JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SPEECH RECOGNITION USING DENOISING TECHINQUES

[1]N M Ramalingeswararao,[2]Kunaparaju Naga Durga Saranya,[3]Raavipati Sirisha, [4]Lakshman Surya Kumar Muppiri ,[5]Korada Ventaka Sai Sankeerthana

[1]**AssistantProfessor,** Dept of ECE, Godavari Institute of Engineering and Technology(A),Rajahmundry,AP
[2,3,4,5]**Students,** Dept of ECE, Godavari Institute of Engineering and Technology (A),Rajahmundry,AP

Abstract-Computers often interact with humans through textual or graphical user interfaces as opposed to spoken human-computer dialogue. Automatic processes and any kind of password-protected computer software all use speech recognition. Research into voice recognition technology has increased greatly in recent years due to the proliferation of multimedia. Based on the decision-making processes used, its many applications for speaker identification fall under the headings of speech identification and voice verification. To demonstrate the potential for collaboration between people and machines in the advancement of technology, this project makes use of a vocal user interface. We are doing some signal deterioration using Gaussian noise and also editing some speech. Using Wavelet Transforms, we can get rid of the backdrop. Two or more signals were compared using correlation in MATLAB to determine which was the most reliable for use in creating the model. In this way, we may create a model where machines can recognise and respond appropriately to various types of instructions.

Keywords: Gaussian Noise, wavelets, symlet, coiflet, correlation.

1. INTRODUCTION

The use of language has a crucial role in the success or failure of a conversation. Speech communication and analysis have been greatly facilitated by the proliferation of digital data storage devices, as well as mobile and landline telephones. The term and the building blocks of speech recognition were developed in the early 1960s as a result of studies on voiceprint analysis, a concept similar to fingerprint analysis. In 1984, the science fiction novel "Star Trek to George Orwell's" introduced the concept of a computer able to detect the human voice.

People who have physical impairments, such as blindness or hearing, may now easily communicate with machines because to advancements in voice recognition technology. Therefore, the brain will use physiologic cues to decipher a tracheal voice. In this paper, we'll take a look at the five cornerstones of speech recognition and the research that's gone into them so far. Include voice enhancement, degradation, format analysis, pitch analysis, and speech editing.

Look at editing as a test for your skills. Search and destroy with your red pen or pencil like a kid playing a video game. Use the "find" function to look for problematic areas in your speech if you wrote it on a word processor. Just press a button and see them

vanish into thin air. The following things are detrimental to effective communication and should be eliminated immediately. Now, get their addresses and go to work on your objective! Overuse of punctuation is present. Oftentimes, emphasis devices like exclamation marks, italics, and underlining are used to convey feelings that aren't present in the actual text. It may be helpful to rewrite the piece using stronger language to convey your ideas.

Specifically, the term "that." I want you to read the text without it. Most of the time, the meaning will remain unchanged. If you want a better speech, change the first sentence to "Zapping difficulty areas indicates that you'll have a better speech." The overuse of the word "I." Using the pronoun "I" at the beginning of every sentence is boring. Keep in mind that the use of one's name or the pronoun "you" is quite endearing to other people. Similar to, "If you take my counsel, you will rule the world," rather than "I would like to offer you some guidance." Modifiers with a "- ly" ending. Adverbs tend to make your speech monotonous, thus it's best to avoid using them. Explain the process to create a more engaging speech. Consider rephrasing the alien invasion as "The alien oozed into each computer, seizing control of the planet one terminal at a time" rather than "The alien silently took over the globe." Connectors." The conjunctions "and," "but," "then," "furthermore," "possibly," "however," and "because" "in speeches often serve no purpose and should be avoided.

If you take out the transitional phrases, you'll have two whole thoughts instead of one weak one. Use of pronouns like "he," "she," and "their." In contexts when there is more than one entity being discussed, pronouns may get muddled. While you may realise that "she" refers to the family dog and not your own mother, your listeners may be confused. Commonly overused amorphous nouns and verbs include "rather," "thing," "plenty," "very few," "stuff," and "some." Those are the work of a sloppy or inexperienced speechwriter. It's worth your effort to discover better, more accurate phrases. Clutter your home no more. Inappropriate language. Remove any overtly sexist or racist language and replace it with language that will resonate with your target demographic.

Vocal pitch is the defining factor in how much emphasis is placed on a word or phrase. Either at the level of a single word or a whole phrase, its significance remains the same. The vocal cord in use, and the frequency at which it vibrates, determine the typical pitch of a human voice. Length, thickness, and tension are the only factors that influence the vibrating frequency of the voice cords. Phonology in English Language Teaching: An International Approach, by Martha Pennington, claims that a person's typical singing pitch is determined by the size of their vocal cords. A man's voice will typically be lower or deeper in pitch than a woman's or child's because their vocal chords are naturally longer in males than in women and children.

The pitch shift is caused by the vibrating of the vocal chords. When the vocal folds vibrate more rapidly, the voice becomes higher in pitch. In most cases, a person's audible pitch range will not exceed the upper limit of 200 hertz (Hz). It is possible to infer the speaker's attitude toward the material being conveyed or the listener standing in front of them based on the utterance's pitch range.

The speaker's modal pitch, which is the unmarked mid-range, reflects the straightforward, unobtrusive tone in which the remark is made. The informational contrast in the statements is shown by the increased volume.

To put it simply, formants are the high-energy frequency peaks in the spectrum. Vowels have a lot of them. There are resonances in the vocal tract that correlate to each formant (roughly speaking, the spectrum has a formant every 1000 Hz). We may think of formants as a kind of filter.

The term "formant" refers to the wide spectral maximum produced by an acoustic resonance of the human vocal tract, and is used in speech science and phonetics. The term "formant" is often used to describe a wide peak (or local maximum) in the spectral distribution in the field of acoustics. With this definition, the formant frequency for harmonic sounds

is oftentimes the frequency of the harmonic whose amplitude is greatest when the sound is resonated. The key distinction between the two definitions is whether "formants" are used to describe the sound's manufacturing processes or the sound itself. Only by chance do harmonics coincide with the resonance frequency, otherwise the frequency of a spectral peak would be slightly off from the corresponding frequency.

2. PROPOSED SYSTEM

Family of wavelets is a new wavelet family that we're developing; it includes several different wavelets, including Daubichies, Coiflet, and Symlet, among others. Reducing signal noise with wavelets is gaining popularity. The versatility of wavelets lies in their ability to filter out unwanted noise from a variety of signals due to their varying orders and kinds. In comparison to other filters, the output curve of one using wavelets is much softer and more curved. With so many different kinds of wavelets available, there's a whole new world of possibilities opening up. To filter out unwanted background noise, signals may be fragmented into a range of granularities. We use wavelets to represent the signal and filter out the background noise. We have implemented wavelets and are comparing them over a range of signal-to-noise ratios. We will be able to successfully apply wavelets to the problem of noise elimination, but the top spot will go to a comparative analysis.

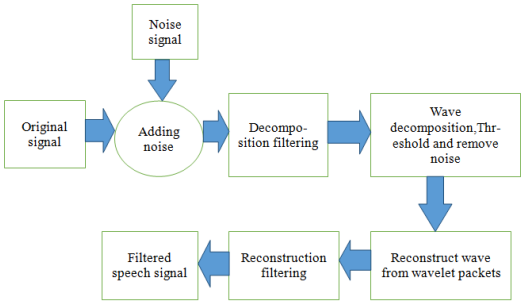

Fig1: System Architecture

3. RESULTS

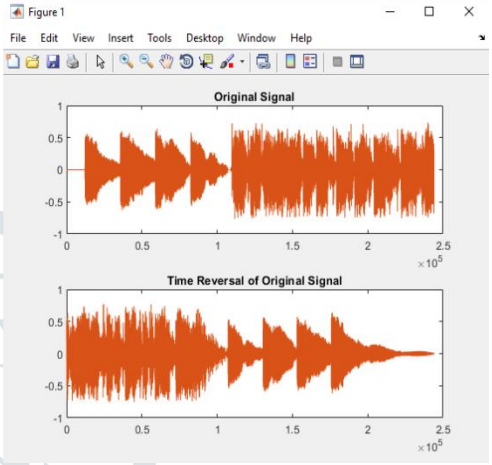

Fig 2: Original Signal and Time Reversal Representation of it

Original signal is generated based on audio signal which is given as input.

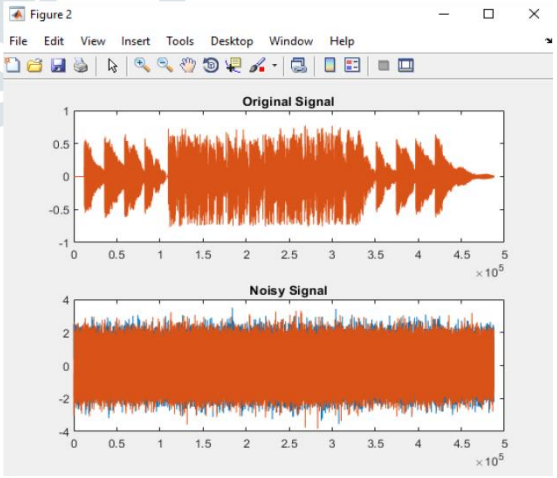

Fig3: Original Signal and Noisy Signal

Gaussian noise is added to the original signal to create the noisy signal then compare both the signals.

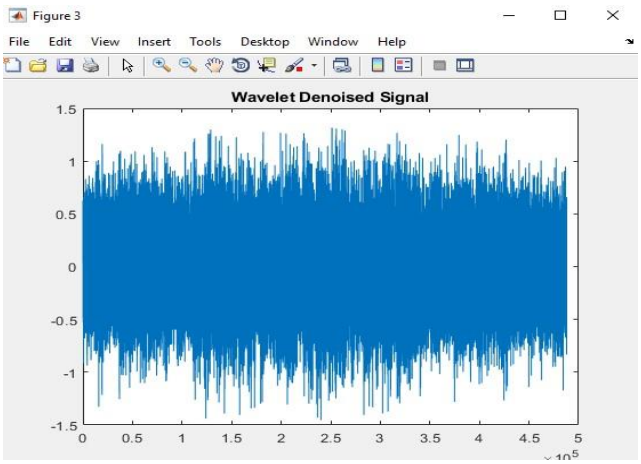

Fig4: Wavelet De-noised Signal

Decompose the noisy signal using the wavelet transforms. Threshold the wavelet coefficients to remove the noise.

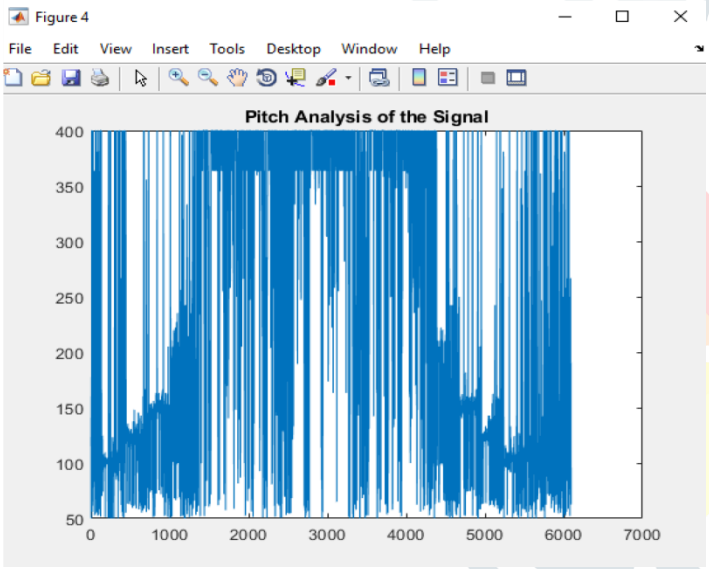

Fig5: Pitch Analysis Waveform

Apply pitch detection algorithm to the signal,such as correlation method.Analyze the pitch detection results to determine the fundamental frequency of the signal.

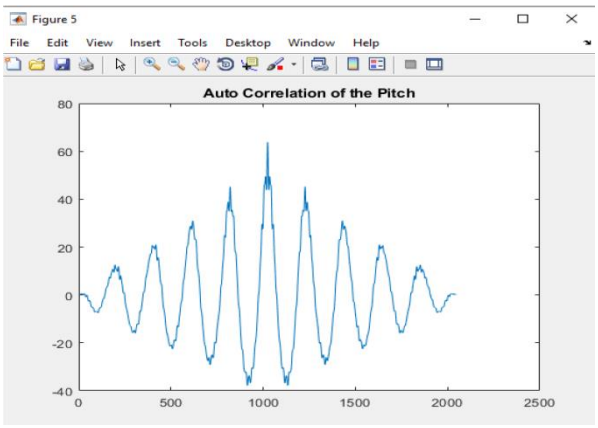

Fig6: Representation of Auto-Correlation of Pitch of the Signal

The result of the auto-correlation is a correlation function that represents the strength of the correlation between the original signal and the delayed version of it.

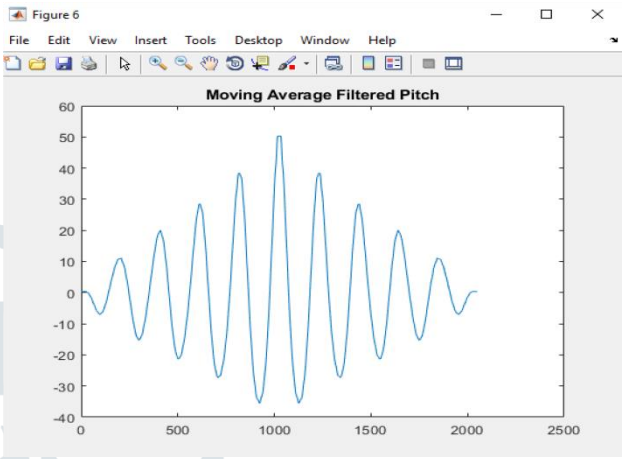

Fig7: Filtered with Moving Average Filter of the Pitch of the Signal

Moving average filter method is used for smoothing of a signal to reduce the noise and unwanted fluctuations.

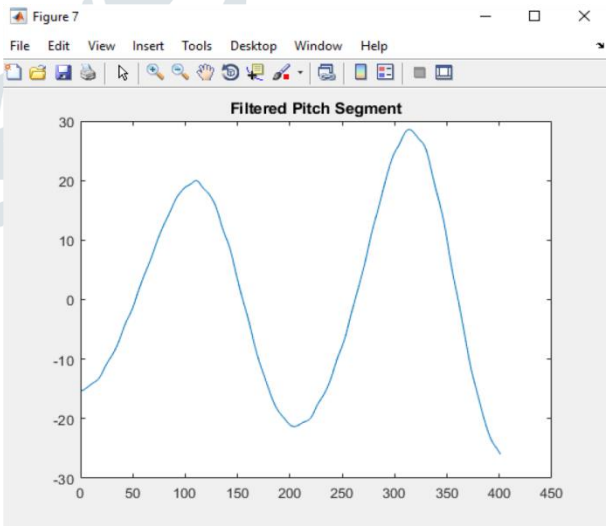

Fig8: Filtered Pitch Segment

The purpose of filtering pitch segment is to reduce noise and obtain a more stable representation of the pitch.

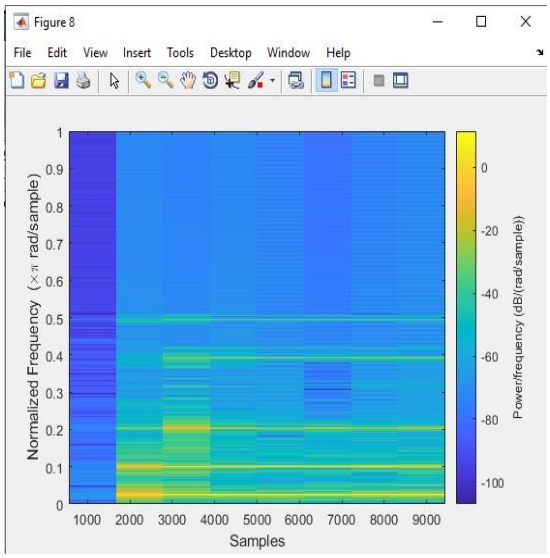

Fig9: Spectrogram of the Pitch of the Signal

Spectrogram provides a time-frequency representation of the signal, allowing one to observe how the pitch evolves over time and across different frequency ranges.

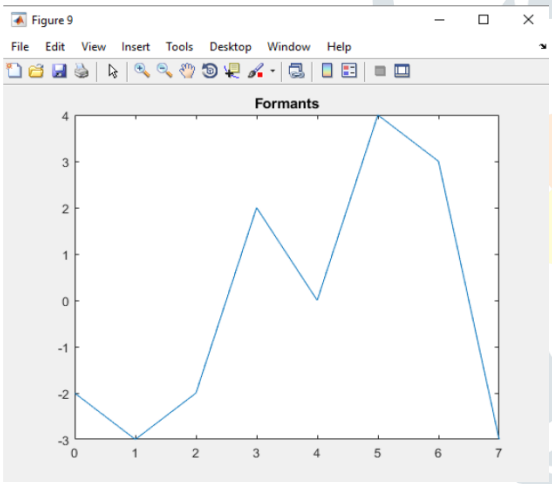

Fig10: Representation of Formants

Formants can be represented as peaks in the frequency spectrum of a speech signal.

4. CONCLUSION

Wavelets have been used effectively to denoise voice signals. Incorporating Gaussian noise into noisy audio (in wavelet form) is discussed, and a practical method for doing so is provided in this study. In this case, Wavelet Transforms will be able to effectively eliminate the backdrop. Cross correlation was introduced in MATLAB to compare two or more signals and identify the most correct one during the development of the model. Implementing wavelets for noise suppression will be a success, but the top spot will go to a competitor.

FUTURE SCOPE

Various types of wavelet based algorithms can be developed for image compression and denoising with improved performance.

REFERENCES

[1] Dr. Mahesh S. Chavan, Mrs ManjushaN.Chavan, Dr.M.S.Gaikwad,“ Studies on Implementation of Wavelet for Denoising Speech Signal ”,International Journal of Computer Applications-June 2010, Volume 3 – No.2.

[2] Dr.Parvinder Singh, Dinesh Singh, Deepak Sethi, “Reduction of Noise from Speech Signal using Haar and Biorthogonal Wavelet ” , International Journal of Electronics & Communication Technology-sept 2011, Vol. 2, Issue 3.

[3] S.Manikandan, “Speech enhancement based on wavelet denoising” Academic Open Internet Journal”,Volume 17, 2006.

[4] AseemSaxena, Amit Kumar Sinha, ShashankChakrawarti, SurabhiCharu, “SPEECH RECOGNITION USING MATLAB”, International Journal of Advances In Computer Science and Cloud Computing, Volume- 1, Issue-2, Nov-2013.

[5] Mr. Maruti Saundade Mr. Pandurang Kurle, “Speech Recognition using Digital Signal Processing”,International Journal of Electronics, Communication & Soft Computing Science and Engineering, Volume 2, Issue 6.

[6] Satish Kumar, O. P. Singh, G. R. Mishra , Saurabh Kumar Mishra , Akanksha Trivedi,

"Speech Compression and Enhancement using Wavelet Coders", International Journal of Electronics Communication and Computer Engineering, Volume 3, Issue 6.

[7] V.S.R Kumari, Dileep Kumar Devarakonda, "A Wavelet Based Denoising of Speech Signal", International Journal of Engineering Trends and Technology, Volume 5 number 2 - Nov 2013.

[8] E. Darren Ellis Department of Computer and Electrical Engineering – University ofTennessee, Knoxville Tennessee 37996 topic on "Design of a Speaker Recognition Code usingMATLAB"

[9] Topic on "Controlling of Device through Voice Recognition Using Matlab1" ISSN NO: 2250-3536 VOLUME 2, ISSUE

[10] Topic on "Extraction of Pitch and Formants and its Analysis to identify 3 different emotional states of a person" ijcsi.org/papers/IJCSI-9-4-1-296-299.pdf

[11] Topic on "Speech Recognition using Digital Signal Processing" ISSN: 2277-9477, Volume 2, Issue 6.