NOISE REDUCTION IN TIME-FREQUENCY DOMAIN

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÖZDEN KALYONCU

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

SEPTEMBER 2007

Approval of the thesis:

NOISE REDUCTION IN TIME-FREQUENCY DOMAIN

submitted by **ÖZDEN KALYONCU** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. İsmet Erkmen _____
Head of Department, **Electrical and Electronics Engineering**

Prof. Dr. Zafer Ünver _____
Supervisor, **Electrical and Electronics Engineering Dept., METU**

Examining Committee Members:

Prof. Dr. Kemal Leblebicioğlu _____
Electrical and Electronics Engineering Dept., METU

Prof. Dr. Zafer Ünver _____
Electrical and Electronics Engineering Dept., METU

Prof. Dr. Engin Tuncer _____
Electrical and Electronics Engineering Dept., METU

Assoc. Prof. Dr. Tolga Çiloğlu _____
Electrical and Electronics Engineering Dept., METU

Dr. Toygar Birinci _____
ASELSAN Inc.

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Özden Kalyoncu

Signature :

ABSTRACT

NOISE REDUCTION IN TIME-FREQUENCY DOMAIN

Kalyoncu, Özden

M. S., Department of Electrical and Electronics Engineering
Supervisor : Prof. Dr. Zafer Ünver

September 2007, 113 pages

In this thesis work, time-frequency filtering of nonstationary signals in noise using Wigner-Ville Distribution is investigated. Continuous-time, discrete-time and discrete Wigner Ville Distribution definitions, their relations, and properties are given.

Time-Frequency Peak Filtering Method is presented. The effects of different parameters on the performance of the method are investigated, and the results are presented.

Time-Varying Wiener Filter is presented. Using simulations it is shown that the performance of the filter is good at SNR levels down to -5 dB. It is proposed and shown that the performance of the filter improves by using Support Vector Machines.

The presented time-frequency filtering techniques are applied on test signals and on a real world signal. The results obtained by the two methods and also by classical zero-phase low-pass filtering are compared. It is observed that for low sampling

rates Time-Varying Wiener Filter, and for high sampling rates Time-Frequency Peak Filter performs better.

Keywords : Wigner-Ville Distribution, time-frequency filtering, Support Vector Machines

ÖZ

ZAMAN-FREKANS BÖLGESİNDE GÜRÜLTÜ AZALTIMI

Kalyoncu, Özden

Yüksek Lisans, Elektrik Elektronik Mühendisliği Bölümü
Tez Yöneticisi : Prof. Dr. Zafer Ünver

Eylül 2007, 113 sayfa

Bu tez çalışmasında gürültü içindeki durağan olmayan sinyallerin Wigner-Ville Dağılımı kullanılarak zaman-frekans süzgeçlenmesi araştırılmıştır. Sürekli zaman, ayrık zaman ve ayrık Wigner-Ville Dağılımı tanımları, ilişkileri ve özellikleri verilmiştir.

Zaman-Frekans Tepe Süzgeçleme Yöntemi tanıtılmıştır. Değişik parametrelerin yöntemin başarımı üzerindeki etkileri araştırılmış ve sonuçlar sunulmuştur.

Zamanla Değişen Wiener Süzgeç tanıtılmıştır. Yapılan benzetimlerle süzgeç başarımının -5 dB SNR seviyesine kadar iyi olduğu gösterilmiştir. Destek Vektör Makinalarının kullanımının süzgecin başarımını iyileştireceği önerilmiş ve gösterilmiştir.

Sunulan zaman-frekans süzgeçleme yöntemleri test sinyalleri ve gerçek bir sinyal örneği üzerine uygulanmıştır. İki yöntem tarafından ve klasik fazı sıfır alçak geçiren süzgeçleme ile elde edilen sonuçlar karşılaştırılmıştır. Düşük örnekleme hızlarında

Zamanla Değişen Wiener Süzgecinin, yüksek örnekleme hızlarında ise Zaman-Frekans Tepe Süzgeçlemenin daha başarılı olduğu gözlenmiştir.

Anahtar Kelimeler : Wigner-Ville Dağılımı, zaman-frekans süzgeçleme, Destek Vektör Makinaları

To My Family

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Prof. Dr. Zafer Ünver for his guidance, advice, criticism, encouragement, endless patience and insight throughout the completion of the thesis.

I am indebted to all of my friends and colleagues for their support and encouragements. I am also grateful to ASELSAN Inc. for the facilities that made my work easier.

My family, no words can help me to express my feelings, but at least I can say that I am grateful to you for the life you provide to me.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Nonstationary signals have great importance in our daily life, since they occur naturally in many real world events. In biomedical applications signals under concern are usually nonstationary, like brain and heart signals. In radar applications nonstationary signals are used in order to increase radar resolution and decrease probability of intercept by other systems like jammers, [35].

The analysis of these kinds of signals is an important topic in signal processing. In particular, one wants to get information about the variation of the spectral content of the analyzed signal with respect to time. Traditional analysis methods like Fourier Transform (FT) based methods are not sufficient tools; because they do not give any feeling about the change of the spectral content of the signal with time. Hence, different analysis methods are needed.

In literature there are time-frequency analysis tools developed to analyze nonstationary signals. These tools are actually transformations which transform signal under analysis to a 2D time-frequency plane. In other words, they make an assignment of spectral components to time. These analysis tools can be grouped basically into two as linear and quadratic ones [34].

Linear Time Frequency Analysis Methods usually use an analysis window which moves in time. The signal to be analyzed is multiplied with the analysis window and the FT of the product is computed. Since, the analysis window suppresses

the signal outside of a certain region in time, the FT gives the local spectrum. The most known linear time frequency analysis method is the Short Time Fourier Transform (STFT) [37]. When the analysis window is chosen to be a Gaussian window, the transformation is known as Gabor Transform, because Gabor introduced STFT with this particular analysis window [38]. Even though the method and the computations are simple, the main draw back of Linear Time Frequency Analysis methods is the dependency of time-frequency resolution on the type and length of the analysis window [37]. The uncertainty principle states that time and frequency localizations are inversely proportional. If time localization is improved then frequency localization is degraded and vice versa. The window which gives the best frequency localization for a given time localization is the Gaussian window which is used by the Gabor Transform.

Cohen introduced a general class of quadratic time-frequency distributions, where the time dependent auto-correlation function of the signal to be analyzed is multiplied with a kernel, and the FT of the product is computed [34]. The kernel is a function of time and frequency lags instead of time and frequency, so the transformation is shift invariant. Quadratic Time Frequency Distributions do not suffer from uncertainty principle; their time-frequency resolution is not limited. But, on the other hand, they suffer from the cross terms [37]. Being quadratic transformations the time-frequency transformation of sum of two signals is not the sum of the time-frequency transformations of the individual signals, but there is a third term known as the cross-term. The cross-term's location is between the two auto terms in time and frequency. Also cross-term contains amplitude modulation whose modulation frequency increases when the distance between auto-terms in time or frequency increases. For a signal containing N auto-terms, there will be N*(N-1)/2 cross terms, that is the number of cross-terms increases quadratically. There are different methods to eliminate these cross-terms. Cohen used the kernel function that actually serves as a lowpass filter. It is successful in eliminating cross-terms to some degree; its main drawback is that it decreases the time-frequency resolution of the transformation. When this kernel is set to unity then there is no

filtering and the resulting distribution is known as Wigner-Ville Distribution (WVD).

Filtering of nonstationary signals in noise has been an important topic. Traditional filtering methods are used for filtering of nonstationary signals in noise, like adaptive and fixed filtering methods [36]. Although adaptive methods, like Least Mean Squares (LMS) approach based methods and Kalman Filter, are generally superior compared to fixed methods, they perform poor for nonstationary signals whose spectral content changes quickly with time. Additionally, adaptive methods need signal modeling for optimum performance; but when the structure of the desired signal is unknown, they will give suboptimal results and even in some cases wrong results. Since the spectral content of these signals changes with time, the information about the spectral variation with time which can be obtained by time-frequency analysis can be used in order to improve the filtering performance.

WVD was not used in practical applications beforehand because of its high computational cost. Nowadays, with the use of powerful computers WVD can be used in different signal processing areas. One of these application areas is the time-varying filtering of nonstationary signals in noise [39]. A time-varying filter is the one whose frequency response changes with time. For signals whose spectral content changes with time, a time-varying filter whose frequency response variation adapted to the variation of the desired signal performs better as compared to the time-invariant filter whose frequency response is fixed.

Time-varying filtering is also known as time-frequency filtering. When the signal under concern is nonstationary, it is advantageous to transform the noisy signal to joint time-frequency plane by using an appropriate time-frequency transformation. By doing so, the information about the variation of signal spectrum with time can be obtained. The noise spreads to entire time-frequency plane; whereas the signal part concentrates on certain time-frequency regions. The filtering is performed using the signal concentrated time-frequency regions. The filtering performance depends on how good the signal is localized to certain time-frequency regions and

how good these regions are detected. Time-frequency localization depends on the type of the time-frequency analysis method. The WVD is preferred in this application instead of linear time-frequency analysis methods because of its high time-frequency localization as compared to linear ones. (Linear time-frequency transformations have some advantages such as low computational cost and easiness to recover the signal [39].)

This thesis mainly concentrates on time-frequency filtering of signals whose spectral content quickly changes with time using WVD. To start with, the WVD and its properties are investigated; the pseudo and robust forms, and discrete formulations of the distribution are studied. Next, computation of the WVD of a signal from noisy observations and estimation of the instantaneous frequency (IF) of the signal from the estimated WVD are studied. Two different time-frequency filtering methods, Time-Frequency Peak Filtering (TFPF) Method and Time-Varying Wiener Filter (TVWF), are investigated. Both methods compute the WVD of the noisy signal and estimate the IF of the signal as main filtering steps. The TFPF method first encodes the noisy signal as an analytic frequency modulated signal and then performs the filtering steps, consequently the estimated IF is actually the signal estimate multiplied with a scalar. The TVWF uses the IF estimate to form a time-frequency mask and uses this mask to recover the desired signal. The filtering performances of the methods are compared with each other and with a classical zero-phase low-pass filter using frequency modulated test signals and real life data. The performance of the second method is improved by using Support Vector Machines (SVM) Method applied in IF estimation phase of the algorithm, which is actually a pattern recognition technique based on statistical learning theory. The filtering algorithms are implemented and tested in MATLAB® environment.

The thesis is organized in six chapters. Chapter 2 deals with the continuous WVD by providing the necessary background information and with the discretization of WVD by giving discrete-time and discrete WVD definitions and providing the

relations among them. In Chapter 3, the TFPF Method is described by introducing the basic concepts and the discrete time algorithm. In addition, the performance of the algorithm is tested. In Chapter 4, the second time-frequency filtering method is described by introducing the derivations of the TVWF and its time-frequency formulation. Besides, the IF estimation of noisy signals using WVD is discussed and the improvement using SVM method is evaluated by comparing the simulation results. Chapter 5 presents and explains the simulation results in which the performances of the time-frequency filtering methods with each other and with a classical low-pass zero-phase filter are compared using different test signals and real-world data. Finally, Chapter 6 summarizes the thesis and presents the conclusions.

CHAPTER 2

WIGNER-VILLE DISTRIBUTION

The WVD was introduced by Wigner in the field of quantum mechanics. Then Ville brought the concept into the signal processing field. Claasen and Mecklenbrauker published three papers in which signal theoretic properties of the distribution were given, a proposal for discrete-time and a proposal for windowed form of the distribution were presented and the relation of the distribution with other time-frequency distributions were mentioned consecutively [28]. The importance of the WVD arises from the fact that it combines the temporal and the spectral analyses of a signal by transforming the signal to a joint time-frequency plane. But there is something important to mention before continuing: the WVD is known as a distribution because it is supposed to reflect the distribution of the signal energy in the time-frequency plane. However, the WVD can not be interpreted pointwise as a distribution of energy because it can also take on negative values. (The WVD of Gaussian and chirp signals are always positive.)

In the first section of this chapter continuous WVD and its properties are discussed; and in the second section the discussions about the discretization of WVD are presented.

2.1 CONTINUOUS WIGNER-VILLE DISTRIBUTION

The continuous WVD of a signal $x(t)$ is defined by

$$W_{xx}(t,\omega) \stackrel{\Delta}{=} \int x(t+\tau/2)\, x^*(t-\tau/2) \exp(-j\omega\tau)\, d\tau \tag{2-1}$$

where the integral is from minus infinity to plus infinity. The dual form is

$$W_{xx}(t,\omega) = 1/2\pi \int X(\omega - v/2)\, X^*(\omega + v/2) \exp(-jvt)\, dv \tag{2-2}$$

where $X(\omega)$ is the FT of $x(t)$.

Actually the WVD is the FT of the time-dependent auto-correlation of the signal $x(t)$ defined by $R_{xx}(t,\tau) = x(t+\tau/2)x^*(t-\tau/2)$. $R_{xx}(t,\tau)$ is also called a bilinear data product by different authors [7].

Some mathematical properties of WVD are given briefly in APPENDIX A. Even though the WVD has many desired properties, there is an important disadvantage of the WVD known as cross-terms which becomes visible in case of multi-component signals. Let $x(t)$ be the sum of N signals

$$x(t) = \sum_{k=1}^{N} x_k(t) \tag{2-3}$$

The WVD of $x(t)$ is

$$W_{xx}(t,\omega) = \sum_{k=1}^{N} W_{x_k x_k}(t,\omega) + \sum_{k=1}^{N} \sum_{d>k}^{N} W_{x_k x_d}(t,\omega) \tag{2-4}$$

where $W_{x_k x_d}(t,\omega)$ is a cross WVD and known as cross-term; so the WVD of sum of signals consists of the WVD of the individual signals plus the cross WVDs. As the number of auto-terms is increased, the time-frequency plane becomes more complex and interpretation of the distribution becomes more difficult, because cross-terms and auto-terms can be located in the same time-frequency regions.

Figure 2-1 shows the cross-term geometry of the WVD. It has been shown that while the cross term's envelope depends on the signal, the cross term's time–frequency location and oscillation geometry merely depend on the time–frequency locations of the interfering signal terms [22]. As it was mentioned above the cross-term is located in between the auto-terms in the time-frequency plane; that is to say $t_{12} = (t_1 + t_2)/2$ and $\omega_{12} = (\omega_1 + \omega_2)/2$; and the cross-term contains amplitude modulation whose modulation frequency increases when the distance between auto-terms in time or frequency increases. More information about cross-terms can be found in [23], [24].

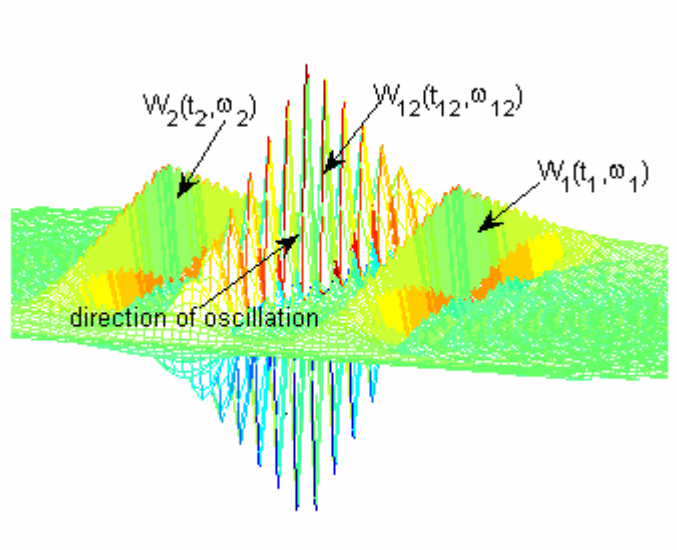

Figure 2-1 Cross-term geometry of the WVD.

Recall that the WVD can not be interpreted pointwise as an energy distribution because it can also take on negative values, which results in negative energy which is physically meaningless. Actually this negative energy arises from the cross-terms. Consider a two-component signal

$$x(t) = a_1 \exp(j\omega_1 t) + a_2 \exp(j\omega_2 t) \tag{2-5}$$

The WVD of this signal is

$$W_{xx}(t,\omega) = a_1^2\,\delta(\omega-\omega_1) + a_2^2\,\delta(\omega-\omega_2) + 2\,a_1 a_2\cos(t\,(\omega_1-\omega_2))\,\delta(\omega-\frac{\omega_1+\omega_2}{2}) \tag{2-6}$$

The first two terms are the auto-terms and the third term is the cross-term, a cosine wave which can take on negative values. However, the integral of this cross-term over frequency is zero, and it does not influence the total signal energy.

In literature there is an intensive research to eliminate or at least to reduce the effects of these cross-terms. The main approach is to use a kernel function which actually serves as a low-pass filter to reduce the cross-terms. The design of these kernels is not an easy topic, since the direction of oscillation and the oscillation frequency depend on the type and the time-frequency location of the input signals [22], [23], [24]. So, a general kernel design for all types of signals is a challenging problem. Additionally, when the signals are located close in time or frequency or time and frequency, the oscillation frequency decreases which means that lowpass filter passes these cross terms. Another drawback of using smoothing kernels is that they decrease the time-frequency resolution of the distribution by broadening the auto (signal) terms. Finally, sometimes, use of these kernels results in loss of some desired mathematical properties of the WVD. An interesting and different approach is proposed in [22] which uses a nonlinear median type filter to eliminate the cross-terms.

In literature the WVD of noisy signals is also intensively studied. It has been shown that the WVD of such signals can not be directly used as an estimate of the WVD of the noise free signal; this is so because the variance of WVD of a signal under additive white Gaussian noise goes to infinity, and a time window must be used to make the variance finite [2].

The windowed form of the WVD is known a pseudo WVD (pWVD) given as

$$W^{p}{}_{xx}(t,\omega) = \int h(\tau)\,x(t+\tau/2)\,x^{*}(t-\tau/2)\exp(-j\omega\tau)\,d\tau \tag{2-7}$$

where $h(\tau)$ is a real, to avoid frequency shifts, and symmetrical, to avoid time shifts, window. The pWVD corresponds approximately to the WVD of the signal. The relation between pWVD and WVD can be shown using (A-5) as

$$W^{p}{}_{xx}(t,\omega) = H(\omega) * W_{xx}(t,\omega) \tag{2-8}$$

where $H(\omega)$ is the FT of $h(t)$.

(2-8) indicates that the use of a window decreases the frequency resolution of the distribution.

In recent years robust time-frequency signal transforms are introduced which are concerned with the estimation of the WVD of a signal from noisy observations of the signal. It has been shown that the WVD of a signal in additive white Gaussian noise, computed directly, is actually the Maximum Likelihood (ML) estimate of the WVD of the noise free signal [6], [10]. The derivations are given in APPENDIX B. This means that for the signal in Gaussian noise the standard signal transformation will give the best estimate of the WVD of the noise free signal. Unfortunately, the ML estimators are quite sensitive to the variation of the noise probability density function, which means that for a non-Gaussian noise the standard signal transformations will give worse results as compared to the signal transformations obtained as the ML estimate for that noise type. This fact motivates the introduction of robust signal transformations which are introduced for a class of noises by computing the ML estimate of the transformation for the worst noise (the noise with the longest tail) in that class. The Laplacian noise is used with absolute error as the loss function, since it is the worst noise for numerous forms of impulsive noises.

Robust signal transformations will give worse results compared to the ML estimators obtained for the signals in additive Gaussian noise; on the other hand, the improvement for impulsive noises is significant as compared to the ML estimator for Gaussian noise.

Robust time-frequency transformations are introduced in order to work under impulsive noise conditions. Signals are often influenced by impulsive noise; especially this is the case for the WVD; because, the resultant noise, after the WVD is computed, is a mixture of impulsive and Gaussian noises even when the input noise is purely Gaussian noise. More information about robust time-frequency transformations can be found in [1]-[6], [10], [25], [26].

2.2 DISCRETE WIGNER-VILLE DISTRIBUTION

Several different approaches are proposed to discretize WVD to obtain alias free discrete WVD [15]-[21]. A summary of the works done to discretize WVD are given in APPENDIX C in a chronological order. In the first and second parts of this section the definitions of the discrete-time and the discrete WVD and their relations with the continuous-time WVD are given.

2.2.1 The Discrete-Time WVD

In this section the discrete-time WVD definition which is used in the thesis is given. The continuous-time WVD can be written in the following form with a change of variable

$$W_{x_c x_c}(t,\omega) = 2\int x_c(t+\tau)\, x_c^*(t-\tau)\exp(-j2\omega\tau)\, d\tau \tag{2-9}$$

Define

$$\phi_{x_t x_t}(\tau) \overset{\Delta}{=} x_c(t+\tau)\, x_c^*(t-\tau) \tag{2-10}$$

with FT $\Phi_{x_t x_t}(\omega)$ where the lower script t indicates the time variable; then

$$W_{x_c x_c}(t,\omega) = 2\,\Phi_{x_t x_t}(2\omega) \tag{2-11}$$

Assume that $\phi_{x_t x_t}(\tau)$ is sampled with sampling frequency $\omega_s = 2\pi/T$ yielding $\hat{\phi}_{x_t x_t}[k] \stackrel{\Delta}{=} x_c(t+kT)\, x_c^*(t-kT)$ with discrete-time FT $\hat{\Phi}_{x_t x_t}(\theta)$. From the derivations of the sampling theorem it follows that

$$\hat{\Phi}_{x_t x_t}(\omega T) = \frac{1}{T}\sum_r \Phi_{x_t x_t}(\omega + r\omega_s) \tag{2-12}$$

where the summation is from minus infinity to plus infinity.

If $x_c(t)$ is bandlimited to ω_N rad/sec, then $\phi_{x_t x_t}(\tau)$ is bandlimited to $2\omega_N$ rad/sec. Therefore, if $\omega_s \geq 4\omega_N$ aliasing is avoided and $\Phi_{x_t x_t}(\omega)$ can be obtained from $\hat{\Phi}_{x_t x_t}(\omega T)$, i.e., $\phi_{x_t x_t}(\tau)$ can be obtained from $\hat{\phi}_{x_t x_t}[k]$.

Let

$$x[n] \stackrel{\Delta}{=} x_c(nT) \tag{2-13}$$

then

$$\phi_{x_n x_n}[k] \stackrel{\Delta}{=} \hat{\phi}_{x_{nT} x_{nT}}[k] = x_c(nT+kT)\, x_c^*(nT-kT) = x_c[n+k]\, x_c^*[n-k] \tag{2-14}$$

The relation between the FTs of $\phi_{x_n x_n}[k]$ and $\phi_{x_{nT} x_{nT}}(\tau)$ is

$$\Phi_{x_n x_n}(\theta) = \hat{\Phi}_{x_{nT} x_{nT}}(\theta) = \frac{1}{T}\sum_r \Phi_{x_{nT} x_{nT}}\left(\frac{\theta}{T} + r\frac{2\pi}{T}\right) \tag{2-15}$$

In conclusion if $\omega_s \geq 4\omega_N$, $\Phi_{x_{nT} x_{nT}}\left(\frac{\theta}{T}\right)$ can be obtained from $\Phi_{x_n x_n}(\theta)$, i.e., $\phi_{x_{nT} x_{nT}}(\tau)$ can be obtained from $\phi_{x_n x_n}[k]$.

Let us now give the discrete-time WVD definition and show the relation with the continuous-time WVD. The discrete-time WVD of sequence $x[n]$ is defined as

$$W_{x_n x_n}[n,\theta) \stackrel{\Delta}{=} 2\sum_k x[n+k]x^*[n-k]\exp(-j2k\theta) \tag{2-16}$$

As seen from (2-16) the discrete-time WVD is periodic in θ with period π. To show the relation between the continuous-time and discrete-time definitions let us proceed as follows: since

$$W_{x_n x_n}[n,\theta) = 2\,\Phi_{x_n x_n}(2\theta) \tag{2-17}$$

using (2-15) we obtain

$$W_{x_n x_n}[n,\theta) = \frac{2}{T}\sum_r \Phi_{x_{nT} x_{nT}}\left(\frac{2\theta}{T} + r\frac{2\pi}{T}\right) \tag{2-18}$$

Finally by inserting (2-11) into (2-18) we obtain

$$W_{x_n x_n}[n,\theta) = \frac{1}{T}\sum_r W_{x_c x_c}\left(nT, \frac{\theta}{T} + r\frac{\pi}{T}\right) \tag{2-19}$$

As seen from (2-19) the discrete-time WVD is composed of the replicas of the continuous-time WVD. If $x_c(t)$ is bandlimited to ω_N rad/sec, $W_{x_c x_c}(nT,\omega)$ is bandlimited to ω_N rad/sec too; therefore, if $\pi/T \geq 2\omega_N$, aliasing is avoided and $W_{x_c x_c}(nT,\theta/T)$ can be obtained from $W_{x_n x_n}[n,\theta)$. This result also shows that the band limited continuous-time signal must be sampled at a rate at least two times the Nyquist rate, since the sampling frequency should be set to $\omega_s \geq 4\,\omega_N$ to avoid aliasing and make reconstruction possible.

2.2.2 The Discrete WVD

Consider a sequence $x[n]$ which is zero for $n < 0$ *and* $n \geq N$ where N is an even integer. $\phi_{x_n x_n}[k] = x[n+k]x^*[n-k]$ will be zero for $n < 0$ *and* $n \geq N$ and for $k < \max\{-n,\, n-N+1\}$ *and* $k > \min\{n,\, N-1-n\}$. Hence $\phi_{x_n x_n}[k]$ will be possibly

nonzero only for $0 \le n < N$ and $\max\{-n, n-(N-1)\} \le k \le \min\{n, -n+(N-1)\}$. For $n = N/2-1$ the maximum interval $-(N/2-1) \le k \le N/2-1$ is achieved. The discrete WVD of $x[n]$ is defined as

$$W_{x_n x_n}[n,m] = 2 \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \phi_{x_n x_n}[k] \exp(-jkm\frac{2\pi}{N}) \tag{2-20}$$

for $0 \le n < N$ and $0 \le m < N$. (In (2-20) the lower limit in the summation is decreased by 1, which has no effect.) Define a sequence $s_{x_n x_n}[p]$

$$s_{x_n x_n}[p] \stackrel{\Delta}{=} \phi_{x_n x_n}[k]\Big|_{k=p-\frac{N}{2}} = x[n-\frac{N}{2}+p]\, x^*[n+\frac{N}{2}-p] \tag{2-21}$$

The discrete WVD can be defined using $s_{x_n x_n}[p]$ as follows:

$$W_{x_n x_n}[n,m] = 2\sum_{p=0}^{N-1} s_{x_n x_n}[p] \exp(-j(p-\frac{N}{2})m\frac{2\pi}{N}) \tag{2-22}$$

(2-22) can be simplified to yield

$$W_{x_n x_n}[n,m] = 2(-1)^m \sum_{p=0}^{N-1} s_{x_n x_n}[p] \exp(-jpm\frac{2\pi}{N}) \tag{2-23}$$

The summation term in (2-23) is actually the discrete FT $S_{x_n x_n}[m]$ of $s_{x_n x_n}[p]$ which is equal to

$$S_{x_n x_n}[m] = \frac{1}{2}(-1)^m W_{x_n x_n}[n,m] \tag{2-24}$$

So, when $W_{x_n x_n}[n,m]$ is given, the sequence $s_{x_n x_n}[p]$ can be obtained from $S_{x_n x_n}[m]$ by inverse discrete FT; and the sequence $\phi_{x_n x_n}[k]$ can be obtained from $s_{x_n x_n}[p]$.

Using (2-20) and (2-16) it is obvious that

$$W_{x_n x_n}[n,m] = W_{x_n x_n}[n,\theta)\Big|_{\theta=\frac{m\pi}{N}} \tag{2-25}$$

(2-25) shows that the discrete WVD can be obtained from the discrete-time WVD by sampling the frequency in N discrete frequencies in the region $0 \le \theta < \pi$.

CHAPTER 3

TIME-FREQUENCY PEAK FILTERING

In many signal processing applications, like radar, sonar and biomedical applications, the signals under concern can not be always observed directly; moreover, they are tried to be obtained from noisy or distorted observations. There are a variety of approaches on the estimation of the desired signal which are mainly based on the minimization of the estimation error energy. However, for a nonstationary signal with rapidly varying frequency content, the estimation problem becomes more complex, and new filtering methods have to be investigated. For this kind of signals time-frequency filtering methods are proposed. Although there are different time-frequency filtering algorithms, the main filtering steps are the same and which can be summarized as follows: first a noisy signal is transformed to the joint time-frequency plane where the noise spreads to the entire plane and the signal is concentrated to certain time-frequency regions known as regions of support of the signal; secondly, the signal concentrated time-frequency regions are detected; and finally, the desired signal is synthesized using signal concentrated regions. The main differences between different time-frequency filtering algorithms are the type of the time-frequency distribution (TFD), and the detection and the synthesis methods used in the algorithms.

The straightforward time-frequency filtering algorithm is performed using a 2-D time-frequency mask which masks the noise concentrated regions and passes signal concentrated regions. If the region of support of the signal is known beforehand, it

can be used directly; otherwise, it has to be estimated. Although, region of support estimation is an easy problem for high SNR cases, it is very problematic for low SNR cases and complex algorithms have to be used to increase the filtering performance.

In this chapter an alternative time-frequency filtering algorithm, Time-Frequency Peak Filtering (TFPF), which is proposed by Boualem Boashash and Mostefa Mesbah is described [7]. The difference of this algorithm from the existing ones arises from the fact that the noisy signal is encoded as instantaneous frequency (IF) of an analytic frequency modulated (FM) signal, and the FM signal is transformed to the joint time-frequency plane; then to recover the desired signal, IF estimation is performed which actually gives the desired signal estimate multiplied with a constant scalar. To do so, an appropriate TFD is used; and since it is shown that TFDs concentrate signal energy at and around IF on the time-frequency plane, to get the IF estimate peak detection is used on the TFD data.

Among different TFDs, WVD is chosen in the TFPF algorithm due to its high time-frequency resolution. However, WVD can not be used in peak detection directly, because it is shown that when the peak detection is applied to WVD, the IF estimates will be biased for IF laws higher than linear [7], [5]. On the other hand, it is also shown that for deterministic band limited nonstationary multi component signals in additive white Gaussian noise (AWGN), the IF estimation using pseudo WVD (pWVD) will be approximately unbiased for a certain window length which depends on the maximum signal frequency and the sampling frequency. So, pWVD is used in TFPF algorithm to reduce the IF estimation bias.

In the first section, IF estimation error bias and variance and the reduced bias window length are derived; in the second section, the discrete time TFPF algorithm is given; and finally in the last section, the performance of the algorithm and the simulation results are evaluated.

3.1 BASIC CONCEPTS

3.1.1 IF Estimation Error Bias

Let the signal model be $s(t) = x(t) + n(t)$, where $n(t)$ is AWGN. After the encoding process $z_s(t)$ is formed as

$$\begin{aligned} z_s(t) &= \exp(j2\pi\mu\int_0^t s(\lambda)\,d\lambda) = \exp(j2\pi\mu\int_0^t x(\lambda)\,d\lambda)\exp(j2\pi\mu\int_0^t n(\lambda)\,d\lambda) \\ &= z_x(t)\,z_n(t) \end{aligned} \tag{3-1}$$

where μ is the frequency deviation constant (it has nothing to do with the IF estimation; it is used to control the bandwidth of frequency modulation to avoid aliasing).

As seen from (3-1) encoding process converts additive noise $n(t)$ to multiplicative noise $z_n(t)$. From (A-5) the WVD of $z_s(t)$ is

$$W_{z_s}(t,\omega) = W_{z_x}(t,\omega) *_\omega W_{z_n}(t,\omega) \tag{3-2}$$

obtained from the convolution of the WVD of $z_x(t)$ and $z_n(t)$ along frequency. It is evident that $W_{z_n}(t,\omega)$ spreads $W_{z_x}(t,\omega)$ through convolution, so the IF estimation bias depends on both $z_x(t)$ and $W_{z_n}(t,\omega)$. In [7] it is shown that when $n(t)$ is AWGN, $W_{z_n}(t,\omega)$ has no effect on the IF estimation bias, since $W_{z_x}(t,\omega)$ is lowpass with a maximum at zero frequency, and the bias only depends on $x(t)$.

In [5] IF estimation error bias and variance are studied in case of additive noise, and it is shown that the IF estimation bias does not depend on the additive noise; it only depends on the desired signal and on the window. Although, we have a

multiplicative noise in our problem, since the encoded noise $z_n(t)$ obtained from AWGN has no effect on the IF estimation error bias, the results of [5] can be used.

Let

$$z_x(t) = \exp(j\phi(t)) \tag{3-3}$$

where $\phi(t) = 2\pi\mu\int_0^t x(\lambda)\,d\lambda$. Assume that the signal is sampled with period T.

The IF $\omega_i(t)$ of $z_x(t)$ is $2\pi\mu\, x(t)$. The general form of shift covariant TFDs (Cohen's Class) in discrete time domain can be written in the following form,

$$C_y(t,\omega;\varphi) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\varphi(mT,nT) \\ y(t+mT+nT)\,y^*(t+mT-nT)\exp(-j2\omega nT) \tag{3-4}$$

where $\varphi(t,\tau)$ is the TFD kernel in the time-lag domain, which determines the TFD characteristics. When the kernel has finite length in both time and lag axes, it will give the pseudo form of the TFDs. The notation $\varphi_h(mT,nT) = (T/h)^2\,\varphi(mT/h, nT/h)$ is used for a finite length kernel whose length is h; the constant term $(T/h)^2$ is used to make the sum of $\varphi_h(mT,nT)$ over time and lag h independent (this constant term has nothing to do in IF estimation; it is given for the sake of completeness). Additionally, it is assumed that the kernel is a symmetric function in both time and lag axes which is the case for most of the commonly used TFD kernels.

By inserting $z_x(t)$ instead of $y(t)$ following results are obtained [5]

$$C_f(t,\omega;\varphi_h) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\varphi_h(mT,nT)\,\exp(-j2\omega nT) \\ \exp(j\phi(t+mT+nT))\exp(-j\phi(t+mt-nT)) \tag{3-5}$$

Expanding $\phi(t+mT\pm nT)$ around t by the Taylor series

$$\begin{aligned}\phi(t+mT\pm nT)=\phi(t)+\phi^{'}(t)(mT\pm nT)+\phi^{(2)}(t)(mT\pm nT)^2/2!+\\ ...+\phi^{(n)}(t)(mT\pm nT)^n/n!+...\end{aligned} \tag{3-6}$$

yields

$$\begin{aligned}C_f(t,\omega;\varphi_h)=|A(t)|^2\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\varphi_h(mT,nT)\\ \exp(-j\left[2nT(\omega-\phi^{'}(t))-2mnT^2\phi^{(2)}(t)-\Delta\phi(t,mT,nT)\right])\end{aligned} \tag{3-7}$$

$\Delta\phi(t,mT,nT)$ is a residue of phase which is equal to

$$\Delta\phi(t,mT,nT)=\sum_{s=3}^{\infty}\phi^{(s)}(t)[(mT+nT)^s-(mT-nT)^s]/s! \tag{3-8}$$

As seen from (3-7) the TFD will have a maximum at the point $\omega(t)=\phi^{'}(t)$, which is the IF of the signal $z_x(t)$, if $\phi^{(s)}(t)=0$ for $s\geq 2$; otherwise, there are oscillatory terms. This shows that the estimated IF using peak detection will be equal to the true IF for signals with linear or lower IF law, and for such signals the IF estimation error bias will be zero.

Using above results the IF estimate $\hat{\omega}_i(t)$ can be found using

$$\hat{\omega}_i(t)=\arg\left[\max_{\omega\in W}\{C_x(t,\omega;\varphi_h)\}\right] \tag{3-9}$$

where $W=\{\omega;\ 0\leq|\omega|\leq\pi/2T\}$ is a basic frequency interval, and (3-9) can be solved by taking the partial derivate of $C_x(t,\omega;\varphi_h)$ with respect to ω.

The IF estimation error produced at time instant t is given by

$$\Delta\hat{\omega}_i(t)=\omega(t)-\hat{\omega}_i(t) \tag{3-10}$$

It is shown in [5] that the estimation error bias can be obtained using the following formula

$$bias = E\{\Delta\hat{\omega}_i(t)\} = \frac{P_h(t)}{2R_h(t)} \tag{3-11}$$

where

$$R_h(t) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\varphi_h(mT,nT)(nT)^2\exp(j2\phi^{(2)}(t)(mT)(nT)) \tag{3-12}$$

$$P_h(t) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\varphi_h(mT,nT)\Delta\phi(t,mT,nT)(nT)\exp(j2\phi^{(2)}(t)(mT)(nT)) \tag{3-13}$$

The bias formula shows that the bias will be zero for signals with linear or lower IF law, since $P_h(t)$ will be zero for these signals.

This bias formula is for the general TFDs. To get the result for special TFDs, their kernels can be inserted into the above formula. This is done for pseudo WVD whose kernel is equal to $\varphi_h(mT,nT) = w_h(mT)\delta(m+n)w_h(nT)$ where $w_h(nT)$ is a real-valued even window function which is equal to $(T/h)w(nT/h)$, and the IF estimation error bias is equal to [5]

$$bias(t,h) = \frac{B_h(0,4)}{6B_h(0,2)}w^{(2)}(t) = \frac{1}{6}\frac{M_4^{w^2}}{M_2^{w^2}}w^{(2)}(t)h^2 \tag{3-14}$$

where $B_h(k,l) = \sum_n\sum_m \varphi_h(mT,nT)(mT)^k(nT)^l$ and $M_r^w = \int_{-1/2}^{1/2} w(\tau)\tau^r d\tau$.

As seen from (3-14) the error bias is directly proportional to the square of the window length which means that to decrease the error bias the window length must be decreased. However, the window length does not only effect the IF estimation error bias, it also effects the time-frequency resolution of the distribution; when the

window length is decreased the time-frequency resolution becomes worse. Therefore, the window length can not be decreased freely. The derivations to determine the window length in TFPF algorithm will be given in the next section.

3.1.2 Worst-Case Window Length for Reduced Error Bias

Recall that for signals with IF laws lower than quadratic, peak detection using WVD produces unbiased IF estimates. Using this information, if the length of the window used in pWVD is adjusted such that the IF of the signal is almost linear in the window, then the IF estimate using pWVD will be unbiased too. For signals whose IF changes rapidly, short windows must be used to get linear regions of the IF. But decreasing window length results in reduced time-frequency resolution, so how much does this window have to be shrunk?

To obtain the relation between the window length and the IF estimation bias, the signal $x[n] = \cos(\phi_x[n])$ which is encoded as

$$z_x[n] = \exp(j2\pi\mu \sum_{m=0}^{n} \cos(\phi_x[m])) \tag{3-15}$$

is used. Without loss of generality a rectangular window is used. It is assumed that the window length h_w is such that $x[n] \approx \alpha n + C$ within the window. At the peaks or valleys of $x[n]$ the maximum deviation from linearity occurs and the validity of this assumption is weak at these peak or valley points m_p corresponding to f_p which is the maximum frequency of $x[n]$. Since the maximum deviation of the signal IF from linearity gives the maximum bias, this point will give the worst-case window length.

To get the worst-case window length, encode the signal with constant frequency f_p [7]

$$z_x(n) = \exp(j2\pi\mu\sum_{m=0}^{n}\cos(2\pi\frac{f_p}{f_s}m)) \tag{3-16}$$

The bilinear data product at time m_p and lag h is obtained as

$$\begin{aligned} K_{z_x}(m_p,h) &= z_x(m_p + h/2)z_x^*(m_p - h/2) \\ &= \exp(j2\pi\mu\sum_{m=m_p-h/2+1}^{m_p+h/2}\cos(2\pi\frac{f_p}{f_s}m)) \end{aligned} \tag{3-17}$$

Using the central finite difference (CFD) estimator, the IF of $K_{z_x}(m_p,h)$ is obtained as a function of m_p and h which is given as [7]

$$f_i(m_p,h) = \mu\cos(2\pi\frac{f_p}{f_s}m_p)\cos(\pi\frac{f_p}{f_s}\frac{h}{2}) \tag{3-18}$$

The maximum value for h_w is obtained from the deviation of $f_i(m_p,h)$ from $\mu x(n)$ which is limited as

$$\left|\mu\cos(2\pi\frac{f_p}{f_s}m_p)\cos(\pi\frac{f_p}{f_s}\frac{h}{2}) - \mu\cos(2\pi\frac{f_p}{f_s}m_p)\right| \le \varepsilon\left|\mu\cos(2\pi\frac{f_p}{f_s}m_p)\right| \tag{3-19}$$

The extent of h_w for unbiased IF estimates is obtained from (3-19) as

$$\left|1-\cos(\pi\frac{f_p}{f_s}\frac{h_w}{2})\right| \le \varepsilon \tag{3-20}$$

where f_s is the sampling frequency, f_p is the maximum frequency of the signal and ε is a constant chosen to limit the bias. The desired solution of (3-20) is the smallest positive one. So h_w can be obtained from

$$1 \le h_w \le \arccos(1-\varepsilon)\frac{2f_s}{\pi f_p} \tag{3-21}$$

(3-21) shows that the worst case window length is directly proportional to the f_s/f_p ratio. In section 3.1.1 it is shown that to reduce the IF estimation error bias the window must be narrowed which results in reduced time-frequency resolution. However, (3-21) says that for a given error tolerance ε, the worst case window length can be increased by increasing the f_s/f_p ratio which results in improved resolution and suppressed error bias below the tolerance; and the result is a better IF estimation performance. Therefore, the sampling frequency plays an important role on the effectiveness of the TFPF algorithm.

In the TFPF algorithm the worst case window length is used which results in decreased error bias and reduced time-frequency resolution (localization) for the entire signal. Since, the IF of the signal is not always high, at time instants where IF is low higher window lengths can be used which gives better time-frequency localization. Therefore, an adaptive algorithm for window length selection seems to perform better compared to fixed worst-case window length selection. There are publications in literature concerned with the adaptive and data-driven window length selection [1], [3], [4]; however, these publications provide solutions to the problem for additive white noise and try to reduce the bias and variance of the WVD; but in this problem the resultant noise is multiplicative and nonwhite and the bias and variance of the IF estimation error is tried to be decreased.

3.2 DISCRETE TIME TFPF ALGORITHM

The sampled noisy observed signal is modeled by the following equation

$$s[n] = x[n] + w[n] = \sum_{k=1}^{p} x_k[n] + w[n] \qquad (3\text{-}22)$$

for $0 \le n \le N-1$ where $x_k[n]$'s are band limited nonstationary deterministic components that may have overlapping frequency spectra, $w[n]$ is AWGN and N is the number of signal samples. Figure 3-1 shows the algorithm steps.

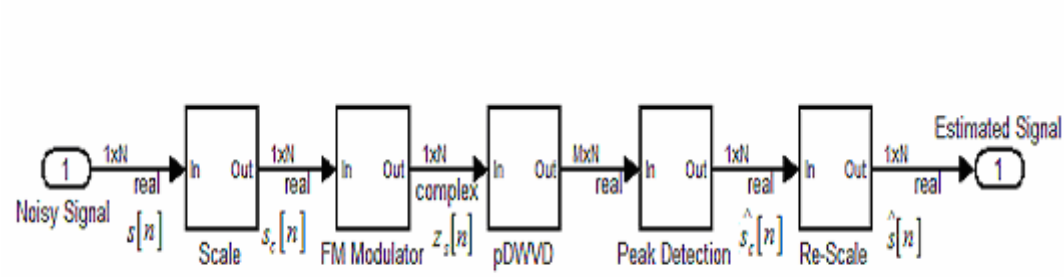

Figure 3-1 Discrete-time TFPF algorithm steps

From the observed signal $s[n]$, the analytic FM signal is obtained using

$$z_s[n] = \exp(j2\pi\mu \sum_{m=0}^{n} s_c[m]) \tag{3-23}$$

where

$$s_c[m] = (a-b)\frac{s[m] - \min[s[m]]}{\max[s[m]] - \min[s[m]]} + b \tag{3-24}$$

is the scaled version of $s[m]$ to avoid aliasing after frequency modulation.

The parameters satisfy $0.5 \geq a = \max[s_c[m]] > b = \min[s_c[m]] \geq 0$ and are chosen to provide suitable frequency limits on the encoded signal. After this scaling, the signal amplitude is limited to the interval [a, b], and using the frequency deviation constant μ the maximum and minimum IF frequency limits can be set.

The discrete pWVD (pDWVD) of the analytic FM signal $z_s[n]$ is computed. To do so, using the windowed form of the discrete WVD; first the bilinear data product $K_{z_s}(m_p, h)$ at time m_p is computed for integer values $h/2$ using the symmetric window $g(h) = g(-h)$ whose length is the worst case window length h_w:

$$K_{z_s}(m_p,h) = z_s(m_p + h/2)\, z_s^*(m_p - h/2)\, g(h/2)\, g(-h/2) \qquad (3\text{-}25)$$

and then the DFT of the $K_{z_s}(m_p,h)$ is computed. Finally, the peak values along frequency axis for time sample m_p are detected to get the IF estimate $\hat{s}_c[m_p]$. The peak detection is done for $0 \le m_p \le N-1$.

The signal estimate is obtained by back scaling the obtained estimates using the following formula

$$\hat{s}[m] = (\hat{s}_c[m] - b)(\max[s[m]] - \min[s[m]])/(a-b) + \min[s[m]] \qquad (3\text{-}26)$$

In the low SNR case the algorithm is applied iteratively. After the estimated signal samples are obtained, filtering continues or terminates if the number of successive iterations is larger than a specified number or the difference between the outputs of the successive iterations is smaller than a specified value.

3.3 THE EFFECTS OF WINDOW AND DFT LENGTHS AND THE NUMBER OF ITERATIONS ON THE PERFORMANCE OF THE ALGORITHM

In this section performance of the TFPF algorithm is investigated for different window lengths, different DFT lengths and different number of iterations by simulating the algorithm. In the simulations a linear frequency modulated test signal $x[m] = \cos(0.005m + 7.5*10^{-7} m^2)$ is used; the signal length is 4096 points. The performance of the algorithm is evaluated using the error bias and variance after the TFPF filtering algorithm.

3.3.1 Effects of Window Length on the Performance of the Algorithm

The performance of the algorithm is investigated for different window lengths in pDWVD computation.

In the first simulation the noise free case is investigated. The window length is varied in between 10 and 400 in steps of 10, and Figure 3-2 is obtained. As seen from the figure the error variance increases with the increase in window length, and after window length reaches 200 it is almost constant. The error bias decreases in magnitude with the increase in window length until window length reaches to 200, but unlike the error variance it is not constant.

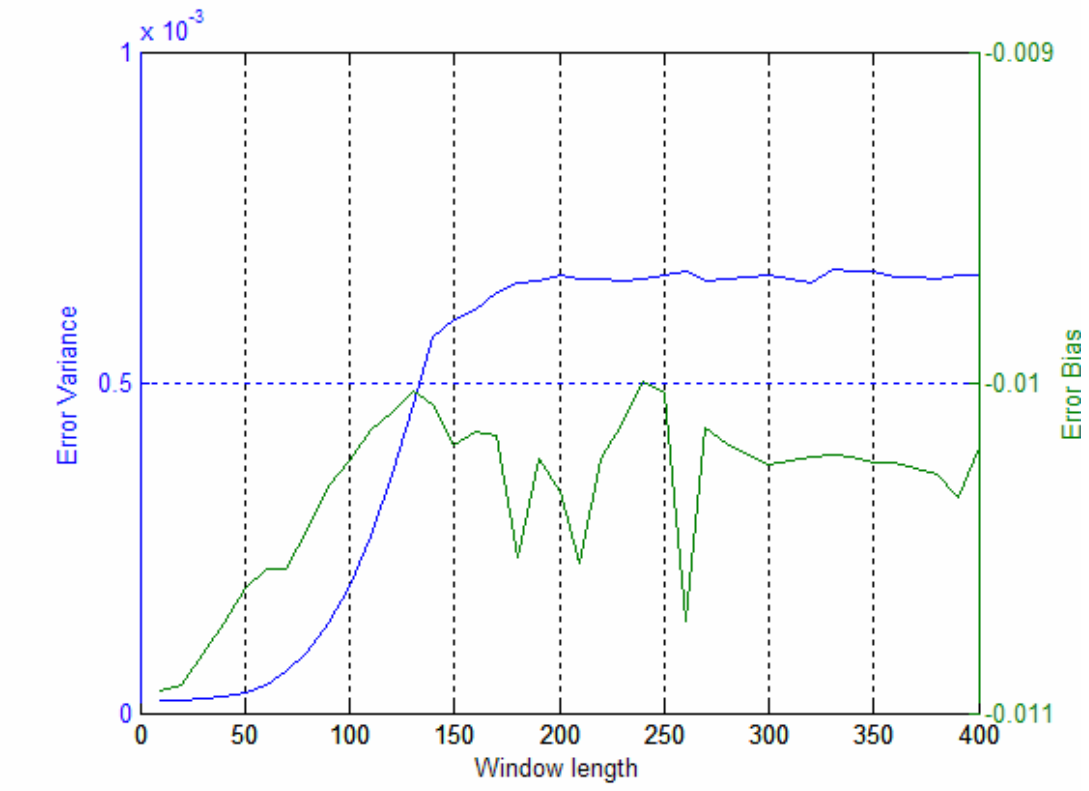

Figure 3-2 Error bias (green) and variance (blue) as functions of window length.

In the second simulation the noisy case is investigated. Additive white Gaussian zero mean noise is added to the signal, the variance of the noise is adjusted such that the SNR is -9 dB. In this simulation the window length is varied in between 10 and 500 in steps of 10, and Figure 3-3 is obtained. As seen from the figure, the error bias increases in magnitude with the increase in window length as expected; and the error variance decreases until window length is 170 and starts to increase after the window length is 250. Actually, this result says that to get the minimum error variance a window length close to the worst case window length which is 228 must be used.

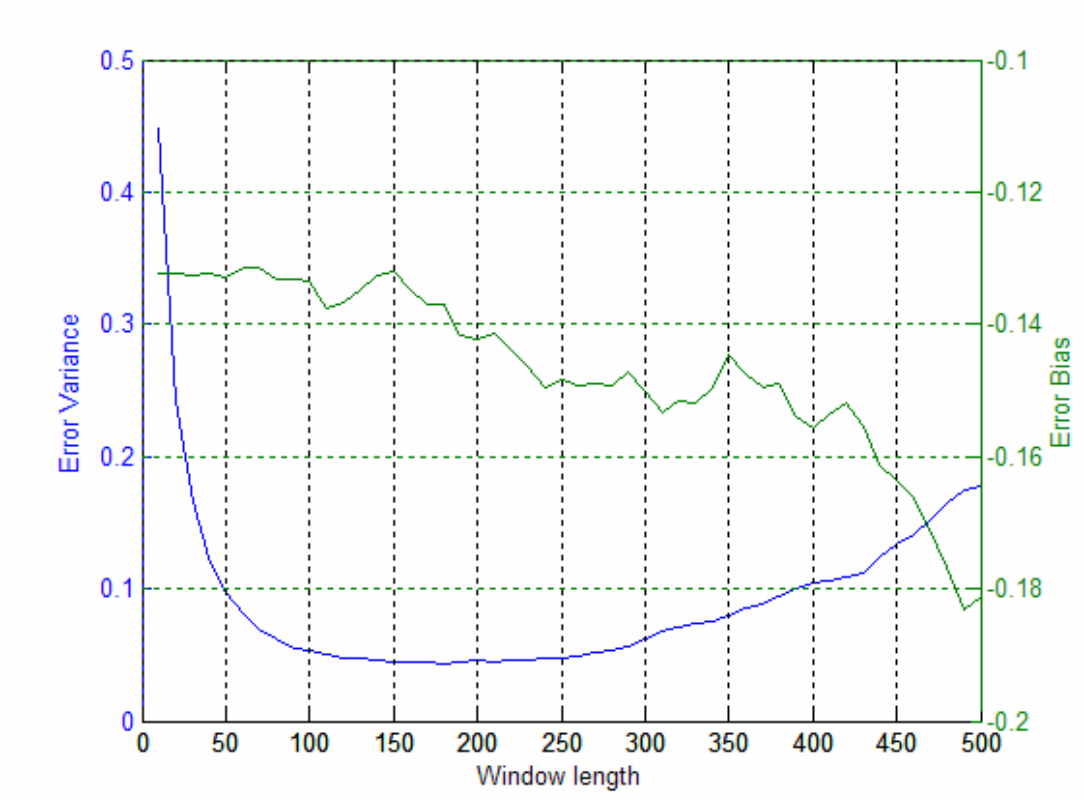

Figure 3-3 Error bias (green) and variance (blue) as functions of window length, SNR = -9 dB.

3.3.2 Effects of DFT Length on the Performance of the Algorithm

The performance of the algorithm is investigated when the number of DFT points is changed.

In the simulation the noise free case is investigated. Figure 3-4 and Figure 3-5 show the error variance and the error bias as functions of the DFT length, respectively. As seen from the figure both of them decrease in magnitude with the increase in DFT length, and after a certain value, for this example 512, they do not change much.

In the noisy case both the error variance and the error bias show the similar behavior with different variance and bias values; so the results of the noisy case are not included.

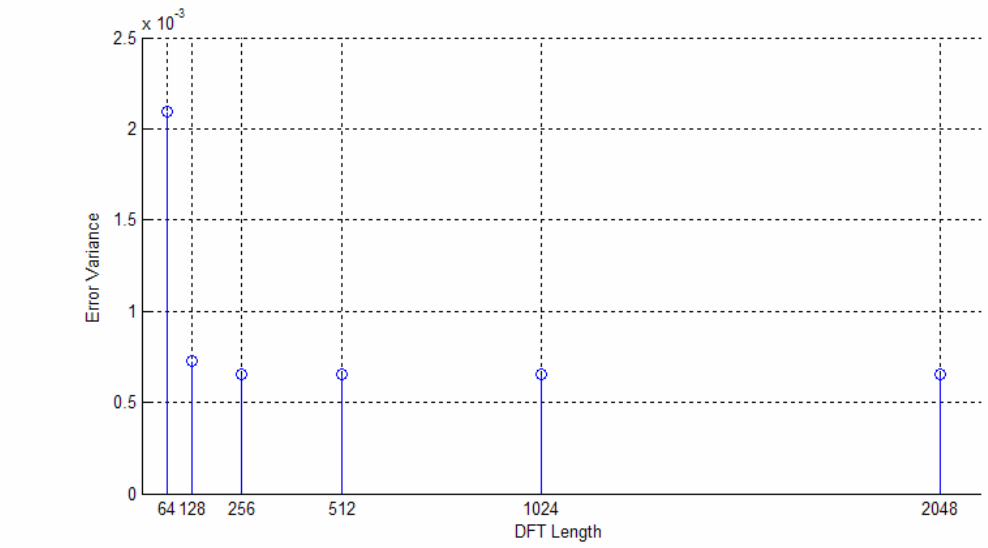

Figure 3-4 Error Variance as a function of DFT length.

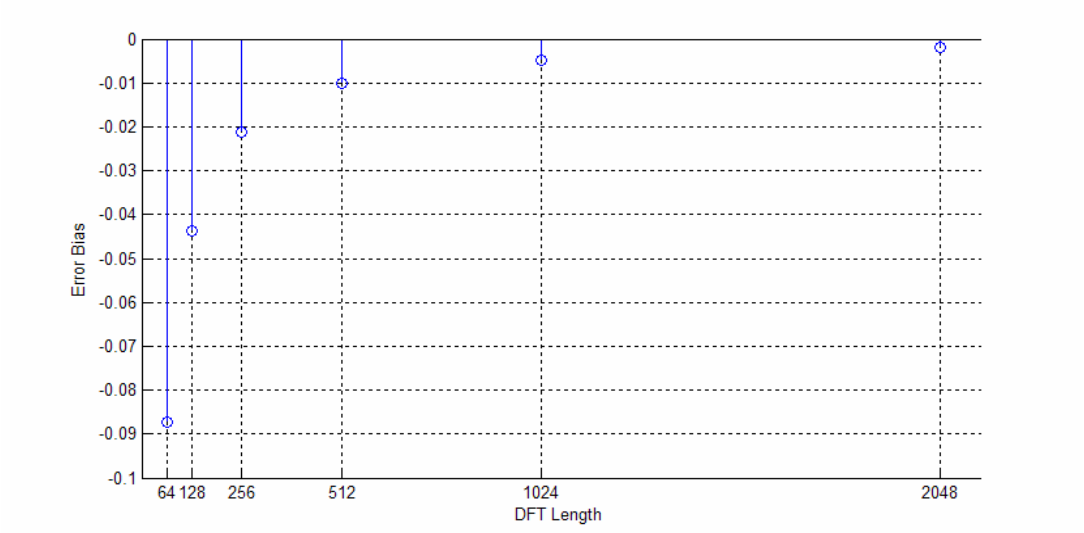

Figure 3-5 Error Bias as a function of DFT length.

3.3.3 Effects of Number of Iterations on the Performance of the Algorithm

The performance of the algorithm is investigated when the number of iterations is increased. The SNR is -9 dB.

Figure 3-6 shows the error variance and the error bias as functions of iteration number when the window length is set to 228. As seen from the figure, in the first steps of the iteration the error variance decreases, then it starts to increase. On the other hand, the error bias increases in magnitude with the increase in iteration number.

Figure 3-7 shows the error variance and the error bias as functions of iteration number when the window length is set to 48. The similar behaviour is observed; but when the window length is decreased it takes more iterations to reach the minimum error variance level.

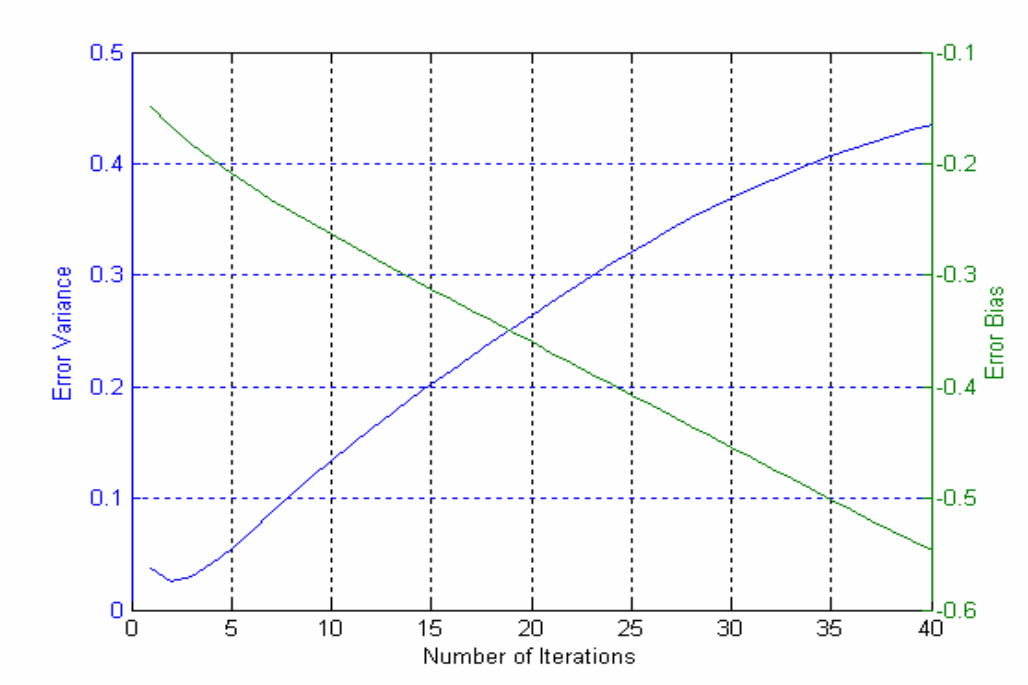

Figure 3-6 The variations of error variance (blue) and bias (green) as functions of number of iterations, the window length is set to 228.

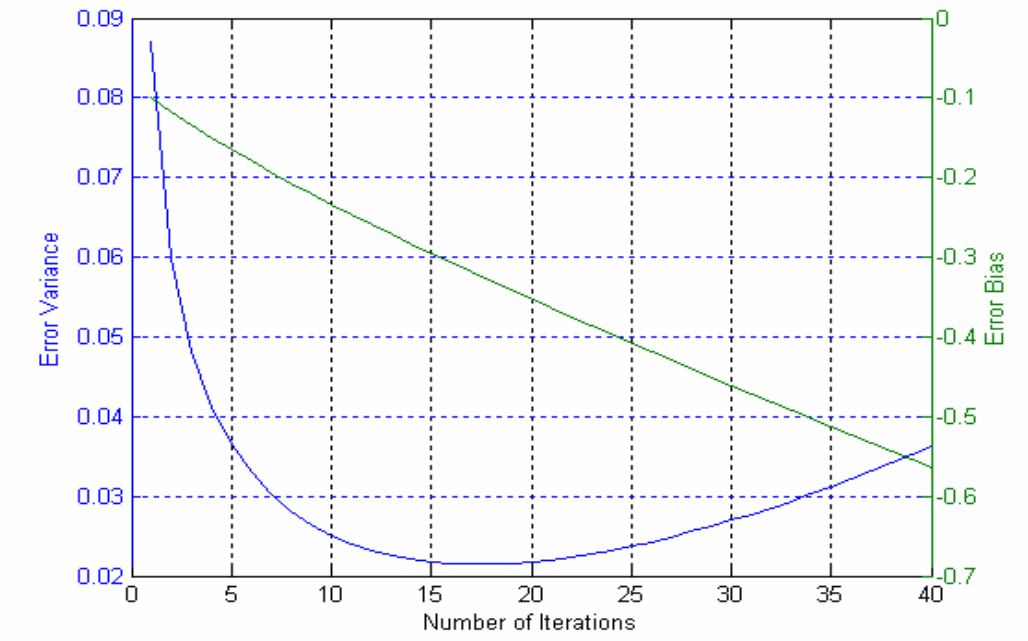

Figure 3-7 The variations of error variance (blue) and bias (green) as functions of number of iterations, the window length is set to 48.

Figure 3-8 shows the error variance as a function of number of iterations for different window lengths. As seen from the figure the minimum variance and the corresponding iteration number is different for different window lengths; additionally, for small window lengths it takes more iterations to attain the minimum variance level.

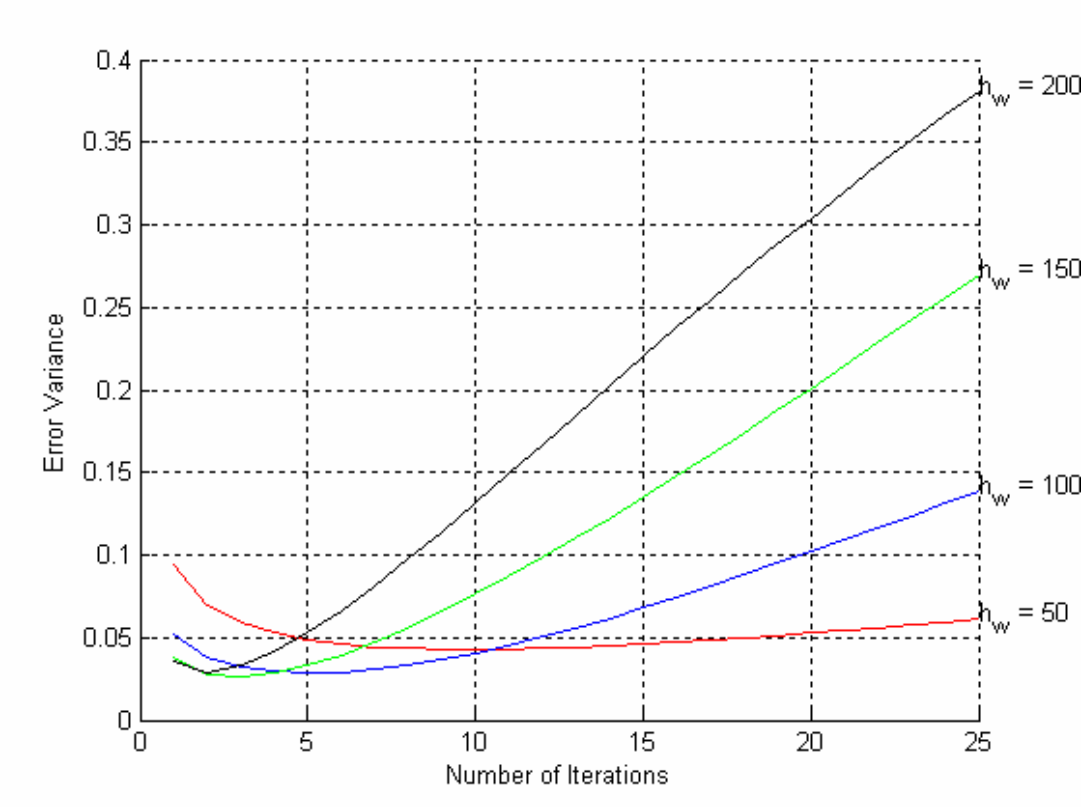

Figure 3-8 The variations of error variance as a function of number of iterations for different window lengths.

3.3.4 The Performance of the Algorithm for Different SNR Values

The performance of the algorithm for different SNR values is investigated; the window length is set to 228 and the DFT length is set to 512 using the results of the previous simulations. The output SNR is $10\log_{10}(P_s/\sigma_e)$ where P_s is the signal power, and σ_e is the variance of the estimation error.

Figure 3-9 shows the results of the simulation; as seen from the figure the TFPF algorithm increases the SNR by approximately 16 dB.

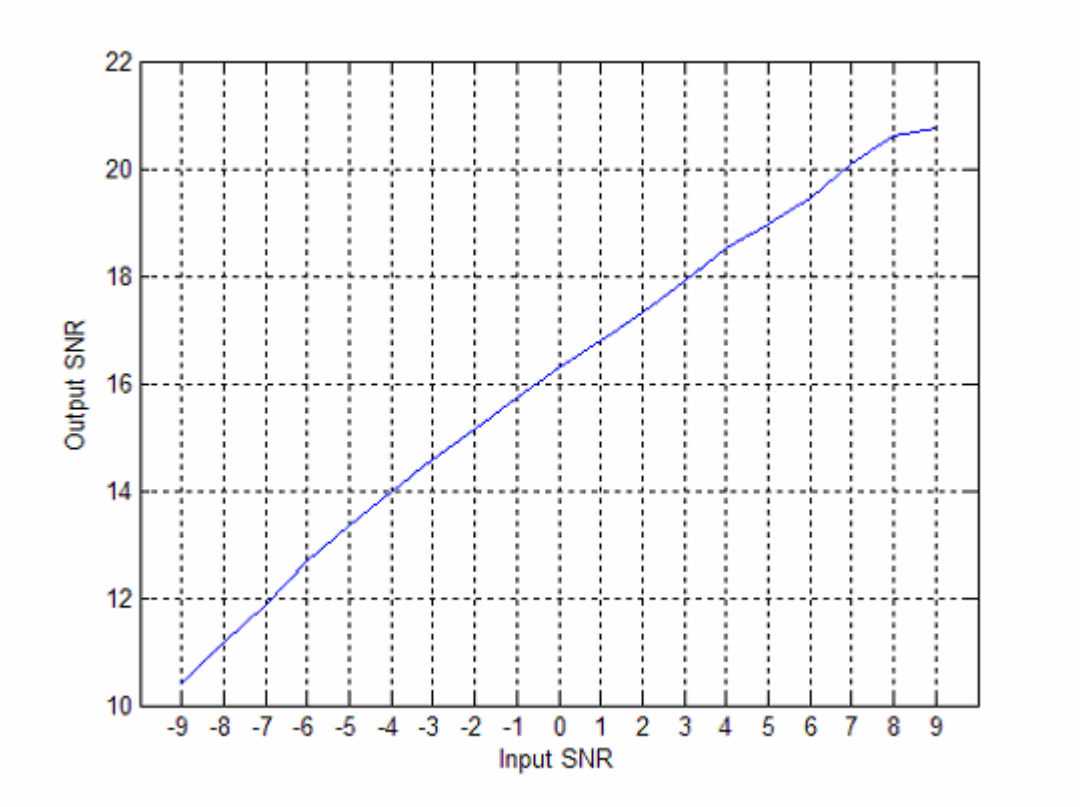

Figure 3-9 Input SNR vs output SNR (in dB).

3.3.5 Results

It is observed that as the number of iterations is increased up to a certain number, which depends on the window length, the error variance decreases. The error variance increases when the number of iterations are further increased. Therefore, continuing the algorithm for a predefined number of iterations or stopping the iterations when the difference between successive iterations is decreased to some predefined level may result in a poor filtering performance. This is so because the difference between successive iterations may never decrease to that predefined level, which results in infinite number of iterations, or the predefined iteration

number may be higher than the iteration number where the error variance starts increasing. In order to obtain a better performance, we propose to stop the algorithm when the difference between successive iterations starts to increase.

It is observed that a further increase of the DFT length above a certain value does not make a significant improvement to the performance of the algorithm. Moreover, increasing the DFT length decreases the speed of the algorithm. Therefore, the DFT length should be set to a value after which its increase does not improve the performance very much. For the signal used in the simulations the DFT length is set to 512.

For the noise free case, it is observed that the increase in the window length increases the error variance, and after a certain value the error variance stays approximately constant. However, when a signal contains AWGN, the increase in the window length up to the worst-case length decreases the error variance, and when the window length exceeds that value the error variance starts to increase.

Window length selection is important for the performance of the algorithm. Fast varying signals can not be recovered using a long window; on the other hand, short windows degrade the performance of the TFPF method. The maximum window length will be equal to 1 when $f_p = 0.4 f_s$ and $\varepsilon = 0.2$, which means that if the maximum frequency of the desired signal is higher than $0.4 f_s$ the estimated signal will be biased. Moreover, when $\varepsilon = 0.05$ the maximum frequency of the desired signal should not exceed $0.2 f_s$ in order to obtain an unbiased signal estimate. When short window lengths have to be used, the iterative algorithm should be preferred. To recover the signal the algorithm needs too much iterations as the window length is shortened. The number of iterations can be decreased using a signal and time dependent window length for a signal which stays a little time at high frequencies and most of the time at low frequencies.

When the window length is decreased aliasing starts; but since the TFPF algorithm uses an analytic FM signal to encode the input signal, it is known that only one part

(positive frequency part) of the time-frequency plane is used; so aliasing is not a problem when the window length is small.

Noting that the analytic FM signal is used to decode the input signal, there is no need to over sample the input signal by a factor of 2 in order to avoid aliasing in pDWVD. Nyquist sampling is sufficient in the TFPF method, but because of the window length selection problem, the method will give unsatisfactory results for the signals whose maximum frequency exceeds $0.4f_s$ (or $0.2f_s$ depending on ε). Also, there is no cross-term elimination problem in this application because of the encoding procedure. Since there is only one frequency term after encoding, the cross-terms do not appear.

The TFPF filtering algorithm is successful, as seen from the results, for signal enhancement problem. Especially the performance of the iterative algorithm is good at SNR levels down to -9 dB.

It is worth adding that this method can be used as a pre-processing method; because the noise can be removed from the input signal using this method. However, when the problem is filtering one signal from a sum of signals in AWGN, this method can not solve the problem.

CHAPTER 4

SUB-OPTIMAL TIME-VARYING FILTERING USING WIGNER-VILLE DISTRIBUTION

Filtering of a desired signal from noisy observations is one of the most important problems in many signal processing applications. There are different approaches proposed to solve the filtering problem for different kinds of signals and application areas. Even though the simplest solution is the use of a classical filter whose pass-band covers the frequency band of the desired signal, it is not always the optimum solution. Although, for stationary signals a better solution is the use of a Wiener Filter which produces the best signal estimate in the minimum mean square error sense, for nonstationary signals adaptive filtering methods, which are based on least-mean-squares, recursive-least-mean-squares or Kalman Filter, are preferable. However, in [7], it is claimed that adaptive filtering methods are not sufficient for nonstationary signals whose frequency content changes quickly with time. Additionally, these adaptive methods need signal modeling for optimal performance, but when the structure of the desired signal is unknown, they will give suboptimal results and even in some cases wrong results.

In this chapter a filtering algorithm based on time-varying Wiener Filter which is proposed for nonstationary signals with rapidly varying frequency content is introduced. In the first two sections the derivations of the time-varying Wiener Filter and its time-frequency formulation using WVD are given; the next section

deals with the instantaneous frequency (IF) estimation problem. In the last section the proposed filtering algorithm is given, and the performance of the proposed filter is investigated.

4.1 TIME-VARYING WIENER FILTER

Consider the noisy observed signal

$$x(t) = s(t) + n(t) \tag{4-1}$$

where *s(t)* is a zero-mean, nonstationary real or complex random process with known correlation function $r_s(t,t') = E\{s(t)s^*(t')\}$, $n(t)$ is an additive zero-mean nonstationary real or complex random noise with known correlation function $r_n(t,t') = E\{n(t)n^*(t')\}$; $s(t)$ and $n(t)$ are uncorrelated processes. Let H be a linear time-varying system with kernel $h(t,t')$, then the estimated signal $\hat{s}(t)$ is obtained using

$$\hat{s}(t) = (Hx)(t) = \int_{t'} h(t,t')x(t')dt' . \tag{4-2}$$

The estimation error is

$$e(t) = s(t) - \hat{s}(t) . \tag{4-3}$$

The performance measure is the mean-square error as in the Wiener Filter derivation for stationary random processes; and as it is mentioned in [9], the expected error energy is minimized when the error is orthogonal to the observed noisy signal:

$$E\{(s(t) - (Hx)(t))\ x^*(t-\tau)\} = 0 . \tag{4-4}$$

Inserting (4-1) in (4-4) we obtain

$$E\left\{s(t)\left(s^*(t)+n^*(t)\right)\right\}-E\left\{\left(H(s+n)(t)\right)\left(s^*(t)+n^*(t)\right)\right\}=0\,. \tag{4-5}$$

Using the fact that signal and noise are uncorrelated and defining a correlation operator R which is a self-adjoint, positive (semi)-definite linear operator whose kernel is the correlation function [8], (4-5) can be reduced to

$$R_s - HR_s - HR_n = 0\,. \tag{4-6}$$

From (4-6) the time-varying linear system operator is obtained with minimal rank as

$$H = R_s(R_s + R_n)^{-1} \tag{4-7}$$

where $(R_s + R_n)^{-1}$ is the pseudo inverse of $(R_s + R_n)$.

4.2 TIME-FREQUENCY FORMULATION OF THE TIME VARYING WIENER FILTER

The time-varying filtering equation (4-2) is used with a little modification, as is proposed in [6] and [9]. The modified filtering equation is

$$\hat{s}(t)=\int_{\tau} h(t+\tau/2,t-\tau/2)x(t+\tau)d\tau\,. \tag{4-8}$$

The reason for the modification is: when there is no noise on the signal the ordinary filtering equation produces the signal with an amplitude variation and phase deviation; on the other hand, the modified filtering equation eliminates the phase deviation [9].

The time-varying transfer function of a linear time-varying operator H with kernel $h(t+\tau/2,t-\tau/2)$ is defined as (which is actually known as the Weyl Symbol (WS) [8])

$$L_H(t,\omega)=\int h(t+\tau/2,t-\tau/2)\exp(-j\omega\tau)\,d\tau \tag{4-9}$$

The kernel can be obtained by inverse FT

$$h(t+\tau/2, t-\tau/2) = \frac{1}{2\pi}\int L_H(t,\omega)\exp(j\omega\tau)\,d\omega \tag{4-10}$$

Inserting (4-10) into (4-8) the following filtering equation is obtained

$$\hat{s}(t) = \frac{1}{2\pi}\int L_H(t,\omega)\,X(\omega)\exp(j\omega t)\,d\omega \tag{4-11}$$

where $X(\omega)$ is the FT of the observed signal $x(t)$. As seen from (4-11) the time-varying filtering operation is equivalent to inverse FT of the product of the time-varying filter transfer function with the Fourier Transform of the observed signal. However, in practical applications the observation time is limited and pseudo forms of (4-8) and (4-11)

$$\begin{aligned} \hat{s}(t) &= \int h(t+\tau/2, t-\tau/2)\; x(t+\tau)\,w(\tau)\,d\tau \\ \hat{s}(t) &= \frac{1}{2\pi}\int L_H(t,\omega)\; STFT_x(t,\omega)\; d\omega \end{aligned} \tag{4-12}$$

are used, where $w(t)$ is the window function and $STFT_x(t,\omega)$ is the Short Time Fourier Transform of the observed signal $x(t)$.

As seen from (4-12) to get the estimated signal the STFT of the observed signal and the WS, which is actually the time-varying transfer function, of the linear time-varying filter are needed. It is easy to compute the STFT of the observed signal; and for optimal filtering the WS of the filter is obtained from the time-varying Wiener Filter derivation.

Recall that the time-varying Wiener Filter is obtained from (4-6) or equivalently from (4-7). Using these equations the time-frequency formulation of the Wiener Filter is proposed to be [8]

$$L_H(t,\omega) \approx \frac{\overline{W}_s(t,\omega)}{\overline{W}_s(t,\omega) + \overline{W}_n(t,\omega)} \tag{4-13}$$

As seen from (4-13) to get the time-varying transfer function, the WVD of the desired signal is needed which is unknown. Therefore, for practical use (4-13) can be simplified as

$$L_H(t,\omega) = \begin{cases} 1, & (t,\omega) \in R_s \\ 0, & (t,\omega) \notin R_s \end{cases} \tag{4-14}$$

where R_s is the region of support of the signal, that is to say R_s is the time-frequency region where the signal is present. (4-14) follows from the fact that on the time-frequency locations where the signal is dominant L_H is approximately 1, and on the time-frequency locations where the noise is dominant L_H is approximately 0. So, to get L_H, the region of support of the noisy signal has to be determined which is equivalent to the instantaneous frequency (IF) estimation. Hence the problem of L_H determination can be solved using the solutions proposed for IF estimation of noisy signals using WVD.

To use in implementations (4-12) can be discretized as [9]

$$\hat{s}[n] = \sum_k L_H[n,k]\ STFT_x[n,k] \tag{4-15}$$

4.3 IF ESTIMATION OF NOISY SIGNALS USING WVD

There are different publications in literature which are concerned with the problem of IF estimation from noisy observations using WVD [1], [3], [4], [5] and [6]. The main approach to obtain the IF estimates from WVD is known as the peak detection method in which the peak values over the frequency are detected for every time sample. In the following, the approach proposed in [1] is examined.

In [1] it is claimed that the lag window length is one of the most important parameters in optimizing the pseudo WVD of the discrete time noisy signals. In other words, it is claimed that by use of the appropriate lag window length, the optimum WVD can be obtained for noisy signals. However, in order to get the appropriate window length the unknown derivatives of the WVD are needed to get the bias value; and in practical applications it is useless. So, a simple adaptive algorithm for the efficient time-frequency representation of noisy signals which uses only the noisy estimate of WVD and the analytical formula for the variance of this estimate is proposed.

In Section 4.3.1 the derivation of the optimum window length is given; and in Section 4.3.2 the proposed algorithm is presented, and the performance of the algorithm is evaluated.

4.3.1 Optimum Window Length for Pseudo DTWVD

The DTWVD of a discrete time noisy signal $x[n] = f[n] + v[n]$ is defined as

$$W_{xx}[n,\theta] = \sum_{k=-N/2}^{N/2-1} x[n+k]\,x^*[n-k]\exp(-j2\theta k) \tag{4-16}$$

$f[n]$ is assumed to be deterministic and the noise is complex-valued white, $v[n] = v_1[n] + jv_2[n]$ with

$$\begin{aligned} &E\{v[n]\} = 0 \\ &E\{v[n]v^*[m]\} = \sigma_v^2\,\delta[n-m] \\ &E\{v_1[n_1]v_2[n_2]\} = 0 \end{aligned} \tag{4-17}$$

Let $\Delta W_{xx}[n,\theta] = W_{xx}[n,\theta] - W_{ff}[n,\theta]$ be the estimation error; then the bias of the estimate is equal to $E\{\Delta W_{xx}[n,\theta]\} = \sigma_v^2$. The variance of the DTWVD is $\sigma_{xx}^2 = E\{W_{xx}[n,\theta]W_{xx}^*[n,\theta]\} - E\{W_{xx}[n,\theta]\}\;E\{W_{xx}^*[n,\theta]\}$. It has two components

$$\sigma_{xx}^2 = \sigma_{fv}^2 + \sigma_{vv}^2 \tag{4-18}$$

where $\sigma_{fv}^2 = \sigma_v^2 \sum_{k=-N/2}^{k=N/2-1} \left| f(n+k) \right|^2 + \left| f(n-k) \right|^2$ and $\sigma_{vv}^2 = N\sigma_v^4$.

As seen from the obtained results the bias is a constant for every *(n,θ)* pair for the unwindowed DTWVD and only depends on the noise; and the variance goes to infinity as N goes to infinity.

When a window is used in the computation of the DTWVD (pseudo DTWVD) the following results are obtained for the bias and variance

$$W_{xx}[n,\theta;N] = \sum_{k=-\infty}^{\infty} w(k)\, w(-k)\, x(n+k)\, x^*(n-k) \exp(-j2\theta k) \tag{4-19}$$

$$E\{W_{xx}[n,\theta]\} = W_{ff}[n,\theta] *_\theta F_w(2\theta) + \sigma_v^2 w^2(0) \tag{4-20}$$

where $F_w(\theta)$ is the FT of the symmetric window $w_e(k) = w(k)w(-k)$; and $*_\theta$ denotes the convolution with respect to θ.

As seen from (4-20) the bias term has two components; one of them depends on the signal and the other term depends on the noise. It is assumed that w(0) does not depend on N, the window length, so the noise dependent term is omitted in the following analysis. Using the Taylor series expansion for $W_{xx}[n,\theta]$

$$W_{ff}[n,\theta] *_\theta F_w(2\theta) \cong W_{ff}[n,\theta] + \frac{1}{8} \frac{\partial^2 W_{ff}[n,\theta]}{\partial \theta^2} m_2 \tag{4-21}$$

where $m_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \theta^2 F_w(\theta) d\theta$.

So, the signal dependent bias term is obtained as

$$bias_{signal}[n,\theta;N] = \frac{1}{8}\frac{\partial^2 W_{ff}[n,\theta]}{\partial\theta^2} m_2 \tag{4-22}$$

For a nonrectangular window the two components of the variance in more general form are

$$\sigma_{fv}^2 = \sigma_v^2 \sum_{k=-N/2}^{N/2-1} w^2[k]w^2[-k]\,[\,|f[n+k]|^2 + |f[n-k]|^2\,] \tag{4-23}$$

and

$$\sigma_{vv}^2 = \sigma_v^2 \sum_{k=-N/2}^{N/2-1} w^2[k]w^2[-k] \tag{4-24}$$

The optimal window length is obtained by minimizing the mean squared error defined by

$$e^2[n,\theta;N] = bias_{signal}{}^2[n,\theta;N] + \sigma_{xx}^2 \tag{4-25}$$

It is obtained for FM signal with slowly varying amplitude $A[n]$ as in [1]

$$N_{opt} = \sqrt[5]{\frac{2(\frac{\partial^2 W_{ff}[n,\theta]}{\partial^2\theta})^2\pi^4}{3\sigma_v^2(2A^2[n]+\sigma_v^2)}} \tag{4-26}$$

As seen from the formula for optimal window length (4-26), the second derivative of the DWVD of the signal only part is needed which can not be obtained in practical applications; hence the optimal window length can not be computed in practical applications.

4.3.2 Two Window Length Algorithm

As it is shown in Section 4.3.1 to suppress the variance of the DWVD a window must be used in computations whose length depends on the signal and the second

derivative of the DWVD of the signal, the computation of which is not possible in practical applications where only the noisy observations of the signal are available. Hence an algorithm which computes the DWVD of the noisy signal for two different window lengths and makes a selection which depends on the differences between the DWVDs and the estimated variance is proposed in [1]. Actually the proposed algorithm uses M-many window lengths to find the one which is as close as possible to the optimal window length; however, in practical applications the use of two window lengths is sufficient.

The general form of the algorithm which uses M-many windows can be summarized as follows:

- Define a dyadic set **N** = {N_s | $N_s = N_{s-1}/2$, s = 1,2,3, …M}. Compute the DWVD for all N_s to obtain $W_{xx}[n,\theta;N_s]$

- The optimal window length $N_s[n,\theta]$ for every $[n,\theta]$ is determined by the largest s, s = 1,2, … J, when

$$\left|W_{xx}[n,\theta;N_s]-W_{xx}[n,\theta;N_{s+1}]\right| \le (K+\Delta K)(\sigma_{xx}(N_s)+\sigma_{xx}(N_{s+1}))$$

  is still satisfied.

It is shown that for the FM signal with slowly varying amplitude $A[n]$, the variance is

$$\sigma_{xx}^2 = E_w\sigma_v^2(2A^2[n]+\sigma_v^2) \tag{4-27}$$

As seen from (4-27) for variance computation the amplitude variation of the signal and the variance of the noise with the window energy are needed. The window energy can be computed from the type of the window used in the pseudo DWVD computation, but the other parameters should be estimated. For high noise case the following estimation can be used [1]:

$$\sigma_v^2(2A^2[n]+\sigma_v^2) \cong (A^2+\sigma_v^2) = \frac{1}{N^2}\left[\sum_{n=1}^{N}|x[n]|^2\right]^2 \qquad (4\text{-}28)$$

In the two window length algorithm, the window lengths are chosen such that N1<<N2. For N1 the variance of the DWVD is small and for N2 the bias of the DWVD is small. So, depending on the difference between the DWVDs the small variance (N1) or the small bias (N2) is favored. The algorithm chooses the small window length for the regions of noise only part and chooses the large window length for the regions of signal part.

If there is no noise, and the difference between the DWVDs of consecutive window lengths is not equal to zero, the algorithm chooses the larger window length to achieve higher concentration.

Figure 4-1 shows the IF estimation algorithm steps. To get the IF estimates from the pseudo DWVD, the peak detection algorithm is used. The peak values along the frequency axis for every time sample are detected. The resultant noise on the IF estimate is impulsive; hence to reduce the effect of the impulsive noise 1-D median filter is applied to the IF estimates. Median filtering is a nonlinear technique that applies a sliding window to a sequence. The median filter replaces the center value in the window with the median value of all the points within the window.

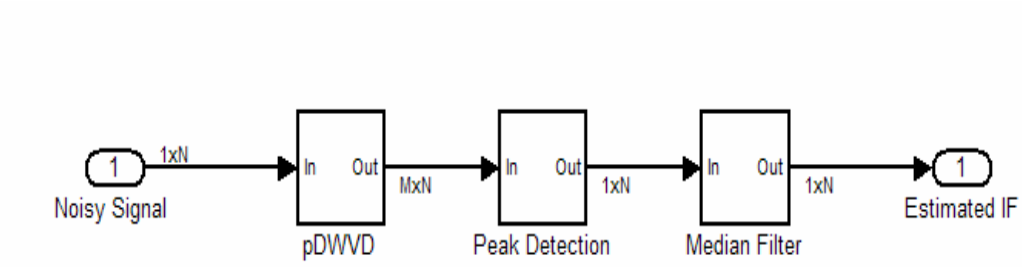

Figure 4-1 The IF estimation algorithm steps.

4.3.2.1 Simulation Results of the Algorithm

The simulations are done for a linear FM signal for two different SNR levels. The window lengths are N1 = 64 and N2 = 512. The obtained IF estimates before and after the median filtering are given.

4.3.2.1.1 Linear FM Signal

The test signal is $x[n] = \exp(j1600(nT - 0.05)^2) + A\,g[n]$

where T is the sampling period which is 1/2048 sec; $g[n]$ is the additive white Gaussian noise, and A is the amplitude of the noise which is used to adjust the SNR. The number of signal samples is 2048.

Figure 4-2 shows the IF of the linear FM signal.

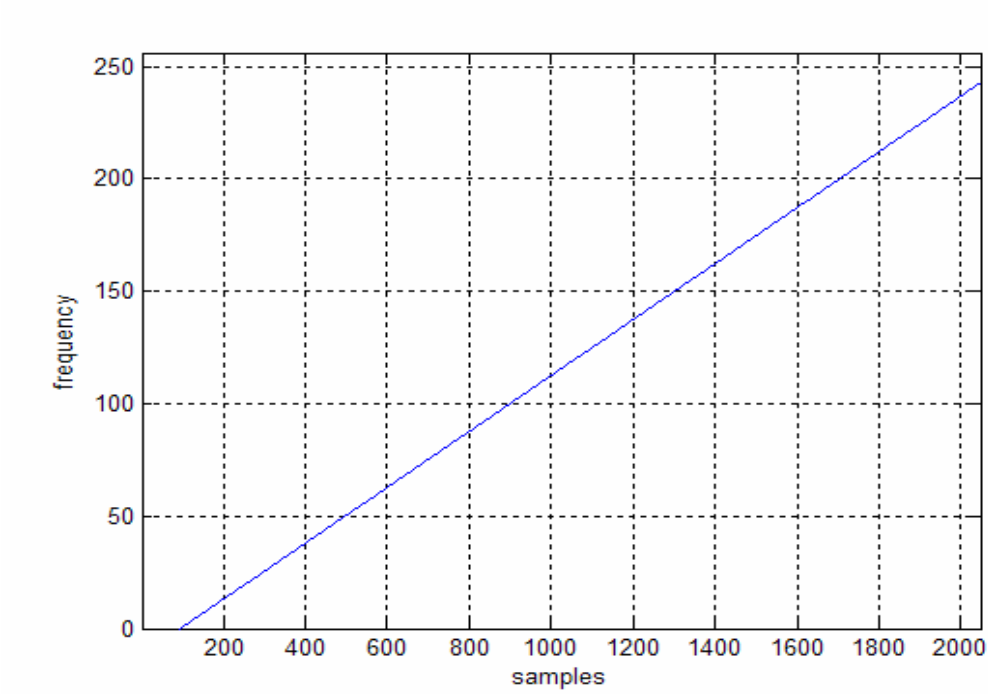

Figure 4-2 IF of the test signal.

Figure 4-3 shows the output of the two window algorithm when the SNR is set to -5 dB. The IF estimate obtained using peak detection before median filtering is shown in Figure 4-4. As seen from the figure the noise on the estimated IF is impulsive.

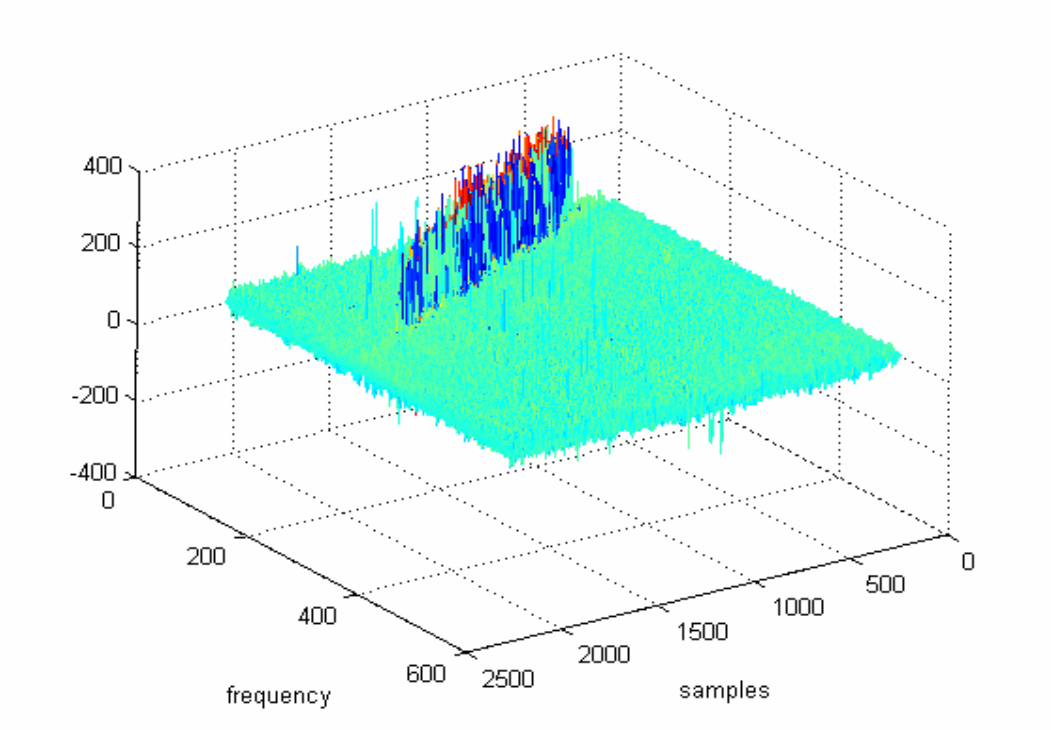

Figure 4-3 Pseudo DWVD with two window lengths, SNR = -5 dB.

The obtained IF estimate after median filtering are shown in Figure 4-5. When Figure 4-4 and Figure 4-5 are compared it is seen that the median filter smoothes the estimated IF and suppresses the noise. The estimates at the beginning and at the end of the signal are not satisfactory; this is because of the window. At these points the window does not cover the signal completely; the signal sinks in noise on the time-frequency plane and can not be detected using peak detection.

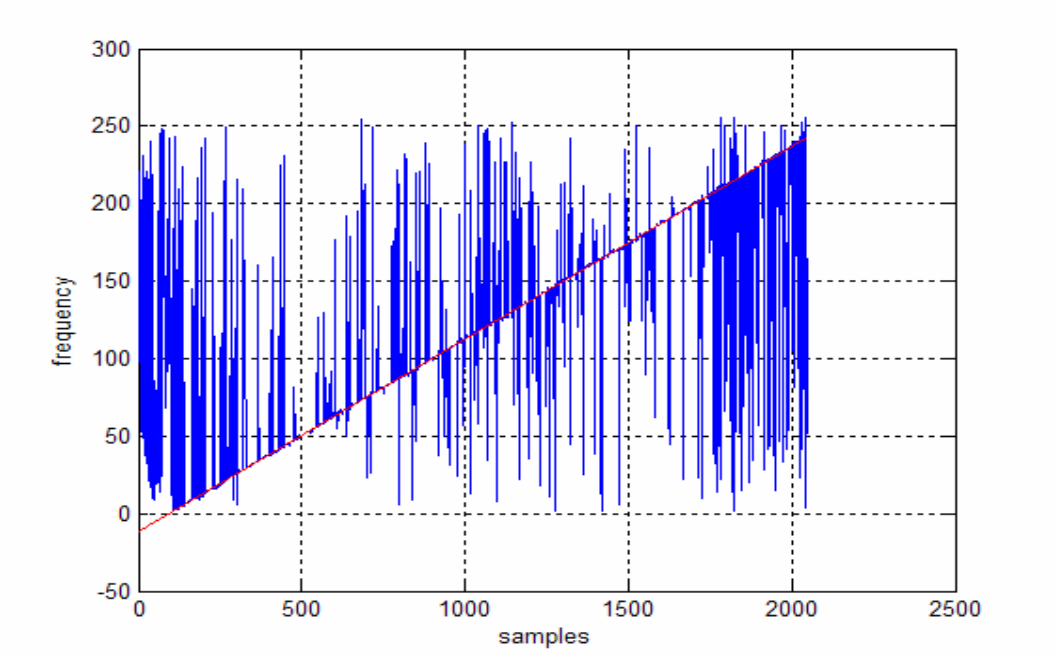

Figure 4-4 The original IF (red) and the IF estimate (blue) obtained using peak detection algorithm before median filtering, SNR = -5 dB.

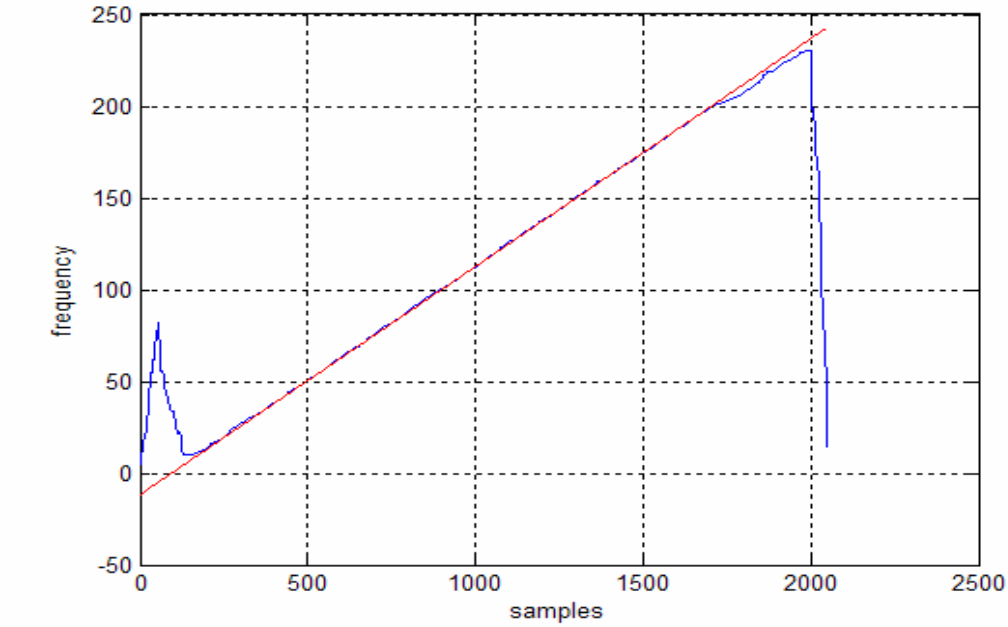

Figure 4-5 The original IF (red) and the IF estimate (blue) obtained using peak detection after median filtering, SNR = -5 dB.

The obtained results when SNR is -10 dB are shown in Figure 4-6 and Figure 4-7. The results indicate that the method is not a satisfactory when SNR is decreased to -10 dB.

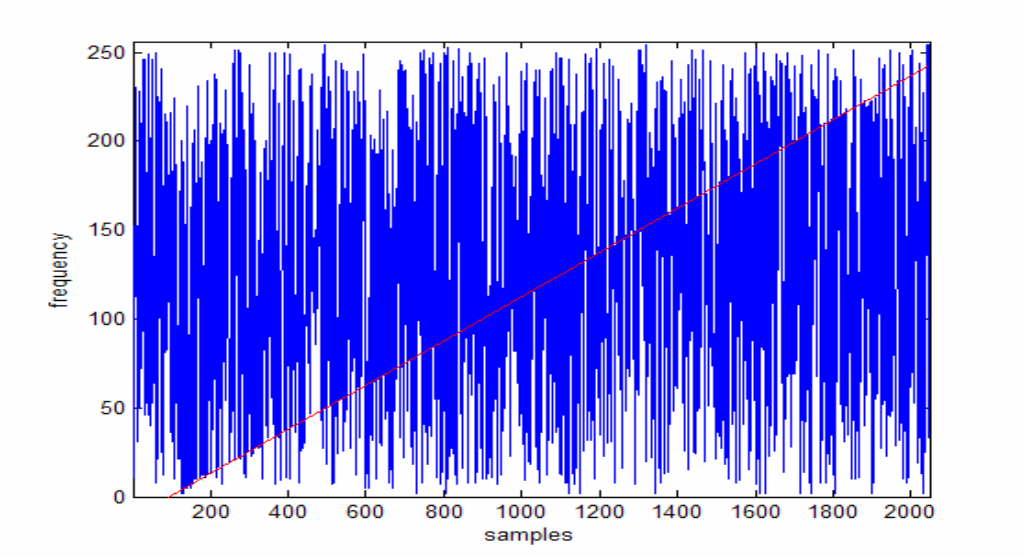

Figure 4-6 The original IF (red) and the IF estimate (blue) obtained using peak detection algorithm before median filtering, SNR = -10 dB.

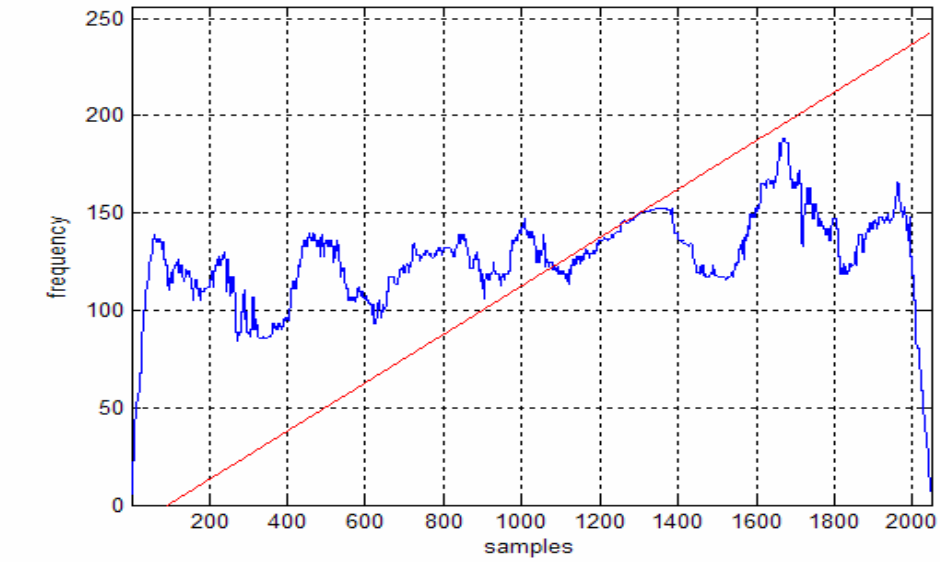

Figure 4-7 The original IF (red) and the IF estimate (blue) obtained using peak detection after median filtering, SNR = -10 dB.

4.3.3 IF Estimation Using Support Vector Machines

As it is shown in Section 4.3.2 the method of IF estimation from the output of the two window length algorithm using peak detection performs well down to SNR levels as low as -5 dB; but it does not give satisfactory results when SNR is decreased to -10 dB level. The performance of the algorithm can be improved using Support Vector Machines (SVM) which is actually a pattern recognition method.

4.3.3.1 Support Vector Machines Method

"An SVM is a general algorithm based on guaranteed risk bounds of statistical learning theory, i.e., the so-called structural risk minimization principle. It is a learning machine capable of implementing a set of functions that approximate best the supervisor's response with an expected risk bounded by the sum of the empirical risk and the Vapnik-Chervonenkis (VC) confidence. The latter is a bound on the generalization ability of the learning machine that depends on the so-called VC dimension of the set of functions implemented by the machine. SVMs can be used to solve pattern recognition, regression estimation and density estimation problems. SVMs have found numerous applications such as in optical character recognition object detection, face verification, text categorization, engine knock detection, bioinformatics, and database marketing and so on."[14].

SVMs are formed using a training data set. The training set contains the correct and false data vectors; from these vectors data pairs are formed. Each data pair contains one training vector and a constant indicating the training vector's class ("+1" indicates correct class, and "-1" indicates false class). Simply, SVM is a match filter which matches to the correct training vectors as much as possible and mismatches to the false training vectors as much as possible. To find such an SVM a quadratic equation must be solved. Depending on the training set which must resemble the signal types in the real application area, the SVM can be classified as separable, linearly nonseparable and nonlinear SVMs.

In the following the SVM for linearly nonseparable case is explained. The details about the formulations for linear separable, nonseparable and nonlinear cases can be found in [12], [13] and [14].

Let $\{x_i, y_i\}$ *for* $i = 1,2,...,L$ are given where $\mathbf{x}_i$'s are the training vectors and $y_i \in \{-1,1\}$ depends on the class of the training vector.

There exists a vector $\mathbf{w}$ which satisfies

$$\mathbf{w}^T \mathbf{x}_i + b \geq 1 - \varepsilon_i \qquad for \quad y_i = 1 \tag{4-29}$$

and

$$\mathbf{w}^T \mathbf{x}_i + b \leq -1 + \varepsilon_i \qquad for \quad y_i = -1 \tag{4-30}$$

where b is the bias term; ε_i is a nonnegative slack variable *for* $i = 1,2,...,L$; $\mathbf{w}$ is known as the support vector. We are trying to find the optimal hyperplane which maximizes the margin between itself and the vectors of two classes in the training set, in other words, which is equidistant from both classes. The so called generalized optimal hyperplane is determined by the vector $\mathbf{w}_0$ that minimizes the functional

$$J(\mathbf{w}, b, \varepsilon) = \mathbf{w}^T \mathbf{w} + C(\sum_{i=1}^{L} \varepsilon_i)^k \tag{4-31}$$

subject to

$$\begin{aligned} & y_i (\mathbf{w}^T \mathbf{x}_i + b) - 1 + \varepsilon_i \geq 0 \\ & \varepsilon_i \geq 0 \end{aligned} \tag{4-32}$$

where k is a nonnegative integer. C is a constant chosen by the user that defines the cost of constraint violations. For larger values of C the assigned penalty on the errors will be higher.

k is chosen to be 1 to get advantage that neither ε_i nor their Lagrange multipliers appear in the Wolfe dual problem.

The Lagrangian of the optimization problem is given by

$$L(\mathbf{w},b,\boldsymbol{\alpha},\varepsilon,\boldsymbol{\mu}) = \mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{L}\varepsilon_i - \sum_{i=1}^{L}\alpha_i\{y_i(\mathbf{w}^T\mathbf{x}_i + b) + \varepsilon_i - 1\} - \sum_{i=1}^{L}\mu_i\varepsilon_i \quad (4\text{-}33)$$

where μ_i's and α_i's are the nonnegative Lagrange multipliers.

To find $\mathbf{w}_0$ the Lagrange multipliers those maximize the Wolfe dual problem

$$W(\boldsymbol{\alpha}) = 1^T\boldsymbol{\alpha} - \frac{1}{4}\boldsymbol{\alpha}^T\mathbf{H}\boldsymbol{\alpha} \qquad where \qquad H_{ij} = y_i y_j (x_i^T x_j) \quad (4\text{-}34)$$

subject to $0 \le \alpha_i \le C \quad for\ i = 1,2,...,L$ and $\sum_{i=1}^{L}\alpha_i y_i = 0$ have to be found.

$\mathbf{w}_0$ is obtained as

$$\mathbf{w}_o = \frac{1}{2}\sum_{i=1}^{L}\alpha_{i,o} y_i x_i \quad (4\text{-}35)$$

The bias term b_0 is obtained as

$$b_o = \frac{1}{|I|}\sum_{i\in I}(y_i - \sum_{j=1}^{L} y_j \alpha_j x_i^t x_j) \quad (4\text{-}36)$$

where $I = \{i : 0 \le \alpha_i \le C\}$; $|I|$ denotes the cardinality of the set *I*.

4.3.3.1.1 The Proposed Algorithm

Let us say we have an image of dimensions N_1 x N_2 (WVD of the test data), and a mask matrix which has the same dimension. The image contains the auto-term and

cross-term regions. The entries of the mask matrix are ±1. Let W be the image matrix, and M be the mask matrix. If W_{nm} is an element of the auto-term region, then M_{nm} is equal to 1; if W_{nm} is not an element of the auto-term region, then M_{nm} is equal to -1. The training vectors are formed using a KxK sub-window. By moving the sub-window over the image, $L=K^2$ many values are obtained. By putting the columns of the sub-window one after another, the training vectors $x_i's$, are formed for $i = (n-1)*N_2+m$. The y_i values are obtained from M_{nm} for $i = (n-1)*N_2+m$. Then using (4-34) the α values are obtained. w_0 is obtained from (4-35), and b_o is obtained from (4-36). The equations are solved using Matlab® and its Optimization Toolbox.

After the support vector w_0 and the bias term b_o are obtained, it is applied to the output of the two window algorithm for each time-frequency point; and the ones which passes the threshold are accepted as signal region, and the ones which can not pass the threshold are accepted as noise region and masked out to obtain a clean time-frequency plane. There is one point to be careful about the bias value: it has to be updated depending on the signal energy, because the support vector is trained using a signal with unity amplitude. Then peak detection is applied to the output of the SVM method. Figure 4-8 shows the IF estimation algorithm steps with the SVM method.

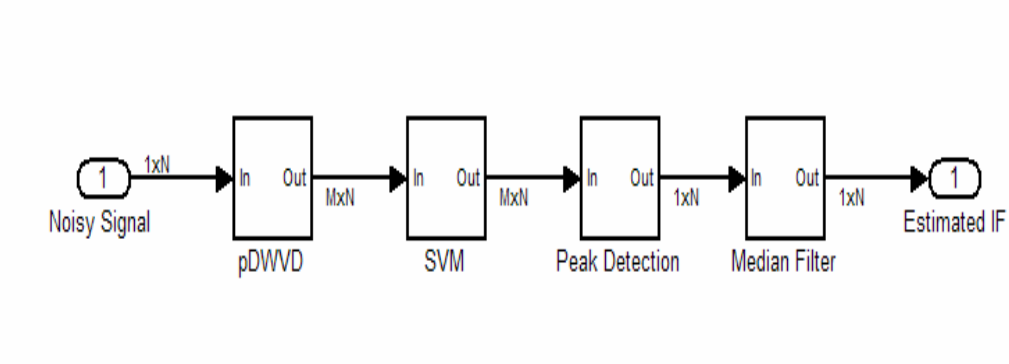

Figure 4-8 The IF estimation algorithm steps with SVM

4.3.3.2 Simulation Results of the Proposed Algorithm

The simulations are done for a linear FM signal for two different SNR levels. The window lengths are N1 = 64 and N2 = 512. The obtained IF estimates before and after median filtering are given.

4.3.3.2.1 Linear FM Signal

The same signal and the noise are used in the simulation as in Section 4.3.2.1.1.

For SNR = -5 dB case the obtained IF estimate after SVM method without median filter is shown in Figure 4-9. As seen form the figure for SNR = -5 dB this method does not need median filtering, and performs better compared to the method without SVM.

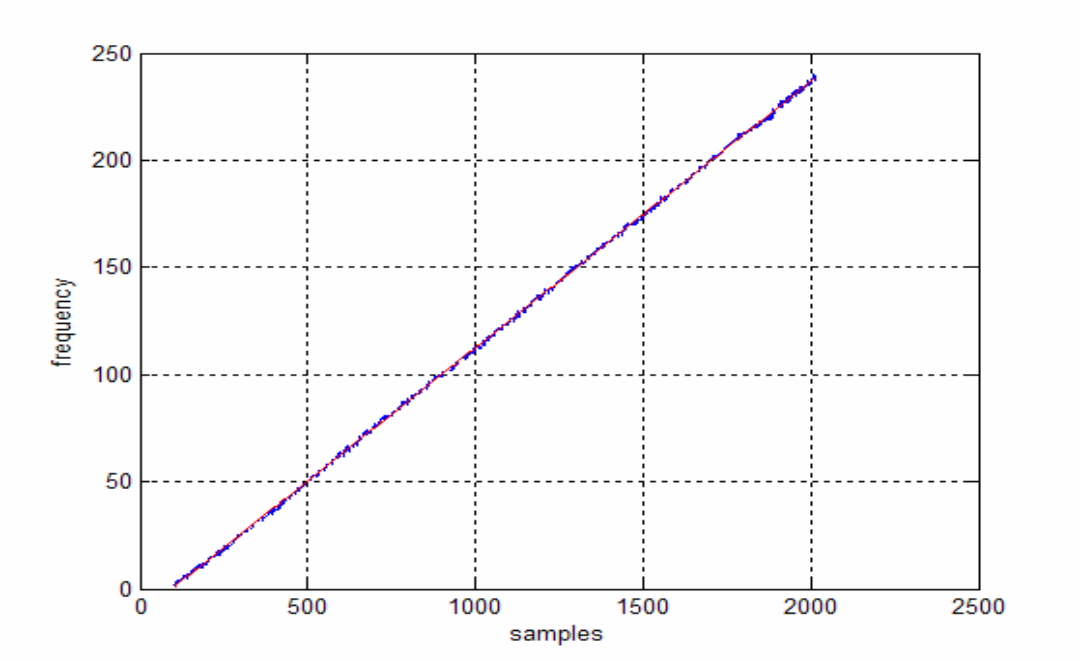

Figure 4-9 The original IF (red) and the IF estimate (blue) obtained after SVM method before median filtering, SNR = -5 dB.

For SNR = -10 dB case the obtained IF estimate after SVM method without median filter is shown in Figure 4-10. As seen from the figure the obtained IF estimates at SNR = -10 dB are comparable with the IF estimates obtained without SVM at SNR = -5 dB.

Figure 4-11 shows the IF estimates with median filtering after the SVM method. When the obtained results from the method without SVM and from the method with SVM are compared, it is seen that the SVM method makes a 5 dB improvement in IF estimation.

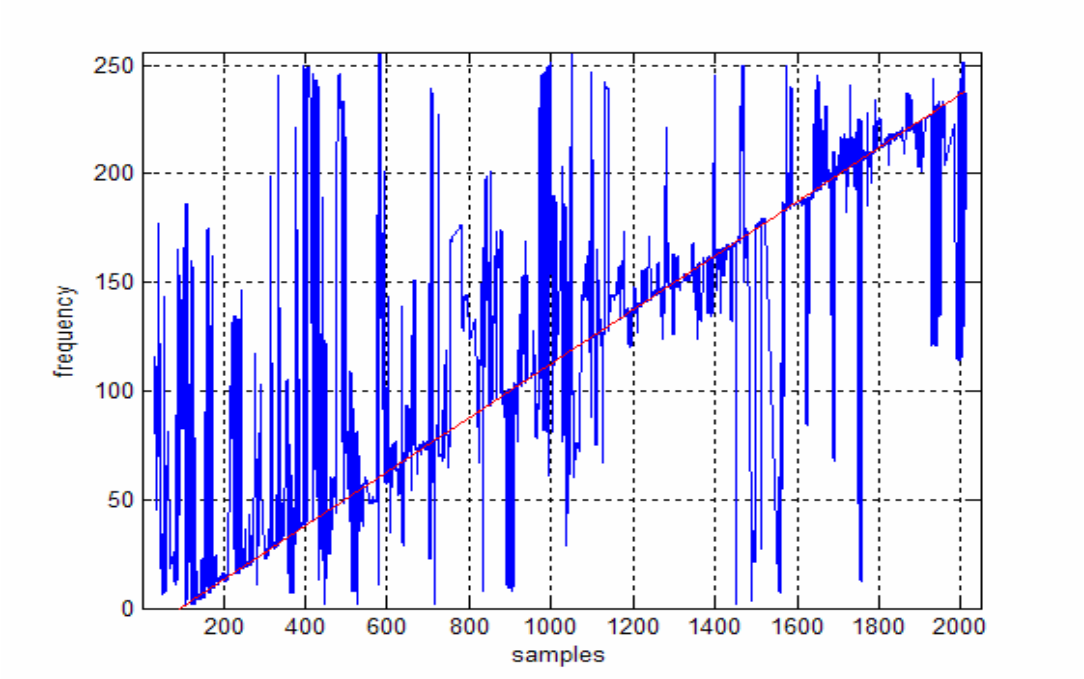

Figure 4-10 The original IF (red) and the IF estimate (blue) obtained after SVM method before median filtering, SNR = -10 dB.

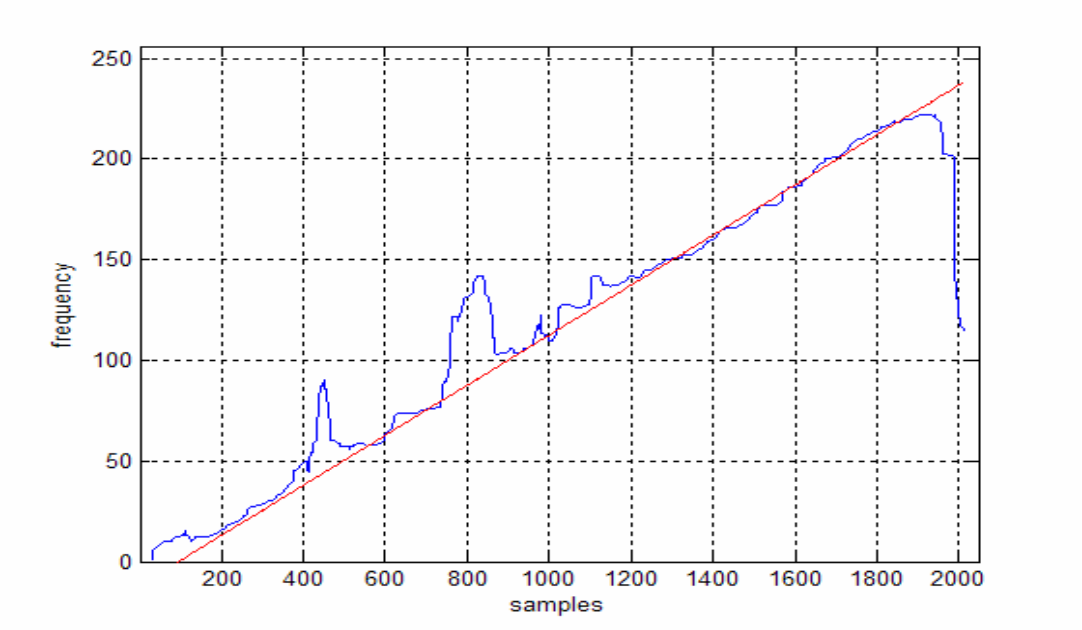

Figure 4-11 The original IF (red) and the IF estimate (blue) obtained after SVM method after median filtering, SNR = -10 dB.

4.4 THE FILTERING ALGORITHM

There are two main steps in the proposed time-varying filtering algorithm. The first step concerns the time-frequency mask, $L_H[n,k]$, computation and the second step concerns the reconstruction of the estimated signal from the time-frequency plane. Figure 4-12 shows the time-varying filtering algorithm steps.

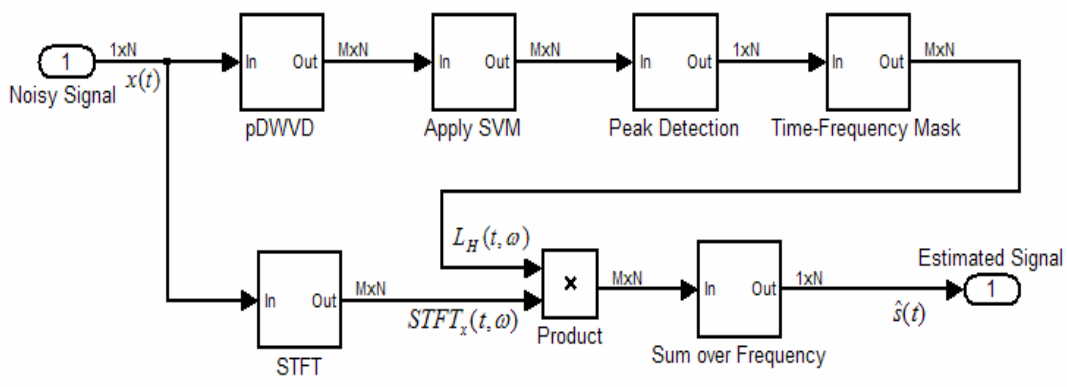

Figure 4-12 The time-varying filtering algorithm steps.

$L_H[n,k]$ is obtained from the IF estimate of the noisy signal which is computed as explained in Section 4.3. The SVM method is used to improve the performance of the IF estimation part of the algorithm. However, the SVM method does not give the IF estimate for some time samples on which the noise is dominant. To improve the filtering performance the IF estimates for these time samples are obtained using linear interpolation. The performance of the algorithm with and without the SVM method and interpolation are investigated in Section 4.4.1. The bias value of the SVM method is found by trial and error and is not changed in the simulations for different SNR values.

Estimated signal is reconstructed using equation (4-15). Even if there is no noise on the observed signal and the IF of the signal is perfectly known, the signal can not be perfectly reconstructed since the estimated signal is amplitude modulated, the reason of which is explained in [9]. (Recall that WVD concentrates signal energy at and around the IF on the time-frequency plane. At a given time instant the maximum energy is found on the IF frequency but this does not mean that the energy level at other frequencies are zero. Additionally, the energy level on the IF frequency changes with time which means that for different time samples the maximum value is different; so for a given time instant if the signal energy only at the IF frequency are summed, this will cause amplitude modulation. In fact to recover the signal at that time instant the signal energy over the entire frequency axis must be summed. To decrease this effect the time-frequency mask is widen around the IF frequency in positive and negative directions by 5 frequency samples; since most of the signal energy is concentrated around the IF on the time-frequency plane. Actually the number of frequency samples depends on the length and the type of the window used in the STFT computation. In this application a rectangular window of length 64 is used; the DFT length is 512 which results in 5 frequency samples.)

4.4.1 The Simulation Results of the Proposed Time-Varying Filtering Algorithm

The sinusoidal FM signal is used in the simulations which is obtained using the following formulas

$$IF = \sin[\,2\pi\, 0.0012207\, n]$$

$$IFc = (as - bs)(IF - \min(IF))/(\max(IF) - \min(IF)) + bs$$

$$x[n] = \exp(\, j2\pi k_f \sum_{m=0}^{n} IFc[m]) + A\, g[n]$$

where $as = 0.4$, $bs = 0.1$ and $k_f = 0.2$. $g[n]$ is the additive white Gaussian noise; A is the amplitude of the noise which is used to adjust the SNR. The simulations are done at SNR = -5 dB and SNR = -10 dB.

In the first simulation the signal is estimated from the observed signal without noise and with apriori IF information. The time-frequency mask is widen around the IF in positive and negative directions by 5 frequency samples.

Figure 4-13 shows first 850 samples of the output of the proposed algorithm when there is no noise on the signal. As it is described in Section 4.4 the signal is recovered with an amplitude modulation on it, but the frequency information is preserved. When the estimation error is defined as the difference between the estimated signal and the original signal, the output SNR with known IF is 3.8 dB.

In the second simulation the SNR = -5 dB case is tested. Figure 4-14 shows the output of the filter when the IF of the signal is known apriori. Figure 4-15 shows the output of the filter when the IF of the signal is estimated using the two window algorithm, peak detection and median filtering. As seen from the figure, the result is acceptable when compared with the IF known case. The output SNR with IF estimation is 2.6 dB.

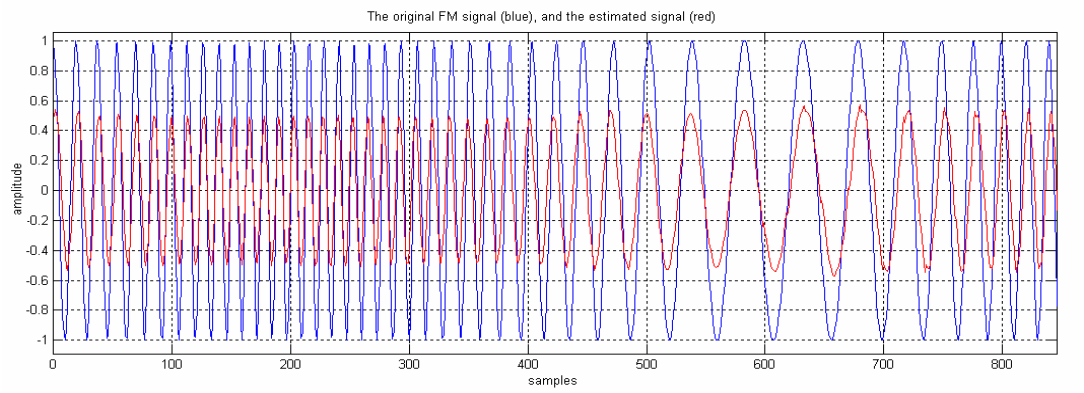

Figure 4-13 The original FM signal (blue), and the estimated signal (red).

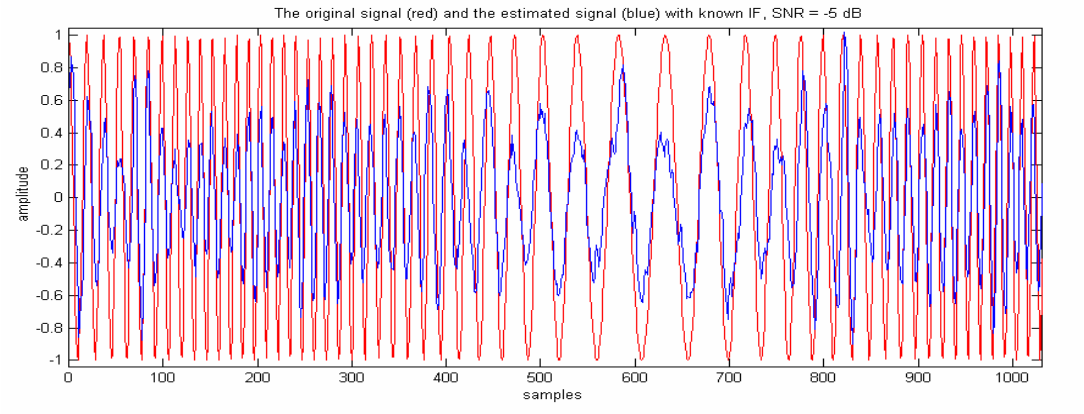

Figure 4-14 The original signal (red) and the estimated signal (blue) with known IF, SNR = -5 dB.

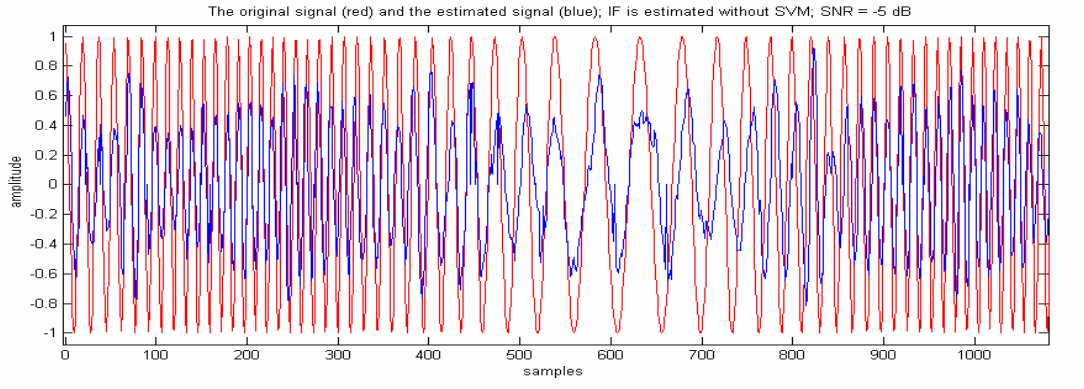

Figure 4-15 The original signal (red) and the estimated signal (blue); IF is estimated without SVM; SNR = -5 dB.

Figure 4-16 shows the pDWVD after SVM Method is applied. As seen from the figure, a cleaner time-frequency plane is obtained by the help of the SVM Method; however there are gaps along the time axis, which are the high noise influence points and eliminated by SVM Method; so for these time samples IF estimates are not computed.

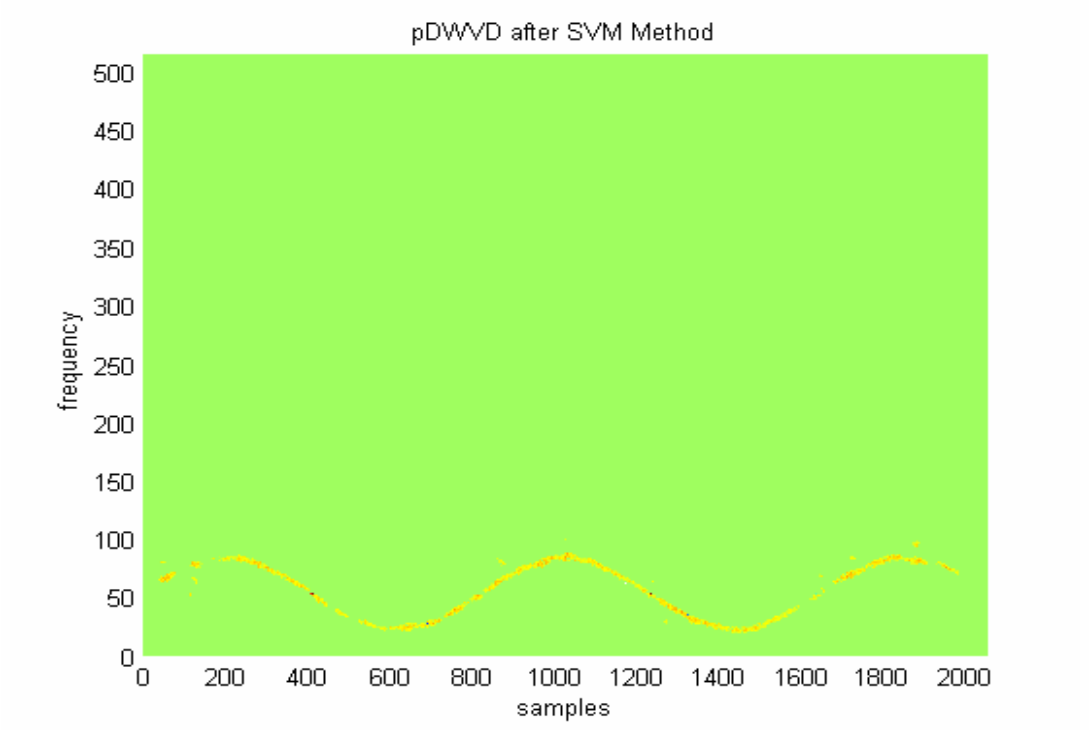

Figure 4-16 pDWVD after SVM method is applied, SNR = -5 dB.

When IF estimate obtained from SVM method is directly applied, the filter output is shown in Figure 4-17. As seen from the figure for some time durations the filter output is zero; these points are the eliminated points by SVM Method. Although these points are eliminated by SVM Method this does not mean that the signal is absent for these points; because the duration of the gaps is smaller compared to the window widths used in two-window algorithm; so performance of the algorithm can be improved if IF estimates for those time samples are obtained. This is done by linear interpolation; that is to say the jumps in time samples for the IF estimates are

detected and the IF estimates are calculated using linear interpolation for those time duration. The obtained result using linear interpolation is shown in Figure 4-18; as seen from the figure those time gaps are filled. The output SNR with IF estimation using SVM Method is 2.9 dB and the output SNR with IF estimation using SVM Method with linear interpolation is 3.6 dB. When the obtained SNR values are compared it is seen that the results of IF known case and IF estimated using SVM Method with linear interpolation case are nearly equal.

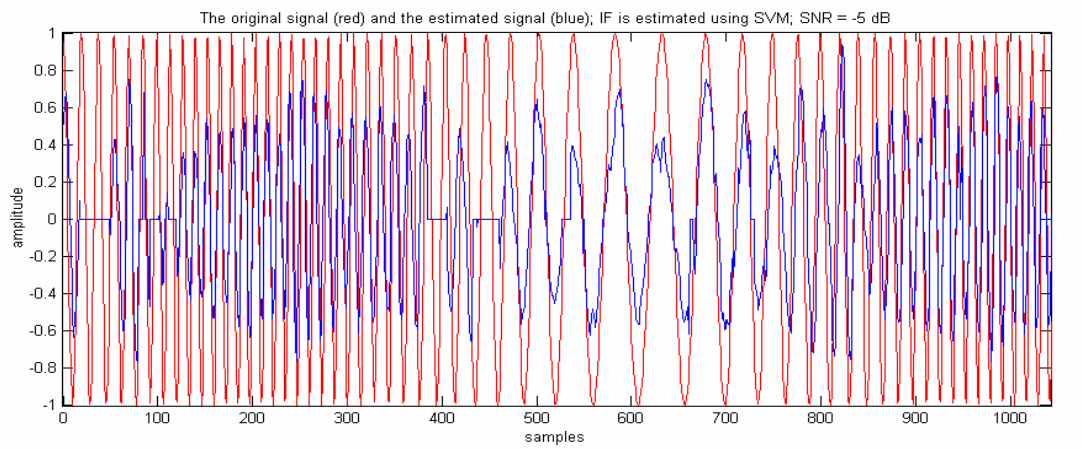

Figure 4-17 The original signal (red) and the estimated signal (blue); IF is estimated using SVM; SNR = -5 dB.

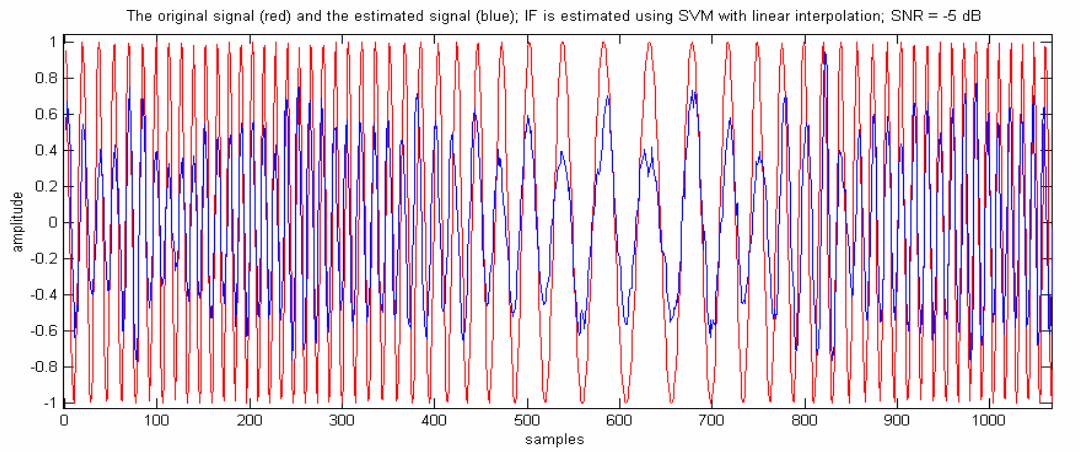

Figure 4-18 The original signal (red) and the estimated signal (blue); IF is estimated using SVM with linear interpolation; SNR = -5 dB.

In the third simulation, SNR = -10 dB case is tested. Figure 4-19 shows the IF estimate obtained from two window algorithm (blue signal), and after median filter (red signal) without SVM method. The green signal is IF of the signal, as seen from the figure and as it is shown in Section 4.3.2.1.1 the algorithm does not work for SNR = -10 dB case.

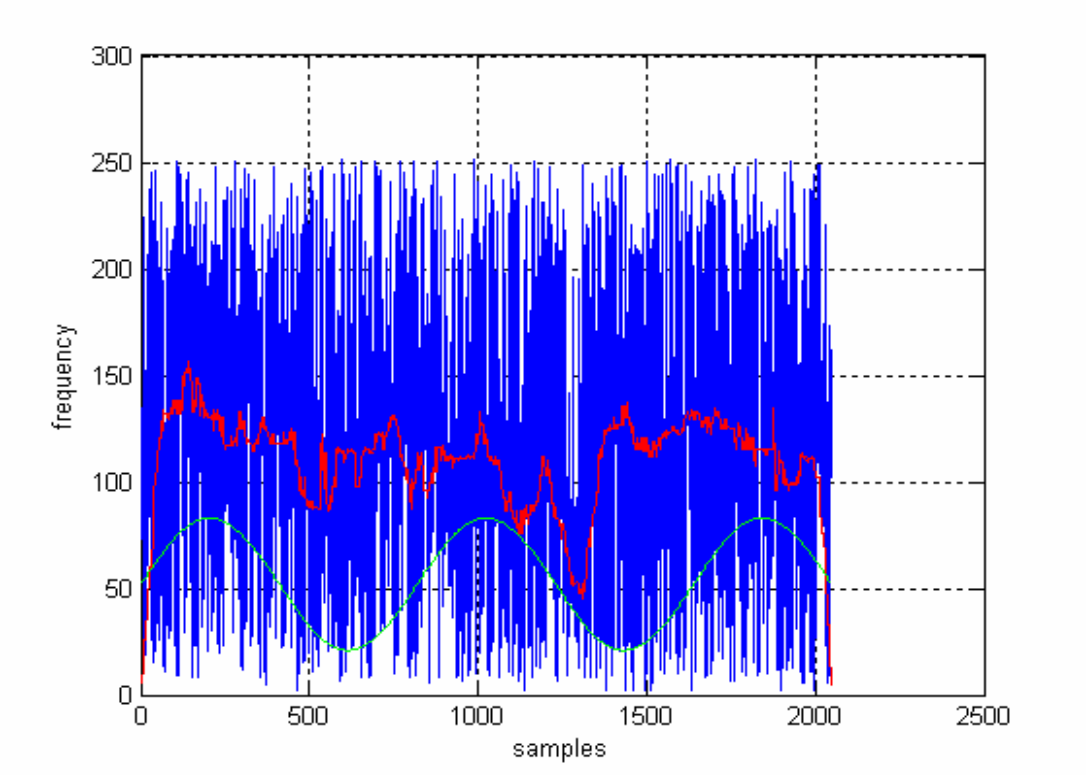

Figure 4-19 IF estimate obtained without SVM (blue); after median filter (red); original IF (green); SNR = -10 dB.

Figure 4-20 shows the output of the filter when the IF of the signal is known apriori. The output SNR of the IF known case is 3.3 dB.

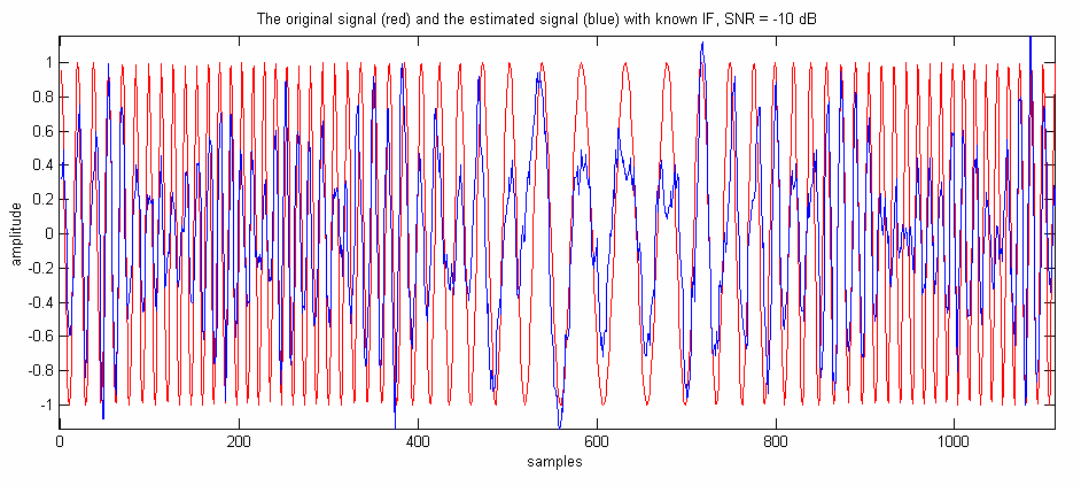

Figure 4-20 The original signal (red) and the estimated signal (blue) with known IF, SNR = -10 dB.

Figure 4-21 shows the pDWVD after SVM Method is applied as seen from the figure a cleaner time-frequency plane is obtained; however there are more time gaps compared to SNR= -5 dB case. This is also seen from Figure 4-22 which shows the filter output when IF estimate obtained from SVM Method is directly used in mask computation. The output SNR of the IF estimated using SVM Method case is 1.9 dB. Figure 4-23 shows the output of the filter when the missing IF estimates are obtained using linear interpolation; as seen from the figure the time gaps are filled after linear interpolation which improves performance of the algorithm. The output SNR of the IF estimated using SVM Method with linear interpolation case is 2.6 dB. When the output SNR obtained from IF estimated using SVM Method with linear interpolation case is compared with the output SNR obtained from IF known case, it is seen that the difference is less than 1 dB.

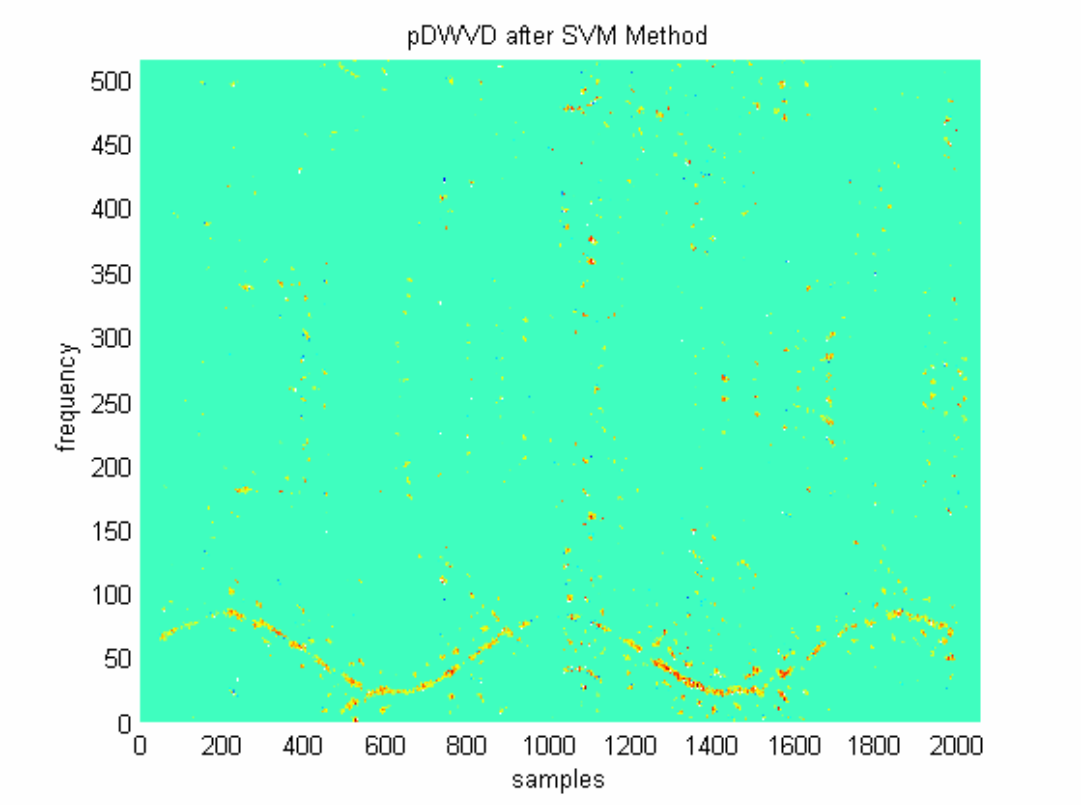

Figure 4-21 pDWVD after SVM Method is applied, SNR = -10 dB.

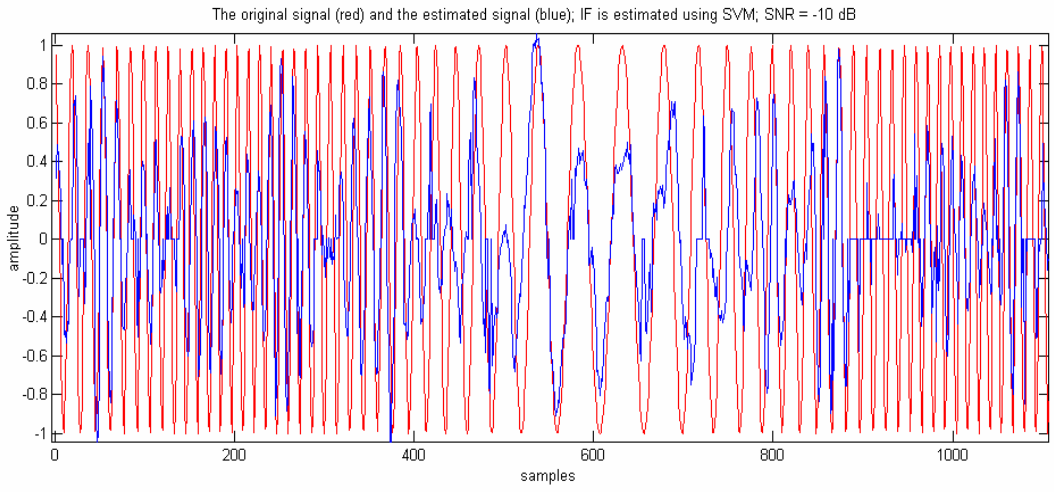

Figure 4-22 The original signal (red) and the estimated signal (blue); IF is estimated using SVM; SNR = -10 dB.

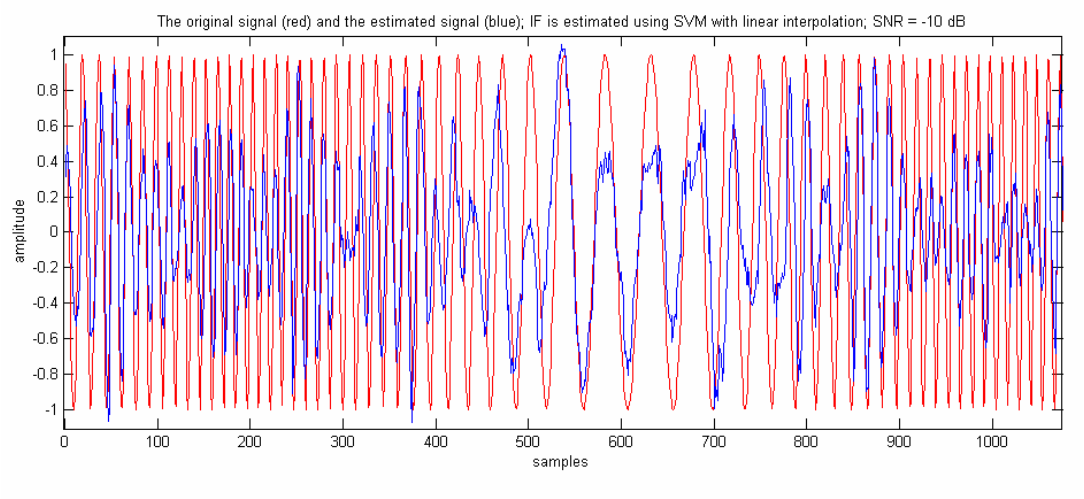

Figure 4-23 The original signal (red) and the estimated signal (blue); IF is estimated using SVM with linear interpolation; SNR = -10 dB.

4.5 CONCLUSION

In this chapter the derivations of the time-varying Wiener Filter and its time-frequency formulation using WVD are given. According to the derived formulas, it is seen that to perform filtering operation, time-varying transfer function of the linear time-varying filtering operator and the STFT of the noisy signal have to be computed. The computation of the latter is simple; however, the determination of the time-varying transfer function is crucial, because if it is incorrectly determined, the desired signal can not be filtered from noise. It is shown that the time-varying transfer function is approximately given by the region of support of the desired signal; hence the determination of time-varying transfer function is equivalent to the determination of region of support of the signal, which is equivalent to the determination of IF of the signal. (Recall that WVD distribution concentrates the signal energy at and around the IF of the signal on the time-frequency plane.)

To determine the IF of the signal, the method which is proposed in [1] is investigated. It is shown that the method gives satisfactory results at SNR levels down to -5 dB. To improve the performance of the method, the use of SVM is

proposed. It is shown that with SVM without any need of median filtering, the performance of the algorithm is better at SNR levels down to -5 dB. Furthermore, it is shown that with SVM and median filtering the algorithm gives satisfactory results at SNR levels down to -10 dB.

The main disadvantage of SVM method is that for some time samples it makes incorrect decisions and does not give the IF estimates. For these time samples the output of the filter is zero which degrades the performance of the algorithm. To overcome this disadvantage the missing IF estimates are obtained by linear interpolation. To do so the jumps in time samples for IF estimates are detected and these time gaps are filled by a linear curve. This improves the performance of the algorithm, and it is shown that at SNR levels down to -5 dB the output of this algorithm is comparable with the output of the filter with apriori known IF.

Another disadvantage of the SVM method is the bias value determination. In the training phase of the method a bias value is computed for optimum decision; however, this value depends on the training signal energy. When the signal energy changes, the bias value has to be updated. In this work, a bias value is set by trial and error, which degrades the decision performance.

The use of SVM method can improve the filtering performance for mono-component signals. However, the SVM method increases the computational cost; since for every time-frequency point a decision has to be made. The performance of the algorithm is not tested for multi-component signals, since the use of peak detection allows the detection of the stronger signal. However, if the number of signal components is known, peak detection can also be used for multi-component signals.

CHAPTER 5

COMPARISON OF METHODS

The performances of TFPF method, SVM based method and classical zero-phase low-pass filter are compared using real, frequency modulated test signals. The performances are evaluated by comparing the output SNR values of the methods under the same input SNR.

Let the continuous time test signal be of the form

$$s(t) = \cos(\phi(t)) \tag{5-1}$$

The instantaneous frequency (IF) of $s(t)$ is given by

$$f_i(t) = \frac{1}{2\pi}\frac{d}{dt}\phi(t) \tag{5-2}$$

Consider an FM signal

$$s(t) = \cos(2\pi f_c t + 2\pi k_f \int_0^t m(t')dt') \tag{5-3}$$

Where k_f is the frequency sensitivity; the IF of the FM signal is

$$f_i(t) = f_c + k_f m(t) \tag{5-4}$$

The frequency deviation Δf is defined as

$$\Delta f = k_f \max(m(t)) \tag{5-5}$$

which represents the maximum deviation of the IF of the FM signal from the carrier frequency f_c. The ratio of the frequency deviation to the bandwidth B_m of $m(t)$ is defined as the modulation index β which is given by

$$\beta = \Delta f / B_m \tag{5-6}$$

The bandwidth of an FM is signal is infinite, since it contains an infinite number of side frequencies which are separated from the carrier frequency by integer multiples of B_m in positive and negative directions. However the side frequencies that are separated from the carrier frequency by an amount greater than the total frequency deviation decrease toward zero rapidly. So, the bandwidth of the signal is grater than the frequency deviation but it is eventually limited. For large values of β (compared to 1) the bandwidth of $s(t)$ (the FM signal) approaches and is slightly greater than the total frequency deviation $2\Delta f$; and this case is known as wide-band FM. On the other hand, for small values of β (compared to 1), the spectrum of $s(t)$ is effectively limited to the band $f_c \pm B_m$, and bandwidth of $s(t)$ approaches $2B_m$; and this case is known as narrow-band FM. So, an approximate formula for the effective bandwidth of $s(t)$ can be defined as

$$B_s \cong 2(\Delta f + B_m) \tag{5-7}$$

which is known as Carson's rule.

In the following simulations discrete time test signals are used which are obtained from the continuous time signal which is contained in additive zero-mean white Gaussian noise by sampling the signal with sampling period T.

The discrete time test signal is of the form

$$x[n] = s[n] + v[n] \tag{5-8}$$

where $s[n]$ is the desired signal with average power P_s and $v[n]$ is the additive zero-mean white Gaussian noise with variance σ_v^2, the SNR at the filter input is

$$SNR_i = 10\log_{10}(P_s / \sigma_v^2) \tag{5-9}$$

The filtered signal is $\hat{x}[n]$ and the error signal is $e[n] = s[n] - \hat{x}[n]$ with variance σ_e^2, the SNR at the filter output is

$$SNR_o = 10\log_{10}(P_s / \sigma_e^2) \tag{5-10}$$

5.1 SIMULATION 1: Linear FM Signal, High Sampling Frequency

The continuous time noise free test signal is

$$s(t) = \begin{cases} \cos(0.005t + 7.5 * 10^{-7} t^2) & for \quad 0 \le t \le 4096\,\text{sec} \\ 0 & otherwise \end{cases} \tag{5-11}$$

The effective bandwidth of the signal which is calculated using Carson's Rule is 0.0018Hz. Figure 5-1 shows the magnitude spectrum of the noise free signal, which is computed using Mathametica software. As seen from the figure, the calculated effective bandwidth of the signal is a sufficient bandwidth approximation for this signal.

It is assumed that the signal is observed only for the time duration $0 \le t \le 4096\,\text{sec}$; and sampled with sampling period $T = 1\,\text{sec}$ which is approximately 250 times the Nyquist rate.

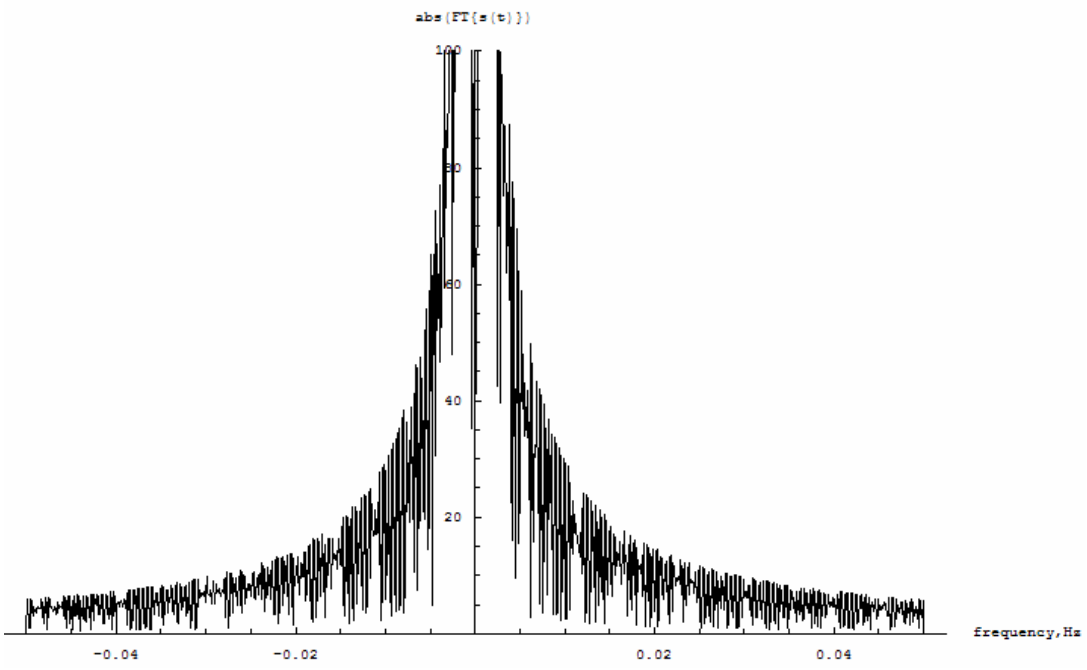

Figure 5-1 Magnitude spectrum of the noise free signal.

The sampled test signal with additive zero-mean white Gaussian noise is of the form

$$x[n] = 2\cos\left[0.005n + 7.5 * 10^{-7} n^2\right] + v[n] \, for \quad n = 0:4095 \tag{5-12}$$

The variance of $v[n]$ is adjusted such that $SNR_i = -5dB$.

Figure 5-2 shows the output of the TFPF Method; the output SNR is approximately 12 dB. Figure 5-3 shows the output of the SVM based Method; the output SNR is approximately 8 dB. Figure 5-4 shows the output of the zero-phase low-pass filter; the output SNR is approximately 7 dB. When the output SNR values are compared it is seen that TFPF Method performs better compared to the others methods for this signal and sampling rate.

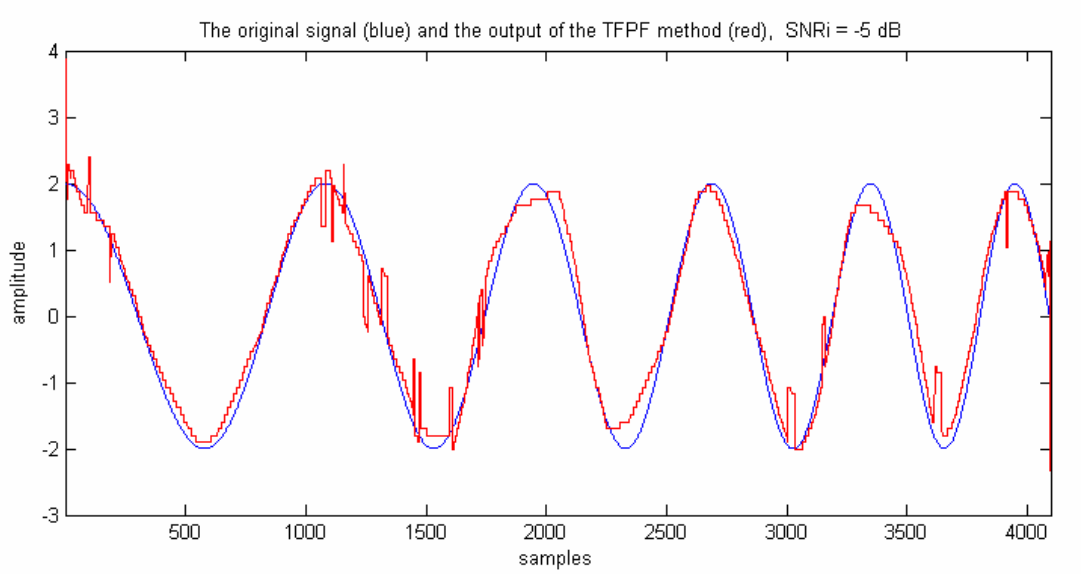

Figure 5-2 The original signal (blue) and the output of the TFPF method (red), SNR_i = -5 dB.

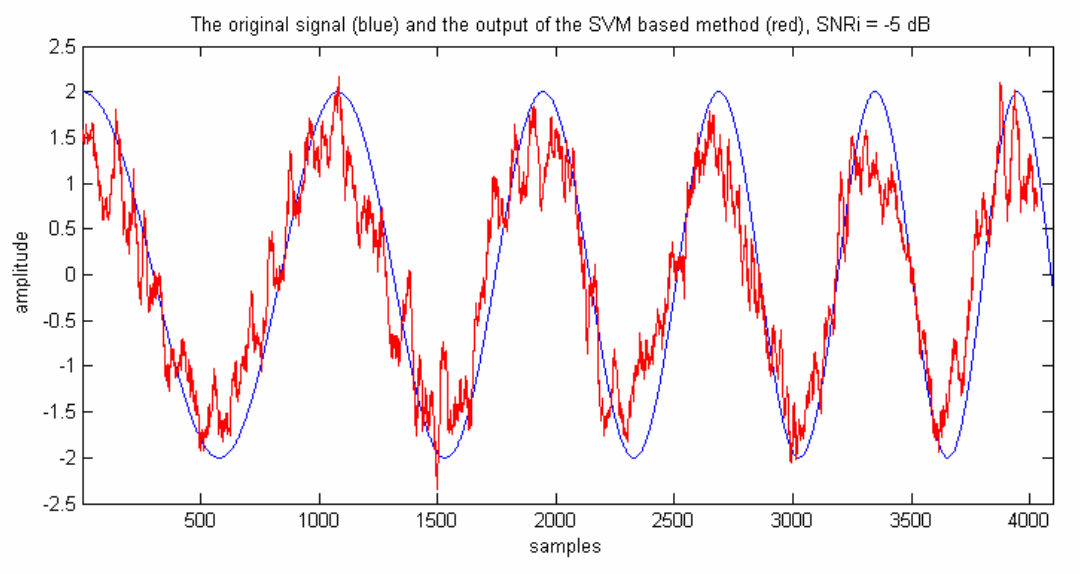

Figure 5-3 The original signal (blue) and the output of the SVM based method (red), SNR_i = -5 dB.

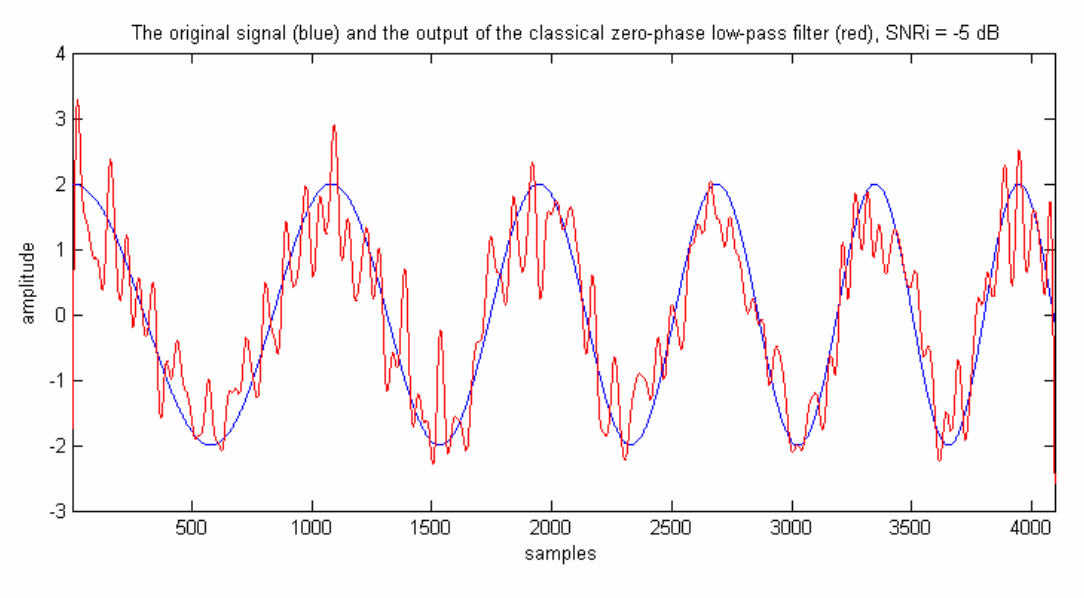

Figure 5-4 The original signal (blue) and the output of the classical zero-phase low-pass filter (red), SNR_i = -5 dB.

5.2 SIMULATION 2: Linear FM Signal, Low Sampling Frequency

The continuous time noise free test signal is

$$s(t) = \begin{cases} \cos(0.01t + 0.7754/4096t^2) & for \quad 0 \le t \le 2048\,\text{sec} \\ 0 & otherwise \end{cases} \tag{5-13}$$

The effective bandwidth of the signal which is calculated using Carson's Rule is 0.125Hz. Figure 5-5 shows the magnitude spectrum of the noise free signal, which is computed using Mathametica software. As seen from the figure, the calculated effective bandwidth of the signal is a sufficient bandwidth approximation for this signal.

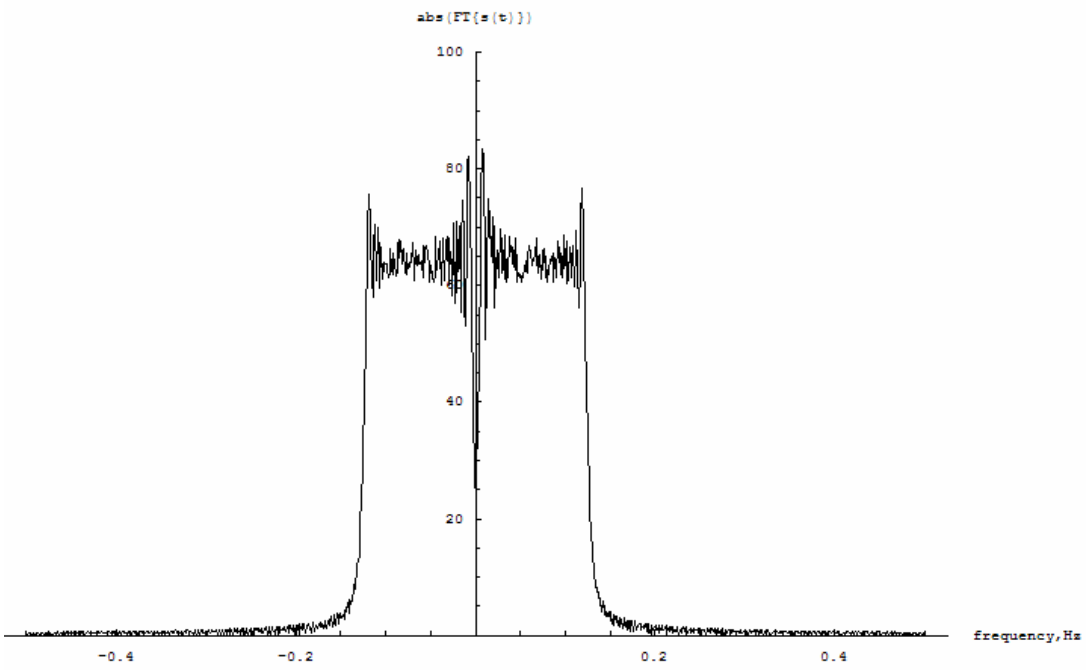

Figure 5-5 Magnitude spectrum of the noise free signal.

It is assumed that the signal is observed only for the time duration $0 \le t \le 2048\,\text{sec}$; and sampled with sampling period $T = 1/4\,\text{sec}$ which is approximately 16 times the Nyquist rate.

The sampled test signal with additive zero-mean white Gaussian noise is of the form

$$x[n] = 2\cos\left[\frac{0.01}{4}n + \frac{0.7754}{4096*16}n^2\right] + v[n]\,for \quad n = 0:8191 \qquad (5\text{-}14)$$

The variance of $v[n]$ is adjusted such that $SNR_i = -5dB$.

The maximum IF of the signal is 0.125Hz, and the worst case window length of the TFPF method is 16.

Figure 5-6 shows the output of the TFPF method. The output SNR is 5 dB.

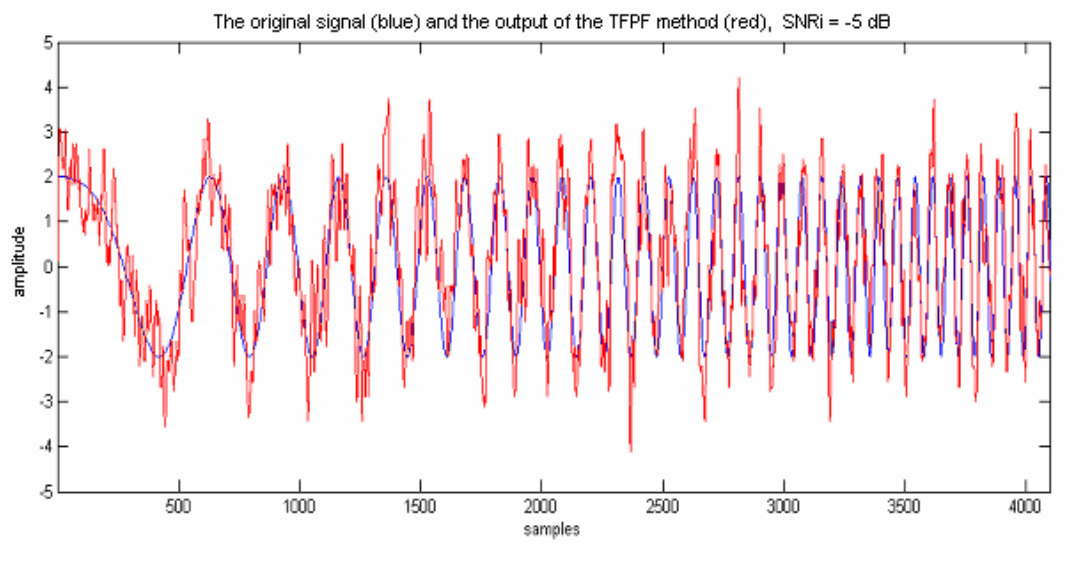

Figure 5-6 The original signal (blue) and the output of the TFPF method (red), SNR_i = -5 dB.

Figure 5-7 shows the output of the SVM based method. The output SNR is approximately 4.9 dB.

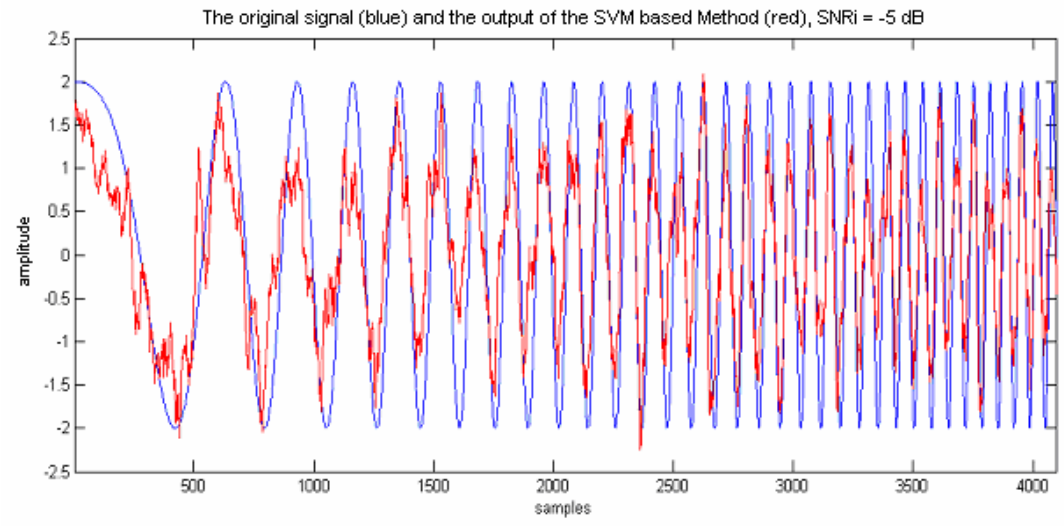

Figure 5-7 The original signal (blue) and the output of the SVM based Method (red), SNR_i = -5 dB.

Figure 5-8 shows the output of the classical zero-phase low-pass filter. The output SNR is approximately 5 dB.

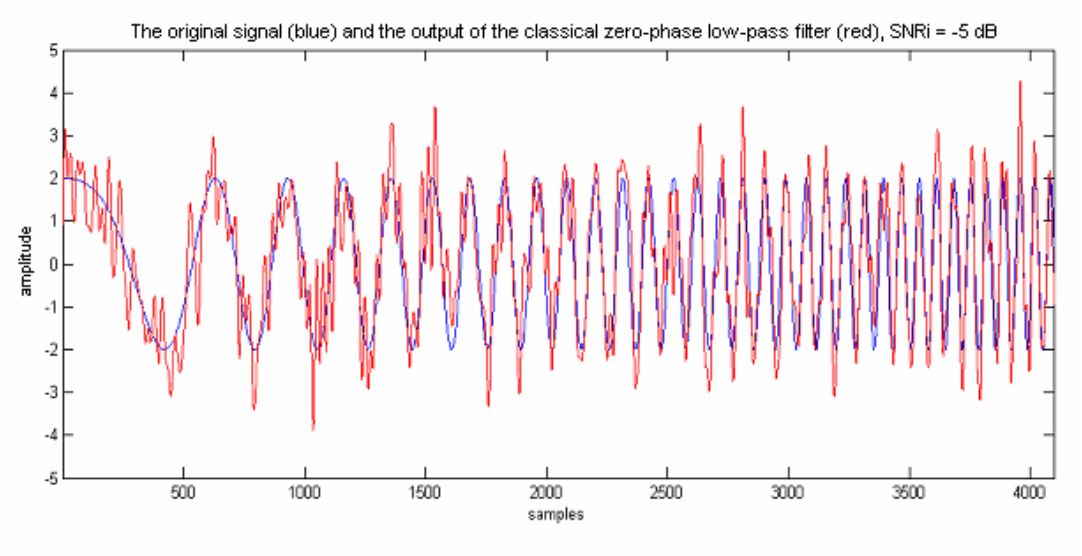

Figure 5-8 The original signal (blue) and the output of the classical zero-phase low-pass filter (red), SNR_i = -5 dB.

It is now assumed that the signal is observed only for the time duration $0 \le t \le 2048\,\text{sec}$; and sampled with sampling period $T = 1\text{sec}$ which is approximately 4 times the Nyquist rate. For this simulation the sampling rate is decreased to see the effects of the sampling rate on the performances of the algorithms.

The sampled test signal with additive zero-mean white Gaussian noise is of the form

$$x[n] = 2\cos\left[0.01n + \frac{0.7754}{4096}n^2\right] + v[n]\,for \quad n = 0:2047 \qquad (5\text{-}15)$$

The variance of $v[n]$ is adjusted such that $SNR_i = -5dB$.

Notice that the sampling rate is lowered 4 times but it is still 4 times higher than the Nyquist rate and still avoids aliasing. The same simulations are done for this lower sampling rate. This time the worst case window length of TFPF method is reduced to 4 from 16.

Figure 5-9 shows the output of the TFPF Method; the output SNR is approximately -1.7 dB. Figure 5-10 shows the output of the SVM based method; the output SNR is approximately 4.2 dB. Figure 5-11 shows the output of the zero-phase low-pass filter; the output SNR is approximately 0 dB.

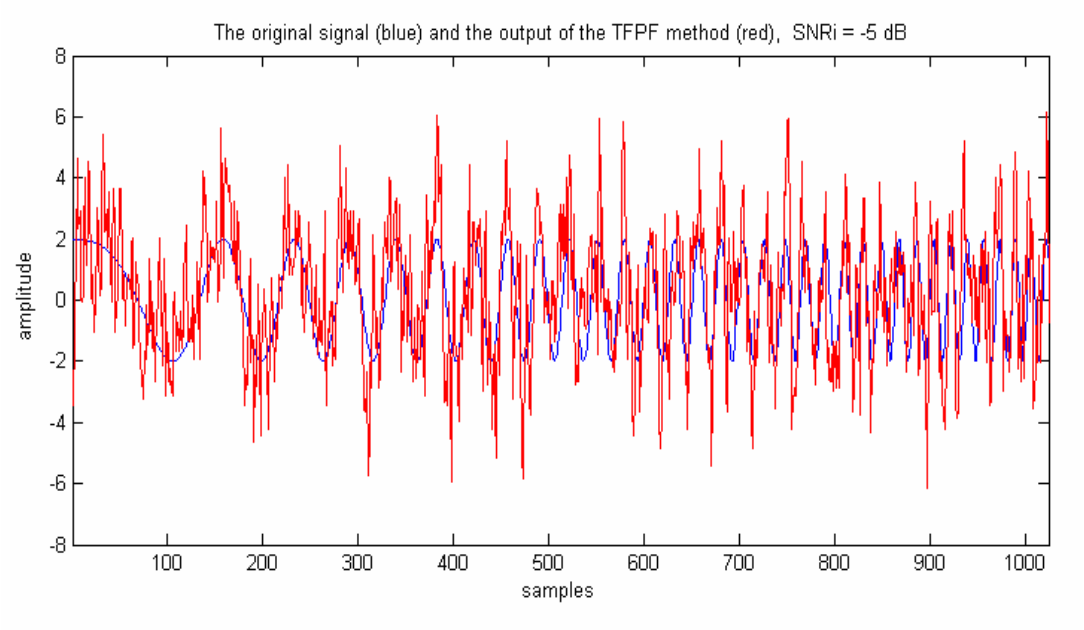

Figure 5-9 The original signal (blue) and the output of the TFPF method (red), SNR_i = -5 dB.

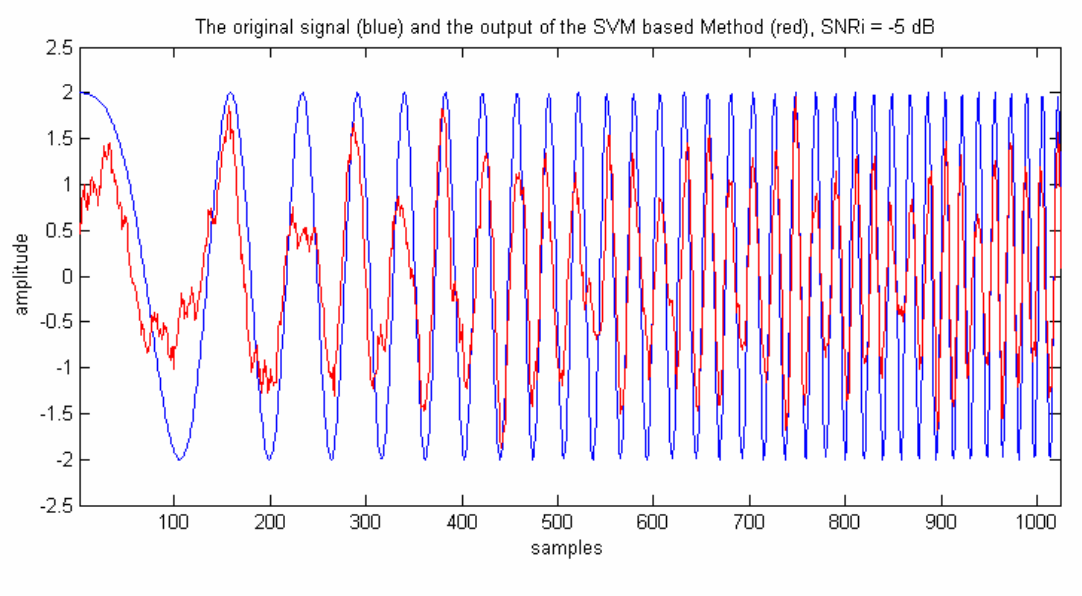

Figure 5-10 The original signal (blue) and the output of the SVM based Method (red), SNR_i = -5 dB.

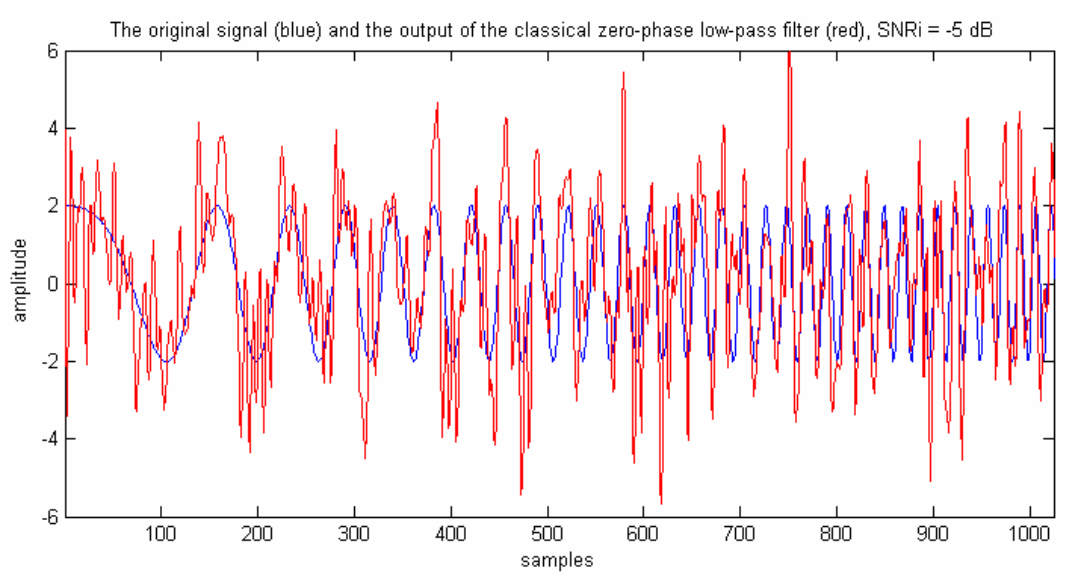

Figure 5-11 The original signal (blue) and the output of the classical zero-phase low-pass filter (red), SNR_i = -5 dB.

As seen from the obtained results when the sampling rate is decreased to $1/4$ of the original sampling rate, the output SNR of the TFPF Method decreases by approximately 7 dB, the output of zero-phase low-pass filter decreases by approximately 5 dB; on the other hand the output SNR of the SVM based method does not change more than 1 dB.

5.3 SIMULATION 3: 4-FSK Signal

The continuous time noise free test signal is a 4-FSK signal which is of the form

$$s(t) = 2\sum_{k=1}^{P} \cos(2\pi f_c t + a_k 2\pi f_k t) w(t - kT_c) \qquad (5\text{-}16)$$

where P is the number of hops which is set to 8 in the simulation, f_c is the carrier frequency which is set to 0, T_c is the duration between hops which is set to 40ms, f_k is the minimum frequency change which is set to 100Hz, $w(t)$ is a rectangular pulse of length T_c, and a_k's are integers in between 1 and 4 for 4-FSK which are set to 1, 2, 1, 4, 4, 3, 1 and 2.

The effective bandwidth of the signal which is calculated using Carson's Rule is approximately 400Hz. Figure 5-12 shows the magnitude spectrum of the noise free signal, which is computed using Mathametica software. As seen from the figure, the calculated effective bandwidth of the signal is a sufficient bandwidth approximation for this signal.

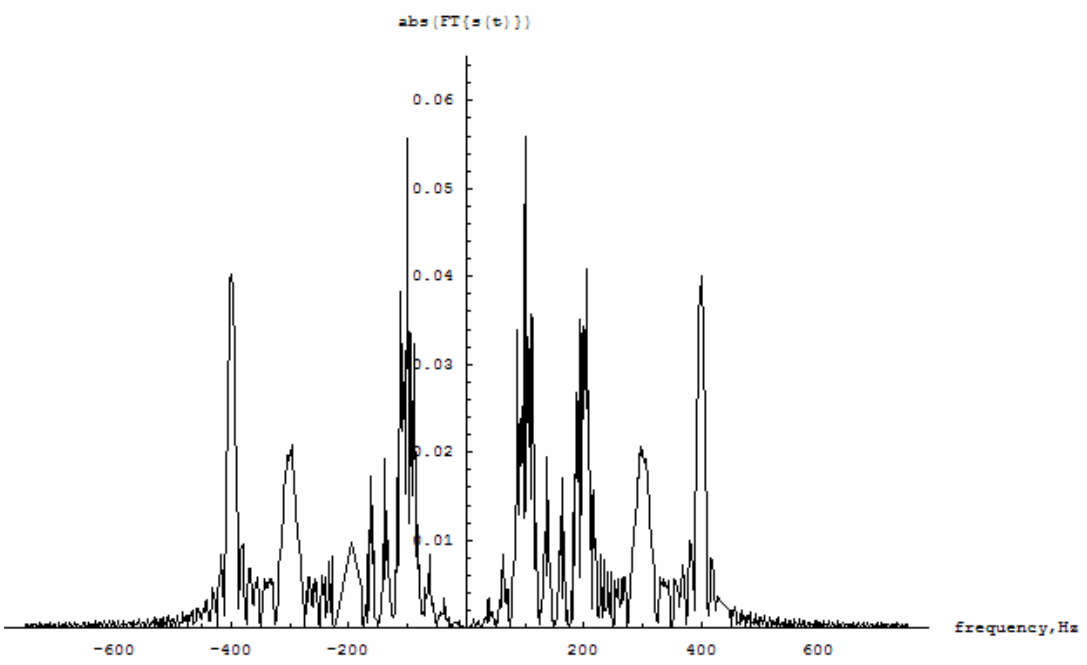

Figure 5-12 Magnitude spectrum of the noise free signal.

It is assumed that the signal is observed only for the time duration $0 \le t \le 320ms$; and sampled with sampling period $T = 50\mu\sec$. The sampled test signal with additive zero-mean white Gaussian noise is of the form

$$s[n] = 2\sum_{k=1}^{P} \cos(a_k 2\pi f_k 50e-6*n) w(50e-6*t-kT_c) + v[n] \quad for \quad n = 0:6399 \quad (5\text{-}17)$$

The variance of $v[n]$ is adjusted such that $SNR_i = -5dB$.

The maximum IF of the signal is 400Hz, and the worst case window length of the TFPF method is 20.

Figure 5-13 shows the output of the TFPF method. The output SNR is 6.8 dB.

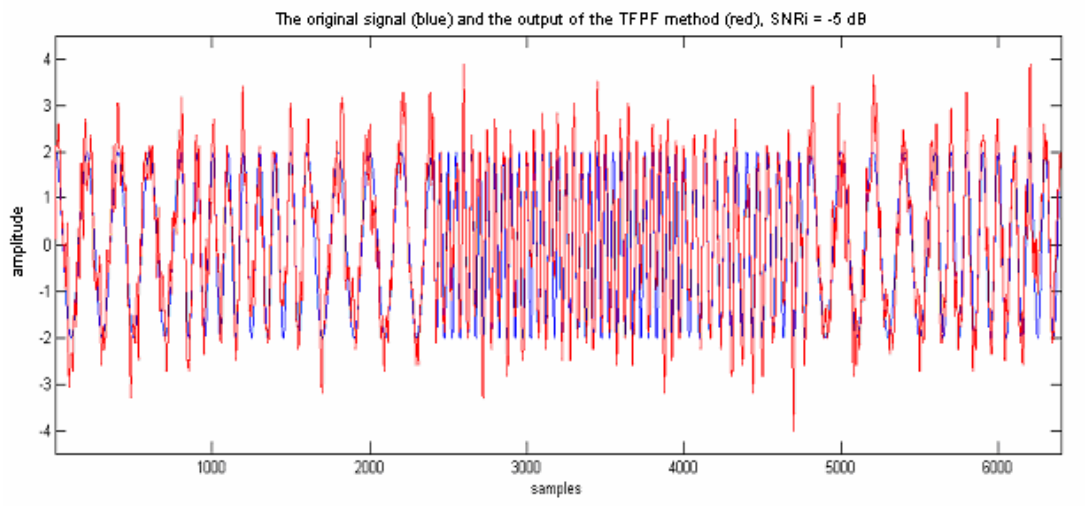

Figure 5-13 The original signal (blue) and the output of the TFPF method (red), SNR_i = -5 dB.

Figure 5-14 shows the output of the SVM based method, the output SNR is 4.9 dB. Figure 5-15 shows the output of the classical zero-phase low-pass filter. The output SNR is 4.8 dB.

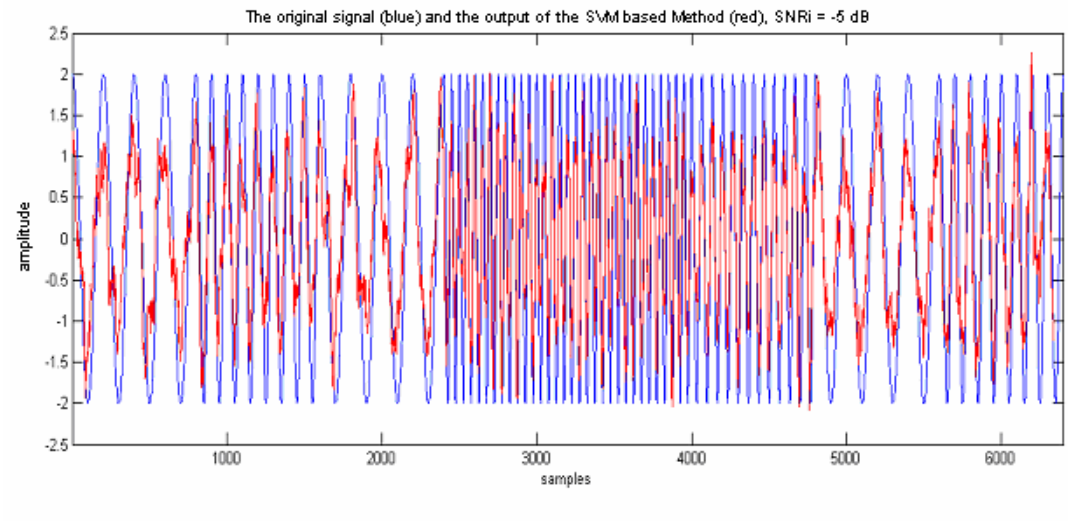

Figure 5-14 The original signal (blue) and the output of the SVM based Method (red), SNR_i = -5 dB.

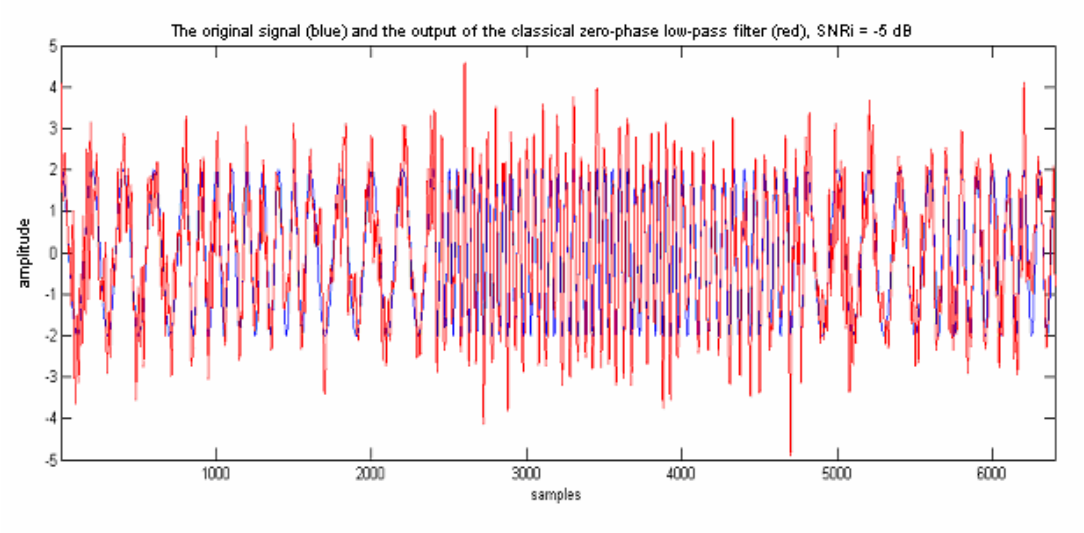

Figure 5-15 The original signal (blue) and the output of the classical zero-phase low-pass filter (red), SNR_i = -5 dB.

It is now assumed that the signal is observed only for the time duration $0 \le t \le 320ms$; and sampled with sampling period $T = 100\mu \sec$.

The sampled test signal with additive zero-mean white Gaussian noise is of the form

$$s[n] = 2\sum_{k=1}^{P} \cos(a_k 2\pi f_k 100e-6 * n) w(100e-6 * t - kT_c) + v[n] \quad for \quad n = 0:3199 \quad (5\text{-}18)$$

Notice that the sampling rate is lowered 2 times but it is still higher than the Nyquist rate and still avoids aliasing. The same simulations are done for this lower sampling rate. This time the worst case window length of TFPF method is reduced to 10 from 20.

Figure 5-16 shows the output of the TFPF method when the sampling rate is reduced to half, the output SNR is 2.9 dB. Recall that the output SNR was 6.8 dB, so, reducing sampling rate to half results in approximately 4 dB SNR decrease.

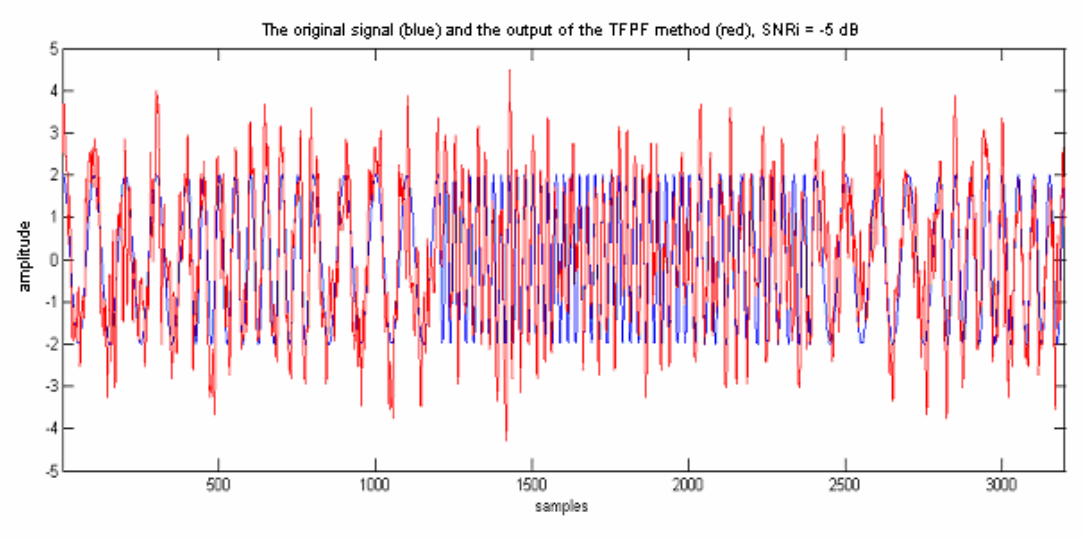

Figure 5-16 The original signal (blue) and the output of the TFPF method (red), SNR_i = -5 dB.

Figure 5-17 shows the output of the SVM based method, the output SNR is 4.15 dB. Recall that the output SNR was 4.9 dB, so reducing the sampling rate does not make a significant degradation in the performance of the SVM based method.

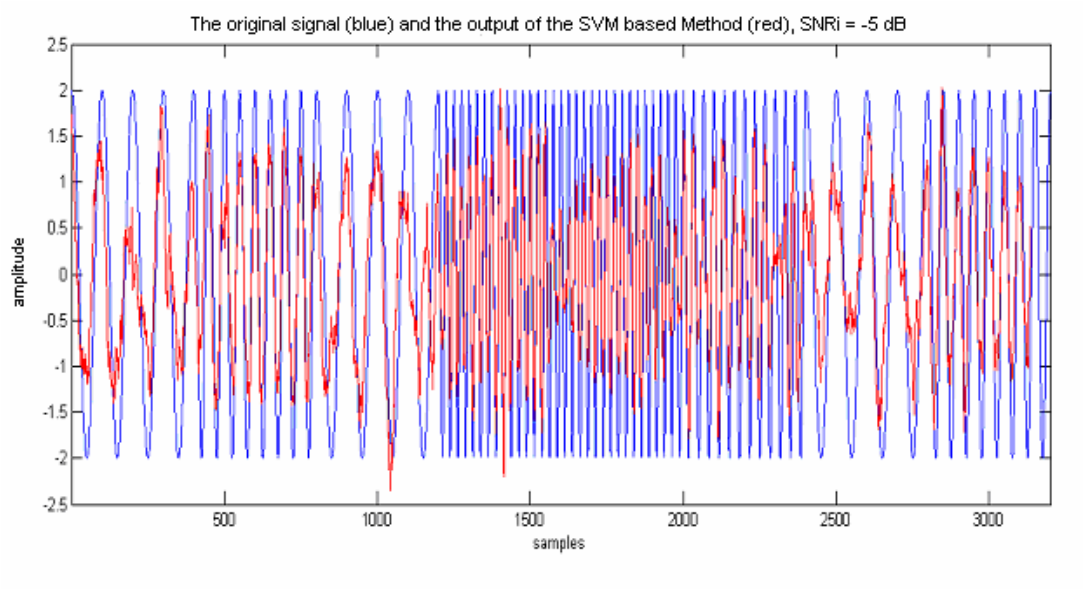

Figure 5-17 The original signal (blue) and the output of the SVM based Method (red), SNR_i = -5 dB.

Figure 5-18 shows the output of the classical zero-phase low-pass filter, the output SNR is 2.5 dB. Recall that the output SNR was approximately 5 dB, so reducing the sampling rate results in approximately 2.5 dB SNR decrease.

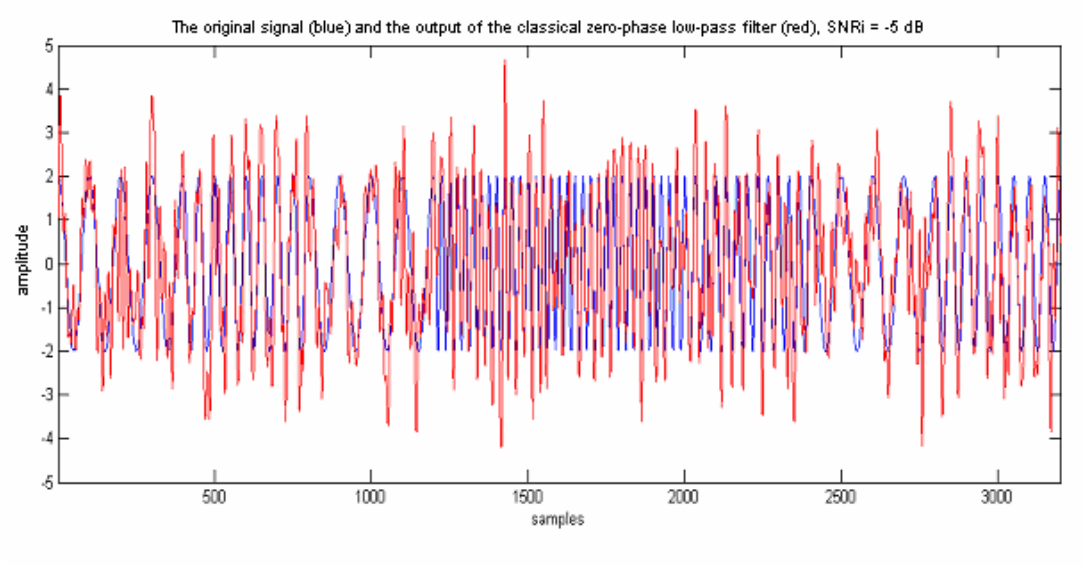

Figure 5-18 The original signal (blue) and the output of the classical zero-phase low-pass filter (red), SNR_i = -5 dB.

When the results for high and low sampling rates are compared it is seen that the output SNR's of the TFPF method and classical zero-phase low-pass filter decrease when the sampling rate is decreased, on the other hand the decrease in the output SNR of the SVM based method is negligible compared to the decrease in the output SNR of the TFPF method and low-pass filter.

5.4 SIMULATION 4: Real World Data

In this simulation a real world data which is digitized echolocation pulse emitted by the Large Brown Bat, Eptesicus Fuscus is used. There are 400 samples; the sampling period was 7 microseconds. However, the signal was sampled at the Nyquist rate and to avoid aliasing its rate is increased using interpolation. The

signal has two simultaneous frequency components. Therefore, SVM based method is modified in order to work with multi-component signals; and it is assumed that the number of simultaneous frequency components is known beforehand. The test signal is very clean and it can be assumed that it is noise free. Although, in the first part of the simulation the signal is used directly, in the second part of the simulation noise is added to the signal to see the performances of the algorithms under noisy case. Figure 5-19 shows the test signal.

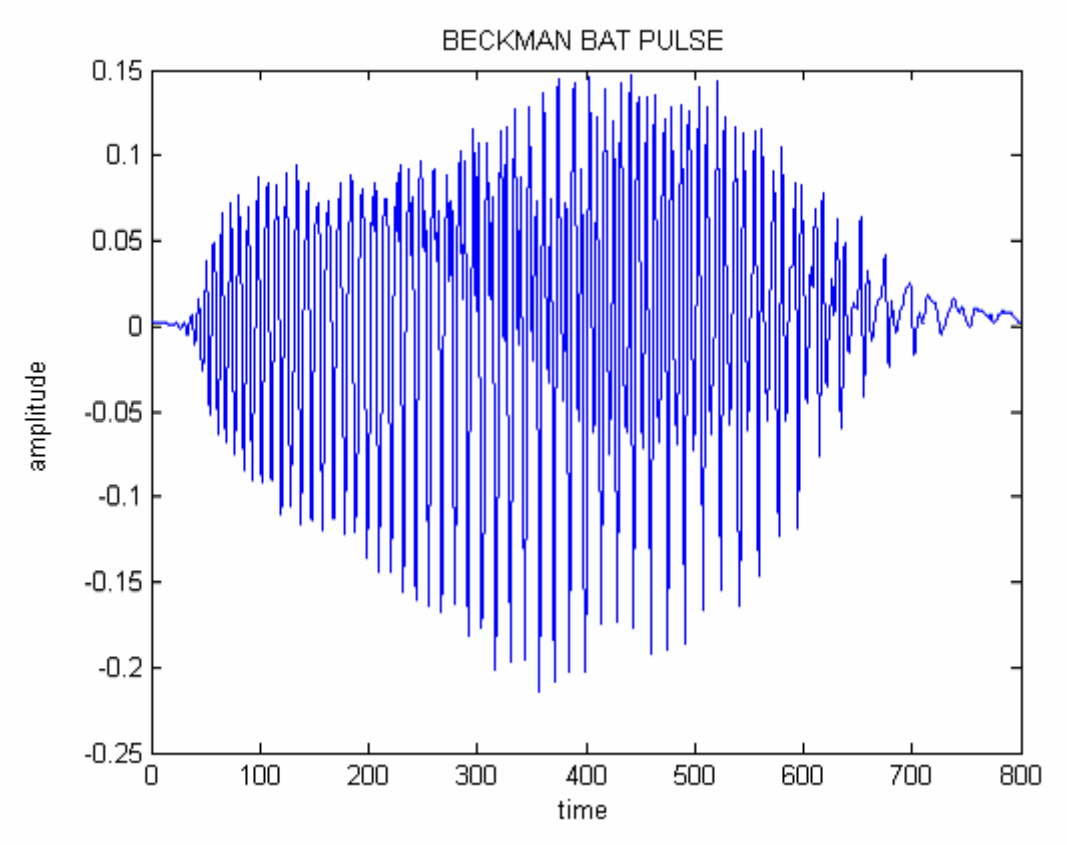

Figure 5-19 The bat signal.

Figure 5-20 shows the pDWVD of the test signal. Figure 5-21 shows the detected auto-term regions (actually the IF of the test signal) after SVM method is applied; as seen from the figure region of support of the signal is correctly detected.

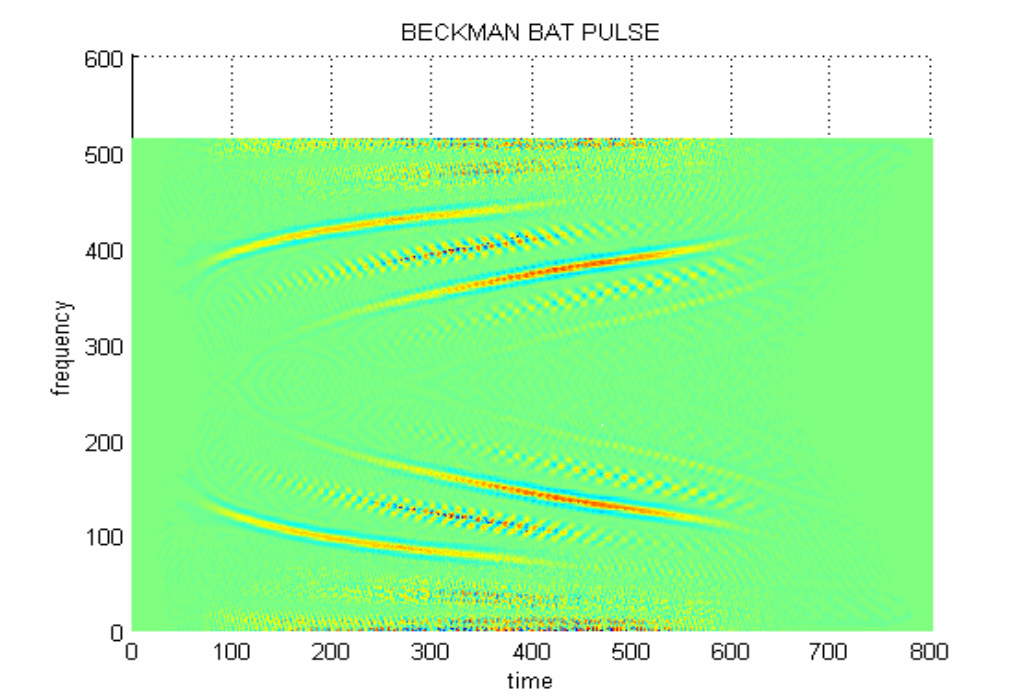

Figure 5-20 pDWVD of the test signal.

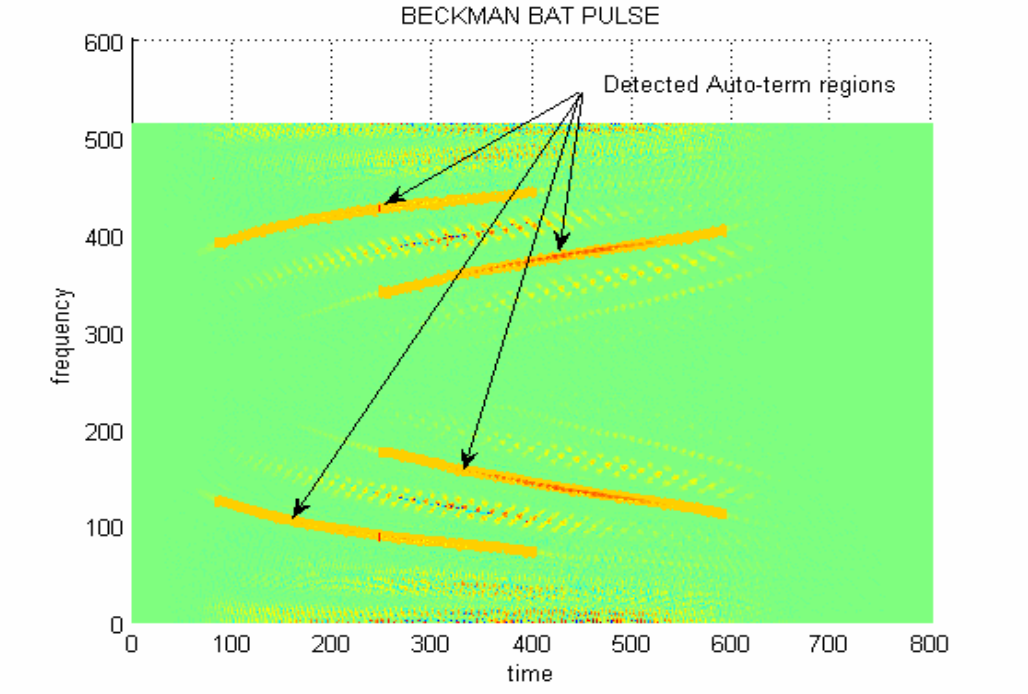

Figure 5-21 Detected auto-term regions (orange) after SVM.

Figure 5-22 shows the estimated and the original signals. Although, the region of support information is correctly found, the signal is estimated with an amplitude modulation as expected.

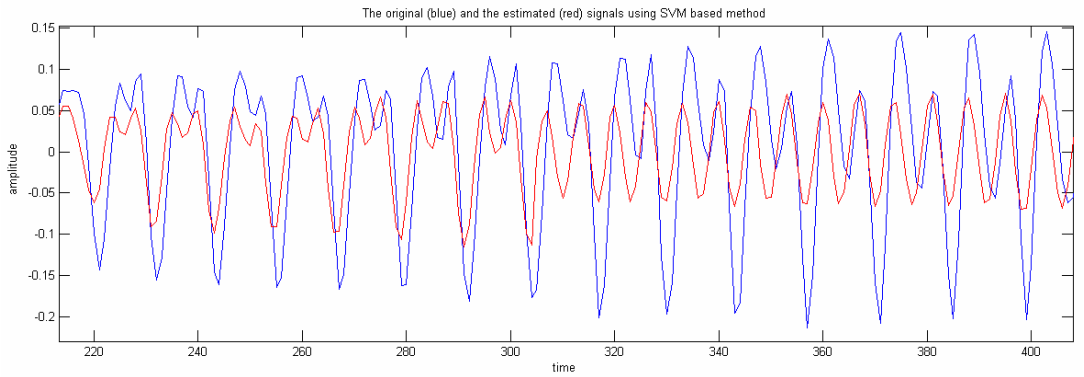

Figure 5-22 The original (blue) and the estimated (red) signals using SVM based method.

Figure 5-23 shows the original signal and the estimated signal using TFPF method. Although, the worst case window length is too small because of the low sampling rate, the signal is correctly estimated since there is no noise on the signal. When the performances of the algorithms are compared, it is seen that TFPF method is a better choice when there is no noise on the signal, because SVM based method produces amplitude modulated estimates, on the other hand, TFPF method not.

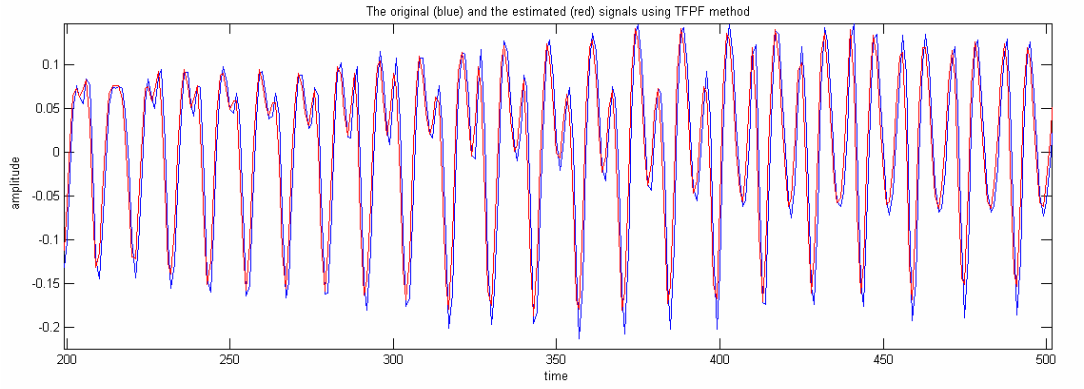

Figure 5-23 The original (blue) and the estimated (red) signals using TFPF method.

To compare the methods under noisy case, noise is added to the test signal; the SNR is approximately -5 dB. Figure 5-24 shows the noisy bat signal.

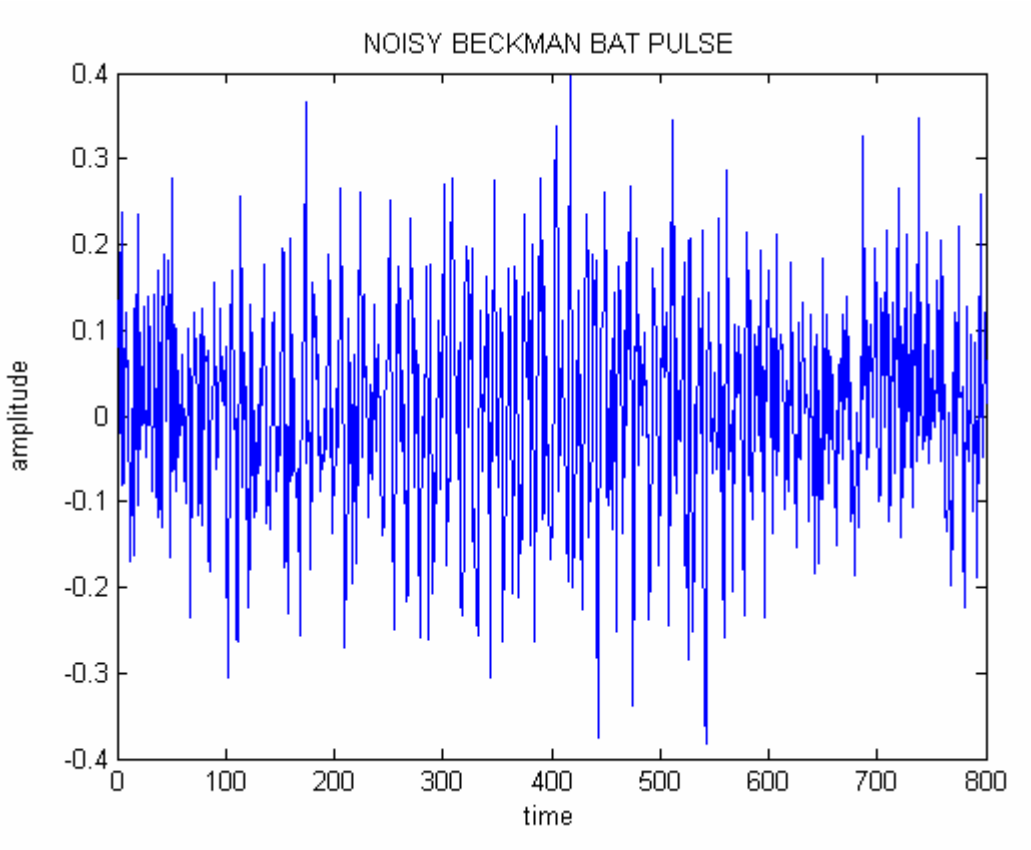

Figure 5-24 The noisy bat signal.

Figure 5-25 shows the pDWVD of the noisy test signal; and Figure 5-26 shows the detected auto-term regions after SVM method. When compared to the noise free case there are extra detected regions in the noisy case.

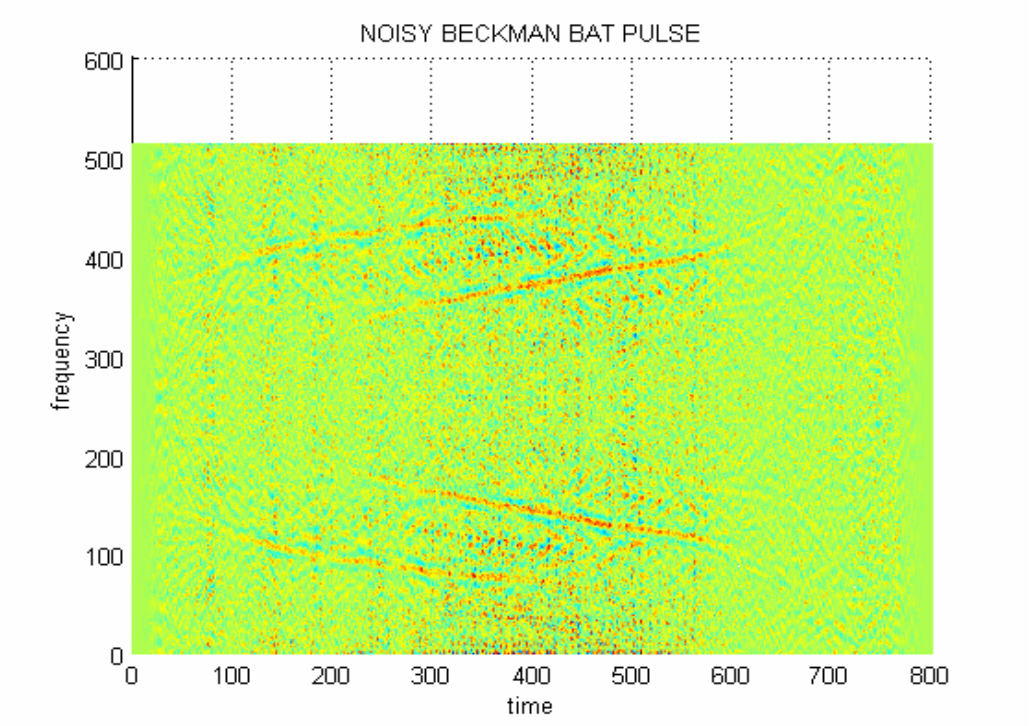

Figure 5-25 pDWVD of the noisy test signal.

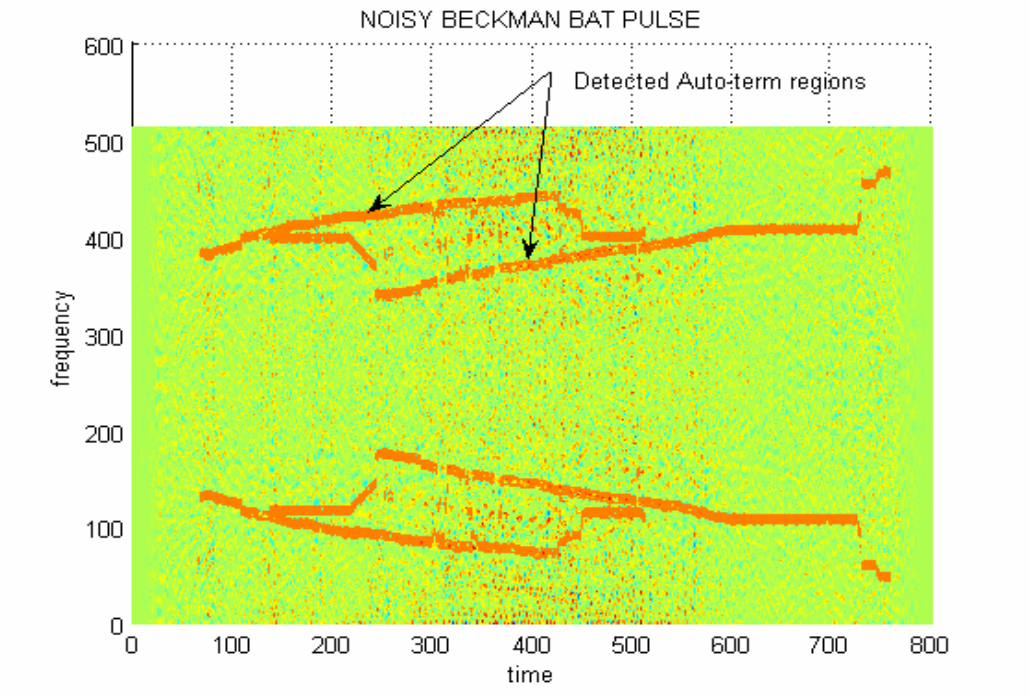

Figure 5-26 Detected auto-term regions (orange) after SVM.

Figure 5-27 shows the output of the SVM based method. The output SNR is approximately 1.6 dB.

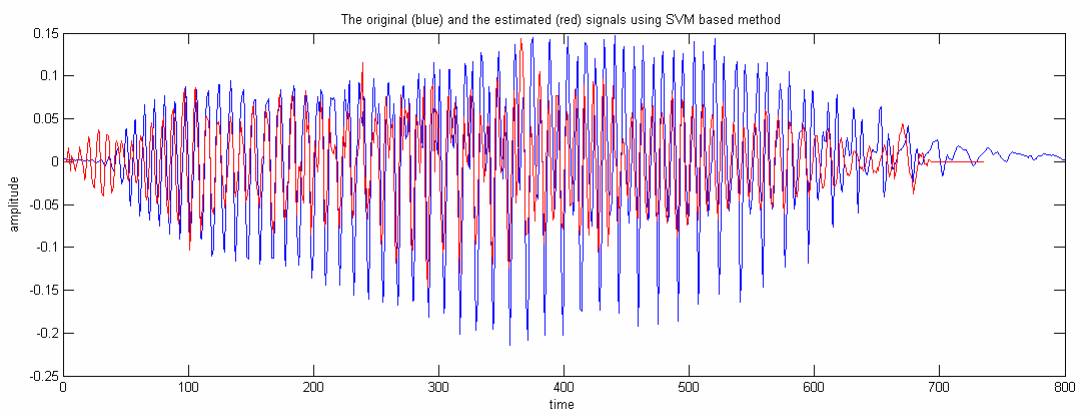

Figure 5-27 The original (blue) and the estimated (red) signals using SVM based method.

Figure 5-28 shows the output of the TFPF method. The output SNR is approximately -1dB.

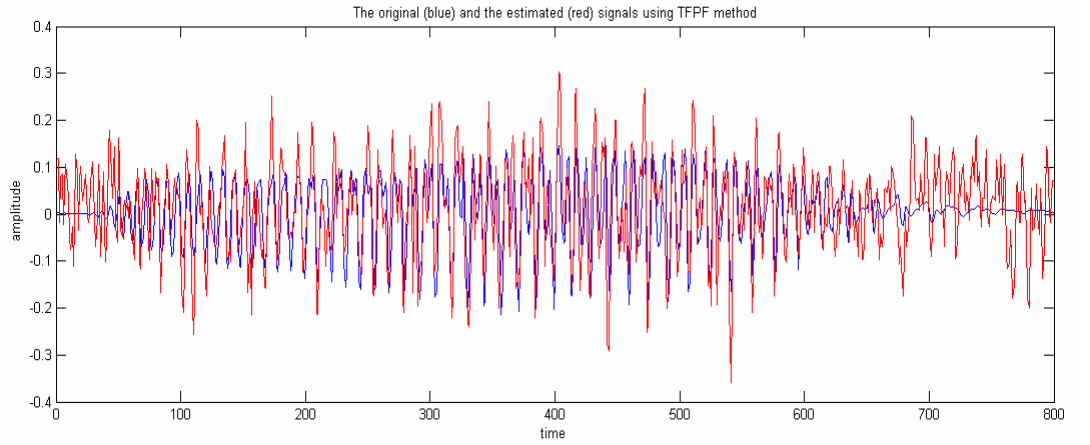

Figure 5-28 The original (blue) and the estimated (red) signals using TFPF method.

In the noisy case SVM based method performs better than the TFPF method.

5.5 RESULTS

Although the simulations are done to determine the best filtering method over the investigated methods, the results show that it can not be concluded that one of the methods performs better than the others under all conditions. The best method is determined by the sampling rate (not by the signal and/or the noise type).

It is observed that for high sampling rates the TFPF method performs better than the other methods; and the performances of the SVM based method and the zero-phase low-pass filter are close to each other. For low sampling rates the performance of the SVM based method is the best, and the performance of the TFPF method is the worst among them.

It is also observed that the performance of the TFPF method mainly depends on the window length which is directly proportional to the sampling rate and inversely proportional to the maximum frequency of the desired signal. When the maximum frequency of the desired signal is kept constant and the sampling rate is increased, the worst-case window length is increased improving the performance; on the other hand, when the sampling rate is decreased, the worst-case window length is decreased degrading the performance. It is seen that the output SNR is approximately decreased by 4 dB when the sampling rate is halved. In addition, it is observed that the performance of the SVM based method is less affected from the sampling rate changes as compared to the performances of the TFPF method and the zero-phase low-pass filter. (The decrease in the output SNR of the SVM based method is less than 1 dB when the sampling rate is halved.)

In case of multi-component signals which contain simultaneous frequency components, the SVM based method needs modification; on the other hand, the TFPF method does not. Actually, the modification is simple and the peak searching is applied to find the peaks where the number of peaks is equal to the number of the simultaneous frequency components, instead of searching for a single peak. So, the number of simultaneous frequency components must be known beforehand, or it

must be detected. In this thesis it is assumed that it is known apriori. It is observed that for the signals which contain many simultaneous frequency components, like brain, heart or sound signals, the performance of the SVM based method degrades. The main reasons for this performance degradation are the number of cross terms which make the time-frequency plane complicated and the differences between the energy levels of the frequency components. However, the SVM based method is not proposed for such signals. Therefore, the real world data, the bat signal, used in the simulation is chosen carefully to make a fair comparison. The bat signal contains only two simultaneous frequency components, it is almost noise free and the frequency components are continuous, i.e., there are no gaps. The obtained results are consistent with the results obtained from the test signals.

CHAPTER 6

CONCLUSIONS

The main interest of this thesis is on the time-frequency filtering of nonstationary signals using WVD. As a first step the WVD is investigated and its properties, advantages and disadvantages are presented. The discrete-time and discrete WVD definitions are given and the relations between them and between continuous-time WVD are shown.

The WVD of noisy signals and the robust and the pseudo forms of the WVD are investigated. The obtained results are presented: the ML estimate of the WVD of a signal under Gaussian noise is equal to the WVD itself; the variance of the discrete-time WVD of a signal under Gaussian noise goes to infinity as the number of signal samples used in the computation goes to infinity; and the pseudo form of the WVD must be used to make variance finite.

Two time-frequency filtering methods are investigated and the performances of the methods are compared using test signals and real world data. Both methods use the pseudo WVD as the time-frequency analysis tool and peak detection to estimate the IF from the pseudo WVD. The main difference between them is on the signal detection and the reconstruction.

First filtering method is the TFPF. In this method the noisy signal is encoded as the IF of an analytic FM signal, and the FM signal is transformed to the time-frequency plane using the pseudo WVD. Then to recover the desired signal, the IF estimation

is performed which actually gives the desired signal estimate multiplied with a constant scalar. So, the desired signal is obtained from the estimated IF data without using the signal synthesis from the time-frequency plane methods which decreases the complexity of the TFPF and increases the signal reconstruction performance. The cross-terms are not a problem for this method, since it works on analytic signals. The effects of different parameters on the performance of the method are investigated and the results are presented. The most important parameter of the method is the length of the window used in the pseudo WVD which is inversely proportional to the maximum frequency of the desired signal and directly proportional to the sampling frequency.

The IF estimation error bias is directly proportional to the window length; on the other hand, the time-frequency resolution is inversely proportional to this length. The window length is computed to limit the error bias; and it is tried to be decreased. However, for a given error bias tolerance, by increasing the sampling frequency the window length can be increased, which results in an increase in the time-frequency resolution at the same time keeping the error bias low. It is observed that by doing so the performance of the method can be improved.

It is shown that to obtain the minimum error variance, the window length must be close to the worst case window length; lower or higher window lengths result in higher error variances. It is also shown that as the number of iterations is increased, the error variance starts to increase after an iteration number which depends on the window length; and it is proposed to stop the iterations when the difference between successive iterations starts to increase to obtain a better performance.

Although the performance of the method is good at low SNR levels, to compute the worst case window length the maximum frequency of the desired signal must be known. In the thesis it is assumed that it is known beforehand. However, when there is no apriori information about the desired signal, the maximum frequency of the desired signal has to be estimated from the noisy observation which requires additional computation.

In the first part of the method the noisy signal is encoded using FM modulation. To limit the effective bandwidth of the FM signal the maximum and the minimum values of the noisy signal must be known; since the effective bandwidth of an FM signal is directly proportional to the maximum value of the modulating signal.

The TFPF method is suitable for multi-component signals which contain simultaneous frequency components.

The second filtering method is the sub-optimal time-varying Wiener filter. In this method the noisy signal is transformed to the time-frequency plane using the pseudo WVD, and then peak detection is performed to detect the desired signal's region of support. From the detected region of support a time-frequency mask is computed and applied to the STFT of the noisy signal; and the desired signal is obtained by summing the magnitude squares of the STFT (spectrogram) along the frequency axis.

The time-varying Wiener filter and the time-frequency formulation of the filter are presented and it is shown that the time-varying transfer function of the filter can be reduced to a time-frequency mask which is equal to 1 on the region of support of the signal and 0 otherwise which gives a sub-optimal result.

The window length used in the pseudo WVD is important, since the variance of the WVD is directly and the bias of the WVD is inversely proportional to the window length. It is shown that to get the optimum window length the unknown derivatives of the WVD of the desired signal are needed. However, in practical applications it is impossible to compute the window length. Therefore, the pseudo WVD is computed for two different window lengths and by comparing their difference and the estimated variance value, one of them is favored.

It is shown that for an analytic mono-component signal, the performance of the method is good for SNR levels down to -5 dB. To increase the performance of the method, the SVM method is applied to the IF estimation phase of the algorithm which is actually a pattern recognition technique. It is shown that for SNR levels

down to -5 dB the estimated signal which is obtained using the SVM method is comparable to the estimated signal which is obtained with known region of support. It is also shown that the method gives satisfactory results at SNR levels down to -10 dB. The main drawback of the method is the threshold setting. In the training phase of the SVM an optimum threshold value is computed. However, this value depends on the signal energy and it must be updated for signals with different energies.

The SVM method increases the computational cost of the algorithm, since for every time-frequency point a decision is made and the noise regions are tried to be masked out. For some time-frequency points the SVM method can make wrong decisions depending on the noise influence, and for these points time gaps or sharp jumps in the IF estimation occur. The sharp jumps are smoothed using median filter, and the time gaps are filled using linear interpolation. However, time gaps also occur when there is no signal at these time instants. Therefore, a decision has to be made to fill or not to fill the time gaps. The decision is made depending on the length of the time gap; if it is lower than half of the minimum window length used in the pseudo WVD computation, the time gaps are filled. By doing so the performance of the algorithm is improved.

Another drawback of the algorithm is the method used in signal reconstruction. Even if there is no noise on the signal, the estimated signal is amplitude modulated. This is because of the masking in STFT of the signal. The signal energy is spread to entire time-frequency plane, and to obtain the signal correctly it has to be summed over the entire frequency axis; however, because of the masking only the masked regions are summed.

The algorithm needs modification to work with multi-component signals which contain simultaneous frequency components. The simplest modification is searching for the peaks whose number is equal to the number of simultaneous components.

The performances of the filtering algorithms are compared using test signals and real world data. It is seen that when the sampling frequency is increased the TFPF method performs better than the time-varying Wiener filter with SVM method; on the other hand, when the sampling frequency is decreased the SVM based method is better. The TFPF method is also better in high SNR cases, since the output of the SVM based method contains an amplitude modulation even if there is no noise on the signal.

Some of the topics remained as future work are noted below:

- Computing worst case window length in the TFPF method by taking the IF estimation error variance into account.
- Computing an adaptive worst case window length in the TFPF method.
- Estimating the maximum frequency of the desired signal from the noisy observations automatically in the TFPF method.
- Improving the signal synthesis method in time-varying Wiener filter.
- Setting the threshold of the SVM method automatically.

REFERENCES

[1] LJubisa Stankovic and Vladimir Katkovnik, “The Wigner Distribution of Noisy Signals with Adaptive Time-Frequency Varying Window,” IEEE Transaction on Signal Processing, vol. 47, no. 4, pp. 1099-1108, April 1999.

[2] LJubisa Stankovic and Srdjan Stankovic, “Wigner Distribution of Noisy Signals,” IEEE Transactions on Signal Processing, vol. 41, no. 2, pp. 956-960, February 1993.

[3] Vladimir Katkovnik and Ljubisa Stankovic, “Instantaneous Frequency Estimation Using the Wigner Distribution with Varying and Data-Driven Window Length,” IEEE Transactions on Signal Processing, vol. 46, pp. 2315–2325, September 1998.

[4] LJubisa Stankovic and Vladimir Katkovnik, “Instantaneous Frequency Estimation Using Higher Order L-Wigner Distributions with Data-Driven Order and Window Length,” IEEE Transactions on Information Theory, vol. 46, no.1, pp. 302-311, January 2000.

[5] Veselin N. Ivanovic, Milos Dakovic, and Ljubisa Stankovic, “Performance of Quadratic Time-Frequency Distributions as Instantaneous Frequency Estimators,” IEEE Transactions on Signal Processing, vol. 51, no. 1, pp. 77-89, January 2003.

[6] Igor Djurovic and LJubisa Stankovic, “Robust Wigner Distribution with Application to the Instantaneous Frequency Estimation,” IEEE Transactions on Signal Processing, pp. 2985-2993, vol. 49, December 2001.

[7] Boualem Boashash and Mostefa Mesbah, “Signal Enhancement by Time-Frequency Peak Filtering,” IEEE Transactions on Signal Processing, vol. 52, no. 4, pp. 929-937, April 2004.

[8] Franz Hlawatsch, Gerald Matz, Heinrich Kirchauer and Werner Kozek, “Time-Frequency Formulation, Design, and Implementation of Time-Varying Optimal Filters for Signal Estimation,” IEEE Transactions on Signal Processing, pp. 1417-1432, vol. 48, May 2000.

[9] LJubisa Stankovic, Srdjan Stankovic and Igor Djurovic, “Space/Spatial-Frequency Analysis Based Filtering,” IEEE Transactions on Signal Processing, pp. 2343-2352, vol. 48, August 2000.

[10] LJubisa Stankovic an Igor Djurovic, “Robust Time-Frequency Analysis: Definitions and Realizations,” EUSIPCO 2004.

[11] Franz Hlawatsch and Werner Kozek, “Time-Frequency Projection Filters and Time-Frequency Signal Expansions,” IEEE Transactions on Signal Processing, pp. 3321-3334, vol. 42, December 1994.

[12] Pai-Hsuen Chen, Chih-Jen Lin, and Bernhard Schölkopf, “A Tutorial on v-Support Vector Machines,”.

[13] Christopher J. C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, Data Mining and Knowledge Discovery, pp. 121-167, Kluwer Academic Publishers, 1998.

[14] Constantine Kotropoulos and Ioannis Pitas, “Signal Detection Using Support Vector Machines in the Presence of Ultrasonic Speckle,” Proceedings of SPIE, pp.304-315, vol. 4687, Medical Imaging 2002.

[15] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, “The Wigner distribution—Part II: Discrete-time signals,” Philips J. Res., vol. 35, pp. 276–300, 1980.

[16] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, “The aliasing problem in discrete-time Wigner distributions,” IEEE Transactions Acoustics, Speech, Signal Processing, vol. ASSP-31, pp. 1067–1072, October 1983.

[17] Françoise Peyrin and Remy Prost, “A Unified Definition for Discrete-Time, Discrete-Frequency, and Discrete-Time/Frequency Wigner Distributions,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-34, no. 4, pp. 858-867, August 1986.

[18] A. H. Nuttall, “Alias-free smoothed Wigner distribution function for discrete-time samples,” Tech. Rep. 8785, Naval Underwater Syst. Cent. (NUSC), New London, CT, Oct. 1990.

[19] Antonio H. Costa and G. Faye Boudreaux-Bartels, “An Overview of Aliasing Errors in Discrete Time-Frequency Representations,” IEEE Transaction on Signal Processing, vol. 47, no. 5, pp. 1463-1474, May 1999.

[20] M. S. Richman, T. W. Parks, and R. G. Shenoy, “Discrete-time, discrete-frequency, time-frequency analysis,” IEEE Trans. Signal Processing, vol. 46, pp. 1517–1527, June 1998.

[21] John O' Toole, Mostefa Mesbah and Boualem Boashash, “A Discrete Time and Frequency Wigner-Ville Distribution: Properties and Implementation,”

[22] Gonzalo R. Arce and Syed Rashid Hasan, “Elimination of Interference Terms of the Discrete Wigner Distribution Using Nonlinear Filtering,” IEEE Transactions on Signal Processing, pp. 2321-2331, vol. 48, August 2000.

[23] F. Hlawatsch and P. Flandrin, “The interference structure of the Wigner distribution and related time-frequency signal representations,” in The Wigner Distribution—Theory and Applications in Signal Processing. Amsterdam, The Netherlands: Elsevier, 1997.

[24] P. Flandrin, “Some features of time-frequency representations of multicomponent signals,” in Proc. IEEE ICASSP, pp. 41.B.4.1–41.B.4.4, 1984.

[25] Igor Djurovic, Ljubisa Stankovic, Johann F. Böhme, “Estimation of FM signal parameters in impulse noise environments,” in Signal Processing 85, pp. 821-835, 2005.

[26] Hiroshi Ijima, Akira Ohsumi, Igor Djurovic, “Parameter Estimation of FM Signals in Random Noise using Wigner Distribution,” SICE Annual Conference in Sapporo, August 4-6, 2004.

[27] P. J. Huber, Robust Statistics. New York: Wiley, 1981.

[28] Franz Hlawatsch, Gerald Matz, “Wigner Distributions (nearly) everywhere: time-frequency analysis of signals, systems, random processes, signal spaces, and frames,” in Signal Processing 83, pp. 1355-1378, 2003.

[29] D. S. K. Chan, “A nonaliased discrete-time Wigner distribution for time frequency signal analysis,” in Proc. IEEE ICASSP, pp. 1333–1336, 1982.

[30] M. A. Poletti, “The development of a discrete transform for the Wigner distribution and ambiguity function,” J. Acoust. Soc. Amer., vol. 84, pp. 238–252, July 1988.

[31] J. Jeong and W. J. Williams, “Alias-free generalized discrete-time time frequency distributions,” IEEE Trans. Signal Processing, vol. 40, pp. 2757–2765, Nov. 1992.

[32] E. C. Bekir, “A contribution to the unaliased discrete-time Wigner distribution,” J. Acoust. Soc. Amer., vol. 93, pp. 363–371, Jan. 1993.

[33] J. C. O’Neill and W. J. Williams, “New properties for discrete, bilinear time-frequency distributions,” in Proc. IEEE-SP Int. Symp. TFTS, Paris, France, pp. 505–508, June 1996.

[34] Victor C. Chen, Hao Ling, Time-Frequency Transforms for Radar Imaging and Signal Analysis, Artech House 2002.

[35] Merrill I. Skolnik, Introduction to Radar Systems, Mc Graw Hill 2001.

[36] Monson H. Hayes, Statistical Digital Signal Processing and Modelling, John Wiley and Sons 1996.

[37] Alfred Mertins, Signal Analysis Wavelets, Filter Banks, Time-Frequency Transforms and Applications, John Wiley and Sons 1999.

[38] Jonathan Stein, Digital Signal Processing, John Wiley and Sons 2000.

[39] Papandreou-Suppapola, Antonia, Applications in Time-Frequency Signal Processing, CRC Press 2003.

APPENDIX A

SOME MATHEMATICAL PROPERTIES OF WVD

Some mathematical properties of WVD will be given briefly;

- WVD of an arbitrary signal is always real. If the signal is also real then the transform is an even function of frequency.

$$W_{xx}(t,w) = W_{xx}^{*}(t,w)$$
$$\text{if} \quad x(t) = x^{*}(t), \qquad \text{then} \quad W_{xx}(t,w) = W_{xx}^{*}(t,w) = W_{xx}(t,-w) \tag{A-1}$$

- Its integral over frequency gives the temporal energy density.

$$\left|x(t)\right|^{2} = 1/2\pi \int_{-\infty}^{+\infty} W_{xx}(t,w)dw \tag{A-2}$$

- Its integral over time gives the energy density spectrum.

$$\left|X(w)\right|^{2} = \int_{-\infty}^{+\infty} W_{xx}(t,w)dt \tag{A-3}$$

- Its integral over time and frequency gives the signal energy.

$$E_{x} = \int_{-\infty}^{+\infty} \left|x(t)\right|^{2} dt = 1/2\pi \int_{-\infty-\infty}^{+\infty+\infty} W_{xx}(t,w)dwdt \tag{A-4}$$

- The WVD of product of two signals is equal to the convolution of the WVD's of the individual signals with respect to frequency.

$$\begin{aligned} &if \quad z(t) = x(t)y(t) \\ &then \quad W_{zz}(t,f) = \int W_{xx}(t,\gamma)W_{yy}(t,f-\gamma)d\gamma = W_{xx}(t,f) *_f W_{yy}(t,f) \end{aligned} \tag{A-5}$$

- The WVD of convolution of two signals is equal to the convolution of the WVD's of the individual signals with respect to time

$$\begin{aligned} &if \quad z(t) = x(t) * y(t) \\ &then \quad W_{zz}(t,f) = \int W_{xx}(\tau,f)W_{yy}(t-\tau,f)d\tau = W_{xx}(t,f) *_t W_{yy}(t,f) \end{aligned} \tag{A-6}$$

- If the signal is restricted only to a certain time interval, then the WVD of the signal is also restricted to that time interval.

$$\begin{aligned} &if \quad x(t) = 0 \quad for \quad t < t_1 \quad and/or \quad t > t_2, \\ &then \quad W_{xx}(t,w) = 0 \quad for \quad t < t_1 \quad and/or \quad t > t_2 \end{aligned} \tag{A-7}$$

- If the signal is restricted only to a certain frequency interval, then the WVD of the signal is also restricted to that frequency interval.

$$\begin{aligned} &if \quad X(w) = 0 \quad for \quad w < w_1 \quad and/or \quad w > w_2, \\ &then \quad W_{xx}(t,w) = 0 \quad for \quad w < w_1 \quad and/or \quad w > w_2 \end{aligned} \tag{A-8}$$

- A time shift of the signal leads to a time shift of the WVD.

$$x(t-t_0) \rightarrow W_{xx}(t-t_0,w) \tag{A-9}$$

- A modulation of the signal leads to a frequency shift of the WVD.

$$x(t)\exp(jw_0t) \rightarrow W_{xx}(t,w-w_0) \tag{A-10}$$

- WVD satisfies the Moyal's Formula, which shows that the squared magnitude of the inner product of two signals is equivalent to the inner product of WVD's of the signals.

$$\left| \int_{-\infty}^{+\infty} x(t) y^*(t) dt \right|^2 = 1/2\pi \int_{-\infty-\infty}^{+\infty+\infty} W_{xx}(t,w) W_{yy}(t,w) dt dw \qquad \text{(A-11)}$$

- The signal can be perfectly reconstructed from the WVD by an inverse Fourier Transform along frequency but with a constant $x^*(0)$ multiplicative term.

APPENDIX B

ML ESTIMATE OF WVD

Let the noisy signal be

$$x(t) = s(t) + v(t) \tag{B-1}$$

The aim is the obtain $f(t) \approx s(t)$,so minimize the function

$$L=\int F(\left| f(t) - x(\tau) \right|) d\tau \tag{B-2}$$

where $F(.)$ is loss function. To obtain the ML estimator,

$$F(e) = -\log(p_v(e)) \tag{B-3}$$

is used as the loss function, where $p_v(e)$ is the PDF of the noise, [27]. For Gaussian noises the loss function reduces to

$$F(e) = \left| e \right|^2 \tag{B-4}$$

To obtain the ML estimate forms of the WVD, consider the following error function

$$e(t, w, \tau) = x(t + \tau/2)\, x^*(t - \tau/2) - m \exp(jw\tau) \tag{B-5}$$

and minimize

$$L(t,w;m) = \int F(\,e(t,w,\tau)\,)\,d\tau \tag{B-6}$$

with respect to m.

$$\left.\frac{\partial L(t,w;m)}{m^*}\right|_{m=W(t,w)} = 0 \tag{B-7}$$

Insert the Gaussian noise loss function into (B-6) to get

$$L(t,w,\tau) = \int \left| x(t+\tau/2)\,x^*(t-\tau/2) - m\exp(jw\tau) \right|^2 d\tau \tag{B-8}$$

To obtain the complex derivative let $m = a + jb$, then

$$\frac{\partial L(t,w;m)}{m^*} = \frac{\partial L(t,w;m)}{a} + j\frac{\partial L(t,w;m)}{b} = 0 \tag{B-9}$$

The derivatives with respect to a and b are

$$\begin{aligned}\frac{\partial L(t,w;\tau)}{a} = \int \big[& x(t+\tau/2)x^*(t-\tau/2)\exp(-jw\tau) \\ & - x^*(t+\tau/2)x(t-\tau/2)\exp(jw\tau)\big]\,d\tau + 2a\end{aligned} \tag{B-10}$$

$$\begin{aligned}\frac{\partial L(t,w;\tau)}{b} = \int \big[& jx(t+\tau/2)x^*(t-\tau/2)\exp(-jw\tau) \\ & - jx^*(t+\tau/2)x(t-\tau/2)\exp(jw\tau)\big]d\tau + j2b\end{aligned} \tag{B-11}$$

Insert (B-10) and (B-11) into (B-9) to obtain

$$m\big|_{m=W(t,w)} = \int x(t+\tau/2)\,x^*(t-\tau/2)\exp(-jw\tau)\,d\tau \tag{B-12}$$

So, the ML estimate of the WVD of a signal in additive Gaussian noise is equal to the WVD of the noisy signal.

APPENDIX C

SURVEY OF THE DISCRETIZATION EFFORTS

Discrete time WVD (DTWVD) of a sampled signal was first defined by Claasen and Mecklenbräuker [15], [16]. It is proposed that, real signals must be over sampled at least by a factor of 2 times the Nyquist rate in order to avoid aliasing. This means that in order to obtain alias free DTWVD, the number of signal samples have to be doubled, which results in an increase in the computations of DTWVD by a factor of 4.

The DTWVD definition proposed in [15], [16] is

$$W^{CM}{}_{xx}(n,\omega) = 2\sum_{k} x[n+k]x^{*}[n-k]\exp(-j2k\omega) \qquad \text{(C-1)}$$

The discrete WVD of a finite length signal $x[n]$ can be obtained from (C-1) by discretizing the frequency as

$$W^{CM}{}_{xx}[n,m] = \frac{2}{N}\sum_{k=-N/2+1}^{N/2-1} x[n+k]x^{*}[n-k]\exp(-j\frac{4\pi}{N}mk) \qquad \text{(C-2)}$$

which is computed for $0 \le n \le N-1$ and $0 \le m \le M-1$, where N is the number of signal samples and M ($\ge$N) is an appropriate number of discrete frequencies.

Figure C-1 shows the DWVD of a test signal obtained using (C-2).

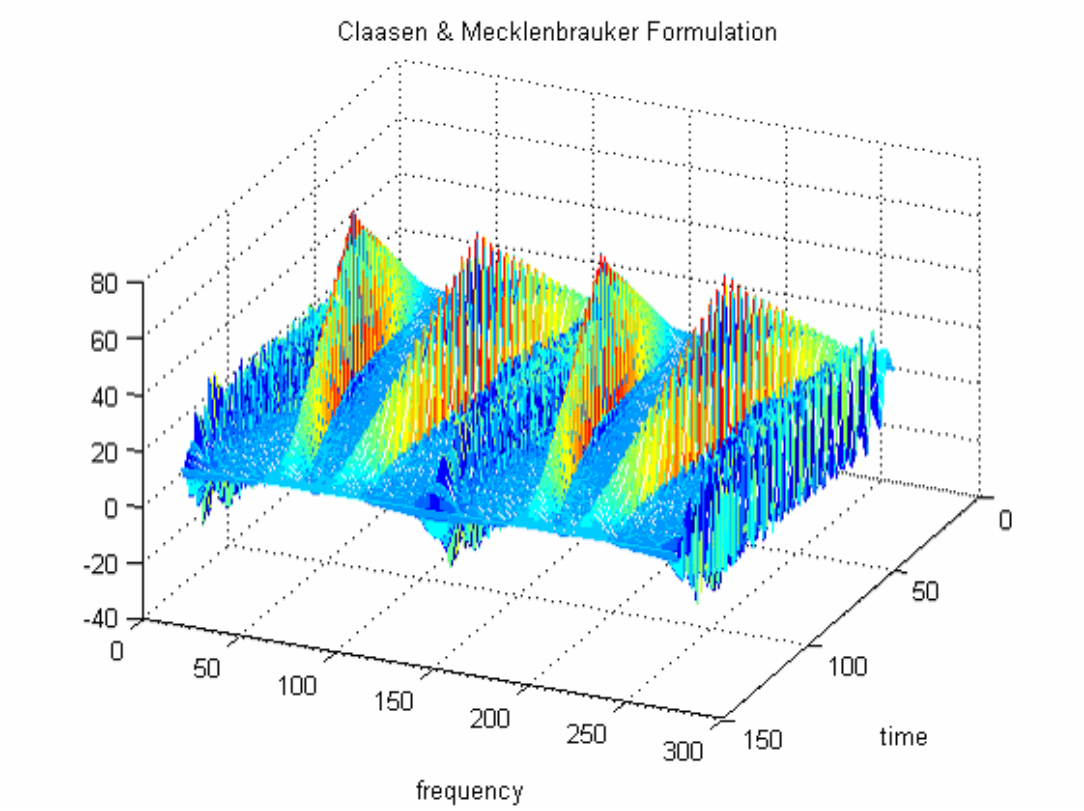

Figure C-1 DWVD of a test signal obtained using (C-2).

Another important contribution came from Peyrin and Prost who shown that it is necessary to sample the signal at least twice the Nyquist rate in order to obtain alias free Discrete WVD (DWVD) in 1986, [17]. In [17], discrete-time, discrete-frequency and discrete-time discrete-frequency (discrete) definitions for WVD are proposed. The proposed DWVD definition is of the form:

Let x_r be a periodic signal restricted to an interval of length NT

$$x_r(t) = \sum_l x(t - NlT) \qquad and \qquad x_k = x_r(kT) \tag{C-3}$$

Then discrete time, discrete frequency WVD (DWVD) of x is

$$W^{PP}{}_{xx}[n,m] = \frac{1}{2N}\sum_{k=0}^{N-1} x_k x_{n-k}^{*} \exp(-j\frac{\pi m}{N}(2k-n)) \qquad \text{(C-4)}$$

Figure C-2 shows the DWVD of a test signal obtained using (C-4).

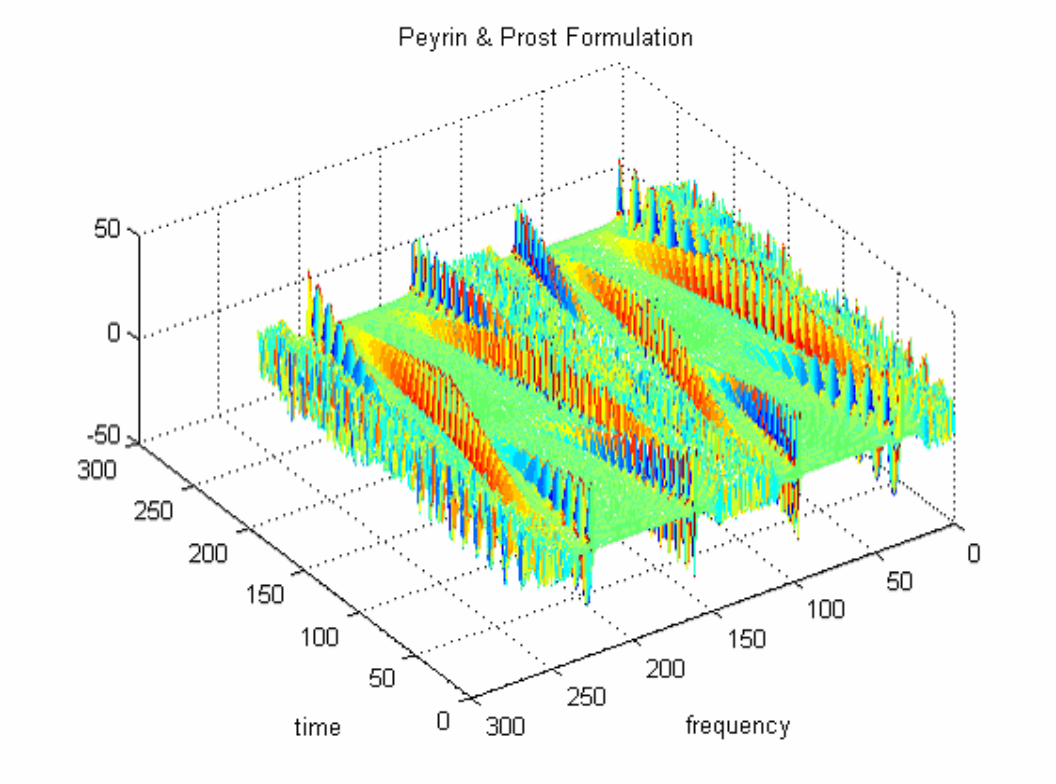

Figure C-2 DWVD of the test signal obtained using (C-4).

The distribution is computed for $0 \le n \le 2N-1$ *and* $0 \le m \le 2N-1$. Fortunately, because of the form of the proposed definition it is sufficient to compute the distribution on an N by N grid. The remaining parts of the distribution can be obtained from this N by N grid by multiplying corresponding entries with a phase term which results in ±1.

In 1990, Nuttall proposed to use the dual form of the continuous time WVD to derive discrete-time and discrete-frequency definition of the distribution, [18]. The Nyquist rate sampled signal is used in the computation and the number of signal

samples is doubled in the DFT computation phase of the algorithm by zero padding the signal. In [19], it is shown that the proposed definition is always alias free for a signal sampled at the Nyquist rate.

The proposed definition is of the form:

Given the signal samples $x[n] \quad for \quad 0 \le n \le M-1$ and a suitable DFT length $N > M$, the DFT $X[k]$ of signal $x[n]$ is computed using

$$X[k] = \sum_{n=0}^{M-1} x[n] \exp(-j2\pi n \frac{k}{2N}) \quad for \quad -N \le k \le N-1 \tag{C-5}$$

Then DWVD of $x[n]$ is

$$W^{N}{}_{xx}[\frac{nT}{2}, \frac{k}{2NT}] = \frac{T}{N} \sum_{m=-l}^{l} X[k+l] X^{*}[k-l] \exp(-j2\pi m \frac{n}{2N})$$

$$for \quad l = \begin{cases} N-k-1 \quad for \quad k \ge 0 \\ \\ N+k \quad for \quad k < 0 \end{cases} \tag{C-6}$$

The distribution is computed for $0 \le n \le 2N-1 \quad and \quad -N \le k \le N-1$. Figure C-3 shows the DWVD of a test signal obtained using (C-6).

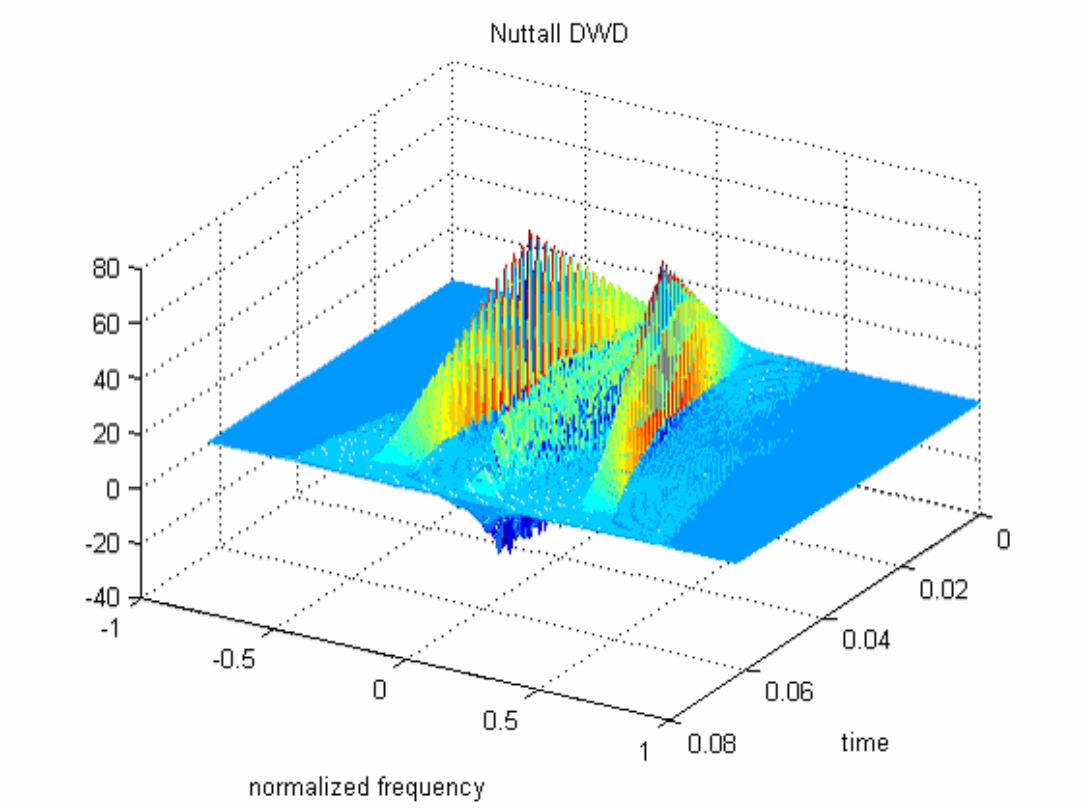

Figure C-3 DWVD of the test signal obtained using (C-6).

In 1998 Michael S. Richman, Thomas W. Parks, and Ramachandra G. Shenoy, proposed to obtain the discrete time discrete frequency definition for WVD using Group Representation Theory, [20]. The distribution is computed on an N by N grid (N is the number of signal samples) and the resulting distribution is a full band representation that is to say the frequency components are computed in the normalized range [-1/2,1/2]. The main drawback of the definition is, it depends on the signal length; for even length and odd length signals it gives different results. Additionally the resultant distribution also depends on the signal frequency for odd length signal. The definition works well for odd length signals with frequency at 1/M where M is an integer and divides N (N is the length of the signal). For even length signal the resultant distribution contains cross-terms at every point of the distribution but it does not depend on the frequency content of the signal.

The proposed definition is of the form

$$W^{G}{}_{xx}[n,k] = \frac{1}{N}\sum_{\tau=0}^{N-1}\sum_{v=0}^{N-1}\sum_{l=0}^{N-1} pN \exp(-j\frac{2\pi}{N}(nv + k\tau - vl))x(\mathrm{mod}(l+\tau, N))x^{*}(l) \quad \text{(C-7)}$$

where pN is a phase term which is defined as

$$for \quad N = odd \quad pN = \exp(j\frac{2\pi}{N}\mathrm{mod}(\frac{v\tau}{2}, N))$$

$$for \quad N = even \quad pN = \begin{cases} \exp(j\frac{\pi}{N}\mathrm{mod}(v\tau, N)) & for \quad \mathrm{mod}(v\tau, N) < \frac{N}{2} \\ \exp(j\frac{\pi}{N}(\mathrm{mod}(v\tau, N) - N)) & for \quad \mathrm{mod}(v\tau, N) \geq \frac{N}{2} \end{cases} \quad \text{(C-8)}$$

In 2005, John O' Toole, Mostefa Mesbah and Boualem Boashash proposed a new definition to obtain DWVD which is based on the definition offered by Peyrin and Prost, [21]. To obtain the modified definition it is assumed that the signal is not periodic.

The proposed definition is of the form

$$W^{M}{}_{xx}[\frac{n}{2f_s}, \frac{kf_s}{2N}] = \exp(j\frac{\pi}{N}kn)\sum_{m=l_1}^{l_2} x[m]x^{*}[n-m]\exp(-j\frac{2\pi}{N}km) \quad \text{(C-9)}$$

for $l_1 = \max\{0, n-(N-1)\}$ and $l_2 = \min\{n, N-1\}$. Figure C-4 shows the DWVD of a test signal obtained using (C-9). When the resultant distribution is compared to the distribution obtained using (C-4), it is seen that it gives a cleaner time-frequency plane for the same test signal.

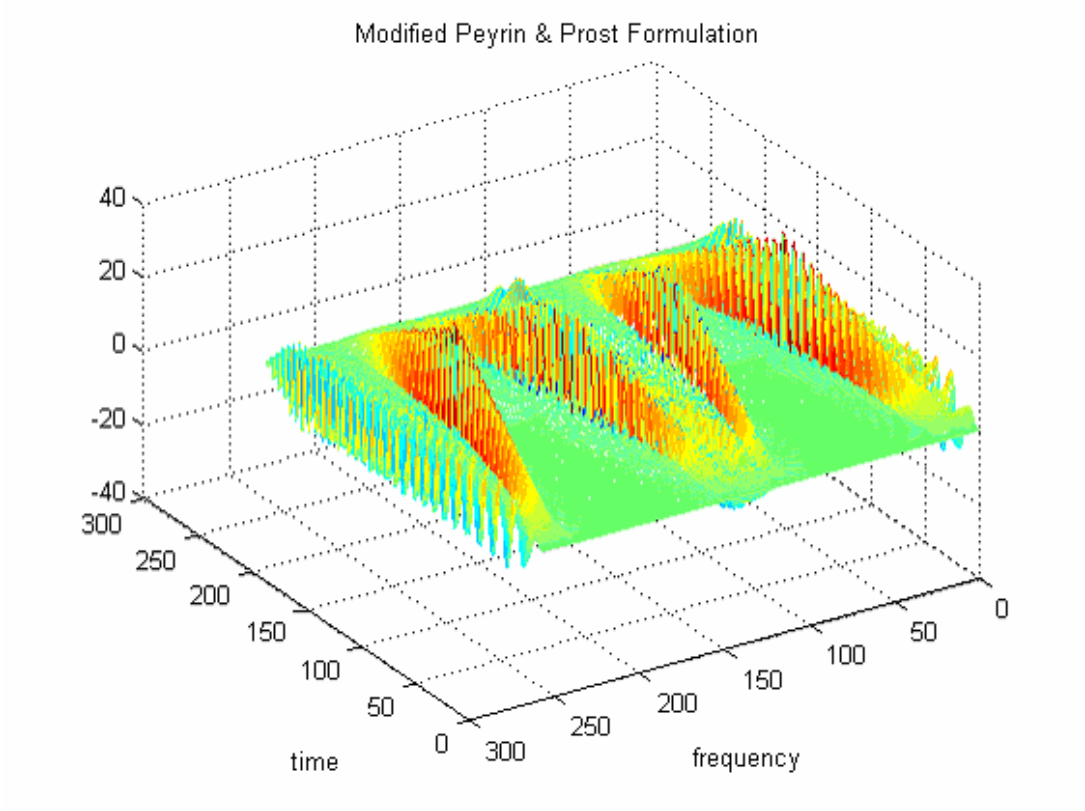

Figure C-4 DWVD of the test signal obtained using (C-9).

The discrete-time discrete-frequency WVD definitions explained in this section are only a limited number of the proposed definitions; interested readers can refer to [29]-[33], for more information in discretization of WVD.