International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 7 (2015) pp. 16787-16796
© Research India Publications
http://www.ripublication.com

Removal Of Harmonic Distortion In Audio Signals Using Wavelet Packet Decomposition

T Anusuya[1], R Pratheepa [1], V Elamaran[2]

[1]Department of Electronics and Communication Engg.,
School of EEE, SASTRA University
Thanjavur, Tamilnadu, India
anusuya140@gmail.com
[2]Assistant Professor, Dept., of ECE,
School of EEE, SASTRA University
Thanjavur, Tamilnadu, India
elamaran@ece.sastra.edu

ABSTRACT

The audio signals are frequently affected by a non-linear distortion called harmonic distortion. The non-linear distortion may occur due to any of the amplifier components such as diodes, transistor etc. The presence of harmonic distortion in audio signals degrades its perceptual quality and makes the sound very harsh to hear. Audio signals can be analyzed using various transform techniques. Wavelet analysis is appropriate for analyzing audio signals since it is well suited for analyzing non-stationary signals. An approach using wavelet packet decomposition is suggested to remove the harmonic interference. At a particular constraint condition, the wavelet packet decomposition coefficients correspond to harmonic interference in each subband become constant, which can be exposed as a baseline. By shifting the baseline back to zero the harmonic interference can be removed. The effectiveness of the approach is justified by estimating peak signal to noise ratio (PSNR) for various audio signals. The experimental results are shown using matlab simulation software.

Keywords: Audio analysis, clipping, non-stationary signals, harmonic interference, PSNR, wavelet packet decomposition.

INTRODUCTION

When an audio amplifier attempts to create an output signal beyond its capacity, it simply “clips” at its maximum capacity and produces a clipped form of the signal.

The clipping of signal introduces harmonics at higher frequencies in the frequency domain [1]. The presence of these additional frequency components reduces the quality of audio signal [2].

In order to get rid of the harmonic distortion, audio signals have to be interpreted in frequency domain. Various transform techniques such as Fourier Transform (FT), Short-Time Fourier Transform (STFT), wavelet transform are available to analyze audio signals in spectral domain. In Fourier Transform, a signal can be viewed in frequency domain by breaking down the signal into its constituent spectral components. But it cannot give access to the spectral variations during a particular time interval. Hence, FT is not apposite for non-stationary signals [3]. Audio signals are not always stationary. STFT can be able to analyze non-stationary signals by using finite length window function. But resolution problems will occur since the size of the analyzing window remains unchanged throughout the analysis. Wavelet transform resolve this resolution problem by using appropriate window sizes for different regions of the signal. So it is flexible to analyze both stationary and non-stationary signals. Therefore, in the midst of these transformations wavelet transform is well suited to scrutinize audio signals [4].

Wavelet packet decomposition is a discrete wavelet transform provides a powerful tool for signal analysis [5]. Using suitable wavelet, signal is decomposed into approximate and detail subbands. The low frequency components which give the signal its identity are preserved in approximate coefficients. The high frequency components together with harmonic interference and other noisy components are represented in detail coefficients. These subbands are further decomposed and the decomposition continues till the cut-off frequency of approximate subband becomes smaller than the fundamental frequency of the harmonics. The harmonic distortion can be eliminated simply by thresholding the detail coefficients when the signal to noise ratio (SNR) is large enough. The degree of distortion increases with decrease in the level of SNR. The bottom level subbands have to be chosen for processing to remove the harmonics [6].

Mean of the wavelet coefficients in a subband can be referred as baseline of the coefficients [7]. If the signal is contaminated by harmonics, non-zero baseline may appear in detail subband. The harmonic interference is solely responsible for the occurrence of non-zero baseline, when a special constraint condition is satisfied. The harmonics can be removed by shifting the baseline back to zero. The shifting of baseline does not disturb the original signal components.

This paper is presented as follows: fundamentals of wavelet transform are described in section II. Section III explains the removal of harmonics by shifting the baseline in detailed manner. The simulation results are discussed in section IV and the paper is concluded in section V.

II. FUNDAMENTALS OF WAVELET TRANSFORM

Wavelet is a small waveform or a brief oscillation of limited duration and irregular shape. The wavelet is capable of analyzing a localized area of a large signal since it has a time localization property [8]. In wavelet analysis, a signal is decomposed into shifted and scaled versions of the mother wavelet. A function which has a wavy appearance and has most of the energy confined in a finite duration can be able to act as a wavelet. Continuous Wavelet Transform (CWT) is nothing but a measure of correlation between the signal and a wavelet function.

In Discrete Wavelet Transform (DWT), a signal is decomposed by passing it through series of low pass and high pass filters. The low pass filter extracts the coarse signal information called approximation. On the other hand, the high pass filter extracts the sharp transitions and noisy components of the signal called detail. After the filtering process the resultant signals are downsampled to maintain good resolution. The decomposition process can be iterated by further decomposing the approximation and it is continued until the desired level is reached.

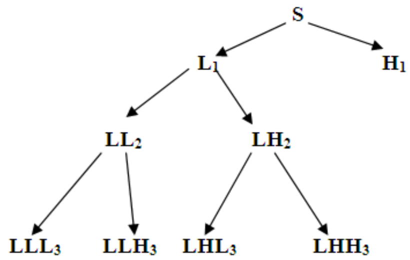

Fig 1: Three level wavelet decomposition tree

Octave-band tree structure can be used to represent wavelet decomposition. Simple tree structure is shown in Fig 1. Bottom up approach is used to reconstruct the original signal from the decomposed coefficients.

Wavelet packet decomposition is a generalization of wavelet decomposition method. Wavelet packet is a collection of wavelet functions. The wavelet packets provide more flexibility in terms of scale and position than a wavelet. In wavelet packet transform the decomposition process is iterated by decomposing both approximation and detail subbands. This yields 2^L nodes as a result, where *'L'* denotes the deepest level of decomposition. The wavelet packet decomposition tree structure can is shown in Fig 2.

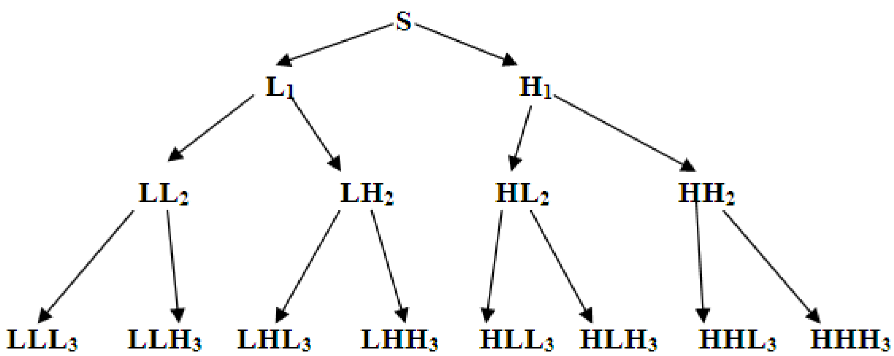

Fig 2: Three level wavelet packet decomposition tree

The original signal can be reconstructed from the approximate and detail coefficients by using bottom-up approach as in wavelet decomposition method.

Harmonic interference in audio signals can be removed by using this wavelet packet transform. Noisy components in the signal can be removed by thresholding the coefficients in the deepest level. Thresholding is nothing but reducing the size of the signal by removing or by reducing the quantity of coefficients below the threshold value. But this process is effective only if the SNR is large enough. Otherwise, thresholding the detail coefficients may distort the original signal. A distortion free approach can be obtained, if a particular constraint among the sampling rate, fundamental frequency of harmonics and the decomposition level is limited.

III. REMOVAL OF HARMONICS BY BASELINE SHIFTING

Consider a signal *s*, contaminated by the harmonics h_j can be described as,

$$y = s + h_j = s + \sum_{j=1}^{K} h_j \qquad (1)$$

where, *h* denotes the sum of the harmonics, *K* denotes the highest order of the harmonics and

$$h_j = a_j \sin(2\pi f_j t + \varphi_j) \qquad (2)$$

a_j, f_j, and φ_j represents the amplitude, frequency and phase of the *j*th harmonic. $f_j = i.f_1$ where, f_1 denotes the fundamental frequency of the harmonics.

In wavelet packet decomposition, let '*l*' be the level of decomposition then $l \le L$, where, $L \le log_2 N$ denotes the deepest level of the tree. Each coefficient in C_l (of *y*) at level *l* is the sum of two transform coefficients such as *s* and *h*, it can be written as,

$$C_l[n] = Cs_l[n] + Ch_l[n] \qquad (3)$$

In equation (3), the coefficients *C* represent the wavelet packet coefficients, while Cs_l and Ch_l are the transform coefficients of *s* and *h* respectively.

The coefficients C_l are considered to be obtained by filtering the contaminated signal i.e. sum of original signal *s* and the harmonics h_j through a bandpass filter followed by a factor 2^l subsampling. If the frequency of h_j is within the passband, h_j can be retained into C_l. The retained h_j is sinusoidal since the filter is linear. Other harmonics which are not within the passband of the filter bank may also be retained by C_l due to the spectral leakage of the filter.

(a) BASELINE OF THE WAVELET PACKET COEFFICIENTS

Mean of the wavelet packet coefficients in each subband can be referred as a baseline. The harmonic interference has to be eliminated at the deepest level *L* only hence the baseline of the wavelet packet coefficients at level *L* is sufficient for further processing. Due to 2^L subsampling at level *L*, only $1/2^L$ samples of filtered harmonics are detained by C_l. The detained samples do not depend on the relationship between the fundamental frequency f_1 of the harmonics, the sampling rate *Fs* and the level of decomposition *L*.

If f_1, Fs and L satisfy,

$$F_s = f_1 . 2^L \tag{4}$$

the harmonic components become constant. The transform coefficients of the harmonic interference in each subband at level L appears as constant if (4) is satisfied. This is because, after 2^L subsampling, for every j cycle of the jth frequency component only one sample can be retained in Ch_l. The subsampling takes place at the same phase of each sinusoidal wave corresponding to each frequency of the harmonics. As a consequence the harmonic interference becomes constant and the baseline of the coefficients can be shifted upwards or downwards with respect to the value of the constant. If the boundary effect [9] is ignored, Cs_l is zero mean for all the subbands in the deepest level L except for the left most subband in the decomposition tree, i.e. for instance, **LLL_3** subband in Fig (2). Hence non-zero baseline could appear only due to the presence of harmonic interference. If the non-zero baseline is moved back to zero, Ch_l in (4) can be removed without disturbing the transform coefficients of the original signal.

The approximate subband **LLL_3** in Fig (2) is the resultant of low pass filter which contains most of the original signal component. Only very small amount of the harmonics is possessed by that subband. Hence it is not necessary to deal with the approximate coefficients [7].

In real time, the sampling rate of the audio signal may not satisfy (4). To handle this situation the sampling rate of the signal can be changed in the way to satisfy equation (4) before acquisition. Otherwise, the signal can be resampled and the sampling rate is altered to the nearest value that satisfies (4). The altered sampling rate has to be larger than the original in order to prevent the loss in signal detail.

(b) ESTIMATION AND SHIFTING OF BASELINE

The baseline of the wavelet packet coefficients in each subband can be obtained by calculating the mean of the coefficients.

$$m_l = 2^l/N \sum_{n=0}^{N/2^l-1} C_l[n] \tag{5}$$

where, m_l is the baseline of the coefficients C_l, l denotes the level of decomposition and N represents the signal length.

The transform coefficients of the sharp transition details and the boundary coefficients due to boundary effect may affect the baseline estimation. To reduce the impact, transform coefficients related to sharp details and boundary effect has to be ignored in the baseline estimation. The Donoho's threshold [10] has to be calculated to differentiate these coefficients.

$$t = \sigma\sqrt{2\log N} \tag{6}$$

where, σ denotes the standard deviation of the transform coefficients, N represents the signal length.

The baseline has to be estimated in an iterative manner. The steps involved in estimating the baseline is as follows,

i. Estimate *m* using all the subband coefficients in the level *L*.
ii. Calculate σ and *t*.
iii. Limit the range of the coefficients within the interval [m-t, m+t]
iv. Reconstruct the coefficients
v. Repeat steps (i) to (iv) for the retained coefficients.

The iteration has to be continued till the relative error between the estimated baseline and the expected baseline approach near zero. The expected baseline is estimated with the harmonics. This baseline estimation procedure is not applicable for left most approximate coefficient since most of the signal components are preserved in this subband.

(C) IMPLEMENTATION OF HARMONICS REMOVAL

The harmonics in the audio signal can be removed by implementing the following steps

i. Find the sampling rate F_s of the contaminated signal and change (if necessary) in terms of (4).
ii. Determine the level of decomposition *L* with respect to the constraint condition (4).
iii. Select the wavelet packet and decompose the signal to the determined level *L*.
iv. Calculate the baseline of the wavelet packet coefficients in each subband at level *L*.
v. Shift the baseline back to zero.
vi. Reconstruct the signal.
vii. Change the sampling rate to its original.

The audio signals may have more channels. In those cases the baseline has to be estimated separately for each channel. For example, stereophonic audio signals divided across two channels; therefore the baseline has to be estimated separately for two channels before shifting and reconstruction.

IV. SIMULATION RESULTS

A stereophonic audio signal is taken and the pictorial representation of the signal is shown in fig (3(a)). The harmonics is added and the corrupted signal is shown in fig (3(b)). With respect to (4), by considering the sampling frequency of the signal and the fundamental frequency of the harmonics the depth of decomposition is chosen as three.

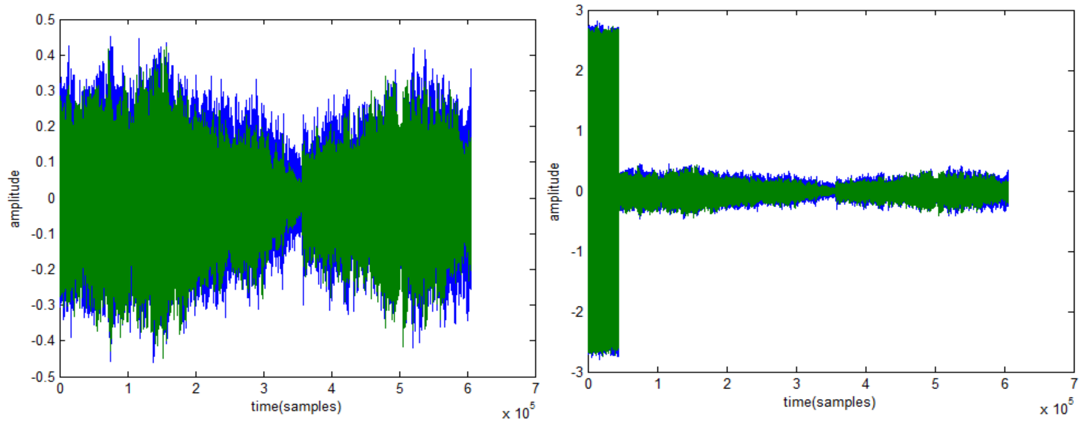

Fig 3: (a) Original audio signal. (b) Contaminated audio signal.

The wavelet packet obtained from Daubechies wavelet "db1" [11] is selected for analysis. The level 3 decomposition is performed on the contaminated signal and 2^3 subbands are obtained the deepest level. The wavelet packet decomposition coefficients at level 3 are shown in fig (4). In order to remove the harmonics the baseline has to be estimated for all the subband coefficients in the deepest level.

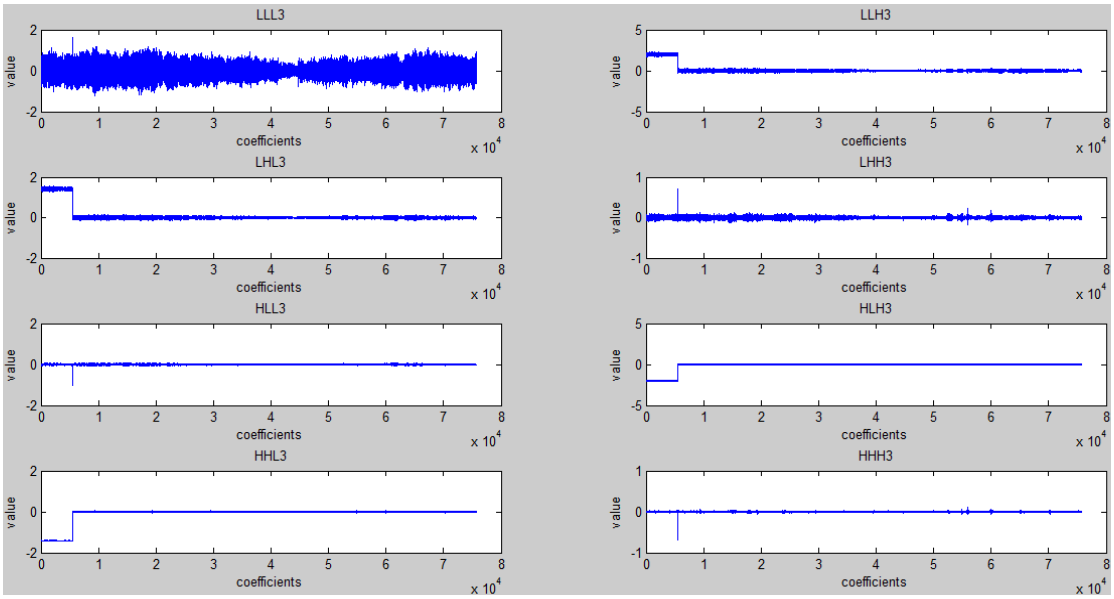

Fig 4: wavelet packet decomposition coefficients at level 3

Let m_a be the estimated baseline and m_e be the expected baseline, the relative error is defined as,

$$r\,(m_a) = \frac{m_a - m_e}{std(y)}\,X\,100\% \quad (7)$$

The baseline estimation has to be iterated until the relative error approach near zero. Then the signal can be reconstructed. The reconstructed signal after the removal of harmonics is shown in fig (5). There is no visible difference while comparing the contaminated and reconstructed audio signals.

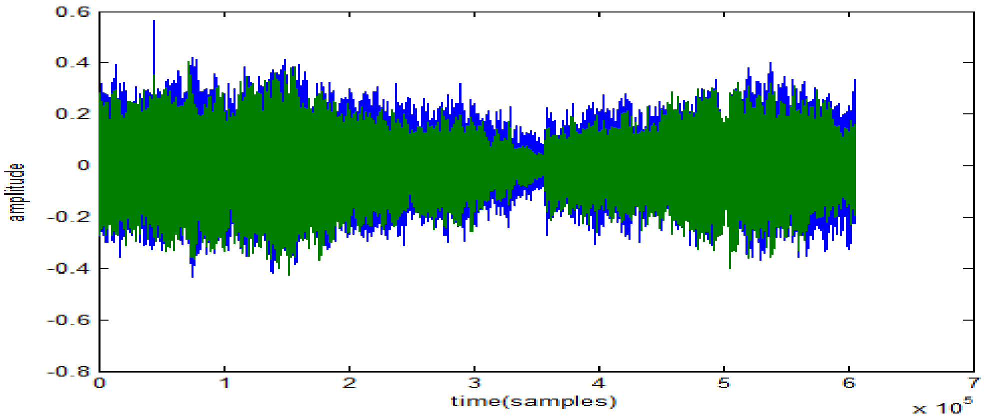

Fig 5: Reconstructed audio signal after the removal of harmonics

Peak signal to noise ratio (PSNR) is estimated as a performance metric in order to determine the performance of the system. Let *A* be the original audio and *B* be the corrupted audio or reconstructed audio. Signals *A* and *B* are having two channels, *i* be one channel and *j* be another channel. The PSNR can be defined as,

$$PSNR = 20*log10\,\frac{length(A)}{\sqrt{MSE}}\,dB \quad (8)$$

Where,

$$MSE = \frac{1}{length(i)*length(j)}\sum_{ij}(A(i,j)-B(i,j))^2 \quad (9)$$

The PSNR of the audio signal in fig (3(b)) is 51.73 dB whereas SNR of the harmonics removed audio signal in fig (5) is 79.27dB. Various music signals and speech signals are tested and the PSNR is tabulated as shown in Table 1 and Table 2.

Table 1. PSNR results for music.

Test signals	PSNR (dB)	
	Before removal of harmonics	**After removal of harmonics**
Music 1	52.27	73.99
Music 2	54.51	81.20
Music 3	54.85	83.18
Music 4	54.84	77.20

Table 2: PSNR results for speech

Test signals	PSNR (dB)	
	Before removal of harmonics	**After removal of harmonics**
speech 1	53.67	74.94
speech 2	52.58	72.34
speech 3	51.61	65.37
speech 4	47.16	77.58

CONCLUSION

Using wavelet packet decomposition, an approach has been conferred to remove harmonic distortion from the audio signal. Since the harmonic interference is eliminated by shifting the baseline without disturbing the signal components, the approach can be a distortion-free approach. The performance of the approach is justified by estimating PSNR for various test signals. The experimental results show the effectiveness of this approach.

REFERENCES

[1] S Temme, "Application Note - Audio distortion measurements", Bruel & Kjar, 1992 - diy-audio.narod.ru

[2] Bruno Defraene, Toon van Waterschoot, Hans Joachim Ferreau, Moritz Diehl and Marc Moonen, "Perception-based Clipping of Audio Signals", Proceedings of the 18th European Signal Processing Conference (EUSIPCO 2010), pp. 517-521.

[3] Robi Polikar, "The Wavelet Tutorial", second edition, http://web.iitd.ac.in/~sumeet/WaveletTutorial.pdf

[4] G. Tzanetakis, G. Essl, P. Cook, "Audio analysis using the discrete wavelet transform", in: C.E.D’ Attellis, V.V. Kluev, N. Mastorakis (Eds.), Mathematics and Simlulation with Biological Economical and Musicoacoustical Applications, WSES Press, New York, 2001, pp. 318–323.

[5] S. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” IEEE Trans. Pattern Anal. Machine Intell., vol.11, no. 7, pp. 674–693, Jul. 1989.

[6] Ronald R Coifman, Yves Meyer, Steven Quake, M.Victor Wickerhauser, "Signal Processing and Compression with Wavelet Packets", preprint, Yale Univ, 1990.
[7] Lijun Xu, "Cancellation of Harmonic Interference by Baseline Shifting of Wavelet Packet Decomposition Coefficients", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 53, NO. 1, JANUARY 2005
[8] Lee A. Barford, R. Shane Fazzio, David R. Smith, "An Introduction to Wavelets", IEEE Computational Science & Engineering, 2(2), pp. 50-61, 1995.
[9] S. Mallat, A Wavelet Tour of Signal Processing. San Diego, CA: Academic,1998.
[10] D. L. Donoho, “De-noising by soft-thresholding,” IEEE Trans. Inf.Theory, vol. 41, no. 3, pp. 613–627, Jun. 1995.
[11] I. Daubechies, “Orthonormal bases of compactly supported wavelets,” Commun. Pure. Applied Math., vol. 41, pp. 909–996, 1988.