Time Frequency Analysis for Motion Magnification and Detection

Thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science (by Research)
in
Electronics and Communication Engineering

by

Sushma M
201032006
sushma.m@research.iiit.ac.in

International Institute of Information Technology
Hyderabad - 500 032, INDIA
May 2015

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "Time Frequency Analysis for Motion Magnification and Detection" by Sushma M (201032006), has been carried out under my supervision and is not submitted elsewhere for a degree.

Date	Advisor: Prof. Jayanthi Sivaswamy
Date	Advisor: Dr. Anubha Gupta

To My Family

Acknowledgements

I am deeply indebted to my advisors Prof. Jayanthi Sivaswamy and Dr. Anubha Gupta for their dedication, encouragement, motivation and also for inspiring guidance and supervision throughout my thesis work. I give my warmest gratitude to Satyanarayana (CVIT) for his unreserved support related to administrative matters. Thanks to Ravi Shekhar for helping and encouraging me throughout my research work. Thanks to Sumit, Ravi Shankar, Swagatika, Vikram and SRK for making life enjoyable during my stay at IIIT.

I would also like to express my gratitude towards Mom, Dad, my sister Reshma and my brother Suman Chaithanya for their support.

Abstract

Motion can be defined as change in position of an object of interest with respect to time. This thesis explores the methods of analyzing motion using time frequency analysis. In this thesis, we address two problems: (i) Small Motion Magnification in Videos and (ii) Motion Detection in Perfusion Weighted Imaging (PWI).

Human eye and its brain interface can visualize or detect the motion within a certain range of spatial and temporal frequencies. But in most of the cases, it might be possible that frequencies which are below this range also can have useful information. We can simplify this by saying that there can be small motions which are not visible to the naked eye. Even though these small motions are difficult to detect, they may contain useful information. In first part of thesis, we present a semi-automated method to magnify small motions in videos. This method amplifies invisible or hidden motions in videos. To achieve motion magnification, we process the spatial and temporal information obtained from the video itself. Advantage of this work is that it is application independent. Proposed technique estimates required parameters to get desirable results. We demonstrate performance on a few videos. Motion magnification performance is equivalent to existing manual methods.

In second part of thesis, we present a novel automated method to detect motion in perfusion weighted images (PWI), which is a type of magnetic resonance imaging (MRI). In PWI, blood perfusion is measured by injecting an exogenous tracer called bolus into the blood flow of a patient and then tracking it in the brain. PWI requires a long data acquisition time to form a time series of volumes. Hence, motion occurs due to patient's unavoidable movements during a scan, which in turn results into motion corrupted data. There is a necessity of detection of these motion artifacts on captured data for correct disease diagnosis. In PWI, intensity profile gets disturbed due to occurrence of motion and/or bolus passage through the blood vessels. In this work, we propose an efficient time-frequency analysis based motion detection method. We show that proposed method is computationally inexpensive and fast. This method is evaluated on a DSC-MRI sequence with simulated motion of different degrees. We show that our approach detects motion in a few seconds.

Contents

Chapter Page

1 Introduction . . . 1
1.1 Problem Overview . . . 3
1.2 Contribution . . . 4
1.3 Thesis Organization . . . 4

2 Background and Previous Work . . . 5
2.1 Stockwell Transform and its Applications . . . 5
2.1.1 Evolution of Stockwell Transform . . . 5
2.1.1.1 Fourier Transform . . . 5
2.1.1.2 Short Time Fourier Transform . . . 6
2.1.1.3 Gabor Transform . . . 8
2.1.1.4 Wigner Ville Distribution . . . 8
2.1.1.5 Continuous Wavelet Transform . . . 8
2.1.1.6 Stockwell Transform . . . 9
2.1.1.7 Discussion . . . 10
2.1.2 Applications of ST . . . 12
2.2 Small Motion Magnification in Videos . . . 13
2.3 Motion Detection in Perfusion Weighted MRI for Brain . . . 14
2.3.1 Perfusion Weighted MRI . . . 14
2.3.1.1 Dynamic Susceptibility Contrast Imaging (DSC) . . . 14
2.3.1.2 Dynamic Contrast Enhanced Imaging (DCE) . . . 15
2.3.2 Motion Detection . . . 16
2.4 Summary and Comments . . . 17

3 Estimation of Parameters for Automated Magnification of Small Motions in Videos using ST . 18
3.1 Introduction . . . 18
3.2 Background . . . 19
3.2.1 Eulerian Video Magnification . . . 19
3.2.1.1 First Order Motion . . . 20
3.2.1.2 Bounds . . . 21
3.2.1.3 Multiscale Analysis . . . 22
3.2.2 Discussion . . . 23
3.3 Proposed Method for Estimating Parameters for Automated Magnification of Small Motions in Videos . . . 23
3.3.1 Landmark Detection . . . 24

3.3.2 Time-Frequency Analysis . . . 25
3.3.2.1 Stockwell Transform . . . 25
3.3.3 Parameter Estimation . . . 26
3.3.4 Motion Magnification . . . 28
3.4 Experiments and Results . . . 28
3.4.1 Experimental Setup . . . 28
3.4.2 Results . . . 30
3.5 Summary . . . 30

4 Motion Detection in Perfusion Weighted MRI using ST . . . 31
4.1 Introduction . . . 31
4.2 Phase Correlation based Motion Detection . . . 32
4.3 Challenges in Motion Detection in PWI . . . 34
4.4 Stockwell Transform . . . 36
4.5 Proposed Motion Detection Method . . . 38
4.5.1 Pre-processing . . . 39
4.5.2 Estimation of Landmark Pixels . . . 39
4.5.3 Detection of Presence of Motion . . . 40
4.5.4 Time-Frequency Analysis . . . 42
4.5.5 Motion Detection . . . 42
4.6 Experiments and Results . . . 43
4.7 Summary . . . 45

5 Conclusions and Future Work . . . 47
5.1 Summary and Conclusion . . . 47
5.2 Future Work . . . 48

Bibliography . . . 50

List of Figures

Figure Page

1.1 Extraction of time series. A video $\{I_t\}_{t=1}^{T}$ with 'k' landmarks ($\{h_i(t)\}_{i=1}^{k}$) is shown. Time series ($\{h_i(t)\}_{i=1}^{k}$) at these 'k' landmark pixels are extracted. Landmarks and corresponding time series are shown in same colour. 2

2.1 Chirp Signal. 6
2.2 Different Time Frequency Representations for a Chirp Signal. 11

3.1 Overview of the Eulerian video magnification method. This Figure is adapted from [75]. The input video sequence is decomposed into different spatial frequency bands, and the same temporal filter is applied to all bands. The filtered spatial bands are then amplified by a given factor α, added back to the original signal, and collapsed to generate the output video. 19
3.2 Approximation of spatial translation using temporal filtering. This Figure is adapted from [75]. 20
3.3 Motion magnification factor, α, as function of spatial wavelength λ. This Figure is adapted from [75]. Magnification factor is fixed to α for spatial bands that are within derived bound (Equation 3.16), and is attenuated linearly for higher spatial frequencies. 23
3.4 Overview of the proposed framework. Initially, parameters (ω_l, ω_h, α, λ_c) are estimated from the landmark pixels ($\{l_i\}_{i=1}^{k}$) determined from the first two frames of input video ($\{I_t\}_{t=1}^{2}$). Then input video ($\{I_t\}_{t=1}^{N}$) is divided into spatial bands, which are pixel-wise filtered with same temporal filter and amplified by magnification factor, α. Reconstruction ($\{\tilde{I}_t\}_{t=1}^{N}$) is done by adding these amplified bands along with original spatial bands. 24
3.5 Estimation of landmark pixels. 25
3.6 Extraction of time series. 25
3.7 Applying Stockwell transform on time series. 26
3.8 An example of Stockwell transform. Time signal shown in left is $h[0:63] = \cos(2\pi t * 6/128), h[64:127] = \cos(2\pi t * 25/128), h[20:30] = h[20:30] + 0.5 * \cos(2\pi t * 52/128)$. It contains a low-frequency signal for the first half, a middle-frequency signal for the second half and a high-frequency burst at $t = 20$. All these frequencies are clearly visible along with time location in Stockwell transform as shown in right (bright pixels indicate high strength of transform). 27
3.9 Sample frames from videos used in experimentation (a)*baby* (b)*guitar* (c) *shadow* (d)*face2* (e)*wrist* (f)*baby2* (g)*face* (h)*camera* (i)*subway*. All these are shown in proportion with the size of the corresponding video. 28

4.1 Motion Detection Scheme for a DSC-MRI time-series. This Figure is adapted from [55]. 32
4.2 Gamma-variate-function (GVF) fitting on the mean-intensity curve of a DSC-MRI time-series. This Figure is adapted from [55]. 33
4.3 Stockwell transform for an impulse function. Impulse function shown in left is $h[0:39] = 0, h[20] = 1$. Stockwell transform is shown in right. Note that bright pixels indicate high strength of transform. Here, bright pixels are at $t = 20$. 35
4.4 Motion Detection in both bolus and non-bolus stages. Central slices $\left(\{C_t\}_{t=1}^{N}\right)$ of a PWI MRI series are pre-processed to get noise-free images $\left(\{I_t\}_{t=1}^{N}\right)$. Then, two sets of one dimensional time series $\{h_{l_1}(t)\}_{l_1=1}^{L_1}$ and $\{h_{l_2}(t)\}_{l_2=1}^{L_2}$ at two sets of landmark pixels $\{p_{l_1}\}_{l_1=1}^{L_1}$ and $\{p_{l_2}\}_{l_2=1}^{L_2}$ are extracted from difference of consecutive pre-processed slices $\left(\{D_{t,t+1}\}_{t=1}^{N-1}\right)$. These two time series are analyzed to know whether there is motion present or not. if motion is present, time-frequency analysis $(\{S_{l_1}(\tau, f)\}_{l_1=1}^{L_1})$ is used to determine the locations of corrupted slices $\left(\{l_m\}_{m=c}^{c+M-1}\right)$. 39
4.5 Central slices with no motion corruption and corresponding time series $h_1(t)$ and $h_2(t)$. $h_1(t)$ and $h_2(t)$ are shown in red and blue colors respectively. Here, conditions involved are, (i) $h_1(t) > h_2(t)$ and $h_2(t) = 0$; (ii) $h_2(t) > h_1(t)$. 41
4.6 Central slices with for motion in bolus affected slices and corresponding time series $h_1(t)$ and $h_2(t)$. Here, conditions involved are, (i) $h_1(t) \geq h_2(t)$ and $h_2(t) > 0$; (ii) $h_2(t) > h_1(t)$. 41
4.7 Central slices with motion in both bolus affected and non-affected slices and corresponding time series $h_1(t)$ and $h_2(t)$. Here, conditions involved are, (i) $h_1(t) \geq h_2(t)$ and $h_2(t) > 0$; (ii)$h_2(t) > h_1(t)$. 42
4.8 DSC-MRI time series. It shows central slices of 40 volumes from top to bottom and left to right. Here, central slices corresponding to non-bolus phase are from 1 to 9 and from 21 to 40. -15^0 rotation is added to volumes 3, 4, 5, 6, 7, 12, 13, 14, 15, 27, 28, 29, 30, 31 and 32 respectively and their corresponding central slices are shown in boxes. Motion in non-bolus phases (3, 4, 5, 6, 7, 27, 28, 29, 30, 31 and 32) and bolus phases (12, 13, 14 and 15) are shown in red and blue boxes respectively. 43

List of Tables

Table Page

3.1 Video Clips used in Testing. 29
3.2 Estimated parameters for magnification factor α, cut-off wavelength λ_c, lower cut-off frequency for temporal filter ω_h, higher cut-off frequency for temporal filter ω_h. 29
3.3 Temporal bandwidth parameters (lower cut-off frequency ω_h and higher cut-off frequency ω_h) with proposed method and [75]. 29
3.4 Magnification parameters (magnification parameter α and cut-off wavelength λ_c) with proposed method and [75]. 30

4.1 Evaluation of Proposed Motion Detection Method. 44
4.2 Evaluation of Proposed Motion Detection Method. 44
4.3 Evaluation of Proposed Motion Detection Method in terms of Detection Accuracy. Here, slice resolution 128×128 and block size 32×32 are used in [26] to get more accuracy. 45
4.4 Evaluation of Proposed Motion Detection Method in terms of Time Taken. Here, time taken for [55] is shown according to slice resolution and block size. Minimum time taken is 7.68 sec with slice resolution 32×32 and block size 8×8, while maximum time taken is 132.21 sec with slice resolution 128×128 and block size 32×32. This time taken is reported only for phase correlation. Thus, it does not include the computation time for intensity correction of transit set. 45

Chapter 1

Introduction

Motion can be defined as change in position of an object of interest with respect to time. It can be described in terms of velocity, acceleration, displacement and time. In this thesis, we attempt to analyze motion using time-frequency representation.

Need for Time-Frequency Analysis

In general, a signal is represented in two forms, i.e., one is in time domain $h(t)$ and other in frequency domain $H(f)$. In both forms, time (t) and frequency (f) variables are treated as mutually exclusive to obtain representation in terms of one variable by integrating over other variable. This means one variable is getting excluded. Hence, each of these representations of the signal are non-localized with respect to the excluded variable. In other terms, time domain representation is obtained by averaging the values of the frequency domain representation at all frequencies and frequency domain representation is obtained by averaging the values of the frequency domain representation at all time instances. Time domain representation hides information about frequency while frequency domain representation hides information about time. Therefore, we need a representation of a signal as two variable function whose domain is the two-dimensional space (t, f). Its constant-t cross section shows the frequencies present at time t while its constant-f cross section shows the time instances present at frequency f. Such a representation is called as 'Time-Frequency Representation' [8]. In this representation, variables t and f are not mutually exclusive but are present together. Some uses of time-frequency representation are listed below:

- Analyze the raw signal in (t, f) domain to identify its characteristics like time variation, frequency variation, number of components, relative amplitude etc.
- Separate the components from each other and even from background noise by filtering in (t, f) domain.
- Analyze specific components separately such as,
  - Track instantaneous amplitude,

– Track instantaneous frequency,

– Track instantaneous bandwidth etc.

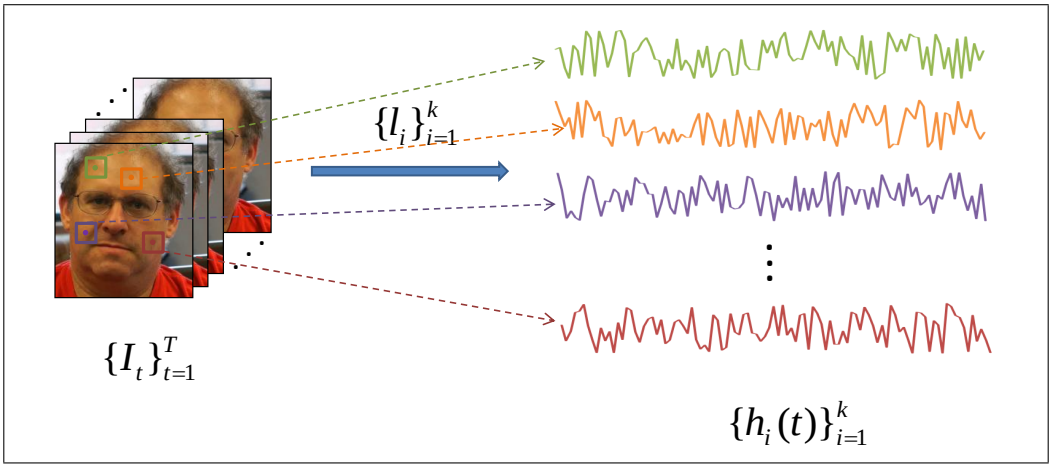


Figure 1.1 Extraction of time series. A video $\{I_t\}_{t=1}^{T}$ with 'k' landmarks $(\{h_i(t)\}_{i=1}^{k})$ is shown. Time series $(\{h_i(t)\}_{i=1}^{k})$ at these 'k' landmark pixels are extracted. Landmarks and corresponding time series are shown in same colour.

Any image data collected over time such as video $I(m, n; t)$, $(m \in [1, M], n \in [1, N], t \in [1, T])$ can be projected as several one dimensional time series. Let us consider a video of 'T' frames (images), $\{I_t\}_{t=1}^{T}$. Each frame is of size $M \times N$. Therefore, $M \times N \times T$ voxels should be considered to process the video. If a set of voxels follow a specific property, we can detect those voxels, i.e., landmarks, $\{l_i\}_{i=1}^{k}$ and process those voxels only. Here, $k \leq M \times N$. Since motion is change in position of an object of interest with respect to time, we extract time series $(\{h_i(t)\}_{i=1}^{k})$ at all these landmarks. Thus, only 'k' time series of length 'T' are processed instead of $M \times N \times T$ voxels. Thus, processing in one dimension is computationally effective compared to that in three dimensions. This motivated us to analyze motion in one dimension due to the fact that motion is a change in position of an object with respect to time. In this thesis, we utilized time-frequency representation for two applications:

- Small Motion Magnification in Videos
- Motion Detection in Perfusion Weighted Imaging (PWI)

Small Motion Magnification in Videos

Human eye and its brain interface can visualize or detect the motion within a certain range of spatial and temporal frequencies. But in most of the cases, it might be possible that frequencies which are below this range also can have useful information. We can simplify this by saying that there can be small motions which are not visible to the naked eye. Even though these small motions are difficult to detect, they may contain useful information. In this thesis, we utilized time-frequency representation to

track instantaneous frequencies and thus estimate desirable parameters automatically to magnify small motions.

There are several applications for motion magnification, for example, visualization, physical diagnosis, pre-measurement planning for precise physical measurements, and surveillance etc. [42]. Therefore, there is a large scope for research to develop a system to amplify or magnify small motions.

Motion Detection in PWI

In perfusion weighted magnetic resonance imaging (PWI), blood perfusion is measured by injecting an exogenous tracer called bolus into the blood flow of a patient and then tracking it in the brain. PWI requires a long data acquisition time to form a time series of volumes, $I(m, n, l; t)$. Hence, motion often occurs between volumes due to a patient's unavoidable movements during a scan, which in turn results into motion corrupted data. There is a necessity of detection and subsequent correction of these motion artifacts on captured data for correct disease diagnosis. In PWI, intensity profile gets disturbed due to occurrence of motion and/or bolus passage through the blood vessels. Even though PWI scans consist of volumes of two dimensional images, they are acquired over the time. Therefore, we can extract one dimensional time series from volumes. Here, we used time-frequency representation to process these extracted one dimensional time series in order to detect the motion.

1.1 Problem Overview

In this work, we have proposed how time frequency analysis can be used for different types of motion analysis. It is shown on two problems, i.e., (i) Estimation of Parameters for Automated Magnification of Small Motions in Videos and (ii) Motion detection in perfusion weighted MRI.

In [75] , an attempt is made to magnify small motions in videos which otherwise will be invisible. In this, videos are projected as one dimensional time series for processing. To magnify small motions, given video is decomposed spatially. Then user has to set parameters for filtering and magnification. In our work, we estimate these parameters automatically using time-frequency analysis. Main contribution is to magnify the small motions in video automatically. It consists of the following two contributions: (1) Estimate bandwidth for temporal filter and (2) Estimate the magnification parameters, automatically.

As explained above, PWI volume series can be converted into one dimensional series. Motion detection using one dimensional time series is obviously faster compared to that of two dimensional scans. These facts motivated us to analyze the PWI data in terms of one dimensional time sequences because we believe that frequency of time series will vary when there are motion artifacts. An efficient time-frequency analysis based motion detection method is proposed and it is computationally inexpensive and fast.

1.2 Contribution

In this research work, we focus on analyzing motion using time-frequency representation. Main contributions are the following:

- Proposed a semi-automated method to magnify small motions in videos. This method amplifies invisible or hidden motions in videos. To achieve motion magnification, we process the spatial and temporal information obtained from the video itself. Main contribution lies in temporal analysis. Here, we used a time-frequency representation for temporal analysis and from which we estimate the required parameters. Advantage of this work is that it is application independent. Proposed technique estimates required parameters to get desirable results.

- Proposed a novel automated method to detect motion corrupted volumes in PWI using time-frequency analysis. This is more efficient than an existing method which uses phase correlation [26]. In [26], given PWI data is divided into 3 sets according to bolus status: (i) pre-wash-in, (ii) transit and, (iii) post-wash-out sets. Intensity correction is applied to transit set volumes. Then, phase correlation is performed to detect motion. In our work, we detect motion without any explicit intensity correction. We show that proposed time-frequency analysis based motion detection method outperforms compared to that of [26] in terms of accuracy and computational efficiency.

1.3 Thesis Organization

This thesis is organized in five chapters. Chapter 1 provides an overview of the general background and the problem setting. Motivation behind the present work and the major contributions are also briefly described. In Chapter 2, we present the state of the art in small motion magnification in videos and motion detection in PWI.

An efficient technique for magnifying small motions in videos is presented in Chapter 3. We have presented motion magnification based on time-frequency analysis. We demonstrate performance on 9 different videos. Motion magnification performance is equivalent to existing manual methods.

Motion detection and subsequent detection in perfusion weighted MRI is presented in Chapter 4. We show that proposed method is computationally inexpensive and fast. This method is evaluated on a PWI sequence with simulated motion of different degrees. We show that our approach detects motion in a few seconds.

Finally, Chapter 5 contains the conclusions of this thesis and future work.

Chapter 2

Background and Previous Work

This chapter provides sufficient background related to this thesis. It contains discussions on (i) different time-frequency representations of a signal and show how Stockwell transform outperforms other time-frequency distributions, (ii) video magnification, i.e., how small motions have been magnified in previous works and, (iii) fundamentals of perfusion weighted magnetic resonance imaging and different techniques used to detect motion artifacts occurred due to patient's motion in MRI scanner.

2.1 Stockwell Transform and its Applications

2.1.1 Evolution of Stockwell Transform

Stockwell Transform (ST) is a time-frequency representation which has been shown as significant improvement over existing techniques for localizing spectral information. In this section, an overview of ST is presented in terms of its efficiency in time-frequency localization. First, we start with the most popular temporal analysis, i.e., Fourier transform and discuss various time-frequency representations. Since discussing reverse transformation will be beyond the scope of this thesis, we explain every transform in terms of forward transformation only.

2.1.1.1 Fourier Transform

Fourier transform (FT) [23] [9] is one of the most popular frequency analysis technique. It transforms time domain signal to frequency domain signals by using a complex sinusoid $e^{i2\pi ft}$ as a basis.

For a given signal $h(t)$, continuous Fourier transform is given as,

$$\begin{aligned} H(f) &= \langle h(t), e^{i2\pi ft} \rangle \\ &= \int_{-\infty}^{\infty} h(t) e^{-i2\pi ft} dt \end{aligned} \tag{2.1}$$

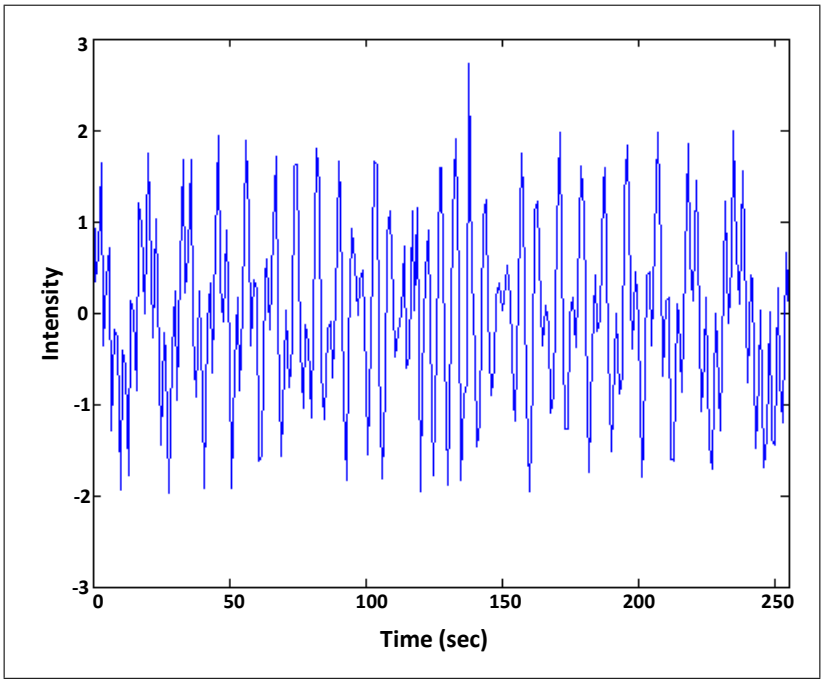


Figure 2.1 Chirp Signal.

where f is frequency and $\langle ., . \rangle$ denotes inner product. This inner product notation is used later also in this chapter. Even though FT is a popular frequency domain representation, it has the following disadvantages [27]:

- FT fails to estimate fractional frequencies. For a signal with fractional frequencies, FT spreads the spectrum to other frequencies which are not actually present in the given signal.
- FT can not help to estimate which frequency exists at which time, i.e., it fails to the frequency content in the temporal dimension.
- FT of a signal consisting of a particular frequency for short duration, i.e., a non stationary signal, can not give information about that particular frequency.

To overcome these disadvantages, a combined time-frequency representation was proposed.

2.1.1.2 Short Time Fourier Transform

Short time Fourier transform (STFT) [2] is the first joint time-frequency representation. Basic idea of STFT is to break up the signal into small time segments and apply Fourier transform (FT) to analyze each segment so that summation of such spectra gives information about how frequency is varying with respect to time.

For a given signal $h(t)$, short time Fourier Transform is given as,

$$\begin{aligned} H_{ST}(\tau, f) &= \langle h(t), w(t-\tau)e^{i2\pi ft} \rangle \\ &= \int_{-\infty}^{\infty} h(t)w(t-\tau)e^{-i2\pi ft}dt \end{aligned} \tag{2.2}$$

where τ is time shift parameter, f is frequency and $w(t)$ is window. Window width is chosen such that the windowed signal segment can be assumed to be stationary so that it makes the signal more or less unaltered around time τ but surpasses the signal for times distant from the time of interest.

The time-frequency distribution obtained from STFT is called a spectrogram. The spectrogram of signal $h(t)$ is given as,

$$SP(\tau, f) = |STFT(\tau, f)|^2 \tag{2.3}$$

where $|.|$ denotes the absolute value.

Time localization refers how well signal variations can be represented in time domain whereas frequency localization refers how well variations in frequencies of signal can be represented in frequency domain. We achieve good time localization with a narrow window in time domain, i.e., wide window in frequency domain and similarly, we achieve good frequency localization with a wide window in time domain, i.e., narrow window in frequency domain. Since the window can not be made arbitrarily narrow, there is a trade-off between time and frequency localization in the spectrogram for a given window. For a given window $w(t)$, trade-off between time and frequency resolution can be explained by the uncertainty principle [14] which is as given below,

$$\sigma_t^2 \sigma_f^2 \geq \frac{1}{4} \tag{2.4}$$

where, σ_t is standard deviation of $w(t)$ in time domain, σ_f is standard deviation of $w(t)$ in frequency domain (i.e., $W(f)$ Fourier transform of $w(t)$). σ_t is defined as,

$$\sigma_t = \sqrt{E[(t-\mu_t)^2]} = \sqrt{\int_{-\infty}^{\infty} (t-\mu_t)^2 w(t)dt} \tag{2.5}$$

where, $E[.]$ represents expectation and $\mu_t = \int_{-\infty}^{\infty} tw(t)dt$ represents mean value.

σ_f is defined as,

$$\sigma_f = \sqrt{E[(f-\mu_f)^2]} = \sqrt{\int_{-\infty}^{\infty} (f-\mu_f)^2 W(f)df} \tag{2.6}$$

where, $E[.]$ represents expectation and $\mu_f = \int_{-\infty}^{\infty} fW(f)df$ represents mean value.

2.1.1.3 Gabor Transform

Gabor transform is a special case of STFT with Gaussian window. Gaussian window satisfies lower bound of uncertainty principle, thus giving optimal resolution in both time and frequency domains.

For a given signal $h(t)$, Gabor transform [25] is given as,

$$\begin{aligned} H_G(\tau, f) &= \langle h(t), e^{-\pi(t-\tau)^2} e^{i2\pi ft} \rangle \\ &= \int_{-\infty}^{\infty} h(t) e^{-\pi(t-\tau)^2} e^{-i2\pi ft} dt \end{aligned} \tag{2.7}$$

where τ is time shift parameter and f is frequency.

2.1.1.4 Wigner Ville Distribution

The Wigner-Ville distribution (WVD) is one in which signal itself is used to define the window. For a given signal $h(t)$, WVD [72] [67] [7] is given as,

$$\begin{aligned} H_{WV}(\tau, f) &= \langle h(\tau - \frac{1}{2}t), h(\tau + \frac{1}{2}t) e^{i2\pi ft} \rangle \\ &= \int_{-\infty}^{\infty} h(\tau - \frac{1}{2}t) h(\tau + \frac{1}{2}t) e^{-i2\pi ft} dt \end{aligned} \tag{2.8}$$

where τ is time shift parameter and f is frequency.

Advantage of Wigner-Ville distribution (WVD) over the spectrogram is that we do not have to choose a window. WVD is better than any spectrogram with a particular window. But it has disadvantage that it produces cross terms in time-frequency distribution which are unavoidable in case of a signal consisting of summation of several signals.

2.1.1.5 Continuous Wavelet Transform

A wavelet is a continuous time signal which satisfies the following properties.

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \tag{2.9}$$

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \tag{2.10}$$

$$\int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} dw = C < \infty \tag{2.11}$$

where $\psi(t)$ is called as mother wavelet and $\Psi(\omega)$ is Fourier transform of $\psi(t)$. Equation 2.11 is called admissibilty condition which should be satisfied by mother wavelet in order to reconstruct the signal. Continuous wavelet transform (CWT) [29] of a square integrable signal $h(t)$ is defined as,

$$\begin{aligned} H_{CW}(a,b) &= \langle h(t), \psi_{a,b}(t) \rangle \\ &= \int_{-\infty}^{\infty} h(t) \psi^*_{a,b}(t) dt \end{aligned} \tag{2.12}$$

where a is dilation parameter, b is translation parameter, $*$ denotes complex conjugation and $\psi^*_{a,b}(t)$ is a family of wavelets generated by dilating and translating mother wavelet as below,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi(\frac{t-b}{a}) \tag{2.13}$$

As shown above, CWT is obtained by the inner product of the signal and dilations and translations of the mother wavelet. CWT is represented as a time scale plot, where scale is the inverse of frequency. At a low scale (high frequency), CWT offers high time resolution whereas at higher scales (lower frequencies) it shows high frequency resolution.

2.1.1.6 Stockwell Transform

Stockwell transform [63] was introduced as an extension of CWT. It is based on a moving and scalable Gaussian window. It has some desirable characteristics over CWT as it is unique in providing frequency dependent resolution. It is due to the fact that the modulating sinusoids are fixed w.r.t. time axis, whereas the localizing scalable Gaussian window dilates and translates. For a given signal $h(t)$, Stockwell transform (ST) is given as,

$$\begin{aligned} S(\tau, f) &= \langle h(t), w(\tau - t, f) e^{i2\pi f t} \rangle \\ &= \int_{-\infty}^{\infty} h(t) w(\tau - t, f) e^{-i2\pi f t} dt \end{aligned} \tag{2.14}$$

where $w(t, f)$ is defined as,

$$w(t, f) = \frac{|f|}{\sqrt{2\pi}} e^{-t^2 f^2 / 2}, \tag{2.15}$$

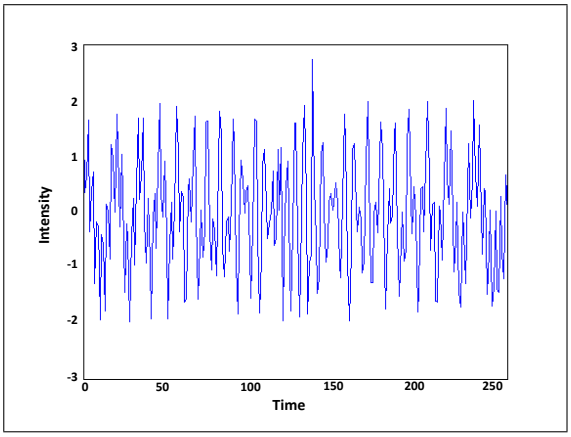
$w(t, f)$ denotes the window, f denotes the frequency, τ denotes time shift parameter, and $|.|$ denotes absolute value.

2.1.1.7 Discussion

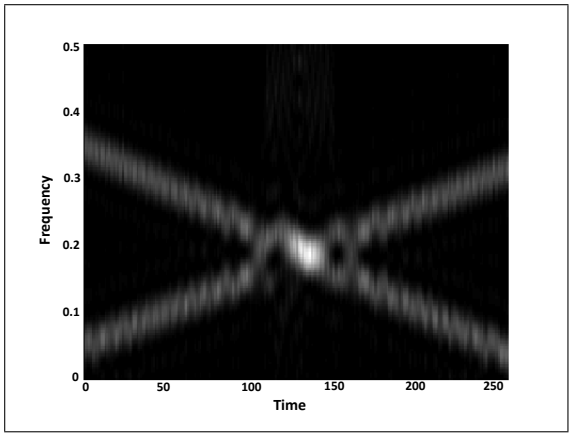
To distinguish between different time-frequency representations mentioned above, let us consider an example of synthetic signal as shown in Figure 2.1. This synthetic signal consists of two cross chirps and two frequency bursts.

$$h(t) = \begin{cases} \cos(2\pi(10+t/7)*t/256) + \cos(2\pi(256/2.8-t/6)*t/256) & \text{if } 0 \le t < 114 \text{ and } 142 < t \le 255 \\ \cos(2\pi(10+t/7)*t/256) + \cos(2\pi(256/2.8-t/6)*t/256) + \cos(2\pi t*0.42) & \text{if } 114 \le t \le 122 \\ \cos(2\pi(10+t/7)*t/256) + \cos(2\pi(256/2.8-t/6)*t/256) + \cos(2\pi t*0.42) + \cos(2\pi t*0.42) & \text{if } 134 \le t \le 142 \end{cases} \tag{2.16}$$

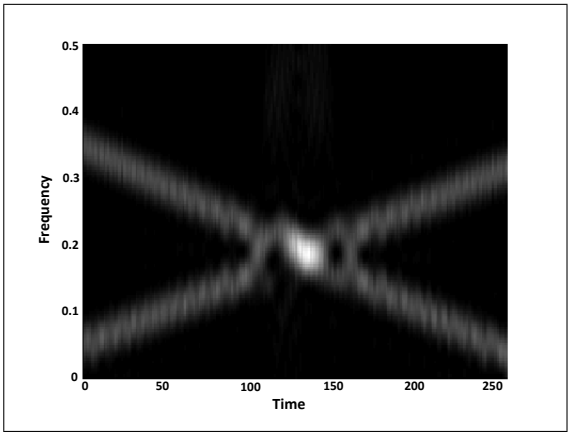
Different time frequency representations for the above synthesized signal are shown in Figure 2.2. All these are generated in MATLAB. Figure 2.2(b) shows spectrogram of signal (Figure 2.1) with boxcar window of length 20 units. Here, both chirps are detected but the two high frequency bursts are not detected. The effect is similar in case of Gabor transform with window of length 20 units (See Figure 2.2(c)). Both STFT and Gabor transform do not have sufficient time resolution to resolve the two signals. Wigner-Ville distribution detected both chirps with very good resolution but failed to detect high frequency bursts and one can see cross terms effect (See Figure 2.2(d)). Even in case of CWT, both chirps are detected but high frequency bursts could not be detected (See Figure 2.2(e)). Here, Morlet wavelet is used as mother wavelet as shown in Figure. Stockwell transform is shown in Figure 2.2(f). In this case, both chirps are detected along with two high-frequency bursts. Hence, we can say Stockwell transform is better for time-frequency localization compared to other time-frequency representations.



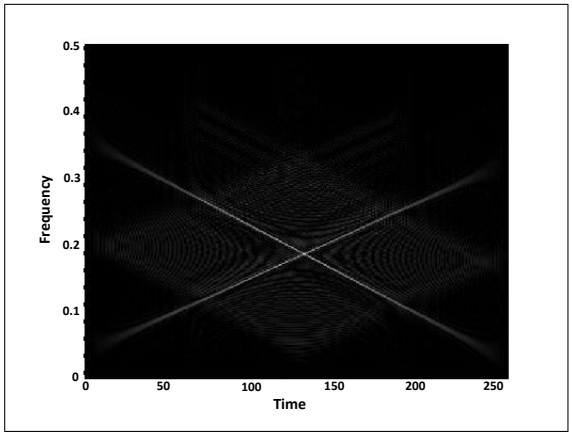
(a) Chirp Signal.



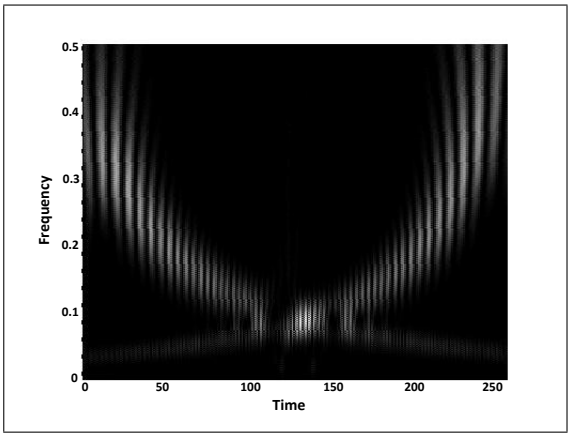
(b) STFT with boxcar window.



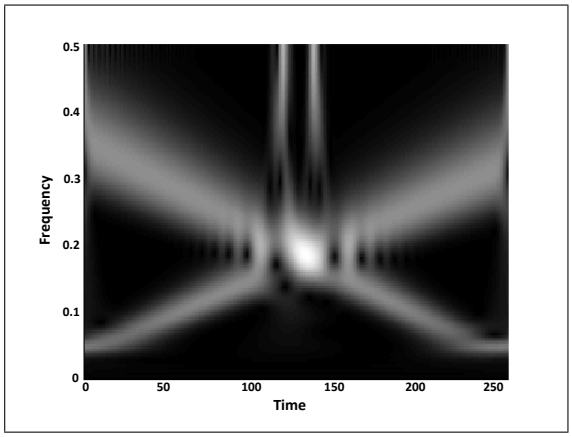
(c) STFT with Gaussian window (Gabor Transform).



(d) WVD.



(e) CWT with Morlet wavelet as mother wavelet.



(f) Stockwell Transform.

Figure 2.2 Different Time Frequency Representations for a Chirp Signal.

2.1.2 Applications of ST

Stockwell transform has been used in various fields [69]. A few of them will be discussed in this section.

Stockwell transform was initially developed for analyzing geophysics data [62]. In Geophysics, S transform has extensively been used. For example, In [49], S transform is used for analyzing polarization and filtering of three component signals.

[44] presents functional MRI (fMRI) cluster analysis using Stockwell ransform. In this, Stockwell transform is used studying time-frequency characteristics of wavelet based functional clutters. It shows the application of the Stockwell transform to the characterization of voxel-type signals obtained by the application of the wavelet packet algorithm. These results showed that the dynamic behavior of the average signal for each cluster can be defined by a time-frequency map. Moreover, a measure of the interdependence of the instantaneous phase between clusters is obtained from the cross ST spectrum.

In the recent past, ST has been used for the analysis of MRI data. In [28] [46], ST is used to remove artifacts in functional MRI (fMRI) time courses due to which brain activity detection is improved. In this work, one dimensional Fourier transforms (FTs) are performed on raw image data to obtain phase profiles. The time series of phase magnitude for each and every point in the phase profile is then subjected to the ST to obtain a time-frequency spectrum. The temporal location of an artifact is identified based on the magnitude of a frequency component relative to the median magnitude of that frequency occurrence over all time points. After each artifact, frequency is removed by replacing its magnitude with the median magnitude, an inverse ST is applied to regain the MR signal. Brain activity detection within fMRI datasets is improved by significantly reducing image artifacts that overlap anatomical regions of interest. The major advantage of ST-filtering is that artifact frequencies may be removed within a narrow time-window, while preserving the frequency information at all other time points.

In [15], an approach for power quality analysis using Stockwell transform is proposed. The local spectral information of the wavelet transform can, with slight modification, be used to perform local cross spectral analysis with very good time resolution. The phase correction absolutely references the phase of the wavelet transform to the zero time point, thus assuring that the amplitude peaks are regions of stationary phase. The excellent timefrequency resolution characteristic of the S transform makes it an attractive candidate for analysis of power system disturbance signals. Several power quality problems are analyzed using both the S transform and discrete wavelet transform, showing clearly the advantage of the S transform in detecting, localizing, and classifying the power quality problems. In [60] also, Stockwell transform is used for power quality analysis.

In [4], a technique to extract palm-print features for recognition is proposed. Here, features are extracted based on instantaneous-phase difference obtained using Stockwell transform of overlapping circular-strips. [37] presents a technique for image compression using S transform. It shows that S transform offers better image compression compared to wavelet transforms. [19] presents a method for characterizing image texture based on two dimensional Stockwell transform. It describes an approach

to obtain local spatial frequency information for an image and show that this information can be used to characterize the horizontal and vertical frequency patterns in synthetic images. This method provides the computational efficiency and multi-scale information of wavelet transforms, while providing texture features in terms of Fourier frequencies. It outperforms leading wavelet based texture analysis methods.

In [76], polar version of ST is used to analyze the texture patterns in MRI for the diagnosis of multiple sclerosis. [77] discusses the effectiveness of ST for medical imaging and shows how to enhance fMRI time courses by removing frequency artifacts which are introduced due to patient's quick breathing.

Stockwell transform is used even in seismogram analysis [18] [17] [47], analysis of engine induction noise in acceleration [30] and for analysis of EEG signals [58] [51] [1] [50].

2.2 Small Motion Magnification in Videos

In this section, we describe the hidden motions in videos and previous methods which attempt to magnify those motions.

Human eye and its brain interface can visualize or detect the motion within a certain range of spatial and temporal frequencies. But in most of the cases, it might be possible that frequencies which are below this range also can have useful information. We can simplify this by saying that there can be small motions which are not visible to the naked eye. Even though these small motions are difficult to detect, they may contain useful information.

There are several applications for motion magnification, for example, visualization, physical diagnosis, pre-measurement planning for precise physical measurements, and surveillance etc. [42]. Therefore, there is a large scope for research to develop a system to amplify or magnify small motions.

There are previous works in this direction which attempted to reveal invisible motions in videos. [42] analyzes and amplifies subtle motions and visualize deformations that would otherwise be invisible. [68] propose using the cartoon animation filter to create perceptually appealing motion exaggeration. As such, they rely on accurate motion estimation, which is computationally expensive and difficult to make artifact-free, especially at regions of occlusion boundaries and complicated motions.

Temporal processing has been used previously to extract invisible signals [52] and to smooth motions [24]. In [52], a heart rate is extracted from a video of a face based on the temporal variation of the skin color, which is normally invisible to the human eye. They focus on extracting a single number, whereas we use localized spatial pooling and bandpass filtering to extract and reveal visually the signal corresponding to the pulse. This primal domain analysis allows to amplify and visualize the pulse signal at each location on the face. This has important potential monitoring and diagnostic applications to medicine, where, for example, the asymmetry in facial blood flow can be a symptom of arterial problems. In [24] per-pixel temporal filters are used to dampen temporal aliasing of motion in videos.

In [3, 39], human motions are generated by reusing the captured motion to create new motions. In [75], small motion is magnified without tracking motion. However, this approach needs users to provide a set of parameters as input for every video in order to magnify motion. For unknown video,

this consumes time to find desirable parameters. While all these techniques require user interaction in some or the other way, we are not aware of any previous work addressing automated approach to magnify small motions.

2.3 Motion Detection in Perfusion Weighted MRI for Brain

In this section, we give brief description on perfusion weighted MRI and its types. Then we explain the necessity to detect motion artifacts prior to motion correction.

2.3.1 Perfusion Weighted MRI

Magnetic resonance imaging (MRI) has been emerging as an efficient tool in clinical practice for the analysis of brain functions through several metabolic parameters. There are two types of MRI, namely, diffusion weighted imaging (DWI) and perfusion weighted imaging (PWI). PWI has been used extensively for the evaluation of tissue after acute stroke, non-invasive histologic assessment of tumors and evaluation of neurodegenerative conditions such as Alzheimers disease [48]. Diffusion weighted images are obtained by incorporating strong agnetic field gradient pulses into an imaging pulse sequence. In DWI, structures with fast diffusion are dark due to the fact that these structures are subject to greater signal attenuation, whereas structures with slow diffusion are bright. In PWI, an exogenous tracer is introduced into the blood circulation and its concentration in a tissue is monitored in a tissue over time. Blood flow to the corresponding tissue can be determined by obtaining the rate of delivery of the tracer. In clinical practice, Gadolinium-DTPA is used as the exogenous tracer. This tracer induces a difference in magnetic susceptibility between the blood compartment and the brain tissue, when it reaches the brain capillaries. Here, difference in magnetic susceptibility occurs because the tracer can not penetrate the blood-brain barrier. Diffusion of water through the internal gradients produce a low signal attenuation [16].

There are two types of PWI: (i) dynamic susceptibility contrast (DSC) imaging, and (ii) dynamic contrast enhanced (DCE) $T1$ weighted imaging. DSC is most widely used for the brain, while DCE is most widely used in the rest of the body though its experimental and research use is increasing in brain. Here, we present these types of PWI from a review paper [53].

2.3.1.1 Dynamic Susceptibility Contrast Imaging (DSC)

This imaging is called as $T2^*$ imaging, which refers to gradient echo sequences. At high concentrations, exogenous tracer induces substantial $T2^*$ shortening, resulting first in loss and then recovering of the signal as the tracer is distributed or diluted. DSC MR imaging can be performed by using either a gradient-echo or a spin-echo pulse sequence [12]. Gradient-echo DSC sequences tend to be more sensitive to larger vessels, such as veins, in the imaged region. Spin-echo DSC techniques tend to

show greater sensitivity to smaller vessels (and therefore are more representative of capillary density) or abnormal (for example, tumor specific) vessels.

It should be noted that $T2^*$ effects extend beyond the borders of the blood vessels into the surrounding tissues; this characteristic is important when there is little leakage of contrast agent into the surrounding tissue, such as with an intact blood-brain barrier [5]. These methods can be safely used when the rate of vascular leakage is low. However, when the rate of leakage is high, relative cerebral blood volume (rCBV) mapping results can be underestimated for two reasons: (a) The $T2^*$ susceptibility effect is reduced as the gradient of contrast agent is reduced, and (b) there is also signal enhancement due to T1 shortening effects of contrast material in the extravascular extracellular space (EES) [73]. For this purpose, T1-insensitive sequences, small flip angle, or dual-echo approaches [20] are used, as is presaturation of the EES with administration of a preinjection dose of contrast material [33]. Postprocessing mathematic corrections are also frequently used [12].

Although at this point a relatively large number of studies in which rCBV was measured have been performed, no single standard technique for rCBV measurement has been established. A number of rCBV measurement methods exist, including placement of a single region of interest and calculation of the mean of repeated rCBV measurements, but few studies have been performed on the reproducibility of rCBV measurements [71]. Although reproducibility with some of these techniques appears to be acceptable for present clinical purposes, it remains to be seen how well suited these measurements are for the assessment of moderate changes in rCBV after such interventions as antiangiogenesis therapy, i.e., how the biologic variation in these measurements compares with changes due to therapy. Permeability can also be assessed with the use of DSC images, which would allow one to obtain both permeability and rCBV measurements from the same infusion of contrast material [54]. However, the technique is not considered to be valid under conditions in which a very high degree of contrast material leakage is present, which is a limitation in many cases [70].

2.3.1.2 Dynamic Contrast Enhanced Imaging (DCE)

Dynamic contrast-enhanced MR imaging approaches are based on $T1$ shortening produced by an infusion of paramagnetic contrast material [65]. $T1$-based changes are primarily a result of contrast material diffusion into the EES. Dynamic imaging is typically performed during an interval of approximately 510 minutes rather than during the first pass of the bolus. However, a $T1$-based first pass approach has also been proposed and tested [41][40][31]. In addition, the actual T1 values of the tissues at baseline (before contrast material infusion) are required for most analysis algorithms in order to perform the pharmacokinetic analysis [22]. This calculation can be performed by using a series of T1-weighted images obtained at different flip angles. To maintain the required temporal resolution, three-dimensional imaging schemes are usually used (which have the drawback of a longer acquisition time than two-dimensional schemes), and arterial input functions are generally obtained in the center of the acquisition volumes to reduce end-section effects. Because dynamic contrast-enhanced approaches rely on T1 shortening of the EES to develop signal intensity and image contrast, these methods are less opti-

mal than $T2^*$-weighted methods, in which the rate of leakage into the EES is low. As might be expected, data acquisition parameters can influence data analysis and need to be optimized [22]. Quantification of absolute cerebral blood volume (rather than rCBV) can be obtained with perfusion CT (Computed Tomography) through use of venous and arterial input functions [34]; however, the exact method by which quantification of absolute cerebral blood volume can be obtained using perfusion MR imaging is still a matter of active investigation.

PWI has been used for the evaluation of functioning of brain through assessment of several metabolic parameters. In PWI, cerebral perfusion is used as a metabolic parameter, which explains the blood passage through the vascular system of the brain. An exogenous tracer called bolus is injected into the blood flow of a patient and then cerebral perfusion is measured by the analysis of hemodynamic time-to-signal intensity curve generated when bolus passes through the brain.

2.3.2 Motion Detection

In PWI, a time series of volumes are formed in a long acquisition time. Patient often has difficulty in staying still during this period. Therefore, it is more likely that patient may move unavoidably during scanning which in turn results into motion artifacts in scans. There is a need of detection and subsequent correction of these motion artifacts. There are works in medical imaging, for example [10], [32], [61] addressed this problem in terms of registration of whole time series to a reference volume. Motion correction of the 4D data can be seen as alignment of motion corrupted volumes to "stationary" volumes. The standard approach to motion correction of 4D perfusion MRI data perform is via an alignment of motion corrupted volumes. Thus, there is no explicit detection and subsequent correction of motion. For DSC-MRI, motion correction techniques include registering the time series to either a single volume or the mean volume of the entire time-series data [32]. [10] includes a model of dynamic contrast in an iterative registration process for tracking tumour motion. 3D rigid registration via cylindrical phase correlation has been proposed in [6] which is capable of handling highly misaligned volumes and is noise resilient. In DCE-MRI, motion correction has been done via registration techniques including a rigid body model [36], nonrigid B-spline [66], maximization of a special Gibbs energy function using a gradient descent algorithm [21], mutual information with a spatial transformation model [56], an intensity correction model [38], group-wise registration [35] or progressive principal component registration [45]. In general, motion correction has been found to be a time-limiting step (90% of processing time) in a PWI analysis pipeline [64]. Since 3D registration is computationally intensive [59] it would be more efficient if only a subset of volumes need to be aligned instead of every volume in the entire time-series data. In general, if a time-series has N phases (or volumes), all the phases are not corrupted by motion. Hence, it is worthwhile to first detect the subset $\leq N$ volumes that is affected by motion and subsequently correct this subset.

All the above methods do not detect motion. Hence, non-corrupted volumes are also registered which makes the process computationally expensive and it is obvious that these volumes do not need any correction [55]. Therefore, it is preferable to have a prior knowledge about motion corrupted volumes.

To the best of our knowledge, only [26] and [55] have detected motion prior to correction. Motion correction is typically the rate limiting step in processing as each volume has to be registered to a reference volume. This is compounded by the dynamically varying contrast in the volume series due to passage of an injected contrast agent. This work presents a two stage motion correction method, consisting of motion detection and a 2-pass registration method for aligning the motion-corrupted volumes. A 2D block-wise phase correlation in central slices is used for the first stage. Alignment employs a strategy which is sensitive to the status of the bolus in the volume and is based on gamma-variate function fitting for intensity correction to handle dynamic contrast in DSC MRI.

2.4 Summary and Comments

In this thesis, we analyze motion using time frequency analysis. For this purpose, we have used Stockwell transform. To the best of our knowledge, any time-frequency analysis or to be specific, Stockwell transform is not used before to magnify small motions in videos and detect motion artifacts in perfusion weighted MRI. Even though data handled in this work is three dimensional, it varies changed with respect to time in case of both videos and perfusion MRI data. Therefore, we project three dimensional data into one dimensional time signals and process the data temporally and then achieve desirable results. Due to reduction in dimensions, proposed approaches are obviously computationally effective and fast.

Chapter 3

Estimation of Parameters for Automated Magnification of Small Motions in Videos using ST

3.1 Introduction

Human eye and its brain interface can visualize or detect the motion within a certain range of spatial and temporal frequencies. But in most of the cases, it might be possible that frequencies which are below this range also can have useful information. We can simplify this by saying that there can be small motions which are not visible to the naked eye. Even though these small motions are difficult to detect, they may contain useful information.

There are several applications for motion magnification, for example, visualization, physical diagnosis, pre-measurement planning for precise physical measurements, and surveillance etc. [42]. Therefore, there is a large scope for research to develop a system to amplify or magnify small motions.

There are previous works in this direction. In [3, 39], human motions are generated by reusing the captured motion to create new motions. In this paper, we propose to process the data obtained from the video and reconstruct the video from the modified data such that new video shows magnified motion. In [75], small motion is magnified without tracking motion. However, this approach needs users to provide a set of parameters as input for every video in order to magnify motion. For unknown video, this consumes time to find desirable parameters. While all these techniques require user interaction in some or the other way, we are not aware of any previous work addressing automated approach to magnify small motions. This motivated us to develop a mechanism to estimate these parameters automatically from the given data. We use a time frequency representation called Stockwell transform for this purpose. We introduce computationally inexpensive techniques to estimate parameters. We illustrate the utility of the proposed method on examples in which small motions were made visible.

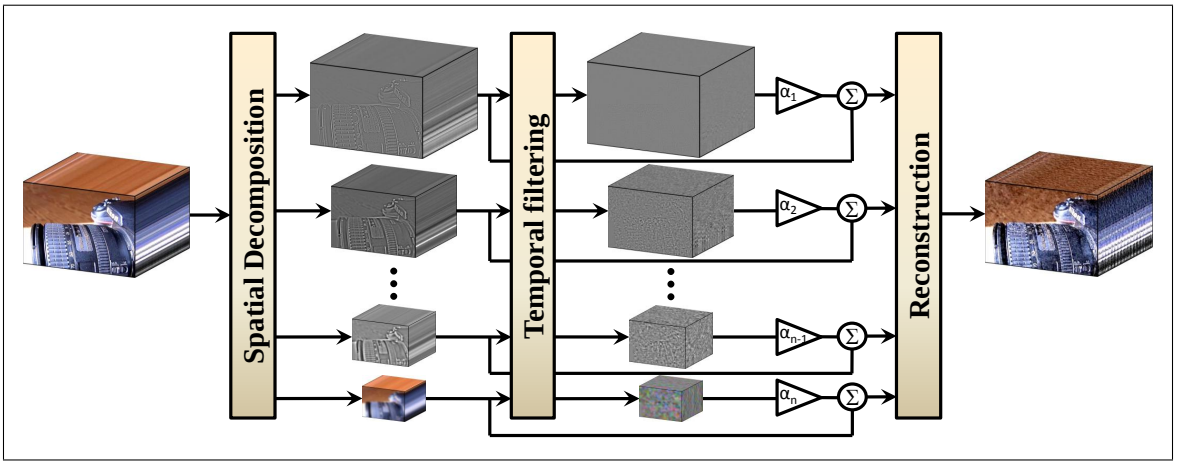


Figure 3.1 Overview of the Eulerian video magnification method. This Figure is adapted from [75]. The input video sequence is decomposed into different spatial frequency bands, and the same temporal filter is applied to all bands. The filtered spatial bands are then amplified by a given factor α, added back to the original signal, and collapsed to generate the output video.

3.2 Background

In this section, we describe Eulerian video magnification [75] [57] as it forms the basis of our work. This method is illustrated in Figure 3.1. Spatial and temporal processing are combined to emphasize subtle temporal changes in a video. Firstly, the video sequence is decomposed into different spatial frequency bands. These bands might be magnified differently because (a) they might exhibit different signal-to-noise ratios, or (b) they might contain spatial frequencies for which the linear approximation used in motion magnification does not hold (Section 3.2.1). In the latter case, the amplification is reduced for these bands to suppress artifacts. Spatial processing is used to increase temporal signal-to-noise ratio by pooling multiple pixels, by which the frames of the video are spatially low-pass filter and downsampled for computational efficiency. A full Laplacian pyramid [11] is constructed for this purpose. Then, temporal processing is performed on each spatial band. The time series corresponding to the value of a pixel is considered in a frequency band and a bandpass filter is applied to extract the frequency bands of interest. The temporal processing is uniform for all spatial levels, and for all pixels within each level. Then the extracted bandpassed signal is multiplied by a magnification factor α. This factor can be specified by the user. Next, the magnified signal is added to the original the spatial pyramid is collapsed to obtain the final output.

3.2.1 Eulerian Video Magnification

Eulerian video magnification depends on the first-order Taylor series expansions. Here, it is shown how temporal processing produces motion magnification.

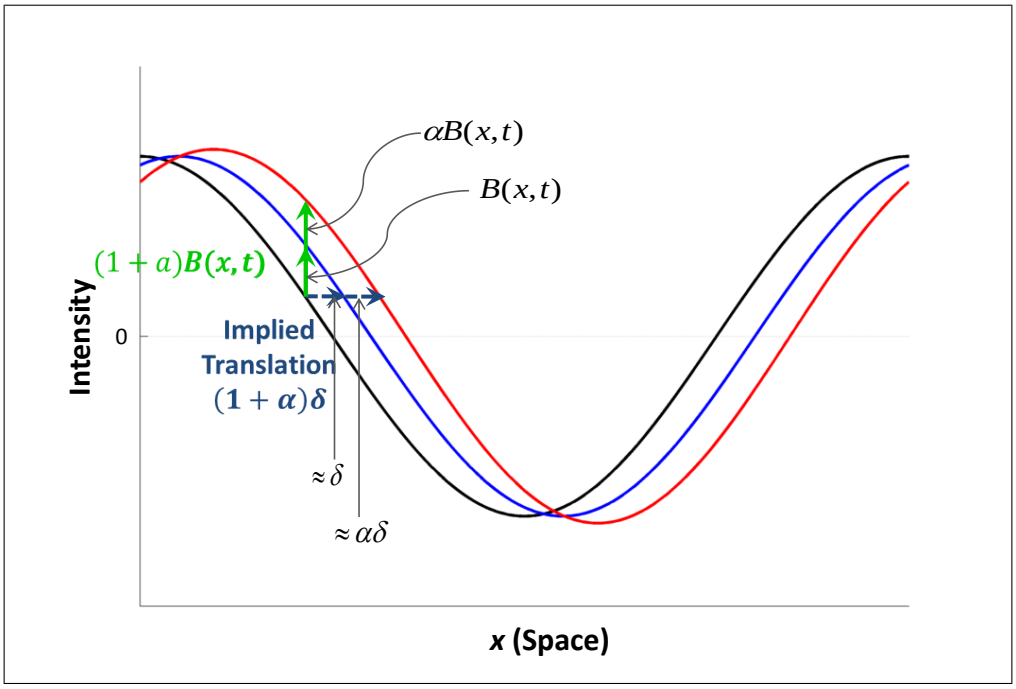


Figure 3.2 Approximation of spatial translation using temporal filtering. This Figure is adapted from [75].

3.2.1.1 First Order Motion

A one dimensional signal undergoing translational motion is considered to describe the relationship between temporal processing and motion magnification. This analysis generalizes directly to locally-translational motion in two dimensional signals. Let $I(x,t)$ denote the image intensity at position x and time t. Since the image undergoes translational motion, the observed intensities can be expressed with respect to a displacement function $\delta(t)$, such that $I(x,t) = f(x+\delta(t))$ and $I(x,0) = f(x)$. The goal of motion magnification is to synthesize the signal

$$\hat{I}(x,t) = f(x+(1+\alpha)\delta(t)) \tag{3.1}$$

for amplification factor α.

Here, the image is approximated by a first-order Taylor series expansion. Then, the image at time t, $f(x+\delta(t))$ in a first-order Taylor expansion about x is expressed as

$$I(x,t) \approx f(x) + \delta(t)\frac{\partial f(x)}{\partial x} \tag{3.2}$$

After applying a broadband temporal bandpass filter to $I(x,t)$, the output will be $B(x,t)$. If the motion signal (displacement function, $\delta(t)$) lies within the passband of this temporal bandpass filter, then $B(x,t)$ can be expressed as

$$B(x,t) = \delta(t)\frac{\partial f(x)}{\partial x} \tag{3.3}$$

This bandpass signal is amplified by a factor α and added to original signal, $I(x,t)$. The reconstructed signal can be represented as

$$\tilde{I}(x,t) = I(x,t) + \alpha B(x,t) \tag{3.4}$$

From Equations 3.2, 3.3 and 3.4, reconstructed signal $\hat{I}(x,t)$ can be approximated as

$$\begin{aligned}
\tilde{I}(x,t) &= I(x,t) + \alpha B(x,t) \\
&\approx \left[f(x) + \delta(t)\frac{\partial f(x)}{\partial x}\right] + \alpha B(x,t) \\
&\approx \left[f(x) + \delta(t)\frac{\partial f(x)}{\partial x}\right] + \alpha\left[\delta(t)\frac{\partial f(x)}{\partial x}\right] \\
&\approx f(x) + (1+\alpha)\delta(t)\frac{\partial f(x)}{\partial x}
\end{aligned} \tag{3.5}$$

Here, it is assumed that first-order Taylor expansion holds for the amplified larger perturbation, $(1+\alpha)\delta(t)$. Thus, reconstructed signal $\hat{I}(x,t)$ (from Equation 3.5) can be approximated as

$$\tilde{I}(x,t) \approx f(x) + (1+\alpha)\delta(t) \tag{3.6}$$

From Equation 3.6, it can be observed that spatial displacement ($\delta(t)$) of the image ($f(x)$) is amplified by a magnitude of '$1+\alpha$' at time 't'. This is shown in Figure 3.2 with a single sinusoid and a relatively small displacement, δ. Here, first-order Taylor series expansion is used to approximate the translated signal at time $t+1$. Therefore, signal is assumed to be translated by $(1+\alpha)\delta$ while adding the amplified signal (by a factor, α) to the original signal $I(x,t)$.

3.2.1.2 Bounds

If an image contains sudden changes, i.e., high spatial frequencies, the first-order Taylor series approximation will be inaccurate for large values of perturbation ($(1+\alpha)\delta(t)$). This is due to the fact that perturbation, $(1+\alpha)\delta(t)$, increases with larger magnification (α) and motion ($\delta(t)$). To overcome this, bounds for magnification factor, α, are derived in terms of spatial frequency, ω. For this purpose, it is assumed that the reconstructed signal, $\tilde{I}(x,t)$, is approximately equal to the true magnified signal, $\hat{I}(x,t)$, as below

$$\tilde{I}(x,t) \approx \hat{I}(x,t) \tag{3.7}$$

Substituting Equations 3.5 and 3.1 in the above Equation 3.7,

$$f(x) + (1+\alpha)\delta(t)\frac{\partial f(x)}{\partial x} \approx f(x + (1+\alpha)\delta(t)) \tag{3.8}$$

Here, it is assumed that $f(x) = \cos\omega(x)$ for spatial frequency ω and $1 + \alpha$ is denoted as β. Then, Equation 3.8 becomes

$$\cos(\omega x) + \beta\delta(t)\frac{\partial \cos(\omega x)}{\partial x} \approx \cos[\omega(x + \beta\delta(t))] \tag{3.9}$$

$$\Rightarrow \quad \cos(\omega x) - \beta\omega\delta(t)\sin(\omega x) \approx \cos(\omega x + \beta\omega\delta(t)) \tag{3.10}$$

$$\Rightarrow \quad \cos(\omega x) - \beta\omega\delta(t)\sin(\omega x) \approx \cos(\omega x)\cos(\beta\omega\delta(t)) - \sin(\omega x)\sin(\beta\omega\delta(t)) \quad (\because \cos(A+B) = \cos A\cos B - \sin A\sin B) \tag{3.11}$$

The Equation 3.11 holds only when

$$\cos(\beta\omega\delta(t)) \approx 1 \tag{3.12}$$

$$\sin(\beta\omega\delta(t)) \approx \beta\omega\delta(t) \tag{3.13}$$

which hold within 10% for $\beta\omega\delta(t) \leq \frac{\pi}{4}$ (the sine term is the leading approximation and $\sin(\frac{\pi}{4}) = 0.9\frac{\pi}{4}$).

$$\therefore \quad \beta\omega\delta(t) \leq \frac{\pi}{4} \tag{3.14}$$

In terms of spatial frequency, $\omega = \frac{2\pi}{\lambda}$, Equation 3.14 becomes

$$\beta\delta(t) \leq \frac{\lambda}{8} \tag{3.15}$$

Re-substituting $\beta = 1 + \alpha$ in Equation 3.15,

$$\boxed{(1+\alpha)\delta(t) \leq \frac{\lambda}{8}} \tag{3.16}$$

The above Equation 3.16 provides the required largest motion magnification factor, α for a given motion, $\delta(t)$ and image spatial wavelength, λ.

3.2.1.3 Multiscale Analysis

The above analysis (Section 3.2.1.2 suggests a scale-varying process. Here, a specified magnification factor, α, is used for some desired band of spatial frequencies. Then, magnification factor, α, is scaled back for the high spatial frequencies where amplification would give undesirable artifacts. Figure 3.3 shows such a modulation scheme for magnification factor, α. The spatial frequency content of the different levels can be estimated using corresponding levels of the Laplacian pyramid.

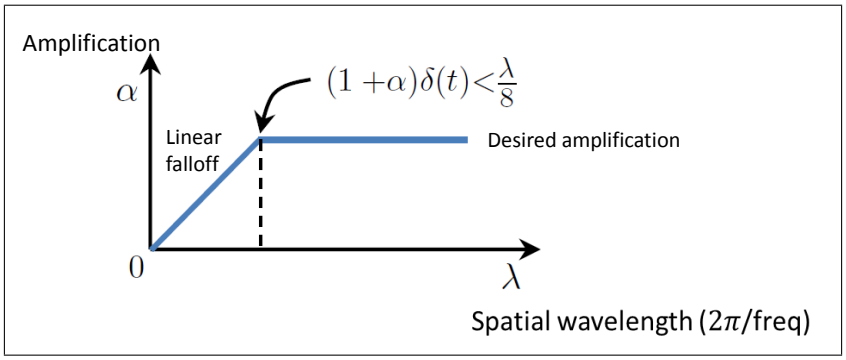


Figure 3.3 Motion magnification factor, α, as function of spatial wavelength λ. This Figure is adapted from [75]. Magnification factor is fixed to α for spatial bands that are within derived bound (Equation 3.16), and is attenuated linearly for higher spatial frequencies.

3.2.2 Discussion

The above explained Eulerian video magnification requires the following.

- Motion type present in video should be known to find the suitable filter, which is used in temporal analysis.
- Band of frequencies present in video should be known to find the filter parameters.
- User has to specify the magnification parameters (cut-off wavelength, λ_c and magnification factor, α).

In our work, we attempted to estimate filter parameters and magnification parameters automatically based on time-frequency analysis. We explain the proposed small motion video magnification.

3.3 Proposed Method for Estimating Parameters for Automated Magnification of Small Motions in Videos

To magnify small motions, in [75], given video is decomposed spatially. Then user has to set parameters for filtering and magnification. In our work, we estimate these parameters automatically. Main contribution of this work is to magnify the small motions in video automatically. It consists of the following two contributions: (1) Estimate bandwidth for temporal filter and (2) Estimate the magnification parameters, automatically. We call proposed method as semi-automated due to the fact that temporal filters used are from [75].

Overview of the proposed framework is given in Figure 3.4. Given an input video of N frames, $\{I_t\}_{t=1}^{N}$, the proposed method consists of the following steps: **(1) Estimation of parameters:** (i) Extract first two frames, $\{I_t\}_{t=1}^{2}$; (ii) Estimate landmark pixels, $\{l_i\}_{i=1}^{k}$; (iii) Extract time series at these 'k' pixels from N frames; (iv) Apply time frequency representation; (v) Estimate parameters for bandpass

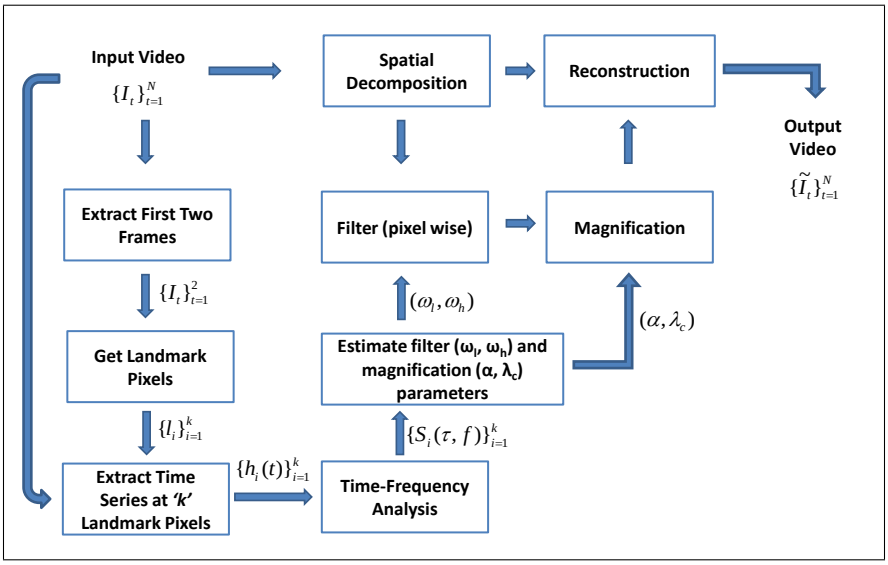


Figure 3.4 Overview of the proposed framework. Initially, parameters (ω_l, ω_h, α, λ_c) are estimated from the landmark pixels ($\{l_i\}_{i=1}^{k}$) determined from the first two frames of input video ($\{I_t\}_{t=1}^{2}$). Then input video ($\{I_t\}_{t=1}^{N}$) is divided into spatial bands, which are pixel-wise filtered with same temporal filter and amplified by magnification factor, α. Reconstruction ($\{\tilde{I}_t\}_{t=1}^{N}$) is done by adding these amplified bands along with original spatial bands.

filter (ω_l, ω_h) and magnification (α, λ_c); **(2) Magnification:** (i) Decompose the given video spatially into different spatial bands; (ii) Apply pixel-wise bandpass filter with ω_l and ω_h as lower and higher cut-off frequencies; (iii) Multiply with magnification factor, α for wavelengths less than λ_c; **(3) Reconstruction:** (i) Reconstruct the video, $\{\tilde{I}_t\}_{t=1}^{N}$ by adding magnified signal to original signal.

3.3.1 Landmark Detection

To get the landmark pixels, it was observed that considering all pixels for determining bandwidth and magnification factor is not efficient due to the fact that (1) the whole process will be time consuming due to the computationally intensive time-frequency analysis and (2) videos, considered for this work, have small motions in few regions and hence, all pixel locations do not necessarily undergo motion. Therefore, we adapted a mechanism to find landmark pixels. These pixels are obtained from the difference of edge maps of first two frames of a given video because the pixels at edges definitely experience motion from one frame to another frame. This is shown in Figure 3.5. Time signals at these landmark pixels are used to determine the parameters (See Figure 3.6).

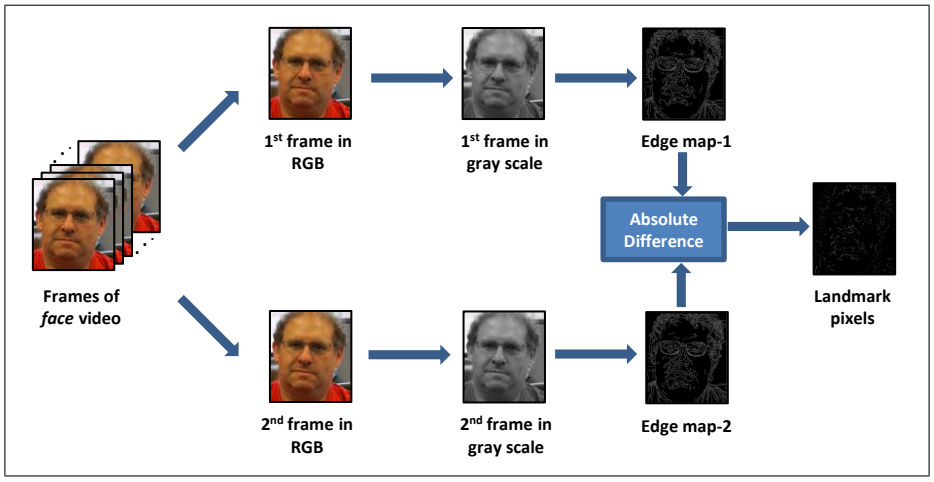


Figure 3.5 Estimation of landmark pixels.

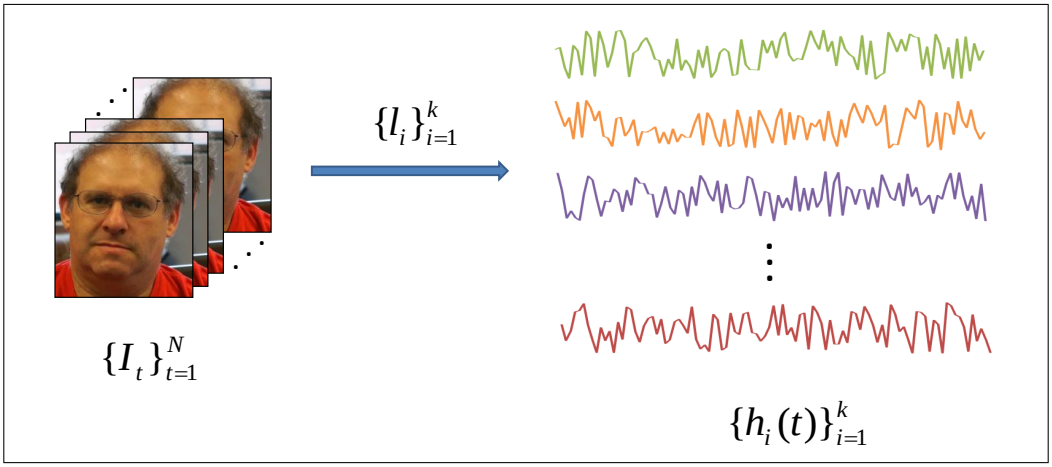


Figure 3.6 Extraction of time series.

3.3.2 Time-Frequency Analysis

3.3.2.1 Stockwell Transform

Even though Fourier transform gives the information about the spectral components in a signal, it fails to locate where those frequencies occur in that signal. So, it is preferable to consider time frequency representation (TFR). Different techniques for time frequency representation have been proposed. A few of them are short time Fourier transform (STFT), Gabor transform, continuous wavelet transform (CWT) and Wigner ville distribution etc. In [63], it was proven that Stockwell transform outperforms all these TFR techniques in localizing time and frequency because it has frequency dependent resolution whereas other transforms have windows of fixed width.

For a given time signal $h(t)$, its Stockwell transform is defined as,

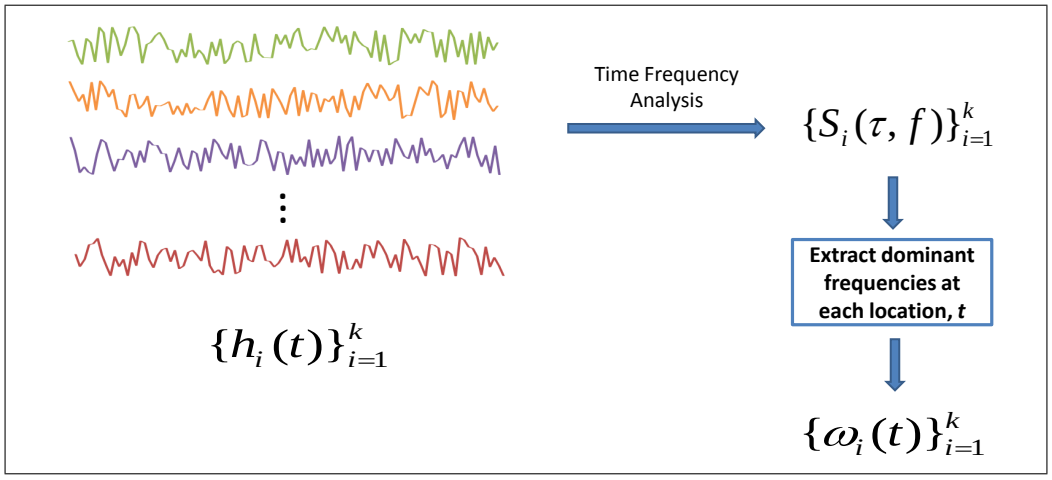


Figure 3.7 Applying Stockwell transform on time series.

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-(\tau-t)^2 f^2/2} e^{-i2\pi f t} dt \tag{3.17}$$

where $h(t)$ is the time signal, f denotes the frequency and τ denotes time shift parameter. An example is shown for a synthetic signal in Fig 4.3. Previous works, for example [52] has shown that temporal processing is done generally to extract invisible information from the signal. In general, spectral band is determined either empirically or assumed depending on application but such approaches fail to find the dominant frequency components automatically. We eliminate such dependency. In this work, we consider time signals from a video sequence, only at the locations of landmark pixels.

3.3.3 Parameter Estimation

We find bandwidth for bandpass filtering the time signals automatically. These limits of frequency are determined using Stockwell transform (See Figure 3.7). It is observed that the minimum and maximum frequencies of mean of frequencies obtained from each time series of a pixel can give useful information for temporal processing. We carried out this processing on time series data in Y space of YIQ colour space.

Lower (ω_l) and higher (ω_h) cut off frequencies are given by,

$$\omega_l = \min_i(\text{mean}_t(\{\omega_i(t)\}_{i=1}^{k})) \tag{3.18}$$

$$\omega_h = \max_i(\text{mean}_t(\{\omega_i(t)\}_{i=1}^{k})) \tag{3.19}$$

From Stockwell transform, we estimate the magnification factor α. It is determined as follows,

$$(1+\alpha)h(t) < \frac{\lambda_c}{8} \tag{3.20}$$

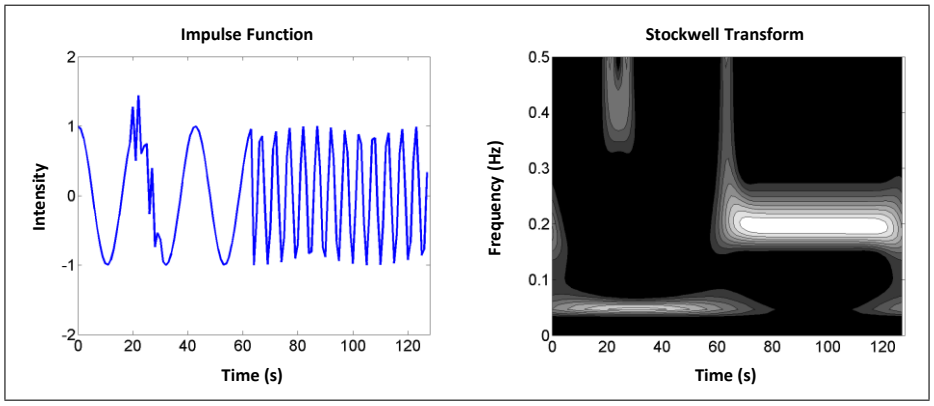


Figure 3.8 An example of Stockwell transform. Time signal shown in left is $h[0:63] = \cos(2\pi t * 6/128), h[64:127] = \cos(2\pi t * 25/128), h[20:30] = h[20:30] + 0.5 * \cos(2\pi t * 52/128)$. It contains a low-frequency signal for the first half, a middle-frequency signal for the second half and a high-frequency burst at $t = 20$. All these frequencies are clearly visible along with time location in Stockwell transform as shown in right (bright pixels indicate high strength of transform).

where $h(t)$ is time signal and λ_c is cut-off wavelength beyond which magnification factor, α is zero. Eq. (3.20) provides the largest magnification factor.

Eq. (3.20) is modified as,

$$(1 + \alpha_i(t))h_i(t) < \frac{\lambda_i(t)}{8} \quad \forall i = 1 \rightarrow k \tag{3.21}$$

where $h_i(t)$ is time signal at i^{th} landmark pixel, $\alpha_i(t)$ provides magnification factors for $h_i(t)$ and $\lambda_i(t)$ is cut-off wavelength corresponding to $h_i(t)$ beyond which magnification factor, $\alpha_i(t)$ is zero.

As explained in Section 4.4, Stockwell transform gives time frequency representation. We utilize this time frequency representation to find desirable magnification factor α. From Stockwell transform, we obtain information about which wavelengths (or frequencies) are occurring at what times. Thus we get $\lambda_i(t)$ from Stockwell transform. These $\{\lambda_i(t)\}_{i=1}^k$ are substituted in Eq. 3.21 to get $\{\alpha_i(t)\}_{i=1}^k$. Since videos involve small motion, we need to consider landmark pixels at which time signals can be used to determine magnification factor. A few of these time signals may contain noise and hence, may give incorrect α and λ_c values. To overcome this problem, we have considered median of all magnification factors and maximum of all wavelengths obtained from Eq. (3.20). This can be shown mathematically as,

$$\lambda_c = \max_i(\max_t(\{\lambda_i(t)\}_{i=1}^k)) \tag{3.22}$$

$$\alpha = \text{median}_i(\max_t(\{\alpha_i(t)\}_{i=1}^k)) \tag{3.23}$$

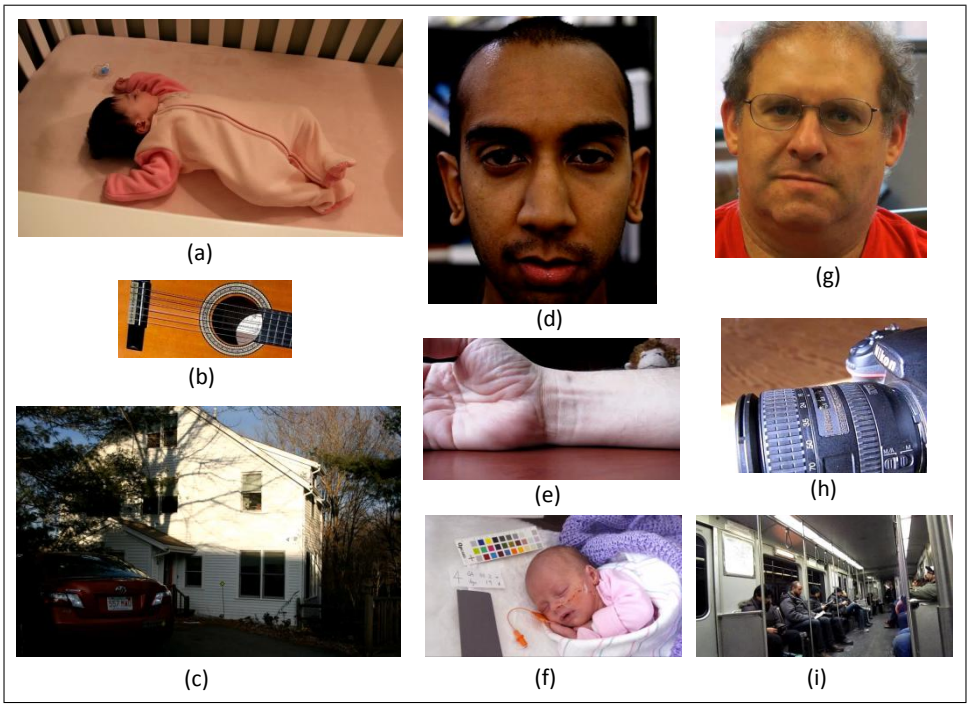


Figure 3.9 Sample frames from videos used in experimentation (a)*baby* (b)*guitar* (c) *shadow* (d)*face2* (e)*wrist* (f)*baby2* (g)*face* (h)*camera* (i)*subway*. All these are shown in proportion with the size of the corresponding video.

3.3.4 Motion Magnification

Spatial decomposition of videos into different spatial frequency bands is done to increase the signal-to- noise ratio using Laplacian pyramid [11]. These frequency bands are filtered and magnified pixel-wise differently according to the level in pyramid [75] using estimated parameters. Reconstruction is done by adding original signal from spatial decomposition and magnified signal.

3.4 Experiments and Results

3.4.1 Experimental Setup

We use data given in [74] for our experiments. This data consists of 9 videos, namely *baby*, *baby2*, *camera*, *face*, *face2*, *guitar*, *shadow*, *subway* and *wrist* as mentioned in Table 3.1. Sample frames from these videos are shown in Figure 3.9. All experiments are implemented using MATLAB on a system with 4GB RAM and Intel® core i5 CPU with 2.5 GHz processor. Every video takes time in the order of a few minutes to compute.

Video	Frame Rate (Hz)	Length (seconds)
baby	30	10
baby2	30	29
camera	300	33
face	30	10
face2	30	10
guitar	600	10
shadow	30	6
subway	30	8
wrist	30	29

Table 3.1 Video Clips used in Testing.

Video	α	λ_c	ω_l(Hz)	ω_h(Hz)
baby	32	30	0.5	3
baby2	30	30	0.6	2.8
camera	7	29	1.5	16.9
face	55	30	0.6	5.6
face2	17	30	1.02	10.56
guitar	101	29	0.77	8.16
shadow	13	30	0.92	5.8
subway	42	29	0.52	4.05
wrist	11	30	0.5	4

Table 3.2 Estimated parameters for magnification factor α, cut-off wavelength λ_c, lower cut-off frequency for temporal filter ω_h, higher cut-off frequency for temporal filter ω_h.

Video	ω_l(Hz)		ω_h(Hz)	
	Proposed	[75]	Proposed	[75]
baby	0.5	0.4	3	3
baby2	0.6	2.33	2.8	2.67
camera	1.5	45	16.9	100
face	0.6	0.83	5.6	1
face2	1.02	0.83	10.56	1
guitar	0.77	72	8.16	92
shadow	0.92	0.5	5.8	10
subway	0.52	3.6	4.05	6.2
wrist	0.5	0.4	4	3

Table 3.3 Temporal bandwidth parameters (lower cut-off frequency ω_h and higher cut-off frequency ω_h) with proposed method and [75].

Video	α		λ_c	
	Proposed	[75]	Proposed	[75]
baby	32	10	30	16
baby2	30	150	30	200
camera	7	120	29	20
face	55	100	30	1000
face2	17	20	30	80
guitar	101	100	29	40
shadow	13	5	30	48
subway	42	60	29	90
wrist	11	10	30	80

Table 3.4 Magnification parameters (magnification parameter α and cut-off wavelength λ_c) with proposed method and [75].

3.4.2 Results

Filter parameters, magnification factor and cut-off wavelength are determined from Stockwell transform by considering only landmark pixels. Obtained magnification factors α and cut-off wavelengths λ_c for each video are listed in Table 3.2. Comparison of temporal parameters and magnification parameters with that of [75] are shown in Table 3.3 and Table 3.4 respectively. Here, evaluation of our method is not in terms of getting exact parameters as in [75], but in terms of qualitative results. Even though estimated parameters are not close to the values mentioned in [75], reconstructed videos are comparable to those of [75]. These results are available on `http://researchweb.iiit.ac.in/~sushma.m/premiResults`.

3.5 Summary

In this chapter, we discussed magnification of small motions in videos automatically by estimating parameters. We used a time frequency representation called Stockwell transform for this purpose. We introduced computationally inexpensive techniques to estimate parameters. We illustrated the utility of the proposed method on examples in which small motions were made visible.

Chapter 4

Motion Detection in Perfusion Weighted MRI using ST

Magnetic resonance imaging (MRI) has been emerging as an efficient tool in clinical practice for the analysis of brain functions through several metabolic parameters. There are two types of MRI, namely, diffusion weighted imaging (DWI) and perfusion weighted imaging (PWI). PWI has been used extensively for the evaluation of tissue after acute stroke, non-invasive histologic assessment of tumors and evaluation of neurodegenerative conditions such as Alzheimers disease [48]. Diffusion weighted images are obtained by incorporating strong agnetic field gradient pulses into an imaging pulse sequence. In DWI, structures with fast diffusion are dark due to the fact that these structures are subject to greater signal attenuation, whereas structures with slow diffusion are bright. In PWI, an exogenous tracer is introduced into the blood circulation and its concentration in a tissue is monitored in a tissue over time. Blood flow to the corresponding tissue can be determined by obtaining the rate of delivery of the tracer. There are two types of PWI: (i) dynamic susceptibility contrast (DSC) imaging, and (ii) dynamic contrast enhanced (DCE) $T1$ weighted imaging. DSC is most widely used for the brain, while DCE is most widely used in the rest of the body though its experimental and research use is increasing in brain. In this work, we address the motion detection problem using DSC-MRI.

4.1 Introduction

In PWI, blood perfusion is measured by injecting an exogenous tracer called bolus into the blood flow of a patient and then tracking it in the brain. PWI requires a long data acquisition time to form a time series of volumes. Hence, motion occurs due to patient's unavoidable movements during a scan, which in turn results into motion corrupted data. There is a necessity of detection of these motion artifacts on captured data for correct disease diagnosis. In PWI, intensity profile gets disturbed due to occurrence of motion and/or bolus passage through the blood vessels. There is no way to distinguish between motion occurrence and bolus passage. In this paper, we propose an efficient time-frequency analysis based motion detection method. We show that proposed method is computationally inexpensive and fast. This method is evaluated on a DSC-MRI sequence with simulated motion of different degrees. We show that our approach detects motion in a few seconds.

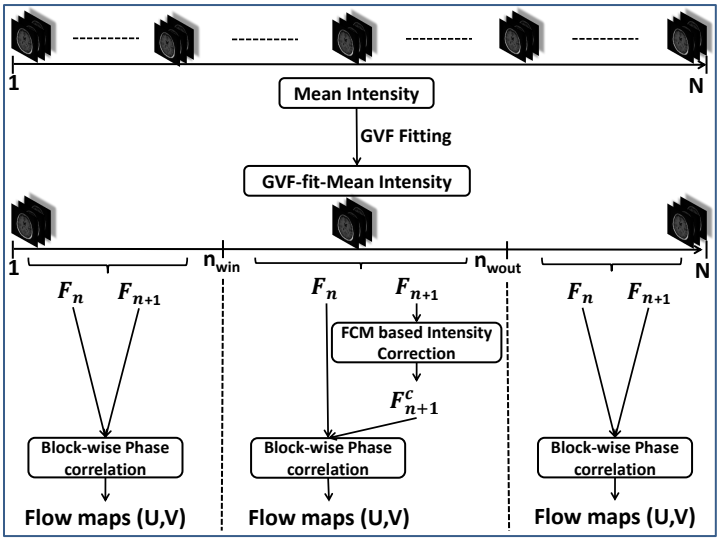


Figure 4.1 Motion Detection Scheme for a DSC-MRI time-series. This Figure is adapted from [55].

In PWI, a time series of volumes are formed in a long acquisition time. Patient often has difficulty in staying still during this period. Therefore, it is more likely that patient may move unavoidably during scanning which in turn results into motion artifacts in scans. There is a need of detection and subsequent correction of these motion artifacts. There are works in medical imaging, for example [10], [32], [61] addressed this problem in terms of registration of whole time series to a reference volume. These methods do not detect motion. Hence, non-corrupted volumes are also registered which makes the process computationally expensive and it is obvious that these volumes do not need any correction [55] . Therefore, it is preferable to have a prior knowledge about motion corrupted volumes.

In DSC-MRI, intensity profile over time should be flat. If there are any disturbances in intensity profile, it can be due to two reasons: (i) passage of bolus through the blood vessels and (ii) motion of the patient during scanning. Therefore, while detecting motion, bolus passage should also be taken care of. However, traditional motion detection methods consider non-uniform intensity variations due to bolus passage as motion corruption. Hence, they may fail to detect motion in perfusion MRI. In [55] and [26], motion is detected by bolus dependent approach. Here, perfusion MRI data is divided into three sets as (i) pre wash-in, (ii) transit and (iii) post wash-out sets. Intensity correction is applied to transit set and then motion is detected in each set differently. In next section, we explain the phase correlatio based motion detection method used in [55] and [26].

4.2 Phase Correlation based Motion Detection

Phase correlation based motion detection in PWI is shown in Figure 4.1. Initially, gamma variate function (GVF) is used to divide the PWI series into 3 sets according to the bolus status as explained below.

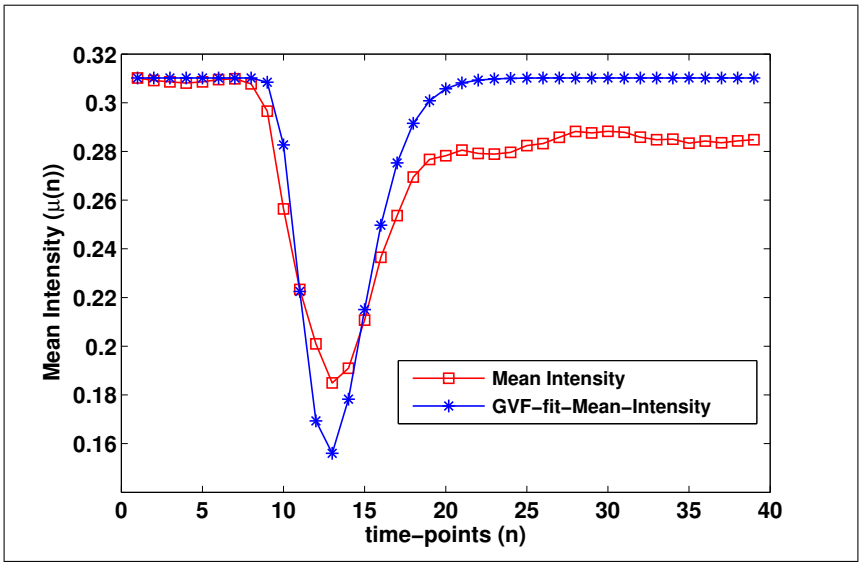


Figure 4.2 Gamma-variate-function (GVF) fitting on the mean-intensity curve of a DSC-MRI time-series. This Figure is adapted from [55].

Gamma-variate function [13] [43] describes the transverse relaxation rate of magnetization with the passage of the bolus. This function is given as:

$$\triangle R_2^*(t) = A(t - t_0)^{\alpha} e^{-\frac{t-t_0}{\beta}}, \quad t > t_0 \tag{4.1}$$

where, $\triangle R_2^*(t)$ is the transverse relaxation rate, t_0 is the wash-in time-point of bolus, and A, α and β are parameters which describe the shape of function. The mean intensity of each volume is computed to derive a curve $Ia(t)$. The wash-in time-point (t_0) is set at time point where mean intensity falls abruptly. Gamma-variate function (GVF) is then fit to $Ia(t)$ to accurately determine the wash-in $(n_{wash-in})$ and wash-out time points $(n_{wash-out})$ (see Figure 4.2). These are the time-points of abrupt change in signal intensities (due to bolus passage) in the GVF-fit-mean intensity curve $(G_a(t))$. Then given PWI series is divided into three sets: (i) pre-wash-in ($[0 n_{wash-in}]$) (ii) transit ($[n_{wash-in} n_{wash-out}]$) and (iii) post-wash-out ($[n_{wash-in} n_{wash-out}]$) sets.

After dividing PWI time series into the above said stages, intensity correction is applied to the transit set of volumes. Fuzzy c-means clustering is used to segment a volume (T) in the transit set as normal (T_a) and bolus affected regions (T_b). Then GVF fitting is used to obtain the intensity changes across time points. The intensity correction is applied to only bolus affected regions. The intensity corrected volume, (T_n^c) is generated as follows,

$$T_b^c(n) = T_b \frac{G_a(n_c)}{G_a(n)} \tag{4.2}$$

$$T^c(n) = T_a(n) \cup T_b^c(n) \tag{4.3}$$

where n_c is the center of time-points from $n_{wash-in}$ to $n_{wash-out}$, i.e., $n_c = \frac{n_{wash-in}+n_{wash-out}}{2}$.

After intensity correction to bolus affected volumes, motion corrupted volumes are detected using phase correlation. Since, there is only inter volume motion is present in PWI series, central slices of all volumes are considered to detect motion. (I_n, I_{n+1}) denote the central slices of pair of adjacent volumes (T_n, T_{n+1}) in a series of N volumes, $\{T_i\}_{i=1}^{N}$. Motion field between these adjacent volumes, (U_n, V_n) is computed using (I_n, I_{n+1}). This motion field will be non-zero, if there is any motion present between them. This motion field is computed by considering a block of pixels $b_n(i,j)$ and $b_{n+1}(i,j)$ around every pixel at location (i,j) in I_n and I_{n+1}. Inter slice intensity variation is compensated by normalizing the blocks by shifting the mean pixel value of every block to zero. Then phase correlation is applied to normalized blocks ($\tilde{b}_n(i,j)$, $\tilde{b}_{n+1}(i,j)$). Locations of maxima of cross power spectrum $G_{\tilde{b_n}\tilde{b_{n+1}}}$ give the flow vector $\vec{r}(i,j) = (u(i,j), v(i,j))$. Then flow maps $U_n = [u(i,j)]$ and $V_n = [v(i,j)]$ for (I_n, I_{n+1}) are given by,

$$(u,v) = argmax_{(i,j)\varepsilon b_{i,j}} \left(\mathfrak{F}^{-1} \left\{ G_{\tilde{b}_n \tilde{b}_{n+1}} \right\} \right) \tag{4.4}$$

where $\mathfrak{F}^{-1}$ denotes the inverse Fourier transform, $G_{\tilde{b}_n \tilde{b}_{n+1}} = \frac{\tilde{B}_n \tilde{B}^*_{n+1}}{|\tilde{B}_n||\tilde{B}^*_{n+1}|}$ and $\tilde{B} = \mathfrak{F}\{\tilde{b}\}$.

Motion corrupted volumes are detected by using total entropy by adding entropies in U_n and V_n. Here, non-zero entropy indicates presence of motion.

In implementation, given 128×128 was downsampled by a factor of 4 to obtain a slice size of 32×32 and block size was chosen as 8×8 for computational efficiency.

4.3 Challenges in Motion Detection in PWI

The following are the challenges in motion detection in PWI.

- Traditional motion correction methods do not detect motion before correction (see Section 4.1).
- [26] detects motion prior to motion correction. However, it has the following disadvantages.
  - It requires identification of bolus stages so that intensity correction can be applied to bolus affected regions. This intensity correction is based on GVF and Fuzzy c-means clustering .
  - There is a trade-off between block size and the computation time. It means that higher the block size, the computation time is more. It can not distinguish no motion case with the motion below 3^0. Sensitivity to detect motion is dependent on slice resolution and block size. Sensitivity is more, if the slice resolution is kept unchanged, i.e., 128×128. But, it increases the computation time.

Even though MRI scans consist of volumes of two dimensional images, they are acquired over the time. Therefore, we can extract one dimensional time series from volumes. Motion detection using

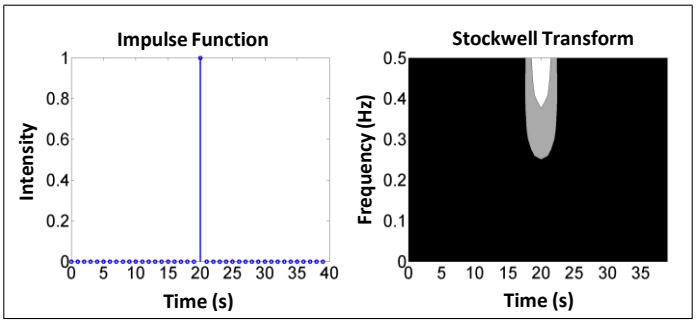


Figure 4.3 Stockwell transform for an impulse function. Impulse function shown in left is $h[0:39] = 0, h[20] = 1$. Stockwell transform is shown in right. Note that bright pixels indicate high strength of transform. Here, bright pixels are at $t = 20$.

one dimensional time series is obviously faster compared to that of two dimensional scans. These facts motivated us to analyze the MRI data in terms of one dimensional time sequences because we believe that frequency of time series will vary when there are motion artifacts. Most popular approach for temporal analysis is Fourier transform. Even though Fourier transform gives the information about the spectral components in a signal, it fails to locate where those frequencies occur in that signal. So, it is preferable to consider time frequency representation (TFR). Different techniques for time frequency representation have been proposed. A few of them are short time Fourier transform (STFT), Gabor transform, continuous wavelet transform (CWT) and Wigner ville distribution etc. In [63], it was proven that Stockwell transform (ST) outperforms all these TFR techniques in localizing time and frequency because it has frequency dependent resolution whereas other transforms have windows of fixed width. ST provides useful phase of the spectrum which is not available from CWT.

In the recent past, ST has been used for the analysis of MRI data. In [28], ST is used to remove artifacts in functional MRI (fMRI) time courses due to which brain activity detection is improved. In [76], polar version of ST is used to analyze the texture patterns in MRI for the diagnosis of multiple sclerosis. [77] discusses the effectiveness of ST for medical imaging and shows how to enhance fMRI time courses by removing frequency artifacts which are introduced due to patient's quick breathing.

In this chapter, we demonstrate how one dimensional ST based framework can be used to detect motion without identifying bolus stages. Given a MRI sequence of volumes, we consider specific key points which are generated by an automated method. Time series are extracted at these key points and ST is applied on them. The process is computationally inexpensive due to the facts that (i) there is no explicit intensity correction for bolus stages, which means proposed method works well regardless of bolus stages, (ii) motion is detected by one dimensional time series instead of two dimensional scans and, (iii) these time series are extracted only at a few key points. Mean time taken for motion detection is around 3 seconds. In next section, we explain Stockwell transform (ST).

4.4 Stockwell Transform

In this section, we discuss details of Stockwell transform (ST) and its suitability for analyzing MRI times sequences. For a given time signal $h(t)$, its Stockwell transform is defined as,

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t)w(\tau - t, f)e^{-i2\pi ft}dt \tag{4.5}$$

where $w(t, f)$ is defined as

$$w(t, f) = \frac{|f|}{\sqrt{2\pi}}e^{-t^2f^2/2}, \tag{4.6}$$

$h(t)$ is the time signal, $w(t, f)$ denotes the window, f denotes the frequency, τ denotes time shift parameter, and $|.|$ denotes absolute value. Since window ($w(t, f)$) is frequency dependent, narrower windows are applied at higher frequencies and broader windows are applied at lower frequencies. Hence, ST is a suitable time-frequency representation for current work.

In our work, intensity profiles over time are obtained from DSC-MRI sequence. There are strong disturbances in these one dimensional time series (intensity profiles) corresponding to motion corrupted volumes. In general, there will be many disturbances in time series with respect to number of continuous corrupted slices in MRI sequence. Since ST is a linear function, we explain it with one single intensity disturbance for simplicity. To show that these intensity variations are well represented by Stockwell transform, we modelled single disturbance in time series as an impulse, $h(t)$ as shown in Eq. 4.7.

$$h(t) = \delta(t - a) \tag{4.7}$$

ST of $h(t)$ is given as

$$S(\tau, f) = \frac{|f|}{\sqrt{2\pi}}e^{-(\tau-a)^2f^2/2}e^{-i2\pi fa} \tag{4.8}$$

Absolute value of $S(\tau, f)$ is given as

$$|S(\tau, f)| = \frac{f}{\sqrt{2\pi}}e^{-(\tau-a)^2f^2/2} \tag{4.9}$$

On taking first and second derivatives of Eq. 4.9 w.r.t. τ, we obtain

$$\frac{\partial|S(\tau, f)|}{\partial\tau} = -\frac{f^3}{\sqrt{2\pi}}(\tau - a)e^{-(\tau-a)^2f^2/2} \tag{4.10}$$

$$\frac{\partial^2|S(\tau, f)|}{\partial\tau^2} = \frac{f^3}{\sqrt{2\pi}}e^{-(\tau-a)^2f^2/2}[-1 + f^2(\tau - a)^2] \tag{4.11}$$

In Eq. 4.9, $|f|$ is replaced with f because $f > 0$. It can be easily observed that maximum of $|S(\tau, f)|$ occurs at $t = a$ because $\frac{\partial|S(\tau,f)|}{\partial\tau}\Big|_{\tau=a} = 0$ and $\frac{\partial^2|S(\tau,f)|}{\partial\tau^2}\Big|_{\tau=a} < 0$. Stockwell transform for $h(t) = \delta(t - a)$ with $a = 20$ is shown in Figure 4.3. There is bright region around $t = 20$ in ST. Even though there is some noise present in the time series, it does not affect ST much because the region around $t = 20$ will still be bright relative to other regions.

Algorithm 1 Motion Detection in PWI.

Input: Central slices, $\{C_t(i,j)\}_{t=1}^{N}$ where $i = 1 \rightarrow A$, $j = 1 \rightarrow B$

1. *Pre-processing:*

   - Pre-process all central slices for removing the noise regions at edge regions to get $\{\tilde{C}_t(i,j)\}_{t=1}^{N}$

     for $t = 1$ to N **do**

     $\tilde{C}_t = C_t$

     $M = mean(C_t)$

     $\tilde{C}_t = 0 \mid \tilde{C}_t < M \quad \forall \quad i,j$

     end for

2. *Find locations of landmark pixels:*

   - Find difference of consecutive slices, $\{D_{t,t+1}\}_{t=1}^{N-1}$

     $D_{t,t+1} = |\tilde{C}_t - \tilde{C}_{t+1}| \quad \forall \quad t = 1 \rightarrow N-1$

   - Sum all $\{D_{t,t+1}\}_{t=1}^{N-1}$, i.e., $D = \sum_{t=1}^{N-1} \{D_{t,t+1}\}_{t=1}^{N-1}$

   - Find landmark pixels, $\{p_l\}_{l=1}^{L} = \{(x_l, y_l)\}_{l=1}^{L} \mid D(x_l, y_l) \neq 0$

3. *Divide landmark pixels into two sets:*

   - Landmarks set-1: $\{p_{l_1}\}_{l_1=1}^{L_1} = \{(x_{l_1}, y_{l_1})\}_{l_1=1}^{L_1}$

     $$\{(x_{l_1}, y_{l_1})\}_{l_1=1}^{L_1} = \left[\underset{(x_l,y_l)}{\arg\min}\, d((x_l,y_l),(x_l,1))\right] \cup \left[\underset{(x_l,y_l)}{\arg\min}\, d((x_l,y_l),(x_l,B))\right] \cup \left[\underset{(x_l,y_l)}{\arg\min}\, d((x_l,y_l),(1,y_l))\right] \cup \left[\underset{(x_l,y_l)}{\arg\min}\, d((x_l,y_l),(A,y_l))\right]$$

   - Landmarks set-2: $\{p_{l_2}\}_{l_2=1}^{L_2} = \{(x_{l_2}, y_{l_2})\}_{l_2=1}^{L_2} = \{(x_l, y_l)\}_{l=1}^{L} \ni \{(x_{l_1}, y_{l_1})\}_{l_1=1}^{L_1}$

4. *Extract time series:*

   - Extract time series $\{h_{l_1}(t)\}_{l_1=1}^{L_1}$ and $\{h_{l_2}(t)\}_{l_2=1}^{L_2}$ at locations $\{p_{l_1}\}_{l_1=1}^{L_1}$ and $\{p_{l_2}\}_{l_2=1}^{L_2}$ respectively, from $\{D_{t,t+1}\}_{t=1}^{N-1}$

     for $l_1 = 1$ to L_1 **do**

     for $t = 1$ to $N-1$ **do**

     $h_{l_1}(t) = D_{t,t+1}(x_{l_1}, y_{l_1})$

     end for

     end for

     for $l_2 = 1$ to L_2 **do**

     for $t = 1$ to $N-1$ **do**

     $h_{l_2}(t) = D_{t,t+1}(x_{l_2}, y_{l_2})$

     end for

     end for

5. Determine whether motion is present or not (see Algorithm 2)

If there is motion present, then follow steps 6 and 7.

6. *Determine Stockwell transform:*

   - Find Stockwell transform of each time series of $\{h_{l_1}(t)\}_{l_1=1}^{L_1}$, $\{S_{l_1}(\tau, f)\}_{l_1=1}^{L_1}$

     for $l_1 = 1$ to L_1 **do**

     $S_{l_1}(\tau, f) = \int_{-\infty}^{\infty} h_{l_1}(t) w(\tau - t, f) e^{-i2\pi f t} dt$

     end for

7. *Estimation of locations of corrupted slices:*

   - Sum all STs, $S(\tau, f) = \sum_{l_1=1}^{L_1} S_j(\tau, f)$
   - Find locations of bright regions, l_c and l_{c+M-1}

Output: Locations of corrupted volumes, $\{l_m\}_{m=c}^{c+M-1}$

4.5 Proposed Motion Detection Method

We propose a novel automated method to detect motion corrupted volumes in perfusion weighted MRI using Stockwell transform. We assume that there is no intra-volume motion in this MRI time series because it takes a few seconds time to scan. Therefore, whole volume is corrupted by same motion. Instead of considering whole volumes to detect motion, we consider only central slice of each volume. Motion can be identified through two ways, i.e., (i) intensity of voxel gets changed and (ii) a voxel comes into the location of another voxel. In this work, we detect motion using first category.

Intensity profile extracted from entire volume series can contain disturbances which maybe due to three reasons: (i) Presence of motion, (ii) Presence of bolus and (iii) Presence of both motion and bolus. We extract two intensity profiles from different sets of landmark pixels to distinguish these three cases. After this, we need to detect motion in cases (i) and (iii).

Overview of the proposed method to detect motion is shown in Fig. 4.4. For a given PWI volume series, we consider central slices to detect the motion in corresponding volumes. The proposed method consists of the following steps: (i) Pre-processing, (ii) Estimation of landmark pixels, (iii) Detection of presence of motion, (iv) Time-frequency analysis of time series extracted at these landmark pixels, and (v) Detection of locations of corrupted slices from time frequency representation along with the extracted time series. Steps (iv) and (v) are used only if there is presence of motion, which can be determined by step-(iii).

Algorithm for proposed motion detection method can be seen in Algorithm 1. If given MRI series contains N volumes, there will be corresponding N central slices, $\{C_t\}_{t=1}^{N}$. Each central slice is of size $A \times B$.

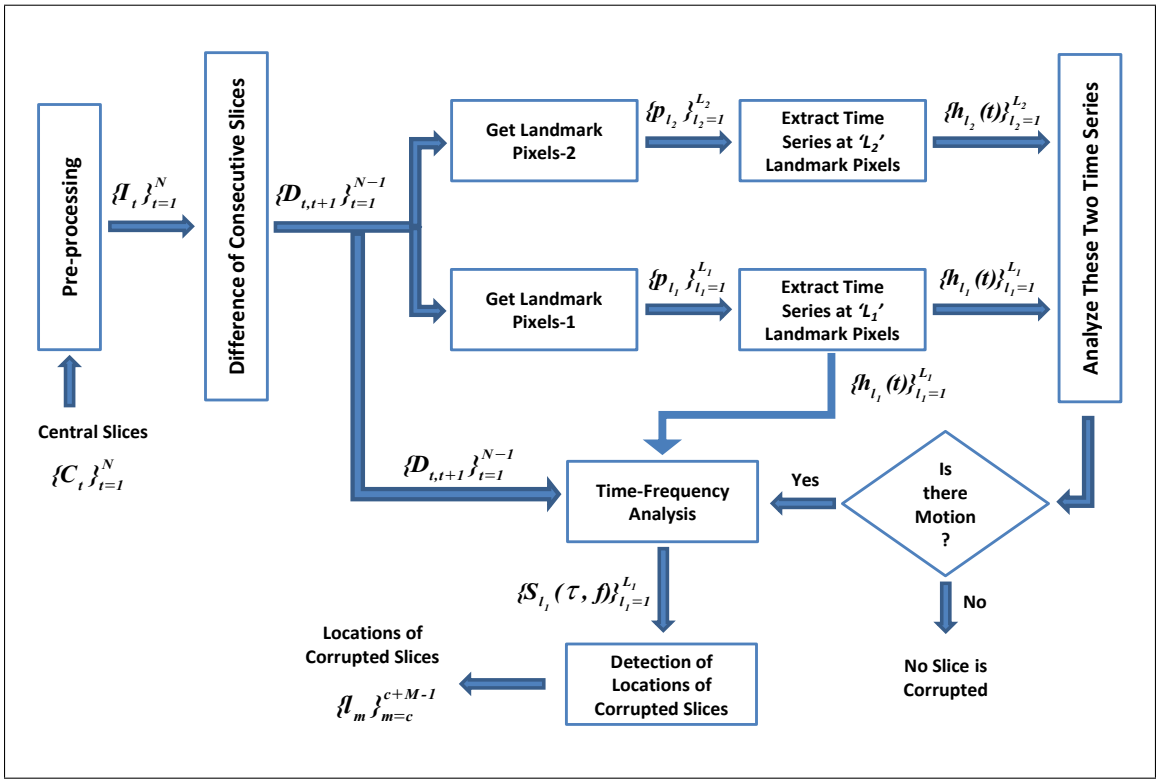


Figure 4.4 Motion Detection in both bolus and non-bolus stages. Central slices $\left(\{C_t\}_{t=1}^{N}\right)$ of a PWI MRI series are pre-processed to get noise-free images $\left(\{I_t\}_{t=1}^{N}\right)$. Then, two sets of one dimensional time series $\{h_{l_1}(t)\}_{l_1=1}^{L_1}$ and $\{h_{l_2}(t)\}_{l_2=1}^{L_2}$ at two sets of landmark pixels $\{p_{l_1}\}_{l_1=1}^{L_1}$ and $\{p_{l_2}\}_{l_2=1}^{L_2}$ are extracted from difference of consecutive pre-processed slices $\left(\{D_{t,t+1}\}_{t=1}^{N-1}\right)$. These two time series are analyzed to know whether there is motion present or not. if motion is present, time-frequency analysis $(\{S_{l_1}(\tau, f)\}_{l_1=1}^{L_1})$ is used to determine the locations of corrupted slices $\left(\{l_m\}_{m=c}^{c+M-1}\right)$.

4.5.1 Pre-processing

We first process all N central slices using intensity based thresholding technique such that the noise regions are discarded while preserving the edges. Here, intensities which are less than mean of all intensities present in slice are discarded to zero. These pre-processed images, $\{\tilde{C}_t\}_{t=1}^{N}$, are then used to detect the motion.

4.5.2 Estimation of Landmark Pixels

To determine the landmark pixels, it was observed that considering all pixels for detecting the motion corrupted slices is not efficient due to the fact that (i) the whole process will be time consuming and (ii) all pixels may not contain information about the corruption. Therefore, we adopted a mechanism to find landmark pixels. These pixels are obtained from the difference of consecutive central slices $\left(\{D_{t,t+1}\}_{t=1}^{N-1}\right)$ of all given volumes because the pixels at edges definitely experience motion from

one to another slice. All these difference maps are summed up $\left(D = \sum_{t=1}^{N-1} \{D_{t,t+1}\}_{t=1}^{N-1}\right)$ and then landmark pixels, $\{p_l\}_{l=1}^{L}$, are selected such that every non-zero pixel can be considered from D.

Here, two sets of landmark pixels, $\{p_{l_1}\}_{l_1=1}^{L_1}$ and $\{p_{l_2}\}_{l_2=1}^{L_2}$, are obtained to differentiate between presence of motion and presence of bolus. First set of landmark pixels ($\{p_{l_1}\}_{l_1=1}^{L_1}$) are selected such that every edge pixel can be considered and second set of landmark pixels ($\{p_{l_2}\}_{l_2=1}^{L_2}$) are selected such that they do not contain edge pixels but contain pixels from bolus affected regions. This can be shown mathematically as, $\{p_{l_2}\}_{l_2=1}^{L_2} = \{p_l\}_{l=1}^{L} - \{p_{l_1}\}_{l_1=1}^{L_1}$. Then, one dimensional time series, $\{h_{l_1}(t)\}_{l_1=1}^{L_1}$ and $\{h_{l_2}(t)\}_{l_2=1}^{L_2}$, at respective L_1 and L_2 landmark pixels are extracted from $\{D_{t,t+1}\}_{t=1}^{N-1}$.

4.5.3 Detection of Presence of Motion

For 'N' number of volumes, there will be N corresponding central slices and thus number of difference of consecutive slices will be $N-1$. Therefore, time series extracted at any landmark pixel is of length $N-1$. As explained before, there are two sets of landmark pixels, denoted as $LM_1 = \{p_{l_1}\}_{l_1=1}^{L_1}$ and $LM_2 = \{p_{l_2}\}_{l_2=1}^{L_2}$. For set LM_1, there are L_1 number of time series extracted from $\{D_{t,t+1}\}_{t=1}^{N-1}$. To reduce the effect of a few insignificant landmark pixels, mean of all these time series is considered. Mathematically, it can be denoted as,

$$h_1(t) = \frac{1}{N-1} \sum_{l_1=1}^{L_1} h_{l_1}(t) \tag{4.12}$$

Similarly for second set of landmark pixels,

$$h_2(t) = \frac{1}{N-1} \sum_{l_2=1}^{L_2} h_{l_2}(t) \tag{4.13}$$

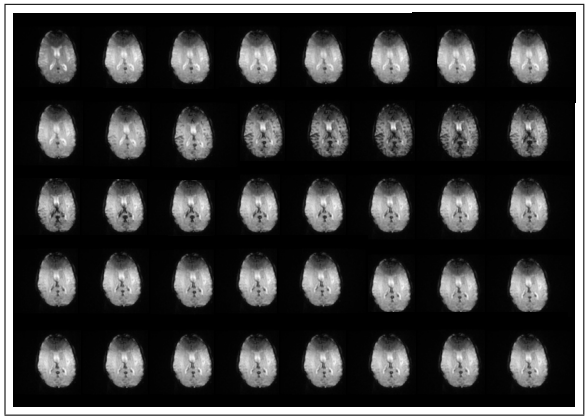
Now, these two time series, i.e., $h_1(t)$ and $h_2(t)$ are analyzed to determine whether the given sequence of volumes is motion corrupted or not. Algorithm for this analysis is given in Algorithm 2. $h_1(t)$ and $h_2(t)$ are pre-processed so that disturbances due to noise can be eliminated to some extent. Since $h_1(t)$ contains information from edges, it gives information about motion. Similarly, $h_2(t)$ contains more information from bolus regions and thus it gives information about presence of bolus. As explained in Algorithm 2, presence of motion is determined from a set of conditions as shown below.

Condition-1 : If $h_1(t) > h_2(t)$ and $h_2(t) = 0$, then there is presence of noise.

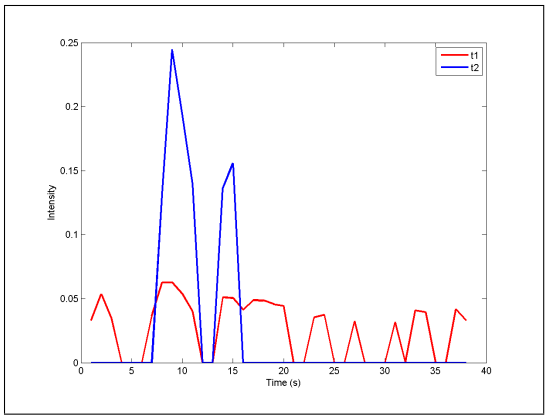
Condition-2 : If $h_2(t) > h_1(t)$, then there is presence of bolus.

Condition-3 : If $h_1(t) \geq h_2(t)$ and $h_2(t) > 0$, then there is presence of motion.

Sample cases for these three conditions are shown in Figures 4.5, 4.6 and 4.7.



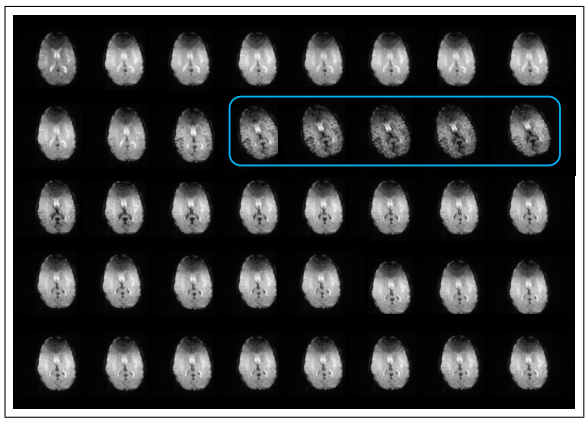
(a) Central slices with no motion.



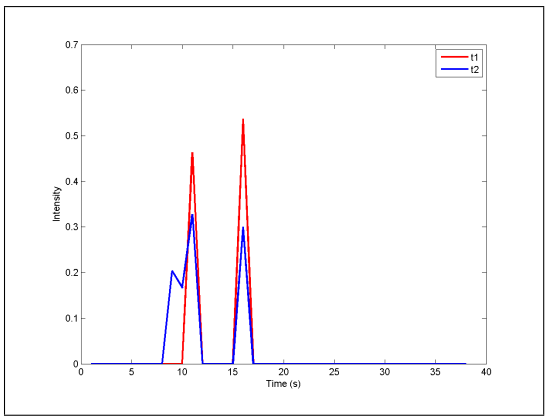
(b) h_1 and h_2 for no motion case. h_1 and h_2 are shown in red and blue colors respectively.

Figure 4.5 Central slices with no motion corruption and corresponding time series $h_1(t)$ and $h_2(t)$. $h_1(t)$ and $h_2(t)$ are shown in red and blue colors respectively. Here, conditions involved are, (i) $h_1(t) > h_2(t)$ and $h_2(t) = 0$; (ii) $h_2(t) > h_1(t)$.

In Figure 4.5(b), there are two conditions present. (i) $h_1(t) > h_2(t)$ and $h_2(t) = 0$ at few time instances which indicates there is presence of noise (***Condition-1***) and (ii) $h_2(t) > h_1(t)$ at remaining time instances which indicates there is presence of bolus (***Condition-2***). Therefore there is no motion corruption in corresponding PWI series. Corresponding central slices are shown in Figure 4.5(a).



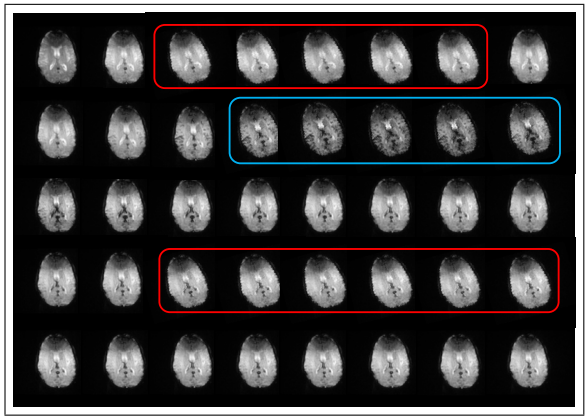
(a) Central slices with motion in bolus affected slices.



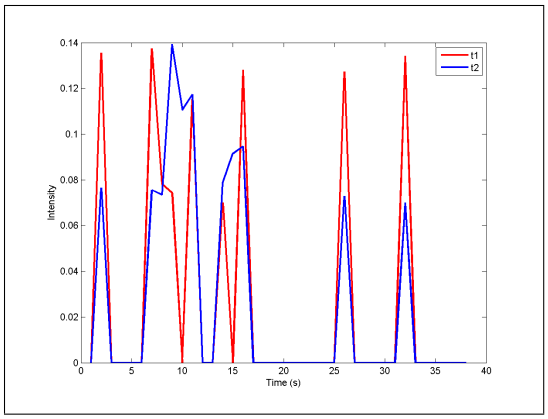
(b) h_1 and h_2 for motion in bolus affected slices case. h_1 and h_2 are shown in red and blue colors respectively.

Figure 4.6 Central slices with for motion in bolus affected slices and corresponding time series $h_1(t)$ and $h_2(t)$. Here, conditions involved are, (i) $h_1(t) \geq h_2(t)$ and $h_2(t) > 0$; (ii) $h_2(t) > h_1(t)$.

In Figure 4.6(b), (i) $h_1(t) \geq h_2(t)$ and $h_2(t) > 0$ at few time instances which indicates there is motion (***Condition-3***) and (ii) $h_2(t) > h_1(t)$ at same time instances and other time instances which indicates bolus is present (***Condition-2***). This means there is motion present in only bolus region of PWI series. Corresponding slices are shown in in Figure 4.6(a).



(a) Central slices with motion in both bolus affected and non-affected slices.



(b) h_1 and h_2 for motion in both bolus affected and non-affected slices case. h_1 and h_2 are shown in red and blue colors respectively.

Figure 4.7 Central slices with motion in both bolus affected and non-affected slices and corresponding time series $h_1(t)$ and $h_2(t)$. Here, conditions involved are, (i) $h_1(t) \geq h_2(t)$ and $h_2(t) > 0$; (ii)$h_2(t) > h_1(t)$.

In Figure 4.7(b), (i) $h_1(t) \geq h_2(t)$ and $h_2(t) > 0$ at few instances which indicates there is motion (***Condition-3***) and (ii) $h_2(t) > h_1(t)$ at some instances (***Condition-2***) other than instances in (i). This means, there is motion present in both bolus and non-bolus phases of PWI series. Corresponding slices are shown in in Figure 4.7(a).

4.5.4 Time-Frequency Analysis

If motion is present (information obtained from Algorithm 2), we detect where motion is present. For this purpose, we use Stockwell transform $\left(\{S_{l_1}(\tau, f)\}_{l_1=1}^{L_1}\right)$ at all L_1 extracted time series. We used first set of landmark pixels LM_1, because it contains more information about presence of motion. There might be still a few of landmark pixels which may not represent the pixels that undergo motion. To take care of this, Stockwell transforms at all landmark pixels can be added to get proper representation so that non-significant landmarks can play negligible role in detecting motion. This summed up Stockwell transform can be denoted as, $S(\tau, f) = \sum_{l_1=1}^{L_1} S_{l_1}(\tau, f)$.

4.5.5 Motion Detection

As explained in Section 4.4, there will be bright region at locations of corrupted slices. If there are M consecutive corrupted volumes, bright region will be around corresponding M locations in ST. We extract those bright regions and locations where those bright regions occur. If there is a bright region from location l_c to l_{c+M-1}, then we can categorize the slices at locations, $\{l_m\}_{m=c}^{c+M-1}$, as corrupted slices and corresponding volumes are motion corrupted.

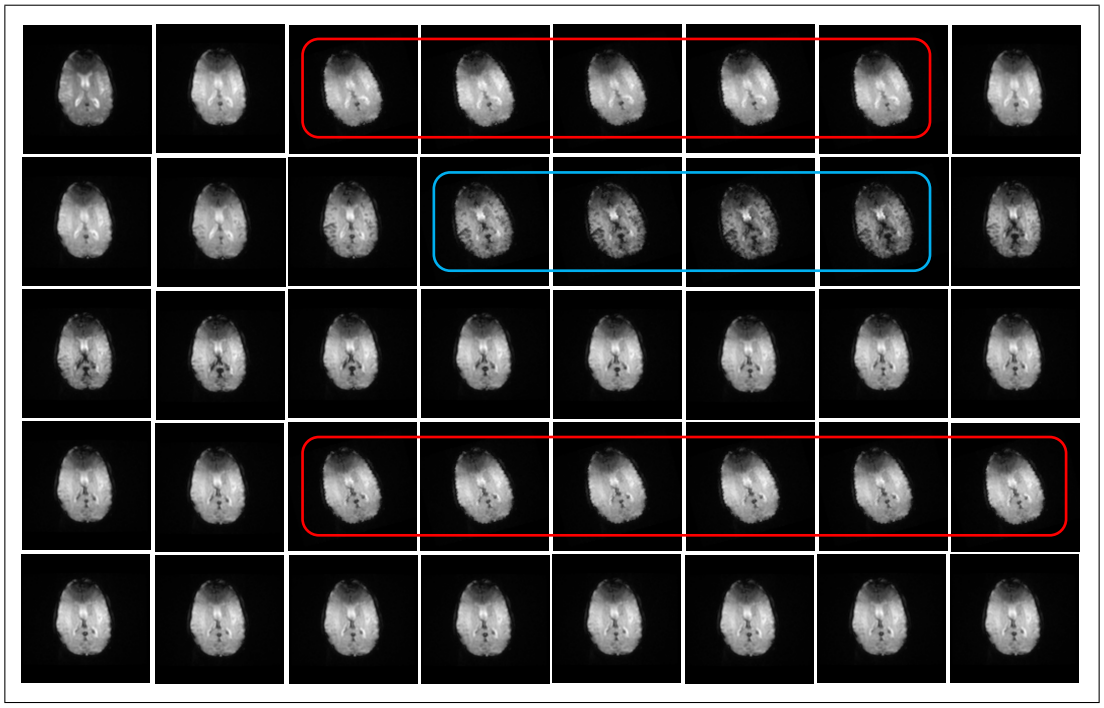


Figure 4.8 DSC-MRI time series. It shows central slices of 40 volumes from top to bottom and left to right. Here, central slices corresponding to non-bolus phase are from 1 to 9 and from 21 to 40. -15^0 rotation is added to volumes 3, 4, 5, 6, 7, 12, 13, 14, 15, 27, 28, 29, 30, 31 and 32 respectively and their corresponding central slices are shown in boxes. Motion in non-bolus phases (3, 4, 5, 6, 7, 27, 28, 29, 30, 31 and 32) and bolus phases (12, 13, 14 and 15) are shown in red and blue boxes respectively.

4.6 Experiments and Results

We have conducted experiments to validate the performance of the proposed framework with a DSC-MRI data obtained from a 1.5T GE MRI scanner. The data details are: number of volumes = 40 (1s/phase), number of slices = 20, dimensions of slice = 128x128 and thickness of slice = 5mm. All experiments are implemented on a system with 4GB RAM and Intel® core i5 CPU with 2.5 GHz processor.

There are 29 non-bolus volumes out of 40 volumes. Here, non-bolus volumes means volumes in which bolus is not present in brain. For our experiments, we introduced 3D rotation to DSC-MRI volumes to simulate motion in transverse plane in the range $[-20^0\ 20^0]$ in random number of volumes. We first determine whether motion is present or not by using one dimensional time series extracted from central slices. If there is motion, we detect motion as explained in Section 4.5. Table 4.1 and Table 4.2 show the performance of our motion detection method with number of corrupted volumes as 5, 10 and 20, 25 respectively. These corrupted volumes are chosen randomly and they are not always consecutive volumes. This randomness reflects the worst possible scenarios during scanning. Here, we considered only rotation because translation inside the scanner is almost impossible due to the structure of MRI scanner. A specific case is shown in Fig. 4.8 where central slices of 15 volumes (3, 4, 5, 6,

Table 4.1 Evaluation of Proposed Motion Detection Method.

# Total Volumes	40			
# Corrupted Volumes	5		10	
Simulated Rotation	# Detected Volumes	Time Taken (in sec)	# Detected Volumes	Time Taken (in sec)
$[-1^0\ 1^0]$	3	2.71	8	2.73
$[-5^0\ 5^0]$	5	2.73	10	2.742
$[-10^0\ 10^0]$	5	2.751	10	2.76
$[-15^0\ 15^0]$	5	2.77	10	2.78
$[-20^0\ 20^0]$	5	2.79	10	2.83

Table 4.2 Evaluation of Proposed Motion Detection Method.

# Total Volumes	40			
# Corrupted Volumes	20		25	
Simulated Rotation	# Detected Volumes	Time Taken (in sec)	# Detected Volumes	Time Taken (in sec)
$[-1^0\ 1^0]$	11	2.74	13	2.75
$[-5^0\ 5^0]$	20	2.77	25	2.91
$[-10^0\ 10^0]$	20	2.78	25	2.93
$[-15^0\ 15^0]$	20	3.12	25	3.14
$[-20^0\ 20^0]$	20	3.17	25	3.45

7, 12, 13, 14, 15, 27, 28, 29, 30, 31 and 32) are corrupted by a rotation of -15^0. In general, different amount of rotation can be possible at different sets of consecutive volumes according to patient's typical movements. For example, as shown in Fig. 4.8, 3^{rd}, 4^{th}, 5^{th}, 6^{th} and 7^{th} volumes can be corrupted by a rotation of -15^0 while 12^{th}, 13^{th}, 14^{th}, 15^{th}, 27^{th}, 28^{th}, 29^{th}, 30^{th}, 31^{st} and 32^{nd} volumes can be corrupted by a different amount of rotation other than -15^0. We have experimented with many such scenarios also and we are still able to achieve similar performance. It can be observed from Table 4.1 and Table 4.2 that except for the range $[-1^0\ 1^0]$, our method is able to detect all corrupted volumes correctly. Even in practical cases, there is less probability that patient can move only 1^0. In case of 25 corrupted volumes, we are able to detect all 25 for $[-10^0\ 10^0]$, $[-15^0\ 15^0]$ and $[-20^0\ 20^0]$, while in [26], 21, 24 and 22 volumes are detected for respective motions (See Table 4.3). Here, slice resolution 128×128 and block size 32×32 are used in [26] to get more accuracy. Average of time taken for all experiments for each case are shown in Table 4.1 and Table 4.2. Time taken to detect motion is from 7.68 to 132.21 seconds (depending on block size) in [55]. In [55], time taken is reported only for phase correlation. Thus, it does not include the computation time for intensity correction of transit set. Time taken for detecting motion using proposed method is around 3 seconds (see Table 4.4). This reduction in time is due to the facts that proposed method detects motion (i) without explicit bolus handling, thus no intensity correction and, (ii) using one dimensional time series instead of two dimensional images. Hence, proposed method outperforms in terms of detection accuracy and computation time.

Table 4.3 Evaluation of Proposed Motion Detection Method in terms of Detection Accuracy. Here, slice resolution 128×128 and block size 32×32 are used in [26] to get more accuracy.

# Total Volumes	# Corrupted Volumes	Simulated Rotation	# Volumes Detected	
			[26]	Proposed
40	25	$[-1^0\ 1^0]$	NA	13
40	25	$[-5^0\ 5^0]$	NA	25
40	25	$[-10^0\ 10^0]$	21	25
40	25	$[-15^0\ 15^0]$	24	25
40	25	$[-20^0\ 20^0]$	22	25

Table 4.4 Evaluation of Proposed Motion Detection Method in terms of Time Taken. Here, time taken for [55] is shown according to slice resolution and block size. Minimum time taken is 7.68 sec with slice resolution 32×32 and block size 8×8, while maximum time taken is 132.21 sec with slice resolution 128×128 and block size 32×32. This time taken is reported only for phase correlation. Thus, it does not include the computation time for intensity correction of transit set.

# Total Volumes	# Corrupted Volumes	Simulated Rotation	Time Taken (sec)	
			[55]	Proposed
40	25	$[-1^0\ 1^0]$	7.68-132.21	2.75
40	25	$[-5^0\ 5^0]$	7.68-132.21	2.91
40	25	$[-10^0\ 10^0]$	7.68-132.21	2.93
40	25	$[-15^0\ 15^0]$	7.68-132.21	3.14
40	25	$[-20^0\ 20^0]$	7.68-132.21	3.45

4.7 Summary

We presented an efficient method to detect motion in perfusion weighted MRI series using time-frequency analysis. Since proposed method detects motion considering one dimensional time series instead of whole 3D volumes, it takes very less time. We used time-frequency analysis called Stockwell transform to extract motion information from one dimensional time series. We provided comparison between proposed method and the existing methods. Proposed method has the advantages: (i) No explicit bolus handling, thus no intensity correction is required, (ii) No use of phase correlation, thus no compromise between block size and speed, (iii) Detection accuracy is more and, (iv) Computational efficiency.

Algorithm 2 Presence of Motion.

Input: Two time series, $h_1(t)$ and $h_2(t)$, where $t = 1 \rightarrow N-1$

1. *Pre-processing:*

   - $M_1 = (max(h_1(t))/2$
   - $M_2 = (max(h_2(t))/2$
   - **for** $t = 1$ to $N-1$ **do**
     if $h_1(t) < M_1$ **then**
     $h_1(t) = 0$
     end if
     if $h_2(t) < M_2$ **then**
     $h_2(t) = 0$
     end if
     end for

2. *Find strong disturbances in both $h_1(t)$ and $h_2(t)$ and they are located at $\{n_i\}_{i=p_1}^{p_P}$*

3. *Determine presence of motion:*

   - $flag = 0$
   - **for** $i = p_1$ to p_P **do**
     if $h_1(n_i) > h_2(n_i)$ and $h_2(n_i) = 0$ **then**
     Disturbance is not due to motion but due to noise.
     else if $h_2(n_i) > h_1(n_i)$ and $h_1(n_i) = 0$ **then**
     Disturbance is not due to motion due to presence of bolus.
     else if $h_1(n_i) \geq h_2(n_i)$ and $h_2(n_i) > 0$ **then**
     Disturbance is due to motion.
     $flag \leftarrow flag + 1$
     end if
     end for
   - **if** $flag > 0$ **then**
     Motion is present.
     else
     Motion is absent.
     end if

Output: Information about presence of motion.

Chapter 5

Conclusions and Future Work

5.1 Summary and Conclusion

In this thesis, we addressed two different problems in motion analysis: (i) Small motion magnification in videos and (ii) Motion detection in perfusion weighted MRI. For analyzing motion in both problems, we used time frequency representation called Stockwell transform.

We proposed semi automated magnification of small motions in videos. Main contribution involves estimation of parameters, namely, bandwidth parameters and magnification parameters. In contrast to earlier methods where parameters are specified by user, we estimated the parameters using Stockwell transform. Video is a collection of time series. A few of these time series are extracted from videos at landmark pixels which represent the pixels that undergo motion. Parameters are estimated from these one dimensional time series. These parameters are incorporated to reconstruct videos with magnified motions. We demonstrated the proposed method on a few videos. However it is observed that noise is introduced in some reconstructed videos.

We have proposed a novel automated approach for motion detection in DSC-MRI perfusion data using time-frequency analysis. Instead of considering all three dimensional volumes or two dimensional images for the process, we used one dimensional time series due to the fact that these scans are acquired over time. For this, three dimensional PWI volume sequences are converted into one dimensional time series. It is observed that central slice of particular volume represents the motion present in the corresponding volume. Hence, all central slices of volumes are considered to form a one dimensional time series. Further, this time series is analyzed to detect motion. This made the proposed method computationally inexpensive. We have demonstrated that motion detection can be performed in automated fashion by using Stockwell transform. We demonstrated our method using 40 volumes of perfusion MRI sequences.

5.2 Future Work

We have presented a semi-automated method for magnifying small motions in videos. We have demonstrated that the parameters required to magnify the motion can be generated in an automated fashion by using the time frequency representation called Stockwell transform. However, it is observed that noise is introduced in some reconstructed videos. Therefore, this work can be extended to reduce this noise. Another future direction of this work is to make this process fully automatic by estimating temporal filters. We believe that this method can have potential applications in medical imaging

We have proposed a novel automated approach for motion detection in DSC-MRI perfusion data using time-frequency analysis. This method detects motion in transverse plane only. Therefore, there is a large scope to detect motion in all other possible directions.

Related Publications

- **Sushma M**, Anubha Gupta and Jayanthi Sivaswamy, **Time-Frequency Analysis based Motion Detection in Perfusion Weighted MRI**, *In Proc. of 4^{th} National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Jodhpur, India, December, 2013*.

- **Sushma M**, Anubha Gupta and Jayanthi Sivaswamy, **Semi-Automated Magnification of Small Motions in Videos**, *In Proc. of 5^{th} International Conf. on Pattern Recognition and Machine Intelligence (PReMI), Kolkata, India, December, 2013*.

Bibliography

[1] L. Ai, J. Nan, and L. hua Pu. Application of S-transform in EEG analysis. In *Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on*, volume 2, pages V2–453. IEEE, 2010. 13

[2] J. B. Allen. Short time spectral analysis, synthesis, and modification by discrete Fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing (ASSP)*, 25(3):235238, 1977. 6

[3] O. Arikan and D. A. Forsyth. Interactive motion generation from examples. In *ACM Transactions on Graphics (TOG)*, volume 21, pages 483–490, 2002. 13, 18

[4] G. Badrinath and P. Gupta. Stockwell transform based palm-print recognition. *Applied Soft Computing*, 11(7):4267–4281, 2011. 12

[5] E. L. Barbier, L. Lamalle, and M. Décorps. Methodology of brain perfusion imaging. *Journal of Magnetic Resonance Imaging*, 13(4):496–520, 2001. 15

[6] J. Bican and J. Flusser. 3D Rigid registration by cylindrical phase correlation method. *Pattern Recognition Letters*, 30:914–921, 2009. 16

[7] B. Boashash. Note on the use of the Wigner distribution for time-frequency signal analysis. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(9):1518–1521, 1988. 8

[8] B. Boashash. *Time frequency analysis*. Access Online via Elsevier, 2003. 1

[9] R. N. Bracewell and R. Bracewell. *The Fourier transform and its applications*, volume 31999. McGraw-Hill New York, 1986. 5

[10] G. A. Buonaccorsi, J. P. O'Connor, A. Caunce, C. Roberts, S. Cheung, Y. Watson, K. Davies, L. Hope, A. Jackson, G. C. Jayson, et al. Tracer kinetic model–driven registration for dynamic contrast-enhanced MRI time-series data. *Magnetic Resonance in Medicine*, 58(5):1010–1019, 2007. 16, 32

[11] P. Burt and E. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983. 19, 28

[12] S. Cha, E. A. Knopp, G. Johnson, S. G. Wetzel, A. W. Litt, and D. Zagzag. Intracranial mass lesions: Dynamic contrast-enhanced susceptibility-weighted echo-planar perfusion MR imaging. *Radiology*, 223(1):11–29, 2002. 14, 15

[13] A. Chan and S. J. Nelson. Simplified gamma-variate fitting of perfusion curves. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1067–1070, 2004. 33

[14] L. Cohen. *Time-frequency analysis*, volume 778. Prentice Hall PTR New Jersey, 1995. 7

[15] P. Dash, B. Panigrahi, and G. Panda. Power quality analysis using S-transform. *IEEE Transactions on Power Delivery*, 18(2):406–411, 2003. 12

[16] P. Dastidar. Overview of neuroradiological MRI. *International Journal of Bioelectromagnetism*, 1(1), 1999. 14

[17] R. Ditommaso, M. Mucciarelli, and F. C. Ponzo. S-transform based filter applied to the analysis of non-linear dynamic behaviour of soil and buildings. In *14th European conference on earthquake engineering. Proceedings volume. Ohrid, Republic of Macedonia*, volume 30, 2010. 13

[18] R. Ditommaso, M. Mucciarelli, and F. C. Ponzo. Analysis of non-stationary structural systems by using a band-variable filter. *Bulletin of Earthquake Engineering*, 10(3):895–911, 2012. 13

[19] S. Drabycz, R. G. Stockwell, and J. R. Mitchell. Image texture characterization using the discrete orthonormal S-transform. *Journal of digital imaging*, 22(6):696–708, 2009. 12

[20] R. Edelman, H. Mattle, D. Atkinson, T. Hill, J. Finn, C. Mayman, M. Ronthal, H. Hoogewoud, and J. Kleefield. Cerebral blood flow: assessment with dynamic contrast-enhanced T2*-weighted MR imaging at 1.5 T. *Radiology*, 176(1):211–220, 1990. 15

[21] A. El-Baz, G. Gimelfarb, and M. A. El-Ghar. New motion correction models for automatic identification of renal transplant rejection. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 235–243. 2007. 16

[22] J. L. Evelhoch. Key factors in the acquisition of contrast kinetic data for oncology. *Journal of Magnetic Resonance Imaging*, 10(3):254–259, 1999. 15, 16

[23] J. B. J. Fourier. *Thorie Analytique de la Chaleur (The Analytical Theory of Heat)*. 1822. 5

[24] M. Fuchs, T. Chen, O. Wang, R. Raskar, H.-P. Seidel, and H. Lensch. Real-time temporal shaping of high-speed video streams. *Computers & Graphics*, 34(5):575–584, 2010. 13

[25] D. Gabor. Theory of communication. *Journal of the Institution of Electrical Engineers*, 94(73):429457, 1946. 8

[26] R. Gautam, J. Sivaswamy, and R. Varma. An efficient, bolus-stage based method for motion correction in perfusion weighted MRI. In *International Conference on Pattern Recognition (ICPR)*, pages 145–148, 2012. xi, 4, 17, 32, 34, 44, 45

[27] N. V. George. *S Transform: Time Frequency Analysis & Filtering*. PhD thesis, 2009. 6

[28] B. G. Goodyear, H. Zhu, R. A. Brown, and J. R. Mitchell. Removal of phase artifacts from fMRI data using a Stockwell transform filter improves brain activity detection. *Magnetic Resonance in Medicine*, 51(1):16–21, 2004. 12, 35

[29] A. Grossmann and J. Morlet. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal of Mathematical Analysis*, 15(4):723–736, 1984. 9

[30] Z. Hao, H. Xu, G. Zheng, and G. Jing. Study on the time-frequency characteristics of engine induction noise in acceleration based on S transform. In *Image and Signal Processing, 2008. CISP'08. Congress on*, volume 5, pages 242–246. IEEE, 2008. 13

[31] H. A. Haroon, D. L. Buckley, T. A. Patankar, G. R. Dow, S. A. Rutherford, D. Baleriaux, and A. Jackson. A comparison of ktrans measurements obtained with conventional and first pass pharmacokinetic models in human gliomas. *Journal of magnetic resonance imaging*, 19(5):527–536, 2004. 15

[32] M. Jenkinson, P. Bannister, M. Brady, and S. Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841, 2002. 16, 32

[33] A. Kassner, D. Annesley, X. Zhu, K. Li, I. Kamaly-Asl, Y. Watson, and A. Jackson. Abnormalities of the contrast re-circulation phase in cerebral tumors demonstrated using dynamic susceptibility contrast-enhanced imaging: A possible marker of vascular tortuosity. *Journal of Magnetic Resonance Imaging*, 11(2):103–113, 2000. 15

[34] S. M. Kealey, V. A. Loving, D. M. Delong, and J. D. Eastwood. User-defined vascular input function curves: Influence on mean perfusion parameter values and signal-to-noise ratio1. *Radiology*, 231(2):587–593, 2004. 16

[35] M. Kim, G. Wu, and D. Shen. Groupwise registration of breast DCE-MR images for accurate tumor measurement. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 598–601, 2011. 16

[36] R. K. Kosior, J. C. Kosior, and R. Frayne. Improved dynamic susceptibility contrast (DSC)-MR perfusion estimates by motion correction. *Journal of Magnetic Resonance Imaging*, 26(4):1167–1172, 2007. 16

[37] J. Ladan and E. R. Vrscay. The Discrete Orthonormal Stockwell Transform and Variations, with Applications to Image Compression. In *Image Analysis and Recognition*, pages 235–244. 2013. 12

[38] A. Lausch, M. Ebrahimi, and A. Martel. Image registration for abdominal dynamic contrast-enhanced magnetic resonance images. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 561–565, 2011. 16

[39] J. Lee, J. Chai, P. S. Reitsma, J. K. Hodgins, and N. S. Pollard. Interactive control of avatars animated with human motion data. In *ACM Transactions on Graphics (TOG)*, volume 21, pages 491–500, 2002. 13, 18

[40] K. Li, X. Zhu, D. Checkley, J. Tessier, V. Hillier, J. Waterton, and A. Jackson. Simultaneous mapping of blood volume and endothelial permeability surface area product in gliomas using iterative analysis of first-pass dynamic contrast enhanced MRI data. *British Journal of Radiology*, 76(901):39–51, 2003. 15

[41] K.-L. Li, X. P. Zhu, J. Waterton, and A. Jackson. Improved 3D quantitative mapping of blood volume and endothelial permeability in brain tumors. *Journal of Magnetic Resonance Imaging*, 12(2):347–357, 2000. 15

[42] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson. Motion magnification. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 519–526, 2005. 3, 13, 18

[43] M. T. Madsen. A simplified formulation of the gamma variate function. *Physics in Medicine and Biology*, 37(7):1597–1600, 1992. 33

[44] A. Medda, J. Billings, and S. Keilholz. fMRI functional cluster analysis using the Stockwell transform. In *Biomedical Engineering Conference (SBEC), 2013 29th Southern*, pages 97–98. IEEE, 2013. 12

[45] A. Melbourne, D. Atkinson, M. White, D. Collins, M. Leach, and D. Hawkes. Registration of dynamic contrast-enhanced MRI using a progressive principal component registration (PPCR). *Physics in medicine and biology*, 52(17):5147, 2007. 16

[46] J. R. Mitchell, T. C. Fong, B. G. Goodyear, H. Zhu, et al. Filtering artifact from fMRI data using the Stockwell transform, Mar. 10 2009. US Patent 7,502,526. 12

[47] M. Mucciarelli, M. Bianca, R. Ditommaso, M. Gallipoli, A. Masi, C. Milkereit, S. Parolai, M. Picozzi, and M. Vona. Far field damage on RC buildings: the case study of navelli during the L'Aquila (Italy) seismic sequence, 2009. *Bulletin of Earthquake Engineering*, 9(1):263–283, 2011. 13

[48] J. R. Petrella and J. M. Provenzale. MR perfusion imaging of the brain: techniques and applications. *American Journal of Roentgenology*, 175(1):207–219, 2000. 14, 31

[49] C. Pinnegar. Polarization analysis and polarization filtering of three-component signals with the time–frequency S transform. *Geophysical Journal International*, 165(2):596–606, 2006. 12

[50] C. R. Pinnegar, H. Khosravani, and P. Federico. Time–frequency phase analysis of ictal EEG recordings with the S-transform. *Biomedical Engineering, IEEE Transactions on*, 56(11):2583–2593, 2009. 13

[51] K.-K. Poh and P. Marziliano. Analysis of neonatal EEG signals using Stockwell transform. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 594–597. IEEE, 2007. 13

[52] M.-Z. Poh, D. J. McDuff, and R. W. Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10):10762–10774, 2010. 13, 26

[53] J. M. Provenzale, S. Mukundan, and D. P. Barboriak. Diffusion-weighted and Perfusion MR Imaging for Brain Tumor Characterization and Assessment of Treatment Response1. *Radiology*, 239(3):632–649, 2006. 14

[54] J. M. Provenzale, G. R. Wang, T. Brenner, J. R. Petrella, and A. G. Sorensen. Comparison of permeability in high-grade and low-grade brain tumors using dynamic susceptibility contrast MR imaging. *American Journal of Roentgenology*, 178(3):711–716, 2002. 15

[55] R. Gautam, J. Sivaswamy, and R. Varma. A method for motion detection and categorization in perfusion weighted MRI. In *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*, pages 6:1–6:8, 2012. x, xi, 16, 17, 32, 33, 44, 45

[56] G. Rohde, A. Barnett, P. Basser, S. Marenco, and C. Pierpaoli. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 51(1):103–114, 2004. 16

[57] M. Rubinstein. *Analysis and Visualization of Temporal Variations in Video*. PhD thesis, Massachusetts Institute of Technology, Feb 2014. 19

[58] G. Rutkowski, K. Patan, and P. Leśniak. Comparison of time-frequency feature extraction methods for EEG signals classification. In *Artificial Intelligence and Soft Computing*, pages 320–329, 2013. 13

[59] D. Shanbhag, R. Mullick, S. Nath, C. Oppenheim, M. Luby, K. Ku, L. Latour, S. Warach, and N. Natural. Impact of motion and symmetry correction on perfusion lesion segmentation in acute ischemic stroke: Quantitative evaluation. *Brain*, 17:4710–4710, 2009. 16

[60] Z. Shi, L. Ruirui, W. Qun, J. Heptol, and Y. Guimin. The research of power quality analysis based on improved S-transform. In *Electronic Measurement & Instruments, 2009. ICEMI'09. 9th International Conference on*, pages 2–477. IEEE, 2009. 12

[61] T. R. Steger and E. F. Jackson. Real-time motion detection of functional MRI data. *Journal of Applied Clinical Medical Physics*, 5(2):64–70, 2004. 16, 32

[62] R. G. Stockwell. S-transform analysis of gravity wave activity from a small scale network of airglow imagers. 1999. 12

[63] R. G. Stockwell, L. Mansinha, and R. P. Lowe. Localization of the complex spectrum: the S transform. *IEEE Transactions on Signal Processing*, 44(4):998–1001, 1996. 9, 25, 35

[64] M. Straka, G. W. Albers, and R. Bammer. Real-time diffusion-perfusion mismatch analysis in acute stroke. *Journal of Magnetic Resonance Imaging*, 32(5):1024–1037, 2010. 16

[65] P. S. Tofts and A. G. Kermode. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. fundamental concepts. *Magnetic Resonance in Medicine*, 17(2):357–367, 1991. 15

[66] J. Tokuda, H. Mamata, R. R. Gill, N. Hata, R. Kikinis, R. F. Padera, R. E. Lenkinski, D. J. Sugarbaker, and H. Hatabu. Impact of nonrigid motion correction technique on pixel-wise pharmacokinetic analysis of free-breathing pulmonary dynamic contrast-enhanced MR imaging. *Journal of Magnetic Resonance Imaging*, 33(4):968–973, 2011. 16

[67] J. Ville. Thorie et applications de la notion de signal analytique. *Cbles et Transmission*, 2(A):61–74, 1948. 8

[68] J. Wang, S. M. Drucker, M. Agrawala, and M. F. Cohen. The cartoon animation filter. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 1169–1173, 2006. 13

[69] Y. Wang. The tutorial: S transform. 12

[70] R. Weisskoff, J. Boxerman, A. Sorensen, S. Kulke, T. Campbell, and B. Rosen. Simultaneous blood volume and permeability mapping using a single Gd-based contrast injection. *International Society for Magnetic Resonance in Medicine*, 279, 1994. 15

[71] S. G. Wetzel, S. Cha, G. Johnson, P. Lee, M. Law, D. L. Kasow, S. D. Pierce, and X. Xue. Relative cerebral blood volume measurements in intracranial mass lesions: Interobserver and intraobserver reproducibility study. *Radiology*, 224(3):797–803, 2002. 15

[72] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40(5):749–759, 1932. 8

[73] J. C. Wong, J. M. Provenzale, and J. R. Petrella. Perfusion MR imaging of brain neoplasms. *American Journal of Roentgenology*, 174(4):1147–1157, 2000. 15

[74] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. http://people.csail.mit.edu/mrub/vidmag/. Accessed: 2013-11-20. 28

[75] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (TOG)*, 31(4):65:1–65:8, 2012. ix, xi, 3, 13, 18, 19, 20, 23, 28, 29, 30

[76] Y. Zhang. MRI texture analysis in multiple sclerosis. *International Journal of Biomedical Imaging*, 2012:2:2–2:2, 2012. 13, 35

[77] H. Zhu, B. Goodyear, M. Lauzon, R. Brown, G. Mayer, A. Law, L. Mansinha, and J. Mitchell. A new local multiscale Fourier analysis for medical imaging. *Medical physics*, 30:1134–41, 2003. 13, 35