Denoising of Speech Signals with Impulse Noise Using Optimal Wavelets

N. Ravisankar [1], D. Leela Rani [2]

[1]Department of ECE, PG Student, Sree Vidyanikethan Engineering College, Tirupati

[2]Department of ECE, Associate Professor Sree Vidyanikethan Engineering College, Tirupati

Abstract: ***Expulsion of drive clamor from discourse in the wavelet space has been discovered to be exceptionally powerful because of the multi-determination property of the wavelet change and the simplicity of uprooting the motivations in that space. A basic variable that influences the execution of the drive evacuation framework is the viability of the motivation location calculation. To this end, we propose another strategy for outlining orthogonal wavelets that are streamlined for distinguishing motivation clamor in discourse. In the technique, the attributes of the motivation clamor and the hidden discourse sign are considered and a curved streamlining issue is detailed for inferring the ideal wavelet for a given bolster size.***

Keywords: impulsive noise detection, wavelet design, speech enhancement, multi-resolution property**,** signal to impulse ratio

1. Introduction

The vicinity of motivation like commotion in discourse can altogether decrease the comprehensibility of discourse and debase programmed discourse acknowledgment (ASR) execution. Motivation commotion is described by short blasts of acoustic vitality having a wide unearthly data transfer capacity and comprising of either disengaged driving forces or a progression of motivations. Normal acoustic motivation commotions incorporate hints of snaps in old phonograph recordings, of downpour drops hitting a hard surface like the windshield of a moving auto, of popping popcorn, of writing on a console, of pointer clicks in autos, etc.

As of late, a few systems for location and/or evacuation of transient furthermore, motivation commotions have been accounted for. In [1], drive commotion was expelled from sound flags by combining various duplicates of the same recording, while in [2] the otherworldly cognizance and consonant property of discourse were utilized to recognize transient commotion from discourse. Established piece handling systems, for example, the STFT calculation alternately the straight expectation (LP) calculation have likewise been utilized to distinguish or uproot drive like sounds [3], [4], [5]. The impulse noise in speech is eliminated by other techniques like using the iterative algorithm [6] and cluster based adaptive fuzzy switching median filter algorithm [7].Notwithstanding, two issues may come about if excellent piece handling procedures are utilized: the primary is deciding the precise position of the drive inside of the investigated information outline – these routines give no direct data about the position of the motivation inside of the dissected casing It is conceivable, on the other hand, to diminish the casing size to accomplish better determination in time; however doing this prompts the second issue where we lose the recurrence determination expected to viably dissect the sign. The wavelet change overcomes both of these troubles due to its multi-determination property [8]. In multi-determination investigation, the window length or wavelet scale for investigating the recurrence parts increments as the recurrence diminishes. This property empowers the wavelet change to have better time determination for higher recurrence parts and better recurrence determination for lower one. Hence, by utilizing the wavelet change we have a relationship between time determination and recurrence determination that is advantageous for distinguishing and uprooting motivation clamor.

The utilization of the Daubechies wavelet has been discovered to be very viable in the recognition and expulsion of drive clamor from discourse or sound [9], [10]. In spite of the fact that such a wavelet may be exceptionally powerful in one application, it may not be as viable in another where the properties of the drive commotion and the hidden sign are diverse. Hence, to empower the originator select the suitable wavelet for a given application, an association between certain wavelet highlights and motivation discovery execution was made in our late work [11].

In that work, we demonstrated how the wavelet motivation discovery highlights are reliant on the qualities of the motivation commotion and the basic flag, and gave a strategy to selecting the most fitting wavelet from an arrangement of predesigned wavelets. The system, in any case, has one disadvantage: the nature of the chose wavelet is subject to the nature of the wavelets inside of the set. On the off chance that none of the wavelets inside of the set are ideal for the given application, the system won't be powerful. In this paper, we try to uproot the disadvantage in our past work [11] by outlining wavelets that are most suitable for a given application. Using the connections between wavelet highlights and motivation identification execution [11], we detailed a streamlining issue for outlining a wavelet of certain bolster estimate that is customized for distinguishing motivations for a given application. The details are encircled as a raised streamlining issue where the arrangement gotten relates to the FIR channel coefficients of an orthogonal wavelet. The ensuing execution examination results with other no doubt understood wavelets demonstrate that the wavelets composed utilizing the proposed system have vastly improved drive recognition highlights.

The paper is composed as takes after. Area 2 abridges about the impulse noise in speech. In Section 4, we create plans to acquire the channel coefficients of the ideal wavelet for a given bolster size. At that point in Section 5, reproduction trials are exhibited to think about the motivation recognition execution of wavelets determined utilizing the proposed technique with other no doubt understood wavelets. Conclusions are attracted Section 6.

2. Impulse noise in speech

Drive commotion is portrayed by short blasts of acoustic vitality having a wide ghostly data transmission and comprising of either detached motivations or a progression of driving forces. Impulse noise is defined as

$$n(t) = \sum_i A_i . P . W_i (t - \tau_i) \tag{1}$$

Where

A_i = the amplitude of impulse i

w_i = the duration of impulse i

t_i = the time of occurrence of impulse i

For impulsive noise the average signal to impulsive noise ratio, averaged over an entire noise sequence including the time instances when the impulses are not present, depends on two parameters one is the average power of each impulsive noise, and second one is the rate of occurrence of impulsive noise. Let *P*impulse denote the average power of each impulse, and *P*signal the signal power. We may define a "local" time-varying signal to impulsive noise

$$SINR(m) = \frac{P_{signal}(m)}{P_{impulse}\, b(m)} \tag{2}$$

Let the parameter α is the fraction of signal samples contaminated by impulsive noise, can be defined as

$$SINR(m) = \frac{P_{signal}(m)}{\alpha P_{impulse}(m)} \tag{3}$$

3. Wavelet Properties and Features for Impulse Detection

In this segment, we abridge the wavelet properties that impact the identification execution and depict a measure for assessing the identification execution. An alluring wavelet for motivation recognition is one that augments the coefficients for the drive with respect to the fundamental flag in the finest scale [9]. Such a wavelet will correspondingly have a high pass investigation channel that boosts the drive commotion with respect to the basic discourse and foundation clamor signals. In the event that Ps (ω) and Pi (ω) are the force ranges of the normal discourse and motivation commotion power, separately, then the proportion between the normal drive commotion power furthermore, discourse control in the finest scale, Ri, is subject to the wavelet high pass investigation channel and given

$$R_i = \frac{{\sigma_i}^2}{{\sigma_s}^2} \tag{4}$$

Where

$$\sigma_i^2 = \int_{-\pi}^{\pi} \left|G(e^{jw})\right|^2 P_i(w) \approx \sum_i \left|G(e^{jw})\right|^2 P_i(w) \tag{5}$$

$$\sigma_s^2 = \int_{-\pi}^{\pi} \left|G(e^{jw})\right|^2 P_s(w) \approx \sum_i \left|G(e^{jw})\right|^2 P_s(w) \tag{6}$$

also, G(z) is the exchange capacity of the wavelet high pass channel. The configuration of an ideal wavelet for identifying the driving forces ought to, hence, looks to augment Ri. The other variable that impacts the location execution is the size of the wavelet bolster, which is reliant on the normal width also, vitality of the motivation clamor [11]. One approach to focus the right wavelet support for a given application is to plan wavelets that augment Ri at different wavelet bolster sizes and afterward select the one with the best identification execution.

4. Derivation of Optimal Wavelet for Impulse Detection

The ideal wavelets are intended to augment the proportion of drive clamor energy to discourse control in the finest scale. In the meantime, the important requirements needed for an orthogonal wavelet need to be forced. On the off chance that H(z) compares to the exchange capacity of a low pass examination channel of an orthogonal wavelet given by

$$H(z) = h(0) + h(1)z^{-1} + \cdots + h(L-1)z^{-(L-1)} \tag{7}$$

at that point the high pass partner, G(z), can be acquired by taking the exchanging flip of H(z) [12]; that is

$$G(z) = -Z^{-(L-1)} H(-Z^{-1}) \tag{8}$$

Where L is thought to be even. To guarantee that the wavelet filter bank is orthogonal, the channel coefficients need to fulfill the twofold movement orthogonality condition [12], given by

$$\sum_n h(n)h(n-2k) = \delta(k), for\ k = 0,1, \ldots . (L/2) - 1 \tag{9}$$

Where δ (k) is the delta capacity. For the presence of the wavelet ψ (t), the accompanying condition should likewise remain constant [13]

$$H(e^{iw}) | (w = 0) = \sum_n h(n) = \sqrt{2} \tag{10}$$

As in the configuration of sign adjusted filter banks, the plan of the streamlining issue turns out to be more tractable in the event that we utilize the autocorrelation succession of the channel coefficients given by

$$r_h(2k) = \begin{cases} \sum_{n=0}^{L-l-1} h(n)h(n+l) & l \geq 0 \\ r_h(-l) & l < 0 \end{cases} \quad (11)$$

In this way, as far as the autocorrelation parameters, the twofold shift orthogonality condition in (9) can be communicated

$$r_h(2k) = \delta(k) for \ k = 0,1, \ldots \ldots, \left\lfloor \frac{L-1}{2} \right\rfloor \quad (12)$$

and the necessary condition is

$$\sum_{m=1}^{L-1} r_h(m) = 0.5 \quad (13)$$

by misusing the orthogonality condition in (9) and the symmetry property in (11). Correspondingly, utilizing (8) and (11) in (5) and (6) the normal force of the drive commotion and discourse in the finest scale are given by

$$\sigma_i^2 \approx \sum_n [r_h(0) + 2\sum_{l-1}^{L-1}(-1)^l r_h(l) \cos(w_n l) P_i(w_n)$$

$$= 1^T C_i A_r \quad (14)$$

$$\sigma_s^2 \approx \sum_n [r_h(0) + 2\sum_{l-1}^{L-1}(-1)^l r_h(l) \cos(w_n l) P_s(w_n)$$

$$= 1^T C_s A_r \quad (15)$$

Where

$$r = [r_h(0) \ldots \ldots r_h(L-1)]^T \quad (16)$$

$$A = \begin{bmatrix} a_{00} & \cdots & a_{0(L-1)} \\ \vdots & \ddots & \vdots \\ a_{(N-1)0} & \cdots & a_{(N-1)(L-1)} \end{bmatrix} \quad (17)$$

$$C_i = diag(c_0^{(i)}, \ldots \ldots, c_{(N-1)}^{(i)}) \quad (18)$$

$$C_i = diag(c_0^{(s)}, \ldots \ldots, c_{(N-1)}^{(s)}) \quad (19)$$

$$a_{nl} = 2(-1)^l \cos(w_n l) \quad (20)$$

$$C_n^{(i)} = P_i(w_n), \quad w_n \in [-\pi, \pi] \quad (21)$$

$$C_n^{(s)} = P_s(w_n), \quad w_n \in [-\pi, \pi] \quad (22)$$

what's more, N is the quantity of tests. The advancement is figured as the minimization of while σ_s^2 keeping σ_i^2 consistent so that Ri in (1) is expanded. Thus, in the wake of joining the twofold movement orthogonality limitation in (12) and the fundamental condition in (13),

Minimize σ_s^2

Subject to: σ_i^2 =constant

$$r_h(2m) = 0 \ for \ k = 0,1, \ldots, \left\lfloor \frac{L-1}{2} \right\rfloor \quad (23)$$

$$r_h(0) = k$$

$$\sum_{m=1}^{L-1} r_h(m) = 0.5k \quad (24)$$

The enhancement issue is given by where k and r_h (m) are advancement variables. Note that the last two balance requirements in (21) guarantee that the important condition in (11) is fulfilled when we situated $r_h(0) = 1$. Supplanting σ_s^2 and σ_i^2 by their network representations, (21) can be communicated as a curved

Minimize $1^T C_s A_r$ (25)

Subject to $1^T C_i A_r = constant$

$$r_h(2m) = 0 \ for \ k = 0,1, \ldots, \left\lfloor \frac{L-1}{2} \right\rfloor \quad (26)$$

$$r_h(0) = k$$

$$\sum_{m=1}^{L-1} r_h(m) = 0.5k$$

$$A_r > 0$$

advancement issue given by where r and k are streamlining variables and $0 \in R^N$. The imbalance imperative in (22) is an energy limitation to guarantee that the greatness is constantly positive. When we acquire the ideal autocorrelation vector ropt, we recoup the base stage low-pass wavelet channel coefficients hmp(n) from ropt utilizing ghastly factorization. The channel coefficients acquired are then suitably scaled so that the fundamental condition in (8), or comparably in (11), is satisfied.

5. Experimental Results

In this section we are taking the speech signal and added the randomly generated impulse noises to the speech signal. After getting the signal with impulse noise, this noisy signal is passed through the designed optimal wavelet. Finally we get the enhanced signal from the denoised signal with better SNR value. To design the wavelet for different support size below order coefficients are used.

Table 1: Low pass filter coefficients

Wavelet type	Wavelet low pass filter coefficients
Proposed-2	0.7071085341784 0.7071050281951
Proposed-4	0.48334487737972 0.83639971558336 0.22380051117744 -0.12933154176747
Proposed-6	0.38024506103769 0.79766077189892 0.43134520739550 -0.14033335583051

	-0.10451985196279 0.04981572983405
Proposed-8	0.309244942584623 0.746498014127496 0.544793281489114 -0.056886405860768 -0.203884318310720 0.041029596831375 0.057037843433025 -0.023619391921051
Proposed-10	0.286710821179274 0.709732716949922 0.581544846361955 0.011959274650484 0.241340577819909 -0.016393787204077 0.122292948688230 0.015163126219072 0.042122417254974 0.0169928630412

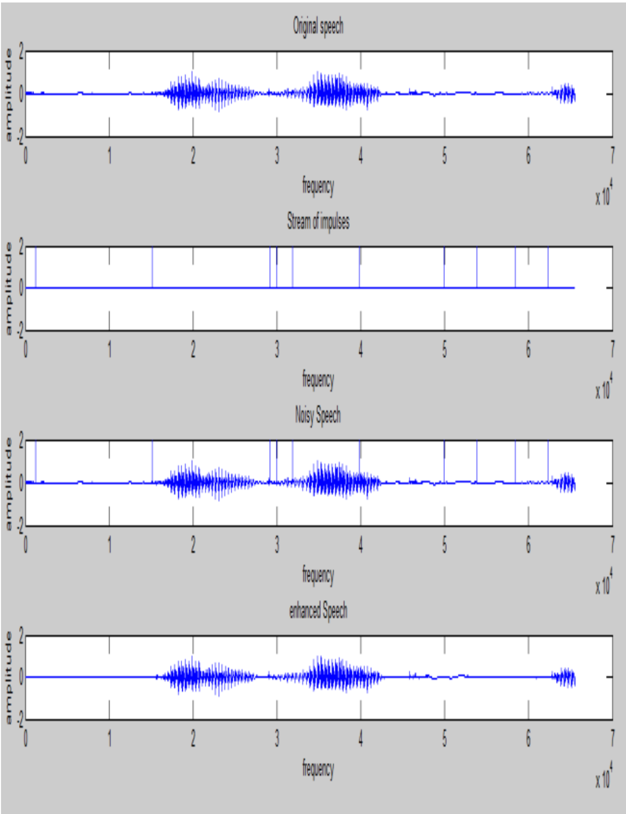

Figure 1: Elimination of impulse noise from speech signal using wavelets

In figure1, initially original signal is taken and sampled at 16 KHz. Then impulses are added randomly by using the impulse noise generator for getting the noisy signal. Then the noisy signal is passed through the designed wavelet with required support length. Finally we get the enhanced signal without impulse noise.

Input signal SNR =110.09988dB
Noisy signal SNR =14.0597dB
Enhanced signal SNR =48.1729dB

6. Conclusion and Future Scope

New method for detecting impulse noise in speech is described by designing the optimal wavelets using the multi-resolution property of wavelets for required wavelet support size. Designing is based on consideration of the impulse noise characteristics and underlying speech signal. The designed optimal wavelets are also used to eliminate the other noises like AWGN, pink noise, babble noise etc.

References

[1] P. Sprechmann, A. Bronstein, J.-M. Morel, and G. Sapiro, "Audio restoration from multiple copies," *Proceedings of ICASSP*, pp. 878-8829 (2013).

[2] C. Zheng, X. Chen, S. Wang, R. Peng, and X. Li, "Delayless method to suppress transient noise using speech properties and spectral coherence," *Proceedings of the 135th AES Convention*, New York, USA (2013).

[3] Z. Liu, A. Subramanya, Z. Zhang, J. Droppo, and A. Acero," Leakage model and teeth clack removal for air-and boneconductive integrated microphones," in *Proceedings of ICASSP*, vol. 1, pp. 1093-1096 (2005).

[4] S. V. Vaseghi and R. Frayling-Cork, "Restoration of old gramophone recordings," *J. Audio Eng. Soc.*, 40, 791-801 (1992).

[5] J. A. Moorer, "Dsp restoration techniques for audio," *Proceedings of ICIP*, vol. 4, pp. 5-8(2007).

[6] Mathew, M.E.Jeevitha,"An impulse noise cancellation using iterative algorithms", *Proceedings of ICECS*,pp 1-6,(2014).

[7] Toh, K.K.V.Isa,"Cluster based adaptive fusing switching median filter for universal impulse noise reduction", *IEEE Proceedings*, pp 2560-2568(2010).

[8] S. Mallat, "A Wavelet Tour of Signal Processing", 2nd ed. (Academic, San Diego), pp. 221-228(1998).

[9] S. Montresor, J. C. Valiere, J. F. Allard, and M. Baudry, "The restoration of old recordings by means of digital techniques," in *Proceedings of the 88th AES Convention*, Montreux, Switzerland (1990).

[10] R. C. Nongpiur, "Impulse noise removal in speech using wavelets," *Proceedings of ICASSP*, pp. 1593-1597(2008).

[11] R. C. Nongpiur and D. J. Shpak, "Impulse-noise suppression in speech using the stationary wavelet transform,"*J. Acoust. Soc. Am.*, 133(2), 866-879 (2013).

[12] G. Strang and T. Nguyen,"Wavelets and filter banks", Wellesley-Cambridge Press (1997).

[13] C.S.Burrus, R.A.Gopinath, and H.Guo, "Introduction to Wavelets and Wavelet Transforms", (Prentice Hall, Upper Saddle River, NJ), pp. 53(1998).

[14] P. Moulin and M. K. Mihcak, "*Theory and design of signal adapted FIR paraunitary filter banks*", *IEEE Trans. Signal Processing*, 46(4), pp 920-929 (1998).

[15] A. H. Sayed and T. Kailath, "A survey of spectral factorization methods", *Numer. Linear Algebra Appl.*, 8, pp 467-496 (2001).

Author Profile

N.Ravisankar is currently pursuing M.Tech Digital Electronics and communication systems in Sree Vidyanikethan Engineering College, Tirupati and he completed his B.Tech in Electronics and communication engineering in Madanapalli Institute of Technology and Science, Madanapalli affiliated to JNTU, Ananthapur. My interests are in Digital electronics and signal processing.

Ms. D.Leela Rani received the M.Tech. Degree from Sri Venkateswara University, Tirupati. She is currently working towards the Ph.D. degree in the Department of Electronics and communication Engineering, SVU College of Engineering, Tirupati. Currently she is working as an Associate professor in Sree Vidyanikethan Engineering College (Autonomous). Her research areas include, Atmospheric Radar Signal Processing.