REDUCED FUZZY RECURSIVE LEAST-SQUARES ALGORITHM FOR REAL TIME ESTIMATION

CHEAH KAH WAI

UNIVERSITI SAINS MALAYSIA

2018

REDUCED FUZZY RECURSIVE LEAST-SQUARES ALGORITHM FOR REAL TIME ESTIMATION

by

CHEAH KAH WAI

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

April 2018

ACKNOWLEDGEMENT

First and foremost, I would like to address my words of deepest appreciation to my supervisor, Associate Professor Dr. Noor Atinah Ahmad for her guidance and her valuable advice during my journey of pursuing my doctorate degree. Although it was challenging at the beginning of research, but I am able to successfully complete my research due to her patience and guidance and this thesis could not have been possible without her. I am grateful to have Associate Professor Dr. Noor Atinah Ahmad to be my supervisor as she consistently steered me to the right direction whenever I needed her guidance.

Next, I would like to express my heartiest gratitude to my family, especially my father and my mother for their unfailing support. At the same time, I wish to apologize for not able to accompany them most of the time due to my study. I would like to honor my parents and gratefully thank them for their encouragement as this accomplishment would not have been possible without them. I am gratefully indebted to my parents for their understanding and always stay beside me during the ups and downs.

On the other hand, I would like to acknowledge Universiti Sains Malaysia and The Ministry of Higher Education Malaysia for their financial support; through The Fundamental Research Grant Scheme for conferences and publications funding and MyBrain15 (MyPhD) for the scholarship of my postgraduate study. Lastly, for those who have ever helped me, I would like to apologize for not able to list all your names here; I must express my profound gratitude for assisting my doctoral study throughout the years.

Thank you.

TABLE OF CONTENTS

CHAPTER 2 – ADAPTIVE FILTERING AND RECURSIVE LEAST-SQUARES

CHAPTER 3 – FUZZY SYSTEM

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MSE	Mean-Square Error
FIR	Finite-duration Impulse Response
IIR	Infinite-duration Impulse Response
FLOP	Floating-point Operation
RLS	Recursive Least-Squares
FRLS	Fuzzy Recursive Least-Squares
LMS	Least-Mean-Square
SVD	Singular Value Decomposition
QRD-RLS	QR Decomposition Based Recursive Least-Squares
IQRD-RLS	Inverse QR Decomposition Based Recursive Least-Squares
IQRD-rFRLS	Inverse QR Decomposition Based Reduced Fuzzy Recursive Least-Squares
SOV	Truncated Second Order Volterra Series Expansion
SOV-RLS	Second Order Volterra Based Recursive Least-Squares
LPC	Linear Predictive Coding
NPC	Nonlinear Predictive Coding
FRLS-LPC	Linear Predictive Coding With Reduced Fuzzy Recursive Least-Squares
SNR	Signal-to-noise Ratio

LIST OF SYMBOLS

z^{-1}	Unit-delay operator
z^{-L}	Series of unit-delay operators
$H(z)$	Transfer function
$x(k)$	Input at time instant k
$w(k)$	Filter coefficient at time instant k
$y(k)$	Filtering output at time instant k
$d(k)$	Desired output at time instant k
$n(k)$	Noise at time instant k
$e(k)$	The *a posteriori* error at time instant k
$\varepsilon(k)$	The *a priori* error at time instant k
$\mathbf{x}(k)$	Column vector consists of the inputs at time instant k
$\mathbf{w}(k)$	Column vector consists of the filter coefficients at time instant k
$\mathbf{R}_D(k)$	Deterministic correlation matrix of the inputs at time instant k
$\mathbf{R}_D{}^{-1}(k)$	Inverse deterministic correlation matrix of the inputs at time instant k
$\widetilde{\mathbf{R}}_D(k)$	Shifted deterministic correlation matrix of the inputs at time instant k
$\mathbf{p}_D(k)$	Deterministic cross-correlation vector between the desired output and the inputs at time instant k
$\mathbf{g}(k)$	Gain vector at time instant k
$\mathbf{L}_D(k)$	Cholesky factor (lower triangular matrix) at time instant k

$\mathbf{L}_D{}^{-1}(k)$	Inverse Cholesky factor at time instant k
$\mathbf{Q}(k)$	Orthogonal matrix at time instant k
$\mathbf{w}_o$	Optimum filter coefficient column vector
$\mathbf{R}$	Correlation matrix of the inputs
$\mathbf{p}$	Cross-correlation vector between the desired output and the inputs
$e_o(k)$	Optimum filtering error using the optimum filter coefficient column vector at time instant k
$y_o(k)$	Optimum filtering output using the optimum filter coefficient column vector at time instant k
$\Delta\mathbf{w}(k)$	Difference of the filter coefficients at time instant k and the optimum filter coefficients
$\emptyset$	Empty set
$\overline{A}$	Complement of set A
$x \in A$	x is the element of set A
$x \notin A$	x is not the element of set A
$A \subseteq B$	Set A is the subset of set B
$A \subset B$	Set A is the proper subset of set B
$A \cup B$	Union of set A and set B
$A \cap B$	Intersection of set A and set B
$\circ$	Fuzzy relation
$\star$	Sup-star composition

$s(A,B)$	Fuzzy union of fuzzy set A and B (or equivalently s-norm)
$t(A,B)$	Fuzzy intersection of fuzzy set A and B (or equivalently t-norm)
$max(a,b)$	Maximum in between value a and b
$min(a,b)$	Minimum in between value a and b
$\mu_A(x)$	Degree of membership of x corresponding to fuzzy set A
$\mu_{A_i{}^j}(x_i(k))$	Degree of membership of x_i at time instant k corresponding to fuzzy set $A_i{}^j$ of j^{th} fuzzy rule
$\mu_{A_{i(t)}}(x_i(k))$	Degree of membership of x_i at time instant k corresponding to fuzzy set $A_{i(t)}$ of t attribute
ζ_i	Total associated attributes of the degree of membership of x_i
$\Delta_{x_1,x_2,x_3,\ldots,x_n}$	The n^{th}-D partitioned space of n cross-attribute inputs
$\theta_j(k)$	Fuzzy parameter / coefficient corresponding to j^{th} fuzzy rule at time instant k
$p_j(k)$	Fuzzy basis function corresponding to j^{th} fuzzy rule at time instant k
$\mu_j(k)$	Reduced fuzzy basis function corresponding to j^{th} fuzzy rule at time instant k
$\mu_d(k)$	Total partitioned space of all fuzzy rules at time instant k
$\bar{\mu}$	Expected mean value of $\mu_d(k)$
$\Theta(k)$	Column vector consists of all the fuzzy parameters at time instant k
$\mathbf{B}(k)$	Column vector consists of the fuzzy basis functions at time instant k

$\mathbf{M}(k)$	Column vector consists of the reduced fuzzy basis functions at time instant k
$\mathbf{R}_{D(FRLS)}(k)$	Deterministic correlation matrix of the fuzzy inputs at time instant k
$\mathbf{p}_{D(FRLS)}(k)$	Deterministic cross-correlation vector between the desired output and the fuzzy inputs at time instant k
$\mathbf{R_B}$	Correlation matrix of the fuzzy inputs
$E(.)$	Statistical expectation
$tr(\mathbf{A})$	Trace of matrix **A**
$cond_2(\mathbf{A})$	Condition number of matrix **A**, with respect to 2-norm
$cond_F(\mathbf{A})$	Condition number of matrix **A**, with respect to Forbenius norm
ξ	Performance function / Objective function of MSE
ξ_{min}	Minimum MSE
ξ_{exc}	Excess MSE
$\mathbf{g_w}$	Gradient vector of MSE
$\mathbf{g}_{\Theta(k)}$	Gradient vector with respect to $\Theta(k)$
$\xi_d(k)$	Deterministic objective function
λ	Forgetting factor
γ	A small constant for the initialization of the algorithm
υ_n	Variance of noise
C_{RLS}	Error conversion factor of the RLS algorithm

C_{FRLS}	Error conversion factor of the FRLS algorithm
$\xi_{dB}(k)$	The MSE in Decibel scale at time instant k
$log_{10}(.)$	Base 10 logarithm scale
$round(.)$	Rounding operation to nearest integer
$cond_{dB}(\mathbf{A})(k)$	Condition number of matrix **A** in Decibel scale at time instant k

ALGORITMA KUASA-DUA TERKECIL REKURSI KABUR TERTURUN UNTUK PENGANGGARAN MASA NYATA

ABSTRAK

Penapisan adaptif merupakan salah satu teknik pemprosesan isyarat dimana ia menganggarkan parameter yang bertujuan mencirikan suatu sistem masa nyata. Penganggaran tersebut melibatkan peminimuman min ralat kuasa-dua di antara output suatu sistem dan output penapisan adaptif. Fungsi pindahan suatu sistem adalah tidak diketahui dan penganggaran struktur fungsi pindahan adalah berdasarkan andaian struktur turas FIR, turas IIR, ataupun turas tidak linear dimana ketiga-tiga struktur turas yang dinyata adalah struktur turas yang lazim dalam penapisan adaptif. Akan tetapi, setiap turas adalah berbeza sesama lain dan ia mempunyai isu yang tersendiri. Turas FIR adalah terhad kepada sistem linear; turas IIR mempunyai prestasi baik dalam sistem suap balik linear tetapi menghadapi masalah ketakstabilan apabila struktur direalisasikan; dan turas tidak linear melibatkan peringkat persamaan yang tinggi dan mempunyai kecenderungan ke sistem bersuasana tak sihat. Didorong oleh pendekatan kabur, tesis ini menyiasat keupayaan sistem kabur dan membentangkan satu pendekatan yang boleh mengatasi masalah-masalah turas yang disebut tadi. Untuk merealisasikan pendekatan tersebut, hubungan antara penapisan adaptif dan sistem kabur dibentuk melalui generalisasi fungsi pindahan dimana pemetaan kabur digunakan untuk menerang hubungan input-output yang diberi oleh penapisan adaptif tanpa andaian struktur turas. Tesis ini melanjutkan kajian atas algoritma kuasa-dua terkecil rekursi kabur (FRLS) Wang dan Mendel (1993a) dan memperkenalkan satu algoritma yang mempunyai struktur pengiraan yang lebih ringkas iaitu algoritma FRLS terturun. Berikutan dengan algoritma FRLS terturun, fungsi asas kabur terturun didefinasikan sebagai input kabur yang mengaplikasikan fungsi keahlian segi tiga. Dalam rujukan literasi,

fungsi keahlian segi tiga menghadapi masalah ketakrifan ruang metrik disebabkan kemungkinan berlakunya kekosongan ruang pemetakan sistem kabur. Akan tetapi, masalah ketakrifan ruang metrik dapat diselesaikan dengan pemetakan Ruspini dan fungsi asas kabur terturun adalah dibukti sebagai kaedah yang lengkap serta mempunyai pemetakan kabur yang kuat. Yang paling penting, fungsi asas kabur terturun dibukti dari segi teori dimana ia juga merupakan penganggar umum sepertimana yang dibukti untuk fungsi asas kabur dalam algoritma FRLS Wang dan Mendel. Kaedah-kaedah penilaian yang lazim untuk algoritma penapisan adaptif telah digunakan untuk menguji dan menilai prestasi algoritma FRLS terturun. Pelbagai simulasi telah dijalankan dan algoritma FRLS terturun menunjukkan kadar penumpuan serta tingkah laku ralat yang baik dimana min ralat kuasa-dua yang minimum dapat dicapai. Dalam pengenalan sistem tidak linear, prestasi algoritma FRLS terturun adalah setanding dengan algoritma FRLS tetapi bilangan operasi titik apungan algoritma FRLS terturun adalah lebih kurang dan ia telah ditunjukkan melalui masa pengiraan yang lebih kurang. Selain itu, algoritma FRLS terturun juga diuji dengan aritmetik kepersisan yang terhingga untuk menyiasat prestasi algoritma dalam sistem bersuasana tak sihat. Keputusan simulasi menunjukkan prestasi algoritma FRLS terturun adalah lebih baik daripada algoritma SOV-RLS dan setanding dengan algoritma yang berdasarkan penguraian QR dimana kestabilan matriks korelasi input telah dijaga. Dalam pengenalan sistem IIR yang melibatkan suap balik, algoritma FRLS terturun juga menunjukkan prestasi yang setanding dengan algoritma Bilinear RLS malahan lebih baik apabila melibatkan ketaklinearan. Akhir sekali, algoritma FRLS terturun dalam penganggaran masa nyata ditunjukkan dengan aplikasi pengenalan sistem suara. Keputusan ekperimen adalah memberansangkan dimana algoritma FRLS terturun berupaya mensintesiskan semula ujaran yang dicemar teruk dengan hanya menggunakan bilangan sample ujaran lepas yang kurang, padahal model pengenalan sistem suara yang sedia ada telah gagal.

REDUCED FUZZY RECURSIVE LEAST-SQUARES ALGORITHM FOR REAL TIME ESTIMATION

ABSTRACT

Adaptive filtering is an online signal processing application that is capable of estimating parameters for the characterization of a real time system. The estimation is done based on the mean-square error (MSE) minimization of the difference between some desired output and the output of the adaptive filter. The unknown transfer function is assumed to be a known structure and is realized by three different structures which are commonly used in conventional adaptive filtering: (i) the finite-duration impulse response (FIR) filter, (ii) the infinite-duration impulse response (IIR) filter, and (iii) the nonlinear filter. However, there are limitations on each structures: (i) the FIR filter is limited for linear and almost linear system, (ii) the IIR filter works well with linear feedback system but encounters some instability issue when implementing the structure in practice, and (iii) the nonlinear filter requires higher order to describe the unknown nonlinearity and it is prone to ill-conditioned system. Motivated by the fuzzy approach, this thesis seeks to investigate the capabilities of the fuzzy system in overcoming the limitations above. By formulating the adaptive filter as a generalized transfer function to include both feedforward and feedback mechanism, an interesting connection between adaptive filtering and fuzzy system is established. It is discovered that fuzzy mapping can be used to realize the unknown input-output relationship in adaptive filtering without prior assumption on the adaptive filter structure. This thesis extends the fuzzy recursive least-squares (FRLS) algorithm proposed by Wang and Mendel (1993a) to develop a more computational simplified algorithm which is the reduced FRLS algorithm. The reduced fuzzy basis function associated with the reduced FRLS algorithm is defined based on the Triangular-shaped membership function. In

literature, the Triangular-shaped membership function may not be well-defined in the metric space due to the possibility of zero total partitioned space. However, this issue is resolved using Ruspini partitioning and the resulting reduced fuzzy basis function is proven to be a complete rule base method and has a strong fuzzy partitioning. Most importantly, a theoretical proof is provided to show that the reduced fuzzy basis function also shares the universal approximation property that characterizes the fuzzy basis function used in the FRLS algorithm by Wang and Mendel. Performance of the reduced FRLS algorithm is tested against a set of standard performance measures which is commonly used to test the performance of adaptive filtering algorithm. Simulations support that the reduced FRLS algorithm has good convergence and error behavior in which the minimum MSE is achievable. In nonlinear system identification, the reduced FRLS algorithm is able to provide comparable performance with the FRLS algorithm but requires less number of floating-point operations and also shown to have less computation time. Particularly in quantized environment where the system is ill-conditioned, the reduced FRLS algorithm gives a more stable performance compared to the SOV-RLS algorithm and it is shown comparable with the QR decomposition based algorithm whereby the condition number of the inputs correlation matrix remains stable throughout the iterations. The reduced FRLS algorithm also shows good performance in IIR system identification involving feedback mechanism where it provides comparable performance with the Bilinear RLS algorithm yet producing a better performance when nonlinearity is involved. Lastly, the reduced FRLS algorithm in real time estimation is presented by speech system identification. The experimental results are promising in which the reduced FRLS algorithm demonstrates the reconstruction of a seriously corrupted speech with only a limited number of past speech samples, while the conventional predictive coding model has failed.

CHAPTER 1

INTRODUCTION

1.1 Introduction to Real Time Estimation

In the real world, systems are said to possess certain characteristics that are usually unknown until certain modeling is performed to describe and estimate the unknown system. For a simple input-output system where the inputs and the outputs are measurable with certainty, estimation can be carried out by using the optimal estimation techniques in an off-line manner. In statistical theory, estimation theory is separated into two distinct approaches: the classical approach and the Bayesian approach (Sayed, 2008). In the classical approach, the unknown quantity to be estimated is modeled as a deterministic but unknown constant. The Bayesian approach, on the other hand, models the unknown quantity as a random variable. Nevertheless, both approaches require the statistical information about the unknown quantity to be estimated (i.e. the mean, the variance, or even the probability density function).

With the advancement in technologies and computing capacity, real time estimation has been in greater demand as the data can now be accessed in a real time manner. As the name suggests, real time data is the immediate information that is delivered without delay and real time estimation is the estimation of the real time data with objective to minimize the estimation error. Different to the conventional approach, real time estimation is carried out in a more demanding on-line manner where the estimation is performed continuously and it goes infinitely. Considering the dynamic of time varying environment in real time estimation, it is practically impossible to obtain the statistical information of the unknown quantity to be estimated, not to mention to store the infinite amount of real time data. Table 1.1 shows the differences between statistical estimation and real time estimation.

There are ways to perform real time estimation and one of the techniques is based on the engineering perspective: signal processing. Here, a signal is a representation of the intelligence to be conveyed to a receiver, while signal processing is a technical tool for the transformation and the manipulation of such representation (Moura et al., 2013). In regard to the real time data, these signals are often acquired from the real time operation such as biomedical systems, speech processing, control systems, communications, computing, geographical monitoring systems, and etc. Although these signals take many different forms, it is recorded as data sequences or in time series representation and serve as the real time data to their respective real time operation. Depending on the user of such signals, the aim and objective of signal processing vary according to the application: from the extraction of useful information contained in the real time system to the estimation and the identification of the unknown system.

Table 1.1: Table of differences between statistical estimation and real time estimation.

Statistical Estimation	**Real Time Estimation**
Number of data is finite due to specific data operation	Number of data is infinite due to real time data operation
Data are stored for its statistical information	Data are continuously collected and only the most recent are temporary stored
Unlimited storage space for limited amount of data	Limited storage space for unlimited amount of data
Stationary processes and data are time-invariant	Dynamic or nonstationary processes and data are time-variant
Statistical information is known by large number of data realization	Statistical information is impossible to know due to the dynamic of real time operation
Perform in an off-line manner	Perform in an on-line manner
Sample data are available for estimation training	Sample data are not available and irrelevant for estimation training
Accurate estimation based on known statistical information	Fast estimation on arbitrary accuracy to capture the fast dynamic of time varying environment
Extensive computation to have accurate estimation	Limited computational load to have fast estimation

1.1.1 Adaptive Filtering as a Real Time Estimation Technique

In the case of digital signal processing, one of the techniques used for estimation is by using filters. Here, filter is a device in software form that is used to extract information about a prescribed quantity of interest from a set of noisy data (Haykin, 1996). From the recording of signals from the sensor, to the transmission of signals through a channel, noise may arise and corrupts the signals, causing the failure in signals interpretation. Therefore, filtering is performed in order to get the information we need in the signals. In other words, filtering is a signal processing operation which objective is to process a signal in order to manipulate the information contained in the signal (Diniz, 2008).

Basically, filters can be classified into two categories: linear filter and nonlinear filter. As the word linear suggests, a filter is classified into linear filter if the output of the filter is a linear function of the observations applied to the filter input; otherwise, the filter is nonlinear (Haykin, 1996). In practice, the signals are often assumed to be a linear filtering problem, whose objective is to provide a solution that is optimum in the mean-square sense. With the availability of the signal and noise's statistical parameters, we want to design a linear filter so that a certain objective function of the error between the desired output (the actual signal output) and the filtering output is minimized. In the case of stationary data, where the specification of the signal and noise are fixed or known, a time-invariant filter can be designed to facilitate the statistical parameters. For this, a common approach which provides an optimum solution in the mean-square sense is the Wiener filter. Mean-square approach is alike to the least-squares approach with the difference of being stochastically implemented. In the error performance surface, Wiener filter gives a point in which the mean-square error (MSE) is minimum and the solution giving that point is called the Wiener solution.

For the design of time-invariant filter, the fixed specification or the statistical characteristics of the signal and noise are needed. In other words, the *a priori* information of the signal and noise are processed to obtain an optimum solution. However, the fixed specification of the signal is unknown in real time processing, especially when the signal is corrupted by noise. This is because the characteristic of noise is rather hard to estimate due to its randomness. In order to capture the unknown underlying dynamic of the signal, a time-variant filter in which can adaptively adjust to have optimum solution in the mean-square sense is needed. By assuming that the signal sequences are ergodic (Diniz, 2008) where the statistical characteristics of the signal can be estimated by a large number of time instant averaging, adaptive filter can be designed to facilitate the statistical parameters. As the word adaptive suggests, adaptive filter is a self-modifying filter that adjusts its coefficients in order to minimize the objective function of the error between the desired output and the filtering output (Apolinario and Netto, 2009).

By utilizing real time data, real time estimation via adaptive filtering can be done using the stochastic approach or the deterministic approach. Since the characteristic of these signals is unknown due to the infeasible of large number of signal realizations, ergodicity is assumed in adaptive filtering to perform the estimation. With the assumption of the unknown signal sequences are stationary and their time averages are identical, stochastic approach used the time averages of the signal to resolve the real time estimation. This can be done by the instantaneous estimation where the MSE is approximated by the instantaneous squared error. While for the deterministic approach, real time estimation is based on the recursive estimation which utilizes the method of least-squares in minimizing the formulated objective function.

1.2 Adaptive Filter Structure

Either the design of the filter is linear, nonlinear, adaptive, or nonadaptive, the objective of filtering is to minimize certain objective function. Let $d(k)$ be the desired output and $y(k)$ be the filtering output at time instant k, if the input $x(k)$ is available, then the objective of filtering is to minimize

$$\xi = E[e^2(k)] \quad , \tag{1.1}$$

where $e(k) = d(k) - y(k)$ is the error between the desired output and the filtering output and $E[.]$ denotes the statistical expectation of the given equation. From (1.1), we denote the performance function, ξ as the mean-square error (MSE). To measure the MSE, large number of signals realization is required and it is impractical for real time estimation. In practice, the performance function is approximated by using several different adaptive filtering approach which will be discussed in Chapter 2. Depending on the adaptation approach, the error signal $e(k)$ is used by the adaptive filter to produce the filter coefficients in which it is updated according to some performance measure. In other words, adaptive filtering will produce the filter coefficients that aims to minimize the objective function, forcing the filtering output to approximate the desired output in a statistical sense (Apolinario and Netto, 2009). Figure 1.1 shows the configuration of adaptive filtering. Referring to the configuration, only two types of data mattered in adaptive filtering: (i) the input $x(k)$, which is collected from the environment, is used to drive the adaptive filter and it can be time varying; and (ii) the desired output $d(k)$, which is also collected from the environment, is the estimation target and its nature depends on the type of adaptive filtering application.

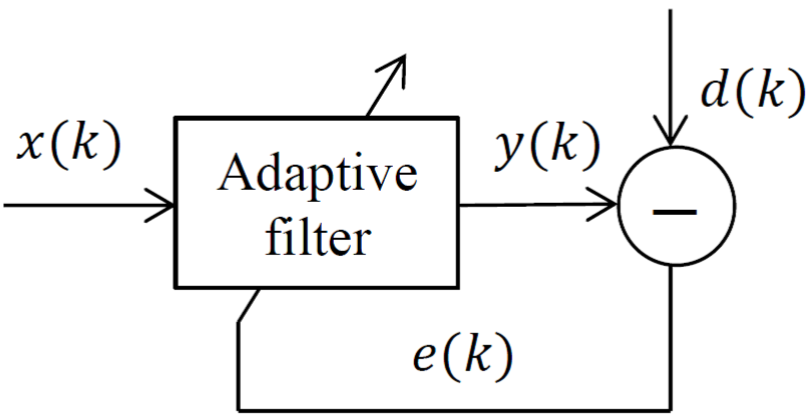

Figure 1.1: The configuration of adaptive filtering.

1.2.1 Finite-duration Impulse Response (FIR) Filter

Let $\mathbf{x}(k)$ be the input column vector and realized through a transversal structure such that $\mathbf{x}(k) = \begin{pmatrix} x(k) & x(k-1) & \dots & x(k-n+1) \end{pmatrix}^T$, where it consists of the current input, $x(k)$ at time instant k and the previous inputs, $x(k-i)$ at time instant $k-i$ for $i = 1, 2, \dots, n-1$. The finite-duration impulse response (FIR) filter is a linear filter where the filtering output, $y(k)$ (refer Figure 1.1) is given by

$$\begin{aligned} y(k) &= \sum_{i=1}^{n} w_i(k)x(k-i+1) \\ &= \begin{pmatrix} w_1(k) & w_2(k) & \dots & w_n(k) \end{pmatrix} \begin{pmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-n+1) \end{pmatrix} \\ &= \mathbf{w}^T(k)\mathbf{x}(k) \ , \end{aligned} \tag{1.2}$$

in which $\mathbf{w}(k) = \begin{pmatrix} w_1(k) & w_2(k) & \dots & w_n(k) \end{pmatrix}^T$ is the column vector consists of the unknown filter coefficients, $w_i(k)$ at time instant k for $i = 1, 2, \dots, n$. The FIR filtering output given by (1.2) is having filter order of $n-1$, with the filter length of n. By the means of adaptive filtering, these filter coefficients are produced to minimize the objective function given by (1.1). From (1.2), the FIR filter is a feedforward and non-recursive filter, where the filtering output does not involve any feedback mechanism (i.e. the computation does not involve previous filtering outputs). Due to this non-recursive structure which only involves feedforward mechanism, FIR filter has a quadratic MSE function with only one minimum point; in which convergence to the global minimum point is ensured. This is the reason why FIR filter is more commonly adopted (Farhang-Boroujeny, 1998).

1.2.2 Infinite-duration Impulse Response (IIR) Filter

Let $\mathbf{y}(k)$ be the column vector consists of the previous filtering outputs such that $\mathbf{y}(k) = \begin{pmatrix} y(k-1) & y(k-2) & \ldots & y(k-n+1) \end{pmatrix}^T$. If the filtering output involves the feedback mechanism, then the filtering output, $y(k)$ of an infinite-duration impulse response (IIR) filter is given by

$$\begin{aligned} y(k) &= \sum_{i=1}^{n} a_i(k)x(k-i+1) - \sum_{i=1}^{n-1} b_i(k)y(k-i) \\ &= \begin{pmatrix} a_1(k) & a_2(k) & \ldots & a_n(k) \end{pmatrix} \begin{pmatrix} x(k) \\ x(k-1) \\ \vdots \\ x(k-n+1) \end{pmatrix} \\ &\quad - \begin{pmatrix} b_1(k) & b_2(k) & \ldots & b_{n-1}(k) \end{pmatrix} \begin{pmatrix} y(k-1) \\ y(k-2) \\ \vdots \\ y(k-n+1) \end{pmatrix} \\ &= \mathbf{a}^T(k)\mathbf{x}(k) - \mathbf{b}^T(k)\mathbf{y}(k) \ , \end{aligned} \tag{1.3}$$

where $\mathbf{a}(k) = \begin{pmatrix} a_1(k) & a_2(k) & \ldots & a_n(k) \end{pmatrix}^T$ is the unknown feedforward filter coefficient column vector, and $\mathbf{b}(k) = \begin{pmatrix} b_1(k) & b_2(k) & \ldots & b_{n-1}(k) \end{pmatrix}^T$ is the unknown feedback filter coefficient column vector. From (1.3), the IIR filter is a recursive filter that consists of the feedback mechanism and this distinguishes it from the FIR filter. The impulse response duration of the IIR filter is infinitely long due to the feedback mechanism, where the computation of current filtering output involves previous filtering outputs. Different from the quadratic MSE function of a FIR filter, the performance function of an IIR filter has many local minima points; in which it may result in local minima convergence (i.e. not the desired global minimum point) and this is the reason why IIR filter is limited in use (Farhang-Boroujeny, 1998).

1.2.3 Discrete-Time Signal Representation and Transfer Function

Consider the IIR filtering output which is given by (1.3), where the FIR filtering output given by (1.2) is a special case of (1.3) when $\mathbf{b}(k)$ is the zero column vector. Rearranging the IIR filtering output as the sequence of delayed input signals, then (1.3) becomes

$$\begin{aligned} & y(k)+b_1(k)y(k-1)+b_2(k)y(k-2)+\ldots+b_{n-1}(k)y(k-n+1) \\ = & a_1(k)x(k)+a_2(k)x(k-1)+a_3(k)x(k-2)+\ldots+a_n(k)x(k-n+1) \ . \end{aligned} \tag{1.4}$$

Let z^{-1} be the unit-delay operator, so that it operates on $x(k)$ resulting $x(k-1)$, and also operates on $y(k)$ resulting $y(k-1)$. In the z-transform domain (Farhang-Boroujeny, 1998), then (1.4) becomes[1]

$$\begin{aligned} & Y(z)+b_1(k)z^{-1}Y(z)+b_2(k)z^{-2}Y(z)+\ldots+b_{n-1}(k)z^{-n+1}Y(z) \\ = & a_1(k)X(z)+a_2(k)z^{-1}X(z)+a_3(k)z^{-2}X(z)+\ldots+a_n(k)z^{-n+1}X(z) \ . \end{aligned} \tag{1.5}$$

By collecting the coefficients of $X(z)$ and $Y(z)$, then

$$Y(z)=\frac{a_1(k)+a_2(k)z^{-1}+a_3(k)z^{-2}+\ldots+a_n(k)z^{-n+1}}{1+b_1(k)z^{-1}+b_2(k)z^{-2}+\ldots+b_{n-1}(k)z^{-n+1}}X(z) \ . \tag{1.6}$$

From (1.6), the input-output filtering operation is given by

$$Y(z)=H(z)X(z) \ , \tag{1.7}$$

where

$$H(z)=\frac{Y(z)}{X(z)}=\frac{a_1(k)+a_2(k)z^{-1}+a_3(k)z^{-2}+\ldots+a_n(k)z^{-n+1}}{1+b_1(k)z^{-1}+b_2(k)z^{-2}+\ldots+b_{n-1}(k)z^{-n+1}}$$

[1]The z-transform properties that used to transform (1.4) to (1.5) are shown in the Appendix A.

is known as the transfer function[2]. Considering the filtering output, $y(k)$ is a linear combination of the filter coefficients and the inputs, (1.7) can be realized through the structure of a FIR filter or an IIR filter. By minimizing the objective function, filter coefficients are produced in the sense that the z-transform of filter coefficients resembles the transfer function, $Z\{w(k)\} = H(z)$.

1.3 Adaptive Filtering Configuration for System Identification

Based on the configuration of adaptive filtering in Figure 1.1, adaptive filter produces a filter coefficient column vector, $\mathbf{w}(k)$ such that the filtering output, $y(k)$ closely approximates the desired output, $d(k)$. Basically, adaptive filtering is categorized into four classes of application; namely, modeling, inverse modeling, linear prediction, and interference cancelation (Farhang-Boroujeny, 1998). Although these four classes of application are different in the set up, nevertheless they share one common feature: the error resulted from the difference of the desired output and the filtering output is used to adjust the filter coefficients so that it is updated according to some performance measure. Under the category of modeling, system identification will be the core application in this thesis among the other applications. The set up of system identification is depicted by Figure 1.2.

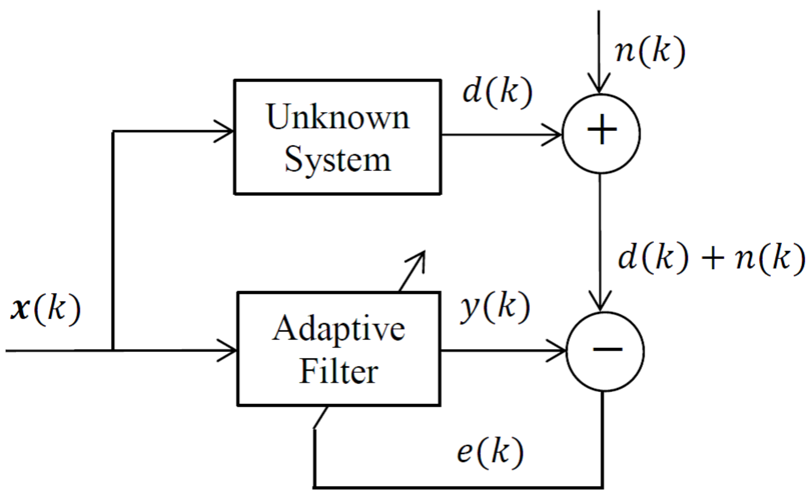

Figure 1.2: The flow diagram for system identification.

[2]The details of input-output characterization using impulse response, $h(k)$ and its z-transform, $Z\{h(k)\} = H(z)$ are shown in the last paragraph of Appendix A.

In literature, system identification is a diverse field that can be presented in various ways. Generally, it consists of three basic entities: (i) the data set, (ii) the model structure, and (iii) the model validation (Ljung, 1999). First of all, the input-output data of the system to be identified are collected, stored, and analyzed. The reason for doing so is to make these data maximally informative so that it could provide *a priori* information and increase the accuracy of estimation. If necessary, data is preprocessed to remove the deficiencies. A model is proposed after careful modeling. At this step, the proposed model is a tool which realize the unknown input-output mapping and it should be fit with the data set. At last, model validation takes place where the performance of the proposed model is monitored when the measured data is being reproduced. Succeeding these three entities, the model is arrived for system identification where it is ready for field run and perform the estimation.

Referring back to Table 1.1, there are differences in between statistical estimation and real time estimation. Similarly, these differences distinguish adaptive system identification (by adaptive filtering) from what is commonly regarded in literature. Under real time estimation, adaptive system identification performs as an on-line model. Instead of collecting, storing, and analyzing for its statistical information, the data are temporary stored and the distant past will be discarded. This is because the data are real time and keep coming in, therefore it is impossible to store all, not to mention to analyze it. In addition, the real time system is dynamic and the analysis is irrelevant due to the data are time-variant. On top of these, fast estimation with minimal computational load is also a concern to capture the fast dynamic of time varying environment.

Based on the configuration in Figure 1.2, a set of filter coefficients is estimated to characterize the unknown system model. Driven by a common set of inputs $\mathbf{x}(k)$, the unknown system supplies the desired output, $d(k)$ while the adaptive filter produces a set of filter coefficients in the sense that the filtering output, $y(k)$ closely resembles the desired output, $d(k)$. In other words, the filter coefficients produced by the adaptive filter should provide a good model that converges to the unknown system.

For a particular time instant, a desired output is obtained in response to the unknown system driven by a common set of inputs. In practice, the procedure of system identification[3] is iterative, not just for a particular time instant but continuous until a satisfactory model is built (or to some satisfactory performance measure). Therefore, the unknown system is dynamic and time varying and it is continuously be driven. The desired output will continuously be produced and as a result of discrete time measurements, a set of real valued data is formed in which it is a collective of desired outputs from a continuously driven system. Up to this point, adaptive filter serves as an online model and the produced filter coefficients are keep updated to the convergence of a good model (this is the reason why we term it as "real time estimation").

Let $\mathbf{x}(k)$ be a set of inputs up to the iteration of time instant k such that $\mathbf{x}(k) = \{x(k), x(k-1), \ldots, x(k-n+1)\}$, where n is the number of filter coefficients used in the adaptive filter. Driven by this common set of inputs, the unknown system supplies the desired output, $d(k)$ while the adaptive filter produces the filtering output, $y(k)$ with the estimated filter coefficients. The filtering output can be either the FIR filtering output given by (1.2) or the IIR filtering output given by (1.3). The difference between the desired output and the filtering output, $d(k) - y(k)$ defines the error signal, $e(k)$.

Referring to Figure 1.2, system identification is often constituted by the corruption of noise, $n(k)$ to the desired output. This noise, $n(k)$ is regarded as the measurement noise and often it is uncorrelated with the input signal. In the case where the order of the adaptive filter is sufficient to model the unknown system, the convergence leads to a zero MSE if there is no measurement noise. However, the measurement noise is unavoidable in practical applications. Under the

[3]Starting from here, all the system identification discussed is adaptive where the estimation is performed via adaptive filtering.

circumstance where the noise is uncorrelated with the input, the convergence of a good model leads to a minimum MSE in which it is the variance of noise, υ_n.

1.4 Literature Review

Let f be a mapping function such that

$$d = f(x_1, x_2, \ldots, x_n) \ , \tag{1.8}$$

where x_i for $i = 1, 2, \ldots, n$ are the inputs and d is the desired output. In adaptive filtering, the mapping function, f is realized through the implementation of either FIR or IIR filter. The implementation of filters is straightforward in conceptual manner, but the numerical issue takes place in practice. In literature, various analysis have been done to study the numerical properties of the adaptive filtering algorithm (Eleftheriou and Falconer, 1986; Ardalan, 1986; Cioffi, 1987; Bottomley and Alexander, 1991; Liavas and Regalia, 1999), as well as to develop a more stabilized variant algorithm (McWhirter, 1983; Alexander and Ghirnikar, 1993).

With the simplicity of the filter structure, the identification using a linear FIR filter are useful when the unknown system is completely characterized by the impulse response. However, the limitation occurs when the system itself is nonlinear in nature. To name a few, digital satellite channels equalization (Im, 1996), acoustic echo cancelation (Stenger et al., 1999), electrocardiogram signal extraction (Shadaydeh et al., 2008), and dynamic system identification (Chen et al., 2010) are the applications that require a nonlinear filter to describe the underlying dynamic and yield a better result. Due to the possibility of higher degree and dimensionality, the analytical nonlinear model is often unavailable and the generalization of its structure is difficult in practice. As a result, various forms of nonlinear adaptive filter exist (Pitas and Venetsanopoulos, 1990) and it is modeled according to the given system.

On the other hand, IIR filter characterizes the unknown system as the pole-zero system[4], where the transfer function involves both feedforward and feedback mechanism. Compared to the FIR filter, additional feedback mechanism is involved for the IIR filter where the present filtering output is dependent on the past filtering outputs. Due to this pole-zero property, IIR filter requires fewer coefficients than the FIR filter to achieve a similar performance. However, IIR filter encounters some issues which do not arise in the FIR filter. Analyses show that, IIR filtering may have biased estimation of coefficients, convergence to local minima, and unstable if the strictly positive real condition of the transfer function is violated (Johnson, 1984; Shynk, 1989; Netto et al., 1995).

1.4.1 Fuzzy System Modeling

In conventional modeling process, specific mathematical model is used to describe the system and numerical data are employed for verification in order to yield a desired result. In contrast to the conventional modeling, fuzzy system modeling is a new modeling paradigm that can utilize both quantitative and qualitative information. While the quantitative information is regarded as the numerical data, the qualitative information is given by human's expertise and knowledge (i.e. the linguistic information). By using fuzzy set, fuzzy logic, and fuzzy rules, fuzzy system incorporates the linguistic information and models them in an implicit linguistic form rather than in an analytical form compared to the conventional modeling (Ying, 2000). As a matter of fact, human expertise and knowledge are powerful where most of the daily life processes are done without any specific mathematical model. All the actions are intuitive but not numerical and fuzzy system modeling provides a platform to incorporate it.

The earliest fuzzy system development can be traced back as early as 1965 in the paper of *Fuzzy Sets* by Zadeh (1965). Compared to the classical set in which an element either belong

[4]The poles and zeros of the transfer function are shown in the Appendix B.

or not belong to a set, fuzzy set theory relieves the restriction by setting the degree of belonging of an element in the closed interval of [0,1]. Due to this transition from 0 to 1, an element can either not belong, sort of belong, partially belong, and etc. until totally belong to a particular set. Taking an example of using linguistic information to describe fast: intuitively it can be a little bit of fast, sort of fast, kind of fast, and very fast. By using the transition from 0 to 1, it can be represented by fuzzy sets and incorporated into the modeling.

By definition (Zadeh, 1965), a fuzzy set A in the universal of discourse X is characterized by a membership function $\mu_A(x)$ which associates with each point x in X a real number in the interval $[0,1]$, with the value of $\mu_A(x)$ at x representing the grade of membership of x in A. With the fuzzy set that calibrates linguistic vagueness into a proper mathematical representation, fuzzy logic comes into operation and models human's sense of words to a decision making process. This leads to the fuzzy algorithm in which is an ordered sequence of instructions and upon execution yields an approximate solution to a specified problem (Zadeh, 1973).

In general, fuzzy system comprises of four principle structures: fuzzification, fuzzy rule base, fuzzy inference engine, and defuzzification (Ying, 2000). With the incorporation of linguistic information, fuzzy system modeling provides a highly approximation to the systems which are too complex or too ill-defined to have a precise mathematical model. Unlike the conventional modeling, the four principle structures of fuzzy system are not a fixed numerical model, and it varies according to the system designer. To name a few, Zadeh (1965), Takagi and Sugeno (1985), Wang (1997), and Dubois and Prade (2000) are among the major contributors in the field of fuzzy system. Modeling in a highly approximate manner, fuzzy system has proved its credibility in various applications such as nonlinear channel equalization (Wang and Mendel, 1993b), image coding (Yu, 1998), nonlinear system modeling (Chen and Xi, 1998), telecommunication receiver (Hu et al., 2005), packet based voice system (Jones et al., 2006), stock market prediction (Sheta, 2006), education grading system (Bai and Chen, 2008), cancer classification (Pham, 2008), and speech coding (Johnny and Mirzaee, 2012).

1.4.2 Fuzzy System Modeling and Adaptive Filtering

In literature, fuzzy system modeling is employed to highly approximate a complex or ill-defined system using fuzzy sets. Given a set of inputs and a desired output, fuzzy system can be designed in two ways: with or without the incorporation of linguistic information. With the linguistic information, fuzzy rules are generated from the examples of operation handled by a human (Wang and Mendel, 1992b), and the obtained results can highly mimic the performance of a human operator. On the other hand, if the fuzzy system is designed without any linguistic information (where the fuzzy rules are meant to be trained), fuzzy system is similar to the adaptive network model (Jang, 1992, 1993), in which the data are processed numerically to yield a highly desirable result.

Although fuzzy system can be designed in numerous ways, there are some numerical studies on the fuzzy system based on the mathematical analysis. Numerous studies lend support to the claim that the fuzzy system are globally stable (Wang, 1993; Thathachar and Viswanath, 1997), uniformly convergent (Zeng and Singh, 1994), and approximate universally (Wang, 1992; Wang and Mendel, 1992a). In an overview study and comparison with various nonlinear black-box modeling techniques, fuzzy system is shown to have the structure of a general model, with the advantage of fuzzy rules to describe some possibly available prior knowledge (Sjoberg et al., 1995).

The earliest development of fuzzy system with learning ability (i.e. similar to adaptive filtering) is proposed by Takagi and Sugeno (1985), where the fuzzy consequent parameters are identified using the Kalman filter algorithm. Following the identification, another fuzzy system with learning ability is introduced by Wang and Mendel (1993b) where the recursive least-squares (RLS) algorithm is used to train the fuzzy system. Comparing with the Takagi and Sugeno's algorithm, Wang and Mendel's algorithm defines and includes the so-called fuzzy basis function, where it has been proven to approximate universally (Wang and Mendel, 1992a). It is noted that the RLS algorithm is an adaptive filtering algorithm and Wang and Mendel (1993b) has successfully employed the fuzzy system into nonlinear channel equalization, in which it is one of the many applications of adaptive filtering. Due to its adaptive learning capability, the RLS algorithm is also widely implemented in the applications of fuzzy consequent parameter identification (Mendel and Mouzouris, 1997; Chen and Xi, 1998; Aliaghasarghamish and Ebrahimi, 2011), Takagi-Sugeno fuzzy model identification (Pan et al., 2010), and interval fuzzy model identification (Khanesar et al., 2010). Furthermore, if the RLS algorithm is extended to the neural fuzzy system, its performance is proven to be competitive compared to the conventional backpropagation learning algorithm (Yeh et al., 2010, 2011; Yeh and Su, 2012).

1.5 Research Methodology

1.5.1 Motivation

The objective of adaptive filtering is to minimize the MSE of the difference between the desired output, $d(k)$ and the filtering output, $y(k)$. The desired output, $d(k)$ is supplied by the unknown system that driven by a set of inputs, $\mathbf{x}(k)$ and adaptive filter produces a set of filter coefficients in the sense that the filtering output, $y(k)$ closely resembles the desired output, $d(k)$. This process can be view as the input-output filtering operation given by (1.7) where the transfer function is an approximation to the unknown system. In practice, the structure of the transfer function is unknown and prior assumption is made for the input-output filtering operation. It is assumed that the transfer function has a known structure whereby it can be realized by three different structures which are commonly used in conventional adaptive filtering: (i) the FIR filter, (ii) the IIR filter, and (iii) the nonlinear filter. However, there are limitations when realizing each structures in practice, as described in previous section of literature review. Motivated by the fuzzy approach, this thesis seeks to investigate the capabilities of the fuzzy system in overcoming the limitations of conventional adaptive filters. As stated in Section 1.4.2, fuzzy system are globally stable, uniformly convergent, and able to approximate universally. Most importantly, the structure of the fuzzy system with learning ability is similar to adaptive filtering and fuzzy mapping can be used to realize the unknown input-output relationship in adaptive filtering without prior assumption on the adaptive filter structure.

1.5.2 Problem Statements

In literature, nonlinear adaptive filtering has been tackle by Wang and Mendel (1993a) in their paper: *Fuzzy Adaptive Filters, with Application to Nonlinear Channel Equalization*. Although Wang and Mendel's work indeed provide a solution to nonlinear adaptive filtering, but at the same time it creates a research gap to the connection in between adaptive filtering and fuzzy system modeling. Based on the previous works by Wang and Mendel, this thesis aims to address the following questions in which arise as the research gap of their work.

In Wang and Mendel's works,

(i) Gaussian membership function is employed to form and define the fuzzy basis function. Can any other form of membership function such as Triangular-shaped membership function be employed to have lower computational cost, better fuzzy partitioning, and a good trade off in nonlinear estimation?

(ii) The proposed algorithm is in the framework of application where the work does not include in depth analysis of the mathematical properties. Can adaptive filtering with fuzzy system be further established with numerical analysis? And its behavior be explained mathematically using the properties of adaptive filtering?

(iii) Fuzzy system is incorporated into adaptive filtering algorithm. Will such incorporation address a better performance in terms of error behavior? And are the common issues of adaptive filtering such as the tracking behavior, the quantization effect, and the nonlinear estimation instability be tackled?

(iv) The proposed algorithm is tested on the application of nonlinear channel equalization. Can it be extended to nonlinear or feedback system identification? And as a different realization besides the characterization using FIR and IIR filter?

1.5.3 Objectives

Among all the various fuzzy models proposed in literature, our research direction is focus on the fuzzy system with learning ability to identify a given unknown system. Based on the Wang and Mendel's algorithm, our research objective is to incorporate fuzzy system into adaptive filtering and our research direction is in two folds: to develop a better algorithm, and perform mathematical analysis on it. Specifically,

(i) To establish a connection between adaptive filtering and fuzzy system modeling and address the incorporation of fuzzy system into existing deterministic approach;

(ii) To propose a reduced fuzzy basis function, which utilizes the Triangular-shaped membership function and the Ruspini partitioning to reduce computational complexity as well as to have a strong fuzzy partitioning and a complete rule base method;

(iii) To prove the universal approximation of the proposed method by using mathematical theorems;

(iv) To study the effect of fuzzification and fuzzy partitioning on the error behavior of the developed algorithm; and

(v) To test the developed algorithm in identifying different structures of adaptive filter realization and compare it with existing RLS-based adaptive filtering algorithm, including the implementation in real time applications.

1.5.4 Contributions

In this thesis, a reduced fuzzy recursive least-squares (FRLS) algorithm is developed by introducing the reduced fuzzy basis function. Due to the development of the reduced FRLS algorithm is rather general in nature, the contribution of this research can be considered as providing an alternative to the existing literature with wide varieties of application.

To be specific in this thesis:

(i) A new reduced FRLS algorithm is proposed by developing the reduced fuzzy basis function. This new algorithm possesses lower computational complexity compared to the existing algorithm by Wang and Mendel and proven to be no compromise on the universal approximation property. In addition, the reduced fuzzy basis function is proven to be a complete rule base method and has a strong fuzzy partitioning.

(ii) Various analysis such as the principle of orthogonality, the error conversion factor, and the error behavior of the algorithm are provided and the research gap left by Wang and Mendel is filled by establishing the connection between adaptive filtering and fuzzy system modeling.

(iii) Compared to conventional adaptive filter realization, the incorporation of reduced fuzzy basis function in adaptive filtering is a different realization using fuzzy system. With this special realization, the reduced FRLS algorithm is shown to have the ability to handle the instability issue that normally encounter by a nonlinear filter or an IIR filter. Significantly, it results in a more stable condition number and its performance is comparable with the QR decomposition (QRD) based algorithm, in which the QRD based algorithm is well known for its stability when dealing with finite precision arithmetic.

(iv) Triangular-shaped membership function is employed in the reduced FRLS algorithm and it has significant advantage when dealing with low nonlinearity system. Although fuzzification by a Triangular-shaped membership function is linear, it doesn't degrade when dealing with high nonlinearity system and it makes no significant difference compared to the existing algorithm by Wang and Mendel that employed Gaussian membership function.

(v) Viability of the proposed algorithm in real life application is shown where the application of adaptive filtering in speech processing – linear predictive coding model with integrated reduced FRLS algorithm is demonstrated.

1.6 Flow of Thesis

The objective of this thesis is to develop the reduced FLRS algorithm, perform various mathematical analysis, and simulate it with real life applications. Basically, this thesis can be divided into few phases: (i) the introductory, (ii) the methodologies, (iii) the development and analysis of the proposed algorithm, (iv) the simulations, and lastly (v) the summary and conclusion.

In the first phase of introductory, Chapter 1: Introduction is included in this phase where the general knowledge regarding adaptive filtering is introduced. This chapter also provides insight on the problems faced by a nonlinear filter and a feedback filter. Motivated by Wang and Mendel's work, the objectives and the expected contributions of this research are stated and lastly, the flow of thesis marks the end of the introductory phase.

Chapter 2: Adaptive Filtering and Recursive Least-Squares and Chapter 3: Fuzzy System made up the second phase of methodologies. In Chapter 2, the RLS algorithm and its related methodologies are included. The derivation of the Wiener filter subsequently lead to the formulation of the RLS algorithm and its properties. To address the adaptive filtering problem when working in finite precision arithmetic, the QRD based algorithm is included; as well as the inclusion of the SOV-RLS algorithm and the Bilinear RLS algorithm to address the instability when dealing with nonlinearity and the feedback mechanism. In Chapter 3, fuzzy system and its methodologies are included where the fuzzy set is defined and its related axioms and operations are shown. Comprised of four principle structures: fuzzification, fuzzy rule base, fuzzy inference engine, and defuzzification, fuzzy system is revealed at the end of chapter on how the crisp output can be obtained.

Chapter 4: Fuzzy Recursive Least-Squares explores the combination of both approaches: the adaptive filtering and the fuzzy system. In this chapter, the reduced FRLS algorithm is developed and analysis based on the mathematical derivations are done to reveal its special properties. The generalization of the transfer function is first derived and fuzzy mapping is employed to solve the parameters in the transformed domain of fuzzy system. Fuzzification using the Triangular-shape membership function and Ruspini partitioning lead to the introduction of reduced fuzzy basis function and as a result, the reduced FRLS algorithm is developed. With lower computational cost, reduced fuzzy basis function is proven to be a universal approximator and possess strong fuzzy properties. Lastly, the principle of orthogonality, the error conversion factor, and the error behavior are formulated and subsequently show the special properties of the reduced FRLS algorithm.

Chapter 5: Simulations and Chapter 6: Real Life Applications made up the fourth phase of simulations. Various simulations are performed in Chapter 5 to verify the theoretical analysis of the reduced FRLS algorithm. The results are promising and agree with the mathematical derivations where the reduced FRLS algorithm is shown comparable to existing adaptive filtering algorithms. Significantly, the reduced FRLS algorithm shows stable performance under finite precision working environment and it is comparable with the QRD based algorithms. Before the concluding remark, the reduced FRLS algorithm is applied using real time data set in Chapter 6. Real life application is performed on the speech system identification where the reduced FRLS algorithm is integrated to have the FRLS-LPC model. The algorithm is shown viable and even more, the FRLS-LPC model is shown to perform well in the environment where the conventional predictive coding models have failed.

Lastly, Chapter 7: Conclusion marks the final phase of summary and conclusion. All the significant results that are developed in this thesis is summarized, followed by the algorithm review and conclusion. Some suggestions are included in the very last part of this thesis for future research direction and to extend current proposed algorithm.

CHAPTER 2

ADAPTIVE FILTERING AND RECURSIVE LEAST-SQUARES

2.1 Principles of Adaptive Filtering

In Chapter 1, some important information such as the adaptive filter structure and the adaptation approach are generally described to give insights into adaptive filtering. In this chapter, the principles of adaptive filtering and the mathematical formulations are described. This includes the derivation of the so-called "Wiener filter" and its solution, as well as the principle of orthogonality. It is important to note that, adaptive filtering evolves around the process of minimizing the mean-square error (MSE), ξ as describe in (1.1) and the derivation of Wiener solution gives the optimum filter coefficients which minimize the MSE. The MSE is formulated in the form of the statistical expectation of the error, $E[e^2(k)]$ which requires a large number of signals realization to produce its exact value. In real time processing such as adaptive filtering, only an estimated value of the MSE is possible to achieve (refer discussion in Section 1.1). To produce an approximation of the Wiener solution, computational methods are based on the type of approximation used to realize the MSE. Two popular approaches (along with their properties) are outlined here: (i) the stochastic approach via the least-mean-square (LMS) algorithm, and (ii) the deterministic approach via the recursive least-squares (RLS) algorithm. A more stable version of the RLS algorithm, namely the QR decomposition based RLS algorithm is also described. The LMS and RLS algorithms are primarily designed to handle linear filtering problems. In order to draw the relevance to the nonlinear and feedback problems, extensions of RLS algorithm to nonlinear and feedback adaptive filtering are discussed towards the end of this chapter.

2.1.1 Wiener Filter

In adaptive filtering, the objective is to minimize the performance function, ξ. Stochastically, the performance function is measured using the MSE and it is given by

$$\xi = E[e^2(k)] = E[d^2(k) - 2d(k)y(k) + y^2(k)] \ , \tag{2.1}$$

where $e(k) = d(k) - y(k)$ is the error between the desired output and the filtering output and $E[.]$ is the statistical expectation[1] of the given equation.

Consider a FIR filter with filter length n (i.e. filter order $n-1$) produces a filtering output given by (1.2), $y(k) = \mathbf{w}^T(k)\mathbf{x}(k)$ where the filtering output, $y(k)$ is a linear combination of the inputs, $x(k-i+1)$ and the filter coefficients, $w_i(k)$ at time instant k for $i = 1, 2, \ldots, n$. By substituting (1.2) into (2.1), the MSE function becomes

$$\begin{aligned} \xi &= E[d^2(k) - 2d(k)\mathbf{w}^T(k)\mathbf{x}(k) + \mathbf{w}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(k) \\ &= E[d^2(k)] - 2\mathbf{w}_o{}^T E[d(k)\mathbf{x}(k)] + \mathbf{w}_o{}^T E[\mathbf{x}(k)\mathbf{x}^T(k)]\mathbf{w}_o \ , \end{aligned} \tag{2.2}$$

where $E[\mathbf{w}(k)] = \mathbf{w}_o$. Here, $\mathbf{w}_o = \begin{pmatrix} w_1 & w_2 & \ldots & w_n \end{pmatrix}^T$ is the optimum filter coefficient column vector consists of all the optimum filter coefficients in which our objective is to minimize ξ with respect to $\mathbf{w}_o$. By the term of (2.2), we can see that ξ is a quadratic function with respect to $\mathbf{w}_o$ and the single global minimum point is achieved when the gradient is equal to zero. By

[1] Given a statistical variable x, the statistical expectation is defined as $E[x] = \int_{-\infty}^{\infty} x p_x dx$ where p_x is the probability density function of x. For the case of a discrete-time wide-sense stationary random variable $x(k)$ in which generated by ergodic discrete-time random real valued processes, then $E[x(k)] = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x(k)$ (Zaknich, 2005).