Computer Systems Science & Engineering
DOI: 10.32604/csse.2023.025969
Article

Tech Science Press

Speech Separation Methodology for Hearing Aid

Joseph Sathiadhas Esra[1,*] and Y. Sukhi[2]

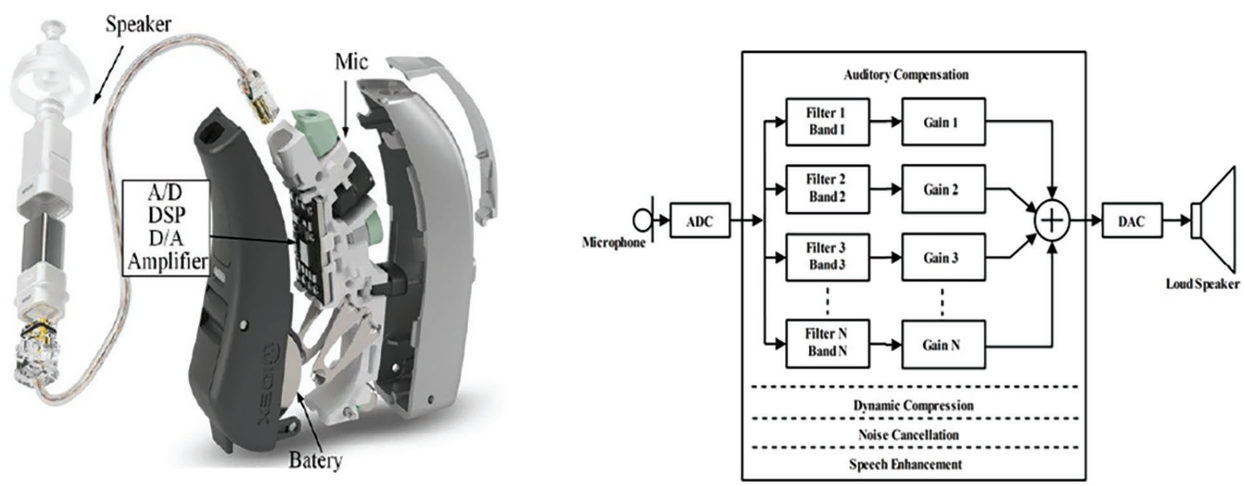
[1]Meenakshi Sundararajan Engineering College, Chennai, Tamil nadu, India
[2]R. M. K. Engineering College, Kavaraipettai, Tamil nadu, India
*Corresponding Author: Joseph Sathiadhas Esra. Email: josephesra.phd@gmail.com
Received: 21 January 2022; Accepted: 03 March 2022

Abstract: In the design of hearing aids (HA), the real-time speech-enhancement is done. The digital hearing aids should provide high signal-to-noise ratio, gain improvement and should eliminate feedback. In generic hearing aids the performance towards different frequencies varies and non uniform. Existing noise cancellation and speech separation methods drops the voice magnitude under the noise environment. The performance of the HA for frequency response is non uniform. Existing noise suppression methods reduce the required signal strength also. So, the performance of uniform sub band analysis is poor when hearing aid is concern. In this paper, a speech separation method using Non-negative Matrix Factorization (NMF) algorithm is proposed for wavelet decomposition. The Proposed non-uniform filter-bank was validated by parameters like band power, Signal-to-noise ratio (SNR), Mean Square Error (MSE), Signal to Noise and Distortion Ratio (SINAD), Spurious-free dynamic range (SFDR), error and time. The speech recordings before and after separation was evaluated for quality using objective speech quality measures International Telecommunication Union -Telecommunication standard ITU-T P.862.

Keywords: Speech separation; wavelet filter; independent component analysis (ICA); non-negative matrix factorization (NMF); fejer-korovkin (FK); signal-to-noise ratio (SNR)

1 Introduction

About 500 million people were suffering from hearing loss. Their quality of life can be enhanced by using Digital Hearing Aid. The hearing aid is used by only 30% of the patients. This percentage can be increased by designing hearing aid device with low noise and improved sound quality. Obviously, the cost is the important factor too. The different hearing aids based on placement Behind-The-Ear (BTE) HA, Receiver-In-Canal (RIC) HA, In-The-Canal (ITC) HA and the In-The-Ear (ITE) HA, have same structures for collection and sound regeneration. The Main functions of hearing aid are shown in Fig. 1 [1]. In the frequency band, the frequency out of audio range was removed using the low pass and high pass filter [2]. This type of noise cancellation will enhance the digital hearing aids that could have a series of advantages such as; high signal-to-noise ratio, higher gain, immune to electromagnetic interference and feedback elimination [3–5]. The components of the digital hearing aids are microphone, an Analog-to-Digital converter (A/D), an amplifier, a Digital Signal Processor (DSP), a Digital-to-Analog converter (D/A), a speaker, and a battery. The existing HA structure adopts a built-in DSP for voice processing. In noise removal process sub-band are selected by using reconfigurable filter bank [6,7]. The Basic hearing aid is given in Fig. 1a Even though the hearing aid manufacturers miniaturized the device, it lags in battery capacity. The improved listening capacity have increased the cost which not affordable by all. Speech separation becomes an important element in binaural HA system. Building an analog filter with multiple stages is difficult when compared to the digital filter [8]. Liu et al. [9] designed a de-noising filter to remove the mixed impulsive and Gaussian noises. The noise suppression circuit is a cascaded filter with similar co-efficients. Chandra Sekhar Yadav et al. [10] tested a Wiener filter to remove the Gaussian noise from the input signal. Abbasa et al. [11] presented the separation of speech mixtures that are often be referred to as the cocktail party problem. The Independent component analysis (ICA) and binary time frequency masking is the two source separation techniques which has been combined to solve the speech separation process [12]. There are several works in literature dedicated for speech separation process which includes fusion framework [13] and NMF-based Target Source Separation [14]. Sean Wood et al. [15] presented a hybrid blind source separation algorithm which was a combination of non-negative matrix factorization (NMF) and generalized cross correlation (GCC) method.



(a) Basic digital hearing aid **(b)** Block diagram of Digital Hearing aid

Figure 1: (a) Basic digital hearing aid, (b) Block diagram of Digital Hearing aid

The Hearing aid should recognize the speech signals out of the environmental noise. Some of the noise is produced by speech babbles, instrumentation noise and other unnecessary sounds. Too much of reverberation will reduce speech intelligibility and the overall sound quality. If the noise magnitude is more than the voice, the efficiency of speech processing unit will be poor. So, the alternative method can be developed for noise reduction in hearing aid. The Noise is removed independently in binaural. Hearing aid (HA), which differs from the two systems since it processes noise independently. In spite of the advances in Hearing aid (HA) technology, improving the speech intelligibility is a challenge. In recent years, hearing aids are connected to android, iPhone Operating System (IOS), and Bluetooth-enabled phones. In most cases the collected sound is directional. If the noise strength is more compared to the required signal, it's very difficult to remove the same. In addition, the frequency response of hearing aids should be uniform throughout the audio spectrum. The Performance of the HA for frequency response is non uniform. Existing noise suppression methods reduce the required signal strength also along with noise. So the performance of uniform sub band analysis is poor when hearing aid is concern. On the other hand, source separation is one of the important issues in the Digital Hearing Aid (DHA). The Microphone in the Digital Hearing Aid (DHA) continuously receives the 'N' number of incoming speech signals. The Hearing-impaired people couldn't understand the collapsed incoming speech signals.

In this research work, sub band analysis using Fejer-Korovkin (FK) wavelet based decomposition Methodology is proposed. The Wavelet based de-noising algorithm minimizes the Gaussian noises present in the input signal. If the number of stages in the filter bank is more, the rejection ratio will be more. It is better in performance when compared to the existing two stage wavelet filter bank and tree structured wavelet filter banks. For speech separation Non-negative Matrix Factorization (NMF) algorithm is used. The Proposed filter bank architecture is compared with other architectures. The Proposed methods Error and time calculation for different divergence are measured. For evaluation of the proposed method and existing method, mixed audio sources were used for speech separation. Different parameters like Signal to Noise Ratio (SNR), Signal to Distortion Ratio (SDR), Signal to Noise and Distortion Ratio (SINAD), Spurious-free dynamic range (SFDR), Mean Square Error (MSE), and Band Power were used for evaluation.

2 Foundation of Filter Bank

The Filter bank is an important functional block in digital hearing aids that decomposes the input signal into different bands [16,17]. The Gain of individual stages can be varied based on the band requirements.

2.1 Wavelet Based Non-Uniform Filter Bank

The Discrete input signal is applied to filter bank to produce a set of sub-band signals [18]. The term uniform filter bank (UFB) is defined to emphasize that all the sub-band signals are at the same rate. Shakya et al. [19] The Discrete Wavelet Transform (DWT) for multi-resolution analysis can be viewed as non-uniform filter bank. In terms of this methodology a low-pass filter corresponds to scaling function and the subsequent high-pass corresponds to wavelet function [20]. The corresponding non-uniform filter bank is possible through repetitive application on the low-pass channel [21].

2.2 Two Channel Wavelet Filter Bank

For speech processing two band ortho normal wavelet is used which can be associated with ortho normal filter bank [22]. Speech processing of hearing aid can be possible using filter bank which has equal bandwidth and perfect reconstruction features. The connection between discrete and continuous wavelet bases enhances the design of the filter banks like equal bandwidth, two-channel, Perfect Reconstruction Quadrature Mirror Filter (PR-QMF) banks [23,24]. Eqs. (1) and (2) represents the ortho-normality state of wavelet and scaling bases with its relations to the discrete-time filter banks.

(1) In intra- and inter-scales, the wavelets are ortho-normal as,

$$\int \beta_{mn}(t)\beta_{m'n'}(t)dt = \theta_{m-m}\theta_{n-n'} \tag{1}$$

(2) The corresponding scaling function of wavelet theory has only intra-scale ortho-normality as,

$$\begin{aligned} &\int \delta_{mn}(t)\delta_{mn'}dt = \theta_{n-n'} \\ &where, \\ &\delta_{mn}(t) = 2^{-m/2}\delta(2^{-m}t - n) \end{aligned} \tag{2}$$

(3) The corresponding property for all values of m, n, m′, and n′ in wavelet and scaling bases is given in Eq. (3).

$$\int \beta_{mn}(t)\delta_{m'n'}(t)dt = 0 \tag{3}$$

In Eq. (4), ho(n), hl(n) represent the two-band discrete-time Perfect Reconstruction Quadrature Mirror Filter (PR-QMF) bank with the additional property of H,(eh) = 0 at o = x.

$$\begin{aligned}\delta(t) &= \sum_n h_0(n)\delta(2t-n) \leftrightarrow \delta(\Omega) = \prod_{k=1}^{\infty} H_0(e^{j\omega 2^k})\\ \beta(t) &= \sum_n h_1(n)\delta(2t-n) \leftrightarrow \beta(\Omega) = H_1(e^{j\omega/2})\prod_{k=2}^{\infty} H_0(e^{j\omega 2^k})\end{aligned} \tag{4}$$

The wavelet δ(t) and scaling functions β(t) are build to maintain ortho-normality. The filter function $h_0(n)$ and $h_l(n)$ have finite duration [25]. The design of a two-band discrete-time para-unitary filter bank are advantages and is been used for channel diagnolization [26].

2.3 Tree Structured Wavelet Filter Bank

In tree-structured filter banks, the inputs pass through two or more filters and the output is downsampled [27]. The Eq. (5) defines the analytic wavelets.

$$\psi_c(t) = \psi_r(t) + j\psi_i(t) \tag{5}$$

Here j represents the unit imaginary.

3 Foundation of NMF

Among the wide variety of sound separation algorithms, the unsupervised Non-negative Matrix Factorization (NMF) dictionary learning algorithm suits well for the delineation of sound mixture [28,29]. The Cost function of Non-negative Matrix Factorization (NMF) will decompose their spectrogram into two non-negative matrices such as; a dictionary matrix W_fd and a coefficient matrix Hdt. Hence the product of both the Non-negative Matrix Factorization (NMF) function is ∧ = WH approximates V. various measures of reconstruction error have been used, several of which generalize to the divergence D_β (V|∧), defined by Eq. (6).

$$D_\beta(\vee|\wedge) = \begin{cases} \dfrac{\vee}{\wedge} - \log\dfrac{\vee}{\wedge} - 1 & if\ \ \beta = 1\\ \vee(\log\vee - \log\wedge) + (\wedge - \vee)) & if\ \ \beta = 0\\ \dfrac{1}{\beta(\beta-1)}(\vee^\beta - \wedge^\beta - \beta\wedge^{\beta-1}(\vee - \wedge))\ otherwise \end{cases} \tag{6}$$

The Speech regeneration may be the right choice of NMF algorithm which includes the Euclidian distance (β = 2) [30] and the generalized Kullback-Leibler divergence (β = 1) [31]. The Coefficient sparsity is used by [32] for Itakura-Saito divergence (β = 0). Hence the normalization will have such kind of l1 norms that is typically used for coefficient sparsity which is shown by [33]. At the point of initialization technique, the multiplicative update rules are then defined as W and H. In Eq. (7) and Eq. (8) the update rules for D_β (V|∧) are given.

$$H \leftarrow H \times \frac{W^T(V \times \wedge^{\beta-2})}{\mathrm{W^T}\wedge^{\beta-1} + \alpha} \tag{7}$$

$$WW \times \frac{(\wedge^{\beta-2} \times V)H^T}{\wedge^{\beta-1}\mathrm{H}^\mathrm{T}} \tag{8}$$

where X is the Hadamard product. Fig. 2 represents the left and right input in the case of stereo audio signals that has been concatenated in time [34] where Fig. 2a shows the single dictionary of spectral atoms which is used to encode both channels via the two coefficient matrices $H_l$dt and $H_r$dt shown in Fig. 2b. Speech emotion recognition was presented in literature [35,36].

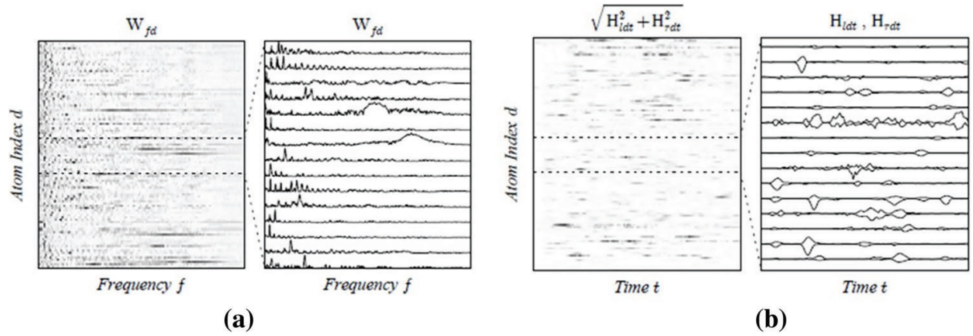


Figure 2: The coefficient of stereo channel with the hearing aid can be represented such that the negative decomposition of left channels is taken on the combination of Hldt and Hrdt. It shows the non-negative matrix factorization (NMF) decomposition of a stereo mixture of speech signals. a) The Non-negative matrix factorization (NMF) dictionary Wfd, with cube root compression applied for clarity, consisting of atoms that are nonnegative functions of frequency. A subset is shown in detail on the right. b) non-negative atom coefficients for the left and right channels

4 Proposed Speech Separation with Filter Bank and NMF

The Proposed speech separation method using filter bank and NMF is shown in Fig. 3.

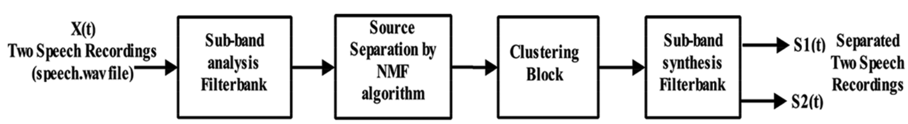


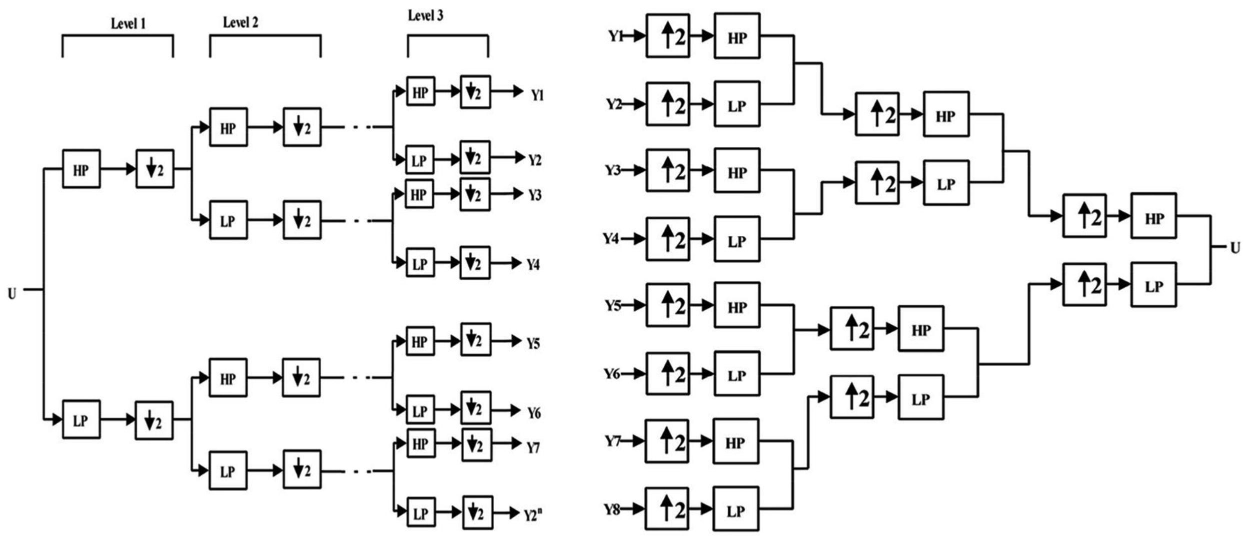
Figure 3: Block diagram of proposed speech separation algorithm

4.1 Sub Band Analysis Filter Bank Using Proposed Fejer-Korovkin Wavelet Filter Bank

The Fejer-Korovkin (FK) wavelet is used to design the proposed filter bank architecture which is shown in Fig. 4. The speech signals are decomposed using the proposed filter bank into non-uniform sub-bands. The Architecture of abalysis and synthesis filter is shown in Figs. 4a and 4b.

In Eq. (9) shows the Fejer-Korovkin (FK) wavelet filter that defines kernel K_n which is used to separate the raw time series into a high frequency (HF) component and a low frequency (LF) component.

$$K_n(\xi) = \begin{cases} \dfrac{2\sin^2(\Pi/(n+2))}{n+2}\left[\dfrac{\cos((n+2)x/2}{\cos(\Pi/(n+2)) - \cos(\xi)}\right]^2, & x \notin \pm\dfrac{\Pi}{n+2} + 2Z\Pi. \\ \dfrac{n+2}{2} & x \in \pm\dfrac{\Pi}{n+2} + 2Z\Pi. \end{cases} \tag{9}$$



(a) Proposed Analysis Filter bank (b) Proposed Synthesis Filter bank

Figure 4: (a) Proposed analysis filter bank, (b) Proposed synthesis filter bank

The $K_n(\xi)$ can be simplified and expressed in Eq. (10) and $\theta_n(k)$ can be expressed in Eq. (11) and Eq. (12).

$$K_n(\xi) = 1 + 2\sum_{k=1}^{n} \theta_n(k)\cos kx \tag{10}$$

With

$$\theta_n(k) = \frac{1}{2(n+2)\sin(\Pi/(n+2))}\left[(n-k+3)\sin\frac{k+1}{n+2}\Pi - (n=k+1)\sin\frac{k-1}{n+2}\Pi\right] \tag{11}$$

$$1 - \theta_n(1) = 1 - \cos(\Pi/(n+2)) = \mathrm{O}(n^{-2}) \tag{12}$$

We define the Fejer-Korovkin filters by Eq. (13).

$$|m_0^n(\xi)|^2 = \frac{1}{2\Pi}\int_{-\Pi/2}^{\Pi/2} K_n(\xi - u)du \tag{13}$$

Then m_0^n has degree n + 1 if n is odd and degree n if n is even. $\gamma_\rho(m_0^n) = O(1/n)$.

The filter bank plot of |m0n|2 for n = 2, 4,….12 with following Fig. 5. Hence the Fejer-Korovkin (FK) kernals $K_{2n} \epsilon K_{p,c}^{2n}$ for n = 1, 2..6 which decreasing on [0,Π] that is noted n all the filters.

The Figs. 6 and 7 show the scaling function and wavelet generated by the Fejer-Korovkin (FK) filter of length 12.

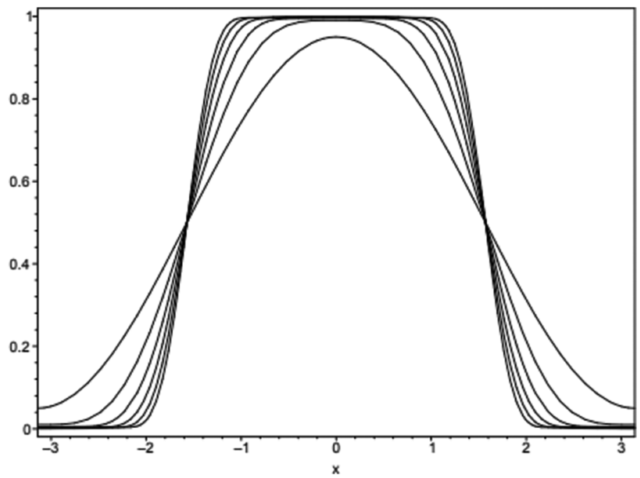


Figure 5: The fejer-korovkin (FK) filters for n = 2, 4,…, 12

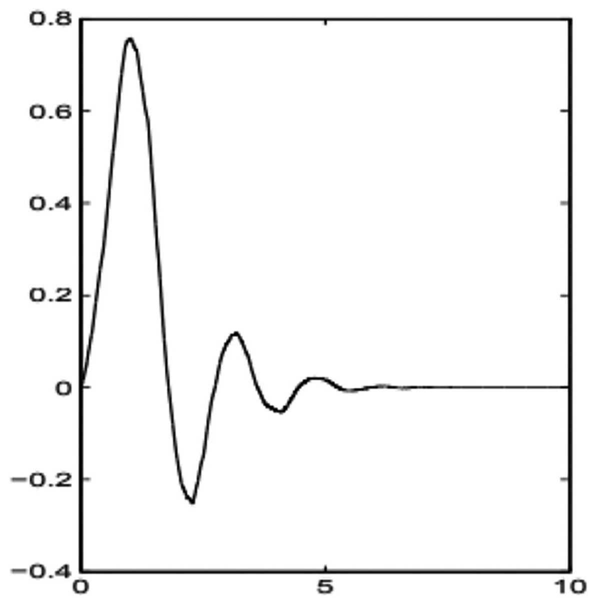


Figure 6: The scaling function generated by the fejer-korovkin filer

4.2 Source Separation by Non-Negative Matrix Factorization (NMF) Algorithm

The Non-negative Matrix Factorization (NMF) is used for blind source separation which is closely approximated by a constant frequency with the magnitude spectrogram X of the mixture. The corresponding audio source is separated into I channels with their corresponding spectrograms C_i, $1 \leq i \leq I$. This Algorithm is based on vector Bi and a time varying gain G_i of the single speech. In Eq. (14) the spectrogram C_i is shown.

$$C_i = B_i G_i \tag{14}$$

The C_i is rank one and the low pass characteristics is found in rows Gi. The Separation of Non-negative Matrix Factorization (NMF) algorithm is improved due to the continuous nature and is shown in Eq. (15).

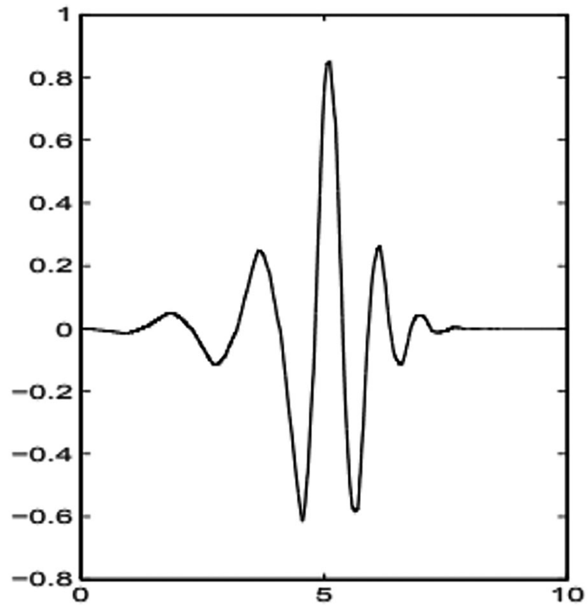


Figure 7: The wavelet function generated by the fejer-korovkin (FK) filer

$$c_t = a \sum_i \frac{\sum_{t=2}^{T} (G(i,\ t) - G(i,\ t-1))^2}{\sum_{t=1}^{T} G^2(i,\ t)} \tag{15}$$

In Eqs. (16)–(20) shows the multiplicative update rules

$$\nabla c_r^+ = B^T 1 \tag{16}$$

$$\nabla c_r^- = B^T \times \frac{X}{BG} \tag{17}$$

$$\nabla c_t^+(i,\ t) = 4a \frac{G(i,\ t)}{\sum_{n=1}^{T} G^2(i,\ n)} \tag{18}$$

$$\nabla c_t^-(i,\ t) = 2a \frac{G(i,\ t-1) + G(i,\ t+1)}{\sum_{n=1}^{T} G^2(i,\ n)} + 2a \frac{\sum_{n=2}^{T} (G(i,\ n) - G(i,\ n-1))^2}{\left(\sum_{n=1}^{T} G^2(i,\ n)\right)^2} \tag{19}$$

$$\mathrm{G} \leftarrow \mathrm{G} \times \frac{\nabla c_r^- + \nabla c_t^-}{\nabla c_r^+ + \nabla c_t^+} \tag{20}$$

The Eqs. (21)–(23) shows the following numerical stability for normalized B_i and G_i in each iteration to ensure equal energy.

$$A_i = \sqrt{\frac{\|G_i\|_2}{\|B_i\|_2}} \tag{21}$$

$$G_i \leftarrow \frac{G_i}{A_i} \tag{22}$$

$$B_i \leftarrow B_i A_i \tag{23}$$

Fig. 8 shows the Non-negative Matrix Factorization (NMF) based source separation method where the separation can be improved by using clustering.

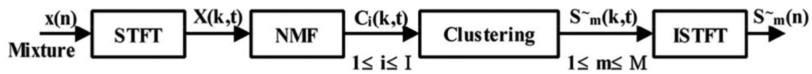


Figure 8: The signal flow of the proposed non-negative matrix factorization (NMF) separation algorithm

4.3 Clustering Algorithm

The Proposed clustering method can be used to cluster any number of sub clusters. For more independent sources the clustering can be done using hierarchical clustering. Here two clusters are created for N number of channels. The clusters m, m˜ ∈ {1, 2} are separated using the vectors a˜, a˜(i) ∈ {1, 2}. The estimated energy E˜m˜ of the spectrograms of both clusters are given by

$$\tilde{E}_{\tilde{m}} = \sum_{i} \sum_{k,t} C_i^2(k,\ t)\delta_{\tilde{m}\tilde{a}(i)} \tag{24}$$

For uncorrelated sources, the energy is estimated from the mixture signals. Further we assume that one cluster corresponds to one source, and the other cluster contains the remaining sources. Therefore, we expect that the first separated source esm1 corresponds to the cluster with lowest energy because the other cluster corresponds to multiple sources:

$$m_1 = \arg\min_{\tilde{m}} \tilde{E}_{\tilde{m}} \tag{25}$$

The Process repeats until all channels are clustered into two. The process terminates once the sources are clustered.

5 Experimental Results

5.1 Analysis of Two Channel Wavelet Filter Bank

Tab. 1 shows the subband analysis of the two channel filter-bank implemented using different wavelets such as db2, haar, coif1, sym2 and dmey. Various parameters such as Signal-to-noise ratio (SNR), SDR, Mean Square Error (MSE) and band power are observed.

Table 1: Sub band analysis of two channel wavelet filter bank with different wavelet

Wavelet used	Separated speech	Frequency (Hz)	SNR	SDR	MSE	Band-power
db2	B1	2500	−6.8218	−6.8243	0.1187	0.0845
	B2	6000	−13.105	−13.106	0.0351	0.0008
haar	B1	2500	−6.7904	−6.7929	0.0343	0.0832
	B2	6000	−6.7390	−6.7398	0.0343	0.0021
coif1	B1	2500	−6.8221	−6.8251	0.0343	0.0845
	B2	6000	−15.946	−15.949	0.0343	0.0008
sym2	B1	2500	−6.8218	−6.8243	0.1187	0.0845

(Continued)

Table 1 (continued)

Wavelet used	Separated speech	Frequency (Hz)	SNR	SDR	MSE	Band-power
	B2	6000	−13.105	−13.106	0.0351	0.0008
dmey	B1	2500	−6.8261	−6.8287	0.0343	0.0845
	B2	6000	−20.090	−20.094	0.0343	0.0006

5.2 Analysis of Tree Structured Wavelet Filter Bank

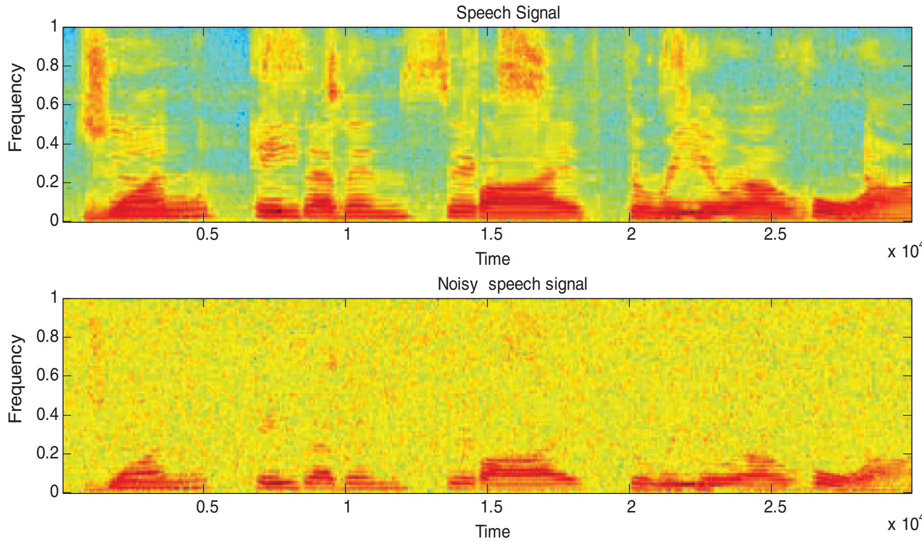
Tab. 2 shows the subband analysis of the Tree structured wavelet filter bank implemented using different wavelets. Various parameters are observed.

Table 2: Subband analysis of tree structured wavelet filter bank with different wavelet

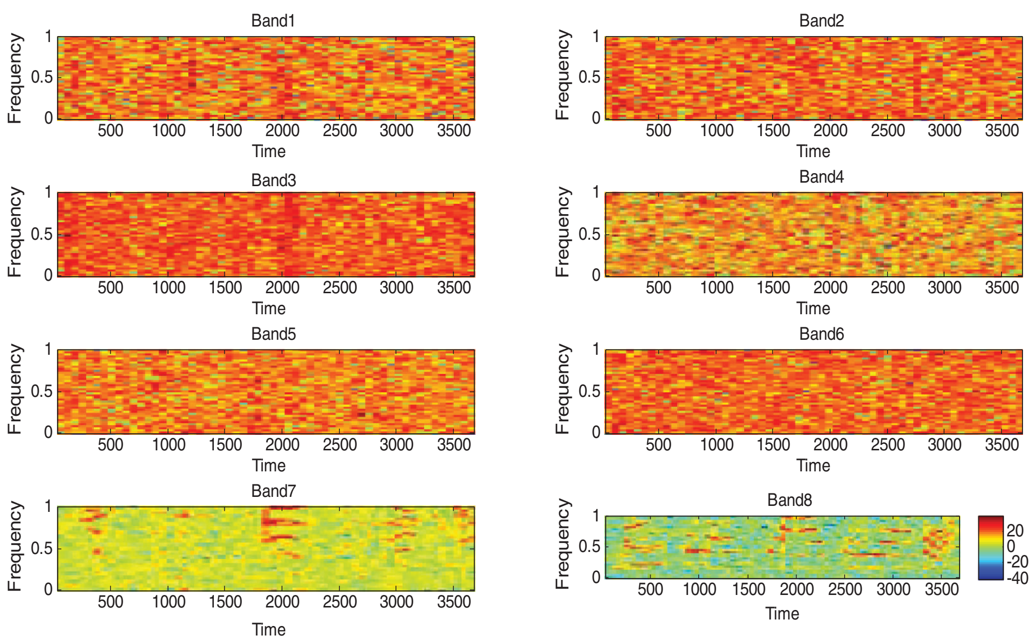
Wavelet used	Separated speech bands	Frequency (Hz)	Band power	SNR	MSE
db2	B1	16000	2.2531e–04	–23.4695	0.0104
	B2	8000	0.0070	−18.3915	0.0220
	B3	4500	0.0422	−16.8541	0.0504
	B4	1200	0.0084	−6.5370	0.0098
	B5	250	0.0198	−6.4164	0.0206
haar	B1	16000	5.3615e–04	−22.2270	0.0107
	B2	8000	0.0083	−18.4683	0.0233
	B3	4500	0.0358	−14.4176	0.0445
	B4	1200	0.0121	−11.4527	0.0133
	B5	250	0.0178	−4.7190	0.0196

5.3 Analysis of Proposed Fejer-Korovkin (FK) Wavelet Filter Bank

The Analysis of the proposed filter bank using Fk wavelet is done by evaluating the spectrum of noisy signal and input Fig. 9a. The Different decomposition stages are presented in Fig. 9b. Tabs. 3–7 presents the various analysis of the subband filter. The Comparative analysis is presented from Figs. 9(c–f).

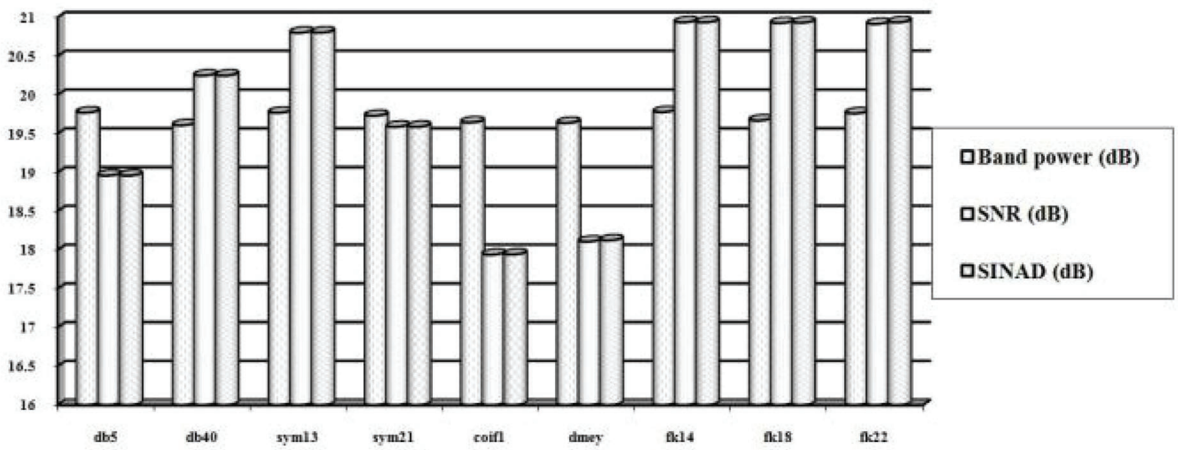


(a) Spectrum of input and noisy speech signal

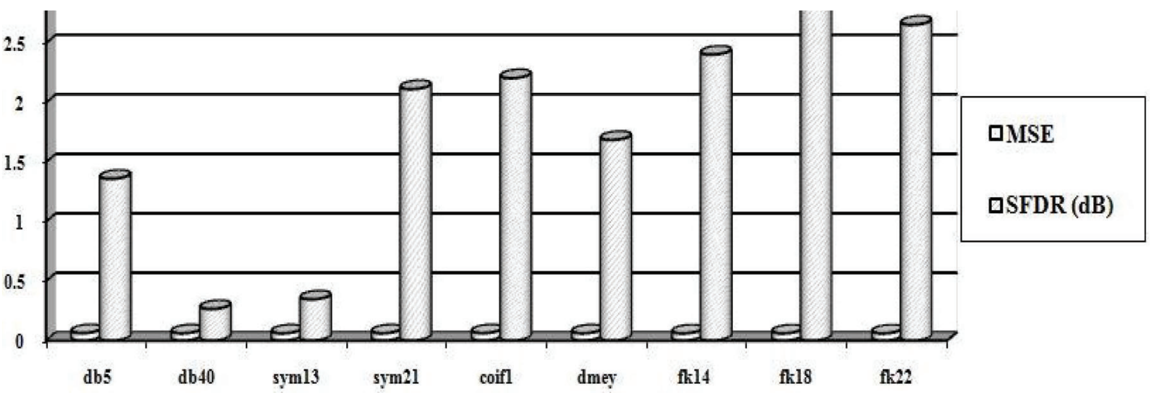


(b) Decomposition of Noised speech signal by using Fejerkorovkin (FK) wavelet ‘fk14’

Figure 9: (Continued)

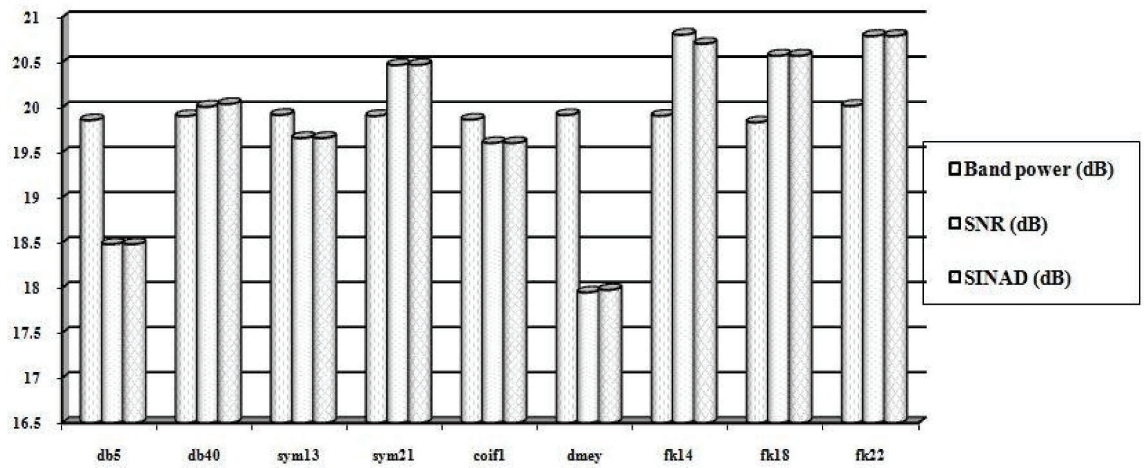


(c) Band power, Signal-to-noise ratio (SNR), Signal to Noise and Distortion Ratio (SINAD) comparison of Subband1 in different wavelet 'fk14'

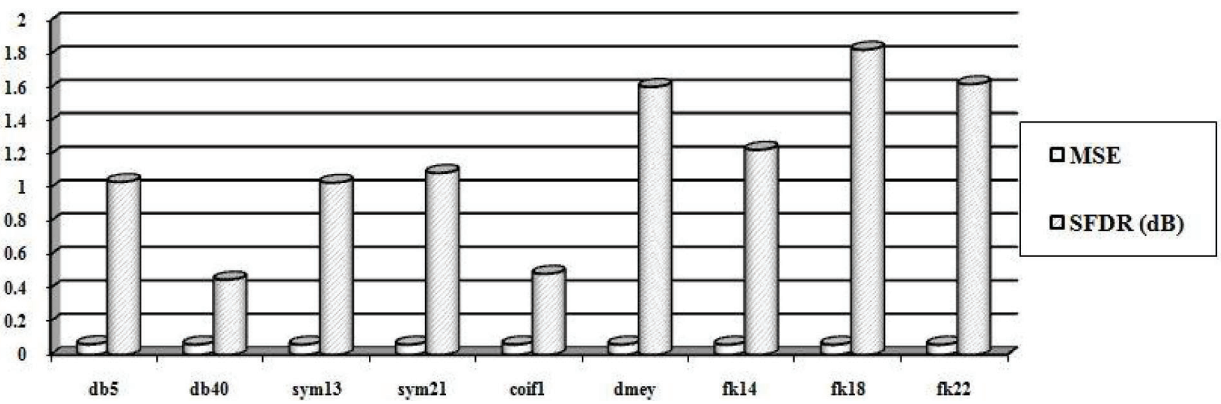


(d) Mean Square Error (MSE) and Spurious-free dynamic range (SFDR) comparison of Subband1 in different wavelet 'fk14'

Figure 9: (Continued)



(e) Band power, Signal-to-noise ratio (SNR), Signal to Noise and Distortion Ratio (SINAD) comparison of Subband2 in different wavelet 'fk14'



(f) Mean Square Error (MSE) and Spurious-free dynamic range (SFDR) comparison of Subband2 in different wavelet fk14'

Figure 9: (a) Spectrum of input and noisy speech signal, (b): Decomposition of noised speech signal by using fejer-korovkin(FK) wavelet 'fk14', (c) Band power, signal-to-noise ratio (SNR), signal to noise and distortion ratio (SINAD) comparison of subband1 in different wavelet 'fk14', (d) Mean square error (MSE) and spurious-free dynamic range (SFDR) comparison of subband1 in different wavelet 'fk14', (e) Band power, signal-to-noise ratio (SNR), signal to noise and distortion ratio (SINAD) comparison of subband2 in different wavelet 'fk14', (f) Mean square error (MSE) and spurious-free dynamic range (SFDR) comparison of subband2 in different wavelet fk14'

Table 3: Band power analysis of 8 sub-bands

Wavelet type	Band power (dB)							
	B1	B2	B3	B4	B5	B6	B7	B8
db5	19.8	19.9	19.5	19.7	19.3	19.8	13.0	5.1
db40	19.6	19.9	19.6	19.9	19.6	19.9	14.3	5.0
sym13	19.8	19.9	19.6	19.8	19.7	19.8	13.6	5.0
sym21	19.7	19.9	19.6	19.8	19.6	19.8	14.0	5.0
coif1	19.6	19.9	19.5	19.6	18.6	18.6	11.9	5.4
dmey	19.6	19.9	19.6	19.9	19.6	19.8	14.3	5.0
fk14	19.8	19.9	19.7	19.9	19.6	19.7	13.9	5.0
fk18	19.7	19.8	19.6	19.7	19.5	19.7	14.3	4.9
fk22	19.8	20.0	19.6	19.8	19.7	19.7	14.3	5.0

Table 4: Signal-to-noise ratio (SNR) analysis of 8 sub-bands

Wavelet type	SNR (dB)							
	B1	B2	B3	B4	B5	B6	B7	B8
db5	9.0	18.5	18.5	19.0	13.8	18.2	4.2	6.0
db40	20.2	20.0	20.2	17.1	17.8	18.3	4.2	5.7
sym13	20.8	19.7	19.5	17.9	16.8	18.6	4.2	5.6
sym21	19.6	20.5	18.7	19.6	18.1	17.4	4.0	5.7
coif1	17.9	19.6	18.4	18.9	9.0	8.7	4.9	6.6
dmey	18.1	17.9	20.1	19.6	17.0	18.0	4.3	5.7
fk14	20.9	20.8	20.4	21.0	19.2	20.2	5.6	6.8
fk18	20.9	20.6	20.6	20.6	19.7	19.1	5.3	6.7
fk22	20.9	20.8	20.8	20.3	19.1	20.9	5.0	6.7

Table 5: Mean square error (MSE) analysis of 8 sub-bands

Wavelet type	MSE							
	B1	B2	B3	B4	B5	B6	B7	B8
db5	0.06	0.06	0.06	0.06	0.06	0.06	0.10	0.36
db40	0.06	0.06	0.06	0.06	0.06	0.06	0.09	0.37
sym13	0.06	0.06	0.06	0.06	0.06	0.06	0.09	0.36
sym21	0.06	0.06	0.06	0.06	0.06	0.06	0.09	0.37

(Continued)

Table 5 (continued)

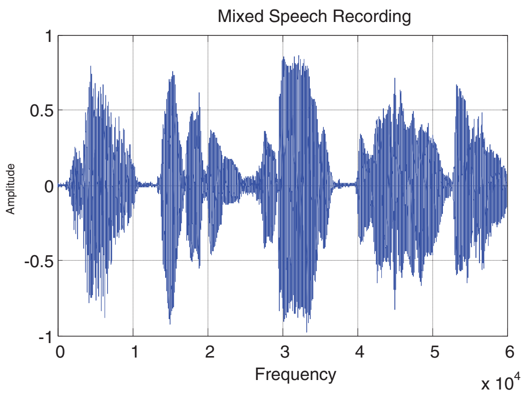
Wavelet type	MSE							
	B1	B2	B3	B4	B5	B6	B7	B8
coif1	0.06	0.06	0.06	0.06	0.06	0.07	0.12	0.38
dmey	0.06	0.06	0.06	0.06	0.06	0.06	0.09	0.37
fk14	0.06	0.06	0.06	0.06	0.06	0.06	0.08	0.34
fk18	0.06	0.06	0.06	0.06	0.06	0.06	0.09	0.34
fk22	0.06	0.06	0.06	0.06	0.06	0.06	0.09	0.34

Table 6: Signal to noise and distortion ratio (SINAD) analysis of 8 sub-bands

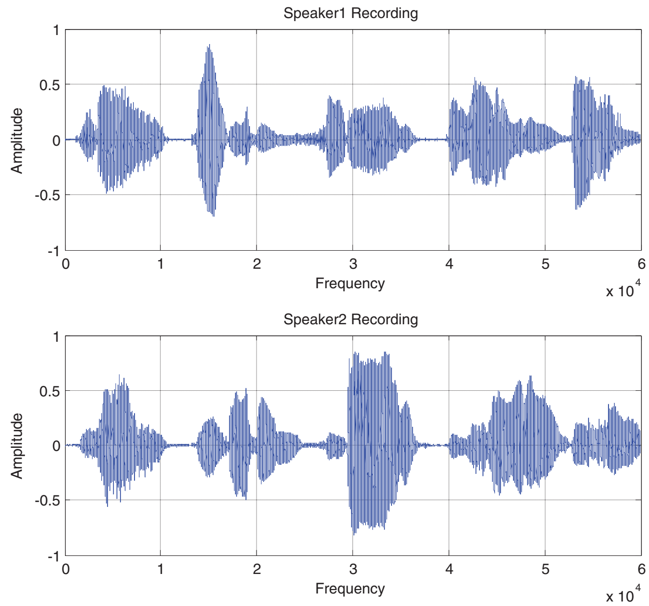
Wavelet type	SINAD (dB)							
	B1	B2	B3	B4	B5	B6	B7	B8
db5	19.0	18.5	18.5	19.0	13.8	18.2	4.2	6.0
db40	20.2	20.0	20.2	17.1	17.8	18.3	4.2	5.7
sym13	20.8	19.7	19.5	17.9	16.8	18.6	4.2	5.6
sym21	19.6	20.5	18.7	19.6	19.2	17.4	4.0	5.7
coif1	17.9	19.6	18.4	18.9	19.0	18.7	4.9	6.6
dmey	18.1	18.0	20.1	19.6	17.0	18.0	4.3	5.7
fk14	20.9	20.7	20.4	20.2	19.2	20.2	5.6	6.8
fk18	20.9	20.6	20.6	20.6	19.7	20.1	5.3	6.7
fk22	20.9	20.8	20.8	20.3	19.2	20.9	5.0	6.7

Table 7: Spurious-free dynamic range (SFDR) analysis of 8 sub-bands

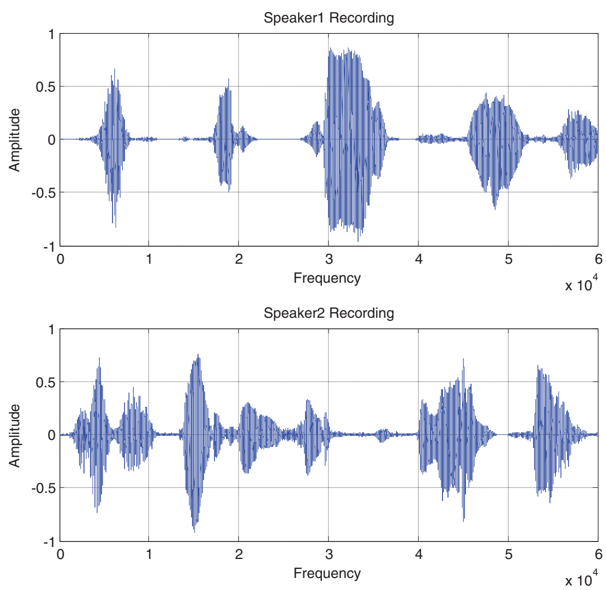
Wavelet type	SFDR							
	B1	B2	B3	B4	B5	B6	B7	B8
db5	1.36	1.03	2.36	1.56	3.94	0.60	7.14	5.36
db40	0.27	0.45	0.57	2.42	2.56	0.64	6.45	1.97
sym13	0.35	1.03	2.00	1.46	4.64	0.58	5.99	5.71
sym21	2.11	1.09	0.69	0.41	1.81	3.12	5.94	5.43
coif1	2.21	0.48	1.95	0.09	4.01	4.10	2.49	5.08
dmey	1.69	1.60	0.79	1.05	0.90	2.76	6.12	5.44
fk14	2.40	1.23	2.38	2.50	2.60	3.81	7.81	6.09
fk18	2.94	1.83	2.38	2.46	2.90	3.15	7.34	5.81
fk22	2.65	1.62	2.42	2.77	2.80	3.12	7.25	5.87



(a) Mixed Audio Source (Mixednumbers.wav file)



(b) Separation of Two speech recordings in mixed audio source by Existing *Independent Component Analysis (ICA)*.



(c) Separation of Two speech recordings in mixed audio source by proposed method.

Figure 10: (a) Mixed audio source (mixednumbers.wav file), (b) Separation of two speech recordings in mixed audio source by existing *independent component analysis (ICA)*, (c) Separation of Two speech recordings in mixed audio source by proposed method

5.4 Independent Component Analysis (ICA)

The Sources in the hearing aid system are statistically independent and the linear mixtures are separated using Independent component analysis (ICA) and Blind Source Separation (BSS). The Experimental setup uses two speech signal and noise. It is a microphone which records with proximities a1 and b1 for male and female voices respectively. Hence they can be source separated using ICA. The linear mixture observed data is given as x (Eq. (26)).

$$x = As \tag{26}$$

where, A is some unknown invertible that mixes the components as A = [a1 b1]. Eq. (27) shows the estimation of underlying source which will construct a new matrix W as linear transformed data,

$$\hat{s} = Wx \tag{27}$$

The Unmixing matrix has the approximate value of A^{-1} so that $\hat{s} \approx s$ for finding the Independent component analysis (ICA). The Mixing source is shown in Fig. 10a and the separated audios are shown in Fig. 10b.

5.5 Analysis of NMF Algorithm for BSS

The Input signals are analyzed using the Non-negative Matrix Factorization (NMF) algorithm and the source matrix are generated W and H. Error and time calculation of Proposed NMF source separation algorithm with different divergence are presented in Tab. 8. Perceptual Evaluation of Speech Quality (PESQ) parameters are analyzed in this work. The performance comparison of the existing and proposed method is shown in Tab. 9.

Table 8: Error and time calculation of proposed non-negative matrix factorization (NMF) source separation algorithm with different divergence

S.No	Divergence	Error	Time (S)
1	nmf_kl	8060.38	0.53
2	nmf_kl_ns	7864.43	0.49
3	nmf_kl_loc	39863.18	0.52
4	nmf_kl_con	7864.55	0.50
5	nmf_euc_orth	108743.00	0.66
6	nmf_euc	104976.70	1.41
7	nmf_convex	119376.86	0.44
8	nmf_beta (β = 0)	9870.46	1.59
9	nmf_amari	5856.42	0.79
10	nmf_euc_sparse_es	270814.32	0.45

Table 9: Performance comparison of existing and proposed wavelets

	Band power (dB)	Signal-to-noise ratio (SNR) (dB)	Mean square error (MSE)	Signal to noise and distortion ratio (SINAD) (dB)	Spurious-free dynamic range (SFDR)
db5	5.1379	6.0394	0.3599	6.0394	5.3621
db40	4.9724	5.6697	0.3696	5.6697	1.9688
sym13	5.0347	5.5948	0.3642	5.5948	5.7113
sym21	4.9620	5.6577	0.3727	5.6577	5.4252
coif1	5.4415	6.5812	0.3770	6.5812	5.0758
dmey	4.9639	5.7350	0.3706	5.7350	5.4405
Proposed fk14	4.9732	6.7527	0.3393	6.7527	6.0928
Proposed fk18	4.9308	6.7338	0.3409	6.7338	5.8103
Proposed fk22	4.9805	6.6878	0.3373	6.6878	5.8748

The Speech recordings before and after separation are evaluated for quality using objective speech quality measures such as ITU-T P.862 for objectivity (Tab. 10).

Table 10: PESQ measure based on the ITU standard P.862

Existing wavelet + non-negative matrix factorization (NMF) separation						Proposed wavelet + non-negative matrix factorization (NMF) separation			
S.No	db5	db40	sym13	sym21	coif1	dmey	fk14	fk18	fk22
Source 1	1.5912	1.3339	1.3922	1.4719	0.8961	1.2298	2.8356	2.0573	2.2219
Source 2	2.2007	0.9160	1.6902	1.0330	2.0850	1.7251	1.8767	2.6213	2.7645

6 Conclusion

The paper presents a new method for speech separation in Hearing aids which provide high signal-to-noise ratio. The Wavelet based decomposition Methodology using Fejer-Korovkin (FK) algorithm is better in performance when compared to the existing decomposition in two stage wavelet filter bank and tree structured wavelet filter bank db. For speech separation Non-negative Matrix Factorization (NMF) algorithm is used. The proposed methods Error and time calculation for different divergence are measured. For evaluation of the proposed method and existing method on speech separation mixed audio sources are used and ITU standard P.862 is utilized for the evaluation. Different parameters like Signal to Noise Ratio (SNR), Signal to Distortion Ratio (SDR), Signal to Noise and Distortion Ratio (SINAD), Spurious-free dynamic range (SFDR), Mean Square Error (MSE), and Band Power were used for evaluation. In future deep learning methods will be proposed for this application. The hardware implementation will be carried out using new semiconductor devices.

Acknowledgement: We acknowledge our family friends and organization for their support in carrying the research work.

Funding Statement: No specific funding for this study is received by the authors for this work.

Conflicts of Interest: The authors declare that we have no conflicts of interest to report regarding the present study.

References

[1] K. J. Ballard, F. S. M. Halaki, K. Paul, A. Daliri and D. A. Robin, "An investigation of compensation and adaptation to auditory perturbations in individuals with acquired apraxia of speech," *Frontiers in Human Neuroscience*, vol. 12, pp. 1–14, 2018.

[2] J. T. George and E. Elias, "A 16-band reconfigurable hearing Aid using variable bandwidth filters," *Global Journal of Researches in Engineering*, vol. 14, no. 1, pp. 1–8, 2014.

[3] L. N. L. Wong, Y. Chen, Q. Wang and V. Kuehnel, "Efficacy of a hearing Aid noise reduction function," *Trends in Hearing*, vol. 22, pp. 1–14, 2018.

[4] H. Levitt, "Noise reduction in hearing aids: A review," *Journal of Rehabilitation Research and Development*, vol. 38, no. 1, pp. 111–121, 2001.

[5] U. Shrawankar and V. Thakare, "Noise estimation and noise removal techniques for speech recognition," *IFIP International Federation for Information Processing*, vol. 340, pp. 336–342, 2010.

[6] G. P. Prajapati and A. Devani, "Review paper on noise reduction using different techniques," *International Research Journal of Engineering and Technology (IRJET)*, vol. 4, no. 3, pp. 522–524, 2017.

[7] Y. Wei and D. Liu, "A reconfigurable digital filterbank for hearing-aid systems with a variety of sound wave decomposition plans," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 6, pp. 1628–1635, 2013.

[8] A. Schasse, "Two-stage filter-bank system for improved single-channel noise reduction in hearing aids," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 2, pp. 383–393, 2015.

[9] H. Liu, R. Zhang, Y. Zhou, X. Jing and T. K. Truong, "Speech denoising using transform domains in the presence of impulsive and Gaussian noises," *IEEE Access*, vol. 5, pp. 21193–21203, 2017.

[10] G. V. P. Chandra Sekhar Yadav, B. A. Krishna and Kamaraju, "Performance of wiener filter and adaptive filter for noise cancellation in real-time environment," *International Journal of Computer Applications*, vol. 97, no. 15, pp. 16–23, 2014.

[11] N. A. M. Abbasa and H. M. Salman, "Independent component analysis based on quantum particle swarm optimization," *Egyptian Informatics Journal*, vol. 19, no. 2, pp. 101–105, 2018.

[12] M. S. Pedersen, D. Wang, J. Larsen and U. Kjems, "Two-microphone separation of speech mixtures," *IEEE Transactions on Neural Networks*, vol. 19, no. 3, pp. 475–492, 2008.

[13] X. Jaureguiberry, E. Vincent and G. Richard, "Fusion methods for speech enhancement and audio source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 7, pp. 1266–1279, 2016.

[14] T. G. Kang, K. Kwon, J. W. Shin and N. S. Kim, "NMF-Based target source separation using deep neural network," *IEEE Signal Processing Letters*, vol. 22, no. 2, pp. 229–233, 2015.

[15] U. N. S. Wood, J. Rouat, S. Dupont and G. Pironkov, "Blind speech separation and enhancement with GCC-NMF," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 4, pp. 745–755, 2017.

[16] S. Raj and A. Shaji, "Design of reconfigurable digital filter bank for hearing Aid," *International Journal of Science and Research (IJSR)*, vol. 5, no. 7, pp. 450–454, 2016.

[17] N. Yang, M. Usman, X. He, M. A. Jan and L. Zhang, "Time-frequency filter bank: A simple approach for audio and music separation," *IEEE Access*, vol. 5, pp. 27114–27125, 2017.

[18] A. S. Kang and R. Vig, "Performance analysis of near perfect reconstruction filter bank in cognitive radio environment," *Interanational Journal of Advanced Networking and Applications*, vol. 8, no. 3, pp. 3070–3083, 2016.

[19] S. Shakya and J. Ogale, "Design and analysis of uniform-band and octave-band tree-structured filter bank," *International Journal of Signal Processing Systems*, vol. 4, no. 2, pp. 162–167, 2016.

[20] I. Missaoui and Z. Lachiri, "Blind speech separation based on undecimated wavelet packetperceptual filterbanks and independent component analysis," *IJCSI International Journal of Computer Science Issues*, vol. 8, no. 1, pp. 265–272, 2011.

[21] K. Tripathi and K. S. Rao, "VEP detection for read, extempore and conversation speech," *IETE Journal of Research*, 2020.

[22] J. Indra, R. K. Shankar, N. Kasthuri and S. Geetha manjuri, "A modified tunable – Q wavelet transform approach for tamil speech enhancement," *IETE Journal of Research*, 2020.

[23] K. K. Shukla and A. K. Tiwari, "Efficient algorithms for discrete wavelet transform," *Springer Briefs in Computer Science*, 2013.

[24] N. Trivedi, V. Kumar, S. Singh, S. Ahuja and R. Chadha, "Speech recognition by wavelet analysis," *International Journal of Computer Applications*, vol. 15, no. 8, pp. 27–32, 2011.

[25] A. Kumar and R. Sunkaria, "Two-channel perfect reconstruction (PR) quadrature mirror filter (QMF) bank design using logarithmic window function and spline function," *Signal, Image and Video Processing*, vol. 10, no. 8, pp. 1473–1480, 2016.

[26] A. N. Akansu and R. A. Haddad, *Multi-resolution Signal Decomposition: Transforms, Subbands, and Wavelets*, Academic Press: Newark, NJ, 2nd ed., 2000.

[27] G. P. Vouras and T. Tran, "Paraunitary filter bank design using derivative constraints," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing - ICASSP '07*, Honolulu, pp. 1453–1456, 2007.

[28] K. M. Ribeiro, R. A. Braga, T. Safadi and G. Horgan, "Comparison between Fourier and wavelets transforms in biospeckle signals," *Applied Mathematics*, vol. 4, pp. 11–22, 2013.

[29] T. Si-Nguyen and W. H. Brian Ng, "Critically sampled discrete wavelet transforms with rational dilation factor of 3/2," in *10th Int. Conf. on Signal Processing Proc.*, Beijing, China, pp. 199–202, 2010.

[30] D. L. Wang and J. Chen, "Supervised speech separation based on deep learning: An overview," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 10, pp. 1–27, 2018.

[31] J. Glass, "Towards unsupervised speech processing," in *11th Int. Conf. on Information Science, Signal Processing and Their Applications (ISSPA)*, Montreal, Quebec, Canada, pp. 1–4, 2012.

[32] Y. Xue, C. S. Tong and T. Li, "Evaluation of distance measures for NMF-based face image applications," *Journal of Computers*, vol. 9, no. 7, pp. 1704–1711, 2014.

[33] Z. Yang, H. Zhang, Z. Yuan and E. Oja, "Kullback-leibler divergence for nonnegative matrix factorization," in *Int. Conf. on Artificial Neural Networks (ICANN)*, Berlin, Heidelberg, pp. 250–257, 2011.

[34] A. Adewusi, K. A. Amusa and A. R. Zubair, "Itakura-saito divergence non negative matrix factorization with application to monaural speech separation," *International Journal of Computer Applications*, vol. 153, no. 9, pp. 17–22, 2016.

[35] Mustaqeem and S. Kwon, "CLSTM: Deep feature-based speech emotion recognition using the hierarchical ConvLSTM network," *Mathematics*, vol. 8, pp. 2133, 2020.

[36] Mustaqeem and S. Kwon, "Optimal feature selection based speech emotion recognition using Two-stream deep convolutional neural network," *International Journal Intelligent System*, vol. 36, no. 9, pp. 1–20, 2021.