ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Speech Recognition on Temporal Bone Based Hearing Aid

Sahana.A, Priyadharsini.K, Srividya.C, Hemalatha.K

UG Scholar, Department of Electronics and Communication Engineering, Velammal Engineering College, Chennai, India

ABSTRACT:Speech recognition is the analysis side of the subject of machine speech processing. The synthesis side might be called speech production. These two taken together allow computers to work with spoken language. My study concentrates on *isolated word speech recognition.* Our topic might better be called *automatic speech recognition* (ASR) by temporal bone. A simple computer experiment, using MATLAB,is described in detail. We experimented with several different recognition algorithms and I used training and testing data. My training and testing data was collected and recorded with both male and female voices**.** In particular, development of the first piezoelectric, micro-electro- mechanical system (MEMS) vibration that satisfies the stringent acoustical specifications of the hearing aid industry will be undertaken. The specific aim to design and to implement the amplifying electronics for readout of the transducer's electrical signal that meet the requirements for HA applications. The technological achievability will be determined by acoustic and electrical testing of the microphone. This will be more robust to the external environment, stable, ease integration with packaging and processing microelectronics, and deliver superior acoustic performance at a smaller size. Once the design of these devices is perfected they can be made at low cost and easily integrated with processing electronics. Our goal is to develop a microphone that provides better acoustic performance than present HAs at a cost comparable to microphones used for cell phones. The availability of robust, stable, and miniature MEMS vibration will improve existing and facilitate the development of new HA systems that utilize multiple microphones (e.g., for noise cancelation) as well as those that may require close integration with downstream circuitry, such might be needed for self-fitting HA's.

KEYWORDS: Automatic Speech Recognition, MEMS Vibration, MEMS fabrication, multiple microphones.

I. INTRODUCTION

The normal recognition of sound occurs in a human when sound waves strike the tympanic membrane and cause that to vibrate. These vibrations are send to the cochlea which is in the inner ear through the tiny bones of the ossicular chain in the middle-ear. Thus the electrical impulses are transmitted through the auditory nerve to the brain. Even though the middle-ear is functioning properly, a hearing loss occurred if the inner ear is damaged.The customary hearing aid is an air-conduction that simply amplifies the sound signal. The microphone picks up the sound signal from the surrounding along with noise, strengthens them and presents them in the ear canals as an air conduction signal.

This amplified sound "overdrives" the ear's sound conducting mechanism. These types of devices offer a small frequency range and also small dynamic range of intensity.Bone conduction hearing aids have been developed for humans because the conventional hearing aids are not satisfactory. A bone conduction device is attached to the skull of the impaired listeners (i.e. the temporal or mastoid bone) and the output from a microphone is amplified and fed into any vibrator which causes a vibration is given to bone. These devices operate over a small dynamic range and are constructed for individuals whose middle ears could not be repaired surgically or for young ones who have deformities in the middle ear cannot be surgically repaired until they get older. Now, these bone conduction devices are uncommon. The another technology comprises of implanting rare earth magnets on the temporal bone and a microphone electronic coil system is employed to cause the magnet to vibrate producing bone conduction hearing. These devices are also infrequently used because of the surgery involved in drilling the bone and place the magnet inside the ear. However, their fidelity is very high. The present technology uses transposing air conduction sound signals in the conventional or audiometric range having frequency range of about 100 to about 10 kHz. These frequencies are transferred to the supersonic range above 20 kHz to about 108 kHz or higher. Then transmit these

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

supersonic frequencies by bone. The hearing aid may shift air conduction sound from the speech frequencies to the supersonic ranges (20Hz to 20kHz) such that the noise burst frequency modulated signals and quiet bursts that relate to speech frequencies will be changed to the supersonic range. These signals are delivered by a bone conduction attachment such as a large fidelity electrical to vibrator transducer. To a greater extent, piezoelectric transducer is connected to the skull for speech recognition on the temporal bone.

While the signal can be managed by analog electronics, the advancement in digitizing have allowed the signal processing to be also done in digital form before modified back to a form that can be exercised by the electrical to vibrating transducer that uses supersonic bone conduction-like signals to the skull. The signals can be cleaned to develop the speech perception by sectioning some frequencies such as frequencies below 500Hz together and attenuating them. But the critical frequencies for speech communication between 500 hertz and 2500 Hz may be resolved so that bitty differences between the frequencies can be detected. There are a number of different modifications of the signals that can be employed giving a number of different options for customizing a hearing aid to an individual. Also, filtering can be used majorly to reduce noise especially in the signal processing of digitized signals. Hearing impaired listeners normally have a difficulty in picking up speech embedded in environment noise. Reduction in noise by signal processing including filtering can be very advantageous on achieving the clarity of the signal.

To overcome the problems in the previously-proposed bone conduction hearing aid, a direct bone conduction hearing aid device has the signal transmitting device is held in place without uncomfortable external devices. This bone conduction device includes signal processing strategy for converting sound into an analog electromagnetic signal and produces the electromagnetic signal through the output transmitter and to be placed supercutaneously on the skull of hearing impaired listeners. The device further having a vibration generating process adapted to be implanted under the skin and secured to a skull bone of the hearing impaired listeners. This second magnet cooperates with the first magnet. First magnet transmits the acoustic signal supercutaneously on the skull of the hearing impaired listeners; second magnet receives the electromagnetic signal from transmitter and vibrates the skull bone in response to such received electromagnetic signal. In this device, vibrations are generated under the skin in response to the analog electromagnetic signal and conducted through the skull bone to stimulate the cochlea to create the sound perception in the hearing impaired listener.

II. AIM OF STUDY

While the proposed model cannot entirely cure hearing loss, deafness or restore one's hearing to normal and also requires surgery. In order to comfort all types of hearing loss patients, our aim is to design and evaluate safety insertion of aid on the temporal bone non-invasively.

III. KEY ELEMENT

The transducer is an electric to vibration type device to apply the supersonic signals as supersonic vibrations to the skull bone, at the mastoid interface. These frequencies are perceived as frequencies within a normal audiometric range by the brain and permit a brilliant understanding of what is being heard in the audiometric range even though the brain receives the signals initially at supersonic frequencies. This is a key element of the invention. Even though the frequencies are changed to supersonic vibration frequencies they can still be interpreted by the brain at audiometric frequencies as speech.

IV. RELATED WORK

1)Signal processing strategy for restoration of cross channel suppression in hearing impaired listeners has two tone suppression techniquehelps the sensorineural hearing impaired listeners.2)A Computationally Efficient Sound EnvironmentClassifier for Hearing Aidsimplemented with the available resources in state-of-the-art digital hearing aids has low error rate and robustness.3)Speech Processing For SensorineuralHearing Impairment has Binaural dichotic presentationimproves speech perception using fixed bandwidth based filters.4)Analysis And Design Of RF Power And Data Link Using Amplitude Modulation Of Class-e For A Novel Bone Conduction Implant provides Bone anchored hearing aid systemtransmits sound via vibrations through the skull directly to the cochlea.5)Low-delay Hearing Aid Based On Cochlear Model With Non-uniform Sub band Acoustic Feedback Cancellation(AFC) performs Spectral gain

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

shaping method along with AFCremoves additional delay.6)Simulation Of Hearing Loss And Hearing Aid Effects Of Electrophysiological Correlates Of Listening Effort performs Wavelet phase synchronization stability(WPSS) has noise masking, discrimination between an easy and difficult listening situation can be achieved.7)Instantaneous Binaural Target PSD Estimation For Hearing Aid Noise Reduction In Complex Acoustic Environments performs integration of target PSD estimator and binaural diffuse noise PSD estimator preserve the inter aural cues of target speech.8)Feature Selection For Sound Classification In Hearing Aids Through Restricted Search Driven By Genetic Algorithms performs genetic algorithm that restricts the space search enhance the listening comprehension when user has a change in environment.

V. PROPOSED TECHNIQUE

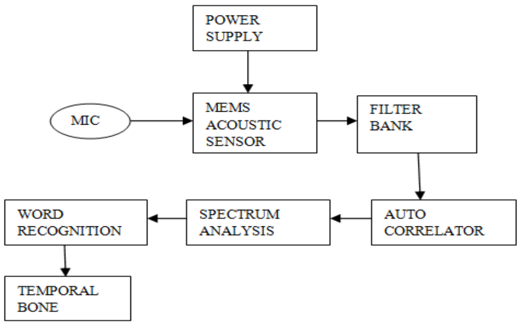
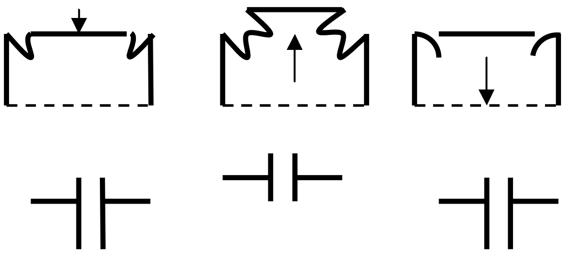


Fig.1 Block diagram of speech recognition hearing aid.

SYSTEM OVERVIEW

Acoustic signal is the mechanical waves such as vibration, sound, ultrasound and infrasound. Microphones are transducers that are transformed from acoustic signals to electric signals that can be processed by the hearing aid's audio signal chain. It uses the single silicon chip and it provides fabrication geometries, excellent stability and repeatability, and low power consumption. The power consumption and noise levels of MEMS microphones have been too high to make them appropriate for use in hearing aids. By giving a constant charge, this capacitance change is converted into an electrical signal. Then the vibrating signal is interfaced with system using 3D-DAC and recorded by using sigview.



Normal Capacitance | ReducedCapacitance due to compression | Increased capacitance due to rarefaction

Fig.2.Capacitance of MEMS microphone varies with acoustic wave

A. Filters

Filtering is the signal conditioning function, usually not all the signal frequency spectrum contains valid data. Sometimes it contains only noise. In 50/60 Hz AC power lines in most environments, will produce noise when amplified.

Time-Frequency Representations

Any signals have been represented in a time domain or in a frequency domain. In the time domain, we can look at an audio signal as magnitudes sampled at given times. In the frequency domain, an audio signal is represented as

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

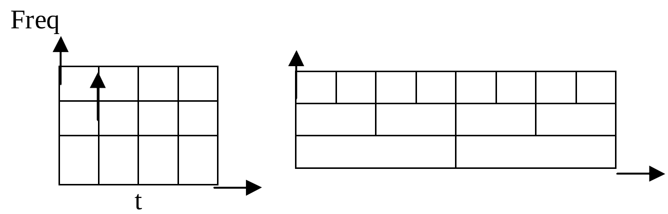
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

magnitudes of sine waves at given frequencies. But for audio analysis, it is quite interesting to have an audio signal over both components, as in temporal view of notes of different pitch. A time-frequency representation is a view of a signal represented over both time and frequency.

i. Wavelet transform

This is used effectively to suppress noises that are out of frequency band of the signal. The resultant signal of this transform is best compared to other technique with minimum loss. The Wavelet Transform gives a solution to the problem of fixed resolution in STFT is multi resolution analysis. The Wavelet Transform uses different window sizes for different regions. Long time intervals require low frequency information and shorter intervals require high frequency information.



Short Time Fourier Transform (fixed resolution) — Wavelet Transform (Multiresolution)

Fig.3 Difference between STFT and WT

Fourier Transform breaks down a signal into its constituent sine signals, Wavelet Transform is the breaking up of a signal into shifted and scaled signal of the original wavelet which is called *mother wavelet.*
The mother wavelet is an oscillating waveform effectively has an average value of zero. While Fourier analysis uses sine signal which are smooth and predictable, wavelets are asymmetric and irregular.
During analysis, a number of Fourier coefficients are obtained, which when multiplied by a sine of its corresponding frequency, yield the constituent sine components of the original signal. The result of the Wavelet analysis is many wavelet coefficients, which are a function of frequency and time domain. Multiplying each Fourier coefficient by the scaled and shifted wavelet yields the constituent wavelets of the original input signal. Scaling a wavelet means stretching or compressing it, while shifting a wavelet means delaying or hastening its onset.

Haar Transform:

Haar wavelet split the input signal into two signals called averages related to approximation coefficients and differences related to detail coefficients
If we have an input signal sj, which has 2j samples sj, k, is split into two signals sj-1 with 2j-1 averages s-j-1,k and dj-1 with 2j-1 differences dj-1,k. The averages sj-1 as a resolution representation of the signal sj and of the differences dj-1 as the information needed to go from the representation back to the original input signal. Applying the same transform to the coarser signal sj-1 itself, and iteratively repeating this process we obtain the averages and differences of successive levels, until obtain the signal s0 on the very coarsest scale, which a single sample s0,0, which is the average of all the samples of the original signal, that is the DC component or zero frequency of the signal.
The whole Haar transform applying a N x N matrix (N = 2n) to the signal sn. The cost of computing the transform requires O (N) operations.TheHaar transform uses a predictor which is correct in case the original signal is a constant. It eliminates the zeroth order correlation. The order of the predictor is one. Similarly the order of the update operator is one as it preserves the zeroth order moment.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

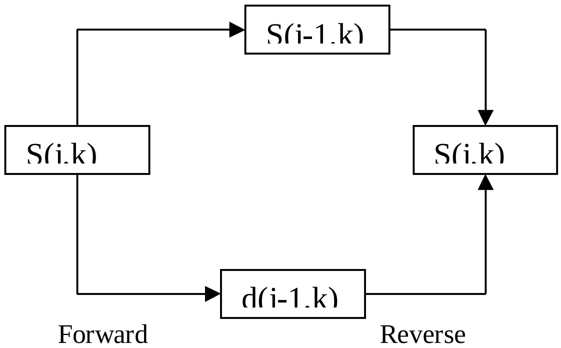


Fig.4 Haar Transform single level step and its inverse

ii. Short Time Fourier Transform

The Fourier Transform is a great tool in signal processing which breaks down a signal into its constituent sine signal of different frequency components. Thus the Fourier analysis transforms our input signal view from time domain to frequency domain. But in Fourier analysis, when a particular event took place (i.e) during transformation to the frequency domain, time domain information is lost. If the signal don't shift much over stationary signals, this drawback isn't that much considered, but most signals contain numerous transitory characteristic or non stationary and we lost a very important information.

The Short-Time Fourier Transform (STFT) was build up to correct this difficulty in Fourier analysis, while adapting the Fourier Transform to analyze the windowed sections of the signal along the time rather than frequency. Thus, the STFT maps the signal into a 2D function of frequency and time in a sort of compromisation between the frequency- and time -based views of a signal.

Even though, the STFT has a fixed resolution we obtain a limited precision information, and that precision is resolved by the size of the window. By this way, a wide window gives better frequency resolution but poor time resolution, while a good time resolution but poor frequency resolution is given by narrower window. These are known as wideband and narrowband transforms.

iii. *Recursive Least squares*

RLS algorithm has high tracking ability and convergence speed but this benefit comes from the high computational cost. At each moment, Recursive least squares (RLS) algorithm performs a precise minimization of the whole of the squares of estimation error. The processing starts with known initial conditions also, the information contained in the new data samples. These equations are to introduce the algorithm P (n) which is the inverse correlation matrix should be made equivalent to where δ is regularization component is a little positive constant.

$$Y(n) = F(n).\ U(n)$$
$$\alpha(n) = G(n) - F(n)\ u(n)$$
$$\pi(n) = P(n-1)\ u(n)$$
$$k(n) = \lambda + \pi(n)\ u(n)$$
$$K(n) = p(n-1)U(n)K(n)$$
$$F(n) = F(n-1) + K(n)\ \alpha(n)$$
$$P1(n-1) = K(n).\ \pi(n)$$
$$P(n) = \{ P(n-1) - P1(n-1) \} / \lambda$$

Where,
F(n) = filter coefficients, K(n) = gain vector,
λ= forgetting factor,
P(n) = inverse correlation matrix of the inputsignal α (n),
π(n) = positive constant.

The Recursive least Square (RLS) algorithm is nothing but subtracting input signals by noise. The RLS Adaptive Filter uses the signal called reference signal on the Input port and the pined signal on the wanted port to

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

facilitate the response of the filter. Since it meets to right filter model, the shifted noise is subtracted and the error signal should contain the original audio input signal. In the model, the signal output at the upper port of the Acoustic signal. Environment subsystem is the foundation of noise. The signal output at the lower port is made out of shaded fuss and a signal from a .wav document file. This specimen model uses an adaptive filter to remove the disturbances or noises from the signal output at the lower port. While generation of signal, we can hear both bustle and a unique individual original sound. Finally the adaptive filter in the model channels so that we just hear the original audio sound.

B. Auto-Correlation:

Measurement of uncertainty and noise makes more difficult to spot oscillatory behavior in a signal, such behavior is an expected one. The autocorrelation sequence of a periodic signal has the same cyclic characteristics as that of the signal itself. Thus, autocorrelation can help to verify the presence of cycles and to determine their durations. Take down the double periodicity of the signal. It restricts the find peaks to look for peaks separated by more than the short period with a minimum height. Sensorineural hearing loss is initially characterized by a loss of sensitivity to sounds that is more severe for low-level signals compared to high-level signals. In order for a hearing aid to restore loudness to normal levels, the gain supplied must change with the level of the input signal. The gain must also adjust fast enough so that the different amounts of gain and noise is provided to successive phonemes of different level. If the gain is reduced for a high-level vowel, the gain increase very quickly so that a following consonant is
audible to the hearing aid listener and vowel does not forward-masked it. This level-dependent gain is known as
syllabic compression or fast-acting and can be designed
in terms of compression ratio, the knee point level at which the
gain transitions from linear to compression set at a low level so that speech falls in the compression region, and the gain at the kneepoint.

C. Spectrum Analyzer

Spectrum analyzer measures the input signal is electrical, but, spectral compositions of the other signals, such as optical light waves and acoustic pressure waves, can be considered through the use of a correct transducer. By analyzing the differentspectra of electrical signals, dominant frequency power, harmonics of the signal, distortion, bandwidth of the signal, and other spectral components of a signal can be examined which are not easily evident in terms of time domain waveforms. As real-time spectrum analyzers, using a hybrid technique where the incoming signal is first converted to a lower frequency using super heterodyne and then analysing using (FFT) technique. The analyzer sample the incoming radio frequency spectrumin the time domain and convert the information to the frequency domain using the FFT process. FFT's are constructed in parallel, without ant gap and overlapped so there are no gaps in the calculated RF spectrum and no information is missed.

D. Word Recognition:

Scientists have proposed new models for word recognition that can be programmed into computers. As a result, Computers can mimic how a human would perceive and react to novel words and language.Correlations also exist between spoken language development, reading ability, and learning disabilities. However, advances in any one of the areas may boost the understanding in inter-related subjects. Finally, the advancement of word recognition may facilitate the breakthrough between how we are reading to learn and how we are learning to read. Through this we can find out the correct word from the vibrating signals.

E.*Computer Simulation*

The visual measure developed with this algorithm, the spectrogram, was formed by a matrix having different columns directly proportional to equalized band levels or
to the energy levels calculated at the each filter’s output. Initially,this computer illustration was an effective simulation for the proposed algorithm before implemented on prosynthesis DSP processor , i.e. this visual measure would allow to test speech processing used sound equalisation and for generating stimulating electrical pulses by displaying band-limits filters and filters outputs. But the popular advantage pointed out from this dynamic illustration was the therapeutic aided-tool which is used to help clinicians during the first stimulation process and during re-education.The columns that are displayed indicated the affected frequency bands to stimulation channels and an estimation of electrical stimulating pulses to be delivered to cochlear nervous endings. It would be better to tryinitially the algorithm by using the already proposed visual reference, that permitted to fix the safety measures to take into account for the

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

electrical stimulating pulse, before being injected in cochlea's biological tissue. Hence, it would appreciably help clinicians when doing more refinement during assisting the different patients.Basically, a typical sounds were composed of numerousharmonics, we use only one harmonic (pure sound) individually , for activating one specific channel. With this first test, all the channels could be verified individually and hence the functionality of the spectrogram.

VI. CONCLUSION AND FUTURE WORK

The ability to recognize amplified speech in quiet is affected by severe to profound hearing loss in the higher frequency region. It appears to be an inability to perceive the spectral cues of speech and data. The cochlea with missing hair cells can destruct the perception of these spectral cues. In such a situation, electric simulation can permit these spectral cues to be transmitted to the brain. Thus the words and sounds are sensed by the hearing impaired listeners without seeing the lip movement by the computers simulation and after that we need to dump the program for each and every sounds and words in DSP processor and place that behind the ear on the temporal bone. Hence the speech is recognized and the patients are closer to achieve the gift of hearing.People with disabilities can benefit from speech recognition programs. For individuals that are Deaf or Hard of Hearing, speech recognition software is used to automatically generate a closed-captioning of conversations such as discussions in conference rooms, classroom lectures, and/or religious services. Speech recognition is also very useful for people who have difficulty using their hands, ranging from mild repetitive stress injuries to involved disabilities that preclude using conventional computer input devices. In fact, people who used the keyboard a lot and developed RSI became an urgent early market for speech recognition. Speech recognition is used in deaf telephony, such as voicemail to text, relay services, and captioned telephone. Individuals with learning disabilities who have problems with thought-to-paper communication can possibly benefit from the software but the technology is not bug proof. Also the whole idea of speak to text can be hard for intellectually disabled person's due to the fact that it is rare that anyone tries to learn the technology to teach the person with the disability. This type of technology can help those with dyslexia but other disabilities are still in question. The effectiveness of the product is the problem that is hindering it being effective. Although a kid may be able to say a word depending on how clear they say it the technology may think they are saying another word and input the wrong one. Giving them more work to fix, causing them to have to take more time with fixingthe wrong word.

REFERENCES

1. Daniel M.Rasetshwane, Michael P.Gorga, "Signal processing strategy for restoration of cross channel suppression in hearing impaired listeners",vol.61,no.1.Oct 2015.
2. Roberto Gil-Pita, David Ayllon, Jose Ranilla,**"**A Computationally Efficient Sound Environment Classifier for Hearing Aids",VOL. 62, NO. 10,Jan 2014.
3. ShraddhaD.sharma, Devendra S. Chaudhari**,"**Speech Processing For Sensorineural Hearing Impairment",vol.3,issue 3,Mar 2013.
4. HamidrezaTaghavi, Bo Hakansson, "Analysis And Design Of RF Power And Data Link Using Amplitude Modulation Of Class-e For A Novel Bone Conduction Implant",vol.59,no.11,Nov 2012**.**
5. Maxim vashkevich, Elias Azarov**,** "Low-delay Hearing Aid Based On Cochlear Model With Non-uniform Sub band Acoustic Feedback Cancellation(AFC)",August 2012.
6. CorannaBernarding, J.Strauss,RonnyHannemann, "Simulation Of Hearing Loss And Hearing Aid Effects Of Electrophysiological Correlates Of Listening Effort ",Sep 2011.
7. A.HomayounKamkar-Parsi ,"Instantaneous Binaural Target PSD Estimation For Hearing Aid Noise Reduction In Complex Acoustic Environments",vol.60,no.4,Apr 2011**.**
8. Enrique Alexandre, Manuel Rosa, "Feature Selection For Sound Classification In Hearing Aids Through Restricted Search Driven By Genetic algorithms",vol.15,no.8,Nov 2007.