ϵ is a small positive number used to prevent the numerical blow-up.

The projection of the curvature on the DOST basis can then be calculated by $d^{TV} = \mathrm{DOST}(\mathrm{curv})$.

7.3   Results

For our tests, we used the synthetic image shown in Figure 7.1. To gain a valuable comparison, we also implemented the wavelet restoration method for wavelet-encoded images. The results from the numerical calculation are shown in Figure 7.2 and Figure 7.3.

As we can observe from Figure 7.2 and Figure 7.3, the DOST restoration has successfully restored more information than the wavelet restoration by offering a higher PSNR and more visible features and clearer edges in the image. In the

Figure 7.1: The original synthetic image for restoration test.

extreme case of randomly losing 90% of the coefficients, even though the original damaged image can hardly be recognized, the DOST restoration method still retains some of the features and edges.

Based on the above experiment, another fact can also be confirmed. When the two methods are faced with equivalently-degraded images (having lost the same percentage of coefficients), the DOST-encoded image recovers more information than the wavelet-encoded image, suggesting that DOST encoding is more resilient to packet loss.

7.4 Conclusion

The DOST restoration algorithm has outperformed wavelet restoration over the corresponding encoding techniques. We conjecture that even better performance can be realized using a more sophisticated restoration method that utilizes the self-similarity in the DOST coefficients.

(a) DOST degraded

(b) DOST restored
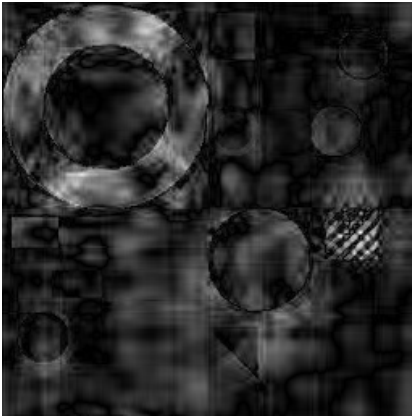
(c) Wavelets degraded
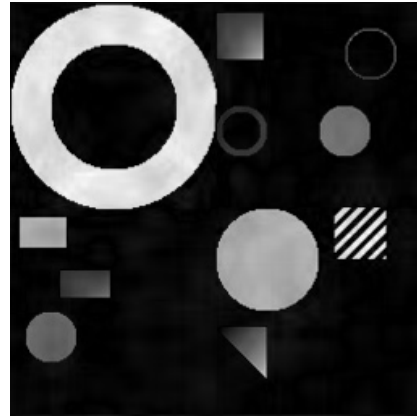
(d) Wavelets restored

Figure 7.2: Image restoration test for randomly losing 50% of the DOST and wavelet coefficients. (a) is the damaged DOST encoded image with PSNR=11.17 and (c) is the damaged wavelet encoded image with PSNR=10.15. (b) is restored using the DOST with PSNR=28.94. (d) is restored using the wavelets with PSNR=26.28. As we can see, the DOST-restored image is also sharper and clearer than the wavelet restored image.
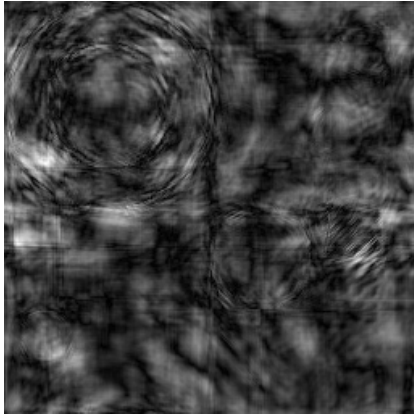
(a) DOST degraded
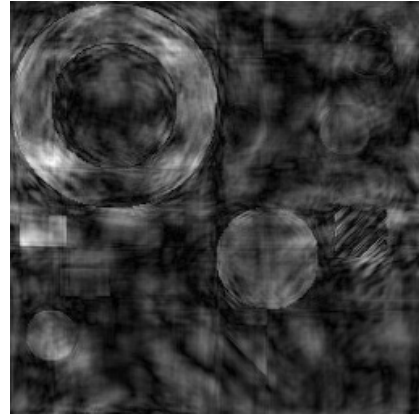
(b) DOST restored
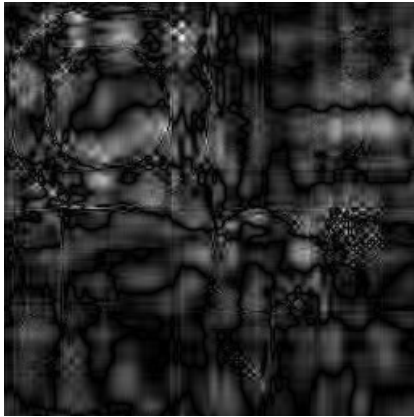
(c) Wavelets degraded
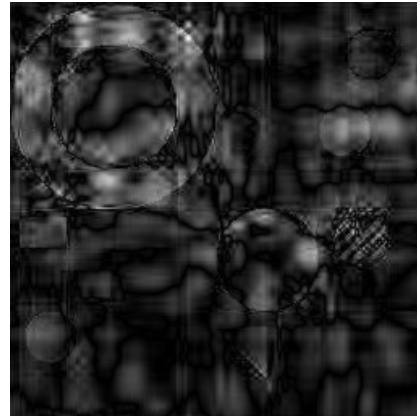
(d)Wavelets restored

Figure 7.3: Image restoration test for randomly losing 90% of the DOST and wavelet coefficients. (a) is the damaged DOST encoded image with PSNR=8.83 and (c) is the damaged wavelet encoded image with PSNR=8.75. (b) is restored using the DOST with PSNR=9.06. (d) is restored using the wavelets with PSNR=8.93. The 90% loss is an extreme test of the restoration, with heavy degradation of information in both images. Even though neither method can restore all features and edges, the DOST restoration method restores more visible image characteristics than the wavelet restoration.

Chapter 8

Conclusions

8.1 Conclusions

The continuous version and discrete version of the Stockwell transform (ST) offer the ability for multiresolution decomposition and provide absolutely-referenced frequency and phase information over the time-frequency domain. We compare the ST to other time-frequency analysis tools, such as the Gabor transform (GT) and the wavelet transform (WT), in terms of definitions, properties and performance in real experiments. Among its many advantages, the ST automatically supplies more resolution over high frequencies due to its self-adjusted Gaussian window. However, the ST of a 1-D signal is represented by a 2-D matrix, and the ST of a 2-D signal is represented by a 4-D matrix. Hence, the ST coefficients contain highly redundant information and require expensive calculation. We proved that (2.6) is an equivalent definition of the ST, enabling the use of the FFT for faster computation of the Stockwell coefficients. During the years of development of this young transform, various research groups have pushed efforts on how to utilize and reasonably visualize these high dimensional Stockwell coefficients. There have

been quite a few successful applications of ST in a variety of scientific fields such as geophysics and medical imaging.

Even though, in quality, the ST outperforms other transforms within the same category, calculation and manipulation of Stockwell coefficients for high-dimensional data sets are particularly difficult, and even impossible for some scenarios. In addressing the redundancy issue of the ST, a feasible strategy of partitioning the time-frequency domain was adopted and applied here to the construction of the basis functions, yielding an orthogonal version of the ST transform, the DOST. The DOST has inherited the multiresolution feature of the ST, but significantly reduces the storage and calculation cost. The DOST transforms an N-tuple 1-D input signal to an N-tuple 2-D output. We discussed 2-D visualization conventions to recover back the 2-D nature of the 1-D output generated by the DOST transform.

The DOST has distinguished itself due to its perfect frequency localization, multiresolution decomposition ability and due to many other advantages in terms of time-frequency analysis. The basis functions of the DOST are complex conjugate to their symmetric counterparts. Thus, for a real-valued input, only half of the DOST coefficients need to be calculated. However, the brute-force algorithm for computing the DOST is $\mathcal{O}(N^2)$ for an input of size N. An efficient calculation strategy was suggested in the original DOST paper [35], but we formulated a matrix decomposition of the DOST, which offers great convenience for its theoretical analysis. The matrix form reveals an $\mathcal{O}(N \log N)$ algorithm for computing the DOST. As shown both in the proof of the complexity and in the numerical experiments, the ratio of the DOST computing time to the FFT computing time is no more than 6. The comparison was run under the Matlab environment, which can be slow at looping.

We analyzed the time-frequency partitioning of the DOST in the context of the sampling theorem; to our knowledge, this is the first such analysis. We also explicitly define an alternative symmetric version of the DOST [43], in which the definition of the DOST basis functions are modified by adjusting the parameters of ν, β and τ. By doing this, we successfully adjusted the partitioning of the frequency axis, while maintaining the complex conjugate symmetry, the orthogonality between basis functions and, more importantly, the fast algorithm. This theoretical branch of the DOST can be extended freely to offer different versions of the DOST with arbitrary frequency windowing and temporal sampling to benefit applications in different situations.

The analysis on the translation property of the DOST is new and opens the door to applications in the field of signal or image tracking. Theories in global translation are perfectly built and sophisticatedly proven. A global translation can be observed in the phase change of one single DOST coefficient. Or, a global translation can be generated by directly manipulating the DOST coefficients of the original image. For local translations, the uncertainty principle between the Fourier domain and the temporal domain confounds our search for a simple and exact solution. The mathematical model we offered might contribute to solving this problem in an approximate sense in the near future.

Researchers are putting more focus on using the DOST and more applications of the DOST are being unearthed. In 2008, Dr. Mitchell and his group published an application using the DOST for image segmentation [16]. In their application, the horizontal and vertical frequency patterns are characterized using the DOST coefficients to differentiate the different image textures. Their result has been shown to be comparably better than the wavelet version of segmentation. In the field of image compression, we verify that the DOST has potential advantages in

giving a better approximation of the original image than the wavelet method [46]. It turns out that the DOST compression can significantly suppress the block phenomenon (commonly found in the wavelet compression) and produces an approximation with higher PSNR. In the modern JPEG compression algorithm using the wavelet, the local wavelet transform is used. It will be exciting to see if the local DOST transform can benefit image compression. The image restoration ability of the DOST has also been investigated (Chapter 6) [44]. The results based on the DOST restoration algorithm offer more recognizable objects and sharper edges than the wavelet algorithm, especially when the loss of the coefficients is more severe.

8.2 Future Work

The absolutely-referenced phase information is one of the many promising properties of the DOST. We will persist in our research of the phase properties of the DOST and apply this knowledge to a wider range of applications in various fields. As mentioned above, the local application of the DOST in an image might have potential and could be expected to achieve better image approximation with lower cost of storage than the wavelet. Considering the close relation between the DOST and the FT, it is conceivable that the DOST could benefit data acquisition in MRI, and could thus reduce the acquisition time and significantly improve the image quality in conjunction with various techniques such as compressive sensing.

At the theoretical stage, even though a comprehensive understanding of the DOST has been achieved in combining the original DOST paper [35] and our thesis here, there are still many deeper fields that can be touched with respect to

the DOST. The inherent nature of the DOST (the independent parameters in the basis, ν, β and τ) offers a great flexibility in the definition of different versions of the DOST. Usually, since each transform (such as, the Fourier transform, the Gabor transform and the wavelet transform) has nearly fixed freedom in the frequency windowing and the temporal sampling, the DOST will head a new tide of freely-adjusting windowing and sampling according to different requirements of different applications. To do this, a sophisticated theoretical study of how to use the sampling theorem to guide new definitions of the DOST needs to be elaborated and will form a fairly large branch in the DOST research of the future.

The explicit relationship between the DOST coefficients before and after the global translation is stated as the DOST shift theorem. This shift theorem offers the straightforward methods to manipulate the DOST coefficients and detect the global translation, and thus highlights its advantage of dealing with the offset in real applications such as Multi-Carrier Modulation. Compared to other multiresolution transformation tools such as the wavelet transform, the DOST might be deeply applied in similar areas. The local translation using the DOST turns out to be a hard problem due to the uncertainty principle between the Fourier domain and the time domain. A certain compromise needs to be achieved to solve this issue. How precisely a local translation can be determined using the DOST coefficients has been left as an interesting question for the future.

With the fast algorithm, the DOST can feasibly be applied in higher dimensional applications. We are interested to see if more higher dimensional applications will come to light and help this powerful young transform grow up.

APPENDIX A

Proof of linearity and modulation properties of the ST

Proof of (2.23), linearity:

Proof.

$$
\begin{aligned}
& \mathcal{S}\{ah(t) + bg(t)\} \\
= \ & \int_{-\infty}^{\infty} (ah(t) + bg(t)) \frac{|f|}{2\pi} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt \\
= \ & a \int_{-\infty}^{\infty} h(t) \frac{|f|}{2\pi} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt + b \int_{-\infty}^{\infty} g(t) \frac{|f|}{2\pi} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt \\
= \ & a\mathcal{S}\{h(t)\} + b\mathcal{S}\{g(t)\}.
\end{aligned}
$$

$\square$

Proof of (2.24), modulation:

Proof.

$$\mathcal{S}\{h(t - t_0)\}$$

$$= \int_{-\infty}^{\infty} h(t - t_0) \frac{|f|}{2\pi} e^{-\frac{(\tau - t)^2 f^2}{2}} e^{-i2\pi ft} dt$$

$$= \int_{-\infty}^{\infty} h(t - t_0) \frac{|f|}{2\pi} e^{-\frac{(\tau - t_0 - (t - t_0))^2 f^2}{2}} e^{-i2\pi f(t - t_0 + t_0)} d(t - t_0)$$

$$= \int_{-\infty}^{\infty} h(s) \frac{|f|}{2\pi} e^{-\frac{(\tau - t_0 - s)^2 f^2}{2}} e^{-i2\pi f(s + t_0)} ds \qquad (\text{Let } s = t - t_0)$$

$$= e^{-i2\pi ft_0} S(\tau - t_0, f).$$

$\square$

APPENDIX B

Fast DOST code and 2-D Visualization Code

Listing 1: 2-D Fast DOST Decomposition.

```matlab
function  D2_coeff = FOST_2D_DEC( S_input )
%       2−Dimensional DOST:
%       tic
    si = size(S_input);
    D2_coeffM = zeros(si);
    D2_coeff = zeros(si);


    %rows first
    for  ii = 1:si(1)
        timeseries = S_input(ii, :);
        D2_coeffM(ii,:) = FOST_1D_DEC(timeseries);
    end
    for  jj = 1:si(2)
```

```
        timeseries = D2_coeffM(:, jj);
        D2_coeff(:, jj) = FOST_1D_DEC(timeseries);
    end
%       Decomposition_Time = toc
end
```

Listing 2: 1-D Fast DOST Decomposition.

```
function bas2 = FOST_1D_DEC(S_input)
    si = size(S_input);
    NN = max(si);
    [vv, bb, tt, num_it, compact_b] = get_vbt(NN);


    %Calculate the FFT
    fftdata = transpose(fftshift(fft(S_input)));


%       tic
    %Use the inverse FFT to calculate the DOST coefficients
    bas2 = zeros(si);
    init_p = 0;
    coof = -pi * 1i;
    for ll = 1:num_it
        step = compact_b(ll);
        %Construct the corresponding Ramp matrix
        Ramp = zeros(1, step);
        for ii = 1:step
            Ramp(ii) = (-1)^ii;%
```

```matlab
        end
    temp_matrix = ...
        ifft(fftdata(init_p+1:init_p+step))*sqrt(step);
    if size(Ramp) == size(temp_matrix)
        bas2(init_p + 1:init_p + step) = ...
            Ramp .* temp_matrix;
    else
        bas2(init_p + 1:init_p + step) = ...
            Ramp .* transpose(temp_matrix);
    end
    if vv(init_p + 1) <0
        bas2(init_p + 1:init_p + step) = ...
            bas2((init_p + step):-1:(init_p + 1));
        Ramp2 = ...
            exp(tt(init_p+(1:step))*2*pi*1i/step);
        if size(Ramp2) == ...
                size(bas2(init_p + 1:init_p + step))
            bas2(init_p + 1:init_p + step) = ...
                bas2(init_p + 1:init_p + step).*Ramp2;
        else
            bas2(init_p + 1:init_p + step) = ...
                bas2(init_p + 1:init_p + step) ...
                .* transpose(Ramp2);
        end
    end
```

```matlab
        clear temp_matrix;

        clear Ramp;

        init_p = init_p + step;

    end

    bas2 = - bas2 / NN;
%       FOST_decom_time = toc
end
```

Listing 3: Parameter Assigning Function.

```matlab
function [v, b, t, num, com_b] = get_vbt(NN);

    count = 0;

    end_P = log2(NN)-1;

    b = zeros(1, NN);

    v = zeros(1, NN);

    t = zeros(1, NN);

    com_b = zeros(1, 2 * end_P + 2);

    num = 0;


    %%%%%%NNegative frequency

    count = count + 1;

    v(count) = -NN/2 + .5; b(count) = 1; t(count) = 0;

    num = num + 1;

    com_b(num) = b(count);


    for p = end_P:-1:2

        freq = -2^(abs(p) - 2) * 3 + 1;
```

107

```matlab
        beta = 2^(abs(p) - 1);
        for tau = beta - 1:-1:0
%          for tau = 0: beta - 1
            count = count + 1;
            v(count) = freq;
            b(count) = beta;
            t(count) = tau;
        end
        num = num + 1;
        com_b(num) = beta;
    end


    count = count + 1;
    v(count) = -0.5; b(count) = 1; t(count) = 0;
    num = num + 1;
    com_b(num) = b(count);


    %%%%%%Positive frequency
    count=count + 1;
    v(count) = 0.5; b(count) = 1; t(count) = 0;
    num = num + 1;
    com_b(num) = b(count);


    count=count + 1;
    v(count) = 1.5; b(count) = 1; t(count) = 0;
```

```
    num = num + 1;
    com_b(num) = b(count);


    for p = 2:end_P
        freq = 2^(abs(p) - 2) * 3;
        beta = 2^(abs(p) - 1);
        for tau=0:beta - 1
            count=count + 1;
            v(count) = freq;
            b(count) = beta;
            t(count) = tau;
        end
        num = num + 1;
        com_b(num) = beta;
    end
end
```

Listing 4: 2-D Visualization of the N-tuple DOST coefficients of an N-tuple signal.

```
function Diag1 = diag_1D(dost_cof)
% This function returns a 2-D matrix according
% to the DOST diagram for visualization.
    NN = max(size(dost_cof));
    Diag1 = zeros(NN);


    [vv, bb, tt, num_it, compact_b] = get_vbt(NN);
    ind_ii = NN;
```

```
ind_jj = 1;

for ll = 1:NN
    hei_ll = bb(ll);
    len_ll = NN/hei_ll;
    for ii = ind_ii:-1:(ind_ii - hei_ll + 1)
        for jj = ind_jj:(ind_jj + len_ll - 1)
            Diag1(ii, jj) = dost_cof(ll);
        end
    end
    if bb(ll) ~= 1 && ll < NN
        if bb(ll) ~= bb(ll + 1)
            ind_ii = ind_ii - hei_ll;
            ind_jj = 1;
        else
            ind_jj = ind_jj + len_ll;
        end
    else
        ind_ii = ind_ii - hei_ll;
        ind_jj = 1;
    end
end

for ii = (NN/2 + 1):NN
    Diag1(ii, 1:NN) = Diag1(ii, NN:-1:1);
```

```matlab
    end

%       figure(1),
%       imshow(Diag1, []);
end
```

Bibliography

[1] R. Ahmad and T. Kundu, *Application of Gabor transform on cylindrical guided waves to detect defects in underground pipes*, Health Monitoring of Structural and Biological Systems, 6532 (2007).

[2] S. Andino, R. Menendez, C. Lantz, O. Blank, C. Michel, and T. Landis, *Non-stationary distributed source approximation: An alternative to improve localization procedures*, Human Brain Mapping, 14 (2001), pp. 81–95.

[3] S. Assous, A. Humeau, M. Tartas, P. Abraham, and J. L'Huillier, *Physiological effects of indomethacin and celecobix: An S-transform laser Doppler flowmetry signal analysis*, Physics in Medicine and Biology, 50 (2005), pp. 1951–1959.

[4] A. Boggess and F. J. Narcowich, *A First Course in Wavelets with Fourier Analysis*, Prentice Hall, 2001. 260 pages.

[5] R. N. Bracewell, *The Fourier Transform and Its Applications*, Graw-Hill, New York, 2nd ed., 1986.

[6] P. Burt and E. Adelson, *The laplacian pyramid as a compact image code*, Communications, IEEE Transactions on, 31 (1983), pp. 532 – 540.

[7] P. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*, Oxford University Press, 1991.

[8] T. F. Chan, J. Shen, and H. min Zhou, *Total variation wavelet inpainting*, J. Math. Imaging Vision, 25 (2006), pp. 107–125.

[9] C. Christopoulos, A. Skodras, and T. Ebrahimi, *The JPEG2000 still image coding system: An overview*, Consumer Electronics, IEEE Transactions, 46 (2000), pp. 1103–1127.

[10] C. K. Chui and A. K. Chan, *Spline-wavelet signal analyzers and methods for processing signals.* Patent, April 1991.

[11] J. W. Cooley and J. W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Mathematics of Computation, 19 (1965), pp. 297–301.

[12] A. Coulson, *A generalization of nonuniform bandpass sampling*, Signal Processing, IEEE Transactions on, 43 (1995), pp. 694 –704.

[13] P. Dash, B. Panigrahi, and G. Panda, *Power quality analysis using S-transform*, IEEE Transactions on Power Dilivery, 18 (2003).

[14] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Communications on Pure and Applied Mathematics, 41 (1988), pp. 909–996.

[15] ⸺, *Ten Lectures on Wavelets*, vol. 61, SIAM, University of Lowell, Philadelphia, 1992.

[16] S. Drabycz, R. G. Stockwell, and J. R. Mitchell, *Image texture characterization using the discrete orthonormal S-transform*, Journal of Digital Imaging, (2008).

[17] M. Eramian, R. A. Schincariol, R. G. Stockwell, R. P. Lowe, and L. Mansinha, *Review of applications of 1-D and 2-D S-transforms*, Proc. SPIE, 3078 (1997), pp. 558–568.

[18] T. C. Farrar and E. D. Becker, *Pulse and Fourier transform NMR: Introduction to Theory and Methods*, Academic Press, September 1971.

[19] D. Gabor, *Theory of communication. part 1: The analysis of information*, Electrical Engineers - Part III: Radio and Communication Engineering, Journal of the Institution of, 93 (1946), pp. 429 –441.

[20] P. C. Gibson, M. P. Lamoureux, and G. F. Margrave, *Letter to the editor: Stockwell and wavelet transforms*, Journal of Fourier Analysis and Applications, 12 (2006), pp. 713–721.

[21] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press Inc, 7th revised ed.

[22] H. Harada and R. Prasad, *Simulation and Software Radio for Mobile Communications*, Artech House, Inc., Norwood, MA, USA, 2002.

[23] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, October 1988.

[24] D. Karamehmedovic, M. K. Lakshmanan, and H. Nikookar, *Performance evaluation of wpmcm with carrier frequency offset and phase noise*, JCM, 4 (2009), pp. 496–508.

[25] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2nd ed., 1999.

[26] L. Mansinha, R. G. Stockwell, R. Lowe, M. Eramian, and R. A. Schincariol, *Local S-spectrum analysis of 1-D and 2-D data*, Phys. Earth Planet. Inter., 103 (1997), p. 329C336.

[27] L. Mansinha, R. G. Stockwell, and R. P. Lowe, *Pattern analysis with two-dimensional spectral localisation: Application of two-dimentional S transforms*, Physica A, 239 (1997), pp. 286–295.

[28] P. McFadden, J. Cook, and L. Forster, *Decomposition of gear vibration signals by the generalised S-transform*, Mechanical Systems and Signal Processing, 13 (1999), pp. 691–707.

[29] E. Merzbacher, *Quantum Mechanics*, John Wiley & Sons, 2nd ed., 1970.

[30] E. G. Merzlyakov, Y. I. Portnyagin, C. Jacobi, N. J. Mitchell, H. G. Muller, A. H. Manson, A. N. Fachrutdinova, W. Singer, and P. Hoffmann, *On the longitudinal structure of the transient day-to-day variation of the semidiurnal tide in the mid-latitude lower thermosphere—i. Winter season*, Ann. Geophys., 19 (2001), pp. 545–562.

[31] S. H. Nawab and T. F. Quatieri, *Short-time Fourier transform*, (1987), pp. 289–337.

[32] G. A. Oldenborger, R. A. Schincariol, and L. Mansinha, *Space-local spectral texture segmentation applied to characterizing the heterogeneity of hydraulic conductivity*, Water Resouces Research, 38 (2002), pp. 1154–1157.

[33] G. A. Oldenborger, R. A. Schincariol, and L. Mansinha, *Radar determination of the spatial structure of hydraulic conductivity*, Ground Water, 41 (2003), pp. 24–32.

[34] C. R. Pinnegar and L. Mansinha, *The Bi-Gaussian S-transform*, SIAM Journal on Scientific Computing, 24 (2002), pp. 1678–1692.

[35] R. G. Stockwell, *A basis for efficient representation of the S-transform*, Digital Signal Processing, 17 (2007), pp. 371–393.

[36] R. G. Stockwell and R. Lowe, *Airglow imaging of gravity waves—1. Results from a small network of OH nightglow scanning imagers*, Journal of the Geophysical Research, 106 (2001), pp. 185–203.

[37] ——, *Airglow imaging of gravity waves—2. Critical layer filtering*, Journal of the Geophysical Research, 106 (2001), pp. 205–220.

[38] R. G. Stockwell, R. P. Lowe, and L. Mansinha, *Localized cross spectral analysis with phase corrected wavelets*, Wavelet Applications III, 2762 (1996), pp. 557–564.

[39] R. G. Stockwell, L. Mansinha, and R. P. Lowe, *Localization of complex spectrum: the S transform*, IEEE Transactions on Signal Processing, 144 (1996), pp. 998–1001.

[40] M. Varanini, G. Paolis, M. Emdin, A. Macerata, S. Pola, M. Cipriani, and C. Marchesi, *Spectral analysis of cardiovascular time series by the S-transform*, Computers in Cardiology, (1997), pp. 383–386.

[41] C. R. Vogel and M. E. Oman, *Iterative methods for total variation denoising*, SIAM J. Sci. Comput, 17 (1996), pp. 227–238.

[42] G. K. Wallace, *The JPEG still picture compression standard*, Source Communications of the ACM, 34 (1991), pp. 30–44.

[43] Y. WANG AND J. ORCHARD, *Symmetric discrete orthonormal Stockwell transform*, American Institute of Physics (AIP) Conference Proceedings, 1048 (2008), pp. 585–588.

[44] ———, *The discrete orthonormal stockwell transform for image restoration*, ICIP'09: Proceedings of the 16th IEEE international conference on Image processing, (2009), pp. 2725–2728.

[45] ———, *Fast discrete orthonormal stockwell transform*, SIAM Journal on Scientific Computing, 31 (2009), pp. 4000–4012.

[46] ———, *On the use of the stockwell transform for image compression*, IS&T/SPIE Electronic Imaging: Algorithms and Systems VII, 7245 (2009), p. 724504.