2016

Time-Frequency Based Methods for Non-Stationary Signal Analysis with Application To EEG Signals

Amal Feltane

University of Rhode Island, amal_feltane@my.uri.edu

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Recommended Citation

Feltane, Amal, "Time-Frequency Based Methods for Non-Stationary Signal Analysis with Application To EEG Signals" (2016). *Open Access Dissertations*. Paper 445.
https://digitalcommons.uri.edu/oa_diss/445

This Dissertation is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

TIME-FREQUENCY BASED METHODS FOR NON-
STATIONARY SIGNAL ANALYSIS WITH APPLICATION
TO EEG SIGNALS
BY
AMAL FELTANE

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
ELECTRICAL ENGINEERING

UNIVERSITY OF RHODE ISLAND

2016

DOCTOR OF PHILOSOPHY DISSERTATION
OF
AMAL FELTANE

APPROVED:

Dissertation Committee:

Major Professor G. Faye Boudreaux-Bartels

Walter G. Besio

Ying Sun

Musa Jouaneh

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2016

ABSTRACT

The analysis of electroencephalogram or EEG plays an important role in diagnosis and detection of brain related disorders like seizures. In this dissertation, we propose three new seizure detection algorithms that can classify seizure from non-seizure data with high accuracy. The first algorithm is based on time-domain features which are the approximate entropy (ApEn), the maximum singular value (MSV) and the median absolute deviation (MAD). These features were fed into the AdaBoost and the Support Vector Machine (SVM) algorithms, which were used to classify the signal as either seizure or non-seizure. The accuracy of these classifications was summarized and compared to different algorithms in the literature.

In the second algorithm, the Rényi entropy was extracted from different spectral components after the EEG signal was decomposed using either Empirical Mode Decomposition (EMD) or the Discrete Dyadic Wavelet Transform (DWT). The k-nearest neighbor (k-NN) classifier was used to classify the seizure segments based on the extracted features. In the third algorithm, we decompose the EEG signal into sub-components occupying different spectral sub-bands using the EMD. A decomposition energy measure was used to discard those sub-components estimated to contain mostly noise. Different time-frequency representations (TFRs) were computed of the remaining sub-components. Local energy measures were estimated and fed into a linear classifier to determine whether or not the EEG signal contained a seizure. The three algorithms were tested on noisy EEG signals from roaming rats as well as the relatively noise free human seizure from a well-known public dataset provided on-line

(Andrzejak et al., 2001). Using Metrics of total Sensitivity, Specificity and Accuracy, it was demonstrated that the proposed algorithms gave either equivalent or superior performance when compared against several other brain seizure algorithms previously reported in the literature.

Furthermore, we propose a new warping function to create a new class of warped Time-Frequency Representations (TFRs) that is a generalization of the previously proposed k^{th} Power Class and Exponential Class TFRs. The new warping function is

$\hat{w}(t) = e^{at} t^{1/k}$. We provide the formulas for the one-to-one derivative warping function and its inverse defined using the Lambert-W function. Examples are provided demonstrating how the new warping function can be successfully used on wide variety of non-linear FM chirp signals to linearize their support in the warped Time-Frequency plane.

An optimization scheme was proposed to find the optimal parameter, “ a ”, of the new warping function for a given non-linear FM chirp signal; algorithms have previously been proposed for finding the k . The performance of the optimization technique was compared to other warped Time-Frequency Representations; the new warped TFRs achieved better linearization in several cases. The new warping function was used to develop a new algorithm which iteratively isolates and separates non-linear FM signal components in a multicomponent signal. The isolated components have negligible interference terms and have energy support concentrated along a curve close to the true instantaneous frequency.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor; Dr. G. Faye Boudreaux-Bartels, for her excellent guidance and continuous support during my Ph.D study and for her motivation, enthusiasm, and immense knowledge that always inspire me. Her mentorship and the freedom she gave me to pursue ideas in my own way helped me all the time to grow both professionally and personally. I am truly fortunate to have her to be my advisors and mentors.

I would like to express the appreciation to Dr. Walter Besio for his guidance and persistent help.

Besides my advisors, I would like to thank the rest of my dissertation committee: Professors: Ying Sun, Musa Jouaneh, Richard J. Vaccaro and Barbara Kaskosz for their time in participating in my comprehensive exam and dissertation defense.

I am also grateful to other faculty and staff members of the Department of Electrical, Computer, and Biomedical Engineering, in particular to Meredith Leach Sanders. I would also like to extend sincere thanks to my colleagues for their support and help during my Ph.D study.

I would like to thank my parents, my brother and my sisters; they are always supporting me and encouraging me with their love and best wishes.

Finally, I would like to thank my husband, Yacine Boudria, for his love, support, understanding and patience throughout the completion of this research.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGMENTS	iv
TABLE OF CONTENTS.....	v
LIST OF TABLES	x
LIST OF FIGURES	xii
CHAPTER 1 INTRODUCTION	1
1.1 Epilepsy and Seizure	1
1.2 Electroencephalogram and Laplacian electroencephalogram	2
1.3 Methods for seizure detection	5
1.4 Thesis contributions	8
1.5 Thesis organization	8
CHAPTER 2 TIME-FREQUENCY SIGNAL PROCESSING	11
2.1 Introduction	11
2.2 The need for time-frequency analysis	11
2.3 Types of time-frequency representations	14
2.3.1 Linear time-frequency representations.....	15
2.3.1.1 Short time-frequency transform (STFT)	15
2.3.1.2 Wavelet transform (WT)	17
2.3.2 Quadratic time-frequency representations	18
2.3.2.1 Wigner-Ville distribution (WVD).....	18
2.3.2.2 Smoothed-Pseudo Wigner-Ville distribution (SPWVD)	20

2.4 A comparison of the existence of cross-terms in the WVD, the Spectrogram and the Scalogram	21
2.5 Reassignment time-frequency representation and Synchrosqueezing transform	23
2.5.1 Reassignment time-frequency representation	23
2.5.2 Synchrosqueezing transform	24
2.6 Summary	26
CHAPTER 3 SEIZURE DETECTION USING THE EEG SIGNAL.....	28
3.1 Introduction	28
3.2 EEG data acquisition.....	28
3.2.1 Rat's dataset	29
3.2.2 Human dataset	32
3.3 Seizure detection using time-domain features	33
3.3.1 Data preprocessing	34
3.3.2 Feature extraction.....	34
3.3.2.1 Feature 1: Approximate Entropy (ApEn).....	34
3.3.2.2 Feature 2: Maximum Singular Value (MSV).....	36
3.3.2.3 Feature 3: Median Absolute Deviation (MAD)	37
3.3.3 Feature vector classification.....	39
3.3.3.1 Classifier 1: Support Vector Machine (SVM)	39
3.3.3.2 Classifier 2: Adaptive Boosting (AdaBoost)	40
3.3.4 Performance evaluation.....	42
3.3.5 Results and discussion	44

3.4 Seizure detection using the Rényi entropy.....	47
3.4.1 Feature extraction.....	48
3.4.1.1 Empirical Mode Decomposition (EMD).....	48
3.4.1.2 Discrete Wavelet transform (DWT).....	51
3.4.1.3 Rényi Entropy	54
3.4.2 Classification and performance calculation	55
3.4.2.1 Ten-fold cross-validation for training and testing.....	55
3.4.2.2 Classification using k-nearest neighbor (k-NN)	56
3.4.3 Performance evaluation.....	57
3.4.4 Results and discussion	57
3.5 Summary	61
CHAPTER 4 SEIZURE DETECTION USING TIME-FREQUENCY	
ENERGY CONCENTRATION.....	62
4.1 Introduction.....	62
4.2 Data description	62
4.3 Model-based seizure detection.....	63
4.4 Feature extraction procedure.....	65
4.4.1 Signal decomposition and IMFs selection.....	65
4.4.2 Time-frequency energy estimation	69
4.5 Classification and performance calculation	71
4.6 Results and discussion	72
4.7 Summary	77

CHAPTER 5 NON-STATIONARY FM CHIRP SIGNAL PROCESSING..

.....	78
5.1 Introduction	78
5.2 Fractional Fourier Transform (FrFT)	79
5.2.1 Chirps and Fractional Fourier Transform	81
5.2.2 The relation between the FrFT and the Wigner-Ville distribution (WVD)	82
5.3 Multi-component FM chirp signals decomposition using the FrFT	84
5.4 Limitations of FrFT-based multi-component signals decomposition algorithm	99
5.5 Time-warping principle.....	103
5.5.1 New warping function.....	108
5.5.2 Performance analysis and parameter optimization	116
5.5.2.1 Rényi entropy of a TFR.....	116
5.5.2.2 Ratio of Norm	117
5.5.2.3 Estimation of optimum warping parameter “a”	118
5.5.2.4 Performance comparison.....	120
5.6 Warping-based multi-component non-linear FM chirp signal decomposition	122
5.7 Futur work.....	132
5.8 Summary	136
CHAPTER 6 CONCLUSION AND FUTURE PERSPECTIVES.....	138
6.1 Conclusion	138

6.2 Futur works	141
BIBLIOGRAPHY	143

LIST OF TABLES

TABLE	PAGE
Table 3.1. Statistical values of extracted features from tEEG human data with and without seizure of Figure 3.3.	45
Table 3.2. A comparison of classification accuracy obtained using tEEG data, rat's data, and EEG data from human	46
Table 3.3. Frequency bands of EEG signal with four levels of DWT decomposition and sampling frequency of 173.6 Hz. D1-D4 are the detail coefficients and A4 is the approximate coefficient.....	52
Table 3.4. The quadratic Rényi entropy values of Intrinsic Mode Functions (IMF1-IMF5) after EMD decomposition on human dataset.....	55
Table 3.5. The quadratic Rényi entropy values of signal sub-components D1-D4 and A4 after DWT on human dataset	56
Table 3.6. A comparison of classification performance achieved by EMD and DWT using k-NN classifier on human dataset.....	57
Table 3.7. A comparison of classification accuracy obtained by our methods versus methods of others' using the dataset described by Andrzejak et al. [58].....	60
Table 4.1. The classification performance using 64-point length for time smoothing window and 128-point length for frequency smoothing window with the five different window lengths (1, 2, 3, 4 and 5 s). sampling frequency: $F_s=128\text{Hz}$	73
Table 4.2. Accuracy obtained using different types of time-frequency distributions (TFDs).....	74

Table 4.3. Performance parameters (Sensitivity (Sens), Specificity (Spec) and Accuracy (Acc)) of the proposed method using different combinations of length of time and frequency smoothing windows	75
Table 5.1. Properties of the Fractional Fourier Transform (FrFT) as related to the Fourier transform (FT) [96]. “I” is the identity matrix	81
Table 5.2. The parameters of signals used to compare the warping operator $w(t) = t^{1/k}$ with the new warping operator $\hat{w}(t) = e^{at}t^{1/k}$	34
Table 5.3. Values of the Rényi entropy (RE), the Ratio of Norm (RN) and the FT maximum peak (Mp) occurring during the search for the optimum value of the new time-warping function parameter “a”. The signal analyzed was $x_2(t) = e^{j2\pi(t^{0.1}+0.85t^{0.7})}$ with $k = 0.7$ in Example 7. Optimal results highlighted in yellow row.....	119
Table 5.4. Performance analysis using two different warping functions. $w(t) = t^{1/k}$ is the power warping function and $\hat{w}(t) = e^{at}t^{1/k}$ is the new warping function.....	121

LIST OF FIGURES

FIGURE	PAGE
Figure 1.1. International 10-20 System of scalp electrode placement. Image obtained from www.sjopat.com/10-20-electrode-placement-pattern	3
Figure 1.2. An example of a seizure pattern in Laplacian EEG (tEEG) signal.....	5
Figure 1.3. An example of conventional disc electrode, (A), and Tri-polar concentric ring electrode, (B)	6
Figure 2.1. mono-component linear FM signal; (A) Represents the time domain representation. (B) Its magnitude spectrum and (C) time-frequency representation..	13
Figure 2.2. multi-component linear FM signal. (A) Represents the time domain representation. (B) Its magnitude spectrum and (C) time-frequency representation..	14
Figure 2.3. Spectrogram using a Hamming window with (A) 15-point length. (B) 64-point length, (C) 128-point length.....	16
Figure 2.4. Wigner-Ville distributions (WVD) of (A) mono-component parabolic chirp signal, (B) Multi-component signal with two cross linear chirp signals.	19
Figure 2.5. Smoothed-pseudo Wigner Ville distributions (SPWVD) of (A) mono-component parabolic chirp signal, (B) Multi-component signal with two cross linear chirp signals	21
Figure 2.6. Reassigned smoothed-pseudo Wigner Ville distributions (RSPWVD) of (A) mono-component parabolic chirp signal, (B) Multi-component signal with two cross linear chirp signals	24
Figure 2.7. Synchrosqueesed Wavelet Transform of (A) mono-component parabolic	

chirp signal, (B) Multi-component signal with two cross linear chirp signals	25
Figure 3.1. Typical tEEG data recording from a rat using TCRE; (A): The location of the tripolar concentric ring electrodes (TCREs) on the rat scalp. Electrode (1) is 10 mm dia. and used for stimulation and recording. Electrodes (2) and (3) are both 6.0 mm dia. and used only for recording. Electrode (r) is the isolated ground. Details of TCRE are shown to the right of the rat head (B): 30 min of data with and without seizure (with 460,800 samples). After 5 min of Baseline recordings the PTZ was administered. Soon after giving the PTZ, usually within 2 min, the rats had their first myoclonic jerk. (C): thirty seconds Baseline data (note that the vertical axis is magnified compared to panels (B) and (D)), (D): thirty seconds Seizure data.....	
	30
Figure 3.2. An example of human EEG signals from each of the five subsets (Z, O, N, F, and S).....	
	33
Figure 3.3. The values of the three features used for seizure detection (MAD in (a), ApEn in (b), and MSV in (c)) using Baseline and Seizure segments of rat data. The oscillation of Seizure data in comparison to the Baseline data explains the high values of MAD and Max singular value. The values of ApEn for Seizure data are low compared to Baseline data	
	38
Figure 3.4. Top of the figure is an example of EEG segments from Z and S sets respectively. Sub-figures (a) – (c) represents the values of the extracted features (ApEn, MSV, and MAD). It is clear that the values of MAD and MSV are higher for Seizure data than for Normal data, but the ApEn for Normal data are higher than those for Seizure.	
	39
Figure 3.5. (A) General diagram of a Support Vector Machine, image obtained from	

[69] (B) An example of classification using the SVM with RBF kernel function, image obtained from [110].	40
Figure 3.6. The AdaBoost classifier Algorithm [71]	43
Figure 3.7. An overview of the proposed methodology for human epileptic seizure detection	48
Figure 3.8. The Empirical mode decomposition (EMD) of non-seizure data (set Z) on left and with seizure (set S) on right, respectively. IMF1-IMF5 are the first five IMFs after EMD decomposition of each segment from the sets A and E. The frequency support of these IMFs decreases from top to bottom. Furthermore, the lower order IMF has fast oscillation mode of the signal.	51
Figure 3.9. Approximate and detail coefficients of non-seizure EEG (set Z) on left and epileptic Seizure EEG (set S) on right. The first row contains the original data. Row 2 contains the approximate coefficients A4. Row 3, 4, 5 and 6 contain the detail coefficients D4, D3, D2 and D1, respectively	53
Figure 4.1. Block diagram of the proposed method for automatic seizure detection	64
Figure 4.2. (Aa) and (Ba) in that order depict a 30-second segment of tEEG recording from rats during non-seizure and seizure period, respectively. The Empirical Mode Decomposition (EMD) of the non-seizure and seizure segments is shown in (Ab) and (Bb) with the IMF1 on the top row and the last IMF in the bottom row. The EMD computed 12 IMFs for the non-seizure data and 10 for the seizure data.	67
Figure 4.3. An example of IMFs selection criteria using equations 4.2 and 4.3. The noise-only model in red and the confidence interval in blue are presented. The black curve represent the energy of each IMF. The IMFs numbers 3 to 10 have energies	

which exceed the confidence interval (or threshold).	68
Figure 4.4. Top plots represent IMF3, IMF4, and IMF5 of non-seizure tEEG signal. Shown at the bottom are the corresponding SPWV TF distributions (two Hamming windows with 64-point length and 128-point length are used respectively, for the time and frequency smoothing windows). TF plots of the three IMFs show that the frequency content of IMF3 is higher than the frequency of IMF4 which has higher frequency content than that of IMF5.....	70
Figure 4.5. Top plots shows IMF3, IMF4, and IMF5 of seizure tEEG signal. Shown at the bottom are the corresponding SPWV TF distributions (two Hamming windows with 64-point length and 128-point length are used respectively, for the time and frequency smoothing windows). It is very clear from these SPWVs, that the frequency of IMF3 is higher than the frequency of IMF4 which has higher frequency content than that of IMF5... ..	71
Figure 5.1. The relation of time-frequency plane (t, f) with the Fractional Fourier domain (u, v) rotated by an angle ϕ	82
Figure 5.2. Linear FM chirp signal: (A) Time domain plot. (B) The WVD. (C) The FT spectrum of the signal in subplot A. (D) FrFT Output of a chirp signal in the first subplot “A” for different values of the FrFT order α . The optimum value of α is 0.1900, where the peak is sharpest. (E) The WVD using the FrFT of the corresponding optimum value $\alpha = 0.1900$	83
Figure 5.3. Block diagram for FrFT-based FM chirp signal extraction.....	87
Figure 5.4. Example 1: <i>multi-component</i> linear FM signal; non-overlapping TF support. (A) Represents the time domain plot, and (B) time-frequency representation	

(the WVD).....89

Figure 5.5. Example 1 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{1opt} = 0.032$ of the first chirp signal. (B) WVD of the FrFT in subplot A. (C) FrFT after applying a rectangular window isolating the sharper spectrum. (D) WVD of the windowed signal in subplot C. (E) WVD of the reconstructed signal of the first chirp signal after applying the inverse FrFT with an order $-\alpha_{1opt} = -0.032$ 90

Figure 5.6. Example 1 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{2opt} = -0.155$ of the second chirp signal. (B) WVD of the FrFT in subplot A. (C) FrFT after applying a rectangular window isolating the sharp peak in the spectrum. (D) WVD of the windowed FrFT in subplot C. (E) WVD of the reconstructed signal of the second chirp signal after applying the inverse FrFT with an order $-\alpha_{2opt} = 0.155$ 91

Figure 5.7. Example 2: *multi-component* crossing linear FM signal; (A) represents the time domain representation, and (B) Time-frequency representation (WVD). 92

Figure 5.8 Example 2 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{1opt} = 0.215$ of the first chirp signal. (B) WVD of the FrFT in subplot A. (C) FrFT after applying a rectangular window for the sharper spectrum. (D) WVD of the windowed signal in subplot C. (E) WVD of the reconstructed signal of the second chirp signal after applying the inverse FrFT with an order $-\alpha_{1opt} = -0.215$..93

Figure 5.9. Example 2 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{2opt} = -0.157$ of the second chirp signal. (B) WVD of the FrFT in

subplot A. (C) FrFT after applying a rectangular window for the sharper spectrum. (D) WVD of the windowed signal in subplot C. (E) WVD of the reconstructed signal of the second chirp signal after applying the inverse FrFT with an order $-\alpha_{2opt} = 0.157$ 94

Figure 5.10. Example 3: *multi-component* non-linear non-crossing FM signal; (A) depicts the time domain representation, and (B) time-frequency representation (The WVD). (C) The FrFT of the signal in (A) after selecting the optimum order $\alpha_{opt} = -0.07$. (D) The WVD of the FrFT in subplot (C).....95

Figure 5.11. Example 3 (continued): (A) FrFT after applying a rectangular window corresponding to the first component. (B) WVD of the first component after applying an inverse FrFT on windowed spectrum. (C) FrFT after applying a rectangular window for the second component. (D) WVD of the second component after applying an inverse FrFT on windowed spectrum..... 96

Figure 5.12. Example 4: Application of the FrFT-based multi-component signal decomposition algorithm on real Bat data. (A) The WVD of the multi-component bat signal; the signal is composed of three chirp components with little overlap in their frequency content. Subplots (B), (C) and (D) show the WVD of the three decomposed components composing the bat signal..... 98

Figure 5.13. Example 5: Decomposition of multi-component overlapping and non-linear chirp signal. (A) The WVD of the multi-component signal. (B) The FT spectrum of the multi-component signal in A. (C) the FrFT with order $\alpha_{1opt} = 0.15$ for the first component $x_1(t)$. (D) The windowed FrFT of the first chirp signal. (E) The WVD of the first reconstructed chirp signal after applying an inverse FrFT with

$-\alpha_{1opt} = -0.15$. (F) The FrFT with order $\alpha_{2opt} = -0.085$ for the second chirp signal $x_2(t)$. (G) The windowed FrFT of the second chirp signal. (H) The WVD of the second reconstructed chirp signal $x_2(t)$ after applying an inverse FrFT with $-\alpha_{2opt} = 0.085$ 100

Figure 5.14. Example 6: power signal $x_1(t) = e^{j2\pi 0.014t^{1.5}}$. (A) The WVD of the original signal $x_1(t)$. (B) The time-warping function $w(t) = t^{1/k}$, for $k=1.5$. (C) The effect of the warping operator on the WVD of the warped signal. (D) The FT spectrum of the warped signal. 107

Figure 5.15. Example 7: two power phase terms $x_2(t) = e^{j2\pi(t^{0.1} + 0.85t^{0.7})}$. (A) the WVD of the original signal $x_2(t)$ in equation 5.19. (B) The instantaneous frequency curve $IF_{x_2}(t) \propto 0.1t^{-0.9} + 0.595t^{-0.3}$. (C) The effect of warping on the WVD using the power time-warping function $w(t) = t^{1/k}$ for $k=0.7$. (D) The FT spectrum of the warped signal with peak equal to 2.128.....109

Figure 5.16. Example 7 (continued) (A) The effect of warping on the WVD using the newly proposed warping function $\hat{w}(t) = e^{at}t^{1/k}$ with $k=0.7$ and $a=5.3214e-5$. (B) The FT spectrum of the warped signal with a peak equal to 2.729..... 112

Figure 5.17. Example 8: $x_3(t) = e^{j2\pi(c_1t + c_2t^2 + c_3t^3)}$, with $c_1 = 0.25$, $c_2 = -9.76e-4$ and $c_3 = 3.204e-6$. (A) The WVD of the original signal $x_3(t)$. (B) Instantaneous frequency curve. (C) The power warped TFR. (D) FT spectrum, respectively obtained using the power warping function $w(t) = t^{1/3}$. (E) The effect of the new time warping

function $\hat{w}(t) = e^{at^{1/3}}$. (F) The corresponding FT spectrum. with $k=3$ and $a=33.33e-4$.
 114

Figure 5.18. Example 9: $x_4(t) = e^{j2\pi(c_1t+c_2t^2+c_3t^k)}$, with $c_1 = 0.4558$, $c_2 = -0.0058$ and $c_3 = 2.2761e-5$. (A) The WVD of the original signal $x_3(t)$. (B) Instantaneous frequency curve. (C) The time-warping using the power warping function $w(t) = t^{1/3}$. (D) The corresponding FT spectrum (E) Time-warping using the new warping function $\hat{w}(t) = e^{at^{1/3}}$. (F) The corresponding FT spectrum. $k=3$ and $a=0.0033$ 115

Figure 5.19. Block diagram for decomposition of multi-component non-linear FM chirp signals using newly proposed time warping-based time-frequency representation algorithm 123

Figure 5.20. Example 10: Non-linear FM chirp signal composed of two components. (A) The WVD of the multi-component signal. (B) The FT spectrum 124

Figure 5.21. Example 10 (continued): Decomposition of multicomponent signal composed of two non-linear chirp components. (A) The FT spectrum of the warped signal $x(t)$ with $k = k_1 = 1.5$ and $a = a_1 = 1e-8$. (B) The FT spectrum after filtering the spectrum peak with larger amplitude. (C) The WVD of the reconstructed chirp component $\hat{x}_1(t)$. (D) The FT spectrum of the warped signal $x(t)$ with $k = k_2 = 1.3$ and $a = a_2 = 3.1e-3$. (E) The FT spectrum after filtering the spectrum peak with larger amplitude. (F) The WVD of the reconstructed chirp component $\hat{x}_2(t)$ 125

Figure 5.22. Comparison between the true instantaneous frequencies IF and the estimated IF of the two chirp components $x_1(t)$ and $x_2(t)$ in subplots A and B,

respectively.	127
Figure 5.23. Example 11: Non-linear FM chirp signal composed of three components.	
(A) The WVD of the multicomponent signal. (B) The corresponding FT spectrum	128
Figure 5.24. Example 11 (continued): Decomposition of multicomponent signal composed of three non-linear chirp components. (A) The FT spectrum of the warped signal $x(t)$ with $k = k_1 = 1.5$ and optimized $a = a_1 = 0$. (B) The FT spectrum after filtering the peak spectrum with larger amplitude. (C) The WVD of the reconstructed chirp component $\hat{x}_A(t)$. (D) The FT spectrum of the warped signal $x(t)$ with $k = k_2 = 1.3$ and $a = a_2 = 3.1e-3$. (E) The FT spectrum after filtering the peak spectrum with larger amplitude. (F) The WVD of the reconstructed chirp component $\hat{x}_B(t)$. (G) The FT spectrum of the warped signal $x(t)$ with $k = k_3 = 3$ and $a = a_3 = 6.7e-3$. (H) The FT spectrum after filtering the peak spectrum with larger amplitude. (I) The WVD of the third reconstructed chirp component $\hat{x}_C(t)$	
	130
Figure 5.25. Comparison between the true instantaneous frequency and the estimated IF of the three chirp components $x_A(t)$, $x_B(t)$ and $x_C(t)$ in subplots A, B and C, respectively	
	132
Figure 5.26. (A) The WVD of the signal $x(t) = e^{j2\pi(c_1t + c_2t^2 + c_3t^3)}$, with $c_1 = 0.25$, $c_2 = -9.76e-4$, $c_3 = 3.204e-6$ and $k=3$. (B) The effect of the new warping function $\hat{w}(t) = e^{at} t^{1/k}$ with $a = 33.33e-4$. (C) Spectrum of the warped signal for several values of k . (D) Zooming of subplot C. (E) The spectrum of the warped signal for several values of k	
	135

CHAPTER 1

INTRODUCTION

1.1 Epilepsy and Seizure

Epilepsy is one of the most common chronic neurological disorders that predispose individuals to experiencing recurrent seizures [1]. Seizures are a sudden, paroxysmal alteration of one or more neurological functions such as motor behavior, and/or autonomic functions [2]. Epilepsy is not a singular disease entity, but a variety of disorders reflecting underlying brain dysfunction that may result from many different causes [3]. Approximately 2% of the world populations are diagnosed with epilepsy. The occurrence of this brain malfunction is unpredictable, and may cause altered perception or behavior as sensory disturbances, or loss of consciousness. The negative influence of uncontrolled seizures, i.e. the patient will experience major limitations in family, social and educational activities, that extend beyond the individual to affect the whole society and may produce irreversible brain damage by time [16].

Seizures are subdivided into two sets: partial and generalized. In partial seizures, a limited brain area is implicated in the epileptic discharge. In contrast, generalized seizures originate from multiple brain regions and are characterized by general neurological symptoms [4]. Epilepsy is commonly treated with anti-epileptic drugs; but for some patients, medications are not enough to restrain their seizures. Thus they are candidates for surgery in order to remove the damaged brain tissue which requires

accurate localization. Because of the unknown time of occurrence of seizures, these patients undergo prolonged monitoring during which a variety of clinical examinations are performed [5].

Different types of seizure detectors were developed [4, 14, 15, 17, 20]. In this thesis, we will illustrate our new developed algorithms for seizure detection using the brain's electrical activity. Moreover, we will demonstrate the feasibility of using our algorithm for accurate and rapidly detecting seizure.

1.2 Electroencephalogram and Laplacian electroencephalogram

The Electroencephalogram or EEG is the recording of electrical activity variations from cortical neuronal activity [6]. The EEG, measured using non-invasive electrodes placed on the scalp, is referred to as a scalp EEG. When an EEG is measured using electrodes placed on the surface of the brain or within the brain it is referred to as intracranial EEG. In this study, scalp EEG signals have been used. The placement of EEG electrodes on the scalp usually follows a standard configuration known as the international 10-20 systems (Figure 1.1) suggested by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IFSECN) [18]. The “10” and “20” indicate that the distance between adjacent electrodes is either 10% or 20% of a specified distance measured using for example the total distance between the front and back or left and right of the head. Each electrode is labeled by specific letters and numbers. The placements of these electrodes are labeled according to the adjacent brain areas: F (frontal), C (central), T (temporal), P (posterior), and O (occipital) [18].

Figure 1.1: International 10-20 System of scalp electrode placement. Image obtained from www.sjopat.com/10-20-electrode-placement-pattern.

The EEG is an important tool in studying and diagnosing neurological disorders such as epilepsy, as it contains valuable information related to the different physiological states of the brain [7]. An abnormal EEG signal displays non-stationary behavior including spikes, sharp waves or spike-and-wave; so patients with epilepsy can have specific features in their EEG [4]. Seizures are manifested in the EEG as paroxysmal events characterized by stereotyped repetitive waveforms that evolve in amplitude, i.e. large amplitude or low amplitude, and high frequency before eventually decaying [8]. Moreover, due to its high temporal resolution and its close relationship to physiological and pathological functions of the brain, the EEG is considered as indispensable for seizure detection which uses significant parameters and dynamical changes of epilepsy patients. The EEG signals recorded during epilepsy can be classified as follows: (1) *interictal EEG* which refers to the period between seizures. This period comprises normal patterns, i.e. alpha rhythm, along with abnormal

patterns, i.e. spikes and high frequency oscillations. The EEG patterns during this period are used by neurologist when diagnosing epilepsy. (2) *preictal EEG*, refers to the period immediately before seizure. (3) *ictal EEG*, refers to the period where seizure patterns emerge. (4) *postictal EEG*, refers to the period immediately after the termination of the seizure [19].

Recently, improvements have been applied to EEG recording techniques, making it more accurate by increasing the spatial resolution. One such improvement is the application of the surface Laplacian to the EEG [9, 10]. The surface potential distribution on the scalp reflects functional activities evolving from the brain [11]. The variation in this surface can be recorded using an array of electrodes on the scalp, and then measuring the voltage between pairs of these electrodes. The scalp surface Laplacian is an alternative method for presenting EEG data with higher spatial resolution. It has been shown that the surface Laplacian (the second spatial derivative) is proportional to the cortical potentials and increases the high spatial frequency components of the brain activity near the electrode [12]. To obtain the Laplacian, a new approach was considered by using unique sensors and instrumentation for recording the signal [9, 13]. The unique sensor configuration which measures the Laplacian potential directly is the Tripolar Concentric Ring Electrode (TCRE) depicted in Figure 1.3-B. Laplacian EEG or tEEG was defined in Besio et al. [9, 13] as $16 * (V_m - V_d) - (V_o - V_d)$ where V_m , V_o and V_d are the voltage of the middle ring, outer ring, and central disc of the tripolar electrode, respectively. Figure 1.3-A shows the traditional disc electrode. Koka and Besio [10] showed that TCRE provides approximately four times improvement in the signal-to-noise ratio, three times

improvement in spatial resolution, and twelve times improvement in mutual information compared to disc electrode signals. The TCRE also exhibit strong attenuation of common mode artifacts [10]. These findings suggest that tEEG may be useful for seizure detection or other neurological disorders analysis [14, 15]. An example of a seizure pattern in Laplacian EEG (tEEG) signal is shown in Figure 1.2. In this thesis, EEG and tEEG signals were used as a tool for seizure detection and analysis.

Figure 1.2 An example of a seizure pattern in Laplacian EEG (tEEG) signal.

1.3 Methods for seizure detection

It is unquestionable that a method capable of detecting the occurrence of seizures with high accuracy would significantly improve the therapeutic possibilities and thereby the quality of life for epilepsy patients [16]. For this purpose and others automatic seizure detection techniques have received intense attention in the recent past. The majority of the existing seizure detection techniques are designed for the

scalp EEG where important information that allows seizure detection are measured and extracted from EEG signals. The usual methods for seizure detection are based on data analysis by observing long recordings of continuous EEG signals. Long-term EEG recordings increase the possibility to capture and analyze seizure events and also augment detection and diagnosis [4]. However, it is a very monotonous and time-consuming process [17], because there are large amounts of data to analyze and the presence of artifacts may lead to false positive detections. For this purpose, automated seizure detection methods with high sensitivity and low false positive rates are of great significance in recognizing and reviewing EEG for epileptic seizure detection.

Figure 1.3 An example of conventional disc electrode, (A), and Tri-polar concentric ring electrode, (B).

Since the 1970s, automated seizure detection has been a challenge with several algorithms and methods developed [4, 7, 14, 15, 20-22] but no detector dominates with excellent sensitivity and specificity. This may be due to noise and artifacts such as eye movement and muscle activity, which make detection more difficult [4]. The

first algorithm developed for automatic seizure detection is the original work of Gotman [17]. Since then, there has been a marked increase in seizure-related research using signal processing techniques used for selection of discriminative features related to the presence of seizure.

The general procedure for developing methods for seizure detection is composed into two main steps: features extraction and classification. The purpose of features extraction is to determine which information or patterns are necessary from the EEG signal to identify a seizure event from non-seizure. Extracted features depend significantly on the method used to extract the hidden information in the signal. Some of the methods are based on morphological characteristics of epileptic EEG recordings such as features for seizure detection [4] where epileptic seizures are often visible in EEG recordings as rhythmic discharges or multiple spikes [4]. For example, the Gotman algorithm successfully detects seizures with sustained rhythmic activity with a fundamental frequency below 20 Hz [17]. He developed an algorithm that first breaks down the EEG signal into half-waves. Then morphological characteristics of these half-waves, such as amplitude and duration, were used to determine the existence of epileptic seizures [17, 23]. For rhythmic discharges: fast Fourier transform based features [4, 24 and 25], frequency domain based features [4, 26-29], time-frequency based features [4, 21, 22, and 30], or wavelet based features [4, 31-35] have often been used.

1.4 Thesis Contributions

The major contributions of the thesis are summarized as follows:

- Developing an algorithm for seizure detection maintaining high sensitivity and specificity with fast execution time using time domain features [14]. The proposed algorithm was tested using human and animal data containing seizure and was compared against popular seizure detection techniques.
- Developing a new technique for seizure detection using time-frequency representation while reducing effects of noise and artifacts [15].
- Developing an algorithm for seizure detection showing the superior performance of the Renyi entropy comparing to other entropies [56].
- Comparing the performance of the Empirical Mode Decomposition (EMD) algorithm and the Discrete Wavelet Transform (DWT) for seizure detection [56].
- Developing a new warping function that can be successfully used to linearize the TF structure of non-linear Frequency Modulated (FM) chirp signal.
- Developing a new warping-based multi-component signal decomposition algorithm to decompose multi-component signals consisting of non-linear FM chirp signals.

1.5 Thesis Organization

The thesis consists of six chapters. The organization of the thesis is as follows:

Chapter 1 is the introductory chapter that reviews the definition and the basic concepts of epilepsy, seizure, EEG and outlines the previous approaches for seizure detection as well as the major contributions of the thesis.

Chapter 2 briefly reviews the fundamentals of time-frequency (TF) signal processing and summarizes the ability of TF in analyzing non-stationary signals, like electroencephalography. The usefulness of time-frequency Reassignment and Synchrosqueezing for cross-terms reduction will be discussed.

Chapter 3 describes two newly developed algorithms for seizure detection using the scalp EEG signal. The first algorithm was based on time-domain features and the second algorithm was based on the measurement of signal complexity using entropies from different spectral components. Using these sets of features, the proposed algorithms can classify the contaminated data with low computational complexity while still maintaining very good accuracy. The performance of the proposed algorithms was evaluated and compared against several other brain seizure algorithms previously reported in the literature.

Chapter 4 describes a new technique for seizure detection using time-frequency signal processing with application methods to reduce noise and artifacts. The obtained results show that the proposed method using time-frequency analysis enhances previous techniques in the literature for seizure detection.

Chapter 5 gives an overview of the Fractional Fourier transform (FrFT) and its properties. In this chapter, the FrFT was used to decompose a multi-component signal

consisting of either linear or non-linear FM chirp components. In this chapter, a brief presentation of the unitary transformation principle was discussed with an introduction of the warping functions which are used to transform the non-linear time-frequency content of several FM chirp signals into an equivalent linear signal. A novel time-warping function was proposed that can be successfully used to linearize the TF structure of a non-linear FM chirp signal with a polynomial phase. It is demonstrated that the inverse of the warping function can be found using the Lambert W function. This warping function and its one-to-one inverse can be used to formulate a new class of warped Time-Frequency Representations (TFRs) that is a generalization of the k^{th} Power Class and the Exponential Class of warped TFRs. A new algorithm using the principle of time warping-based time-frequency representation and parameter optimization was proposed to decompose multi-component signals composed of non-linear FM chirp signals.

Finally, Chapter 6 concludes the thesis and provides suggestions for further research.

CHAPTER 2

TIME-FREQUENCY SIGNAL PROCESSING

2.1 Introduction

The objective of many signal processing applications in the real world is to explore the behavior of recorded signals to analyze its components [107, 109]. The main goal of signal processing is to extract useful information from the signal by transforming it and involves techniques that improve our understanding of information contained in this signal. For example, the EEG has become one of the most important diagnostic tools in the neurophysiology area, but until now, EEG analysis still relies mostly on its visual inspection. For this purpose, different signal processing techniques were used in order to quantify the information of the EEG. Among these, the time-frequency analysis methods emerged as a very powerful tool capable of characterizing the energy of the EEG and other signals over time and frequency. Time-frequency analysis is considered as an area of active research in the signal processing domain. Thus, representing the signal in time-frequency domain is very helpful for concentrating on the components of interest in the signal.

2.2 The need for Time-Frequency Analysis

The traditional way to analyze a signal is time domain analysis or frequency domain analysis. The time domain is a record of the system parameters versus time.

Analyses of a signal in the time domain makes it difficult to provide any information about the distribution of energy over different frequencies and the frequency variations over time. On the other hand, using frequency domain analysis, the information about the frequency content of the signal can be extracted and analyzed. The problem with frequency domain analysis is the difficulty to recognize when these frequency components occurred in time. Therefore, there is a need to describe how the spectral content of a signal changes with time. In contrast with the time and frequency domains for signal analysis, time-frequency (TF) techniques analyze the signal in both time and frequency domains [106-109]. This means that TF techniques show changes of the signal's frequency components with respect to time to yield a potentially more revealing picture of the temporal localization of a signal's spectral components [40]. To overcome the limitations of analysis in the time domain alone or of the frequency domain alone, a number of time-frequency analysis methods have been introduced [40, 48, 106-109].

In real life, in many applications such as seismic surveying, communications, radar, and sonar, the signals are non-stationary, which means that the frequency content of those signals are varying with time. Time-frequency representations (TFRs) have been applied to analyze, modify and synthesize non-stationary or time-varying signals [40, 105, 107-109]. Furthermore, TFRs can handle multi-component signals and outperform existing techniques which are only applicable for mono-component signals [36]. Figure 2.1-(A) and 2.1-(B) shows a mono-component linear FM signal, i.e. chirp signal, and its frequency spectrum, respectively. The chirp has linear increasing frequency modulation from 0.1 to 0.4 Hz. The sampling frequency is $F_s=1$

Hz and the number of frequency points (bins) is $N=128$. The frequency spectrum demonstrates the frequency distribution of the signal for the whole length of the signal without any information of the time of appearance and the duration of each spectral component. The TFR of the signal in subplot (C) suggests more appropriate information about the behavior of the signal, clearly indicating that the spectral content of the signal is changing linearly with time.

Figure 2.1 *mono-component* linear FM signal; (A) represents the time domain. (B) Magnitude spectrum and (C) time-frequency representation.

Figure 2.2-(A) shows the sum of several linear FM signals, otherwise known as a multi-component signal. The first chirp signal has an increasing frequency from 0.01 to 0.25 Hz and the other from 0.35 to 0.5 Hz. Figure 2.2-(B) displays the spectrum of the signal as it appears in the frequency domain. From the time domain, it is difficult to identify the number of signal components, and extremely difficult to determine their nature. One way to determine these signal characteristics is a time-frequency representation which makes signal analysis very easy. A TFR of this signal is shown in Figure 2.2- (C). From this representation, it is evident that there are two chirps

comprising the signal shown in Figure 2.2-(A). Moreover, the linear frequency nature of the chirps is easily determined as well. The intuitive time-frequency representations often can easily be understood by someone with little signal analysis experience.

In addition to all the advantages of TF analysis cited above, TFR can significantly simplify the interpretation of signals due to their ability to display frequency information that changes over time. Thus, TF analysis has been developed for a wide range of problems with signals that contain highly localized events such as bursts, spikes, and discontinuities, which typically occur in EEG signals during seizures. Moreover, using TF analysis, the energy content of an EEG signal can be visualized. This can help understanding the characteristics of the EEG signal in order to determine the best approach to analyze and process the signal.

Figure 2.2 *multi-component* linear FM signal. (A) The time domain plot. (B) Magnitude spectrum and (C) time-frequency representation.

2.3 Types of Time-Frequency Representations

There are many different time-frequency Representations (TFRs) available in the literature for analysis [37-43, 106-109] and each TFR has its own strengths and

weaknesses. Some of these TFRs provide excellent resolution, but take long to calculate. Others provide poor resolution, but can be calculated quickly. Furthermore, some TFRs provide a good balance of speed and resolution, but due to their quadratic nature, have interference terms and cross-terms. TFRs can be divided into three main classes: linear classes, quadratic classes and higher order classes [40, 106-109].

2.3.1 Linear Time-Frequency Representations

Linear TFRs satisfy the linearity principle which states that if $x(t)$ is a linear combination of some signal components, then the TFR of $x(t)$ is also a linear combination of the TFRs of each of the signal components [40, 109]. Two important linear TFR are the short-time Fourier transform (STFT) [40, 107, 108] and the Wavelet Transform (WT) [40, 107, 108].

2.3.1.1 Short time Fourier Transform (STFT)

The Fourier Transform (FT) is a reversible, linear transform with many significant properties. The FT decomposes a signal into a set of weighted harmonic components with fixed frequencies. To compute the Fourier transform, the signal must have a finite energy. The Fourier transform of a signal $x(t)$ is given by:

$$X_x(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt . \quad (2.1)$$

Figure 2.3 Spectrogram using a Hamming window with (A) 15-point length. (B) 64-point length, or (C) 128-point length.

It has been proven that analysis of non-stationary signals by the FT does not bring complete information about spectral content [44, 45]. In order to add time-dependency in the FT, a time-domain signal is divided into a series of small pieces, where the signal is windowed into short segments and then the FT is applied to each segment. The new transformation is called the Short-Time Fourier transform (STFT) and it maps the signal into the TF plane. The STFT assumes that the signal is quasi-stationary, i.e. stationary over the duration of the window. The STFT of a signal $x(t)$ is given by:

$$STFT_x(t, f; h) = \int_{-\infty}^{+\infty} x(\tau) h^*(\tau - t) e^{-j2\pi f \tau} d\tau \quad (2.2)$$

where $h(t)$ is the sliding analysis window and $h(\tau - t)$ is the same window, time reversed and shifted centered by the output time t . The spectrogram can then be obtained by squaring the STFT modulus, $|STFT_x(t, f; h)|^2$; the spectrogram is a quadratic TFR. One problem with the STFT is the width of the window h . There is a trade-off between the choices of good time versus good frequency resolution. Good

time resolution requires short duration windows, whereas good frequency resolution requires long duration windows as shown in Figure 2.3 which represent the spectrogram of a linear FM chirp with increasing frequency from 0.1 to 0.4 Hz. We see that the short window used in (A) produce smearing in the frequency direction whereas the long window used in (C) produces smearing the time direction .Therefore, it is difficult to achieve good localization simultaneously in both the time and the frequency domains since the STFT depends only on one window. This limitation is related to the Heisenberg uncertainty principle [46, 47]. It was proven that a Gaussian window is the only window that meets the lower bound of the Heisenberg principle; the Gabor transform [40] is similar to a STFT computed using a Gaussian window.

2.3.1.2 Wavelet Transform (WT)

The Wavelet transform, a time-scale representation, is another linear TF distribution. The WT is similar to the STFT but instead of a fixed duration window function, the WT uses a varying window length by scaling the axis of the window. The WT of signal $x(t)$ is defined as [40, 108]:

$$WT_x(b, a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) g^* \left(\frac{t-b}{a} \right) dt \quad (2.3)$$

The WT is the convolution of a signal $x(t)$ with a window $g(t)$ shifted in time by “ b ” and dilated by a scale parameter “ a ” [40, 48, 108]. The parameter “ a ” can be chosen such that it is inversely proportional to frequency to obtain a TF representation comparable to the STFT [40]. For this reason, at low frequencies the WT provides high spectral resolution but poor temporal resolution. On the other hand, for high

frequencies, the WT provides high temporal resolution which enables the WT to “zoom in” on singularities. Unfortunately, high frequencies will have poor spectral resolution. This means that for the WT there is also a tradeoff between time and frequency resolution. The energy density function of a WT is defined as $|WT_x(b, a)|^2$ and is called the Scalogram which is a quadratic TFR [40, 108].

2.3.2 Quadratic Time-Frequency Representations

2.3.2.1 Wigner-Ville Distribution (WVD)

The Wigner-Ville distribution (WVD) is considered as a TF representation that attains good tradeoff between time versus frequency resolution [40, 105, 106]. The WVD of a signal $x(t)$ is given by:

$$WVD_x(t, f) = \int_{-\infty}^{+\infty} x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-j2\pi f\tau} d\tau \quad (2.4)$$

where, $x^*(t)$ is the complex conjugate of $x(t)$. The WVD can provide very good resolution in time and frequency of the underlying signal structure because of its interesting properties such as preserving frequency support, instantaneous frequency, group delay, etc. [40, 48, and 49]. However, because of the bilinear nature of the WVD, and due to the existence of negative values, the WVD has misleading TF results in the case of multi-component signals (e.g., the EEG) due to the presence of cross terms and interference terms [40, 48, 105]. For example, the WVD of the multi-component signal $\tilde{x}(t) = x_1(t) + x_2(t)$ is:

$$WVD_{\tilde{x}}(t, f) = WVD_{x_1 + x_2}(t, f) = WVD_{x_1}(t, f) + WVD_{x_2}(t, f) + 2\text{Re}[WVD_{x_1 x_2}(t, f)] \quad (2.5)$$

Figure 2.4 Wigner-Ville distributions (WVD) of (A) mono-component parabolic FM chirp signal, (B) Multi-component signal with two cross linear chirp signals.

The first two terms, $WVD_{x_1}(t, f)$ and $WVD_{x_2}(t, f)$, are the WVD of the signals $x_1(t)$ and $x_2(t)$, respectively, and they are called auto-terms. The last term $WVD_{x_1x_2}(t, f)$ is the cross WVD of $x_1(t)$ and $x_2(t)$; it is given by:

$$WVD_{x_1x_2}(t, f) = \int_{-\infty}^{+\infty} x_1\left(t + \frac{\tau}{2}\right) x_2^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau. \quad (2.6)$$

Cross WVD terms can be reduced by using proper smoothing kernel functions as well as analyzing the analytic signal (instead of the original signal) to solve the problem of cross terms produced by negative frequency components. The analytic signal is given by: $z_x(t) = x(t) + jH[x(t)]$, where $H[.]$ is the Hilbert transform defined as:

$H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(t-\theta)}{\theta} d\theta$. Thus, equation (2.4) will be:

$$WVD_z(t, f) = \int_{-\infty}^{+\infty} z_x\left(t + \frac{\tau}{2}\right) z_x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \quad (2.7)$$

where, $z_x(t)$ is the analytic signal associated with the signal $x(t)$. Figure 2.4 shows two examples of signals with their corresponding WVD plots. The subplot (A) represent a mono-component parabolic FM chirp and subplot (B) represent a multi-component signal with two linear FM chirps, one with increasing frequency support from 0.1 to 0.4 Hz and the other with a decreasing frequency support from 0.4 to 0.1 Hz. The oscillatory terms in (A) are referred to as inner interference terms occurring in a convex mono-component signal whereas those in (B) are referred to as the cross terms between a pair of signal components in a multi-component signal. To avoid the problem of oscillatory cross terms or inner interference terms, smoothed versions of the WVD were introduced [40, 50, 105, 109].

2.3.2.2 Smoothed-Pseudo Wigner-Ville Distribution (SPWVD)

To address the problem of cross-terms suppression while keeping a high TF resolution, other TFRs have been proposed. Among these is the smoothed pseudo WVD (SPWVD) [50]. The SPWVD permits two independent analysis windows, one in time and the other in the frequency domains to improve the readability of the Wigner-Ville distribution. The SPWVD of a signal $x(t)$ is given by:

$$SPWVD_x(t, f; g, h) = \int_{-\infty}^{+\infty} h(\tau) \int_{-\infty}^{+\infty} g(s - t) x(s + \frac{\tau}{2}) x^*(s - \frac{\tau}{2}) ds e^{-j2\pi f\tau} d\tau \quad (2.8)$$

Figure 2.5 Smoothed-pseudo Wigner-Ville distributions (SPWVD) of **(A)** mono-component parabolic chirp signal, **(B)** Multi-component signal with two cross linear chirp signals. The inner interference terms in (A) have been reduced.

where t is the time variable, f is frequency, h is the frequency smoothing window and g is the time smoothing window. Figure 2.5 shows the SPWVD of the same signals as in Figure 2.4. Two Hamming windows with 5-point length and 1023-point length were used respectively, for the time and frequency smoothing windows.

2.4 A comparison of the existence of cross-terms in the WVD, the Spectrogram and the Scalogram

It was shown from the previous section that although the WVD has very important properties compared to other TFRs, the presence of cross-terms makes the WVD difficult to interpret. These cross-terms interfere with, and often mask, the true TF information and could lead to misinterpretations of the signal energy concentration and misreading of the TF signature of the corresponding signal. The cross-terms of the WVD have been extensively analyzed [40, 48, 105]. It has been found that the WD cross-terms lie at the mid-time and mid-frequency of each pair of auto-components;

they are highly oscillatory and can have amplitudes twice as large as the product of the magnitudes of the WVD of the two signals under consideration [40, 48].

Often practitioners use the STFT spectrogram or WT Scalogram claiming they are free of the problems of cross terms found in the WVD. However, Kadambe and Boudreaux-Bartels [48] showed that the cross-terms comparable to those found in the WVD exist when considering the energy distributions of the STFT (spectrogram) and the WT (Scalogram). By deriving the mathematical expressions for the energy distributions of the STFT the authors deduced that [48]: (1) the STFT cross-terms occur at the intersection of the respective transforms of the two signals under consideration, unlike the WD cross-terms which always occur at mid-time and mid-frequency of the two WVD auto components. Thus, the Spectrogram and Scalogram of an 'n' component signal can have a minimum of zero cross-terms and a maximum of $\binom{n}{2}$, unlike the WVD which always has $\binom{n}{2}$ cross-terms. Here, $\binom{n}{2}$ is equal to a combination of n things taken 2 at a time. (2) The STFT cross-terms are oscillatory in nature similar to the WVD cross-terms. (3) The STFT cross-terms can have a maximum magnitude equal to twice the product of the magnitude of the transforms, again similar to the WVD cross-terms. For the $|WT|^2$, and similar to the $|STFT|^2$, cross-terms appear for multi-component signals. Moreover, since the energy distributions of the $|WT|^2$ and $|STFT|^2$ are equivalent to the smoothed affine WVD and a smoothed WVD [48], respectively, then the nature and geometry of the $|WT|^2$ cross-terms are similar to the $|STFT|^2$ cross-terms.

Figure 2.4-B shows the effect of cross-terms of the WVD when applied to multi-component signals. Even though the cross-terms are significant, the frequency support is still well concentrated along the linear instantaneous frequency axis.

2.5 Reassignment Time-Frequency Representation and Synchrosqueezing Transform

2.5.1 Reassignment Time-Frequency Representation

The aim of the Reassignment method is to sharpen the TFR of a signal while keeping the temporal localization correct. This method is well adapted for multi-component signals [51, 108, 109]. Moreover, Reassignment TFRs have been used to diminish cross-terms and enhance time-frequency concentration of auto-terms [51, 52]. It was proven that the Reassignment method perfectly localizes linear chirps while removing most of the inner interference terms [52]. This method has been generalized to all the bilinear time-frequency and time-scale distributions [52]. The reassignment SPWVD (RSPWVD) offers good TF resolution and good interference reduction. The RSPWVD is given by [52]:

$$RSPWVD_x(t', f'; g, h) = \int_{-\infty}^{\infty} SPWVD_x(t, f; g, h) \delta(t' - \hat{t}(t, f)) \delta(f' - \hat{f}(t, f)) dt df \quad (2.9)$$

$$\begin{aligned} \hat{t}(t, f) &= t - \frac{SPWVD_x(t, f; t \times g, h)}{2\pi \times SPWVD_x(t, f; g, h)} \\ \text{with} \quad \hat{f}(t, f) &= f + j \frac{SPWVD_x(t, f; t \times g, \frac{dh}{dt})}{2\pi \times SPWVD_x(t, f; g, h)} \end{aligned} \quad (2.10)$$

where $x(t)$ is the signal, t is the time variable, f is the frequency variable; $g()$ and $h()$ are the smoothing time and frequency windows, respectively. Figure 2.6 shows the

reassignment plot of the SPWVD shown in Figure 2.5. The inner interference terms in Figure 2.6 (A) are significantly reduced while maintaining the concentrated support along the parabolic instantaneous frequency curve.

Figure 2.6 Reassigned Smoothed-Pseudo Wigner-Ville distributions (RSPWVD) of (A) mono-component parabolic chirp signal, (B) Multi-component signal with two cross linear chirp signals.

2.5.2 Synchrosqueezing Transform

Synchrosqueezing, introduced by Maes and Daubechies [55], is an invertible time-frequency analysis tool designed to decompose signals into constituent components with time-varying oscillatory characteristics. The reassignment method is a good analysis tool as it is a powerful representation of multi-component signals that results in high energy concentration in TF plane. However, it cannot be used for synthesis or reconstruction of the signal's individual sub-components. A recent study [53] shows that the Synchrosqueezing transform is an alternative to the Empirical Mode Decomposition (EMD) method [54], used for analyzing and decomposing natural signals, and developed with more theoretical foundation. This means that the

Synchrosqueezing transform can extract and delineate components with time-varying spectra and allow for the reconstruction of these components [51]. Figure 2.7 shows the squared magnitude of Continuous Wavelet Transform (CWT) Synchrosqueezing Transform of a parabolic FM chirp signal, subplot (A), and of a multi-component linear FM chirp signal, subplot (B). The concentration in the TF plane is very sharp and significant cross term reduction has occurred in subplot (A). However the X shape of the two linear FM chirps in subplot (B) is difficult to see for low frequencies.

Figure 2.7 Synchrosqueezed Wavelet Transform Scalogram of (A) mono-component parabolic chirp signal. (B) Multi-component signal with two cross linear chirp signals.

2.6 Summary

Time-frequency representations (TFRs) have found important applications for analysis, synthesis and detection of non-stationary signals by combining time and frequency information. Moreover, TFRs help identify important features of the signal being analyzed, such as, the number of signal components and the regions of energy concentration. The TFRs are divided into three main groups: (1) linear time-frequency distributions such as the Short-Time Fourier transform (STFT) and the Wavelet transform (WT), (2) Quadratic time frequency representation such as the Wigner-Ville distribution (WVD), the smoothed-pseudo Wigner-Ville distribution (SPWVD), the spectrogram, the Scalogram and (3) non-linear iterative algorithms like Reassignment and Synchrosqueezing or higher order TFRs. The linear distributions are typically efficient and easy to construct but create a tradeoff between good time resolution versus good frequency resolution. On the other hand, the quadratic distributions provide better resolution but suffer from higher computational cost and the appearance of inner-interference or cross terms between concave or multiple signal components.

Among the developed TFRs, the WVD has attracted much attention because of its several desirable mathematical properties such as preserving time and frequency marginals and shifts. Moreover, the WVD offers high frequency resolution of mono-component linear FM signal energy. However, in spite of these advantages, the WVD creates spurious frequency information or cross-terms interferences when it is applied to multi-component signals. Although the STFT and WT are linear TFRs, most analysis uses their squared magnitudes, e.g. spectrogram and Scalogram, respectively.

It was proven that the existence of cross-terms in these two transforms are comparable to those found in the WVD, they just occur in different regions of the TF plane.

To deal with the problem of cross-terms, the non-linear and iterative Reassignment time-frequency representation and Synchrosqueezing were introduced as new TF analysis of non-stationary and multi-component signals. The Synchrosqueezing transform is a kind of reassignment method that aims to sharpen a time-frequency representation, while remaining invertible as in the EMD. After all these TFRs we discussed, it is important to note that: *There is no one “best Time-Frequency Method” for all applications, but, there is one best Time-Frequency Method for each application.*

CHAPTER 3

SEIZURE DETECTION USING THE EEG SIGNAL

3.1 Introduction

With a goal of reducing the time-consuming nature of and automating the visual analysis of electroencephalogram or EEG signals during seizure analysis, we developed two different types of algorithms for detecting the occurrence of seizures in data recorded using a scalp EEG. For both algorithms, the detection of seizure events is treated as a classification task to differentiate between data with seizure versus data without seizure. The classification involves two steps. First, figure out which kind of patterns we want to extract from the signal. This step is known as feature extraction; the feature vectors used by the proposed algorithms depend considerably on the method used. Next, a classifier will be used to distinguish the difference between feature vectors and classify the data with seizure from the data without seizure. The achievement of high classification accuracy depends on the features extracted, as well as on the classifier used to determine the class relationship.

3.2 EEG data acquisition

The data used for seizure detection are obtained from two different sources - (1) University of Rhode Island (URI), Kingston, RI USA, recording data from rats, and

(2) Bonn University, Germany, using Human data. The databases from these sources are described in the next section.

3.2.1 Rats' dataset

The first dataset consists of scalp Laplacian EEG (tEEG) data acquired from rats at the University of Rhode Island (URI). The animal protocol used for recording the tEEG signal was approved by the University of Rhode Island IACUC. Approximately 24 h before the induction of seizures caused by pentylenetetrazole (PTZ), an adult male Sprague–Dawley rat weighting 220–320 g was given a combination of 80 mg/kg of ketamine and 12 mg/kg xylazine (i.e.) for anesthesia. The scalp was shaved and prepared with NuPrep abrasive gel (D. O. Weaver & Co., Aurora, CO, and USA). Three tripolar concentric ring electrodes (TCREs) [9, 10] were applied to the scalp (Figure 3.1-A, one for stimulation and two for recording) using conductive paste 0.5 mmTen20, Grass Technologies, RI, USA, and adhered with Teets dental acrylic (Pearson Lab Supply, Sylmar, CA, USA). The TCRE number (1) centered on the top of the head, with diameter equal to 1 cm, and the width of each ring equal to 0.9 mm, was used to record from and stimulate the brain. The front edge of the electrode was placed as close as possible to where the bregma was expected to be. The two other electrodes (2) and (3) with diameter equal 6 mm and ring width equal 0.4 mm were placed bilaterally behind the eyes, but in front of the ears. An isolated ground electrode (r) was placed on the top of the neck behind the ears. The electrodes were made of gold-plated copper. For approximately 24 h the rats were returned to their

cages and allowed food and water ad libitum. On the next afternoon, the rats were placed in a transparent plastic cage and via a commutator and cables (Plastics

Figure 3.1 Typical tEEG data recording from a rat using tripolar concentric ring electrodes (TCRE); (A): The location of the TCRES on the rat scalp. Electrode (1) is 10 mm dia. and used for stimulation and recording. Electrodes (2) and (3) are both 6.0 mm dia. and used only for recording. Electrode (r) is the isolated ground. Details of TCRES are shown to the right of the rat head (B): 30 min of data with and without seizure (with 460,800 samples). After 5 min of Baseline recordings the PTZ was administered. Soon after giving the PTZ, usually within 2 min, the rats had their first myoclonic jerk. (C): thirty seconds Baseline data from early region within dashed rectangle in (B) (note that the vertical axis is magnified compared to panels (B) and (D)), (D): thirty seconds Seizure data from later dashed rectangular region in (B).

One, Roanoke, VA) the electrodes were connected to a multiplexer. This multiplexer was first set to connect the electrodes to a Prep-Check Plus EIM-107 (General Devices, Ridgefield, NJ) to measure the skin-to-electrode impedance. The EEG

recording and the video were started. After 5 min of recording of baseline EEG, the PTZ was given (55 mg/kg, i.e.) to induce a seizure. The EEG signals were pre-amplified (gain 100 and 0.3 Hz high-pass filter) with a custom built preamplifier and then amplified using a Grass NRS2 Neurological Research System with Model 15A54 AC amplifiers (Grass Technologies, West Warwick, RI, USA) with a gain of 1000 and band pass of 1.0 – 100 Hz with the 60 Hz notch filter active, and digitized (16 bits, 256 Samples/s). The digitized signals were stored on a computer for offline analysis using MATLAB (Mathworks Natick, MA, USA). The two differential signals from the electrode elements (outer ring, inner ring, and center disc) were combined using an algorithm to give a Laplacian derivation of the signal as described by Besio [9, 10]. The algorithm weights the difference between the middle ring and center disc sixteen times greater than the difference between the outer ring and the center disc.

The tEEG data of ten rats have been recorded using the procedure described above. The first five minutes from each recorded signal are considered as baseline or non-seizure data. Firstly, the tEEG data were divided into 30 second segments with each segment containing 7680 samples (256 samples per second for 30 s). The selection of the Seizure segments, i.e. segments during which a seizure occurred, was performed by an experienced behavioralist through visual inspections of the video recordings. Because of a large amount of artifacts and noise caused by grooming, chewing, and roaming of the rats during the recording, baseline segments were selected after visual inspection where the tEEG appeared to be calm and artifacts free. Two sets of tEEG data corresponding to the Baseline data and Seizure data were used as the investigational data set for seizure detection. The database included 70 Seizures

segments and 65 Baseline or non-seizure segments comprising the total number of Seizures and Baseline segments recognized by the behavioralist. Figure 3.1 gives an example of dataset decomposition into seizure versus baseline or non-seizure segments.

3.2.2 Human dataset

The second dataset consists of a subset of data recorded on humans described by Andrzejak et al. [2]. The database consists of five sets (denoted as Z, O, N, F and S) each containing 100 single channel EEG segments of 23.6 s duration. The data were all recorded with the same 128-channel amplifier system and digitized at 173.6 Hz sampling rate and 12 bit A/D resolution. The bandwidth of the acquisition system was from 0.5 to 85 Hz. These segments were selected and cut out from continuous multi-channel EEG recordings after visual inspection where the data appear to be artifacts free [2]. Sets Z and O have been recorded extracranially from five healthy volunteers using electrodes placed according to the international 10–20 system locations [18]. Volunteers were relaxed in a wake state with eyes open (Z) and eyes closed (O), respectively. Signals in subsets (F) and (N) have been measured in seizure-free intervals, from five patients in the epileptogenic zone (F) and from the hippocampal formation of the opposite hemisphere of the brain (N). Subset (S) contains seizure activity, selected from all recording sites exhibiting ictal activity. Sets (N), (F) and (S) have been recorded intracranially. More specifically, depth electrodes are implanted symmetrically into the hippocampal formation. EEG segments of subsets (N) and (F) were taken from all contacts of the relevant depth electrode [2]. In addition, strip

electrodes are implanted onto the lateral and basal regions (middle and bottom) of the neocortex. EEG segments of the subsets S were taken from contacts of all electrodes (depth and strip). One EEG segment from each category is shown in Figure 3.2.

Figure 3.2 An example of human EEG signals from each of the five subsets (Z, O, N, F, and S)

3.3 Seizure detection using Time-Domain features

In the EEG recording, seizures appear as a sudden redistribution of spectral energy on a set of EEG channels. To automate the detection of these changes in the scalp EEG signal during seizure activity, we proposed an algorithm which enables us to understand the differences between a seizure period and a non-seizure period of different patients by finding a subset of time-domain features capturing the seizure characteristics of each patient. These features will be used to classify whether a given data segment contained seizures or not. The procedure of the algorithm will be described in the next sections.

3.3.1 Data preprocessing

In this stage, the rats' data were preprocessed before they are used. Before starting the analysis procedure for seizure detection, the tEEG signals are down-sampled from 256 to 128 Hz. This reduced the amount of data samples and thus the computation time of the algorithm. Before down-sampling, the signals were filtered first using an anti-aliasing low-pass filter with a cutoff frequency 64 Hz, to avoid aliasing. The MATLAB function `downsample` was used for the down-sampling procedure.

Note that the human dataset will be used as it is without any preprocessing step.

3.3.2 Feature extraction

A common approach in seizure detection and also in seizure prediction is to extract information or patterns; in other words, features that can characterize seizure morphologies from EEG recordings. Prior to feature extraction, the tEEG data, rats' data, and EEG data, human data, were segmented into one second duration epochs using non-overlapping Hamming window. During this research, we tested different segment lengths, e.g. 0.5s, 1s, 2s,...,. Among those we tested, the one second EEG epoch achieved optimal detection accuracy. Features used were: (1) the approximate entropy (ApEn) [62] (2) the maximum singular value (MSV) [64], and (3) the median absolute deviation (MAD) [65], which will be described in the next subsection.

3.3.2.1 Feature 1: Approximate Entropy (ApEn)

Entropy of a signal is a measure of the information contained in that signal [62]. It follows that entropy is also a measure of uncertainty of random variables or a complexity measure of a dynamical system. ApEn, which is the acronym for the

approximate entropy [20, 61], is a nonlinear dynamical analysis that quantifies the degree of complexity and irregularity in a time series data set such as estimation of regularity in epileptic seizure time series data [62]. ApEn is less sensitive to noise and can be used for short-length data [20, 61]. The described methods by Dimbra et al. [59] and Kumar et al. [61] have shown that the value of the ApEn drops abruptly due to the synchronous discharge of neurons during an epileptic activity. Hence, it is a good feature in the automated detection of seizures.

Given N points, the $ApEn(m, r, N)$ is equal to the negative of the average natural logarithm of the conditional probability that two similar sequences for m points stay analogous, that is, within a tolerance r , at the next point [63]. The process of estimating ApEn is described in the following steps taken from [20, 61]:

1. Let the original signal contain N data points: $x(1), x(2), \dots, x(N)$.
2. length m vectors $\bar{X}(1), \dots, \bar{X}(N - m + 1)$ defined by:

$$\bar{X}(i) = [x(i), x(i+1), \dots, x(i+m-1)] \text{ for } 1 \leq i \leq N - m + 1.$$

Each vector represents m consecutive signal values, starting with the i^{th} data point, $x(i)$.

3. The distance between two vectors $\bar{X}(i)$ and $\bar{X}(j)$ is defined as the maximum absolute difference value between their respective scalar components,

$$d[\bar{X}(i), \bar{X}(j)] = \max_{k=0, \dots, m-1} |x(i+k) - x(j+k)|, \quad (3.1)$$

4. For a given $\bar{X}(i)$, find the value of j , ($j=1, \dots, N-m+1, j \neq i$) so that

$d[\bar{X}(i), \bar{X}(j)] \leq r$, denoted as $N^m(i)$. Then, for $i=1, \dots, N-m+1$:

$$C_i^m(r) = \frac{N^m(i)}{N-m+1},$$

where $C_i^m(r)$ measures the frequency of patterns similar to the one given by the window of length m within a tolerance r . In our case we used a threshold level of r equal to $0.2 * SD$, where SD is the standard deviation of the data sequence $x(\cdot)$.

5. Calculate the average natural logarithm of $C_i^m(r)$ by:

$$\psi_r^m = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r),$$

6. Repeat the method described above (steps 2 - 5) for $(m+1)$, until the final value of ApEn is given by:

$$ApEn(m, r, N) = \psi_r^m - \psi_r^{m+1}. \quad (3.2)$$

To compute ApEn values of a signal with length N , the embedding dimension m is set to 5, and a tolerance window r is set to 0.2 times the standard deviation of the original data sequence. Among the several different values for r and m , we tried, choosing $m=5$ and $r=0.2 * SD$ gave the best detection accuracy was achieved.

3.3.2.2 Feature 2: Maximum Singular Value (MSV)

The singular value decomposition (SVD) is a very powerful technique for recognizing and ordering the dimensions along which data points display the most variation. There are some fundamental ideas behind SVD [64]: (1) taking a high

dimensional, (2) highly variable set of data points and (3) reducing the data to a lower dimensional space. These can make characteristics of the original data more visible from noise. The SVD decomposes a given matrix A into three different matrices as follows:

$$A = U_{mm} \sum_{mn} V_{nn}^T, \quad (3.3)$$

where, U_{mm} and V_{nn} are orthogonal matrices, and $\sum_{mn} = \text{diag}(\sigma_i)$ is a diagonal matrix containing singular values σ_i in descending order of magnitude along its diagonal. The columns of the two matrices U_{mm} and V_{nn} are called the left and right singular vectors, respectively; they are mutually orthogonal. Larger singular values correspond to higher model correlation with the original data. For that reason, the SVD allows the extraction of dominant components from the recorded EEG data by keeping the largest singular values. The dimension of the SVD approximation can be reduced by setting to zero the singular values that are small [64]. In this work, the maximum of the singular values is taken as a feature extracted for seizure detection.

3.3.2.3 Feature 3: Median Absolute Deviation (MAD)

The Median Absolute Deviation (MAD) is a measure of statistical dispersion [65]. The absolute deviation of an element of a data set is the absolute difference between that element and a given point. Common measures of statistical dispersion are the variance and standard deviation. The MAD is considered as a simpler way to quantify variation than using variance or standard deviation.

For dataset $X = x(1), x(2), \dots, x(N)$, the MAD is defined as the median of the absolute deviations from the data's median [65]:

$$MAD = \text{median} (|X - \text{median}(X)|) \quad (3.4)$$

The MAD provides an absolute measure of dispersion that is not affected by outlier data that can throw off statistical analysis based on the mean and standard deviation.

MAD is more flexible with outliers existing in the data. Since large deviations are possible with seizure signals, a median based measure of dispersion will be a good choice for detection [66].

Figure 3.3 The values of the three features used for seizure detection (MAD in (a), ApEn in (b), and MSV in (c)) using Baseline and Seizure segments of rat data shown in the top left and right figures, respectively. The oscillation of Seizure data in comparison to the Baseline data explains the high values of MAD and Max singular value. The values of ApEn for Seizure data are low compared to Baseline data.

3.3.3 Feature vector classification

Figures 3.3 and 3.4 show an example of the extracted features using rat and human data, respectively. Two different classifiers were used to classify the feature vectors to distinguish between seizure and non-seizure activity. The classifiers used are: the Support-Vector Machine (SVM) [68] and the AdaBoost [71].

Figure 3.4 Top two figures are examples of EEG segments from non-seizure (Z) and seizure (S) sets, respectively of human data [2]. Sub-figures (a) – (c) represents the values of the extracted features (ApEn, MSV, and MAD). It is clear that the values of MAD and MSV are higher for Seizure data than for Normal data, but the ApEn for Normal data is higher than those for Seizure.

3.3.3.1 Classifier 1: Support Vector Machine (SVM)

The basics of the SVM algorithm have been developed by Vapnik [68] to analyze data and recognize patterns. The Support Vector Machine has been used for classification and regression analysis [67]. In this dissertation, it is applied for the

classification of extracted features. The classification is done by constructing an N-dimensional hyper-plane that optimally separates the data into two categories by maximizing the distance from the decision boundary to the nearest data-points (called support vectors) in the training data. The MATLAB implementation of the SVM classifier is used for this study (bioinformatics toolbox statistical learning commands (svmtrain, and svmclassify)). Since Seizure and non-Seizure classes are often not linearly separable, we generate non-linear decision boundaries using the Gaussian Radial Basis Function kernel (RBF) [67]. Furthermore the classification results obtained by our algorithm give better results using the RBF kernel function when compared to other SVM kernels. Figure 3.5-A shows the general diagram of SVM and Figure 3.5-B shows an example of classification using the SVM with an RBF kernel.

Figure 3.5 (A) General diagram of a Support Vector Machine; image obtained from [69] (B) An example of classification using the SVM with RBF kernel function; image obtained from [110].

3.3.3.2 Classifier 2: Adaptive Boosting (AdaBoost)

AdaBoost is one of the most popular machine learning algorithms introduced in 1995 by Freund and Schapire [71]; it has surprising good classification results. In

theory, boosting can be used to reduce the error of any weak-learning algorithm that constantly creates classifiers which need to be better than random guessing. Similar to SVM, the AdaBoost algorithm works by combining several votes by using weak learners instead of using support vectors. The AdaBoost.M1 algorithm described in Figure 3.6 was adopted for our work [71]. The algorithm takes as input the training set: $D = \{(f_i, y_i)\}, i = 1, \dots, N$ with $f_i \in F$ (instance space) and each label $y_i \in Y = \{-1, +1\}$ (label set).

In the beginning, the EEG data have to be labeled for training, with the Baseline segments labeled as “-1” and Seizure segments labeled as “+1”. The algorithm repeatedly calls a given weak or learning algorithm in a series of Rounds $t = 1, \dots, T$ and produces a distribution or a set of weights over the training set. The weight of this distribution on training data i on round t is denoted $D_t(i)$. In this work, the decision stumps were used as weak classifiers. A decision stump is a decision tree with one root node and two leaf nodes [71]. It performs a single test on a single attribute with threshold θ . A decision stump can also be used to reduce the error of any Weak-Learning algorithm that constantly creates classifiers which need to be better than random guessing; it is constructed for each feature in the input data as:

$$h(f_i) = \begin{cases} +1 & \text{if } f_i > \theta \\ -1 & \text{else} \end{cases}.$$

For $i = 1, \dots, N$ and $t = 1, \dots, T$, the weak learner’s goal is to find a hypothesis

$$h_t : F \rightarrow Y \text{ which minimizes the training error: } \varepsilon_t = \sum_{i: h_t(f_i) \neq y_i} D_t(i).$$

such that pr is the probability and ε_t is the sum of distribution weights of the

instances misclassified by the hypothesis h_i , (step 3 in the algorithm). And we require that this error be less than $\frac{1}{2}$. This error is measured with respect to the weight of the distribution denoted by $D_i(i)$. The data is re-weighted to increase the “importance” of misclassified samples (rule 5 shown in Figure 3.6). This process continues and at each step the weight of each learner is determined. Finally, the strong classifier is defined as the output of this algorithm.

3.3.4 Performance evaluation

In order to evaluate the performance of the proposed algorithm for seizure detection, the first two-thirds of the feature vectors (for both seizure and non-seizure data) were selected for training while the last one-third was used for testing. The results of classification and the performance of classifiers are expressed in terms of sensitivity, specificity, and accuracy (Table 3.2) which are defined as follows [72]:

- **Sensitivity**: Is the percentage of seizure segments correctly classified by the algorithm.
- **Specificity**: Is the proportion of segments without seizures correctly classified by the algorithm.
- **Accuracy**: Is the percentage of correctly classified segments relative to the total number of segments considering for classification.

$$\text{Sensitivity (Sens)} = \frac{TP}{TP + FN}, \quad (3.5a) \quad \text{Specificity (Spec)} = \frac{TN}{TN + FP}, \quad (3.5b)$$

$$\text{Accuracy (Acc)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.5c)$$

Input: N examples $((f_1, y_1), \dots, (f_N, y_N))$ with $f_i \in F$ and labels $y_i \in Y = \{-1, +1\}$

Weak learner algorithm **WeakLearner**

Integer T specifying number of iterations

Initialize $D_1(i) = 1/N$ for all i

Do for $t = 1, 2, \dots, T$

1. Call Weaklearner,
2. Get a hypothesis $h_t : F \rightarrow Y$,
3. Calculate the error of h_t : $\varepsilon_t = \sum_{i: h_t(f_i) \neq y_i} D_t(i)$. If $\varepsilon_t > 0.5$, then set $T=t-1$ and stop loop.
4. Set $\beta_t = \varepsilon_t / (1 - \varepsilon_t)$,
5. update D_t : $D_{t+1}(i) = \frac{1}{Z_t} (D_t(i) \begin{cases} \beta_t, & h_t(f_i) = y_i \\ 1, & \text{otherwise} \end{cases})$

where $Z_t = \sum_i D_t(i)$ is a normalization constant (chosen so that D_{t+1} will be a proper distribution function).

Output the final hypothesis: $h_{fin}(f) = \arg \max_{y \in Y} \sum_{t: h_t(f)=y} \log \frac{1}{\beta_t}$.

Figure 3.6 The AdaBoost classifier Algorithm [71].

where TP, TN, FN, and FP represent the numbers of true positive, true negative, false negative, and false positive results, respectively. Sensitivity in (3.5a) is the percentage of seizure events correctly identified by the test whereas specificity in (3.5b) is the percentage of non-seizure events correctly identified by the test.

3.3.5 Results and Discussion

The proposed features extraction method is evaluated using two classifiers; SVM and AdaBoost. Each tEEG and EEG segment is divided into epochs of a predetermined length of one second duration. Three statistical features are extracted for each epoch which are: MAD, ApEn, and MSV. Figures 3.3 and Figure 3.4 show the variations in the values of the extracted features during seizure and non-seizure for tEEG and EEG signals, respectively. From both figures, we can observe that the values of MAD and MSV during seizure are larger compared to that of the non-seizure segment. Furthermore, both figures show that the seizure data have smaller ApEn values compared to the non-seizure data due to the data's rhythmicity.

In order to reduce the dimensionality and computational complexity of the proposal algorithm, the following statistical features were evaluated for each one second epoch of extracted feature vectors: mean (Mean), minimum (Min), maximum (Max), and standard deviation (Std). This reduction decreases the classification complexity without losing high performance. Table 3.1 presents the values of Max, Mean, Min and Std of extracted features described above, for Seizure and non-seizure data from Figure 3.3. These values show that there are definite differences between the Seizure and non-seizure tEEG signals. Analysis of Table 3.1 illustrates that the ApEn, the Mean, Min and Std values during seizure activities had lower values compared to Baseline data. Also it is obvious that Seizure data have higher values of MAD and MSV compared to those of non-seizure data. Furthermore, the values are significantly different among the two classes of Seizure and non-seizure or Baseline tEEG signals ($p = 0.038$ for ApEn, and $p < 0.01$ for MAD and MSV). These results suggest that

the tEEG segments with seizure describe a more complex behavior compared to the segments when there are no-seizure events. Thus, these features should be suitable to characterize the tEEG signals.

Table 3.1 Statistical values of extracted features from tEEG human data with and without seizure of Figure 3.3.

Features	Non-seizure data				Seizure data			
	Max	Mean	Min	Std	Max	Mean	Min	Std
Approximate entropy (ApEn)	1.5044	1.2945	0.9708	0.1252	1.6378	1.1846	0.6341	0.2530
Maximum Singular Value (MSV)	0.0289	0.0206	0.0125	0.0036	2.7634	0.5409	0.1109	0.4938
Median Absolute Deviation (MAD)	0.0025	0.0016	0.0011	3.3e-4	0.2740	0.0508	0.0101	0.0491

Two different classifiers were used to classify the extracted feature vectors into seizure and non-seizure segments. The classifiers are SVM and AdaBoost. The result of classification using the rat and human datasets are shown in Table 3.2. The linear classification using the AdaBoost algorithm classified tEEG seizure and non-seizure segments with a very promising accuracy of 84.81%. Comparatively, the non-linear classification using SVM algorithm (with RBF kernel) had an accuracy of 96.51%. Even higher classification accuracies were achieved using the human dataset [58] with 98.44% accuracy via AdaBoost algorithm and 100% using the SVM classifier. This performance is expected since the human data are considered relatively artifact free, whereas our tEEG data recordings using TCRES contain noise and artifacts because the signals are recorded externally from the scalp and the rats are roaming, grooming,

chewing, running, etc. The human dataset [2] was used without any preprocessing or downsampling of the data.

Table 3.2 A comparison of classification accuracy obtained using tEEG data, rat's data, and EEG data from human.

	Sensitivity (%)	Specificity (%)	Accuracy (%)
tEEG data from rats			
AdaBoost	90.91	78.72	84.81
SVM	93.02	100	96.51
Average	91.96	89.36	90.66
EEG data from humans [2]			
AdaBoost	100	96.88	98.44
SVM	100	100	100
Average	100	98.44	99.22

Among the published works for seizure detection, there are a number of methods that deal with detection of epileptic seizure using the *same* on-line EEG database described by Andrzejak et al. [2]. For example Guo et al. [20] applied ApEn derived from a multiwavelet transform to classify EEG signals. Their results were very promising with an overall accuracy from 98.27 to 99.85%. Fathima et al. [73] used Wavelet based features for the classification between non-seizure and seizure EEG signals via a linear classifier with an accuracy of 99.5%. Bedeuzzaman et al. [66] have proposed a time domain analysis of the EEG data set and their classification was based on a linear discriminate function with an accuracy of 100%. Comparing these techniques [20, 66 and 73] with our proposed method, the results obtained from the

evaluation of our method using SVM classifier are better than those obtained by Fathima et al. [73] and Guo et al. [20], and are as good as the results obtained by Bedeuzzaman et al. [66].

The major limitation of our study was that we could not use all the data recorded from rats using the TCRE. We had to pick a small segment of baseline or non-seizure data, 30 seconds duration, which was the least contaminated with artifacts. The artifacts and noise were caused by the rats roaming, grooming, and other behaviors during the recording. Even in the presence of strong artifacts, our methods worked quite well suggesting the algorithm is robust.

3.4 Seizure detection using the Rényi entropy

In our second algorithm proposed for human seizure detection, the Rényi entropy [74] was combined with two different methods: the Empirical Mode Decomposition method (EMD) [54] and the discrete wavelet transform (DWT) [48]. A block diagram describing the different steps of the algorithm is shown in Figure 3.7. Initially, the EEG signal was decomposed into sub-signals using the EMD method or the DWT. Then, the quadratic Rényi entropy was used as a feature. The Rényi entropy was extracted from the first five IMFs when using EMD decomposition or from each sub-signal when using the DWT. Finally, the extracted features are input to a k-nearest neighbor (KNN) classifier [27] to separate the EEG segments into either seizure or non-seizure classes. In this study, we will show that high accuracies are obtained using both techniques (EMD and DWT), which indicate good performance of the proposed method.

Figure 3.7 An overview of the proposed methodology for human epileptic seizure classification.

3.4.1 Feature extraction

Entropy can be used as a complexity measure for analyzing biological signals such as brain signals or EEG, because the brain is a complicated dynamical system with uncertainties [74].

3.4.1.1 Empirical Mode Decomposition (EMD)

EMD has been proposed for the analysis of nonlinear and non-stationary time series [54]. The EMD is a signal processing technique that adaptively decomposes a signal into oscillating components or Intrinsic Mode Functions (IMFs). Each IMF satisfies two basic conditions [54]: (i) the number of extrema and the number of zero crossings must be equal or differ at most by one; (ii) the mean value of the envelope defined by the local maxima and by the local minima must be zero or close to zero at all points. The spectral content of the IMF extracted at a given iteration is lower than that of the IMF extracted from the previous iteration, which permits analyzing the

signal in different frequency bands. The EMD technique has been used in the field of biomedical signal processing, especially for seizure detection in EEG signals [75, 76].

The EMD has several advantages: (1) The EMD algorithm is a decomposition method developed for non-linear and non-stationary signals [54], making it suitable in biomedical engineering applications such as detection of epileptic seizure from EEG signals [75, 76]; (2) EMD can break down complicated signals, without the need of basis functions, into a finite set of band limited IMFs which can be described in the time-frequency domain; and (3) The EMD algorithm is considered as a type of filter bank decomposition method [78] used to isolate different components from multi-component signals like EEG. Moreover, The EMD procedure allows time-frequency analysis of transient signals, which is not the case for stationary Fourier Transform based methods [77]. The EMD algorithm for a given signal $x(t)$ can be summarized in the following steps [78]:

(1) Initialize sifting loop: $j = 1$, $r_0(t) = x(t)$

(2) Set $f_j(t) = r_{j-1}(t)$,

(3) Detect the maxima and minima of $f_j(t)$. Compute the upper and lower envelopes $e_m(t)$ and $e_l(t)$, respectively, using interpolation between the maxima and minima of $f_j(t)$,

(4) Extract the local mean: $m(t) = \frac{e_m(t) + e_l(t)}{2}$,

(5) The IMF should have zero local mean; so subtract the local mean from the original signal: $f_j(t) \leftarrow f_j(t) - m(t)$,

(6) Check the two conditions for an IMF described above to decide whether $f_j(t)$ is an IMF or not, (conditions (i) and (ii)),

(7) $j = j + 1$, if $j \leq N$ go to step (2),

When the first IMF is defined, the smallest temporal scale in the signal $x(t)$ is defined as: $g_1(t) = f_1(t)$.

In order to obtain all the other IMF components, the residue $r_1(t)$ of the data is obtained by subtracting $g_1(t)$ from the signal as $r_1(t) = r_0(t) - g_1(t)$. The sifting process will be continued until the final residue is a constant, monotonic function, or a function with only one extrema from which no more IMFs can be derived. The subsequent basis function and the residues are computed as:

$$r_2(t) = r_1(t) - g_2(t) \quad , \dots \dots \dots , \quad r_N(t) = r_{N-1}(t) - g_N(t)$$

where N is the number of IMFs and $r_N(t)$ is the final residue. At the end of the decomposition, the signal $x(t)$ is given by:

$$x(t) = \sum_{j=1}^N f_j(t) + r_N(t) . \quad (3.6)$$

The adaptive nature of EMD can often provide a better numerical description of temporal patterns in non-stationary and nonlinear time series than traditional methods such as Wavelet Transform and Fourier Transform methods [54]. The EMD decomposition method is well localized in the time-frequency domain and reveals important characteristics of the signal [54]. The MATLAB code used in this paper to perform EMD is available at [88]. All calculations were implemented using MATLAB software version 7.10. In order to have different rhythms (or frequency components)

in the EEG signal, the first five IMFs of each EEG segment were selected and used for further analysis. Figure 3.8 shows the first five IMFs, using the human dataset [2], for non-seizure segments (Z) on the left and seizure segment (S) on the right.

Figure 3.8 The Empirical mode decomposition (EMD) of non-seizure data (set Z) on left and with seizure (set S) on right. IMF1-IMF5 are the first five IMFs after EMD decomposition of each segment from the sets Z and S. The frequency support of these IMFs decreases from top to bottom. Furthermore, the lower order IMFs have the fast oscillation mode of the signal.

3.4.1.2 Discrete Wavelet Transform (DWT)

As we already mentioned in Chapter 2, The Continuous WT (CWT) of a signal $x(t)$ is defined as the affine correlation between $x(t)$ and a scaled and shifted wavelet function $\psi(t)$ as follows [48].

$$CWT_x(a,b) = |a|^{-1/2} \int_{-\infty}^{\infty} x(t) \Psi^* \left(\frac{t-b}{a} \right) dt \quad (3.7)$$

The Discrete Wavelet Transform (DWT) is one of the methods used to examine biomedical signals. It provides both time and frequency information of the signal [79]. The DWT is advantageous over the continuous WT since it allows the complete generation of the original signal without redundancy [79]. The DWT uses discretized parameters, “ a ” and “ b ”, respectively [79]:

$$a = a_0^m, \quad b = n.b_0.a \quad \text{with } m, n \in Z \quad (Z \text{ is a set of all integers}).$$

To obtain an orthogonal basis, one can chose samples on a dyadic grid (base 2):

$$a = 2^m \text{ and } b = n.2^m.$$

Table 3.3 Frequency bands of an EEG signal with four levels of DWT decomposition assuming sampling frequency of 173.6 Hz. $D1 - D4$ are the detail coefficients and $A4$ is the approximation coefficient.

Signal	Frequency bands (Hz)	Decomposition level
D1	43.4-86.8	1
D2	21.7-43.4	2
D3	10.8-21.7	3
D4	5.4-10.8	4
A4	0-5.4	4

Tests were conducted using different types of wavelets, e.g. Coiflet wavelet, Haar wavelets and Symlet wavelet. The discrete Daubechies wavelet of order 2 (db2) was chosen since it gave the best results using our method compared to the other types of wavelet we tested. Selection of the appropriate wavelet and the number of decomposition levels is very important for the analysis of any signal using the DWT [80]. The number of decomposition levels was chosen based on the dominant frequency components of the signal. The frequency bands related to a 4-level DWT decomposition with sampling frequency $f_s = 173.6\text{Hz}$ using the human EEG dataset [2] is shown in Table 3.3, where D1–D4 are the detailed coefficients and A4 is the

Figure 3.9 Discrete Wavelet Transform Approximate and detail coefficients of non-seizure EEG (set Z) on left and epileptic Seizure EEG (set S) on right. The first row contains the original data. Row 2 contains the approximate coefficient A4. Rows 3, 4, 5 and 6 contain the detail coefficients D4, D3, D2 and D1, respectively. Details coefficients D1-D4 have low-scale and high-frequencies, whereas the approximate coefficient A4 has high-scale and low-frequency.

approximate coefficient. It is clear from Table 3.3 that each EEG signal is decomposed into filtered sub-components corresponding to different frequency bands. Figure 3.9 shows the sub-signals after DWT decomposition for both seizure and non-seizure segments, where each sub-signal represents different frequency sub-bands.

3.4.1.3 Rényi Entropy

Rényi entropy [74] based measures have recently been applied to different areas of signal processing and information theory. It has been proven that the Rényi entropy is a very robust measure [81]. The Rényi α -entropy is a generalization of the Shannon entropy and is defined as follows [74]:

$$H_{\alpha}(P) = - \frac{1}{1-\alpha} \log_2 \left(\sum_{i=1}^N p_i^{\alpha} \right) \quad (3.8)$$

where $P = (p_1, \dots, p_N)$ is a finite probability distribution and α is a non-negative real number with $\alpha \geq 0$ and $\alpha \neq 1$. The fundamental properties of Rényi entropy are: (1) the entropy $H_{\alpha}(P)$ tends to the Shannon entropy as the order α tends to one; (2) $H_{\alpha}(P)$ is a non increasing function of α , so $\alpha_1 < \alpha_2 \Rightarrow H_{\alpha_1}(P) \geq H_{\alpha_2}(P)$.

The advantage of Rényi entropy over the Shannon entropy lies in its generality and flexibility due to the parameter α enabling several measurement of uncertainty within a given distribution. Moreover, in contrast to Shannon entropy, there exist efficient methods for computing Rényi entropy for some values of parameter α [83]. In this paper, the Rényi entropy with $\alpha = 2$, called the ***Quadratic Rényi Entropy***, was used. It gave the best classification accuracy of the data analyzed among several values of α ($\alpha = 2, 3, 4, 5$) that were examined.

3.4.2 Classification and Performance Calculation

3.4.2.1 Ten-fold cross-validation for training and testing

There are various methods to split a given dataset into training and testing for EEG Seizure detection. One of the most known methods is a ten-fold cross-validation [82]. A ten-fold cross-validation technique will be applied throughout the training data and used to estimate how well the classification method will classify the new data during the testing period. In ten-fold cross validation, the data set is split into 10 equal sub-set partitions. In each iteration, one of the 10 subsets is used for testing whereas the other 9 subsets are used for training the dataset. The whole procedure is then repeated ten times using a different testing subset. The final result is the average of all the 10 iterations. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once.

Table 3.4 The quadratic Rényi entropy values of Intrinsic Mode Functions (IMF1-IMF5) after EMD decomposition on human dataset.

Subject	Quadratic Rényi Entropy					Mean
	IMF1	IMF2	IMF3	IMF4	IMF5	
Non-seizure (set Z)	-13.4849	-14.4962	-14.2129	-13.3737	-13.6569	-13.8449
	-19.9743	-19.8003	-19.7017	-18.5954	-17.6477	-19.1439
Seizure (set S)						

Table 3.5 The quadratic Rényi entropy values of signal sub-components D1-D4, and A4 after DWT on human dataset.

Subject	Quadratic Rényi Entropy					Mean
	D1	D2	D3	D4	A4	
Non-seizure (set Z)	-15.2143	-10.2599	-12.6259	-14.1820	-14.5023	-13.3569
Seizure (set S)	-19.8281	-14.4539	-17.7013	-19.5416	-19.0559	-18.1162

3.4.2.2 Classification using k-nearest neighbor (k-NN)

The k-NN classifier was used to classify the seizure from non-seizure segments. k-NN is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure [84]. The k-NN approach has recently been recognized as a very important algorithm, due to its high classification accuracy in problems with unknown and abnormal distributions [84]. The k-NN classification finds a group of k objects in the training set that are closest to the test object, and assigns a label on the predominance of a particular class in this neighborhood [84]. Using the k-NN classifier, three important components must be identified: a set of labeled objects, a distance or similarity metric to compute the distance between objects, and the value of k, the number of nearest neighbors. In this chapter, the number of nearest neighbor k was set to be one, because it achieved the optimal detection accuracy of the different values of k (1, 2, 3, 4 and 5) that were examined.

3.4.3 Performance Evaluation

The performance of the proposed method for seizure detection was evaluated using the human dataset [2]. The same algorithm was tested using the approximate entropy (ApEn). The performance results are summarized in Table 3.6. Comparison of the obtained results, using the quadratic Rényi entropy combined with EMD or DWT, with others in the literature for the process of seizure detection using the same dataset was shown in Table 3.7.

Table 3.6 A comparison of classification performance achieved by EMD and DWT using the k-NN classifier on human dataset.

Classification Method	Sensitivity (%)	Specificity (%)	Accuracy (%)
EMD + Approximate Entropy (ApEn)	83.33	90	86.36
DWT + Approximate Entropy (ApEn)	90.80	89.80	90.30
EMD + Quadratic Rényi Entropy	100	100	100
DWT + Quadratic Rényi Entropy	100	99.90	99.95

3.4.4 Results and Discussion

The second proposed algorithm for seizure detection was based on the extraction of the quadratic Rényi entropy as a feature to discriminate between non-seizure and seizure segments after decomposing the signal into sub-signals using two different techniques: EMD which decomposes the signal into intrinsic mode functions (IMFs) and DWT which enable a time-frequency view of the signal. To evaluate the proposed method, two sets of human EEG data (available online in [2]) were used; one set having seizure (set S) events and the other set contain non-seizure data (set Z).

After EMD decomposition, the number of bands or IMFs is dependent on the frequency content of the signal to be analyzed (IMF_1 - IMF_N shown in Figure 3.8, where N is the number of the last IMF which varies from signal to signal depending on the frequency content). Figure 3.8 shows that the frequency content of these IMFs was given from high to low. Furthermore, the lower orders of these IMFs have fast oscillation modes of the signal, whereas the higher order IMFs have the slower oscillation modes for both seizure and non-seizure segments. The first five IMFs were used since they contained significant information with different rhythms.

For the WT, the number of decomposition levels was chosen based on the dominant frequency components of the signal. Similar selection orders as for the EMD was made for DWT. The seizure and non-seizure segments were decomposed into five sub-signals (D1-D4 and A4).

The quadratic Rényi entropy values were calculated for each sub-signal after EMD and DWT decomposition. In order to reduce the dimensionality of the detection algorithm, the mean value of the five quadratic Rényi entropies (for the five IMFs and the five wavelet sub-signals) were computed. Table 3.4 and Table 3.5 list the values of the quadratic Rényi entropy for IMFs and DWT sub-signals. It was found that the values of the quadratic Rényi entropy are higher in Normal segments compared to Seizure segments for both IMFs and Wavelet sub-signals.

The k-nearest neighbor (k-NN) classifier was used to classify the extracted features and calculate the performance of the proposed method. One important concern using k-NN classifier is that the parameter k (number of nearest neighbors) is a variable. In this chapter, good accuracy was achieved with k equal to one. The same

algorithm was tested using the approximate entropy (ApEn). The computed accuracies are presented in Table 3.6. These results demonstrate the following; using the Rényi entropy the total classification accuracy reaches 100% using the EMD method and 99.95% using the DWT. These accuracy performances show that the EMD method using the Rényi entropy has some improved results. Thus, the classification accuracy using the EMD method performed slightly better. Comparing the results obtained using the quadratic Rényi entropy to those using the ApEn, with an accuracy of 86.36% using the EMD and 90.30 using the DWT, it is clear that the algorithm shows the superior performance of the Rényi entropy compared to the ApEn in detecting epileptic seizures. There are many other methods proposed for epileptic seizure detection in the literature. Table 3.7 compares our results with the results of other methods proposed in the literature using the same human dataset in [2]. It is shown in Table 3.7 that the results obtained in this study using quadratic Rényi entropy combined with dyadic wavelet transform indicate an improvement over the other approaches using dyadic WT algorithm for seizure detection [20, 80 and 85]. The EMD algorithm reaches an overall accuracy of 100%, which is the best presented method using this dataset when compared to the other algorithms shown in Table 3.7.

Table 3.7: A comparison of classification accuracy obtained by our methods versus methods of others' using the same human dataset described by Andrzejak et al. [2].

Author (s)	Ref	Method	Classifier	Human Dataset [2]	Accuracy (%)
Fathima et al.	73	Wavelet based features	Linear Classifier	Z, S	99.5
Guo et al.	20	Wavelet transform & approximate entropy	Artificial neural network (ANN)	Z, S	99.85
Nigam & Graupe	28	Nonlinear pre-processing filter	Diagnostic neural network	Z, S	97.2
N. Kannathal et al.	85	Entropy measures	Adaptive Neuro-fuzzy inference system	Z, S	99.22
E. D. Übeyli	87	Burg autoregressive & Moving average & least squares modified Yule-Walker autoregressive	Support vector machines	Z, S	99.56
L. Guo et al.	80	Wavelet transform & Line length	Artificial neural network	Z, S	99.6
Feltane et al.	14	ApEn, MAD, MSV	AdaBoost, SVM	Z, S	99.22
Our work		DWT & Renyi entropy EMD & Renyi entropy	K-nearest neighbor classifier	Z, S	99.95 100

3.5 Summary

This chapter presents two new approaches for automatic seizure detection using two different datasets. The first algorithm is based on the extraction of time-domain features which are the Approximate Entropy (ApEn), the Median Absolute Deviation (MAD) and the Maximum Singular Value (MSV). A linear classifier, Adaptive Boosting (AdaBoost), and a non-linear classifier Support Vector Machine (SVM) were used to classify the feature vectors into seizure and non-seizure segments. The performance of the algorithm was evaluated using two different datasets, one using rat and the other using human data. The obtained results showed that the proposed algorithm for detecting seizure segments was robust; it achieved high accuracy from the tEEG rat dataset even though the data were contaminated with artifacts such as roaming, grooming, etc.

In the second algorithm, we demonstrated the capability of applying quadratic Rényi entropy for the purpose of EEG seizure detection as compared to other entropies. Because of the non-stationarity of the EEG signal, Empirical Mode Decomposition method (EMD) and Discrete Wavelet Transform (DWT) were used to decompose the signal into oscillations and trend components. The human dataset in [58] was used to evaluate the performance of the algorithm using the k-nearest neighbor (k-NN) classifier. The results showed that the performance of EMD is slightly better than DWT, which may be because the signals were relatively noise free. Our results suggest that the approach of using quadratic Rényi entropy combined with EMD and DWT are new methods for detecting human seizures. According to the high accuracy obtained, quadratic Rényi entropy combined with EMD or DWT is effective and applicable for discriminating Seizures from Normal EEG recordings.

CHAPTER 4

SEIZURE DETECTION USING TIME-FREQUENCY ENERGY CONCENTRATION

4.1 Introduction

It was shown in the literature that seizure evolution is, in general, a dynamic and non-stationary process [107, 109]. Moreover, the signal during seizure is composed of multiple frequency components. In this chapter, we demonstrate the suitability of using the estimated energy from a time-frequency (TF) representation to classify EEG seizure versus non-seizure segments; we also compare our approach to other methods. Frequency domain analysis of EEG signals shows that EEG signal activity ranges from almost DC to 100 Hz. Since seizure activity may exist in different frequency ranges, the Empirical Mode Decomposition method (EMD) was used to decompose the signal into different components which helps to detect the abnormal energy that appears in the signal during seizure. The algorithm exhibited a high percentage of classification. To test the proposed algorithm, data from rats with induced seizures were recorded using the tripolar concentric ring electrodes (TCREs).

4.2 Data Description

In order to evaluate the performance of the proposed algorithm for seizure detection, the rats' dataset shown in chapter 3 was used. The data were recorded using

the tripolar concentric ring electrodes (TCRE). Two sets of the Laplacian EEG signal (tEEG) data from ten rats were used. The first set is composed of 65 single-channel tEEG non-seizure segments and the second set is composed of 70 single-channel tEEG Seizure segments.

4.3 Model-based seizure detection

Different methods for automatic seizure detection have been discussed in the literature [20, 28, 66, 73, 85, and 87]. Our proposed algorithm is based upon the decomposition of tEEG segments into several oscillating components via the EMD algorithm [54] followed by TFR (time-frequency representation). A localized energy estimate is extracted and considered as a feature for discrimination between seizure and non-seizure data. An overview of the different steps of the detection process is summarized in Figure 4.1, with a more detailed explanation to follow:

1. Downsample the tEEG signal to reduce the sampling rate from 256 Hz to 128 Hz to reduce the amount of data, and hence, computations without losing important information for analysis. The signal was filtered first using an anti-aliasing low-pass filter with a cutoff frequency 64 Hz to meet the Nyquist criteria and avoid aliasing. The MATLAB function *downsample* was used for the down-sampling procedure.
2. Decompose each tEEG segment into IMFs (Intrinsic Mode Functions) using the EMD algorithm [54]. Figure 4.2 shows the IMF bands for both seizure and non-seizure data. The IMF noise distribution algorithm proposed by Flandrin et al. [89] was used to select three IMFs for seizure detection.

3. The three selected IMFs from each seizure and non-seizure segment were partitioned into one-second epochs using a non-overlapping, sliding Hamming window to avoid redundancy caused by an overlap.

4. The analytic signal for each epoch was used [90] to eliminate the negative frequencies for better cross term reduction in TFRs. For a given signal $x(t)$, the corresponding analytic signal is defined as: $y_x(t) = x(t) + jHT(x(t))$, where $HT()$ is the Hilbert transform.

Figure 4.1 Block diagram of the proposed method for automatic seizure classification.

5. The Smoothed-Pseudo Wigner-Ville distribution (SPWVD) was computed for each analytic epoch [50].

6. An estimate of the localized time-frequency (TF) energy was extracted for each epoch, used as a feature, and classified using a linear classifier to determine whether or not a given epoch contains a seizure or not.

4.4 Feature extraction procedure

The procedure for feature extraction was delineated into three steps. First, the data segments were decomposed into the oscillatory components, known as IMFs using the EMD method, and three were adaptively selected. Then, the SPWV distribution was computed for each epoch for selected IMFs (to be described later). Finally, an estimate of the localized TF energy was calculated and used as a feature to discriminate between seizure and non-seizure events. The method is briefly described in the following sections.

4.4.1 Signal decomposition and IMFs selection

The tEEG seizure and non-seizure segments were iteratively decomposed into IMFs using the EMD algorithm [54]. Figure 4.2-(Aa and Ba) shows an example of a 30-second seizure and non-seizure tEEG segments, respectively. For each segment, 10-12 IMF components were computed using the EMD decomposition method; they are shown in Figure 4.2-Ab and Bb.

To identify which IMFs to use in the proposed analysis, we need to know whether a specific IMF contains useful information or primarily noise. Flandrin et al. [89] developed a statistical model based on energy distribution of the noise between IMFs. The method suggests decomposing the noisy signal into IMFs, and then comparing the IMF energies with the theoretical estimated noise-only IMF energies. The model is based on studying the energy in the modes of fractional Gaussian noise (fGn) after EMD decomposition. fGn is a generalization of white noise; it exhibits a flat spectrum and its statistical properties are determined solely by a scalar parameter H known as the Hurst exponent. In this chapter, we take $H = 0.5$, so the process is reduced to one

of uncorrelated white noise. We consider that the EMD of a discrete-time signal $x[n]$ for $N = 1, \dots, M$ results in a set of k IMFs $f_k[n]$ for $k = 1, \dots, N$. The desired signal $x[n]$ is assumed to be corrupted with white noise (fGn with $H = 0.5$). The energy of the first IMF is:

$$W_H[1] = \sum_{n=1}^M f_1^2[n]. \quad (4.1)$$

The energy of the noise in the other IMFs for a given Hurst exponent $H = 0.5$ is:

$$W_H[k] = \frac{W_H[1]}{0.719} 2.01^{-k}, \quad k = 2, 3, \dots, N. \quad (4.2)$$

Moreover, there is a linear relationship between the logarithm of the confidence interval $T_H[k]$ and the index of the IMF, k , given by:

$$\log_2(\log_2(T_H[k]/W_H[k])) = a_H k + b_H. \quad (4.3)$$

For a confidence interval of 99%: $a_H = 0.46$ and $b_H = -1.92$. The algorithm proposed by Flandrin et al. [89] is the following:

- (I) Assuming that the first IMF captures most of the noise, estimate the noise level in the noisy signal by computing $W_H[1]$ from equation (4.1).
- (II) Estimate the “noise only” model by using equation (4.2).
- (III) Estimate the corresponding model for a chosen confidence interval from equation (4.3).
- (IV) Compute the EMD of the noisy signal, and compare the IMF energies by using the confidence interval as a threshold.
- (V) Compute a partial reconstruction by keeping only the residual and those IMFs whose energy exceeds the threshold (confidence interval).

Figure 4.2 (*Aa*) and (*Ba*) depict a 30-second segment of tEEG recording from rats during non-seizure and seizure period, respectively. The Empirical Mode Decomposition (EMD) of the non-seizure and seizure segments is shown in (*Ab*) and (*Bb*), respectively, with the IMF1 on the top row and the last IMF in the bottom row. The EMD computed 12 IMFs for the non-seizure data and 10 for the seizure data.

In this paper, the Flandrin algorithm for IMF selection [89] was run on each 30-second segment of the signal to identify which three IMFs should be selected for the next algorithmic step of TFR energy feature classification.

In this study, the Flandrin algorithm determined that IMF3 was the first IMF to cross the threshold for the vast majority of the data set used. Figure 4.3 shows an example using a 30-second Seizure segment. The figure shows that the IMF energy increases significantly at IMF3. Consequently, IMF numbers 3, 4, and 5 were used in subsequent analysis to best represent the dataset used in this study and to reduce algorithm computational complexity.

Figure 4.3 An example of IMF selection criteria using equations (4.2) and (4.3). The noise-only model in red and the confidence interval in blue are presented. The black curve represents the energy of each IMF. The IMFs numbers 3 to 10 have energies which exceed the confidence interval (or threshold).

4.4.2 Time-Frequency energy estimation

TF energy distributions are very important in the analysis and processing of non-stationary signals like the EEG [106, 107, 109]. An energy TFR $T_x(t, f)$ combines the concepts of a signal's instantaneous power $p_x(t) = |x(t)|^2$ and spectral energy density $p_x(f) = |X(f)|^2$, where $X(f)$ is the Fourier transform of the signal $x(t)$. The energy of TFR satisfies the following marginal properties [40]:

$$\int T_x(t, f) df = p_x(t) = |x(t)|^2 \quad (4.4)$$

$$\int T_x(t, f) dt = p_x(f) = |X(f)|^2 \quad (4.5)$$

Equations (4.4) and (4.5) indicate that if the TF energy density is integrated along one variable, the energy density corresponding to the other variable can be obtained.

$$E_x = \iint T_x(t, f) dt df = \int |x(t)|^2 dt = \int |X(f)|^2 df \quad (4.6)$$

The total signal energy condition E_x in equation (4.6) is derived by integrating the TFR $T_x(t, f)$ over the entire TF plane. Many TFRs, including the SPWV distribution, do not strictly obey the marginal properties in (4.4 and 4.5); that is, the frequency and time integrals of the distribution do not exactly equal the instantaneous signal power and the spectral energy density, respectively [40]. However, depending upon window choice, some TFRs can still be used to generate estimates of localized signal energy [38, 40]. The total energy can be a good feature to detect signal events in the TFR because the energy during an EEG seizure event is usually larger than that during normal activity [22]. In this chapter, the approximate localized energy extracted from TFRs was calculated for each epoch of the three selected IMFs components, IMF3,

IMF4, and IMF5, to construct the feature vector used for automatic seizure detection. The Smoothed Pseudo Wigner-Ville distribution (SPWVD) was used as the TFR to extract the energy from the selected IMFs. Figure 4.4 and Figure 4.5 use rat data from non-seizure and seizure segments, respectively. They show the three selected IMFs and their SPWVD, respectively.

Figure 4.4 Top plots represent IMF3, IMF4, and IMF5 of a **non-seizure** tEEG rat signal. Shown at the bottom are the corresponding SPWV TF distributions created using two Hamming windows with 64-point length and 128-point length are used respectively, for the time and frequency smoothing windows. TF plots of the three IMFs show that the frequency content of IMF3 is higher than that of IMF4, which in turn, has higher frequency content than that of IMF5.

Figure 4.5 Top plots shows IMF3, IMF4, and IMF5 of **seizure** tEEG rat signal. Shown at the bottom are the corresponding SPWV TF distributions created using two Hamming windows with 64-point length and 128-point length, respectively, for the time and frequency smoothing windows. It is very clear from these SPWVDs, that the frequency content of IMF3 is higher than that of IMF4, which in turn, has higher frequency content than that of IMF5.

4.5 Classification and Performance calculation

After the selection of IMFs from both seizure and non-seizure tEEG segments, each IMF (of 30 second length) is partitioned into epochs of one second length and the energy is calculated for each epoch. This results in a feature vector set consisting of 90 samples (3 IMFs x 30 x 1 second) for each segment. A ten-fold cross-validation technique was applied during the training periods to estimate how well the classification method will classify the new data which were not used during the testing validation period. In ten-fold cross validation, the data set is split into 10 equal sub-set partitions. In each iteration, one of the 10 subsets is used for testing whereas the other

9 subsets are used for training the dataset. The total data were randomly split into ten subsets. The whole procedure is repeated ten times using a different testing subset. The final result is the average of all 10 repetitions. After the training and testing features were selected, they were then applied to a discriminant analysis classifier [91]. Discriminant analysis is a technique used to discriminate a single classification variable using various features. Discriminant analysis also assigns observations to one of the pre-defined groups based on the knowledge of the multi-features [92]. The usefulness of a discriminant model is based on its ability to predict the relationships among known groups in the categories of the dependent variable. In this paper, the MATLAB command “classify” was used and “diagQuadratic” was selected as a discriminant function type [91]. In order to evaluate the performance of the proposed method for seizure detection, the following statistical parameters were calculated [72]: (1) Sensitivity, (2) Specificity, and (3) Accuracy.

4.6 Results and discussion

The proposed algorithm for seizure detection is based on TFR analysis of several oscillating components broken down from the original signal via the EMD algorithm [54]. Localized energy estimates were extracted and considered as features fed into a classifier for discrimination between seizure and non-seizure data. The SPWV distribution was used to calculate the localized energy distribution of the signal.

Table 4.1 The classification performance using 64-point length Hamming window for time smoothing and 128-point length window for frequency smoothing with the five different segment lengths (1 s, 2 s, 3 s, 4 s and 5 s). Sampling Frequency: $F_s=128\text{Hz}$.

	Sensitivity (%)	Specificity (%)	Accuracy (%)
1 s	98.68	98.54	98.61
2 s	97.46	96.17	96.79
3 s	96.53	93.93	95.18
4 s	96.47	93.39	94.87
5 s	98.41	87.24	92.60

After EMD decomposition, the seizure and non-seizure segments may have different total numbers of IMFs. Figure 4.2 shows that the non-seizure segment had twelve IMFs and the Seizure segment had only ten IMFs; this is because the number of IMFs depends on the spectral content of each signal [6]. Furthermore, it can be observed that the frequency content of these IMFs is organized from high to low, and the lower order IMFs capture fast oscillation modes of the signal, whereas the higher order IMFs capture the slow oscillation modes.

The IMF selection method proposed by Flandrin et al. [89] was used to automate selection of the IMFs used in order to reduce the impact of noise. Figure 4.3 show that IMF indices 3 to 10 have energies which exceed the confidence interval (or threshold, blue). So, IMF3, IMF4, and IMF5 were selected for further analysis. The three selected IMFs from each signal were partitioned into one-second epochs using a non-overlapping, sliding Hamming window. Different epoch sizes were tested (1s, 2 s, 3 s, 4 s and 5 s duration), but we found that one-second EEG epochs with a non-overlapping window provided the best detection accuracy of the options tested. Table 4.1 shows the obtained accuracy using different epoch sizes.

Table 4.2 Accuracy obtained using different types of Time-Frequency distributions (TFDs).

	$ \text{STFT} ^2$	WVD	PWVD	SPWVD
Accuracy (%)	89.02	95.86	96.82	98.16

TFR analysis of each epoch was applied using the SPWV distribution. The main reason we used TFR was to have more localized energy concentration. The idea is to analyze behaviors of the energy distribution, i.e., the concentration of energy at certain time instants or certain frequency bands or more generally, in some particular time and frequency region. The localized energy can be a good feature to detect signal events in the SPWVD because the energy in the EEG during seizures is usually larger than during normal activity. Different types of TFRs were tested and the SPWVD was selected since it gave the optimal accuracy of those TFRs tested. Table 4.2 shows the accuracy obtained using Short-time Fourier transform (STFT) squared magnitude (Spectrogram), Wigner-Ville distribution (WVD), Pseudo-Wigner Ville distribution (PWVD) and Smoothed-Pseudo Wigner-Ville distribution (SPWVD). The optimal accuracy was obtained using the SPWVD. Figures 4.4 and 4.5 show examples of SPWVD on 30 s duration seizure and non-seizure segments after the IMFs were selected (IMF3, IMF4 and IMF5). The SPWVD used two Hamming windows with 64-point length for time smoothing window and 128-point length for frequency smoothing window. From these two Figures it can be seen that there is a higher

concentration of energy in seizure segments compared to non-seizure segments and the concentration location depends on the frequency content of each IMF.

Table 4.3. Performance parameters (Sensitivity (Sens), Specificity (Spec) and Accuracy (Acc)) of the proposed method using different combinations for the length of time and frequency smoothing windows.

		Duration length of frequency smoothing window											
		64			128			256			512		
		Sens (%)	Spec (%)	Acc (%)	Sens (%)	Spec (%)	Acc (%)	Sens (%)	Spec (%)	Acc (%)	Sens (%)	Spec (%)	Acc (%)
Duration length of time smoothing window	64	98.32	97.58	97.94	98.99	97.34	98.13	98.68	98.54	98.61	98.5	98.21	98.35
	128	98.17	98.18	98.17	97.99	98.03	98.01	98.98	97.04	97.97	98.99	97.34	98.13
	256	98.54	96.75	97.61	97.80	97.88	97.85	98.98	97.04	97.97	97.33	98.13	97.77
	512	98.89	97.27	98.05	98.13	97.92	98.02	97.96	97.29	97.61	98.32	97.58	97.94

Different lengths of time and frequency smoothing windows (64, 128, 256, and 512 point length) were selected and tested on the proposed method. Table 4.3 shows the computed accuracies using different durations of smoothing windows. It is clear that changing the window length parameters had little effect on the accuracy of the proposed method. The best accuracy obtained was 98.61% achieved using two Hamming windows with 64-point length for time window and 256-point length for frequency window.

These results are very promising, and show that the proposed method has the ability to recognize and classify seizure from non-seizure tEEG segments. The same human dataset [2] used in this study has already been used and tested in our previous work [14] using different features; the obtained accuracy varied between 84.81 and 96.51%. Comparing the results from [14] with the results obtained in this current study

achieving 98.61% accuracy shows that the estimated time-frequency localized energy is a better feature to detect the presence of a seizure in a tEEG signal. Furthermore, the classifiers (SVM and AdaBoost) used in our previous work [14] using the same dataset were tested in this study with the new selected features. The obtained results with the classifier used in this chapter gave better performance than SVM and AdaBoost.

Furthermore, there are other studies based on EMD decomposition for seizure detection such as Tafreshi et al. [76], in which non-Seizure data was distinguished from Seizure data with success rates up to 95.42%. Pachori and Bajaj [75] used the area measured from the trace of the analytic IMFs as features to analyze normal and epileptic human seizure EEG signals and they showed that those calculated areas gave good discrimination performance. Comparing our results to those in the literature which used EMD and TFRs for seizure detection, the high accuracy obtained from our method, with 98.61% accuracy, using data from freely moving rats shows that our proposed method for seizure detection is competitive with other methods disclosed. Again the human data used in other researcher's work often has fewer interfering artifacts than the data recorded while the rats were moving, grooming, etc. which often renders the rat's data more difficult to analyze and classify.

4.7 Summary

In this chapter, we proposed an algorithm for seizure detection based on the recognition of relevant changes of the signal's frequency content in different sub-band levels. For this purpose, time-frequency localization of signal energy was essential. The analysis is performed in three stages: (i) decomposition of tEEG seizure and non-seizure segments into oscillation IMF components using the EMD method; this process helps the analysis of the signal using different frequency levels, (ii) TF analysis of each selected IMF. (iii) Feature extraction. An energy estimate was applied to each selected IMF on specific selected windows length (one second duration) and (iv) classification of the EEG segment (to check the existence of seizure or not on each segment and classify them), was performed using a diagonal-quadratic discriminant analysis classifier. The methods were evaluated using a rat dataset recorded with the new electrodes TCRE. Sensitivity, specificity and the accuracy results were presented and shown that the proposed algorithm is effective to detect seizures from rat's tEEG signals.

CHAPTER 5

NON-STATIONARY FM CHIRP SIGNAL PROCESSING

5.1 Introduction

Frequency modulated (FM) signals, also known as chirp signals, are non-stationary signals that may serve as a paradigm for some non-stationary deterministic signals [93]. Chirps can be observed in a number of important scientific areas including the analysis of echolocation in bats, animal communication (birds, whales, etc.), acoustics (propagation of impulses in dispersive media), seismic surveying, and biology (epileptic seizure activity in EEG data, contraction in EMG, etc.) [57, 107].

In this chapter the Fractional Fourier Transform (FrFT), which is a generalization of the Fourier Transform, was used to decompose multi-component linear and non-linear FM chirp signals. However, it shows some drawbacks when the multi-component signals are composed of non-linear FM chirp signals. To overcome these problems a warping operator was used to linearize the time-frequency behavior of the signal which makes the decomposition easier. For this purpose, a new warping function is proposed along with a parametric optimization scheme. The performance analysis proves that this new algorithm can successfully be used to linearize the TF structure of non-linear FM chirp signals with reduced inner-interference terms.

5.2 Fractional Fourier Transform (FrFT)

The Fractional Fourier Transform [93] was first introduced by Namias in the field of quantum mechanics for solving some classes of differential equations efficiently. Since then, a number of applications of FrFT have been developed. It was discussed in chapter 2 about the importance and ubiquity of the ordinary Fourier transform (FT). The FrFT is a generalization of the ordinary Fourier Transform with more flexibility, rich in theory and comparable computational cost. The FFT has many applications in filter design [96], communications [97], and TFRs [98]. Ozaktas et al. [94, 95] have come up with a discrete implementation of the FrFT, like Cooley-Tukey's FFT. Their algorithm computes the FrFT in $O(N \log N)$ time, which is about the same cost as the ordinary FFT. Thus, FT and FrFT have comparable implementation cost. One important motivation behind the proposed method is the ability of the FrFT to analyze linear chirp signals better than the conventional FT. The FrFT of a given signal $x(t)$ is [93]:

$$FrFT^\phi(t, u) = F^\alpha(x(t)) = X_\phi(u) = \int_{-\infty}^{+\infty} x(t) K_\phi(t, u) dt. \quad (5.1)$$

where, α is a real number called the order of the FrFT, $\phi = \alpha \frac{\pi}{2}$ is the angle of rotation, $F^\alpha(.)$ denotes the FrFT operator, and $K_\phi(t, u)$ is the kernel of the FrFT,

$$K_\phi(t, u) = \begin{cases} \sqrt{\frac{1 - j \cot \phi}{2\pi}} \exp(j \frac{t^2 + u^2}{2} \cot \phi - jtu \csc \phi), & \text{if } \phi \neq n\pi \\ \delta(t - u), & \text{if } \phi = 2n\pi \\ \delta(t + u), & \text{if } \phi = (2n + 1)\pi \end{cases}, \quad (5.2)$$

where n is an integer.

The kernel $K_\phi(t, f)$ has the following properties:

$$K_{-\phi}(t, f) = K_\phi^*(t, f),$$

$$\int_{-\infty}^{+\infty} K_\phi(t, f) K_\phi^*(t, f') dt = \delta(f - f').$$

For $\phi = \alpha \frac{\pi}{2}$, when $\phi = 0$, then $\alpha = 0$, and the FrFT of the signal $x(t)$ is the signal itself i.e. $F^0(x(t)) = x(t)$. When $\phi = \pi/2$, then $\alpha = 1$, the FrFT becomes the FT of the signal $x(t)$, i.e. $F^1(x(t)) = X(f)$. That is why the FrFT is considered as a generalization of the FT. Moreover, the FrFT is linear, commutative and associative. In general, the FrFT computation can be interpreted as multiplication by a chirp signal in one domain followed by a FT, then multiplication by a chirp signal in the transform domain and finally a complex scaling. Table 5.1 summarizes the important properties of FrFT in the time-frequency plane.

$$FrFT^\phi(t, u) = X_\phi(u) = \begin{cases} \sqrt{\frac{1-j\cot\phi}{2\pi}} e^{j\frac{u^2\cot\phi}{2}} \int_{-\infty}^{+\infty} \left[x(t) e^{j\frac{t^2\cot\phi}{2} - jtu\csc\phi} \right] dt, & \phi \neq n\pi \\ x(t) & \phi = 2n\pi \\ x(-t) & \phi = (2n+1)\pi \end{cases} \quad (5.3)$$

In other words, linear FM chirps are similar to basis functions in FrFT computations. This means that the FrFT can be helpful solving problems where linear FM chirps signals are involved. The inverse FrFT is defined as:

$$x(t) = FrFT^{-\phi}(X_\phi(u)) = \int_{-\infty}^{+\infty} X_\phi(u) K_{-\phi}(t, u) du. \quad (5.4)$$

Table 5.1 Properties of the Fractional Fourier Transform (FrFT) as related to the Fourier transform (FT) [96]. “I” is the identity matrix.

Property	Order	FrFT
Identity operator	(Zero rotation) $\alpha = 0, \phi = 0$ $\alpha = 4, \phi = 2\pi$	$FrFT^0 = FrFT^{2\pi} = I$
Fourier operator	($\pi/2$ rotation) $\alpha = 1, \phi = \pi/2$	$FrFT^{\pi/2} = FT$
Reflection operator	(π rotation) $\alpha = 2, \phi = \pi$	$FrFT^\pi = -I$
Inverse FrFT	$\alpha = 3, \phi = 3\pi/2$	$FrFT^{3\pi/2} = FrFT^{-1}$

5.2.1 Chirps and Fractional Fourier Transform

Linear and non-linear chirps are signals with nonstationary instantaneous frequency (IF), which means the IF changes with time. Consider a linear FM chirp, $x(t) = e^{j(c_1 t^2 + bt + c)}$; the chirp rate $2c_1$ is related to the Fractional Fourier transform order α . In this case the FrFT order parameter α can be used to adjust the transform to find out an appropriate energy concentration of the given LFM chirp auto-components.

When the FrFT axis of rotation matches the chirp rate of the signal, the FrFT of the chirp signal will produce an impulse and the magnitude of the FrFT will reach its maximum magnitude [93-97]. The corresponding order of the FrFT is called the optimum α :

$$\alpha_{opt} = \frac{2\phi}{\pi} = \frac{2}{\pi} \tan^{-1} \left(\frac{1}{2c_1} \right) = \frac{2}{\pi} \tan^{-1} \left(\frac{Fs^2 / N}{2c_1} \right) \quad (5.5)$$

with F_s equal to the sampling frequency (in Hz) and N is the number of samples in the chirp signal. Equation (5.5) shows that the appropriate choice of the FrFT fraction order depends on the chirp rate. However, the FrFT optimum order α cannot be found

analytically, especially if we don't have any information about the chirp signal. For this reason, a one dimensional search algorithm must be applied to select the optimal order $\alpha \in [-1,1]$ using a fine spacing to get a good estimate of this order. The optimum value of α will be selected when the energy of the linear FM chirp focuses well in the time-frequency plane and the FrFT spectrum is highly concentrated and reaches a maximum magnitude.

Figure 5.1 The relation of the time-frequency plane (t, f) with the Fractional Fourier domain plane (u, v) rotated by an angle ϕ .

5.2.2 The relation between the FrFT and the Wigner-Ville Distribution (WVD)

The rotation of the time-frequency plane by an angle ϕ in the transformed coordinates (u, v) used in the FrFT is given by:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} t \\ f \end{pmatrix}, \quad \text{where } \phi = \alpha \frac{\pi}{2}.$$

Figure 5.1 shows the relation of time-frequency plane coordinates (t, f) with the Fractional Fourier domain co-ordinates (u, v) rotated by an angle ϕ . It was discussed in

Chapter 2 that the Wigner-Ville distribution (WVD) is considered to be a very important time-frequency analysis tool. The fundamental relationship between the FrFT and the WVD is [115, 116]:

$$WVD_{F^\alpha(x(t))}(u, v) = WVD_{x(t)}(t \cos \phi - f \sin \phi, t \sin \phi + f \cos \phi) \quad (5.6)$$

Figure 5.2 Linear FM chirp signal: (A) Time domain plot. (B) The WVD. (C) The FT spectrum of the signal in subplot A. (D) FrFT Output of a chirp signal in the first subplot “A” for different values of the FrFT order α . The optimum value of α is 0.1900, where the peak is sharpest and tallest. (E) The WVD using the FrFT of the corresponding optimum value $\alpha = 0.1900$.

This corresponds to rotation of the signal $x(t)$ at the origin in a counter-clockwise direction with rotational angle $\phi = \alpha \frac{\pi}{2}$ in the time-frequency domain as shown in Figure 5.1. Figure 5.2-B shows the WVD of a linear chirp signal with an increasing linear frequency modulation from 0.1 to 0.4 Hz, the sampling frequency is $F_s=1$ Hz, and the number of frequency points (or frequency bins) is $N=128$. The time-domain plot of the signal $x(t)$ is shown in Figure 5.2-A.

It is shown in Figure 5.2-C that the Fourier Transform spectrum of the chirp signal is spread out over many frequencies whereas its FrFTs in Figure 5.2-D get sharper as the value of α gets closer to the optimum value. The FrFT for the optimum order $\alpha = 0.1900$ is highly concentrated and appears as an impulse. Moreover, the WVD of the FrFT corresponding to this optimum value is a horizontal linear chirp with concentrated energy in Figure 5.2-E. Thus, the effect of the FrFT on the original signal appears as a rotation of the original WVD. This result demonstrates the work of Almeida who analyzed the relationship between the FRFT and the WVD, and interpreted it as a rotation operator in the time-frequency plane [97]. This rotation will appear as a horizontal chirp when the optimum value of α is used in the FrFT. Hence, it can be said that the WD is rotationally covariant to the effect of the FrFT [99].

5.3 Multi-component FM chirp signals decomposition using the FrFT

One important advantage of time-frequency representations (TFRs) is the capability of analyzing non-stationary signals [100-104]. Among these TFRs, it was proven that the WVD is a very effective representation with high energy resolution in both time and frequency [40]. However, with multi-component signals the WVD

exhibits cross-terms because of the bilinear form in the WVD definition. Hence, a better method is needed for good signal concentration. The FrFT attracts more and more attention in the signal processing literature especially for solving the problems which appear when using multi-component chirp signals since the FrFT is a linear operator. The components in the multi-component chirp signals may overlap in frequency, in time or both. The separation of FM chirp signals using the FrFT is not a new problem and a lot of work has been done on this subject [124-128]. L. Zhang et al. [124] applied the FrFT to estimate the instantaneous frequency of multi-component chirp signals. First, they transformed the multi-component chirp signals through the FrFT, then they found the order α and applied a band-pass filter to extract each component. Finally, the instantaneous frequency of each decomposed single-component chirp signal was estimated. Y. Lu et al. [125] used the FrFT based chirplet signal decomposition algorithm to isolate the dominant chirplet echoes for successive steps in signal decomposition and parameter estimation of ultrasonic signals. They used the FrFT to rotate the signal with an optimal transformed order α based on the highest kurtosis value of the signal in the FrFT domain.

Let us consider a multi-component FM signal

$$x(t) = \sum_{i=1}^n x_i(t) = \sum_{i=1}^n A_i(t) e^{j2\pi\varphi_i(t)} \quad (5.7)$$

where $x(t)$ is a multi-component signal expressed as the sum of “n” components $x_i(t)$ with $i = 1, \dots, n$. These components may overlap in the time-frequency plane. $A_i(t)$ is the magnitude and $\varphi_i(t)$ is the phase for the component $x_i(t)$. Each FM chirp signal $x_i(t)$ has an optimal order α_i for the FrFT used to rotate it. Using the FrFT-based

signal decomposition, one of the FM chirp signals will be rotated optimally while the other components are ignored. A rectangular filter will be applied to the FrFT to separate this rotated component. The different steps, shown in Figure 5.3, are as follows:

- 1- Search algorithm for the optimal transform order $\alpha_{i_{opt}}$, of the chirp component $x_i(t)$, (with $i \leq n$) related to the maximum Fractional FT spectrum $(\max_{\phi_i} |FT(FrFT^{\phi_i}(x(t)))|)$, since $\phi_i = \alpha_i \frac{\pi}{2}$.
- 2- Signal filtering using a Rectangular Window:
 - For linear FM chirp signals: only a small number of samples on either side of the bin location of the peak magnitude of the FrFT of the chirp component will be selected using a rectangular window.
 - For non-linear FM chirp signals: the FrFT of a non-linear FM chirp signal is spread out more than that of a linear FM chirp signal. Hence, the number of samples selected on either side of the peak, via windowing, will be larger for non-linear chirps.
- 3- After doing an inverse FrFT with the optimum order $(-\alpha_{i_{opt}})$, plot the WVD to determine whether the chirp component $x_i(t)$ is well decomposed.
- 4- Obtain the residual signal by subtracting the decomposed chirp component $x_i(t)$ from the signal in the previous iteration.
- 5- These steps will be repeated to iteratively extract each of the chirp components one after the other, $i = 1, \dots, n$.

Figure 5.3 Block diagram for FrFT-based FM chirp signal extraction.

Note: In the next proposed examples (1-3), we will decompose the components by searching in each iteration for the corresponding optimum order α and then apply the other steps of the algorithm. This means we will skip the step of subtracting the decomposed component from the signal in the previous iteration. We do this for demonstration purposes to show how the optimum order α changes by changing the chirp signal we optimize for. For example 4 we will apply the algorithm as described in Figure 5.3.

In the following examples, we will illustrate how the FrFT-based multi-component FM chirp signal decomposition algorithm described in Figure 5.3 works. In example 4, we will apply the algorithm to decompose a real-life bat echolocation chirp signal into three sub-components. In examples 1-3, the length of the rectangular window used to filter the FrFT near the highest amplitude will be selected based on the calculated minimum mean squared error (MSE) [129] value between the original and the retrieved chirp components; this of course, assume we know all the information of the original signal components (for examples 1-3). However, in the real world we can't use the MSE since we won't have information about the original signal, so in our future work, we will try to optimize the length of the rectangular window by finding a relation between the filter length and the chirp signal. For Examples 1-3, MATLAB toolbox in [90] was used to generate the signals.

Figure 5.4 Example 1: *multi-component* linear FM signal; **non-overlapping** TF support. **(A)** Time domain plot. **(B)** Time-frequency representation (the WVD).

Example 1: Figure 5.4-A shows a multi-component signal that is the sum of two linear FM chirp signals. For simplicity in the first example the frequency support regions of the two chirps do not overlap. The first chirp signal has constant amplitude with an increasing frequency support from 0.05 to 0.1 Hz and the second chirp has constant amplitude with decreasing frequency support from 0.45 to 0.2 Hz. The sampling frequency is $F_s=1$ Hz and the number of frequency points is $N=256$. From the time-domain plot of this signal, it is difficult to identify the number of signal components, and extremely difficult to determine their nature. One way to determine these signal characteristics is the time-frequency representation (TFR), which can make signal analysis easier. The WVD of this signal is shown in Figure 5.4-B. From this representation, it is evident that there are two FM chirp components comprising the signal shown in Figure 5.4-A. Moreover, the linear frequency nature of the chirps is easily visualized as well. The oscillatory terms in Figure 5.4-B are cross terms, which occur midway between two signal components in the quadratic WVD.

Figure 5.5 Example 1 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{1opt} = 0.032$ of the first chirp signal. (B) WVD of the FrFT in subplot A. (C) FrFT after applying a rectangular window isolating the sharper spectrum. (D) WVD of the filtered signal in subplot C. (E) WVD of the reconstructed signal of the first chirp signal after applying the inverse FrFT with an order $-\alpha_{1opt} = -0.032$.

From Figure 5.5-A, the selected optimum α_{1opt} for the first chirp signal which yielded the higher peak is equal to $\alpha_{1opt} = 0.032$. Figure 5.5-B shows the rotation of the WVD by the FrFT. After selecting the optimum order for the first chirp, a rectangular window was used to filter out the concentrated narrow pulse with the highest amplitude of this chirp from the fractional Fourier domain.

Figure 5.6 Example 1 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{2_{opt}} = -0.155$ of the second chirp signal. (B) WVD of the FrFT in subplot A. (C) FrFT after applying a rectangular window isolating the sharp peak in the spectrum. (D) WVD of the filtered FrFT in subplot C. (E) WVD of the reconstructed signal component of the second chirp signal after applying the inverse FrFT with an order $-\alpha_{2_{opt}} = 0.155$.

Only the peak and two bins on either side of the peak are retained as shown in Figure 5.5-C. Figure 5.5-D shows the WVD of the windowed component. The effect of the FrFT with an order $\alpha_{1_{opt}} = 0.032$ is shown as a horizontally concentrated chirp in the time-frequency plane. Applying an inverse FrFT with the order value $-\alpha_{1_{opt}} = -0.032$, the first FM chirp signal was easily reconstructed as shown in Figure 5.5-E. Comparing Figure 5.4-B with Figure 5.5-E shows that the extraction algorithm worked well on the low frequency chirp signal.

To decompose the second FM chirp signal, this process will be repeated as shown in Figure 5.6. In this case the concentrated impulse with the highest FrFT peak value will appear when the optimum order value is equal to $\alpha_{2opt} = -0.155$. After selecting the optimum order for the second chirp, a rectangular window was used to filter out the concentrated narrow pulse of this chirp from the Fractional Fourier domain. Only the peak and two bins on either side of the peak are retained as shown in Figure 5.6-C. Figure 5.6-D shows the WVD the windowed signal in subplot C. To reconstruct the second FM chirp signal, an inverse FrFT will be applied with an order $-\alpha_{2opt} = 0.155$. The WVD plot of the second decomposed chirp signal is shown in Figure 5.6-E; the decomposed chirp signal in Figure 5.6-E is a good match to the high frequency chirp signal in Figure 5.4-B.

Figure 5.7 Example 2: *multi-component* crossing linear FM signal; (A) represents the time domain representation, and (B) Time-frequency representation (WVD).

Example 2: Figure 5.7-A depicts a multi-component signal with two crossing linear FM chirps, one with constant amplitude and an increasing frequency support from 0.1 to 0.45 Hz whereas the second has constant amplitude and decreasing frequency

support from 0.3 to 0.05 Hz. The sampling frequency is $F_s=1\text{Hz}$ and the number of frequency points is $N=256$. As in the last example, from the time-domain plot of this signal, it is difficult to identify the number of signal components, and extremely difficult to determine their nature.

Figure 5.8 Example 2 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{1opt} = 0.215$ of the first chirp signal. (B) WVD of the FrFT in subplot A. (C) FrFT after applying a rectangular window for the sharper spectrum. (D) WVD of the filtered signal in subplot C. (E) WVD of the reconstructed signal of the second chirp signal after applying the inverse FrFT with an order $-\alpha_{1opt} = -0.215$.

The FrFT-based multi-component signal decomposition algorithm was used to extract the linear FM chirps that overlapped in the TF plane as shown in Figure 5.7-B. For the first chirp (with an increasing frequency from 0.1 to 0.45 Hz), the optimum order value used was $\alpha_{1opt} = 0.215$. The FrFT and the WVD are shown in Figure 5.8-A and Figure 5.8-B, respectively. A rectangular window was used to filter out the concentrated narrow pulse of this chirp in the Fractional Fourier domain. Only the peak and two bins on either side of the peak were retained as shown in Figure 5.8-C.

Figure 5.9 Example 2 (continued): (A) The FrFT of the multi-component signal with the corresponding $\alpha_{2opt} = -0.157$ of the second chirp signal. (B) WVD of the FrFT in subplot A. (C) FrFT after applying a rectangular window for the sharper spectrum. (D) WVD of the filtered signal in subplot C. (E) WVD of the reconstructed signal of the second chirp signal after applying the inverse FrFT with an order $-\alpha_{2opt} = 0.157$.

The WVD of the windowed signal is shown in Figure 5.8-D. To reconstruct the first chirp signal, an inverse FrFT of the windowed FrFT was applied with $-\alpha_{1opt} = -0.215$. The corresponding WVD plot of the first decomposed chirp signal is shown in Figure 5.8 -E. The chirp in Figure 5.8-E should be compared to the chirp component with decreasing frequency content in Figure 5.7-B.

For the second chirp signal (with decreasing frequency support from 0.3 to 0.05 Hz) the optimum order value was determined to be $\alpha_{2opt} = -0.157$. The whole process used to extract the second chirp signal is shown in Figure 5.9. The same for the second chirp signal, a rectangular window was used to select the concentrated narrow pulse of this chirp from the Fractional Fourier domain. Only the peak and two bins on either

Figure 5.10 Example 3: *multi-component* non-linear non-crossing FM signal; (A) depicts the time domain representation, and (B) time-frequency representation (The WVD). (C) The FrFT of the signal in (A) after selecting the optimum order $\alpha_{opt} = -0.07$. (D) The WVD of the FrFT in subplot (C).

side of the peak are retained as shown in Figure 5.9-C. The WVD of the windowed signal is shown in Figure 5.9-D. To reconstruct the second chirp signal an inverse FrFT was apply with $-\alpha_{2_{opt}} = 0.157$ to the windowed FrFT spectrum. The WVD of the second decomposed chirp signal is shown in Figure 5.9-E and is very similar to the increasing frequency FM chirp component in Figure 5.7-B.

Example 3: In this example we will apply the FrFT-based decomposition algorithm to decompose non-linear and non-overlapping time-frequency support signal. Figure 5.10-A depicts a multi-component signal consisting of two non-overlapping and (non-linear) parabolic FM chirps, one with constant amplitude and decreasing frequency

Figure 5.11 Example 3 (continued): **(A)** FrFT after applying a rectangular window corresponding to the first component. **(B)** WVD of the first component after applying an inverse FrFT on windowed spectrum. **(C)** FrFT after applying a rectangular window for the second component. **(D)** WVD of the second component after applying an inverse FrFT on windowed spectrum.

support from 0.5 to 0.35 Hz and the other with constant amplitude and decreasing frequency support from 0.25 to 0.15 Hz. The sampling frequency is $F_s=1\text{Hz}$ and the number of frequency bins is $N=256$. The WVD plot is shown in Figure 5.10-B. In this example, the optimal order is $\alpha_{1opt} = \alpha_{2opt} = -0.07$, i.e the same for the two signal chirps components. Figure 5.10-C and Figure 5.10-D show the FrFT with an order $\alpha_{1opt} = \alpha_{2opt} = -0.07$ and the corresponding WVD, respectively.

Figures 5.11-A and 5.11-B show the filtered Fractional Fourier spectrum and the corresponding WVD of the first chirp component, respectively. Figures 5.11-C and 5.11-D show the filtered Fractional Fourier spectrum and the corresponding WVD of the second chirp component, respectively. The length of the rectangular window is forty five bins as shown in Figure 5.11-A and 5.11-C for both chirp signals. The FM chirp components are extracted using the inverse FrFT with an order $-\alpha_{1opt} = -\alpha_{2opt} = 0.07$ of the windowed FrFT. Figures 5.11-B and D show that the FrFT-based decomposition algorithm successfully decompose the two non-linear chirp signals, since their spectrum had little overlap as shown in Figure 5.10-C. However, the inner-interference terms still exist on the WVD plot of the decomposed signals as shown in Figures 5.10 and 5.11-B and D.

Example 4: The FrFT-based decomposition algorithm was tested on the real-life bat echolocation chirp signal [122]. The signal is a 400 sample long recording of a bat chirp signal sampled with a sampling period of 7 microseconds. The WVD (Figure 5.12-A) shows this signal to contain three strong, quasi-linear FM chirp components. Figures 5.12-B, 5.12-C and 5.12-D are the WVD of the decomposed components

extracted using the FrFT-based multi-component signal decomposition algorithm in Figure 5.3. The optimal order of the first component, in Figure 5.12-B, was $\alpha_{1opt} = 0.13$. The optimal order of the second component, Figure 5.12-C was $\alpha_{2opt} = 0.175$. And the optimal order of the third component, Figure 5.12-D was $\alpha_{3opt} = 0.18$. The length of the windows were: nine bins, eleven bins and eighteen bins for the first, the second and the third component, respectively. The application of the FrFT-based signal decomposition algorithm reveals that the bat signal mainly contains three chirp signals in time-frequency (TF) domain and can provide a high resolution of their TF representation.

Figure 5.12 Example 4: Application of the FrFT-based multi-component signal decomposition algorithm on real Bat data. (A) The WVD of the multi-component bat signal; the signal is composed of three chirp components with little overlap in their time-frequency content. Subplots (B), (C) and (D) show the WVD of the three decomposed components composing the bat signal.

5.4 Limitations of FrFT-based multi-component signals decomposition algorithm

In the previous section, we have illustrated the methodology of the FrFT in decomposing multi-component FM chirp signals. We also shows how this algorithm can successfully decompose some non-linear and non-crossing FM chirp signals. The algorithm was also applied on a bat echolocation signal which consisted of multiple chirp signals with approximately linearly changing frequency behavior.

In this section, the FrFT-based multi-component signal decomposition will be applied on other types of multi-component FM chirp signals. Example 5 demonstrates the case of a multi-component signal composed of two non-linear FM chirp signals that overlap in time and in frequency.

Example 5: The signal $x(t) = x_1(t) + x_2(t)$, is composed of two non-linear FM chirp components that overlap in time and in frequency. The first component is $x_1(t) = e^{j2\pi 0.014 t^{1.5}}$ and the second component is $x_2(t) = e^{j2\pi (0.4t - 0.05 t^{1.3})}$. The number of frequency bins used is $N=256$ and the sampling frequency is $F_s=1$ Hz. Figure 5.13-A shows the WVD of the chirp mixture. It can be seen that the frequency behavior of the two components changes nonlinearly in time and overlaps in frequency and in time. Figure 5.13-B shows the regular FT spectrum of the multi-component signal. It can be seen that the FT spectrum and the time support of the two components $x_1(t)$ and $x_2(t)$ overlap.

The FrFT-based multi-component signal decomposition algorithm described in Figure 5.3 was applied, by searching over several values for the optimum α . It was found that using the orders $\alpha_{1opt} = 0.15$ and $\alpha_{2opt} = -0.085$, the outputs of the FrFT

Figure 5.13 Example 5: Decomposition of multi-component overlapping and non-linear chirp signal. (A) The WVD of the multi-component signal. (B) The FT spectrum of the multi-component signal in A. (C) the FrFT with order $\alpha_{1opt} = 0.15$ for the first component $x_1(t)$. (D) The windowed FrFT of the first chirp component. (E) The WVD of the first reconstructed chirp component after applying an inverse FrFT with $-\alpha_{1opt} = -0.15$. (F) The FrFT with order $\alpha_{2opt} = -0.085$ for the second chirp component $x_2(t)$. (G) The windowed FrFT of the second chirp component. (H) The WVD of the second reconstructed chirp component $x_2(t)$ after applying an inverse FrFT with $-\alpha_{2opt} = 0.085$.

showed a large peak amplitude (Figure 5.13-C and Figure 5.13-F) for the first chirp component $x_1(t)$ and the second chirp component $x_2(t)$, respectively. Figures 5.13-D and 5.13-G show the filtered Fractional FT spectrum for the first chirp component and the second chirp component, respectively. The length of the rectangular filter used was forty five bins for the first chirp signals and seventy one bins for the second chirp component. To reconstruct the components $x_1(t)$ and $x_2(t)$, an inverse FrFT was applied with $-\alpha_{1opt} = -0.15$ and $-\alpha_{2opt} = 0.085$, respectively, to the filtered FrFT. Figures 5.13-E and 5.13-H show the corresponding WVD of the first and the second reconstructed chirp components, respectively.

The WVD of the first extracted chirp component, in Figure 5.13-E, shows that the algorithm reconstructed the first chirp component $x_1(t)$ along with a part of the second component $x_2(t)$. For the second extracted chirp signal, the WVD in Figure 5.13-H shows that the algorithm reconstructed this signal with a part of the first signal $x_1(t)$ as seen in Figure 5.13-H. This problem appears because of the overlapping time-frequency support of the two FM chirp components $x_1(t)$ and $x_2(t)$. This demonstrates a drawback or the limitation of the FrFT-based multi-component signal decomposition for non-linear FM chirp signals.

From Example 5 we can conclude that the FrFT-based multi-component signal decomposition often fails to extract the components from multi-component signals comprised of non-linear overlapping FM chirp components. This problem demonstrates that the FrFT works well in analyzing linear FM chirp signals, however there are a lot of drawbacks when the FM chirp signals are nonlinear.

To solve the problem of decomposing multi-component signals composed of non-linear FM chirp components, we propose to use the principle of warping-based Time-Frequency Representations [105, 106, 108, 109]. In signal processing, warping operators have been used to build Time-Frequency Representations with reduced inner-interference terms [105, 106, 108, 109]. In this chapter, the principle of time warping will be used to linearize the time-frequency support of non-linear chirp signals. Definitions of time warping with all the equations will be given in the next section.

The principle of the proposed warping-based multi-component FM chirp decomposition algorithm is to apply time warping-based Time-Frequency Representations on the multi-component signal to linearize as much as possible each occurrence of a non-linear FM chirp component we want to extract. As a result of the linearization, the corresponding spectrum will have a narrower, higher amplitude and the WVD will show a linear time-frequency structure. The next step is to filter out the spectrum peak using a rectangular window. The last step is to apply the inverse transform, unwarp the filtered signal and plot the WVD of the extracted non-linear FM chirp component, using the original (unwarped) time-frequency axis.

Comparing the proposed warping-based multi-component signal decomposition algorithm to the FrFT-based multi-component signal decomposition, one can say that the general framework of both algorithms is the same. The principle is to filter out in each iteration the spectrum with the higher amplitude corresponding to a “linearized” FM chirp component.

Another drawback of the FrFT-based decomposition is the existence of the inner-interference terms in the time-frequency plots of the extracted component as seen in Figures 5.11-B and D. This problem appears when the components are non-linear chirps which means that the FrFT-based decomposition cannot reduce the inner-interference terms, which we will demonstrate, is not the case of the proposed time warping-based decomposition.

In the next sections of this chapter, we will explain the new warping function and we will provide some examples demonstrating that the new warping function can work better than the power warping function used in the literature to linearize the time-frequency structure of certain non-linear FM chirp signals. We will also show that the time-frequency distribution of the extracted component using the warping-based decomposition algorithm exhibits reduced inner-interference terms.

5.5 Time-warping principle

One simple and powerful tool used to generate new classes of TFRs is the unitary similarity transformation [105, 106, 108, 109, 118]. It permits construction of new TFRs to match closely any FM chirp signal with one-to-one group delay or instantaneous frequency characteristics [105, 106, 108].

Let $x(t) \in L^2(\mathfrak{R})$ be a signal that is a non-linear function of “ t ”. Mathematically, if $W : L^2(\mathfrak{R}) \rightarrow L^2(\mathfrak{R})$ is the warping operator, its effect on $x(t)$ is given by [105, 106]:

$$W_x(t) = \sqrt{|\dot{w}(t)|} \quad x(w(t)) \quad (5.8)$$

where $w(t)$ is the time-warping function, which is assumed to be a smooth one-to-one function [105], and $\dot{w}(t) = \frac{d}{dt} w(t)$ is its derivative with respect to t . The purpose of the warping is to transform the non-linear time-frequency support of $x(t)$ into an equivalent linear one.

Since the effect of the time-warping function is the “linearization” of the time-frequency structure of the analyzed signal, this means that the degree of linearization can be observed by plotting the TFR of the warped signal. For the time-warping operator $W_x(t)$, the new time-frequency coordinates are related to the standard ones via the transformation [105]:

$$t' = w(t) \quad ; \quad f' = f \dot{w}^{-1}(w(t)). \quad (5.9)$$

where $w^{-1}(t)$ is the inverse function of $w(t)$ and $\dot{w}^{-1}(t) = \frac{d}{dt}[w^{-1}(t)]$ represent its first derivative. The warped-based TFR is obtained by calculating the new TF coordinates related to the standard one via [105, 106]:

$$t'' = w^{-1}(t) \quad ; \quad f'' = f \dot{w}(w^{-1}(t)). \quad (5.10)$$

The inverse warping consists in transforming the coordinates of the TFR plane according to the real time and frequency coordinates using equation (5.10). Using a valid warping transformation, it is always possible to construct quadratic TFRs (QTFRs) for arbitrary one-to-one phase functions [105, 106, 109, 118]. The choice of the function $w(t)$ is critical, especially if the analyzed signal has an unknown nature. This can be solved by estimating the correct model for the signal [108]. The warping functions $w(t) = e^t$ for exponential FM chirps and $w(t) = |t|^k \text{sgn}(t)$ for power FM

chirps, with $k \neq 0$, provide examples of useful warping operators [117, 118]. Different power classes are obtained by varying the parameter k . The affine class of TFRs [108] is obtained when $k=1$. The k^{th} exponential class quadratic time-frequency representations (QTFRs) are obtained when $w(t) = e^{kt}$, with $k \neq 0$. When $k=1$, the k^{th} exponential power yields the exponential class [109, 118]. Using the warping function $w(t) = t^{\frac{1}{k-1}}$, Iacona Cornel et al. [108] proposed a new method to solve the noise robustness performance of classical high-order methods polynomial signal processing. G. Le Touzé et al. [111] used an axis warping operator to realize mode filtering adapted to guided waves. To avoid the problems due to non-linear time-frequency structures, they used a classical unitary operator [105]. For the m^{th} mode, where C_1 is the velocity in the water layer, D the depth guide and R the source-sensor distance, their warping function is $w(t) = \frac{(k^2 t^2 + R^2)^{1/2}}{C_1}$, with $k = D/4(2m-1)$.

Let the signal $x(t)$ be given by:

$$x(t) = A(t)e^{j2\pi\varphi(t)} \quad (5.11)$$

where $A(t)$ is the non-negative amplitude, $\varphi(t)$ is the phase and $IF_x(t) = \frac{d}{dt}\varphi(t)$ is the instantaneous frequency. Ideally, the warping function for the FM signal is defined as the functional inverse of its phase [105]: $w(t) = \varphi^{-1}(t)$. Assuming $\varphi^{-1}(t)$ exists and is one-to-one, we can use it to warp the signal in equation (5.11) to obtain the following form:

$$W_x(t) = A(w(t))e^{j2\pi t}. \quad (5.12)$$

The following FM signal with constant amplitude demonstrates the property of warping operators for time-frequency linearization. Let:

$$x_1(t) = e^{j2\pi ct^k} \quad (5.13)$$

be a power chirp signal with the phase power parameter k . The associated warping operator can be defined as [119]:

$$w(t) = t^{1/k} \quad (5.14)$$

The first-order derivative of the function $w(t)$ is given by:

$$\dot{w}(t) = \frac{d}{dt} w(t) = \frac{1}{k} t^{(\frac{1}{k})-1} \quad (5.15)$$

And the inverse function is define by: $w^{-1}(t) = t^k$. The corresponding warped signal from equation (5.8) is:

$$W_{x_1}(t) = \sqrt{\left| \frac{t^{(1-k)/k}}{k} \right|} e^{j2\pi c(t^{1/k})^k}, \quad \text{with} \quad t \in \mathfrak{R}^+.$$

Thus,

$$W_{x_1}(t) = \sqrt{\left| \frac{t^{(1-k)/k}}{k} \right|} e^{j2\pi ct}. \quad (5.16)$$

For the signal $x_1(t) = e^{j2\pi 0.014 t^{1.5}}$ with $k = 1.5$, the warped signal $W_{x_1}(t)$ is given by:

$$W_{x_1}(t) = \sqrt{\left| \frac{t^{(-0.333)}}{1.5} \right|} e^{j2\pi 0.014 t}. \quad (5.17)$$

Figure 5.14 Example 6: power signal $x_1(t) = e^{j2\pi 0.014 t^{1.5}}$. (A) The WVD of the original signal $x_1(t)$. (B) The time-warping function $w(t) = t^{1/k}$, and its inverse $w^{-1}(t) = t^k$, for $k=1.5$. (C) The effect of the warping operator on the WVD of the warped signal. (D) The FT spectrum of the warped signal.

The WVD of the signal $x_1(t)$ is depicted in Figure 5.14-A. The WVD of the warped signal $W_{x_1}(t)$ and its FT spectrum are shown in Figures 5.14-C and 5.14-D, respectively. Figure 5.14-B shows the time-warping function $w(t) = t^{1/1.5}$ and its inverse $invw(t) = w^{-1}(t) = t^{1.5}$. Application of the warping operator on the power FM signal produces the linearization of the time-frequency support in the TF plane. The time-frequency support of the signal has moved from being concentrated along the curve $IF_{x_1}(t) \propto t^{1.5}$ in Figure 5.14-A to a horizontal line in the time-frequency plane

in Figure 5.14-C. This example demonstrates the capability of the warping operator concept proposed in [105, 118] to “linearize” the time-frequency content of a power signal.

5.5.1 New warping function

One goal of a warping operator is to linearize the TF behavior of a signal [105, 106]. Thus, this linearization should produce a constant instantaneous frequency or a constant group delay. The time warping function $w(t) = t^{1/k}$ is designed to normalize to one the polynomial order of a chirp signal with phase function that is a monomial power polynomial, $e^{j2\pi c t^k}$. In the case of more complicated signals of the form:

$$x(t) = \exp(j2\pi(\sum_{i=1}^N c_i t^{k_i})); \quad (5.18)$$

the purpose of the warping function design is to normalize to one the polynomial highest order in an effort to linearize the time-frequency behavior of the highest order power of “ t ” in the signal phase (as seen in Figure 5.3). However, the effect on the other powers of “ t ” in the phase can appear as non-linear energetic terms in the TF plane. Example 7 explains more about this problem.

Example 7: Figure 5.16 shows the example of a signal containing two power phase terms expressed as:

$$x_2(t) = e^{j2\pi(t^{0.1} + 0.85 t^{0.7})}. \quad (5.19)$$

with $k_1 = 0.1$, $k_2 = 0.7$, $c_1 = 1$ and $c_2 = 0.85$. Note that the powers do not have to be integers, as in the traditional polynomial phase estimation problems [130]. Figure 5.15-A shows the WVD of the original signal $x_2(t)$ (equation 5.19) and Figure 5.15-B

shows the corresponding instantaneous frequency curve.

$IF_{x_2}(t) \propto 0.1t^{-0.9} + 0.595t^{-0.3}$ To warp the signal, the time-warping function

associated with the monomial of highest order, $t^{0.7}$, is as follows: $w(t) = t^{1/0.7}$.

Figure 5.15 Example 7: two power phase terms $x_2(t) = e^{j2\pi(t^{0.1} + 0.85t^{0.7})}$. (A) the WVD of the original signal $x_2(t)$ in equation 5.19. (B) The instantaneous frequency curve $IF_{x_2}(t) \propto 0.1t^{-0.9} + 0.595t^{-0.3}$. (C) The effect of warping on the WVD using the power time-warping function $w(t) = t^{1/k}$ for $k=0.7$. (D) The FT spectrum of the warped signal with peak equal to 2.128.

Figure 5.15-C shows that the application of the warping operator for this signal produces an imperfect linearization near the beginning and the end of the signal. Some artifacts of the time-frequency behavior of the warped signal are due to the existence of another power term $t^{0.1}$ in the phase. Figure 5.15-D shows the FT spectrum with a peak equal to 2.128.

To overcome this limitation, we propose a new warping based method by defining the new warping function as follows:

$$\hat{w}(t) = e^{at} t^{1/k}, \quad a, k \in \mathbb{R} \text{ and } t \in \mathbb{R}^+ \quad (5.20)$$

with “ k ” to equal the power of the phase term and the “ a ” parameter used to “tune” the new warping function. For “ $a=0$ ”, $\hat{w}(t)$ simplifies to the power class time warping function $w(t) = t^{1/k}$. The corresponding derivative of the warping operator

$\dot{\hat{w}}(t) = \frac{d}{dt} \hat{w}(t)$ is:

$$\dot{\hat{w}}(t) = e^{at} t^{\frac{1}{k}} \left[a + \frac{1}{kt} \right]. \quad (5.21)$$

The new warping function is composed of the product of two functions: e^{at} and $t^{1/k}$ with parameters “ a ” and “ k ”, respectively. The function $t^{1/k}$ is the same power class time-warping function used in [105, 106, 108] to linearize the behavior of the time-frequency representation. The function e^{at} is the same k^{th} exponential function class time warping used in [109] with $a=k$.

Here, the inverse time axis warping function $\hat{w}^{-1}(t)$ defined by $\hat{w}(\hat{w}^{-1}(t)) = t$ is given by:

$$\hat{w}^{-1}(t) = \frac{1}{ak} L_w(akt^k). \quad (5.22)$$

were, $L_w(akt^k)$ is the Lambert W function [60]. The Lambert W function (L_w) is defined to be the multi-valued inverse of the function $t \rightarrow te^t$ [60]. Thus, an algorithm implementing the time-warping using the new warping function $\hat{w}(t)$ (in equation (5.20)) can be used to implement the inverse transform or the unwarping described in equation (5.22).

The results of applying the new warping function, $\hat{w}(t)$ show a good linearization of a chirp in TF plane especially when the chirp phase is a monomial or has two power terms. To solve the problems which appear when the signal has multiple powers of time “ t ” in the phase (Figure 5.15-C), we incorporate the function e^{at} into our new warping operator. The motivation for the new warping function came from examining the results produced by warping using one more degree of freedom, the parameter “ a ”, used to “tune” the new warping function to provide an optimal linearization for the given signal. Let’s suppose, for the purposes of illustration, that we somehow know what is the best “ a ” for the next few examples. The optimization of the parameter “ a ” will be described clearly in section 5.5.2. In Figure 5.16 we assume the parameter “ a ” is already optimized and it is equal to $a=5.321e-5$.

Figure 5.16 Example 7 (continued) **(A)** The effect of warping on the WVD using the newly proposed warping function $\hat{w}(t) = e^{at} t^{1/k}$ with $k=0.7$ and $a=5.321e-5$. **(B)** The FT spectrum of the warped signal with a peak equal to 2.729.

Figure 5.16-A shows the result of warping-based time-frequency representation using the new time-warping function in equation (5.20). Comparing Figure 5.15-C and Figure 5.16-A, the linearization has improved using the new warping operator with reduced inner-interference terms compared to the linearization in Figure 5.15-C. Moreover, comparing Figure 5.15-D and 5.16-B the FT spectrum of the newly warped signal has a higher maximum peak which equals to 2.729. The performance of the new warping operator in term, as relates to energy concentration and the presence of inner-interference terms, will be tested later in this chapter. More examples (Examples 8 and 9) will be provided to test the new proposed warping function.

Example 8: The signal $x_3(t) = e^{j2\pi(c_1 t + c_2 t^2 + c_3 t^3)}$, with $c_1 = 0.25$, $c_2 = -9.76e-4$ and $c_3 = 3.204e-6$, is a quadratic FM chirp signal (the phase power parameter $k=3$ is set to the highest power of t in the phase function) with an instantaneous frequency (IF),

$IF_{x_3}(t) = c_1 + 2c_2t + 3c_3t^2$, equal to 0.25 Hz at $t=0$. The TF support of the chirp crosses 0.15 Hz at $t=100$ seconds as shown in Figure 5.17-B, the corresponding WVD is shown in Figure 5.17-A. The sampling frequency is $F_s=1$ Hz, and the number of frequency bins $N=256$. Figure 5.17-C shows the effect of the original power warping using the function $w(t) = t^{1/k}$, for $k=3$, and Figure 5.17-E shows the effect of warping using the new time-warping function $\hat{w}(t) = e^{at^{1/k}}$, with $a = 33.333e-4$ and $k=3$. Comparing the two figures, 5.17-C and E, we can observe that the new time-warping function gives better results, e.g. the TF support is more linear near the origin. These improvements are also confirmed by plotting the FT spectrum, Figures 5.17-D and 5.17-F, respectively, with higher peak spectrum resulting from using the new warping function.

Example 9: The signal $x_4(t) = e^{j2\pi(c_1t + c_2t^2 + c_3t^3)}$, with $k=3$, $c_1 = 0.4558$, $c_2 = -0.0058$ and $c_3 = 2.2761e-5$, is a concave quadratic chirp signal with symmetric instantaneous frequency about midway along the time axis (Figure 5.18-B). The corresponding WVD is shown in Figure 5.18-A. The minimum frequency is 0.08 Hz and the maximum frequency is 0.45 Hz with a sampling frequency $F_s=1$ Hz and the number of frequency bins $N=256$. Notice the oscillatory inner-interference terms inside the concave TF support in Figure 5.18-A. Figure 5.18-C shows the effect of the warping using the power time-warping function $w(t) = t^{1/k}$, for $k=3$, and Figure 5.18-E shows the effect of warping using the new warping function $\hat{w}(t) = e^{at^{1/k}}$, with $k=3$ and $a=0.0033$.

Figure 5.17 Example 8: $x_3(t) = e^{j2\pi(c_1 t + c_2 t^2 + c_3 t^3)}$, with $c_1 = 0.25$, $c_2 = -9.76e-4$ and $c_3 = 3.204e-6$. (A) The WVD of the original signal $x_3(t)$. (B) Instantaneous frequency curve. (C) The power warped TFR. (D) FT spectrum obtained using the power warping function $w(t) = t^{1/k}$. (E) WVD of the new time warping function $\hat{w}(t) = e^{at^{1/k}}$. (F) The corresponding FT spectrum of new warping function. with $k=3$ and $a=33.33e-4$.

Figure 5.18 Example 9: $x_4(t) = e^{j2\pi(c_1t+c_2t^2+c_3t^k)}$, with $c_1 = 0.4558$, $c_2 = -0.0058$ and $c_3 = 2.2761e-5$ and $k=3$. (A) The WVD of the original signal $x_3(t)$. (B) Instantaneous frequency curve. (C) The time-warping using the power warping function $w(t) = t^{1/k}$. (D) The corresponding FT spectrum. (E) Time-warping using the new warping function $\hat{w}(t) = e^{at} t^{1/k}$. (F) The corresponding FT spectrum. $k=3$ and $a=0.0033$.

Comparing the two figures, we can observe that the new warping gives better linearization, particularly near the origin, compared to the power warping function $w(t) = t^{1/k}$ (Figure 5.18-C). Comparing Figures 5.18-D and 5.18-F, we can observe that the peak of the FT spectrum of the newly warping function is larger.

Metrics for comparing the performances of the two warping function will be discussed in the next section (section 5.5.2).

5.5.2 Performance analysis and parameter optimization

It is known that an ideal TFR has few inner-interference terms or cross-terms, high energy concentration and good TF resolution and readability [40, 123]. One ad hoc way to evaluate cross-term suppression and energy concentration in the TF plane is by visual inspection. Alternatively, to estimate the optimum warping parameter “ a ” and to compare the performance of linearization (or warping-based TFRs) between the two warping functions $w(t) = t^{1/k}$ and $\hat{w}(t) = e^{at} t^{1/k}$, we propose using numerical metrics such as the Rényi entropy [74, 114], the ratio of norm-based measurement [121] and the maximum value of the warped signal spectrum.

5.5.2.1 Rényi entropy of a TFR

Entropy is a quantative measure proposed for computing the information content of a signal [74]. By the analogy to probability distributions, minimizing the complexity or information in a particular TFR is similar to maximizing the TFR concentration, or peakiness and often, readability [120]. The Rényi entropies make excellent measures of the information extraction performance of TFRs [74,113,114,120]. The Rényi entropy of a TFR ($TFR_x(t, f)$) is defined as [74, 114]:

$$Entropy_{R\acute{e}nyi}(\sigma) = \frac{1}{1-\sigma} \log_2 \left(\sum_t \sum_f TFR_x^\sigma(t, f) \right), \quad (5.23)$$

where σ is the order of entropy. Since the TFR cross-term oscillatory structure cancels under the integration with odd powers, “ σ ” should be an odd integer value for a general quadratic TFD [58]. In our case the value of the order σ was chosen to be $\sigma = 3$ [113, 58]. In the case of TFRs, high entropy indicates that the energy concentration of an auto-component is low and vice versa, low entropy indicates that the energy concentration of an auto-component is high. In other words, ***a highly concentrated TFR reduces the Rényi entropy.***

5.5.2.2 Ratio of Norm

An efficient quantitative criterion to evaluate the performance of any TFR can be obtained by measuring the energy concentration. The Ratio of Norm (RN) based measures proposed by Jones and Parks [121] was used to discriminate a low resolution TFR from a high resolution TFR. The RN was used to adapt the parameters of an analysis window to maximize the measure of concentration created by dividing the fourth power norm of a TFR by its second power norm. RN is given by [121]:

$$\text{Ratio of Norm (RN)} = \frac{\sum_t \sum_f |TFR_x(t, f)|^4}{\left(\sum_t \sum_f |TFR_x(t, f)|^2 \right)^2}. \quad (5.24)$$

Unlike Rényi entropy, ***higher values of RN indicate that the TFR has better energy concentration and vice versa.***

5.5.2.3 Estimation of optimum warping parameter “a”

The parameter “a” is used to tune the new warping function $\hat{w}(t) = e^{at} t^{1/k}$ to provide an optimal TF linearization for a given signal. To find the optimum value of the parameter “a”, a one dimensional search was done based upon quantitative measures of concentration. In the search algorithm, warping was computed for different values of “a” along a grid and the value that yields the maximum peak value of the FT spectrum, along with the maximum value of the Ratio of Norm and the minimum value of the Rényi entropy was selected. Using the selected optimum value of “a”, it is proposed that the warped signal will have higher energy concentration (higher Ratio of Norm) with minimal inner interference terms (minimum Rényi entropy value) as compared to time-warping using the power function.

To show how the process for choosing the optimum value for the parameter “a” is done, the optimum parameter “a” for the signal $x_2(t) = e^{j2\pi(t^{0.1} + 0.85t^{0.7})}$ will be selected (from Example 7). Notice that the phase is not a monomial, nor are the powers equal to integers as is typically required for polynomial phase estimation algorithms. For $k=0.7$, we tested 10^6 values of “a” on an interval $a \in (0 \ 1)$, using a fine spacing to get a good estimate. If “a=0” we will have the same effect as the power warping function $w(t) = t^{1/k}$. Table 5.3 shows some selected values of “a” of the newly warping function with the corresponding values of the FT spectrum maximum peak (Mp) of the warped signal, the Rényi entropy (RE) and the Ratio of Norm (RN). For this example, it was found that the optimum value of the parameter “a” that has the maximum amplitude of the FT spectrum, the minimum value of the

Table 5.3 Values of the Rényi entropy (RE), the Ratio of Norm (RN) and the FT maximum peak (Mp) occurring during the search for the optimum value of the new time-warping function parameter “ a ”. The signal analyzed was $x_2(t) = e^{j2\pi(t^{0.1}+0.85t^{0.7})}$ with $k = 0.7$ in Example 7. Optimal results are highlighted in yellow row.

The parameter “ a ”	The WVD of the warped signal	Renyi Entropy (Small value is better)	Ratio of Norm (higher value is better)	Max FT(warped(sig)) (higher value is better)
$a = 0.011$		RE=0.1800	RN=0.0032	Mp=0.8725
$a=6.385e-4$		RE=-0.0035	RN=0.0051	Mp=1.1291
$a=1.915e-4$		RE=-0.0113	RN=0.0053	Mp=1.8967
$a=6.385e-5$		RE=-0.0469	RN=0.0058	Mp=2.7186
$a=5.321e-5$		RE=-0.0512	RN=0.0058	Mp=2.7292
$a=5.108e-5$		RE=-0.0431	RN=0.0056	Mp=2.7186
$a=4.257e-5$		RE=0.0016	RN=0.0048	Mp=2.6341

Rényi entropy and the maximum value of Ratio of Norm was equal to $a=5.321e-5$. The same process will be used to select the optimum value of parameter “ a ” for the other examples discussed in this chapter (Examples 6, 8 and 9). These values are shown in Table 5.4.

5.5.2.4 Performance comparison

To compare the performance of the new proposed warping function to a traditional warping function, the Rényi entropy, the Ratio of Norm and the maximum amplitude of the FT of the warped signal were calculated for several signals. The results are shown in Table 5.4.

Table 5.4 shows that the new warping function, $\hat{w}(t) = e^{at} t^{1/k}$, produces the same result as the power warping function, $w(t) = t^{1/k}$, when the signal has only a single power term in the phase function, i.e. monomial phase. For example, examine the results for the case of the signal $x_1(t) = e^{j2\pi 0.014 t^{1.5}}$ with $k=1.5$ (already discussed in in Example 6); the table entries are identical for $w(t)$ and $\hat{w}(t)$, where the new $\hat{w}(t) = e^{at} t^{1/k}$ simplifies when “ $a=0$ ” to the warping function $w(t) = t^{1/k}$. Moreover, the values in the subsequent rows show that using the new warping function optimized for the parameter “ a ” gives smaller values for the Rényi entropy (which means reduced inner interference terms) and larger values for Ratio of Norm (which means higher energy concentration) for the case of signals with more than one power term in the phase function; this in turn means better linearization.

Discussing the other advantages of the new proposed time-warping function, we can say that the new warping function offers one more degree of freedom, the parameter “ a ”, which helps to improve the linearization. Moreover, the new warping function can linearize with better performance many types of signals whether the order of the power term in the phase function is either integer or real, which is not the case of the polynomial Wigner-Ville distribution (PWVD) method [112, 123]. The PWVD estimated the instantaneous frequency (IF) of nonlinear FM signals that had polynomial phase where the order of the nonlinearity of the PWVD should be an even integer.

Table 5.4 Performance analysis using two different warping functions. $w(t) = t^{1/k}$ is the power warping function and $\hat{w}(t) = e^{at} t^{1/k}$ is the new warping function.

The signal $x_i(t)$	Warping function	Renyi Entropy (Small value is better)	Ratio of Norm (higher value is better)	Max(FT(warped(sig))) (higher value is better)
$x_1(t) = e^{j2\pi c_1 t^{1.5}}$ $c=0.014, k=1.5,$ $a=0.$	$w(t) = t^{1/k}$	0.0163	0.0024	0.7232
	$\hat{w}(t) = e^{at} t^{1/k}$	0.0163	0.0024	0.7232
$x_2(t) = e^{j2\pi(c_1 t^{0.1} + c_2 t^{0.7})}$ $c_1=1, c_2=0.85,$ $k=0.7, a=5.321e-5$	$w(t) = t^{1/k}$	-0.0123	0.0051	2.128
	$\hat{w}(t) = e^{at} t^{1/k}$	-0.0512	0.0058	2.7292
$x_3(t) = e^{j2\pi(c_1 t + c_2 t^2 + c_3 t^3)}$ $c_1=0.25, c_2=-9.76e-4,$ $c_3=3.204e-6,$ $k=3, a=33.33e-4$	$w(t) = t^{1/k}$	0.1949	0.0015	0.1076
	$\hat{w}(t) = e^{at} t^{1/k}$	0.1069	0.0022	0.2190
$x_3(t) = e^{j2\pi(c_1 t + c_2 t^2 + c_3 t^3)}$ $c_1=0.4558,$ $c_2=-0.0058,$ $c_3=2.2761e-5,$ $k=3, a=0.0033$	$w(t) = t^{1/k}$	0.6436	6.4633e-4	0.0714
	$\hat{w}(t) = e^{at} t^{1/k}$	0.3448	0.0012	0.1959

In practical applications, the use of the new warping function requires setting up a search for the parameter “ a ”. In this chapter, we assumed that the order of the signal “ k ” was known.

5.6 Warping-based multi-component non-linear FM chirp signal decomposition

The principle of the new algorithm used to decompose a multi-component signal composed of the sum of non-linear FM chirp signals is based on the concept of warping. In this case, time warping was used to linearize the time-frequency behavior of the non-linear FM chirp components. The result will be a FT spectrum with larger, more concentrated peak amplitudes and a more linear time-frequency structure. After that, a rectangular window will be applied to filter out the peak in the spectrum with the largest amplitude. An inverse transform will be applied. The inverse time warping in (equation (5.22)) will be applied to the filtered signal to reconstruct the time-frequency structure of the decomposed component. The algorithm will be iterated until each component is isolated. The next examples will illustrate how the warping-based multi-component signal decomposition algorithm works.

Figure 5.19 Block diagram for decomposition of multi-component non-linear FM chirp signals using newly proposed time warping-based time-frequency representation algorithm.

Example 10: In this example we will try to solve the problem caused by overlapping or crossing TF support discussed in example 5 by applying the warping-based signal decomposition. The signal is composed of two non-linear chirp components overlapping in time and in frequency. The first component is $x_1(t) = e^{j2\pi 0.014t^{1.5}}$ and the second signal is $x_2(t) = e^{j2\pi(0.4t - 0.05t^{1.3})}$. Figures 5.20-A and 5.20-B show the WVD and the FT spectrum, respectively of the multi-component signal $x(t) = x_1(t) + x_2(t)$. We already discussed in Example 5 how the FrFT-based decomposition algorithm failed to decompose this signal.

Figure 5.20 Example 10: Non-linear FM chirp signal composed of two components. (A) The WVD of the multi-component signal. (B) The FT spectrum.

Applying the new warping-based signal decomposition algorithm, shown in Figure 5.19, the different steps to extract the two components $x_1(t)$ and $x_2(t)$ are explained in Figure 5.21. In the first step, the signal $x(t)$ is warped considering the characteristic of the first chirp component $x_1(t)$, which means select the signal order $k = k_1 = 1.5$

and find the optimum value of the warping function parameter $a = a_1 = 0$. The result is a spectrum with larger amplitude. This corresponds to the FT seen in Figure 5.21-A. A rectangular window of six bins was used to filter the spectrum with the largest amplitude, the result is shown in Figure 5.21-B. The WVD of the extracted component $\hat{x}_1(t)$ after applying the inverse transform is shown in Figure 5.21-C.

Figure 5.21 Example 10 (continued): Decomposition of multi-component signal composed of two non-linear chirp components. (A) The FT spectrum of the warped signal $x(t)$ with $k = k_1 = 1.5$ and $a = a_1 = 0$. (B) The FT spectrum after filtering the spectrum peak with larger amplitude. (C) The WVD of the reconstructed chirp signal $\hat{x}_1(t)$. (D) The FT spectrum of the warped signal $x(t)$ with $k = k_2 = 1.3$ and $a = a_2 = 3.1e - 3$. (E) The FT spectrum after filtering the spectrum peak with larger amplitude. (F) The WVD of the reconstructed chirp signal $\hat{x}_2(t)$.

The same process was applied to decompose the second chirp component $x_2(t)$. This time the signal $x(t)$ is warped considering the characteristic of the second chirp signal $x_2(t)$, which means select the signal order $k = k_2 = 1.3$ and find the optimum value of the warping function $a = a_2 = 3.1e - 3$. The result is a spectrum with larger amplitude corresponding to the second signal as seen in Figure 5.21-D. A rectangular window of four bins was used to filter the spectrum peak with the largest amplitude; the result is shown in Figure 5.21-E. The WVD of the new reconstructed signal $\hat{x}_2(t)$ is plotted after applying the inverse transform as shown in Figure 5.21-F.

The WVD plots in Figures 5.21-C and F have reduced inner-interference terms with good resolution considering that the two components $x_1(t)$ and $x_2(t)$ have non-linear instantaneous frequency which is an advantage of the warping-based decomposition algorithm. Comparing Figures 5.21-C and F to Figures 5.13-E and H, we see that the new algorithm does a better job than the FrFT decomposition algorithm. To compare the reconstructed chirp components to the originals, we plot the true instantaneous frequency (IF) of the two chirp components which is defined as $IF_{x_i}(t) = \frac{1}{2\pi} \frac{d}{dt}(\phi_i(t))$, with $\phi_i(t)$ is the phase of the component $x_i(t)$. For the signal $x_1(t) = e^{j2\pi 0.014 t^{1.5}}$, the IF law is: $IF_{x_1}(t) = 0.0210t^{1.5}$; and for the signal $x_2(t) = e^{j2\pi(0.4t - 0.05t^{1.3})}$, the IF law is: $IF_{x_2}(t) = 0.4 - 0.065t^{1.3}$.

The estimated IF is calculated as: $\hat{IF}_{\hat{x}_i}(t) = \arg\{\max_f(WVD_{\hat{x}_i}(t, f))\}$ where $\hat{x}_i(t)$ is the reconstructed signal with $i=1,2$. Figures 5.22-A and B show the corresponding plots of the true and the estimated instantaneous frequency for the two chirp signals $x_1(t)$ and

$x_2(t)$, respectively. To measure the estimation quality, the mean squared error (MSE) [129] was calculated between the true and the estimated IF. The calculated values are $MSE_1=4.2925e-6$ and $MSE_2=2.7254e-5$ for the first and the second chirp signals, respectively.

Example 11: In this example, the multi-component signal is composed of three non-

Figure 5.22 Comparison between the true IF and the estimated IF of the two chirp components $x_1(t)$ and $x_2(t)$ in subplots A and B, respectively.

linear chirp signals that overlap in time and in frequency. The first component is

$x_A(t) = e^{j2\pi 0.0208 t^{1.5}}$, the second component is $x_B(t) = e^{j2\pi(0.4t-0.05t^{1.3})}$ and the third

component is $x_C(t) = e^{j2\pi(c_1t+c_2t^2+c_3t^3)}$, with $c_1 = 0.001$, $c_2 = -1.176e-5$ and

$c_3 = 7.66e-6$. Figures 5.23-A and 5.23-B show the WVD and the FT spectrum,

respectively of the multi-component signal $x(t) = x_A(t) + x_B(t) + x_C(t)$. We can

observe the overlapping of the FT spectrum of the three non-linear chirp components

comprising the signal $x(t)$; which means that we cannot apply the FrFT-based signal

decomposition algorithm.

Figure 5.23 Example 11: Non-linear FM chirp signal composed of three components. (A) The WVD of the multi-component signal. (B) The corresponding FT spectrum.

Applying the warping-based signal decomposition algorithm, the different steps to extract the three components $x_A(t)$, $x_B(t)$ and $x_C(t)$ are explained in Figure 5.24. In the first step, the multi-component signal $x(t)$ is warped considering the characteristic of the first chirp component $x_A(t)$, which means select the signal order $k = k_A = 1.5$ and the optimum value of the warping function parameter $a = a_A = 0$. The result is a FT spectrum with largest peak amplitude corresponding to the component $x_A(t)$ since it matches the order $k = k_A = 1.5$ and the optimum value of the warping function “ a ”, as seen in Figure 5.24-A. Since $x_A(t)$ has monomial phase, it is important to note that the search algorithm found $a=0$, reducing the new warping operator $\hat{w}(t) = e^{at} t^{1/k}$ to the traditional power operator $w(t) = t^{1/k}$. A rectangular window of six bins was used to filter the peak of the spectrum with the highest

amplitude as shown in Figure 5.24-B. The WVD of the extracted estimate $\hat{x}_A(t)$ is plotted after applying the inverse transform as shown in Figure 5.24-C.

The same process was applied to decompose the second chirp component $x_B(t)$. This time the signal $x(t)$ is warped considering the characteristic of the second chirp signal $x_B(t)$, which means select the signal order $k = k_B = 1.3$ and have the search algorithm find the optimum value of the warping function parameter $a = a_B = 3.1e - 3$. The result is a spectrum with larger amplitude corresponding the second signal $x_B(t)$ since it matches the order $k = k_B = 1.3$ and the optimum parameter “ a ” of the warping function as seen in Figure 5.24-D. A rectangular window of eleven bins was used to filter the peak spectrum with the highest amplitude as shown in Figure 5.22-E. The WVD of the extracted estimate $\hat{x}_B(t)$ is plotted after applying the inverse transform and is shown in Figure 5.24-F.

To decompose the third component $x_C(t)$, the signal $x(t)$ is warped considering the characteristic of the third chirp component $x_C(t)$, with the signal order $k = k_C = 3$ and the optimum value of the warping function parameter $a = a_C = 6.7e - 3$. The result is a spectrum with larger amplitude corresponding to the component $x_C(t)$ as seen in Figure 5.24-G. A rectangular window of three bins was used to mask out the peak spectrum with the highest amplitude; the result is shown in Figure 5.24-H. The WVD of the extracted component $\hat{x}_C(t)$ is plotted after applying the inverse transform as shown in Figure 5.24-I.

Figure 5.24 Example 11 (continued): Decomposition of multi-component signal composed of three non-linear chirp components. (A) The FT spectrum of the warped signal $x(t)$ with $k = k_A = 1.5$ and optimized $a = a_A = 0$. (B) The FT spectrum after filtering the peak spectrum with larger amplitude. (C) The WVD of the reconstructed chirp component $\hat{x}_A(t)$. (D) The FT spectrum of the warped signal $x(t)$ with $k = k_B = 1.3$ and optimized $a = a_B = 3.1e - 3$. (E) The FT spectrum after filtering the peak spectrum with larger amplitude. (F) The WVD of the reconstructed chirp component $\hat{x}_B(t)$. (G) The FT spectrum of the warped signal $x(t)$ with $k = k_C = 3$ and optimized $a = a_C = 6.7e - 3$. (H) The FT spectrum after filtering the peak spectrum with larger amplitude. (I) The WVD of the third reconstructed chirp component $\hat{x}_C(t)$.

To compare the reconstructed chirp components to the original, we plot the true instantaneous frequency (IF) of the three chirp components which is defined as $IF_{x_i}(t) = \frac{1}{2\pi} \frac{d}{dt}(\phi_i(t))$, where $\phi_i(t)$ is the phase of the component $x_i(t)$. For the component $x_A(t) = e^{j2\pi 0.0208 t^{1.5}}$, the IF law is: $IF_{x_A}(t) = 0.0312 t^{0.5}$; For the component $x_B(t) = e^{j2\pi(0.4t - 0.05t^{1.3})}$, the IF law is: $IF_{x_B}(t) = 0.4 - 0.065t^{0.3}$, For the component $x_C(t) = e^{j2\pi(c_1t + c_2t^2 + c_3t^3)}$, the IF law is: $IF_{x_C}(t) = c_1 + 2c_2t + 3c_3t^2$, with $c_1 = 0.001$, $c_2 = -1.176e-5$ and $c_3 = 7.66e-6$.

The estimated IF is calculated as: $\hat{IF}_{x_i}(t) = \arg\{\max_f(WVD_{x_i}(t, f))\}$ where $\hat{x}_i(t)$ is the reconstructed component with $i=A, B, C$. Figures 5.25-A, B and C show the corresponding plots of the true and the estimated instantaneous frequency for the three chirp components $x_A(t)$, $x_B(t)$ and $x_C(t)$, respectively, and their estimated values. The mean squared error (MSE) [129] was calculated between the true and the estimated IF. The calculated values are $MSE_A = 8.542e-6$, $MSE_B = 4.837e-5$ and $MSE_C = 7.188e-6$ for the first, the second and the third chirp components, respectively.

Figure 5.25 Comparison between the true instantaneous frequency (True IF) and the estimated IF of the three chirp components $x_A(t)$, $x_B(t)$ and $x_C(t)$ in subplots A, B, and C, respectively.

5.6 Future works

The new proposed warping function compared to the power warping function improved linearizing signals with polynomial phase. However, in some cases this linearization is imperfect as seen in Example 9. We also showed how the readability of the unwarped signals is improved with much reduced inner-interferences as seen in Examples 10 and 11. Nevertheless, we mention that this new algorithm requires the

knowledge of the signal power order “ k ” which characterizes the signals being analyzed.

From Examples 10 and 11, we have shown that in each step in which we want to decompose a component from the multi-component signal, we need to match its highest power order “ k ” and the corresponding warping parameter “ a ”. This means that there is a relation between the optimization of the time warping and the signal order. Some works have already been done in the literature using the warping function $w(t) = t^{1/k}$ with unknown k [86]. In their work, the authors used an unsupervised approach based on the Hilbert transform to estimate a non-linear monomial power warping function which characterizes the time frequency content of the signal [86]. Since we already demonstrated with some examples in this work the dependence of the warping function to the order of the signal, we propose in the future to develop new techniques which help to estimate the parameters such as the order of non-linear chirp signals using the warping principle.

In the next example, we will explain some preliminary results to show that the new warping can be a useful technique to estimate the signal order. We already discussed in this chapter that the purpose of the warping the signal was to linearize the time-frequency structure of this signal. This leads to a Fourier spectrum with larger amplitude and a narrow peak compared to other non-linear signals. In the next example, we consider the signal $x(t) = e^{j2\pi(c_1 t + c_2 t^2 + c_3 t^3)}$, with $c_1 = 0.25$, $c_2 = -9.76e-4$, $c_3 = 3.204e-6$ and $k=3$. The WVD plot is shown in Figure 5.26-A, Figure 5.26-B shows the effect of the new warping function $\hat{w}(t) = e^{at^{1/k}}$ with $a = 33.33e-4$ found by the proposed search algorithm with a prior knowledge that $k=3$. In this example, we

will select different values of the order $k \in [0 \ 3.5]$, and plot the corresponding spectrum of the warped signal for different values of k . Figures 5.26-C and D show how the spectrum peak of the warped signal become more sharp each time the order of the signal approaches the true value, $k=3$. From Figure 5.26-E, one can observe that the plot is concentrated vary between $k = [2.9 \ 3.1]$ represented by a black ellipse in the Figure. This is very near the correct answer of $k=3$. These preliminary results hint that the warping principle may be useful to estimate the highest power parameter of non-linear signals.

Figure 5.26 (A) the WVD of the signal $x(t) = e^{j2\pi(c_1 t + c_2 t^2 + c_3 t^3)}$, with $c_1 = 0.25$, $c_2 = -9.76e-4$, $c_3 = 3.204e-6$ and $k=3$. (B) The effect of the new warping function $\hat{w}(t) = e^{at} t^{1/k}$ with $a = 33.33e-4$. (C) Spectrum of the warped signal for several values of k . (D) Zooming of subplot C. (E) The spectrum of the warped signal for several values of k along the vertical axis.

5.8 Summary

In this chapter, a review of how the FrFT has been used to decompose multi-component FM chirp signals was provided. However, FrFT-based multi-component signal decomposition fails to decompose non-linear overlapping FM chirp signals. To overcome this problem, we proposed a new algorithm based on time warped Time-Frequency Representations. In this chapter, a new warping function $\hat{w}(t) = e^{at} t^{1/k}$ was used to linearize the time-frequency structure of the non-linear FM chirp signal to help extract non-linear FM chirp signals whether or not they overlapped. The inverse of the new warping function was found to be $\hat{w}^{-1}(t) = \frac{1}{ak} L_w(akt^k)$, where “ L_w ” is the Lambert- W function [131]. A search algorithm based upon the metrics of Rényi Entropy, Ratio of Norm and the maximum FT magnitude was proposed to select the optimal value of the parameter “ a ”.

A comparison of the new warping function to the previously proposed power warping function $w(t) = t^{1/k}$ was provided. The new warping function shows a good improvement in linearizing non-linear FM chirp signals, even in the cases when the order of the phase function is non-integer or the signal phase is not a monomial (single power term “ t^k ” in the phase function). It also reduces inner-interference terms in TFRs. Several examples were given demonstrating the efficacy of the algorithm and comparing it to the traditional power warped TFR as well as k^{th} power exponential.

It was discussed earlier in this chapter that the length of the rectangular window used to filter the peak spectrum with higher amplitude was set depending on the minimum MSE calculated between the original signal and the retrieved signal. It was

also assumed that the user knows a priori the number of signal components. However, when one does not know any information about the original signal, this method of determining the length of the window is not practical. So in our future work we will try to automate the selection of this length by finding a relationship between the type of the chirp signal and the width of the window used. Future research also includes applying rank estimation techniques to determine the number of signal components, as well as algorithms to estimate the power parameter, k .

CHAPTER 6

CONCLUSION AND FUTURE PERSPECTIVES

6.1 Conclusion

Signal Processing is a topic that was used long time ago and still is used in almost all areas including biomedical and electrical domains. One focus of this dissertation was to develop and test new algorithms for epileptic seizure detection and compare the efficacy of the algorithms in data recorded using conventional disc electrodes (EEG) versus novel tripolar concentric ring disc electrodes (tEEG). These algorithms involved different signal processing techniques which showed that the proposed algorithms are competitive to those in the literature and some of them outperform those techniques.

The design process of the developed algorithms involves the determination of discriminating boundaries between seizure and non-seizure EEG data by parametric representation of seizure EEG in classification algorithms. In this thesis, the first proposed algorithm used time-domain features, such as which were extracted to characterize the seizure from non-seizure segments. The algorithm was evaluated using human and rat data. The results showed the robustness of time-domain features; high accuracy was achieved using the rats' dataset even though the data were highly contaminated with motion artifacts.

It has been shown that entropy can be used as a complexity measure for analyzing biological signals such as brain signals or EEG because the brain is a

complicated dynamical system with uncertainties. We have developed a two stage seizure detection technique to classify seizure from non-seizure segments. In the first stage, a signal was decomposed into different frequency sub-bands using either the Empirical Mode Decomposition (EMD) method or the discrete dyadic Wavelet transform (DWT). In the second stage, the Rényi entropy was computed and input into a classifier to discriminate between seizure and non-seizure activities. According to the high accuracy obtained, Rényi entropy combined with EMD or DWT is effective and applicable for discriminating Seizures from Normal EEG recordings.

Because of the non-stationary behavior of EEG signal during seizure activities, time-frequency signal analysis techniques were used. These techniques use joint time and frequency domain representations of a signal to yield more revealing pictures of the temporal localization of the signal's time-varying spectral component. Our proposed algorithm was based on the automated recognition of changes in the signal's energy in different decomposition levels. For this purpose the EMD algorithm was used to decompose the signal into sublevels, called Intrinsic Mode Functions (IMFs), each with different frequency support. It was shown that the result after applying the TF process outperformed those using just time-domain analysis for seizure detection especially after using Flandrin's method to select the IMFs with the most important information.

The main limitations of our proposed algorithms used for seizure detection can be summarized as follow:

- (1) The datasets used to prove the performance of the proposed algorithms for seizure detection. For example it was mentioned in chapter 3 that the rats' dataset was

minimized into just seizure and non-seizure segments because of the grooming and moving of rats which infect the data by noise. (2) The proposed algorithms for seizure detection were based on the selection of several parameters, such as the window length used to extract the features, the window length of the Smoothed-Pseudo Wigner-Ville distribution, the value of k of the k -NN classifier and others. All of these parameters must be automatically updated as new data arrive in real time processing.

The focus of this dissertation was not limited only on the development of new algorithms for seizure detection, but also to explore and develop new methods for signal processing such as decomposition of multi-component signals. For this purpose, a review was provided of the Fractional Fourier Transform (FrFT), which is a generalization of Fourier transform, as a flexible approach for the time-frequency analysis of multi-component signals. It is known that fractional Fourier transform can be successfully used to decompose multi-component signals composed of linear FM chirp signals. However, the FrFT-based decomposition fails to decompose multi-component signals composed of non-linear FM chirp signals, especially if these signals overlap in time and in frequency.

Signals encountered in real applications often involve frequency modulation (FM) chirp components which have non-linear instantaneous frequency. Some transformations such as time/frequency warping were used to change non-linear time-frequency structures into linear ones. The mathematical interpretation of time warping is just the operation of replacing the time dependence of a signal with a warping function. The results will be a linearization of the time-frequency structure of the warped signal wherever the signal phase is matched to the warping. In this thesis a

new time-warping function, $\hat{w}(t) = e^{at^{1/k}}$, based time-frequency representation was proposed. The performance analysis proves that the new warping function can be successfully used to linearize the TF structure of non-linear FM chirp signal. To overcome the limitations of FrFT-based decomposition algorithm, a new algorithm was proposed based on the new time warping operator which was used to linearize the time-frequency behavior of the signal. This process made the decomposition easier.

6.2 Future works

In this study, the algorithms proposed for seizure detection were focused on an offline analysis of EEG signals recorded during seizure events. Prediction and detection of seizures from an online recorded EEG can significantly improve the life of patients with seizure by preventing the negative influence of uncontrolled seizures. Hence, an automated seizure warning algorithm is desired that can analyze the EEG signals directly from patients with the objective of predicting the appearance of seizure during real time recording.

To analyze non-linear chirp signals, the FrFT equation for linear chirp signals has been adopted. In this case the FrFT spectrum of non-linear chirps were spread out which means more bins have to be retained on both sides of the FrFT peak. Future works will consider how to adapt the FrFT for non-linear chirp signals.

It was discussed in chapter 5 that the length of the rectangular window used to mask out the spectrum with higher amplitude was set depending on the minimum MSE calculated between the original signal and the retrieved signal. However, when one does not know any information about the original signal, this method of

determining the length of the window is practical. So in our future work we will try to automate this length by finding a relationship between the type of the chirp signal and the width of the window used. Future research also includes applying warping-based analysis to estimate the signal order, k , and the number of signal components, n , as well as finding new warping operators that can exploit the Lambert- W function.

BIBLIOGRAPHY

- [1] M. Adjouadi, M. Cabrerizo, M. Ayala, and N. Mirkovic., "Seizing lesions in 3-D," *IEEE Potentials*, vol. 24, pp. 11–17, 2006.
- [2] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indication of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: dependence on recording region and brain state," *Physical Review E*, vol. 64:061907, 2001.
- [3] R. S. Fisher, G. Harding, G. Erba, G. L. Barkley, and A. Wilkins, "Photic-and pattern-induced seizures: a review for the epilepsy foundation of America working group," *Epilepsia*, vol. 46, pp. 1426-1441, 2005.
- [4] A. T. Tzallas M. G. Tsipouras, D. G. Tsalikakis, E. C. Karvounis, L. Astrakas, S. Konitsiotis, and M. Tzaphlidou, Automated epileptic seizure detection methods: a review study, In: Epilepsy-histological, electroencephalographic and psychological aspects, *INTECH Open Access Publisher*, edited by D. Stevanovic, pp. 75–98, 2012.
- [5] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lenhnertz, "Seizure prediction: the long and winding road," *Brain*, vol. 130, pp. 314–333, 2006.
- [6] R. Cooper, J. W. Osselton and J. C. Shaw, EEG technology, *Boston: Butterworth-Heinemann's, London*, 1980.
- [7] J. M. Rho, and R. Sankar, Epilepsy: mechanisms, models, and translational perspectives, *CRC Press, New York*, 684 pp., 2010.
- [8] N. B. Karayiannis, A. Mukherjee, J. R. Glover, P. Y. Ktonas, J. D. Frost, Jr., R. A. Hrachovy, and E. M. Mizrahi, "Detection of pseudo-sinusoidal epileptic seizure segments in the neonatal EEG by cascading a rule-based algorithm with a neural network," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 633–641, 2006.
- [9] W. Besio, K. Koka, R. Aakula, and W. Dai, "Tri-polar concentric electrode development for high resolution EEG Laplacian electroencephalography using tri-polar concentric ring electrodes," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 926-933, 2006.
- [10] K. Koka and W. G. Besio, "Improvement of spatial selectivity and decrease of mutual information of tri-polar concentric ring electrodes," *Journal of Neuroscience Methods*, vol. 165, pp. 216–222, 2007.

- [11] J. B. Ochoa, "EEG signal classification for brain computer interface application," *Ecole Polytechnique Federale de Lausanne*, vol. 7, pp. 1–72, 2002.
- [12] F. Babiloni, C. Babiloni, L. Fattorini, F. Carducci, P. Onorati, and A. Urbano, "Performances of surface Laplacian estimators: a study of simulated and real scalp potential distributions," *Brain Topography*, vol. 8, pp. 35–45, 1995.
- [13] W. Besio, H. Cao, and P. Zhou, "Application of tripolar concentric electrodes and pre-feature selection algorithm for brain-computer interface," *IEEE Transactions on Neural System and Rehabilitation Engineering*, vol. 16, pp.191–194, 2008.
- [14] A. Feltane, G. Faye Boudreaux-Bartels and W. G. Besio. "Automatic Seizure Detection in Rats Using Laplacian EEG and Verification with Human Seizure Signals." *Annals of Biomedical Engineering*, vol. 41, pp. 645-654, 2013.
- [15] A. Feltane, G. F. Boudreaux Bartels, Y. Boudria, and W. Besio, "Seizure detection using empirical mode decomposition and time-frequency energy concentration," *IEEE Signal Processing in Medicine and Biology Symposium, New York City*, pp. 6, 2013.
- [16] C. E. Elger, "Future trends in epileptology," *Current Opinion in Neurology*, vol. 14, pp. 185–186, 2001.
- [17] J. Gotman, "Automatic recognition of epileptic seizures in the EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 54, pp. 530–540, 1982.
- [18] M. B. Connolly, F. W. Sharbrough, and P. K. Wong, Electrical fields and recording techniques, In J. S. Ebersole and T. A. Pedley, editors, Chapter 4 in *Current Practice of Clinical Electroencephalography*, Lippincott Williams & Wilkins, 3rd edition, pp. 72–99, 2003.
- [19] J. Gotman, "Automatic recognition of interictal spikes," *Electroencephalography and Clinical Neurophysiology*, vol. 37, pp. 93-114, 1985.
- [20] L. Guo, D. Rivero and A. Pazos, "Epileptic seizure detection using multi-wavelet transform based approximate entropy and artificial neural networks," *Journal of Neuroscience Methods*, vol. 193, pp. 156–163, 2010.
- [21] H. Hassanpour, M. Mesbah, and B. Bouashash, "Time frequency feature extraction of newborn EEG seizure using SVD-based techniques," *EURASIP Journal on Applied Signal Processing*, vol. 16, pp. 2544–2554, 2004.

- [22] A. Tzalles, M. Tsipouras, and D. Fotiadis, "Epileptic seizure detection in EEGs using time–frequency analysis," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, pp. 703-710, 2009.
- [23] J. Gotman, "Automatic detection of seizures and spikes," *Journal of Clinical Neurophysiology*, vol. 16, pp. 130-140, 1999.
- [24] K. Polat and S. Gunes, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform," *Applied Mathematics and Computation*, vol. 187, pp. 1017–1026, 2007.
- [25] K. Polat, and S. Gunes, "Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals," *Expert Systems with Applications*, vol. 34, pp. 2039-2048, 2008.
- [26] A. Alkan, E. Koklukaya, and A. Subasi, "Automatic seizure detection in EEG using logistic regression and artificial neural network," *Journal of Neuroscience Methods*, vol. 148, pp. 167-176, 2005.
- [27] K. C. Chua, V.Chandran, R. Acharya, and C. M. Lim, "Automatic identification of epilepsy by HOS and power spectrum parameters using EEG signals: a comparative study," *Engineering in Medicine and Biology Society, EMBS, 30th Annual International Conference of the IEEE*, pp. 3824-3827, 2008.
- [28] V. P. Nigam and D. Graupe, "A neural-network-based detection of epilepsy," *Neurological Research*, vol. 26, pp. 55-60, 2004.
- [29] E. D. Ubeyli, "Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals," *Expert Systems with Applications*, vol. 37, pp. 233–239, 2010.
- [30] A. Subasi and I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Systems with Applications*, vol. 37, pp. 8659–8666, 2010.
- [31] I. Guler and E. D. Ubeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *Journal of Neuroscience Methods*, vol. 148, pp. 113-121, 2005.
- [32] I. Guler and E. D. Ubeyli, "Multiclass support vector machines for EEG-signals classification," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, pp. 117-126, 2007.

- [33] E. D. Ubeyli, "Wavelet/mixture of experts network structure for EEG classification," *Expert Systems with Applications*, vol. 37, pp. 1954-1962, 2008.
- [34] E. D. Ubeyli, "Combined neural network model employing wavelet coefficients for EEG signals classification," *Digital Signal Processing Journal*, vol. 19, pp. 297-308, 2009.
- [35] E. D. Ubeyli, "Probabilistic neural networks combined with wavelet coefficients for analysis of EEG signals," *Expert Systems*, vol. 26, pp. 147-159, 2009.
- [36] V. Clarkson, "Efficient single frequency estimator," in *Proceedings International Symposium in Signal Processing and Application, Gold Coast. Australia*, 4 p., 1992.
- [37] S. Qian, Introduction to Time-frequency and wavelet transforms, *Prentice-Hall, New York, USA*, 2002.
- [38] L. Cohen, "Time-Frequency Distributions -A Review," *Proceedings of IEEE*, vol. 77, pp. 941-981, 1989.
- [39] L. Cohen, Time-frequency analysis, *Prentice-Hall Signal Processing, New York, USA*, 1995.
- [40] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and Quadratic Time-Frequency Signal Representations," *IEEE Signal Processing Magazine*, pp. 21-67, 1992.
- [41] V. C. Chen and H. Ling, Time-Frequency Transforms for Radar Imaging and Signal Analysis, *Artech House, Boston, MA, USA*, 2002.
- [42] S. Qian, and D. Chen, Joint time-frequency analysis: methods and applications, *Prentice-Hall Inc., New York, USA*, 1996.
- [43] T. Thayaparan, "Linear and quadratic time-frequency representations," *Defense Research Establishment, OTTAWA, (ONTARIO Canada, DREO TM-2000-080)*, 2000.
- [44] E. O. Brigham, "The Fast Fourier Transform and Applications," *Englewood Cliffs, NJ: Prentice Hall*, vol. 1, 1988.
- [45] B. Escudie, C. Gazanhes, H. Tachoire, and V. Torra, "Des Cordes aux Ondelettes," *Publication De l'Université de Provence*, 2001.

- [46] D. C. Cassidy, "Answer to the question: When did the indeterminacy principle become the uncertainty principle," *American Journal of Physics*, vol. 66, pp. 278–279, 1998.
- [47] P. Gibbins, "A note on quantum logic and the uncertainty principle," *Philosophy of Science*, vol. 48, pp. 122–126, 1981.
- [48] S. Kadambe and G. F. Boudreaux-Bartels, "A comparison of the existence of cross terms in the Wigner distribution and the squared magnitude of the Wavelet Transform and the Short Time Fourier transform" *IEEE Transactions on Signal Processing*, vol. 40, pp. 2498–2517, 1992.
- [49] T. A. Classen and F. G. Mecklenbrauker, "The Wigner distribution-a tool for time-frequency signal analysis Part I: continuous – time signals," *Philips Journal Res.*, vol. 35, pp. 217–250, 1980.
- [50] E. F. Velez, R. G. Absher, "Transient analysis of speech signal using the Wigner time frequency representation," *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Glasgow, Scotland, pp. 2242–2245, May 1989.
- [51] F. Auger, P. Flandrin, Y. T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H. Wu, "Time-Frequency Reassignment and Synchrosqueezing: An overview," *IEEE Signal Processing Magazine*, vol. 30, pp. 32–41, 2013.
- [52] F. Auger and P. Flandrin, "Improving the Readability of Time–Frequency and Time-Scale Representations by the Reassignment Method," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1068–1089, May 1995.
- [53] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, vol. 30, pp. 243–261, 2011.
- [54] N. E. Huang, Z. Shen, S. R. Long, Wu, M. L., H. H. Shih, Q. Zheng, Yen N. C., C. C. Tung, and H. H. Liu, "The Empirical Mode Decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis," *In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, The Royal Society, 1998.
- [55] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," *Wavelet in Medicine and Biology*, pp. 527–546, 1996.
- [56] A. Feltane, G. F. Boudreaux-Bartels, J. Gaitanis, Y. Boudria, and W. Besio, "Human seizure detection using quadratic Rényi entropy," in *6th International*

IEEE/EMBS Conference on Neural Engineering (NER), San Diego, CA , USA, pp. 815-818, 2013.

- [57] A. Feltane, G. F. Boudreaux-Bartels, Y. Boudria and W. Besio, "Analyzing the presence of chirp signals in the electroencephalogram during seizure using the reassignment time-frequency representation and the Hough transform," in *6th International IEEE/EMBS Conference on Neural Engineering (NER)*, San Diego, CA , USA, pp. 186-189, 2013.
- [58] B. Boashash, *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, Academic Press, 2015.
- [59] L. Diambra, J. de Figueiredo, and C. Malta, "Epileptic activity recognition in EEG recording," *Physical A: Statistical Mechanics and its Applications*, vol. 273, pp. 495–505, 1999.
- [60] R. M. Corless, G. H. Gonnet, D.E.G. Hare, D. J. Jeffrey, and Knuth, "On the Lambert W Function," *Advances in Computational Mathematics*, vol. 5, pp. 329-359, 1996.
- [61] Y. Kumar and M. L. Dewal, "Complexity measures for normal and epileptic EEG signals using ApEn, SampEn and SEN," *International Journal of Computer and Communication Technology (IJCCT)*, vol. 2, pp. 6–12, 2011.
- [62] N. Radhakrishnan and B. Gangadhar, "Estimating regularity in epileptic seizure time-series data: a complexity measure approach," *IEEE Engineering in Medicine and Biology Magazine*, vol. 17, pp. 89-94, 1998.
- [63] J. S. Richman and J. R. Moorman, "Physiological time series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, pp. H2039-H2049, 2000.
- [64] D. Kalman, "A singularly valuable decomposition: the SVD of a matrix," *College Math Journal*, vol. 27, pp. 2–23, 1996.
- [65] P. J. Rouss and C. Croux, "Alternatives to the median absolute deviation," *Journal of the American Statistical Association*, vol. 88, pp. 1273–1283, 1993.
- [66] M. Bedeuzzaman, O. Farooq and Y. U. Khan, "Dispersion measures and entropy for seizure detection," *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP)*, Prague, Czech Republic, pp. 673–676, 2011.
- [67] S. R. Gunn, "Support vector machines for classification and regression," *Faculty of Engineering, Science and Mathematics School of Electronics and Computer Science*, University of Southampton, 1998.

- [68] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, New York, 2000.
- [69] K. P. Soman, S. Diwakar and V. Ajay, *Data mining: theory and practice*, PHI Learning Pvt. Ltd., 2006.
- [70] W. Xue, Wei, and Xiao-Wei Sun. "Multiple targets detection method based on binary hough transform and adaptive time-frequency filtering," *Progress In Electromagnetics Research*, vol. 74, pp. 309-317, 2007.
- [71] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, pp. 119-139, 1997.
- [72] A. G. Lalkhen and A. McCluskey, "Clinical tests: sensitivity and specificity," *Continuing Education in Anesthesia, Critical Care & Pain*, vol. 8, pp. 221-223, 2008.
- [73] T. Fathima, M. Bedeuzzaman, O. Farooq, and Y. U. Khan, "Wavelet based features for epileptic seizure detection," *MES Journal of Technology and Management*, vol. 2, pp. 108–112, 2011.
- [74] A. Rényi, "On measure of entropy and information," *4th Berkeley Symposium, Math Statistical and Probability*, vol. 1, pp. 547-561, 1961.
- [75] R. B. Pachori and V. Bajaj, "Analysis of normal and epileptic seizure EEG signals using empirical mode decomposition," *Computer Methods and Programs in Biomedicine*, vol. 104, pp. 373- 381, 2011.
- [76] A. K. Tafreshi, A. M. Nasrabadi and A. H. Omidvarnia, "Epileptic Seizure Detection using Empirical Mode Decomposition," *IEEE International Signal Processing and Information Technology, ISSPIT*, pp. 238-242, 2008.
- [77] A. Roy, "Signal Analysis using Raised Cosine Empirical Mode Decomposition," *PHD Dissertation, The Pennsylvania State University*, Philadelphia, PA, 2011.
- [78] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters*, vol. 11, pp. 112-114, 2004.
- [79] M. Akay and C. Mello, "Wavelets for biomedical signal processing," *In Engineering in Medicine and Biology Society, Proceedings of the 19th Annual International Conference of the IEEE*, vol. 6, pp. 2688-2691, 1997.

- [80] L. Guo, D. Rivero, J. Dorado, J. R. Rabuñal and A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks," *Neuroscience Methods*, vol. 191, pp. 101-109, 2010.
- [81] S. Aviyente, "Information Theoretic Measures for Quantifying the Integration of Neural Activity," in *Information Theory and Application Workshop*, La Jolla, CA , pp. 20-26, 2007.
- [82] Y. Song and P. Liò, "A new approach for epileptic seizure detection: sample entropy based features extraction and extreme learning machine," *Journal of Biomedical Science and Engineering*, vol. 3, pp. 556-567, 2010.
- [83] P. Grassberger and I. Procaccia, "Estimation of the Kolmogorov entropy from a chaotic signal," *Physical Review*, vol. 28, pp. 2591-2593, 1983.
- [84] S. Zhang, "KNN-CF Approach: Incorporating Certainty Factor to kNN Classification," *IEEE Intelligent Informatics Bulletin*, vol. 11, pp. 24-33, 2010.
- [85] N. Kannathal, M. L. Choo, U. R. Acharya and P. K. Sadasivan, "Entropies for detection of epilepsy in EEG," *Computer Methods and Programs in Biomedicine*, vol. 80, pp. 187-194, 2005.
- [86] A. Jarrot, P. Oonincx, C. Ioana and A. Quinquis, "Identifying non-linear fractional chirps using unsupervised Hilbert approach, *14th European IEEE In Signal Processing Conference, Florence, Italy*, pp. 1-5, 2006.
- [87] E. D. Ubeyli, "Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals," *Expert Systems with Applications*, vol. 37, pp. 233-239, 2010.
- [88] Empirical Mode Decomposition Algorithm.
<http://perso.ens-lyon.fr/patrick.flandrin/emd.html>. Accessed on May 11, 2011.
- [89] P. Flandrin, P. Goncalves, and G. Rilling, "EMD equivalent filter banks, from interpretation to application," *Chapter in Hilbert-Huang Transform and its Applications, published by: World Scientific Publishing Co. Pte. Ltd*, Singapore, pp. 57-74, 2005.
- [90] F. Auger, P. Flandrin, P. Gonçalves, and O. Lemoine, "Time-Frequency Toolbox," *France: CNRS/Rice University*, 1995–1996. [Online]. Available: <http://tftb.nongnu.org>.
- [91] W. J. Krzanowski, *Principles of Multivariate Analysis: A User's Perspective*, New York: Oxford University Press, 1988.

- [92] G. C. J. Fernandez, "Discriminant analysis, a powerful classification technique in data mining," In *Proceeding of the SAS Users International Conference*, pp. 247-256, 2002.
- [93] V. Namias, "The Fractional order Fourier Transform and its application to quantum mechanics," *IMA Journal of Applied Mathematics*, vol. 25, pp. 241-265, 1980.
- [94] H. Ozaktas, Z. Zalevsky and A. Kutay, *FrFT with applications in optics and signal processing*, John Wiley and Sons, Chichester, 2001.
- [95] H. Ozaktas, O. Arikan, A. Kutay and G. Bozdagi, "Digital Computation of the Fractional Fourier Transform," *IEEE Transactions on Signal Processing*, vol. 44, pp. 2141-2150, 1996.
- [96] E. Sejdic, I. Djurovic, L. Stankovic, "Fractional Fourier transform as a signal processing tool: An overview of recent developments," *Signal Processing*, vol. 91, pp. 1351-1369, 2011.
- [97] L. B. Almeida, "The Fractional Fourier transform and time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 42, pp. 3084-3091, 1994.
- [98] F. Hlawatsch, T. G. Manickam, R. L. Urbanke, W. Jones, "Smoothed pseudo-Wigner distribution, Choi-Williams distribution, and cone-kernel representation: Ambiguity-domain analysis and experimental comparison," *Signal Processing*, vol. 43, pp. 149-168, 1995.
- [99] O. Akay, and G. F. Boudreaux-Bartels, "Fractional convolution and correlation via operator methods and an application to detection of linear FM signals," *IEEE Transactions on Signal Processing*, vol. 49, pp. 979-993, 2001.
- [100] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-34, pp. 442-451, 1986.
- [101] B. Boashash, "Theory of quadratic TFDs," in *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, B. Boashash, Ed., chapter 3, Elsevier, Oxford, pp. 59-76, 2003.
- [102] B. Boashash, Time-frequency signal analysis, in *Advances in Spectrum Estimation and Array Processing*, S. Haykin, Ed., vol. 1, chapter 9, Englewood Cliffs, NJ: Prentice-Hall, pp. 418-517, 1991.
- [103] P. Flandrin, Patrick, Time-frequency/time-scale analysis, Wavelet analysis and its Applications, vol. 10, Academic Press, 1998.

- [104] A. Papandreou-Suppappola, Applications in time-frequency signal processing, vol. 10, *CRC Press*, 2010.
- [105] R. Baraniuk, "Unitary Equivalence: a new twist on signal processing, " *IEEE Transactions on Signal Processing*, vol. 43, pp. 2269-2282, 1995.
- [106] G. R. Baraniuk, "Warped perspectives in time-frequency analysis," *Proceedings of the IEEE-SP International Symposium*, pp. 528-531, 1994.
- [107] H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay, The Fractional Fourier Transform with Applications in Optics and Signal Processing. *New York: Wiley*, 2001.
- [108] C. Ioana, C. Cornu and A. Quinquis, "Polynomial phase signal processing via warped high-order ambiguity function," *In 12th European Signal Processing Conference, Vienna , Austria*, pp. 1159-1162, 2004.
- [109] A. Papandreou-Suppappola and G. F. Boudreaux-Bartels, "The exponential class and generalized time-shift covariant quadratic time-frequency representations," *in Proceedings IEEE International Symposium on Time-Frequency/Time-Scale Analysis, (Paris, France)*, pp. 429-432, 1996.
- [110] B. Barkat and B. Boashash, "Instantaneous frequency estimation of polynomial FM signals using the peak of the PWVD: Statistical performance in the presence of additive Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 47), pp. 2480-2490, 1999.
- [111] G. Le Touzé, J. I. Mars and J. L. Lacoume, " Matched time-frequency representations and warping operator for modal filtering," *In EURASIP Signal and Image Processing Conference (EUSIPCO 06), Florence, Italy*, pp. 1-5, 2006.
- [112] B. Barkat and B. Boashash, "Design of higher-order polynomial Wigner-Ville distributions," *IEEE Transaction on Signal Processing*, vol. 47, pp. 2608–2611, 1999.
- [113] P. Flandrin, R. G. Baraniuk, O. Michel, "Time-frequency complexity and information," *Proceeding of IEEE International Conference on Acoustics and Speech Signal Processing, (ICASSP), Adelaide, Australia*, vol.3, pp. 329-332, 1994.
- [114] W. J. Williams, M. Brown, A. Hero, "Uncertainty, information and time frequency distributions," *In SPIE, Advanced Signal Processing Algorithms*, vol. 1556, pp. 144-156, 1991.

- [115] S. C. Pei and J. J. Ding, "Fractional Fourier Transform, Wigner Distribution and filter design for stationary random processes," *IEEE Transactions on Signal Processing*, vol. 58, pp. 4079-4092, 2010.
- [116] A. W. Lohmann, "Image rotation, Wigner rotation, and the fractional Fourier transform," *Journal of the Optical Society of America A (JOSA A)*, vol. 10, pp. 2181-2186, 1993.
- [117] A. Papandreou, S. M. Kay, and G. F. Boudreaux-Bartels, "The use of hyperbolic time-frequency representations for optimum detection and parameter estimation of hyperbolic chirps," in *Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis (TFTS '94)*, Philadelphia, PA, USA, pp. 369-372, 1994.
- [118] F. Hlawatsch, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels, "The power classes-quadratic time-frequency representations with scale covariance and dispersive time-shift covariance," *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 3067-3083, 1999.
- [119] T. Twaroch and F. Hlawatsch, "Modulation and warping operators in joint signal analysis," in *Proc. IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis (TFTS)*, Pittsburgh, PA, USA, pp. 9-12, 1998.
- [120] R. G. Baraniuk, P. Flandrin, A. J. Janssen and O. J. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Transactions on Information Theory*, vol. 47, pp. 1391-1409, 2001.
- [121] D. Jones and T. Parks, "A high resolution data-adaptive time frequency representation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 2127-2135, 1990.
- [122] Bat Echolocation Chirp Signal.
<http://dsp.rice.edu/software/bat-echolocation-chirp>. Accessed on Jun 14, 2010.
- [123] B. Boashash, N. A. Khan and T. Ben-Jabeur, "Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review," *Digital Signal Processing*, vol. 40, pp. 1-30, 2015.
- [124] L. Zhang, S. Liu and J. Wu, "Instantaneous Frequency Estimation of Multi-component Chirp Signals Based on FrFT," In *Proceedings of the 2012 International Conference on Computer Application and System Modeling*. Atlantis Press, 4 p., 2012.
- [125] Y. Lu, A. Kasaeifard, E. Oruklu and J. Saniie, "Fractional Fourier Transform for ultrasonic Chirplet signal decomposition," *Advances in Acoustics and Vibration*, pp. 1-13, 2012.

- [126] C. Capus, Y. Rzhanov and L. Linnett, "The analysis of multiple chirp signals" *IEEE Seminar on Time-scale and Time-frequency Analysis and Applications*, London, pp. 1-4, 2000.
- [127] L. Durak and S. Aldirmaz, "Adaptive Fractional Fourier domain filtering", *Signal Processing*, pp. 1188-1196, 2009.
- [128] D. Cowell and S. Freear, "Separation of overlapping linear frequency modulated (LFM) signals using the Fractional Fourier Transform," *IEEE Transactions on Ultrasonic, Ferroelectrics and Frequency Control*, vol. 57, No. 10, pp. 2324-2333, 2010.
- [129] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, pp. 98-117, 2009.