Novel Speech Processing Algorithm for Perception Improvement and Needed Research for Hearing Impaired

Bhagyashree M. Magdum and Pravin A. Dhulekar

Abstract Now-a-days, huge amount of people face the problems of hearing impairment. Hearing defines the capacity to recognize sound. People facing problems of hearing impairment has obscurity in classification of sound due to hearing defects. It is not easy to mimic the performance of human auditory system in its whole and thus reimburse for the hearing impairment. Due to the current technological advancement in signal processing area, a quite simple artificial hearing implant can be designed, thus we can achieve improvement in perception of the impaireds. For the design of artificial hearing implant, human auditory system is the top model to begin with. This paper focuses on the mechanism of human auditory system, types of hearing impairments, etc.

Keywords Hearing mechanism · Hearing loss · Hearing implants · Dyadic filters

1 Introduction

1.1 Hearing Mechanism

Outer ear, middle ear and inner ear are the three main divisions of human ear as shown in Fig. 1 [1]. Pinna and the ear canal together form the outer ear. Pinna helps in transforming the incoming signal at higher frequencies and thus helps in finding out the origin, direction and distance of the sound signal. The ear canal defends the ear from distinct bodies such as dust, dirt particles, etc. and keep the canal moist. Ear canal behaves as a filter and boosts the intensity of sound [1]. Eardrum splits the outer and middle ear.

B.M. Magdum (✉) · P.A. Dhulekar
Sandip Institute of Technology and Research Center, Nasik, India
e-mail: bhagyamagdum@gmail.com

P.A. Dhulekar
e-mail: pravin.dhulekar@sitrc.org

S.C. Satapathy et al. (eds.), *Proceedings of the International Conference on Data Engineering and Communication Technology*, Advances in Intelligent Systems and Computing 468, DOI 10.1007/978-981-10-1675-2_80

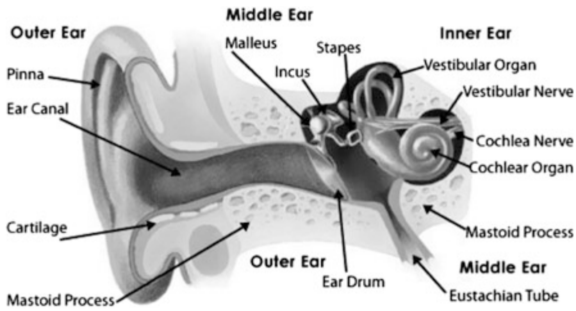

Fig. 1 Human auditory system

Middle ear is an air-filled cavity which comprises of ear drum, three ossicles and middle ear muscles. Ossicles are further subdivided as the malleus, the incus and the stapes. Ear drum is attached to Malleus at one end and the incuses attached at the other end. Incus makes contact with the stapes, the footplates of which are attached to the membrane of oval window (beginning point of the inner ear). The middle ear provides impedance matching and decreases the quantity of reflecting sound [1]. Due to impedance matching intensity of sound is about 25 dB in the middle ear. The increase in sound intensity is uneven over all the frequencies. For intense sound contraction of the muscles takes place to which the ossicles are attached and decreases transmission of sound, this is known as acoustic reflex, which plays an crucial role in defending the inner ear from extremely intense sound. Due to slow activation of acoustic reflex, impulsive signals are likely to reach the inner ear and cause damage. Reflex reduces the self-generated sounds that are emitted during chewing and before vocalization [1]. Inner ear is also called as cochlea which is a snail shaped cavity having spherical arrangement, filled with fluid [2]. The beginning of cochlea is called as oval window. Due to the incoming sound signal the oval window is set in motion due to which a pressure difference is applied athwart basilar membrane. The movement that take place depends on the involuntary properties of the basilar membrane and the rate of the incoming signal fluctuates from bottom to peak. The high-rate signals generate highest displacement at the bottom while low-rate signals generate highest displacement at the peak [1]. The organ of corti contains large number of sensory hair cells, which is situated in between basilar and tectorial membrane. The inner hair cells contains 3000–3500 hairs cells, each cell contains about 40 hairs and the outer hair cells contains 20000–25000 hairs cells, each cell contains about 140 hairs, respectively. [1]. The sound information is conveyed by means of internal hair cells which excite the tactile neurons. Outer hair cells helps in improving responses of the basilar membrane which generate high acuity, and fine-tuning to the basilar membrane [1].

2 Possible Impairment in Auditory System and Causes

Hearing impairment is generally categorized as conductive and sensorineural depending upon the location of damage in the auditory system [1]. Conductive impairment occurs due to damage in the outer ear, ear drum or middle ear, due to which the transmission of sound to the inner ear gets blocked. It generates attenuation of the stimulus and enhance both the hearing threshold levels and discomfortable loudness levels. Conductive impairment can be treated medically or surgically in most cases [1]. When the transduction mechanism of the inner ear gets damaged then such a impairment is known as Sensorineural hearing impairment. The injury caused due to deficiency in the hearing part is known as cochlear impairment [1] and the injury taking place due to defects in the auditory nerve is retrocochlear hearing impairment. The cochlear hearing impairment is associated with absence of cochlear hair cells. Impairment caused due to ageing causes malfunctions in auditory system such as stiffening of basilar membrane, of the cochlea which causes loss of hair cells in the organ of Corti and deterioration of neurons in the intact auditory system. Impairment is bilaterally symmetrical, with greater hearing impairment at high frequencies, and increased difficulty in understanding speech. Initially, the hearing loss may be mild with a loss of sensitivity in the high frequencies that continues to increase with increasing age. When hearing loss occurs due to both conductive and sensorineural loss, in other words there may be any defect in the outer/middle or inner ear or auditory nerve is called as mixed hearing loss.

3 Related Work on Hearing Aids

Hearing impairment can be partially compensated using hearing aid which amplifies acoustic signals with a frequency/gain characteristic which best reimburses for the deficiency in hearing [3]. Conventional hearing aids are usually used for addressing the problem of elevated thresholds. The hearing aids comprises of a microphone, electronic filter, control for the fine-tuning and frequency response, an earphone, a battery for supplying the power source, and elastic tubing and an ear mold, for merging the response of the earphone to the peripheral ear channel. Conventional hearing aids are categorized by their size and the way of wearing, as body-worn hearing aid, eyeglass hearing aid, etc. [2–4]. The BTE Hearing aids are the electronic components which are snowed under a tiny oval case that is fitted at the back of the ear. Receiver distributes the responses to the ear canal through a flexible tube terminating in an ear mold. These are the hearing aids most commonly used severe to profound hearing impaired. ITE hearing aid is made up of a small plastic case containing all the components, which is worn in the external section of the peripheral ear and the cochlea. Such aid is normally used by mild to moderate and moderately severe hearing impaired. ITC aid is the smallest in size, among all hearing aids that fits entirely in the ear canal. Because of small size, available power output is low. Hence mild or moderate, flat, and gradual sloping hearings impaired generally use these aids [4]. The compression techniques are used in hearing aids to overcome the problem of acoustic amplification, which compose weak signals audible to intense signals but makes uncomfortably loud due to loudness recruitment in sensorineural hearing impairment. In these techniques the amplification decreases with intensity, such that the wide dynamic range of input signal acquired compression in a smaller dynamic range of the output [3]. An epochal method is projected where the critical bands were compacted inorder to regulate the outline of the acoustic filters of hearing impaired persons. The compression using Fourier transform-based approach get better quality and intelligibility of speech for hearing impaired persons. The compression rates vary between 20 to 40 % depending individual shaped auditory filter [5]. Frequency transposition technique is a method in which energy in the high frequency region is transposed to low frequency region. The effects of the simultaneous and non simultaneous masking have been compensated by various speech enhancing schemes in case of the impairments in the inner ear. Some methods which use alteration of energy for certain frequency components are used in order to reduce the spectral masking. Schemes for enhancing the speech perception for the persons with reduced temporal resolution are based on the use of clear speech [2–4]. Masking effects plays an vital role in our day to day life. When two people are having conversion in silent environment very small amount of speech power is required for both the persons to hear and understand each other. However, the speech in the crowded shopping mall is completely impossible to hear. A signal is most likely to get masked by another signal with frequency components which are near to, or the same as, that of the signal. When masking event occurs among any two signals which appear at the same time is called simultaneous masking or spectral masking, on the other hand, if the signals are comparatively postponed in time it is called temporal masking or non simultaneous masking [6].

4 Proposed System

The proposed work is divided into two sections, first Decomposition by Dyadic Analysis Filter Bank represented in Fig. 2. And second Reconstruction by Synthesis of Dyadic Analysis Filter Bank represented in Fig. 3.

4.1 Decomposition

Figure 2 represents the decomposition part of the proposed system. First, the own recorded VCV speech signals are acquired. The sampling rate of the signal is lowered in the process of decimation. Then the signals are downsampled by factor

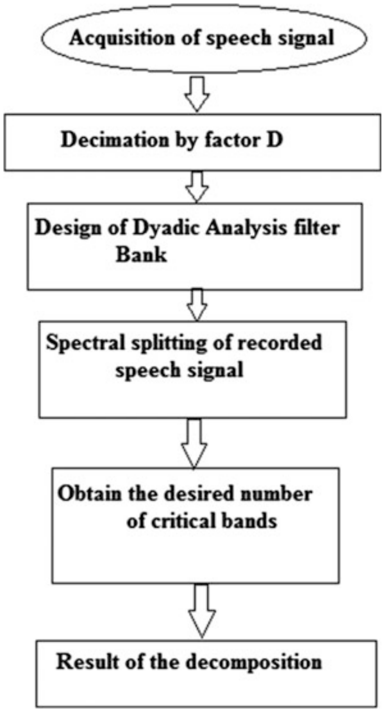

Fig. 2 Flowchart for decomposition

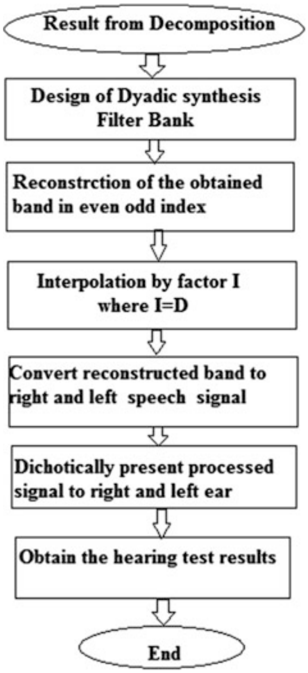

Fig. 3 Flowchart for reconstruction

Table 1 Bands as per constant, critical and 1/3 octave bandwidth

Band No.	Constant bandwidth pass band (kHz)	Critical bandwidth pass band (kHz)	1/3 Octave bandwidth
1	0.001–0.028	0.01–0.20	0.0708–0.089
2	0.027–0.056	0.20–0.30	0.089–0.112
3	0.056–0.084	0.30–0.40	0.112–0.141
4	0.084–1.111	0.40–0.51	0.141–0.178
5	1.111–1.396	0.51–0.63	0.178–0.224
6	1.396–1.167	0.63–0.77	0.224–0.282
7	1.167–1.950	0.77–0.92	0.282–0.355
8	1.950–2.227	0.92–1.08	0.355–0.447
9	2.227–2.505	1.08–1.27	0.447–0.562
10	2.505–2.782	1.27–1.48	0.562–0.708
11	2.782–3.059	1.48–1.72	0.708–0.891
12	3.059–3.336	1.72–2.00	0.891–1.120
13	3.336–3.613	2.00–2.32	1.120–1.410
14	3.613–3.891	2.32–2.70	1.410–1.780
15	3.891–4.168	2.70–3.15	1.780–2.240
16	4.168–4.445	3.15–3.70	2.240–2.820
17	4.445–4.722	3.70–4.40	2.820–3.550
18	4.722–5.000	4.40–5.00	3.550–4.470
19			4.470–5.000

D, here each subband has half bandwidth and half sample rate. Next Design of Dyadic Analysis Filter Bank is done. By filtering the signals by using the filter banks these speech signals are spectrally splitted. Then the desired number of critical bands are obtained which are listed below in Table 1. At last the desired output of the decomposition part is obtained.

4.2 Reconstruction

Figure 3 represents the part B that is the Reconstruction part of the proposed system. In decomposition stage we decimate the speech signals here we reconstruct the signals. Dyadic Synthesis Filter bank reconstruct the signals in even and odd index. The next process is upsampling of the signal where I = D. Next, the reconstructed bands are converted to right and left speech signal. These signals are presented in even and odd terms i.e. if there are 16 bands then eight bands are presented to right ear and next eight bands to left ear. This presentation of different signals to different ears is known as dichotic presentation of the speech signals..

5 Results and Discussion

For achieving the above mentioned objectives, the algorithm is developed to obtain a desired number of bands of selected speech material for binaural dichotic presentation. These bands are obtained using comb filters which gives sharp transitions between selected bands. For implementing this strategy MATLAB software is used along with its Simulink expansion. In the first stage, 18 bands of Constant Bandwidth, nineteen bands of 1/3 Octave bandwidths and eighteen bands of critical bandwidth are obtained by using a couple of corresponding comb filters. There are nine pass bands analogous to auditory critical bandwidth filters in the corresponding comb filters. These bands are shown in the following tables as:

Simulation results obtained after applying above scheme of analysis on sample speech signal is shown below.

The first quadrant of Fig. 4 shows the unprocessed sample speech signal, 2nd and 3rd quadrant shows the result of splitted speech having 9 bands each with constant bandwidth.

Figure 5 represents the unprocessed sample speech signal in 1st quadrant while 2nd and 3rd quadrant shows splitted speech having 9 bands each based on critical bandwidth.

Similarly, Fig. 6 represents unprocessed speech signal and 19 bands of processed speech based on 1/3 octave bandwidth.

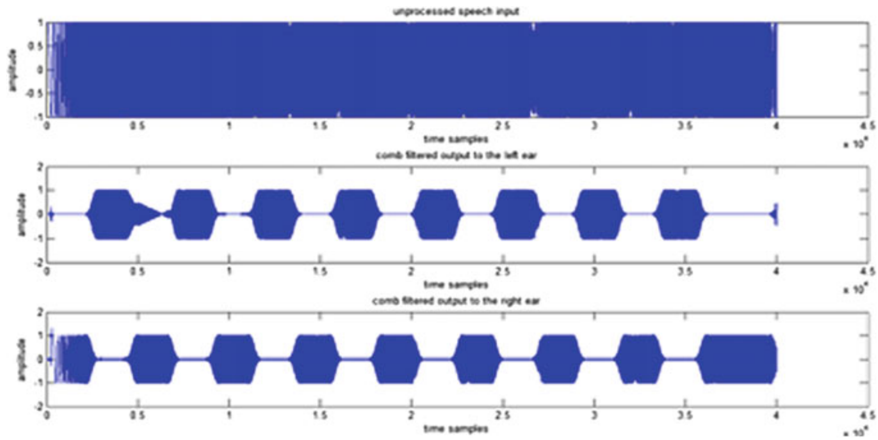

Fig. 4 Bands of processed speech signal based on constant bandwidth

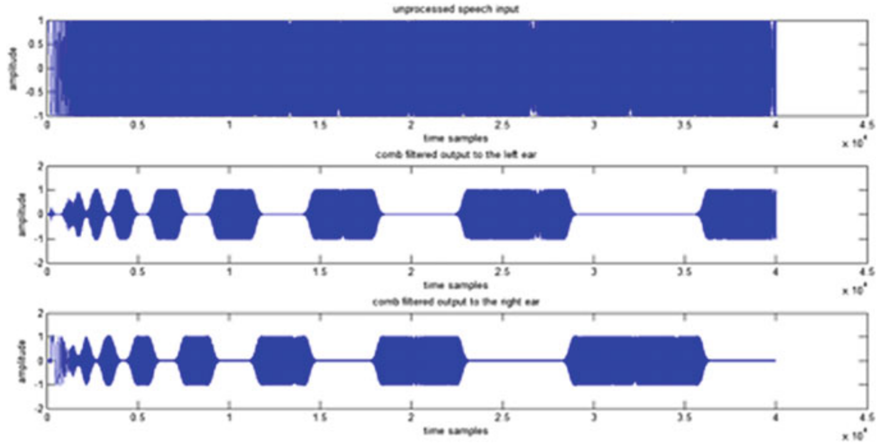

Fig. 5 Bands of processed speech signal based on critical bandwidth

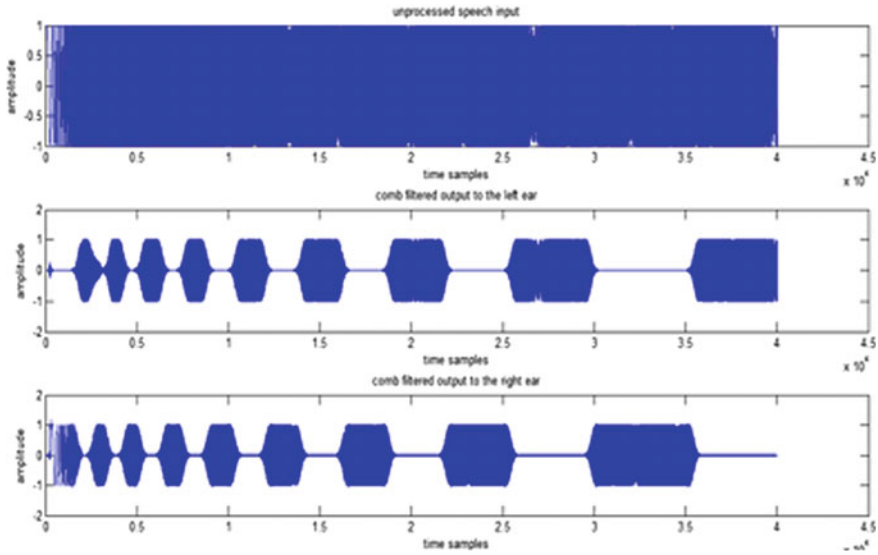

Fig. 6 Bands of processed speech signal based on 1/3 Octave bandwidth

6 Conclusion

This paper gives a brief study of the Human Auditory system, Types of hearing losses, Different hearing aids and measures to overcome the hearing losses. The proposed algorithm is useful in removing the masking effect caused due to the improper frequency selection of the people suffering from hearing loss. To obtain the desired result, a novel algorithm is designed. Eighteen bands of Constant Bandwidth, 19 bands of 1/3 Octave bandwidths and 18 bands of critical bandwidth using pair of complementary comb filters from the filter sets are obtained successfully.

References

1. Moore, B. C. J.: An Introduction to Psychology of Hearing, 4th Ed. (Academic, London), 1997.
2. Loizou, P. C., Mani, Arunvijay, and Dorman, M. F.: Dichotic speech recognition in noise using reduced spectral cues, J. Acoust. Soc. Am. vol. 114(1), 475–483, 2003.
3. GL Ward, MD MacAllister.: Apparatus and method for conveying amplified sound to ear, Google Patent, 1993.
4. CHABA.: Speech-perception aids for hearing-impaired people: Current status and needed research, J. Acoust. Soc. Am. vol. 90, 637–683, 1991.
5. Yasu, K., Kobayashi, K., Hishitani, M., Arai, T.: and Murahara Y. "Critical band-based frequency compression for digital hearing aids, J. Acoust. Sci. and Tech. vol. 25, 61–63, 2004.
6. Peng Dai, Frank Rudzicz, Ing Yann Soon, Alex Mihailidis, Huijun Ding.: 2D Psychoacoustic modeling of equivalent masking for automatic speech recognition, ELSEVIER, 2015.