Scientific
Research
Publishing

Design of Optimized Wavelet Packet Algorithm to Improve Perception of Sensorineural Hearing Impaired

Jayant J. Chopade[1], Niteenkumar P. Futane[2]

[1]SPPU Pune, Matoshri College of Engineering and Research, Nashik, India
[2]Government College of Engineering and Research, Avasari, India
Email: jchopade@yahoo.com, niteen_futane10@rediffmail.com

Abstract

A novel optimized wavelet packet algorithm is proposed to improve the perception of sensorineural hearing-impaired people. In this work, we have developed optimized wavelet packet along with, biorthogonal wavelet basis functions using MATLAB Code. Here, we have created eight bands based on auditory filters of quasi octave bandwidth. Evaluation was carried out by conducting listening tests on seven subjects with bilateral mild to severe sensorineural hearing loss. The speech material used for the listening test consisted of a set of fifteen nonsense syllables in VCV context. The test results show that the proposed algorithm improves the recognition score, speech quality and transmission of overall feature specifically over the unprocessed signal. The response time also reduces significantly.

Keywords

Sensorineural, Dichotic, Binaural, Masking, Wavelet Transform

1. Introduction

For sensorineural hearing impaired people, the auditory filters are wider than normal in increased spectral masking [1]. Masking proceeds primarily at peripheral level of ear and splitting of speech into two complementary signals thereby presenting them dichotically to diverse ears which might help in reducing the effect of increased masking in persons with sensorineural hearing impairment with residual hearing [1]-[3]. Our ear uses wavelet transform while analyzing sound, at least in the very first stage [4]. The wavelet transform is used in signal processing, due to the capability of wavelet transforms to existing a time-frequency (or time-scale) repre-

sentation of signals as the wavelet transform uses a variable-width window (narrow at high frequencies and wide at low frequencies).

Wavelet analysis is equivalent to a bank of band pass filters. The wavelet filter bank allows a better representation of both the temporal and the place pitch in the speech signals. Nogueira *et al*. [5] have designed a WP filter bank and incorporated it into a commercial ACE (Advanced Combinational Encoder) strategy for speech processing in cochlear implants. Averaged results of speech intelligibility tests have shown that the mixed WP filter-bank leads to significantly better speech perception performance than the fast Fourier transform (FFT) as used in the commercial ACE strategy. Yao, J. *et al*. [6] investigated the application of an improved signal processing method called bionic wavelet transform (BWT). Authors have concluded that application of the BWT in cochlear implants has a number of advantages, including improved recognition rates for vowels and consonants, reduction in the number of channels in cochlear implant, reduction in the average stimulation duration of words, better noise tolerance and higher speech intelligibility rates. Abhjit Karmarkar *et al*. [7] have proposed a criterion to select the optimal wavelet packet based on the Zwicker's model critical band structure. Authors obtained optimal WP tree for different sampling frequencies and results are compared with other CB motivated WP trees. M. T. Kolte *et al*. [8] showed that the modified wavelet packets algorithm based on auditory critical bandwidth, resulted the relative improvements in recognition scores for processed scheme of wavelet packets were 3.33% to 22.23%.

The objective of our work is to minimize the effect of spectral masking in sensorineural hearing impaired with better perception using minimum number of channels. Modified wavelet algorithm using ten bands is proposed in [8]. In this investigation, we have developed optimized wavelet packet algorithm biorthogonal wavelet family using MATLAB Code. We have created eight bands based on auditory filters of quasi octave bandwidth. [9] [10]. Four alternate bands are combined for even-odd dichotic presentation. The inverse wavelet packet transform were used to produce speech components from the wavelet packet representation. Wavelet coefficients are being employed in order to synthesize the speech components.

The paper is planned into four sections. Section 1 introduces the need of the proposed system and also reviews the different techniques proposed by the different researchers to overcome various problems related to the hearing impaired using wavelet transform. Section 2 discusses the design of optimized wavelet packet. Section 3 includes listening tests for evaluation. Listening test results and discussion are presented in section 4. Section 5 concludes this paper.

2. Optimized Wavelet Packets

The handling scheme is developed as spectral splitting with optimized wavelets packets based on eight frequency bands as the performance by hearing-impaired subjects saturated around eight channels, while performance by normal-hearing subjects is sustained to 12 - 16 channels in higher background noise [11]. The number of channels desired to obtain high levels of speech understanding is still the subject of discussion [12]. MATLAB code was developed based on optimized wavelet packet with biorthogonal wavelet functions. Biorthogonal wavelets chosen such that symmetry and exact reconstruction are possible using FIR filters. The inverse wavelet packet transform was used to synthesize speech components from the wavelet packet representation. To produce the speech component, wavelet coefficients are used.

The wavelet packet decomposition produces generic analysis signals that give richer signal analysis. The nodes of wavelet packet decomposition are known as wavelet packet atoms. Each wavelet packet atom is indexed by three parameters namely: scale, position and frequency. Unlike conventional wavelet transform, which is employed on low pass bands iteratively, the wavelet packet analysis is employed on both low pass (approximations) and high pass (details) sub bands. The conventional wavelet transform can offer (n+1) possible ways to analyze signal when the "n" decomposition levels are utilized. For wavelet packet analysis, for "n" level decomposition, it yields $2^{2^{n-1}}$ ways to encode the signal [13]-[15]. A generic wavelet packet analysis is shown in the **Figure 1**.

In the notation $W_{j,n}$ where, j stands for scaling factor and n denotes frequency parameter, the representative equations are given by (1) and (2).

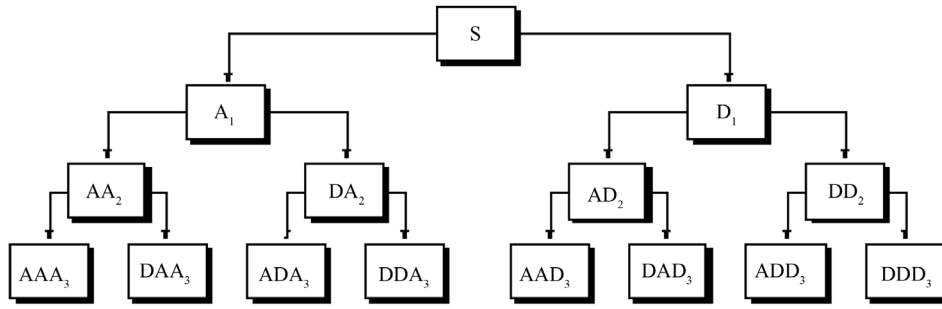$$W_{0,0} = \left(\phi\left(x - k\right), k \in Z \right) \tag{1}$$

and

Figure 1. A generic wavelet packet analysis.

$$W_{1,1} = \left(\psi \left(\frac{x}{2} - k \right), k \in Z \right) \tag{2}$$

In optimized wavelet packet, we have applied discrete wavelet transform at first level of decomposition and wavelet packet for further three levels to obtain eight bands having quasi octave bandwidth. For each level, we have applied biorthogonal family with same order of decomposition. Following **Figure 2** shows optimized wavelet packet tree.

Following **Table 1** shows that all the eight band in alternate fashion for even-odd index with centre and pass band frequency for each band in KHz.

The stepwise workflow of the new approach of optimized wavelet packet is presented in the following algorithm.

Pseudo Algorithm

- Read audio input signal $x(n)$ of length N.
- Perform wavelet packet decomposition of $x(n)$ up to level 4 as directed in **Figure 1**.
- Construct the optimized wavelet packet tree by rejoining following nodes of the original tree T: 11, 12, 13, and 14 and 9, 10, 5, 6. Thus optimized tree will have only eight nodes as shown in **Figure 2**.
- Selectively reconstruct the optimized wavelet tree to get two output signals—one for left ear and other for right ear, as follows:
- In optimized tree, make all four approximate coefficients nodes numbered 15, 17, 9, and 5, zero while keeping detail coefficients nodes as it is and reconstructed the tree.
- In optimized tree, make all four detail coefficients nodes numbered 16, 18, 10, and 6, zero keeping approximate coefficients nodes as it is and reconstructed that tree.

3. Listening Tests for Evaluation

The assessment was carried out by conducting listening tests on seven subjects with bilateral mild to severe sensorineural hearing loss. The speech material used for the listening test consisted of a set of fifteen nonsense syllables in VCV context with consonants /p, b, t, d, k, g, m, n, s, z, f, v, r, l, y/ and vowel /a/ as in "farmer". Responses were tabulated in the form of confusion matrix and response time was also been recorded. Confusion matrices were used for calculating recognition scores and relative transmitted information. Further, the consonants were clustered according to the articulatory features [16] and the contribution of different features was analyzed. The features selected for this study were voicing (voiced: /b d g m n z v r l y/ and unvoiced: /p t k s f/), place (front: /p b m f v/, middle: /t d n s z r l/, and back: /k g y/), manner (oral stop: /p b t d k g l y/, fricative: /s z f v r/, and nasals: /m n/), nasality (oral: /p b t d k g s z f v r l y/, nasal: /m n/), frication (stop: /p b t d k g m n l y/, fricative: /s z f v r/), and duration (short: /p b t d k g m n f v l/ and long: /s z r y/).

4. Listening Tests Results and Discussion

Listening tests were conducted to measure three performance parameters that are recognition score, response time and information transmission analysis. Comparative analysis of these parameters for processed and unpro-

Table 1. Eight bands for odd-even presentation.

Filter for left ear			Filter for right ear		
Band	Centre frequency KHz	Passband frequency KHz	Band	Centre frequency KHz	Passband frequency KHz
1	0.15625	0 - 0.3125	2	0.46875	0.3125 - 0.625
3	0.78125	0.625 - 0.9375	4	1.0937	0.9375 - 1.25
5	1.5625	1.25 - 1.875	6	2.1875	1.875 - 2.5
7	3.125	2.5 - 3.75	8	4.375	3.75 - 5
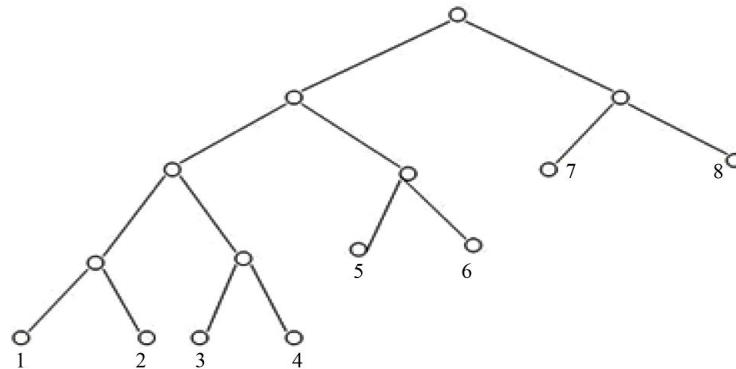
Figure 2. Optimized wavelet packet tree.

cessed signal was evaluated. The detailed analysis of these results is shown in following subsections.

4.1. Recognition Score

Figure 3 provides percentage recognition scores (%) acquired from the confusion matrix. For the impaired subjects, the recognition score for unprocessed signal varies from 48.33% to 90%, and for processed signal recognition score varies from 53.33% to 93.33%. The average values observed as 66.17% and 74.58% for unprocessed and processed signals. The average relative improvement observed was 8.40%. **Figure 4** shows the graphical representation of relative improvement in percentage recognition scores with respect to unprocessed signal.

4.2. Response Time

Response time is the time interval between speech materials presented dichotically to subjects and the response given by subjects. The response time for unprocessed signal varies from 4.08 to 8.6 seconds, and for processed signal, it varies from 3.9 to 8.2 seconds. The relative decrease in response time varies from 4.65% to 44.20%. The observed average value for processed signal was 5.01 Sec. **Figure 5** shows the comparative the results for unprocessed and processed signals. **Figure 6** shows relative decrease in response time. Response time reduces significantly showing reduction in load on perception process.

4.3. Information Transmission Analysis

Relative information transmission is used to measure the transmission performance in the context of specific features. The overall information transmitted and information transmitted for specific features were obtained for all subjects. The average overall information transmitted for Bi-ortho filter was observed as 78.62% for unprocessed signal and 85.31% for processed signal. The average relative improvement in overall information transmission observed as 6.69%. The contribution of all the six features to overall improvement was indicated by information transmission analysis. In addition, the improvement observed for the place feature. Since, the place information is linked with frequency resolving ability of auditory process, the effect of spectral masking has been reduced. Relative information transmitted for consonantal features is given in **Table 2** and plotted in

Table 2. Relative information transmitted for consonantal features.

Features	Subjects	US	PS-Bior2p4	Features	Subjects	US	PS-Bior2p4
Overall	AP	71.15	80.28	Continuance	AP	33.15	35.61
	GK	89.77	96.59		GK	41.12	69.63
	SK	62	78.91		SK	26.02	35.83
	KB	74.19	70.05		KB	11.42	29.1
	PD	74.69	86.36		PD	42.53	35.61
	PN	94.24	93.54		PN	44.43	41.27
	RB	84.3	91.47		RB	39.84	81.95
Features	**Subjects**	**US**	**PS-Bior2p4**	**Features**	**Subjects**	**US**	**PS-Bior2p4**
Duration	AP	41.12	59.28	Frication	AP	26.46	48.96
	GK	58	100		GK	53.77	71.86
	SK	41.29	35.61		SK	21.17	26.43
	KB	44.43	60.42		KB	10.52	23.72
	PD	76.74	100		PD	57.73	64.9
	PN	76.35	75.96		PN	48.96	64.77
	RB	41.16	71.23		RB	38.72	82.45
Features	**Subjects**	**US**	**PS-Bior2p4**	**Features**	**Subjects**	**US**	**PS-Bior2p4**
Manner	AP	40.45	55.85	Nasality	AP	57.1	67.57
	GK	70.73	82.06		GK	100	100
	SK	40.37	48.23		SK	71.4	78.76
	KB	15.55	29.11		KB	20.52	24.48
	PD	64.02	66.16		PD	74.54	67.57
	PN	68.05	78.39		PN	100	100
	RB	50.54	79.89		RB	67.57	78.76
Features	**Subjects**	**US**	**PS-Bior2p4**	**Features**	**Subjects**	**US**	**PS-Bior2p4**
Place	AP	27.64	44.6	Voicing	AP	100	100
	GK	61.11	91.34		GK	71.86	100
	SK	18.52	32.5		SK	63.98	78.94
	KB	25.06	41.97		KB	21.8	32.75
	PD	79.67	100		PD	37.01	71.68
	PN	89.3	84.23		PN	100	100
	RB	52.58	74.7		RB	86.94	100

Percentage Recognition Scores

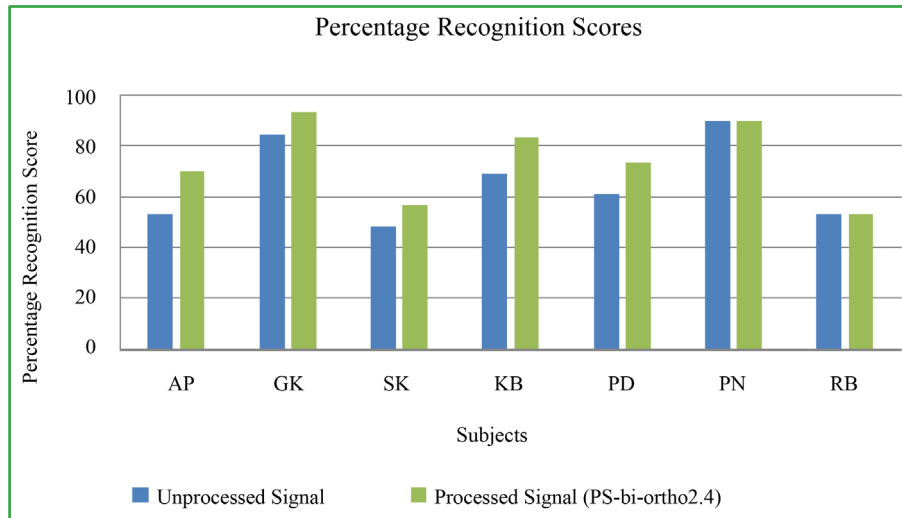**Figure 3.** Comparative results of percentage recognition scores of unprocessed and processed signal.

Relative Improvement in percentage recognition score
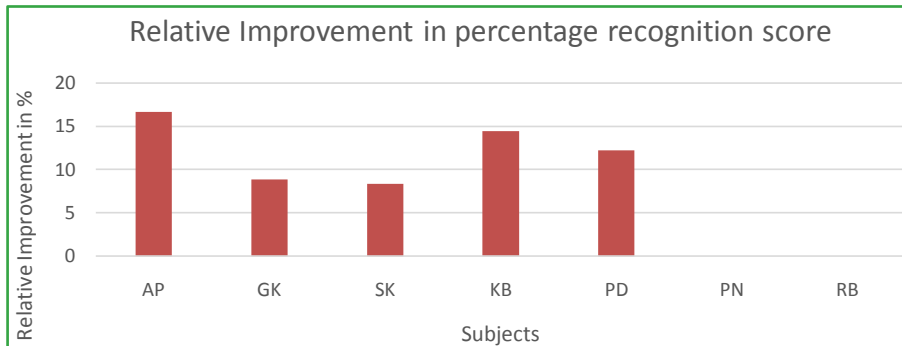

Figure 4. Relative Improvement in % with respect to unprocessed signal.
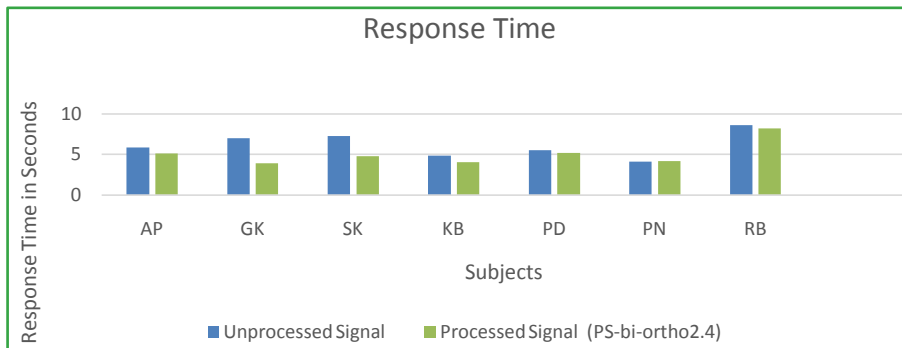
Response Time

Figure 5. Comparative result of response time of unprocessed and processed signal.

Figures 7-14.

5. Conclusion

An optimized wavelet packet using biorthogonal family based on auditory critical bandwidth is designed and implemented in the MATLAB. The experimentation results shows that signal processing scheme resulted in improvement in overall speech reception quality and significantly, improvement was recorded in recognition scores. Response time reduces significantly showing reduction in load on perception process. The contribution
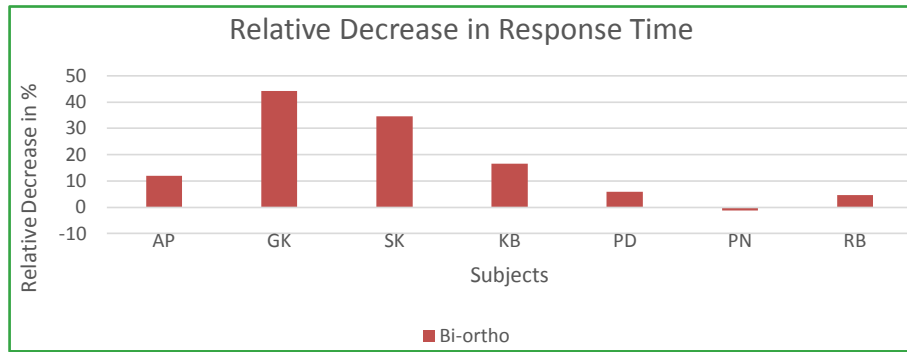
Figure 6. Relative decreases in response time.
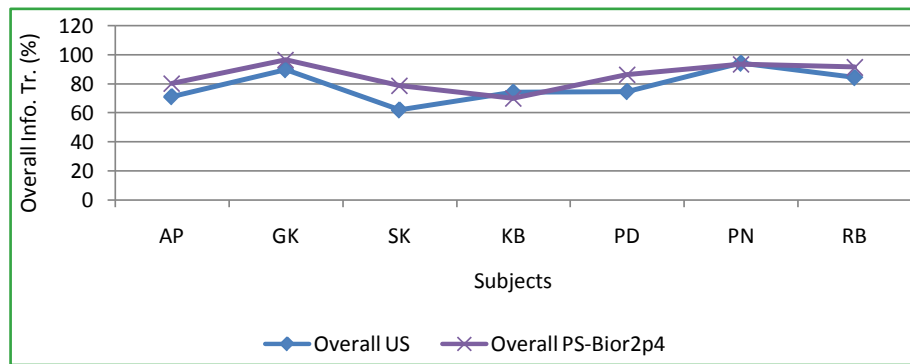
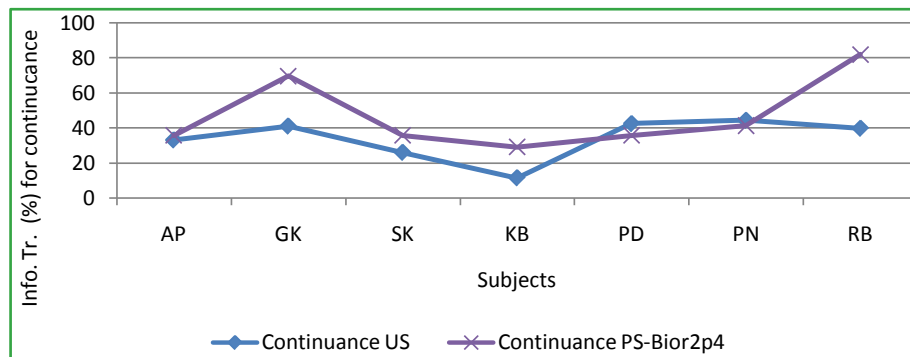**Figure 7.** Information transmitted for overall feature.

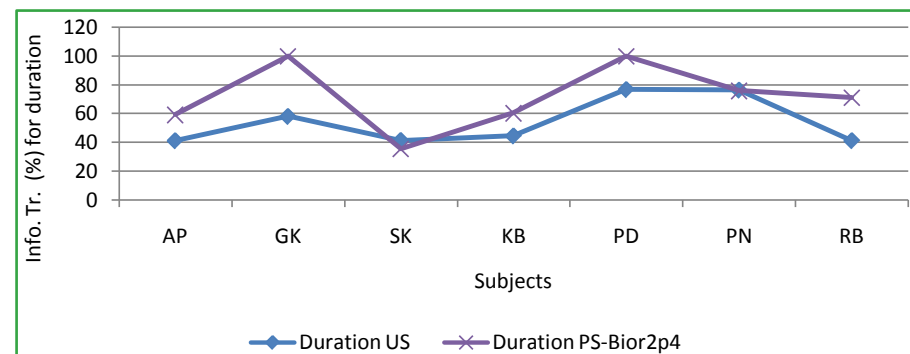**Figure 8.** Information transmitted for continuance feature.

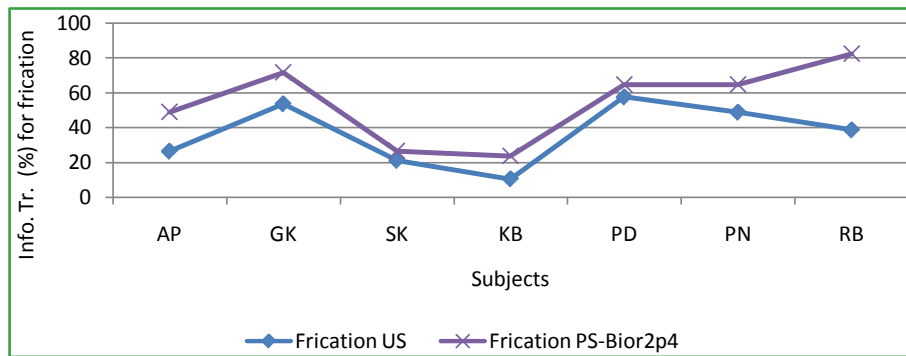**Figure 9.** Information transmitted for duration feature.

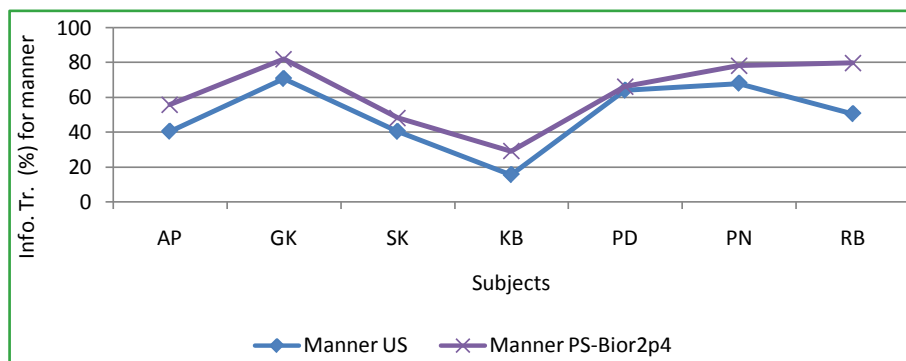**Figure 10.** Information transmitted for frication feature.

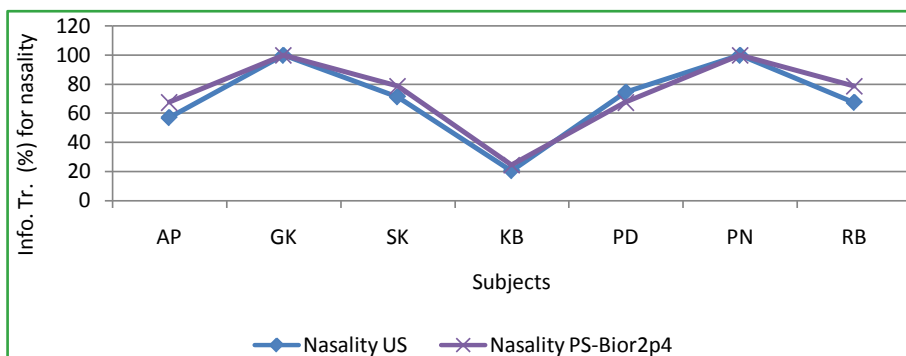**Figure 11.** Information transmitted for manner feature.

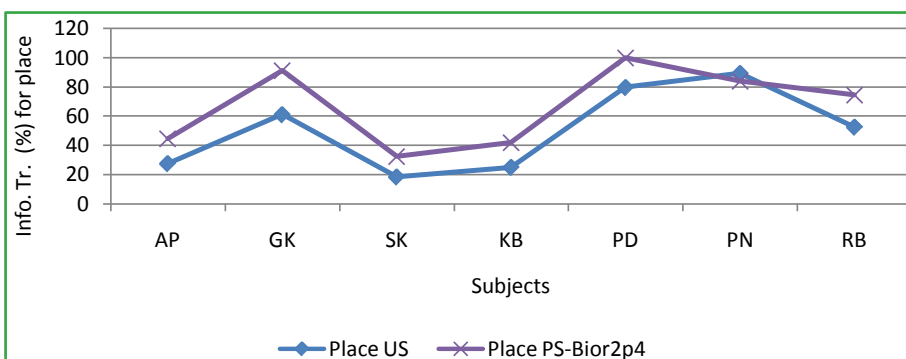**Figure 12.** Information transmitted for nasality feature.

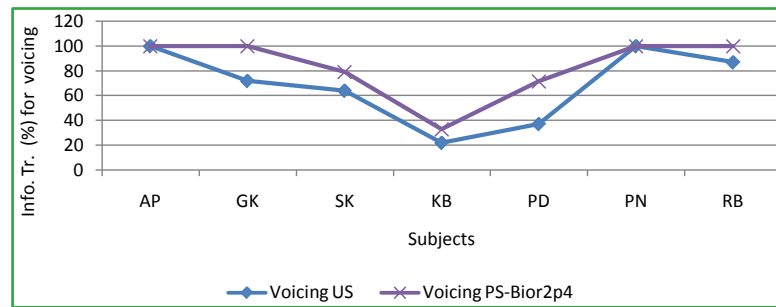**Figure 13.** Information transmitted for place feature.

Figure 14. Information transmitted for voicing feature.

of all the six features to overall improvement was indicated by information transmission analysis. In addition, the improvement was observed for place feature. Since the place information is linked to frequency resolving capacity of auditory process, the effect of spectral masking has been reduced.

References

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

[2] Kulkarni, P.N. and Pandey, P.C. (2008) Optimizing the Comb Filters for Spectral Splitting of Speech to Reduce the Effect of Spectral Masking. *IEEE-International Conference on Signal Processing*, *Communications and Networking*, Madras Institute of Technology, Anna University, Chennai, 4-6 January 2008, 69-73. http://dx.doi.org/10.1109/icscn.2008.4447163

[3] Chaudhari, D.S. and Pandey, P.C (1998) Dichotic Presentation of Speech Signal with Critical Band Filtering for Improving Speech Perception. *Proc. IEEE Int. Conf. Acoust.*, *Speech and Signal Processing* (*ICASSP*'98), Seattle, Washington, AE 3.1.

[4] Daubechies, I. (1992) Ten Lectures on Wavelets. Vol. 61. Society for Industrial and Applied Mathematics, Philadelphia.

[5] Nogueira, W., Giese, A., Edler, B. and Buchner, A. (2006) Wavelet Packet Filterbank for Speech Processing Strategies in Cochlear Implants. *Proceedings of IEEE International Conference on Acoustic Speech*, *Signal Processing* (*ICASSP*'06), **5**, 14-19.

[6] Yao, J. and Zhang, Y. (2002) The Application of Bionic Wavelet Transform to Speech Signal Processing in Cochlear Implants using Neural Network Simulations. *IEEE Transactions on Biomedical Engineering*, **49**, 1299-1309. http://dx.doi.org/10.1109/TBME.2002.804590

[7] Karmarkal, A., Kumar, A. and Patney, R.K. (2007) Design of Optimal Wavelet Packet Trees Based on Auditory Perception Criterion. *IEEE Signal Processing Letters*, **14**, 240-243.

[8] Kolte, M.T. and Chaudhari, D.S. (2010) Evaluation of Speech Processing Schemes to Improve Perception of Sensorineural Hearing Impaired. *Current Science*, **98**, 613-615.

[9] Zwicker, E.W. (1961) Subdivision of Audible Frequency Rangeinto Critical Bands (Freqenzgruppen). *Journal of the Acoustical Society of America*, **33**, 248. http://dx.doi.org/10.1121/1.1908630

[10] Chopade, J.J. and Futane, N.P. (2015) Wavelet Based Scheme to Improve Performance of Hearing under Noisy Environment. *International Journal of Computer Applications*, **130**, 57-61.

[11] Baskent, D. (2006) Speech Recognition in Normal Hearing and Sensorineural Hearing Loss as a Function of the Number of Spectral Channels. *Journal of the Acoustical Society of America*, **120**, 2908-2925. http://dx.doi.org/10.1121/1.2354017

[12] Loizou, P.C., Mani, A. and Dorman, M.F. (2003) Dichotic Speech Recognition in Noise Using Reduced Spectral Cues. *Journal of the Acoustical Society of America*, **114**, 475-483. http://dx.doi.org/10.1121/1.1582861

[13] Soman, K.P. and Ramachandran, K.I. (2004) Insight into Wavelets from Theory to Practice. Prentice-Hall of India, New Delhi.

[14] Rao, R.M. and Bopardikar, A.S. (2001) Wavelet Transform Introduction to Theory and Applications. Addison Wesley Longmman Pte. Ltd., Delhi.

[15] Proakis, J.G. and Manolakis, D.G (1997) Digital Signal Processing Principles, Algoritms, and Applications. Prentice Hall, New Delhi.

[16] Miller, G.A. and Nicely, P.E. (1955) An Analysis of Perceptual Confusions among Some English Consonants. *Journal of the Acoustical Society of America*, **27**, 338-352. http://dx.doi.org/10.1121/1.1907526