Global Journal of Computer Science and Technology
Volume 11 Issue 7 Version 1.0 May 2011
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
ISSN: 0975-4172 & Print ISSN: 0975-4350

Improve Speech Enhancement Using Weiner Filtering

By S.China Venkateswarlu, Dr. K.Satya Prasad, Dr. A.SubbaRami Reddy

Abstract- Speech enhancement aims to improve speech quality by using various algorithms. It may sound simple, but what is meant by the word quality. It can be at least clarity and intelligibility, pleasantness, or compatibility with some other method in speech processing. Wiener filter are rather simple and workable, but after the estimation of the background noise, one neglects the fact that the signal is actually speech. Furthermore, the phase component of the signal is left untouched. However, this is perhaps not such a bad problem; after all, human ear is not very sensitive to phase changes. The third restriction in spectral subtraction methods is the processing of the speech signal in frames, so the Proceeding from one frame to another must be handled with care to avoid discontinuities. Noise reduction is a key-point of speech enhancement systems in hands-free communications. A number of techniques have been already developed in the frequency domain such as an optimal short-time spectral amplitude estimator proposed by Ephraim and Malah including the estimation of the a priori signal-to-noise ratio. This approach reduces significantly the disturbing noise and provides enhanced speech with colorless residual noise. In this paper, we propose a technique based on a Wiener filtering under uncertainty of signal presence in the noisy observation. Two different estimators of the a priori signal-to-noise ratio are tested and compared. The main interest of this approach comes from its low complexity. In this paper we demonstrate the application of weiner filter for a speech signal using Matlab 7.1 and signal processing toolbox.

Keywords: Communication, Enhancement, Intelligibility, Matlab, Speech, Wiener filter

GJCST Classification: H.5.2, I.2.7

IMPROVE SPEECH ENHANCEMENT USING WEINER FILTERING

Strictly as per the compliance and regulations of:

Improve Speech Enhancement Using Weiner Filtering

S.China Venkateswarlu[α], Dr. K.Satya Prasad[Ω], Dr. A.SubbaRami Reddy[β]

Abstract- Speech enhancement aims to improve speech quality by using various algorithms. It may sound simple, but what is meant by the word quality. It can be at least clarity and intelligibility, pleasantness, or compatibility with some other method in speech processing. Wiener filter are rather simple and workable, but after the estimation of the background noise, one neglects the fact that the signal is actually speech. Furthermore, the phase component of the signal is left untouched. However, this is perhaps not such a bad problem; after all, human ear is not very sensitive to phase changes. The third restriction in spectral subtraction methods is the processing of the speech signal in frames, so the Proceeding from one frame to another must be handled with care to avoid discontinuities. Noise reduction is a key-point of speech enhancement systems in hands-free communications. A number of techniques have been already developed in the frequency domain such as an optimal short-time spectral amplitude estimator proposed by Ephraim and Malah including the estimation of the a priori signal-to-noise ratio. This approach reduces significantly the disturbing noise and provides enhanced speech with colorless residual noise. In this paper, we propose a technique based on a Wiener filtering under uncertainty of signal presence in the noisy observation. Two different estimators of the a priori signal-to-noise ratio are tested and compared. The main interest of this approach comes from its low complexity. In this paper we demonstrate the application of weiner filter for a speech signal using Matlab 7.1 and signal processing toolbox.

Keywords- Communication, Enhancement, Intelligibility, Matlab, Speech, Wiener filter

I. Introduction

Speech is a form of communication in every day life. It existed since human civilizations began and even till now, speech is applied to high technological telecommunication systems. A particular field which I personally feel will excel will be speech signal processing in the world of telecommunications. As applications like cellular and satellite technology are getting popular among mankind, human beings tend to demand more advance technology and are in search of improved applications. For this reason, researchers are looking closely into the four generic attributes of speech coding. They are complexity, quality, bit rate and delay. Other issues like robustness to transmission errors, multistage encoding/decoding, and accommodation of non-voice signals such as in-band signaling and voice band modem data play an important role in coding of speech as well .

In order to understand these processes, both human and machine speeches have to be studied carefully on the structures and functions of spoken language:

How we produce and perceive it and how speech technology may assist us in communication. Therefore in this project, we will be looking more into speech processing with the aid of an interesting technology known as the Weiner Filter for speech processing. Presently, this technique is widely used in the field of signal processing. Speech processing has been a growing and dynamic field for more than two decades and there is every indication that this growth will continue and even accelerate. During this growth there has been a close relationship between the development of new algorithms and theoretical results, new filtering techniques are also of consideration to the success of speech processing. One of the common adaptive filtering techniques that are applied to speech is the Wiener filter. This filter is capable of estimating errors however at only very slow computations. On the other hand, the Kalman filter suppresses this disadvantage.

As widely known to the world, weiner filtering techniques are used on GPS (Global Positioning System) and INS (Inertial Navigation System). They are widely used for speech signal coding applications. Due to its accurate estimation characteristic, electrical engineers are picturing the Weiner filter as a design tool for speech, whereby it can estimate and resolve errors that are contained in speech after passing through a distorted channel. Due to this motivating fact, there are many ways a Weiner filter can be tuned to suit engineering applications such as network telephony and even satellite phone conferencing. Knowing the fact

The work is carried out through the research facility at the Department of Computer Science & Engineering and Department of Electronics & Communication Engineering, HITS College of Engineering, Bogaram(V), Keesara(M), R.R.District,A.P., and Dept., ECE/CSE from JNTUK. The Authors also would like to thank the authorities of JNTU Kakinada, SKIT Chittoor for encouraging this research work.

About[α]- HITS College of Engineering is a RESEARCH SCHOLAR, from JNTU Kakinada, A.P., INDIA., RESEARCH SCHOLAR ROLL NO: 09022P0426, phone: 08415-242500, fax: 08415-24500252, Mobile No: +91 9666819730, +91 9030076793;
E-mail: cvenkateswarlus@gmail.com
About[Ω]- Professor & DE from JNTU Kakinada, Kakinada.and Research Supervisor from JNTUCollege of Engineering Kakinada..
E-mail: prasad_kodati@yahoo.co.in.
About[β]- Professor & Principal from SKIT-Chittor.and Research Supervisor from JNTUK.
E mail: principalskit@gmail.com

that preserving information, which is contained in speech, is of extreme importance, the availability of signal filters such as the Weiner filter is of great importance.

In this paper, the primary goal is to design a MATLAB based simulator for processing of speech together with the aid of the Weiner filtering technique and to obtain a reconstructed speech signal, which is similar to the input speech signal. To achieve these results, sample speeches were obtained. These were modeled as an autoregressive (AR) process and represented in the state-space domain by the Weiner filter. The original speech signal and the reconstructed speech signal obtained from the output of the filter were compared. The idea of this comparison is to pursue an output speech signal, which is similar to the original one. It was concluded that Weiner filtering is a good constructing method for speech.

II. Speech Processing

The term speech processing basically refers to the scientific discipline concerning the analysis and processing of speech signals in order to achieve the best benefit in various practical scenarios. The field of speech processing is, at present, undergoing a rapid growth in terms of both performance and applications. This is stimulated by the advances being made in the field of microelectronics, computation and algorithm design. Nevertheless, speech processing still covers an extremely broad area, which relates to the following three engineering applications:

- Speech Coding and transmission that is mainly concerned with man-to man voice Communication
- Speech Synthesis which deals with machine-to-man communications;
- Speech Recognition relating to man-to machine communication.

a) *Speech Coding*

Speech coding or compression is the field concerned with compact digital representations of speech signals for the purpose of efficient transmission or storage. The central objective is to represent a signal with a minimum number of bits while maintaining perceptual quality. Current applications for speech and audio coding algorithms include cellular and personal communications networks (PCNs), teleconferencing, desktop multi-media systems, and secure communications.

b) *Speech Synthesis*

The process that involves the conversion of a command sequence or input text (words or sentences) into speech waveform using algorithms and previously coded speech data is known as speech synthesis. The inputting of text can be processed through by keyboard, optical character recognition, or from a previously stored database. A speech synthesizer can be characterized by the size of the speech units they concatenate to yield the output speech as well as by the method used to code, store and synthesize the speech. If large speech units are involved, such as phrases and sentences, high-quality output speech (with large memory requirements) can be achieved. On the contrary, efficient coding methods can be used for reducing memory needs, but these usually degrade speech quality.

c) *Speech Recognition*

Speech or voice recognition is the ability of a machine or program to recognize and carry out voice commands or take dictation. On the whole, speech recognition involves the ability to match a voice pattern against a provided or acquired vocabulary. A limited vocabulary is mostly provided with a product and the user can record additional words. On the other hand, sophisticated software has the ability to accept natural speech (meaning speech as we usually speak it rather than carefully-spoken speech).

Speech information can be observed and processed only in the form of sound waveforms. It is an essential for speech signal to be reconstructed properly.

d) *Speech Production*

Speech is the acoustic product of voluntary and well-controlled movement of a vocal mechanism of a human. During the generation of speech, air is inhaled into the human lungs by expanding the rib cage and drawing it in via the nasal cavity, velum and trachea it is then expelled back into the air by contracting the rib cage and increasing the lung pressure. During the expulsion of air, the air travels from the lungs and passes through vocal cords which are the two symmetric pieces of ligaments and muscles located in the larynx on the trachea. Speech is produced by the vibration of the vocal cords. Before the expulsion of air, the larynx is initially closed. When the pressure produced by the expelled air is sufficient, the vocal cords are pushed apart, allowing air to pass through. The vocal cords close upon the decrease in air flow. This relaxation cycle is repeated with generation frequencies in the range of 80Hz – 300Hz. The generation of this frequency depends on the speaker's age, sex, stress and emotions. This succession of the glottis openings and closure generates quasi-periodic pulses of air after the vocal cords.

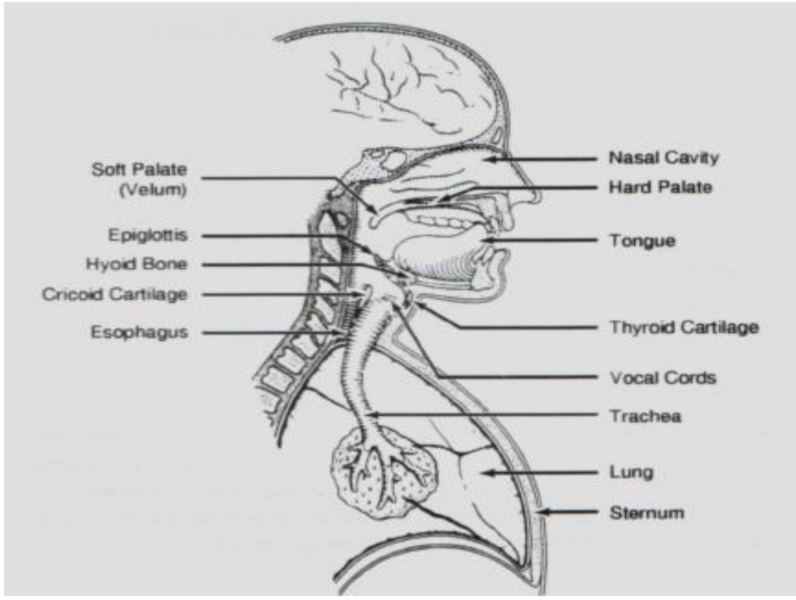

Figure1: Speech -acoustic product of voluntary and well controlled movement of a vocal mechanism of a human

The speech signal is a time varying signal whose signal characteristics represent the different speech sounds produced. There are three ways of labelling events in speech. First is the silence state in which no speech is produced. Second state is the unvoiced state in which the vocal cords are not vibrating, thus the output speech waveform is aperiodic and random in nature. The last state is the voiced state in which the vocal cords are vibrating periodically when air is expelled from the lungs. This results in the output speech being quasi-periodic- shows a speech waveform with unvoiced and voiced state.

Speech is produced as a sequence of sounds. The type of sound produced depends on shape of the vocal tract. The vocal tract starts from the opening of the vocal cords to the end of the lips. Its cross sectional area depends on the position of the tongue, lips, jaw and velum. Therefore the tongue, lips, jaw and velum play an important part in the production of speech.

III. Hearing and Perception

Audible sounds are transmitted to the human ears through the vibration of the particles in the air. Human ears consist of three parts, the outer ear, the middle ear and the inner ear. The function of the outer ear is to direct speech pressure variations toward the eardrum where the middle ear converts the pressure variations into mechanical motion. The mechanical motion is then transmitted to the inner ear, which transforms these motion into electrical potentials that passes through the auditory nerve, cortex and then to the brain . Figure below shows the schematic diagram of the human ear.

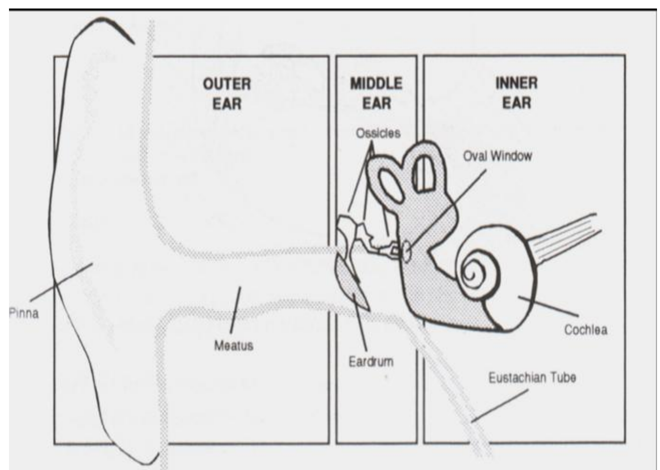

Figure2: The schematic diagram of the human ear

IV. Speech Waveform

The speech waveform needs to be converted into digital format before it is suitable for processing in the speech recognition system. The raw speech waveform is in the analog format before conversion. The conversion of analog signal to digital signal involves three phases, mainly the sampling, quantisation and coding phase. In the sampling phase, the analog signal is being transformed from a waveform that is continuous in time to a discrete signal. A discrete signal refers to the sequence of samples that are discrete in time. In the quantisation phase, an approximate sampled value of a variable is converted into one of the finite values contained in a code set. These two stages allow the speech waveform to be represented by a sequence of values with each of these values belonging to the set of finites values. After passing through the sampling and quantization stage, the signal is then coded in the

Global Journal of Computer Science and Technology Volume XI Issue VII Version I May 2011

May 2011

Global Journal of Computer Science and Technology Volume XI Issue VII Version I

coding phase. The signal is usually represented by binary code. These three phases needs to be carried out with caution as any miscalculations, over-sampling and quantization noise will result in loss of information. Below are the problems faced by the three phases.

V. Single and Multi-Channel Enhancement

Single channel methods operate on the input obtained from only one microphone. They have been attractive due to cost and size factors, especially in mobile communications. In contrast, multi-channel methods employ an array of two or more microphones to record the noisy signal and exploit the resulting spatial diversity. The two approaches are not necessarily independent, and can be combined to improve performance. For example, in practical diffuse noise environments, the multi-channel enhancement schemes rely on a single-channel post-filter to provide additional noise reduction. We discuss single-channel methods and introduce the contributions of this project towards this area is also included in this document. This section is intended to be a survey on single-channel enhancement algorithms.

VI. Maximum-Likelihood Estimation

Consider the estimation of a parameter μ = [μ1 : : : μp]^T based on a sequence of K observations y = [y(0) : : : y(Ki1)]^T . In ML estimation, μ is treated as a deterministic variable. The ML estimate of μ is the value μ ML that maximizes the likelihood function p(y; μ) defined on the data. ML estimation has several favorable properties, in particular, it is asymptotically unbiased and efficient, i.e., as the number of observations K tends to infinity, the ML estimate is unbiased and achieves the Cramer-Rao lower bound (CRLB). It can be shown that

$$\theta^{\mathrm{ML}} \underset{K\to\infty}{\sim} \mathcal{N}(\theta, \mathbf{I}^{-1}(\theta)), \tag{1}$$

where I(μ) is the p £ p Fisher information matrix whose (i; j)th entry is given by

$$[\mathbf{I}(\theta)]_{ij} = -\mathrm{E}\left[\frac{\partial^2 \ln p(\mathbf{y};\theta)}{\partial\theta_i \partial\theta_j}\right] \tag{2}$$

Thus we have (asymptotically)

$$\begin{aligned} &\text{Unbiased: } \mathrm{E}[\theta^{\mathrm{ML}}] = \theta, \\ &\quad\text{CRLB: } \mathrm{var}(\theta_i^{\mathrm{ML}}) = [\mathbf{I}^{-1}(\theta)]_{ii}. \end{aligned} \tag{3}$$

The maximization of the likelihood function is performed over the domain of μ. In many cases, μ ML cannot be computed in closed form and a numerical solution is obtained instead. Such numerical solutions are typically obtained through iterative maximization procedures such as the Newton-Raphson method or the expectation-maximization (EM) approach. The initial value of the parameter used to start the iterative procedure usually has a strong impact on whether the final estimate results in a local or a global maximum of the likelihood function.

In applications where the parameter μ is known to assume one of a finite set of values, the problems due to the iterative procedures can be avoided by performing the maximization over this finite set. An exhaustive search over the finite parameter space guarantees a global maximum. For speech enhancement, we assume that both speech and noise can be described by independent auto-regressive (AR) processes. The problem is then one of estimating the speech and noise LP coefficients based on the observed noisy speech in an ML framework. The clean speech AR model can be mathematically expressed as

$$x(n) = \sum_{l=1}^{p} a_l x(n-l) + e(n), \tag{4}$$

where a1…. ap are the LP coefficients of order p and e(n) is the prediction error, also referred to as the excitation signal. It is common to model e(n) as a Gaussian random process. The LP analysis is typically performed for each frame of 20-30 ms, within which speech can be assumed to be stationary.

For each frame, the model parameters are the vector of LP coefficients μ = [a1 : : : ap], and the variance of the excitation signal. A similar model can be obtained for the noise signal. The physiology of speech production imposes a constraint on the possible shapes of the speech spectral envelope. Since the spectral envelope is specified by the LP coefficients, this knowledge can be modeled using a sufficiently large codebook of speech LP coefficients obtained from long sequences of training data. Such a-priori information about the LP coefficients of speech has been exploited successfully in speech coding using trained codebooks.

Similarly, noise LP coefficients can also be modeled based on training sequences for different noise types. Thus, it is sufficient to perform the maximization over the speech and noise codebooks.

We characterize the speech and noise power spectra, which can be used to construct a Wiener filter to obtain the enhanced speech signal. Given the noisy data, the excitation variances maximizing the likelihood are determined for each pair of speech and noise LP coefficients from the codebooks. This is done for all combinations of codebook pairs, and the most likely codebook combination, together with the optimal excitation variances, is obtained. Since this optimization is performed on a frame-by-frame basis, good performance is achieved in non-stationary noise environments.

Apart from restricting the search space, using a codebook in the ML estimation has an additional benefit in applications where a codebook index needs to be transmitted over a network, e.g., in speech coding. In this case, the likelihood function can be interpreted as a modified distortion criterion.

VII. Bayesian MMSE Estimation

In ML estimation, the parameter μ is treated as a deterministic but unknown constant. In the Bayesian approach, μ is treated as a random variable. The Bayesian methodology allows us to incorporate prior (before observing the data) knowledge about the parameter by assigning a prior pdf to μ.

A cost function is formulated and its expected value, referred to as the Bayesian risk, is minimized. A commonly used cost function is the mean squared error (MSE).

In this case, the Bayesian minimum mean squared error (MMSE) estimate μ^{BY} of μ given the observations y is obtained by minimizing E[(μ i μ^{BY})2], where E is the statistical expectation operator. The expectation is with respect to the joint distribution p(y; μ). Thus, the cost function to be minimized can be written as

$$\begin{aligned}\eta &= \mathrm{E}[(\theta-\theta^{\mathrm{BY}})^2] \\ &= \int\int(\theta-\theta^{\mathrm{BY}})^2 p(\mathbf{y},\theta)d\mathbf{y}d\theta \\ &= \int\left(\int(\theta-\theta^{\mathrm{BY}})^2 p(\theta|\mathbf{y})d\theta\right)p(\mathbf{y})d\mathbf{y},\end{aligned} \tag{5}$$

where the posterior pdf p(μjy) is the pdf of μ after the observation of data. Since p(y) ¸ 0, it is su±cient to minimize the inner integral for each y. An estimate of μ can be found by determining a stationary point of the cost function (setting the derivative of the inner integral to zero). We can write

$$\frac{\partial}{\partial\theta^{\mathrm{BY}}}\int(\theta-\theta^{\mathrm{BY}})^2 p(\theta|\mathbf{y})d\theta = 0 \tag{6}$$

$$\theta^{\mathrm{BY}} = \int\theta p(\theta|\mathbf{y})d\theta = \mathrm{E}[\theta|\mathbf{y}]. \tag{7}$$

E[μjY = y], where y is a realization of the corresponding random variable Y. Using Bayes' rule, the posterior pdf can be written as

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)p(\theta)}{p(\mathbf{y})}, \tag{8}$$

where the denominator p(y) is a normalizing factor, independent of the parameter μ.

We describe a method to obtain Bayesian MMSE estimates of the speech and noise AR parameters. The respective prior pdfs are modeled by codebooks. The integral in (7) is replaced by a summation over the codebook entries. We also consider MMSE estimation of functions of the AR parameters, and one such function is shown to result in the MMSE estimate of the clean speech signal, given the noisy speech. As in the ML case, MMSE estimates of the speech and noise AR parameters are obtained on a frame-by-frame basis, ensuring good performance in non stationary noise.

In the ML estimation framework, one pair of speech and noise codebook vectors was selected as the ML estimate, whereas the Bayesian approach results in a weighted sum of the speech (noise) codebook vectors. The Bayesian method provides a framework to account for both the knowledge provided by the observed data and the prior knowledge.

VIII. Single-Channel Speech Enhancement

Single-channel speech enhancement systems obtain the input signal using only one microphone. This is in contrast to multi-channel systems where the presence of two or more microphones enables both spatial and temporal processing. Single-channel approaches are relevant due to cost and size factors. They achieve noise reduction by exploiting the spectral diversity between the speech and noise signals. Since the frequency spectra of speech and noise often overlap, single-channel methods generally achieve noise reduction at the expense of speech distortion.

The reduction of background noise using single-channel methods requires an estimate of the noise statistics. Early approaches were based on voice activity detectors (VAD), and noise estimates were updated during periods of speech inactivity. Accuracy deteriorates with decreasing signal-to-noise ratios (SNR) and in non-stationary noise. Soft-decision VADs, update the noise statistics even during speech activity.

Since single-channel methods exploit the spectral diversity between the speech and noise signals, it is therefore natural to perform the processing in the frequency domain. Processing is done on short segments of the speech signal, typically of the order of 20 to 30 ms, to ensure that the speech signal satisfies assumptions of wide-sense stationary. The segmentation is performed using a sliding window of finite support. The windowed signal (assuming it is absolute sum able) is transformed to the frequency domain using the discrete short-time Fourier transform (STFT)

$$X_m(k) = \frac{1}{\sqrt{K}} \sum_{n=-\infty}^{\infty} x(n)h(n-m)\exp(-j\frac{2\pi}{K}kn),\ k = 0,1,\ldots,K-1, \ (9) \qquad (9)$$

1. Math and computation
2. Algorithm development
3. Data acquisition
4. Data analysis ,exploration ands visualization
5. Scientific and engineering graphics

MATLAB displays graphs in a special window known as a figure. To create a graph, you need to define a coordinate system. Therefore every graph is placed within axes, which are contained by the figure. The actual visual representation of the data is achieved with graphics objects like lines and surfaces. These objects are drawn within the coordinate system defined by the axes, which MATLAB automatically creates specifically to accommodate the range of the data. The actual data is stored as properties of the graphics objects.

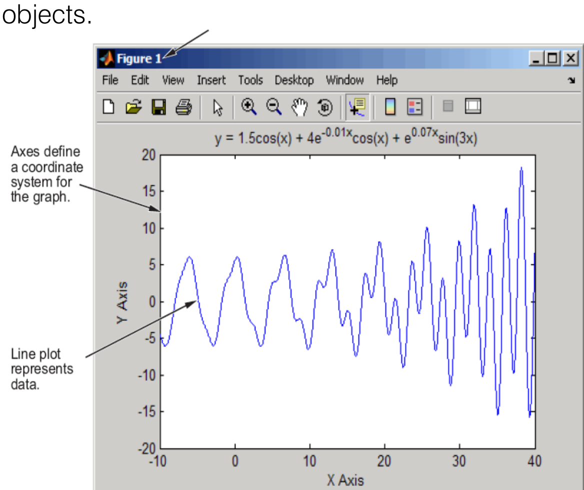

Figure3: Graphical objects

a) *Plotting Tools*

Plotting tools are attached to figures and create an environment for creating Graphs. These tools enable you to do the following:

Select from a wide variety of graph types , Change the type of graph that represents a variable , See and set the properties of graphics objects , Annotate graphs with text, arrows, etc.

• Create and arrange subplots in the figure ,• Drag and drop data into graphs Display the plotting tools from the View menu or by clicking the plotting tools icon in the figure toolbar, as shown in the following picture

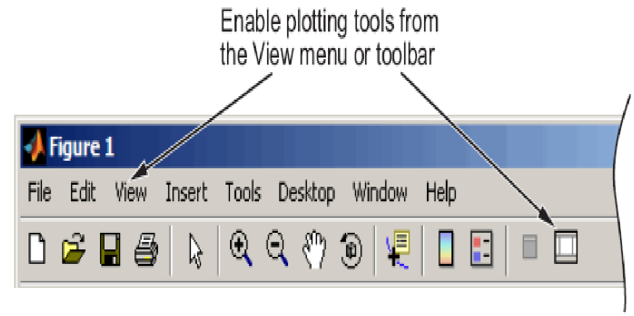

Figure4: Drag and drop data into graphs display

b) *Editor/Debugger*

Use the Editor/Debugger to create and debug M-files, which are programs you write to run MATLAB functions. The Editor/Debugger provides a graphical user interface for text editing, as well as for M-file debugging. To create or edit an M-file use File > New or File > Open, or use the edit function.

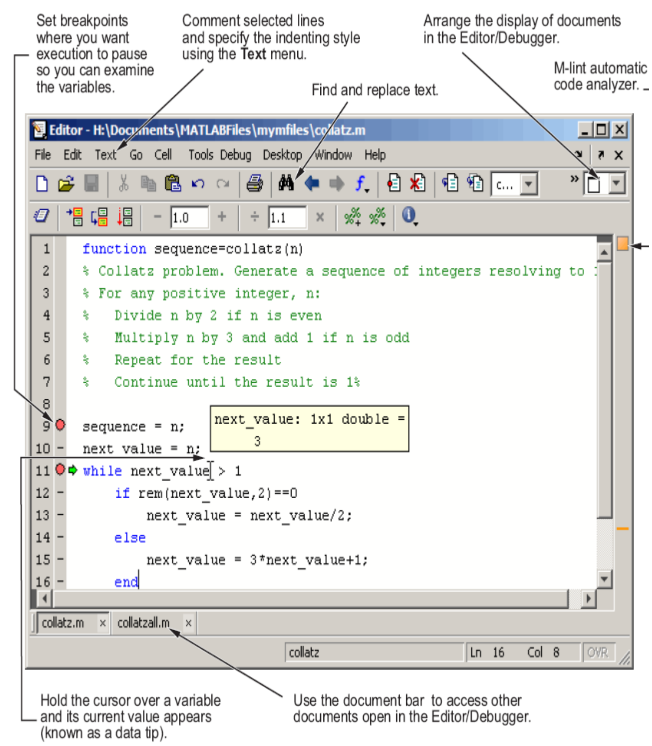

Figure5: M-file debugging-edit function

IX. Results

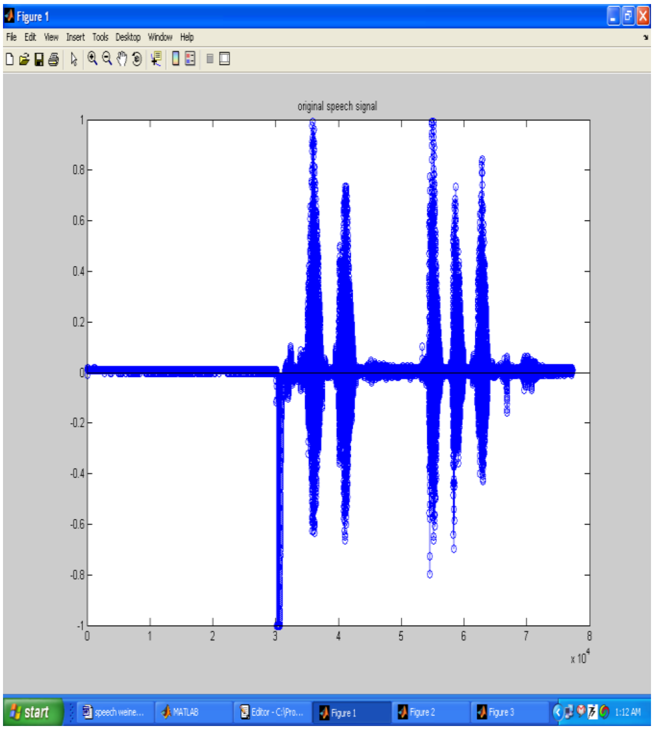

Figure6: Original speech signal

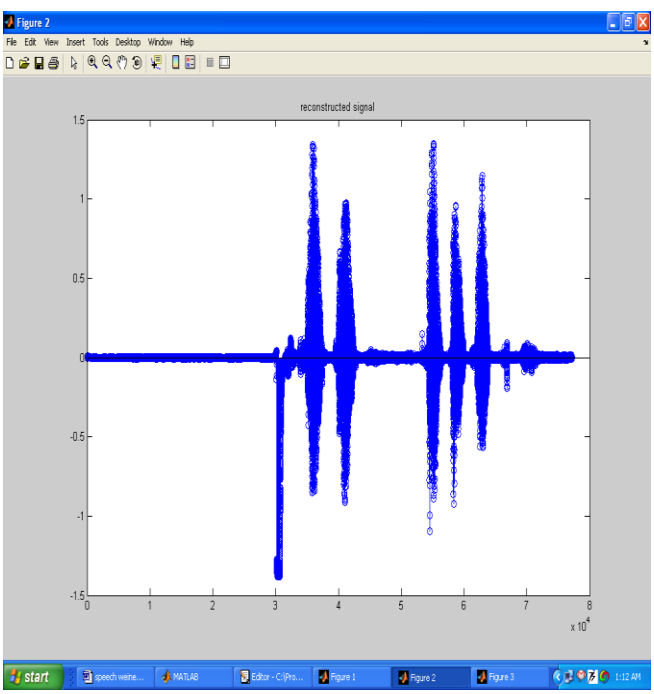

Figure6: Reconstructed signal

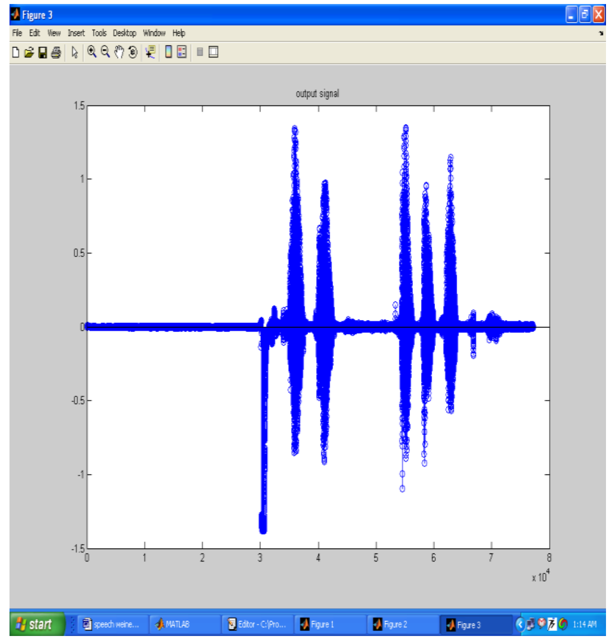

Figure7: Out put signal

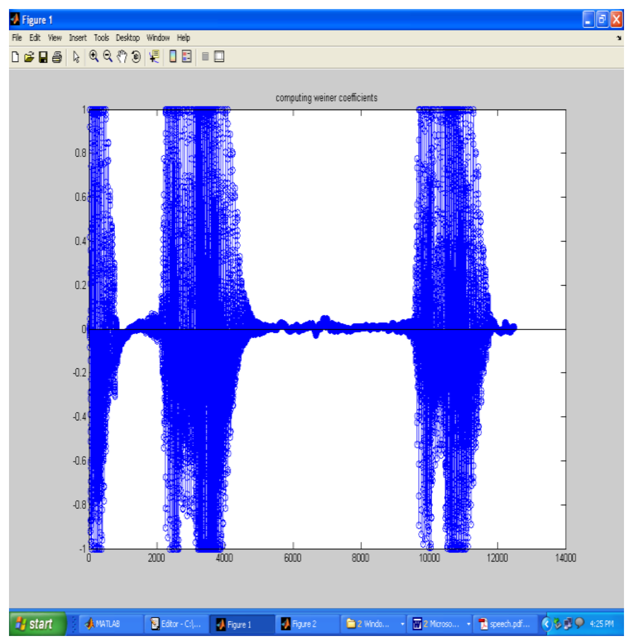

Figure8: Computed Weiner coefficients

Total no.of coefficients 100 The shift length have to be an integer as it is the number of samples. shift length is fixed to 80.

X. Conclusion

In this paper, an implementation of employing Weiner filtering to speech processing had been developed. As has been previously mentioned, the purpose of this approach is to reconstruct an output speech signal by making use of the accurate estimating ability of the Weiner filter. True enough, simulated results from the previous chapter had proven that the

Kalman filter indeed has the ability to estimate accurately. Furthermore, the results have also shown that Weiner filter could be tuned to provide optimal performance. This test is of necessity for the reason that different signals are bound to be similar but not identical. By and large, this thesis has been quite successful in terms of achieving the objectives. Consequently, perception on signal processing and Kalman filter had also been treasured throughout the process. Most importantly, the skill in time management applied during the research of this project had been developed.

XI. Future Developments

The future of Weiner filtering on Speech Processing seems reasonably bright. During the process of this project, many issues have been found to be potential topics for further research work. For that reason, the following issues were raised for further developments:

Speech Compression: The technique of Kalman filtering can be applied to speech coding using Autoregressive (AR) modeling. Since compression is the major issue here, optimal compression cannot be achieve if the entire speech signal used. The best approach is to extract the excitation sequence (white noise) otherwise known as the nonredundant information, which contains the core information of the entire speech signal. Moreover, it is said to beneficial for compression. After which, this white noise can be process through a quantizer and ready to be encoded into suitable bit rates. Quality of speech: Human speech is however difficult to artificially produce. This implies that the quality of speech is yet another major issue. Quality factors to be considering which will affect the speech are somehow complex. For instance, tandem connections, robustness to acoustic inputs, robustness to digital transmission errors as well as delay of transmission are all important factors for thorough investigation.

References Références Referencias

1. S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol.27, pp. 113-120, Apr. 1979.
2. M. Berouti, R. Schwartz and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," Proc. *IEEE Int. Conf. On Acoust. Speech, Signal Procs.*, pp. 208-211, Apr. 1979.
3. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-term spectral amplitude estimator," *IEEE Trans.On Acoust. Speech, Signal Proc.,* vol.ASSP-32, No.6, pp. 1109-1121, Dec.1984.
4. Y. Ephraim and H. Van Trees, "A signal subspace approach for speech enhancement," IEEE *Trans. Speech Audio Procs,* vol. 3, pp. 251-266, Jul. 1995.
5. J. Lim and A. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proc. IEEE,* vol. 67, No. 12, pp. 221-239, Dec. 1979.
6. M. Sambur, "Adaptive noise canceling for speech signals," *IEEE Transactions on Acoustics, Speech and Signal Processing,* vol. 26, p. 419-423, 1978.
7. S. Kamath and P. Loizou, "A multi-band spectral subtraction method for enhancing Speech corrupted by colored noise," Proc ICASSP, pp.IV-4164, May 2002.
8. Alan V. Oppenheim and Ronald W. Schafer. "Discrete-Time Signal Processing".
9. Lawrence R.Rabiner, Ronald W.Schafer. "Digital Processing of Speech Signals". Pearson education

XII. Appendix

M-files for MULTIBAND SPECTRAL SUBTRACTION

Main module

```
%***********************************************
***********************Real is a file containing the words "PHASE DETECTION AND RECOGNITION ARE CHALLENGING TASKS" uttered by a male voice.
%***********************************************
*********************
[signal,fs] = wavread('C:\MATLAB7\work\noisy\real.wav');
% signal = signal(1:45000);
nsignal = signal;
% EACH BAND HAS N/bands FREQUNCY COMPONENTS
each = N/bands;
[seg, nf] = segmenth(ensignal,ovlap,W);
% DETERMINE THE DFT OF NOISY SPEECH
pha = (angle(fft(seg,N,2)));
% ESTIMATE OF NOISE FROM FIRST 10 FRAMES
nmag = (abs(fft(seg(1:10,:),N,2)));
uw = (sum(nmag))/10;
% MAGNITUDE AVERAGING
```

XIII. Acknowledgment

The work is carried out through the research facility at the Department of Computer Science & Engineering and Department of Electronics & Communication Engineering, HITS College of Engineering, Bogaram(V), Keesara(M), R.R.District,A.P., and Dept., ECE/CSE from JNTUK. The Authors also would like to thank the authorities of JNTU Kakinada, SKIT Chittoor for encouraging this research work.