A Novel Approach of Speech Recognition Using System Identification

Ibrahim Sadoon Fatah
Electronic Dep. – Engineering Collage – Daiyla Universit,.
ibrahim1962@engineering.uodiyala.edu.iq

Abstract

With the growing of artificial intelligence and the usage of sound commands the needs for a high accuracy speech recognition increases. Many researches are done in this area using different kinds of methods and approaches. In this research two algorithms have been introduced. The autoregressive system identification and the FIR Wiener filter. The objective of this research is to show the robustness of system identification in terms of speech recognition.Both algorithms have been implemented and tested using MATLAB where the process is done by recording full sentences from different subjects under two conditions which are clear and noisy background. For each sentence, it has been recorded two timesfor each subject; the first one was used for testing and the second sentence was used for validation. The results show that both algorithms are giving an accurate prediction when the used data are from the same subject with clear background. The advantage of system identification over the Weiner filter is shine when using noisy signals. Another advantage of using system identification for speech recognition is it can distinguish the sound difference when same sentence from different subjects is used where the Weiner filter in some cases passes them as from the same subject. This could be a huge issue if the algorithm is used for security reasons.

Keywords: System identification, ARX, FIR filter, Wiener filter, speech recognition, signal processing

Introduction

Speech is the natural tool of communication between individuals. Most speaking skills are learnt during early childhood, and then improved throughout our lives. The human voicecharacteristics differs depending on several factors such as gender, age, emotion and so on.Human voices are varied in terms of volume, accent, speaking speed and pronunciation.Moreover, patterns of spoken language may be loose or twisted during transmission by the echo factors or the dissonance of the background. All these parameters increase the complexity the speech recognition. Speech recognition is one of the most useful technologies of today. Speech recognition is widely applied in the actual world of human language, for example retrieved information. Speech recognition is the foundation of all forms of communication as it is made up of the basic form of conversation. The conversations or speech, as it is known, where the words are transformed into signal waves that are identified by microphone.

Speech recognition has been one of the most popular research fields [1]. There are several applications of speech recognition, which are found foreffectively construct our lives. Mobiles are a good example of such applications where instead of typing,the users can communicate with the device by speaking to it and it will transmit the speech to orders. Speech recognition has also allowed people to control systems using their speech [2]. In speech recognition purpose, various techniques are utilized and they are Dynamic Time Warping (DTW), Neural Network (NN), Vector Quantization (VQ),Cross correlation (CC).Expert System and Hidden Markov Model (HMM).With reduced typing time or operating of system button, speech recognition makes system configuration and management easier [3].It is more convenient to utilize the speech or sound to control the system [4]. It can also lower the cost of industrial production. Utilization of speech recognition system improves the time delay efficiency and makes people's lives more diversified.

In 1930s, the earliest speech machine was introduced and it had limited functions as it could only respond to a small set of words. The machine was responsive to words uttered and speech was produced. Later, many researchers became interested in inventing a speech recognition system. One sample of the machine that was invented by Olson and Belar of RCA Laboratories in the 1950s could identify 10 syllables uttered by a single speaker. MIT Lincoln Lab, invented a speaker-free 10-vowel recognizer. Fig. 1 shows speech and multimodal technology researches' highlights. Advancing from a 1962s speech technology that had setbacks in terms of limited vocabulary and acoustic-phonetics, the technology has advanced further where it has the capacity to process a larger vocabulary using the semantic multimodal dialogue technology.

Today's speech technology has the crucial role in various modern applications. It has moved forward from research to a more profitable application such as Siri of the iOS Apple and Microsoft systems' Cortana. Speech recognition is divided into two types: continuous speech recognition (CSR) and isolated word recognition (IWR). CSR functions of recording the observed sequence corresponds to a full sentence where IWR functions by recording one word at a time from the input signal. IWR is a special case of CSR.

The main objective is to design a proper algorithm that can recognize a full sentence from different subjects

accurately. Many researches have been done in this area and a lot of them has been promising, but there is a limitation in each algorithm. In this research system identification is used to approve that this approach is a robust method for speech recognition and it is more efficient than other algorithms that are used in the speech recognition field.

The rest of this paper will be organized as follows: section two will be the Theory of speech recognition. Section threedescribes the overall methodology of the research. Section four describes the implementation and the results and finally, section five will be the conclusions.

I. THEORY OF SPEECH RECOGNITION
A. Sampling Theory and DC Load Line Removal

One of the most unflavored data in the signals is DC level. It is only preferred when actual analog circuits, signals are used in the conversion of digital data. The frequency domain where the DC line load or DC level is deemed unusable and in certain circumstances of the low frequency band, it affects the signal as it makes interference.This issue is caused by the variance and the signal's mean value when it does not change with time. In the effort to lessen the DC level effects. This allows the DC level to remove the zero frequency components in the documented signal's frequency spectrum. Another vital aspect is the frequency of the sampling that determines the data's quality. Generally, the analog signal is characterized in the following equation:

$$x(t) = \sum_{i=1}^{N} A_i \cos(2\pi f_i + \emptyset_i) \qquad (1)$$

This analog signal is made from many different frequency components. With the assumption that there is a single frequency component does not have a phase shift. The characteristics of the signal are:

$$x(t) = A\cos(2\pi f t) \qquad (2)$$

The computer is unable to process an analog signal; thus, the signal needs to be transformed into a discrete signal. The discrete signal is often considered as a one signal arrangement or a vector. Fig. 2 demonstrates how the analog signal's sampling is converted into a discrete time signal. For example, the sine wave is taken to show the signal's sampling.

The analog signal's duration is T that represents the discrete time signal's sampling period. With the assumption that the analog signal is sampled from time equal to 0, the sampled signal can be written as a vector x(t) = [x (0), x (1), x (2), x (3), x (T-1)]. The mutual link between the time and the frequency is the sampling frequency, Fs = 1/T_s supposing the length of x (t) is T for N original time periods. The link between T and T_s is:

$$T \times T_s = N \times T \qquad (3)$$

Both T and N are integers, so it will be:

$$T / N = T / T_s = F_s / f \qquad (4)$$

In circumstances where the analog signal is accurately sampled with similar sampling space and periodic sampled signal, then N/K is also an integer. In another situation, the sampled signal may also be aperiodic. Based on the sampling theorem (Nyquist theorem), if the sampling frequency is greater than or equivalent twice of the analog signal frequencies maximum capacity, the discrete time signal has the capacity to be used to restructure the original analog signal [5]. The higher sampling frequency gives better sampled signals as there is more data space. In non-telecommunication applications, the subsystem speech recognition may access the better-quality speech, with sampling frequencies is in Kilohertz such as 10KHz, 14KHz and 16KHz. These sampling frequencies provide improved time and frequency resolution [1]. The sampling frequency in this project is set to 16KHz. The recorded signal's length is 30 seconds.

B. Converting from Time Domain to Frequency Domain

The human sounds or any recorded sounds ofreal time is in the time domain and the programs cannot deal with it for that the signal should be converted from time to frequency domain in other word from continues to discrete representation. Many algorithms are used to convert the signal from time to frequency domain. The most popular method is the Fourier transforms. There are two types of Fourier transform to convert the signal from time to frequency domain [6]. The first one is the discrete Fourier transform (DFT) which is done by a long mathematical calculation and the second one is the fast Fourier transform (FFT) which is the same as (DFT) but in the short term of calculation [7]. The DFT is a kind of Fourier Transform where it's used for discrete time domainx(n) instead of continuous time domain x(t). The Fourier Transform equation is:

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{j\omega n} \qquad (5)$$

Fourier Transform is mainly used for transforming the variables from n to ω, which means transforming the signals from time to frequency domain.

FFT is a DFT also which is used for transforming the discrete signal from the time to frequency domain as well. The difference FFT and DFT arethat the FFT is faster and more efficient. There are several ways to increase the DFT efficiency, butstill FFT is the most widely used algorithms[5]. It is convenient to investigate FFT by firstly considering the N point DFT equation since it's still a kind of DFT:

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{kn}, \quad k = 0,1,2,. \qquad (6)$$

Where x(n) must be separated into odd and even. x(odd) = x(2m+1) and x(even) = x(2m) where m = 0, 1, 2, …, N -1 then the N point DFT equation also becomes two parts for each N/2.

The values of the phase do not change by shifting the phase with half period, but, the phase factor's sign will be opposite. This operation, called the phase factor's symmetrical property. The advantage of this process is reducing the calculation of the DFT to N/2 from N point by continuously separating the series to an odd and even series. The total number of the complex multiplication will be reduced by reducing the DFT N point continuously until it reaches to one-point sequence.

C. Wiener Filter

Wiener filter can be utilized for generating an estimation of a preferred or targeted random processing of linear time-invariant filtering of a noted noise process which has known noise spectra and signals. Wiener filter can minimize the mean square error that occurs between the estimation of a random process and the preferred process. The aim of using the Wiener filter is to calculate an estimation of the signal by utilizing a related signal as an input and filtering out known signals to produce the estimates as outputs. The Wiener filter is utilized for filtering the signal from the corrupted noise to stipulate an estimation of the essential signals of interest. The Wiener filter uses a statistical technique whereby a statistical explanation of the theory involved is presented for the minimum mean-square error. In order to produce a desired frequency response, deterministic filters are created. Nevertheless, a Wiener filter uses a different methodology. It is assumed that a person is aware of the spectral features of the original noise and signal. It also assumes that a person will seek the linear time-invariant filter which produces outputs which are as close as possible to the original signal. Wiener filters have the following characteristic [8]:

- Assumption: The additive and signal noise is stationary linear stochastic processes which have identified spectral features or cross-correlation and autocorrelation features.
- Requirement: The filters need to be physically realizable or causal. If this feature is not fulfilled, then the solution will be a non-causal.
- Performance criteria: has minimum mean-square error.

The process of deconvolution usually uses a Wiener filter. Wiener filter can be utilized to solve the following cases:

- A non-causal filter requires an infinite number of past and future data.

$$G(s) = \frac{S_{x,s}(s)}{S_x(s)} e^{\alpha x} \tag{7}$$

Where S is spectra.As long as g(s) is optimal. The usage of the minimum MSE equation is to reduce the equation below:

$$E(e^2) = R_s(0) - \int_{-\infty}^{\infty} g(\tau) R_{x,s}(\tau + \alpha) d$$

Where g(τ) is two sides Laplace transform inverse of G(S).

- A causal filter requires the usage of an infinite amount of past data.

$$G(s) = \frac{H(s)}{S_x^+(s)} \tag{9}$$

Whereby H(s) is made of the casual section of the non-causal.

- The finite impulse response (FIR) case whereby a limited amount of past data is utilized.

The first case can be solved easily, but it still not suitable for real time applications. Wiener's main achievement was solving the case whereby the causality requirement is in effect.

D. System Identification

System identification is a mathematical model implementation based on statistical methods. It provides an efficient fitting model because it contains a pathway of the optimal design [9]. System identification was used mostly with discrete time domain until the last two decades where the researchers start using system identification approaches with continuous time domain [10]. Nowadays, several system identification is used with several continuous time domain applications such as estimating the time delay between two signals and design to supply an artifact robust estimation [11]. System identification does not use for speech recognition in spite of its usage with many signal processing applications. The general idea of system identification is that the two signals are imported into the system one of them is set as an input signal and the second is an output signal which means we are know the input and the output and you as a system predict the fitting or matching or calculate the error between these two signals.

II. RESEARCH METHODOLOGY
A. Database Description

Six young volunteers are invited to do the tests for this research (4 males and 2 females) with a range of 25 ± 3 years old. Because every person hasa different level of sound frequencies and the level of sound can be differ depending on the situation each subject asked to pronounce four different sentences with two times repetition for each sentence. The sentences that are used for this project are (he is good, let us go to the market, wish you safe travel and see you next time). Fig. 3shows an example of the recorded signals (Let us go to the market).

B. Programing Procedure

This researchfocuses on two algorithms that are used in the field of speech recognition and signal processing. MATLAB is used to implement these algorithms for real signals or real timeapplications[12].

The program will ask the subjects to pronounce the sentence and it will record it as a wave signal for two times. The first one is used as reference signals or simulation signal in system identification. The second

signal is used as the target signal or the validation signal in system identification. The procedure of the wiener filter algorithm is shown in Fig. 4.

After finding the filter coefficients for the reference signal and calculate the minimum mean square error for each reference signal. The better estimation should have the smaller minimum mean square error. The procedure of the system identification approach is shown in Fig. 5.

C. ARX Model Order Selection

Each system consists of three things input, output and the black box. Autoregressive with exogenous input (ARX) model is one of the common system identification approaches. The ARX approach can be represented be the following equation:

$$\sum_{k=0}^{N} a\, y[n-k] = \sum_{k=0}^{M} b\, x[n-k] + e[k \qquad (10)$$

Where $x[n]$ is the input, $y[n]$ is the output and $e[k]$ is the error. The system model is identified by the parameters a and b. The second step after identifying the model is the validation. The validation process in system identification is done by setting the first signal as a simulation and the second signal is set for validation which helps to avoid over fitting. To quality check of the model is done by comparing the output to the true system output with the validation data set. In this research, the selected parameters for model determination are ($na = 6$, $nb = 4$ and $nk = 1$). The zeros and poles of the bode diagram, determine the accuracy of selected parameters. The accurate system should not have a zero-poles cancelation in other word there is no overfeeding between them. Fig. 6 shows the bode diagram of the selected model order.

Many other parameters have been examined to get the best system order such as ($na = 10$, $nb = 8$ and $nk = 6$) are examined. Where these parameters produce many poles and zeros which could be affected by the system. The largest number of zero-poles will cause a cancelation. For the same tasted signal in experiment one. Fig. 7 shows an example of fitness function for one of the tests.

From Fig. 7 shows the fitness function of using a model with parameters ($na = 6$, $nb = 4$ and $nk = 1$). It can be seen that the two signals are matched almost typically where the fitness function's matching results is 96.87%. The black signal represents the simulated signal and the blue signal represents the validation signal. Fig. 8 shows how a fitness function can be affected by changing the model order parameters.

Fig. 8 shows how the fitness function could be affected by changing the model parameters. Where, in spite of using the same signal that is used in Fig. 7, the fitness function shows only 11.28% of matching between the two signals. There is a huge difference between the two results which shows how system identification is

sensitive to the parameter selection. The fitness function is not the only parameter that is used for giving an indication of on the system accuracy. The correlation is the other parameter that is used for determining the accuracy of the system. The most accurate system should have highest fitness function and highest possible correlation. Fig. 9 shows the correlation for the model with parameters ($na = 6$, $nb = 4$ and $nk = 1$).

Fig. 9 shows that there is a high correlation relationship between the simulated and validated signals where two dotted lines represent the determents of the cross and auto correlation. Fig. 10 shows the model residuals of the system using ($na = 8$, $nb = 6$ and $nk = 10$) as a model orders.

From Fig. 10 it can be seen that the signals are not correlated to each other where the correlation line is out the determents of the correlation. From the figures (6, 7, 8, 9 and 10) it can be easily selected the accurate model order.

III. RESULTS AND DISCUSSION
A. Results of Wiener Filter Algorithm

Each subject asked to pronounce four sentences. The process of testing the methods, accuracy is by testing the method four times for each subject where in the first test both simulated and validated signals are from the same subject with clear background. The second test is also same subjects pronounce the same sentence twice for simulation and validation but with a noisy background. The third test is using the same sentence for simulation and validation with clear background, but from different subject to test the method robustness in differentiating between different voices. The final round is testing same sentence pronounced by different subjects with noisy background which the most difficult situation. Table 1 shows the overall results for all subjects with all pronounced sentences in four proposed testing processes with the minimum MSE for each sentence and the success probability of this system. Where MMSE means minimum mean square error.

From Fig. 11 shows that the method is giving a good result when both signals are recorded from the same subject with clear background where the error values are too small. The figure shows a dramatic increase in error values when the noisy background is added to the signals. The worst results with higher error values are appearing when the signals are recorded from different subjects with a noisy background. Fig. 12 shows the distribution of the successful probability of detection the sentence.

From Fig. 12 shows the highest probability occurs when the two signals are recorded from the same subject with clear background where the lowest probability occurs when the two signals are recorded from different subjects with a noisy background. In General, there is an opposite relationship between the MMSE and the success

probability. When there is a small error there is a high success probability which means that the method can recognize the speech. From table 1 and figures (11 and 12) we can come out with two findings, which are the method is accurate if the two signals are recorded with clear background in either the two signals are recorded from same or different subjects. The method is failing in detecting the speech when the two signals are recorded with noisy background and. The second finding is when the used signals are from different subjects, butthey have close frequencies to each other the method is not able to differentiate between them. The overall results show that thedetection accuracy is around 50%, which means the method has a huge chance of failure.

B. Results of System Identification

The process of calculating the system identification results differs from the Wiener filter. In system identification, the main factorsofspeech recognition are the fitness function and the correlation properties of the models. System identification can deal with the signals in time domain directly. The system identification, filtering process is more complicated than wiener filter to keep the important information that are required, and get rid of unwanted noise in the signals. Table 2 shows all the overall results for all subjects pronouncing the four subjected sentences in same four cases that are used with wiener filter.

Fig. 13 represents the results of table 2 by showing the distribution of the fitness function values of all subjects.

Fig. 13 shows that the system can detect the noisy signals clearly where the wiener filter failed in this part. Table 2 and Fig. 13shows system identification providesgood results even with a noisy background. The range of fitness function varies between 80% - 99% when both signals are recorded from the same subject which is a very high range where the fitness function values varies between 15% - 40% when the signals are recorded from different subjects.

C. Wiener Filter and System Identification Comparison

The aim of this research is to show the robustness of system identification in the field of speech recognition over the wiener filter. Fig. 14shows the difference between the success probability of the Wiener filter and the fitness function of the system identification.

From Fig. 14it can be seen that the system identification gives a better accuracy percentage over the wiener filter. System identification can detect and recognize the speech in both clear and noisy backgrounds where the wiener filter fails in detecting recognizing the speech with a noisy background. The highest success for wiener filter is 95%, which occurred when the test is done on clear background. The method's accuracy fallen to 65% when it applied on noisy background. Where is

system identification the worst results were 80%, which very good hit rate compare to wiener filter.

In case the simulation and validation signals are recorded from different subjects the results should be in opposite, which means each system most provide lowest possible matching. In some cases of using the wiener filter the matching hits 50% because the wiener filter cannot recognize the difference between the close frequencies to each other. Where the range of match was around 30%. On the other hand, the system identification worst result was 38%, which still better than wiener filter with an overall range below 25%.

Generally, the overall advantages of system identification over wiener filter are system identification have the ability of dealing with time domain signals directly without needing to transform it to the frequency domain unlike the Wiener filter. Another advantage is the dealing with noisy background where system identification does not require signal normalization.

IV.   CONCLUSIONS

Speech recognition is one of the common research areas. Two algorithms are implemented to test the robustness of these algorithms is field of speech recognition. The first algorithm is the wiener filter and the second is ARX system identification. The wiener filter provides the results by calculating the MMSE where the less error means better results. The results show that wiener filter produces good results when both signals are recorded from the same subject with a clear background and its fallen when using noisy background. System identification provides the results depending on the fitness function between the simulated and validated results where the highest fitness function means better results. The results show the robustness of system identification in both clear and noisy backgrounds.

REFERENCES

[1] Joseph W. Picone, Signal Modeling Techniques in Speech, *Systems and Information Sciences Laboratory, Tsukuba Research and Development Center, Tsukuba, Japan,* 1993.

[2] John G. Proakis and Dimitris G. Manolakis, Digital Signal Processing, *Principles, Algorithms, and Applications, 4th edition, Pearson Education Inc., Upper Saddle River,* 1996.

[3] Adam L. Buchsbaum and Raffaele Giancarlo, Algorithmic Aspects in Speech Recognition: An Introduction,*Annual Symposium on Combinatorial Pattern Matching,* 2009.

[4] HartmutTraunmuller and Anders Eriksson, The Frequency Range of the Voice Fundamental in the Speech of Male and Female Adults, *Institution for linguistic, Stockholm, Sweden,* 2009.

[5] John Wiley, Sons, 2006, Inc. Statistical Signal Processing and Modeling, Monson H Hayes, *Georgia I institute of Technology,* 2006.

[6] Fredric J. Harris, On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform, *Proceedings of the IEEE,* VOL 66, No.1, 1978.

[7] Henrik V. Sorensen and C. Sidney Burrus, Efficient Computation of the Short-Time Fast Fourier Transform, Electrical and Computer Engineering Department, Rice University, Houston, 2012.

[8] Heather Sobey, Literature Survey Automatic Speech Recognition, *Department of Computer Science, University of Cape Town,*2011.

[9] G.P. Rao and H. Unbehauen, Identification of Continuous Time Systems,*IEE Proceedings - Control Theory and Applications, Vol. 153, No. 2,* 13th March 2006.

[10] Da Xu, Kathy L. Ryan, Caroline A. Rickards, Guanqun Zhang, Victor A. Convertino and Ramakrishna Mukkamala, Robust Pulse Wave Velocity Estimation by Application of System Identification to Proximal and Distal Arterial Waveforms,*Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),*4th September 2010.

[11] Da Xu, Kathy L. Ryan, Caroline A. Rickards, Guanqun Zhang, Victor A. Convertino and Ramakrishna Mukkamala, Improved Pulse Transit Time Estimation by System Identification Analysis of Proximal and Distal Arterial Waveforms, *Am J Physiol Heart Circ Physiol. Vol. 301, No. 4, PP. 1389-1395.*29th July 2011.

[12] Tingxiao Yang, The Algorithms of Speech Recognition, Programming and Simulating in MATLAB. January 2012.
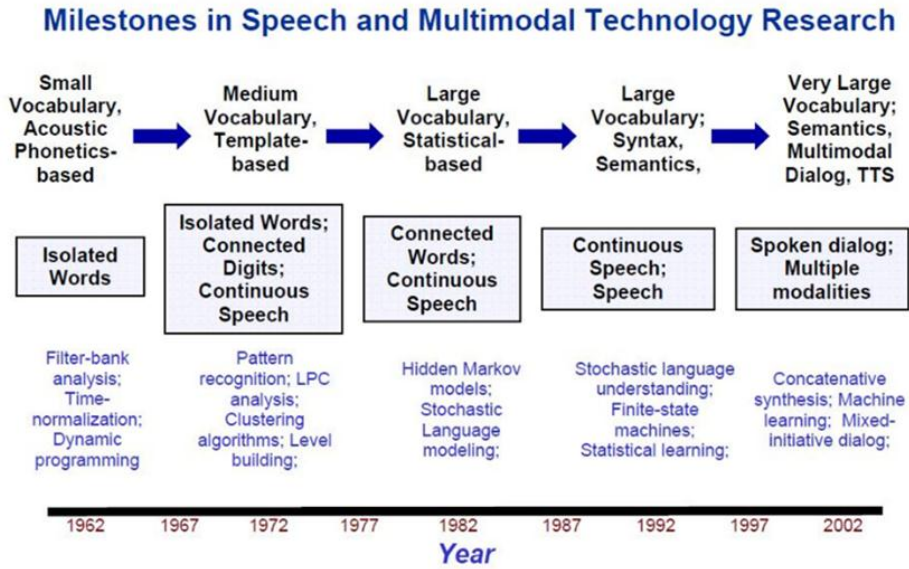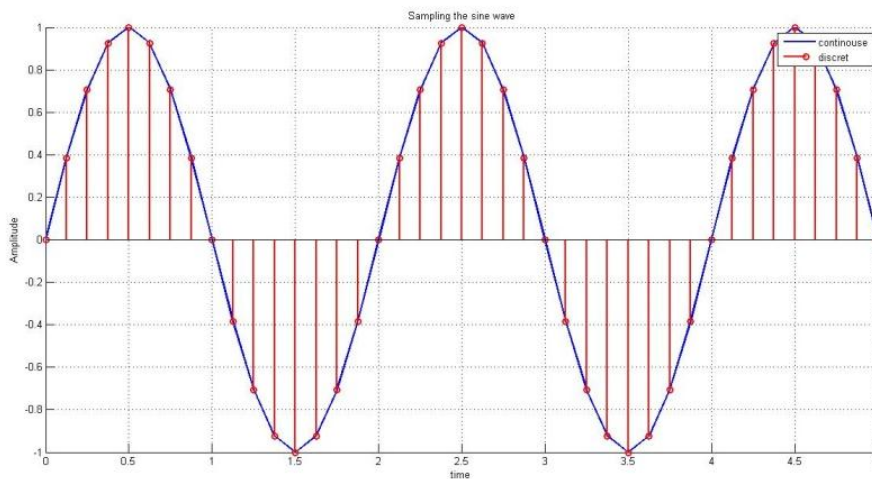
Fig. 1 Speech recognition milestones
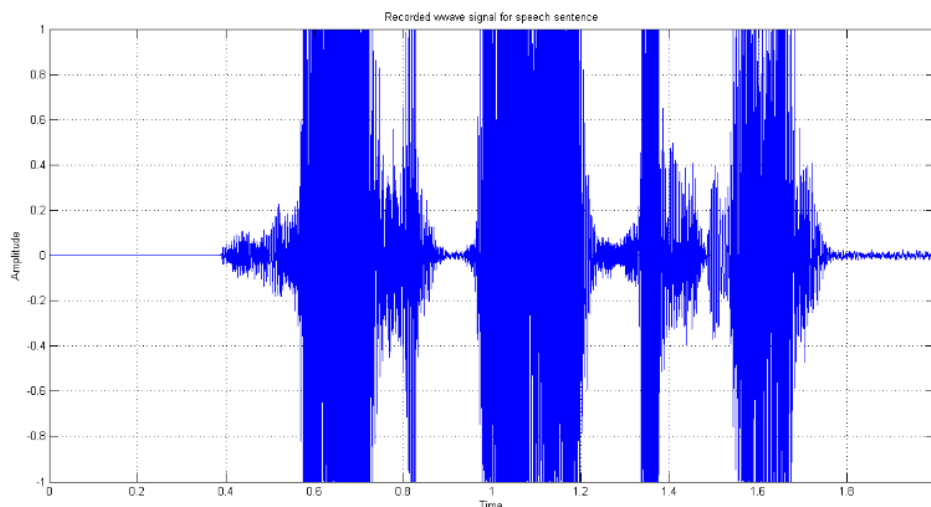

Fig. 2 Analog to discrete signal sampling

Fig. 3 Example of recorded signal

Initialize sampling frequency = 16 KHz
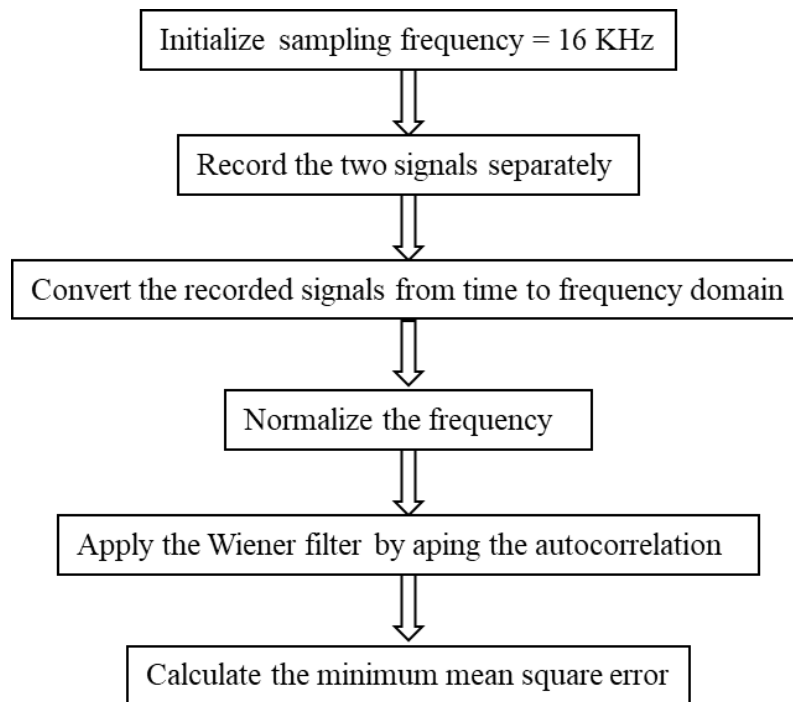
Record the two signals separately

Convert the recorded signals from time to frequency domain

Normalize the frequency

Apply the Wiener filter by aping the autocorrelation

Calculate the minimum mean square error

Fig. 4 Wiener filter procedure

Initialize sampling frequency = 16KHz

Record the two signals separately

Import the two signals to system identification

Sampling the signals with 100 sample rate

Filtering the signal using low pass filter

Detrended the signal to remove DC level

Set the first signal as estimation and the second as validation

Apply ARX model of System Identification

Calculate the results depending on the fitness function and the correlation relationship

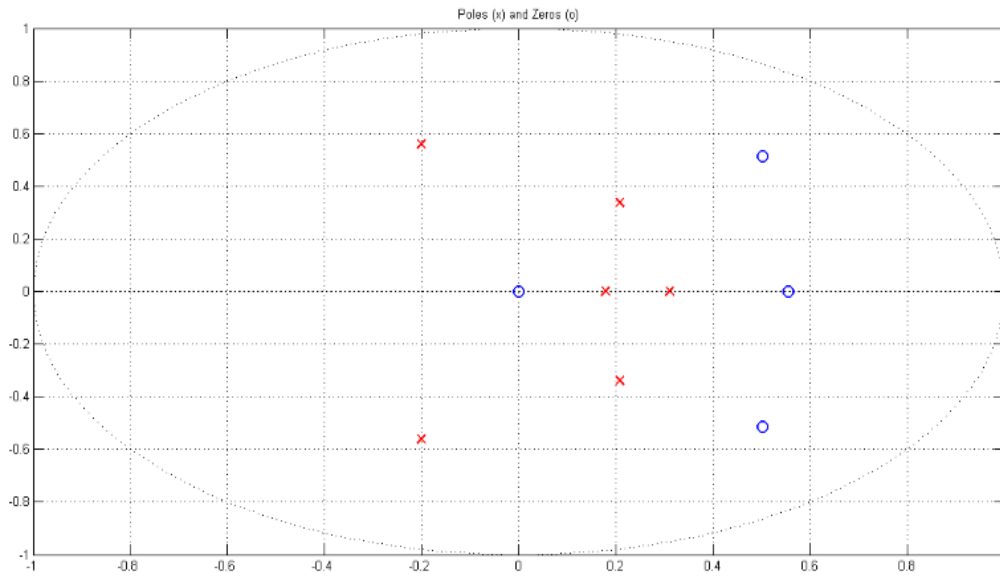Fig. 5 The System Identification procedure

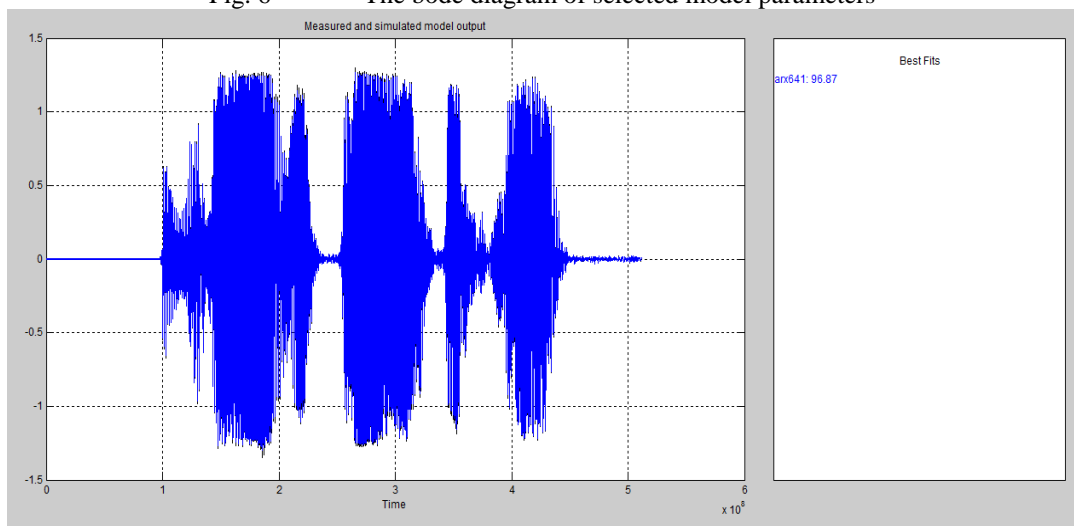Fig. 6   The bode diagram of selected model parameters

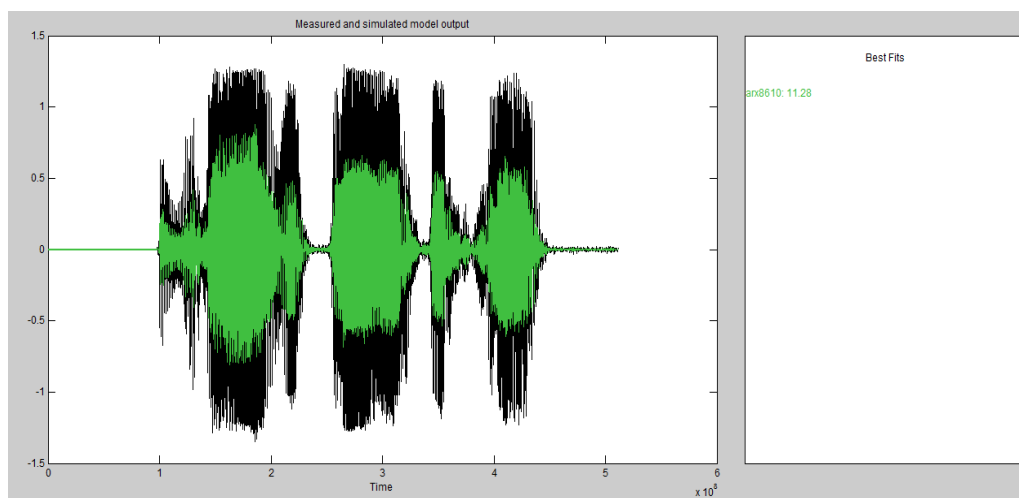Fig. 7   Fitness function of ARX system identification

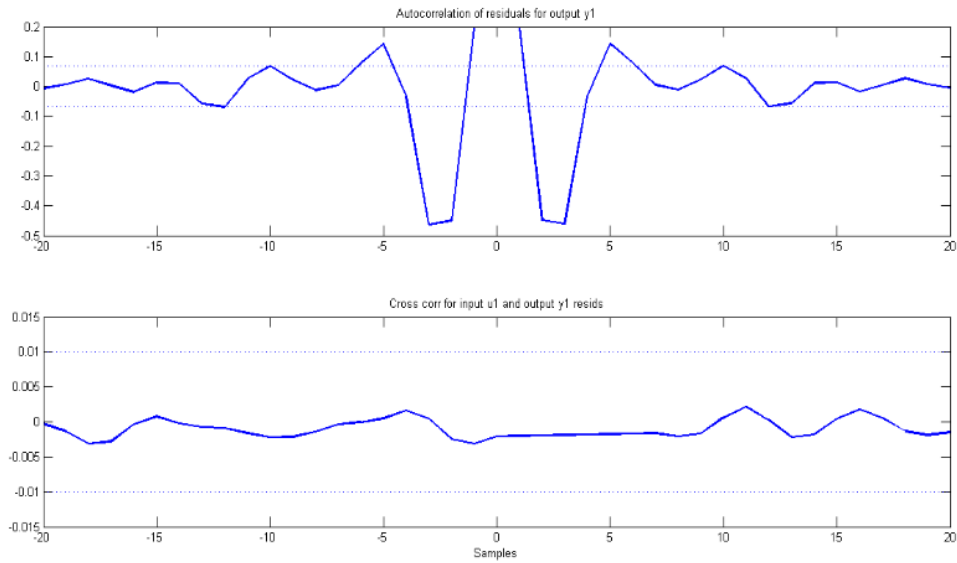Fig. 8   Fitness function withparameters ($na = 8$, $nb = 6$ and $nk = 10$)

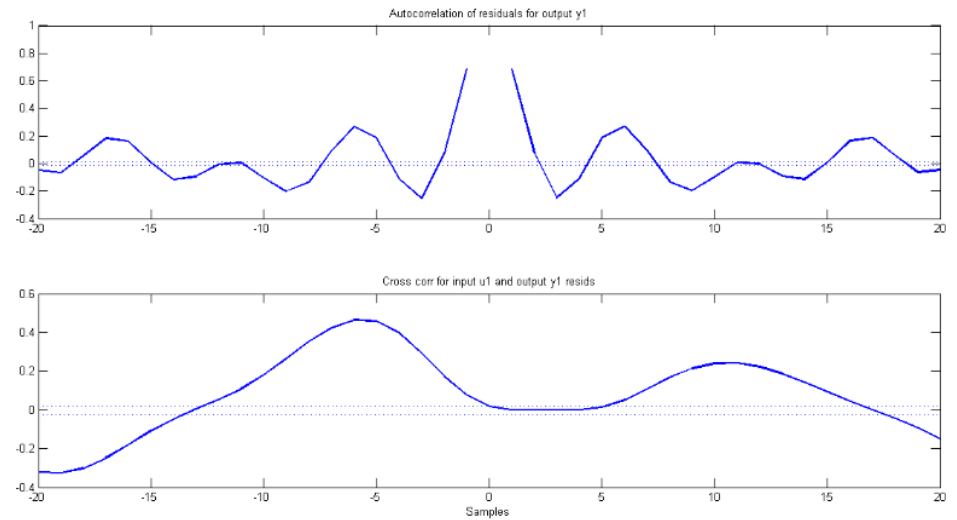Fig. 9   Model residuals of the model with order ($na = 6$, $nb = 4$, $nk = 1$)

Fig. 10 Model residuals of the model with order ($na = 8$, $nb = 6$, $nk = 10$)

Table 1 Wiener Filter Algorithm Results

Subject		Noise	Sentence	MMSE	Success probability
Sim.	Val.				
1	1	Clear	He is good	0.3277	93%
1	1	Noisy	Let us go to the market	2.7586	67%
1	2	Clear	Wish you safe travel	3.3163	51%
1	2	Noisy	See you next time	13.0460	32%
2	2	Clear	He is good	0.5835	87%
2	2	Noisy	Let us go to the market	1.4109	71%
2	3	Clear	Wish you safe travel	4.6609	49%
2	3	Noisy	See you next time	13.0519	34%
3	3	Clear	He is good	0.3277	93%
3	3	Noisy	Let us go to the market	1.1808	72%
3	4	Clear	Wish you safe travel	3.7882	52%
3	4	Noisy	See you next time	14.2020	29%
4	4	Clear	He is good	0.2032	95%
4	4	Noisy	Let us go to the market	1.6448	78%
4	5	Clear	Wish you safe travel	3.4522	55%
4	5	Noisy	See you next time	14.5291	28%
5	5	Clear	He is good	0.4941	91%

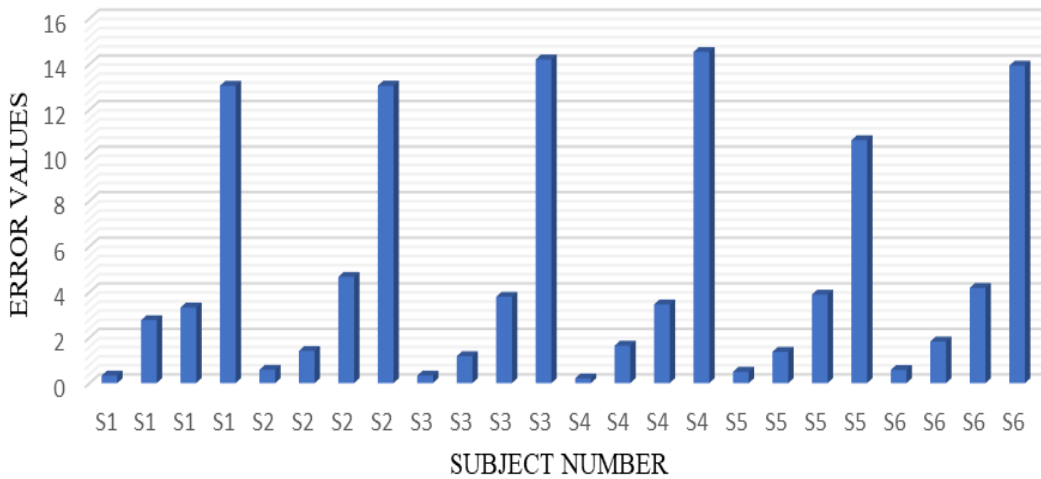| 5	5	Noisy	Let us go to the market	1.3690	70%
5	6	Clear	Wish you safe travel	3.8828	51%
5	6	Noisy	See you next time	10.6575	41%
6	6	Clear	He is good	0.5781	89%
6	6	Noisy	Let us go to the market	1.8230	67%
6	1	Clear	Wish you safe travel	4.1770	50%
6	1	Noisy	See you next time	13.9298	29%

Fig. 11   Minimum mean square error distribution
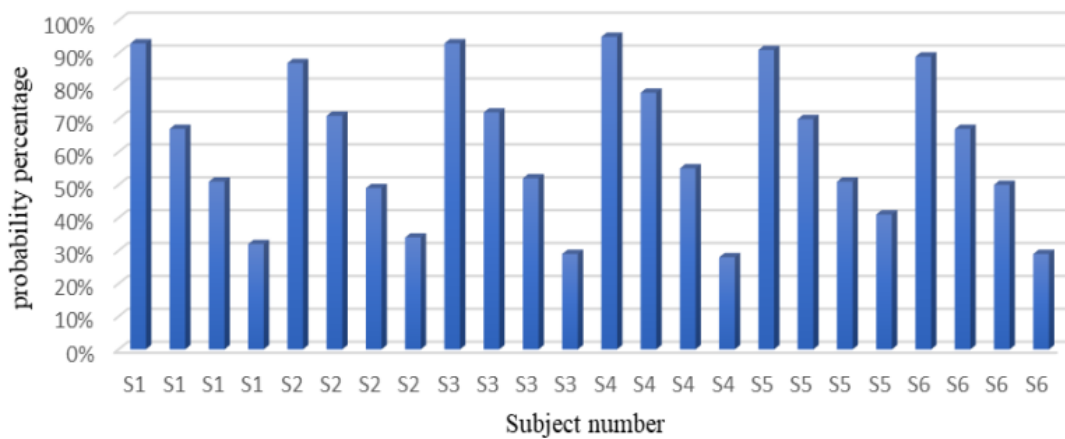

Fig. 12   Successful probabilities for all subjects

Table 2   System Identification Results

Subject		Noise	Sentence	Fitness function
Simu.	Valida.			
1	1	Clear	He is good	98%
1	1	Noisy	Let us go to the market	89%
1	2	Clear	Wish you safe travel	32%
1	2	Noisy	See you next time	12%
2	2	Clear	He is good	96%

| 2	2	Noisy	Let us go to the market	81%
2	3	Clear	Wish you safe travel	38%
2	3	Noisy	See you next time	24%
3	3	Clear	He is good	97%
3	3	Noisy	Let us go to the market	82%
3	4	Clear	Wish you safe travel	39%
3	4	Noisy	See you next time	18%
4	4	Clear	He is good	99%
4	4	Noisy	Let us go to the market	89%
4	5	Clear	Wish you safe travel	38%
4	5	Noisy	See you next time	19%
5	5	Clear	He is good	96%
5	5	Noisy	Let us go to the market	90%
5	6	Clear	Wish you safe travel	31%
5	6	Noisy	See you next time	20%
6	6	Clear	He is good	95%
6	6	Noisy	Let us go to the market	86%
6	1	Clear	Wish you safe travel	20%
6	1	Noisy	See you next time	16%
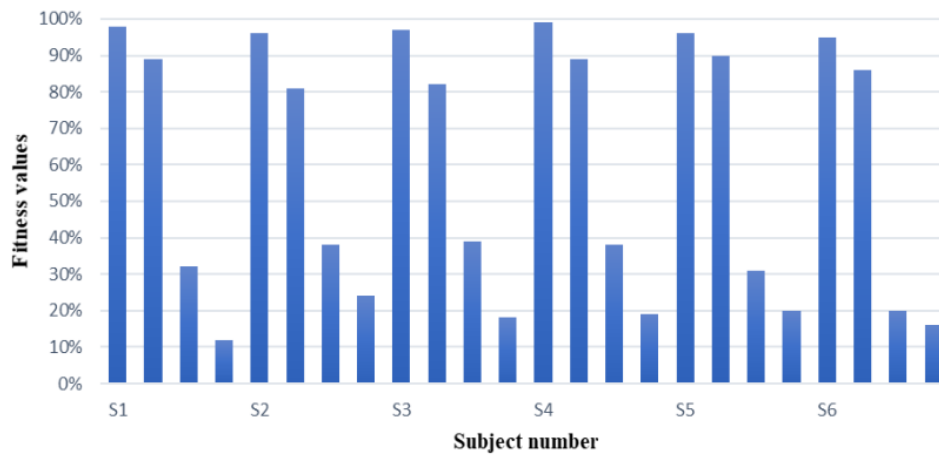
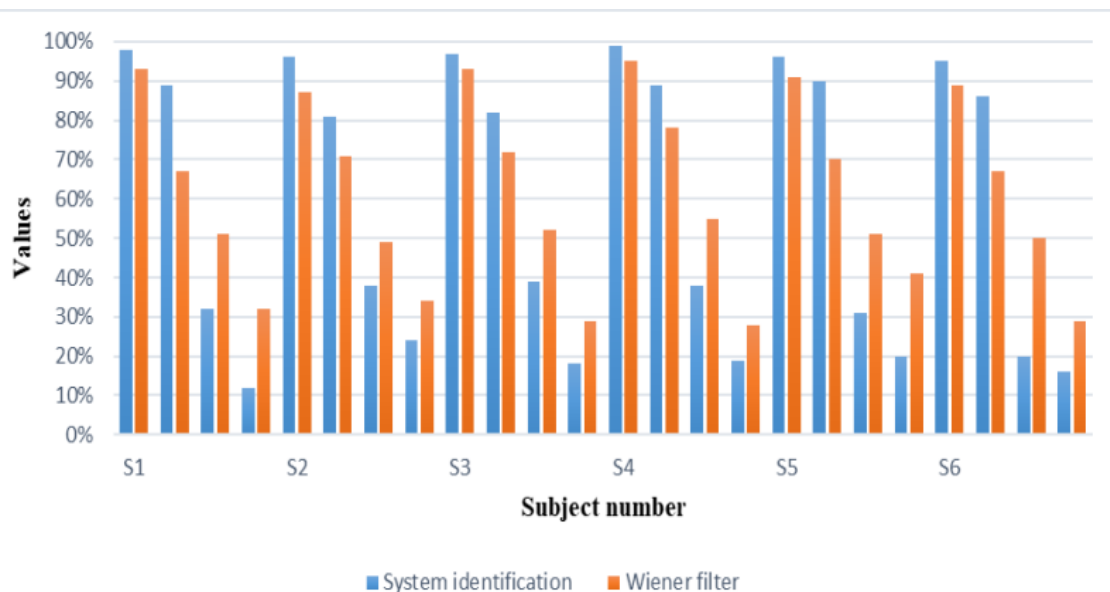Fig. 13   Fitness function values of all subjects

Fig. 14 Wiener filter and System Identification results comparison