International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 17 (2015) pp 38343-38348
© Research India Publications. http://www.ripublication.com

Design of Adaptive Wavelet Algorithm for Audibility Enhancement

J. J. Chopade
Research Scholar, Matoshri College of Engineering and Research Center,
Nashik, India [M.S.]
jchopade@yahoo.com

Dr. N. P. Futane
Department of Electronics and Telecommunication
Government College of Engineering,
Avasari, Pune, India [M.S.]
niteen_futane10@rediffmail.com

Abstract-Several works has been carried out to enhance the audibility of the patients suffering from inner ear impairment based on design of filters having constant, critical and octave bands. Few researchers have adopted wavelet packets with constant order for decomposition and reconstruction at each level. Here we propose a novel scheme of decomposition in which the order of wavelet is adaptive at each stage of decomposition while order of reconstruction is static. The developed algorithm has been tested on sample speech material with different wavelet families like biorthogonal, daubechies, coiflet and symlet. Comparative results of power spectrums and PSNR obtained clearly indicates the outperformance of the proposed scheme over state-of-the-art methods.

General Terms: Wavelet Packets, Decomposition, Reconstruction

Keywords: Adaptive wavelet Packets, Audibility enhancement, Power spectrums, PSNR

1. Introduction

Audibility is the measure of hearing ability of human. It is measured in terms of hearing range and sound pressure level. Hearing range is the audible range of frequencies that can be heard by humans. Generally this ranges from 20 to 20,000 Hz, though there is considerable variation between individuals, especially at high frequencies, and a gradual decline with age is considered normal [1]. Sound pressure level is referred to as sound level with respect to reference level which is 10^{-12} watts per square meter(w/m^2) equivalent to a pressure of 2 x 10^{-5} N/m^2 or 20µPa(micropascal). Thus a sound at 60dB SPL is 60dB higher in level than reference level of 0db has an intensity of 10^{-6} w/m^2 . In fact the average human absolute threshold at 1000Hz is about 6.5db SPL (when listening with one ear). A sound level is specified in this way is referred to as a sensation level(SL). The minimum sound pressure level required to hear a sound is 6.5db while above 100db SPL human ear may damage [2]. Audibility may get affected due to the effect on different region of human ear. Degradation of audibility in the outer and middle ear known as conductive hearing loss are caused due to deformity of outer ear, ear canal, or middle ear, infections in the ear near tympanic membrane and tiny bones, various allergies. While degradation of audibility in the inner ear known as senosorineural hearing loss(described in section II) are caused due to exposure to loud noise, head trauma, autoimmune inner ear disease, hearing loss that runs in the family, aging (presbycusis), malformation of the inner ear, meniere's disease and tumors[3,4].Objective of the proposed work is to enhance the audibility of the human ear which is degraded due to above mention causes by preprocessing the speech signal before feeding to the ear. In the recent year lot of work has been reported in the area of enhancing the audibility of the normal as well as hearing impaired by applying the filters like constant band, critical band, octave band and comb filters. Also few researchers have adopted wavelet packets with constant order for decomposition and reconstruction at each level (described in the section III). Here we propose a novel scheme of decomposition in which the order of wavelet is adaptive at each stage of decomposition while order of reconstruction is static, details of the scheme is discussed in the section IV.

2. Inner ear impairment

Inner ear impairment becomes significant due to the recruitment of loudness and masking effects [5, 6]. These characteristics are explicated in the following subsections.

2.1 Loudness recruitment

Loudness recruitment is an abnormally rapid growth of loudness with the increase in sound intensity. It is possibly associated with hair cell damage and particularly due to damage to the outer hair cells, as outer hair cells increases the sensitivity for low input sound levels, leaving the response to high level sounds unaltered. In an ear with recruitment, the absolute threshold is elevated but the level of uncomforted sound remains almost normal. It results in reduced dynamic hearing range (range between hearing threshold and loudness discomfort level) and loss of speech intelligibility [7].

The equivalent rectangular band of the auditory filter increased with increasing age and the slope of the lower skirt decreased with increasing age. In contrast the slope of the upper skirt was more correlated with absolute threshold, decreasing with increasing absolute threshold [8].

2.2 Frequency selectivity and temporal resolution

Frequency selectivity is the ability to resolve the sinusoidal components in a complex sound and useful in auditory perception. It refers to the ability of the auditory system to resolve among simultaneously presented different frequency components in a complex signal. The masking effect is used to explain and quantify frequency selectivity. Masking is a phenomenon in which threshold of audibility of one signal component is raised by the presence of another component. A signal is most likely to get masked by another signal with frequency components close to, or the same as, that of the signal. It is important to differentiate between frequency selectivity and frequency discrimination. Frequency discrimination is ability of the auditory system to detect changes in successively presented tones that differ in their frequency content over time. It is found that the frequency selectivity is reduced in persons with sensorineural hearing impairment [2, 9].

Peripheral auditory system can be modeled as a bank of band pass filters, called auditory filters with overlapping pass bands. Each location on the basilar membrane behaves like a filter with different centre frequency. The shape of the auditory filter by measuring the threshold of a tone as function of the bandwidth of a band pass noise masker centered at the tone frequency. The threshold of the signal increases at first with an increase in the noise bandwidth, but then remains constant irrespective of increase in noise bandwidth. It's bandwidth at which the signal threshold ceased to increase, as CB. Critical bandwidths are 15–20% of the centre frequency above 1 kHz and nearly constant below 500 Hz. The CB concept is useful to explain the frequency selectivity of the auditory system [10]. Recent calculations approximately for auditory-filter shape are used to derive a simple formula relating the equivalent rectangular bandwidth (ERB) of the auditory filter to centre frequency. The value of the auditory-filter bandwidth continues to decrease as centre frequency decreases below 500 Hz. A formula is also given relating ERB rate to frequency.

Wider auditory filters are observed in listeners with sensorineural hearing impairment. Auditory filters which are broad due to the impairment resulted from increased upward spread of masking (amount of masking raises non linearly on high frequency components sides) in sensorineural impairment [2]. Researchers have reported that, frequency selectivity for listeners with cochlear hearing impairment is inferior to normal, which leads the reducing of speech perception.

Ability to follow variations in time pattern of sounds is known as temporal resolution. It is given by the minimum duration of a silence interval between two stimuli that can be detected. Changes in time pattern of a signal are usually associated with changes in magnitude spectra, which pose difficulty in measuring true value of temporal resolution. The reduced temporal resolution is associated with increased forward and backward masking of weak signal components by adjacent strong ones. In forward masking, signal follows the masker whereas, in backward masking signal precedes masker. The effectiveness of forward masking is within 20ms after onset of masker and it reduces with time while remaining significant up to 200ms. Both forward and backward masking is caused by temporal overlap of cochlear responses [11].

2.3 Effect of damage to hair cells and auditory nerve

The responsibility of hair cells in the transduction mechanism and role of auditory neurons is to convey the information to the brain. Outer hair cells enhance the basilar membrane responses to low level sounds which results in increase in the sensitivity and they also sharpen the tuning (frequency selectivity) of the basilar membrane. Inner hair cells play role in detection of the signal and excitation of the neurons [12].

Damage to the inner hair cells reduces their transduction sensitivity for basilar membrane vibrations, leading to increased hearing thresholds at all the frequencies, without affecting the tuning curves (frequency selectivity). The reduction of sensitivity at low sound levels and loss of frequency selectivity causes due to damage to outer hair cells. Severe spectral masking occurs due to broadening of the tuning curves. Damage to the auditory nerves generally impairs the place coding of the frequencies as the place representation of frequency along the basilar membrane is preserved in the auditory nerve, resulting in reduced frequency selectivity and also in loudness recruitment and reduced dynamic hearing range [13].

3. Preprocessing by filters and wavelet packets

Presenting two different signals to the two ears known as dichotic presentation will significantly reduces the problems associated with inner ear like loudness recruitment, spectral and temporal masking which are detailed in previous section.For dichotic presentation, preprocessing of speech signal by filters having constant bands, critical bands and octave bands as well as the wavelet packet with constant order for decomposition and reconstruction is described in further subsections.

3.1 Filters

For enhancing the audibility which is degraded due to inner ear hearing loss, filters with eighteen bands of Constant Bandwidth, nineteen bands of 1/3 Octave bandwidths and eighteen bands of critical bandwidth are designed. Pair of complementary comb filters from the filter sets is formed [7, 13]. Each comb filter has nine pass bands corresponding to constant, critical and octave bandwidths. These bands are shown in table I.

TABLE 1 Pass bands as per Constant, Critical and Octave Bandwidth (in kHz)

Band No.	**Constant Band**	**Critical Band**	**Octave Band**
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

3.2 Wavelet Packets

Decomposition scheme based on wavelet packet in which ten bands was created to present them dichotically in even and odd manner (5 bands each) to left and right ear concurrently. Fig 3.1 shows the wavelet packet decomposition with constant order.

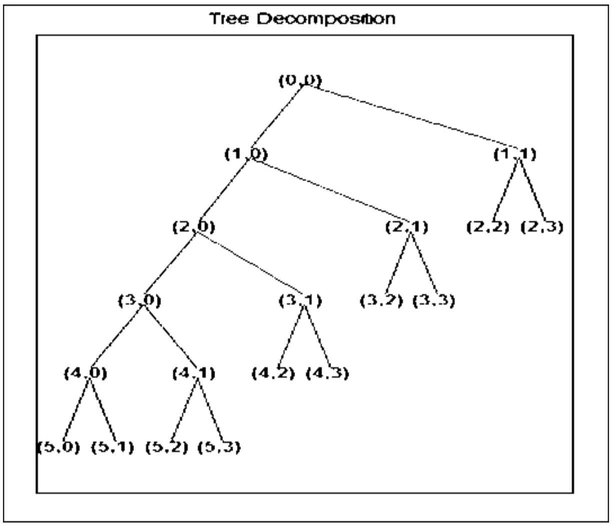

Fig 3.1 Wavelet packet decomposition with constant order

In this work author used same order of wavelet at each level of decomposition and same order is used for the reconstruction as well. Wavelet with adaptive order is discussed in next section.

4. Adaptive wavelet packets

In this proposed scheme we used wavelet packets with adaptive order for decomposition i.e. at each level of decomposition order of wavelet is different. As level of decomposition increases, order of wavelet decreases which means the first level of decomposition has highest order while last level of decomposition has lowest order [14].

For this experimentation wavelet family like duabechies, coiflet and symlet are chosen which gives best result among others while in the fourth wavelet family biorthogonal, order of decomposition and reconstruction is constant i.e. bior 2.4.

In this scheme eight bands are created to present them dichotically in even and odd manner (4 bands each) to left and right ear concurrently. Adaptive wavelet decomposition scheme is shown in fig 4.1 below:

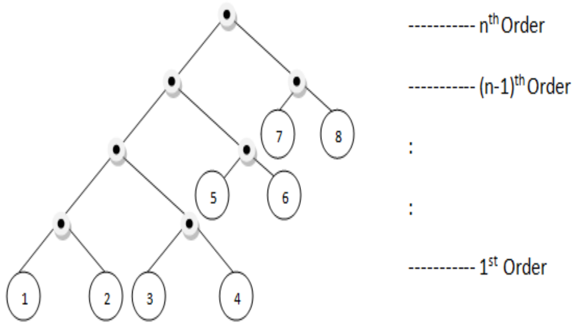

Fig 4.1 Wavelet packet decomposition with Adaptive order

The stepwise workflow of adaptive wavelet packet method for enhancing the audibility of subject having inner ear impairment is shown in following algorithm:

4.1 Algorithm

1. Reading the speech signal
2. Selection of wavelet family
3. Decomposition at first level with order n, if number of decomposition level is n.
4. Decomposition at next levels with order n-1, n-2, n-3…till order become 1.
5. Obtaining the eight bands as per the decomposition tree shown in fig.4.1 by dyadic analysis filter bank.
6. Reconstruction at first level with least order whichever used at the time of decomposition.
7. Apply odd bands (1, 3, 5, 7) to one dyadic synthesis filter bank and even bands (2, 4, 6, 8) to second dyadic synthesis filter bank for reconstruction.
8. Output of first synthesis bank (reconstructed odd bands) is fed to right ear while Output of second synthesis bank (reconstructed even bands) is fed to left ear for dichotic presentation.
9. Further reconstruction is possible by applying output of first and second synthesis filter bank to next synthesis filter bank with same order.

5. Results and discussion

In this work we have applied all wavelet families of which four families biorthogonal, daubechies, coiflet and symlet has given best results in terms of PSNR value and power spectral densities compared to other wavelets. Adaptive wavelet with these four families has been tested on the sample speech material “chirp.wave”. power spectrum magnitude of sample signal is shown in figure 5.1.

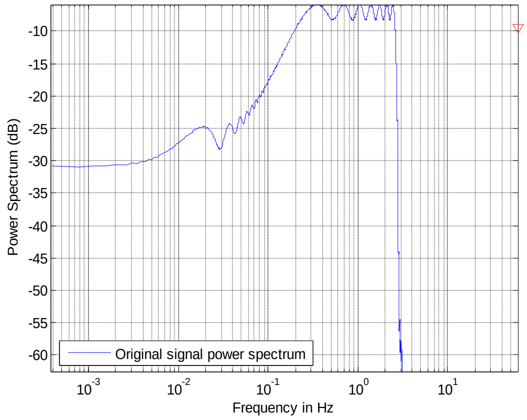

Fig. 5.1 Power Spectrum of Original Signal "Chirp.wav"

In the first phase of testing we have designed the three filters having constant, critical and octave bandwidths as shown in table 1. Out of these three filters octave band filter gives better results, so here we have shown power spectrum results of octave band filter only. Fig 5.2 and Fig 5.3 shows Power Spectrum of Left and Right Ear Signal obtained using 1/3 Octave Band Filter. While Fig.5.4 shows the reconstructed signal.

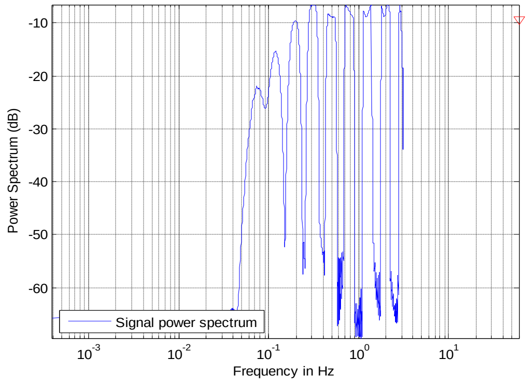

Fig.5.2 Power Spectrum of Left Ear Signal using 1/3 Octave Band Filter

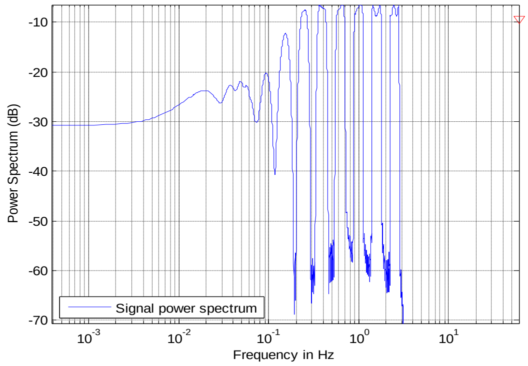

Fig. 5.3 Power Spectrum of Right Ear Signal using 1/3 Octave Band Filter

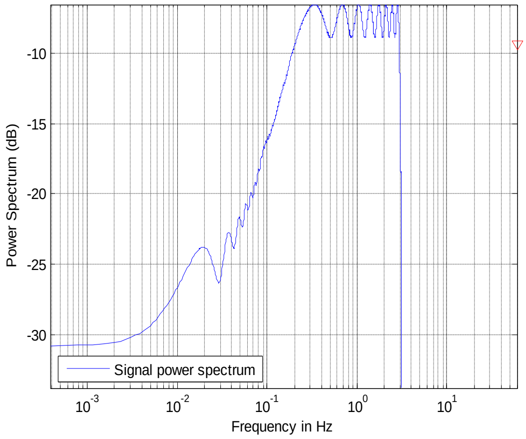

Fig. 5.4 Power Spectrum of Reconstructed Signal using 1/3 Octave Band Filter

In the second phase of testing we designed the adaptive wavelet packet with four families' biorthogonal, daubechies, coiflet and symlet. Out of four families biorthogonal wavelet gives best results, so here we have shown power spectrum results of biorthogonal only. Fig 5.5 and Fig 5.6 shows Power Spectrum of Left and Right Ear Signal obtained using biorthogonal wavelet with adaptive order for decomposition. While Fig.5.7 shows the reconstructed signal.

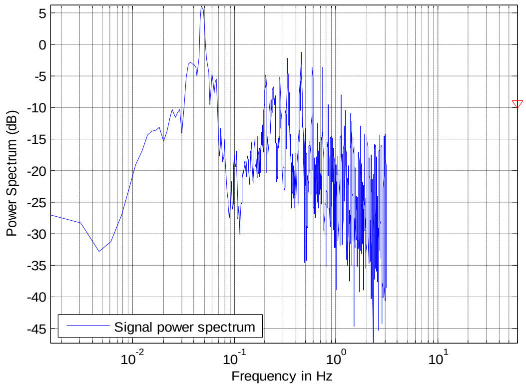

Fig. 5.5 Power Spectrum of Left Ear Signal using Biorthogonal wavelet with adaptive order for decomposition

TABLE 2 PSNR, Mean, Variance and Standard Deviation obtained by applying Filter Bank with constant, critical and octave bandwidth

Type	**PSNR**	**Mean**	**Variance**	**Standard Deviation**
Constant Band	96.3007	0.0025	0.5059	0.7113
Critical Band	96.3180	0.0025	0.4981	0.7057
Octave Band	96.3237	0.0025	0.4954	0.7039

TABLE 3 PSNR, Mean, Variance and Standard Deviation obtained using Adaptive Wavelet Packets with Biorthogonal, Daubechies, Coiflet and Symlet wavelet family.

Type	**PSNR**	**Mean**	**Variance**	**Standard Deviation**
Biorthogonal	97.289	0.0033	0.6101	0.781
Daubechies	96.375	0.0033	0.5010	0.707
Coiflet	96.331	0.0033	0.5007	0.707
Symlet	96.298	0.0033	0.5010	0.707

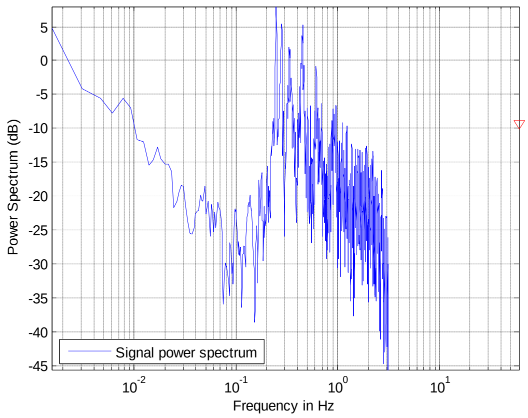

Fig. 5.6 Power Spectrum of Right Ear Signal using Biorthogonal wavelet with adaptive order for decomposition

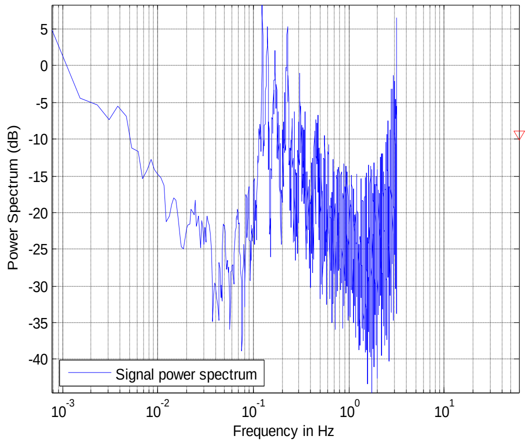

Fig. 5.7 Power Spectrum of Reconstructed Signal using Biorthogonal wavelet with adaptive order for decomposition

Table II shows the result of PSNR, Mean, Variance and Standard Deviation obtained using filter bank with constant, critical and octave bandwidths, while and Table III shows the result of PSNR, Mean, Variance and Standard Deviation obtained using adaptive wavelet packets with Biorthogonal wavelet family.

6. CONCLUSION

The method of adaptive wavelet packet gives improved peak signal to noise ratio as well as least reconstruction error compared to the conventional filter banks. Performance of the proposed scheme is also compared with wavelet packet having constant order at each stage of decomposition. Obtained result shows that varying order of wavelet enhances the audibility of subjects having inner ear impairment specially effect of loudness recruitment, spectral and temporal masking can be minimized.

7. References

[1] Dachasilaruk, S. “Improving speech intelligibility in perceptual wavelet packet-based speech coding for cochlear implants”, 7th International Conference on Biomedical Engineering and Informatics (BMEI), pp.323-328, 2014

[2] Moore, B. C. J. An Introduction to Psychology of Hearing 4th Ed. (Academic, London), 1997.

[3] Andrea V. and Hugh J. M. “Application of Real-Time Loudness Models Can Improve Speech Recognition for Cochlear Implant Users” IEEE Transactions on Neural Systems And Rehabilitation Engineering, Vol. 21, No. 1, 81-87, 2013.

[4] Hongmei Hu ; Jinqiu Sang ; Lutman, M.E. ; Bleeck, S. “Simulation of hearing loss using compressive gammachirp auditory filters”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 5428-5431, 2011

[5] Hines, A., Dublin, Ireland ; Harte, N. “Evaluating sensorineural hearing loss with an auditory nerve model using a mean structural similarity measure”, IEEE 18th European Signal Processing Conference,pp.1924-1928, Aug-2010.

[6] Yao, J., and Zhang, Y. “The Application of Bionic Wavelet Transform to Speech Signal Processing in Cochlear Implants using Neural Network Simulations,” IEEE Transactions on Biomedical Engineering, vol. 49(11), 1299–1309, 2002.

[7] Kambalimath, S.G. ; Pandey, P.C. ; Kulkarni, P.N. ; Mahant-Shetti “FPGA based implementation of comb filters for use in binaural hearing aids for reducing intraspeech spectral masking” Annual IEEE India Conference (INDICON), pp.1-5, 2013

[8] Hairong Jia ; Xueying Zhang ; Chensheng Jin “A speech enhancement method based on wavelet packet and hearing masking effect”, 2nd International Conference on Signal Processing Systems (ICSPS), pp.272-275, 2010.

[9] Medeiros, J.E.G. ; Chrisostomo, L.A.P. ; Meira, G. ; Toledo, Y.C.R “A fully analog low-power wavelet-based Hearing Aid Front-end” IEEE Biomedical Circuits and Systems Conference (BioCAS), pp.242-245, 2013

[10] Derbel, A. ; Ghorbel, M. ; Samet, M. ; Ben Hamida, A. “New algorithm of bionic wavelet transform to DSP-implementation for cochlear implant”, 16th IEEE Mediterranean Electrotechnical Conference (MELECON),pp.1025-1029, 2012

[11] Hongo, S. ;Sakamoto, S. ; Suzuki, Y. “Binaural speech enhancement method by wavelet transform based on

interaural level and argument differences", Wavelet Analysis and Pattern Recognition (ICWAPR), pp.290-295,2012

[12] Kathryn Hopkins , Brian C. J. Moore "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise" J. Acoust. Soc. Am. 130, 334, 2011

[13] Upal Mahbub, Celia Shahnaz "Exploiting Wavelet and Prosody-related Features for the Detection of Voice Disorders", American Journal of Biomedical Engineering and Technology, Vol 2, No 1, 2015

[14] Waldo Nogueira, Andreas geese, Bernd Elder, "Wavelet Packet filterbank for speech processing strategies in cochlear implant" IEEE International Conference on Acoustics, Speech and Signal Processing, 14-19 May 2006