ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Delayless Method to Suppress Transient Noise by Design and Using Optimal Wavelets

Masthan Andela
PG Scholar,
Dept of ECE,
MJR College of Engineering & Technology, Piler, Andhra - Pradesh 517214, India.

N. Ramanjulu
Assistant Professor,
Dept of ECE,
MJR College of Engineering & Technology, Piler, Andhra - Pradesh 517214, India.

K. Pradeep Kumar
HOD,
Dept of ECE,
MJR College of Engineering & Technology, Piler, Andhra - Pradesh 517214, India.

ABSTRACT:

Expulsion of motivation commotion from discourse in the wavelet area has been observed to be extremely compelling due to the multi-determination property of the wavelet change and the simplicity of evacuating the driving forces in that space. A basic factor that influences the execution of the drive evacuation framework is the viability of the motivation detection calculation. To this end, we propose another strategy for outlining orthogonal wavelets that are advanced for distinguishing motivation commotion in discourse. In the technique, the qualities of the drive commotion and the hidden discourse flag are considered and a raised operation timization issue is figured for inferring the ideal wavelet for a given help measure. Execution examination with other surely understood wavelets demonstrate that the wavelets outlined utilizing the proposed technique have much better motivation location properties.

Index Terms:

Impulsive noise detection, wavelet design, speech enhancement.

1. INTRODUCTION:

The nearness of motivation like clamor in discourse can fundamentally re-duce the clarity of discourse and corrupt programmed discourse recognition (ASR) execution. Drive clamor is described by short blasts of acoustic vitality having a wide unearthly data transmission and consisting of either disconnected motivations or a progression of driving forces.

Run of the mill acoustic drive commotions incorporate hints of snaps in old phonograph accounts, of rain 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, et cetera. As of late, a few techniques for location or potentially expulsion of transient and drive clamor have been accounted for. In [1], drive commotion was expelled from sound flags by melding different duplicates of a similar account, while in [2], the otherworldly cognizance and symphonious property of discourse were utilized to recognize transient clamor from discourse. Established piece handling techniques, for example, the STFT algorithm or the direct expectation (LP) calculation have likewise been utilized to recognize or evacuate drive like sounds [3, 4, 5].

In any case, two issues may come about if exemplary square preparing procedures are utilized: the first is deciding the correct position of the motivation inside the broke down information outline – these strategies give no direct information about the position of the drive inside the dissected casing. It is conceivable, nonetheless, to lessen the edge size to accomplish better resolution in time; however doing this prompts the second issue where we lose the recurrence determination expected to successfully break down the signal.

Cite this article as: Masthan Andela, N. Ramanjulu & K. Pradeep Kumar, " Delayless Method to Suppress Transient Noise by Design and Using Optimal Wavelets", International Journal & Magazine of Engineering, Technology, Management and Research, Volume 5, Issue 3, 2018, Page 22-28.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The wavelet change defeats both of these troubles because of its multi-determination property [6]. In multi-determination examination, the window length or wavelet scale for dissecting the recurrence components increments as the recurrence diminishes. This property empowers the wavelet change to have better time determination for higher frequency parts and better recurrence determination for bring down ones. Therefore, by utilizing the wavelet change we have a connection between time determination and recurrence determination that is advantageous for recognizing and evacuating motivation clamor. The utilization of the Daubechies wavelet has been observed to be very effective in the recognition and expulsion of drive clamor from discourse or sound [7, 8]. In spite of the fact that such a wavelet might be exceptionally successful in one application, it may not be very as viable in another where the properties of the motivation commotion and the hidden flag are unique.

Subsequently, to empower the fashioner select the suitable wavelet for a given application, an association between certain wavelet highlights and drive recognition execution was made in our current work [9]. In that work, we indicated how the wavelet motivation recognition highlights are subject to the qualities of the drive commotion and the underlying signal, and gave a methodology to choosing the most appropriate wavelet from an arrangement of pre-outlined wavelets. The strategy, be that as it may, has one downside: the nature of the chose wavelet is subject to the nature of the wavelets inside the set. On the off chance that none of the wavelets inside the set are ideal for the given application, the strategy won't be successful. In this paper, we try to evacuate the downside in our past work [9] by outlining wavelets that are most proper for a given application. Using the connections between wavelet highlights and motivation recognition execution [9], we planned an enhancement issue for outlining a wavelet of certain help measure that is tailored for recognizing driving forces for a given application.

The recipetions are encircled as a raised advancement issue where the solution got relates to the FIR channel coefficients of an orthogonal wavelet. The consequent execution correlation comes about with other understood wavelets demonstrate that the wavelets planned utilizing the proposed technique have much better motivation recognition features. The paper is sorted out as takes after. Segment 2 abridges the wavelet properties that are essential for drive identification and demonstrates their reliance on the idea of the motivation commotion and the under-lying discourse flag. In Section 3, we create plans to get the channel coefficients of the ideal wavelet for a given help measure. At that point in Section 4, reproduction tests are displayed to com-pare the drive recognition execution of wavelets inferred utilizing the proposed strategy with other surely understood wavelets. Conclusions are attracted Section 5.

2. DETECTION OF IMPULSE NOISE FROM SPEECH

In this section, we summarize the wavelet properties that influence the detection performance and describe a measure for evaluating the detection performance.

2.1. Wavelet properties and features for impulse detection

A desirable wavelet for impulse detection is one that maximizes the coefficients for the impulse relative to the underlying signal in the finest scale [9]. Such a wavelet will correspondingly have a high pass analysis filter that maximizes the impulse noise relative to the under-lying speech and background noise signals. If $P_s(\omega)$ and $P_i(\omega)$ are the power spectrums of the average speech and impulse noise power, respectively, then the ratio between the average impulse noise power and speech power in the finest scale, R_i, is dependent on the wavelet high pass analysis filter and given by

$$R_i = \frac{\sigma_i^2}{\sigma_s^2} \qquad (1)$$

where

$$\sigma_i^2 = \int_{-\pi}^{\pi} |G(e^{j\omega})|^2 P_i(\omega) d\omega \approx \sum_i |G(e^{j\omega_i})|^2 P_i(\omega_i) \qquad (2)$$

$$\sigma_s^2 = \int_{-\pi}^{\pi} |G(e^{j\omega})|^2 P_s(\omega) d\omega \approx \sum_i |G(e^{j\omega_i})|^2 P_s(\omega_i) \qquad (3)$$

and G(z) is the transfer function of the wavelet highpass filter. The design of an optimal wavelet for detecting the impulses should, there-fore, seek to maximize R_i. The other factor that influences the detection performance is the size of the wavelet support, which is dependent on the average width and energy of the impulse noise [9]. One way to determine the cor-rect wavelet support for a given application is to design wavelets that maximize R_i at various wavelet support sizes and then select the one with the best detection performance.

2.2. Metrics to evaluate the detection performance

To determine the most appropriate wavelet for impulse detection, we evaluate the discriminatory capability of the wavelet coefficients in the finest scale, with respect to the impulse noise. This is done by using a stability criterion derived from the scatter matrices [9]. For a one-dimensional, two-class scenario, the separability criterion for feature x is given by

$$J = \frac{n_1(m_1 - m)^2 + n_2(m_2 - m)^2}{\sum_{x\in\omega_1}(x - m_1)^2 + \sum_{x\in\omega_2}(x - m_2)^2} \quad (4)$$

where (m_1, n_1) and (m_2, n_2) are the means and number of feature samples for classes ω_1 and ω_2, respectively. It has been shown [9] that a wavelet with a higher value of J will correspondingly have better detection performance.

3. DERIVING THE OPTIMAL AVELETS FOR IMPULSE DETECTION

The optimal wavelets are designed to maximize the ratio of impulse noise power to speech power in the finest scale. At the same time, the necessary constraints required for an orthogonal wavelet need to be imposed.

If H(z) corresponds to the transfer function of a low pass analysis filter of an orthogonal wavelet given by

$$H(z) = h(0) + h(1)z^{-1} + \cdots + h(L-1)z^{-(L-1)} \quad (5)$$

Then the high pass counterpart, G(z), can be obtained by taking the alternating flip of H(z) [10]; that is

$$G(z) = -z^{-(L-1)}H(-z^{-1}) \quad (6)$$

where L is assumed to be even. To ensure that the wavelet filter bank is orthogonal, the filter coefficients need to satisfy the double shift orthogonality condition [10], given by

$$\sum_n h(n)h(n-2k) = \delta(k), \quad \text{for } k = 0, 1, \ldots, (L/2) - 1 \quad (7)$$

where δ(k)is the delta function. For the existence of the waveletψ(t), the following condition must also hold true [11]:

$$H(e^{j\omega})|_{\omega=0} = \sum_n h(n) = \sqrt{2} \quad (8)$$

3. DERIVING THE OPTIMAL WAVELETS FOR IMPULSE DETECTION

The optimal wavelets are designed to maximize the ratio of impulse noise power to speech power in the finest scale. At the same time, the necessary constraints required for an orthogonal wavelet need to be imposed.

If H(z) corresponds to the transfer function of a low pass analysis filter of an orthogonal wavelet given by

$$H(z) = h(0) + h(1)z^{-1} + \cdots + h(L-1)z^{-(L-1)} \quad (5)$$

then the high pass counterpart, G(z), can be obtained by taking the alternating flip of H(z) [10]; that is

$$G(z) = -z^{-(L-1)}H(-z^{-1}) \quad (6)$$

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

where L is assumed to be even. To ensure that the wavelet filterbank is orthogonal, the filter coefficients need to satisfy the *double-shift orthogonality* condition [10], given by

$$\sum_n h(n)h(n-2k) = \delta(k), \ \text{for } k = 0, 1, \ldots, (L/2)-1 \quad (7)$$

where $\delta(k)$ is the delta function. For the existence of the wavelet $\psi(t)$, the following condition must also hold true [11]:

$$H(e^{j\omega})|_{\omega=0} = \sum_n h(n) = \sqrt{2} \quad (8)$$

As in the design of signal-adapted filterbanks by Moulin et al[12], the formulation of the optimization problem becomes more tractable if we use the autocorrelation sequence of the filter coefficients given by

$$r_h(l) = \begin{cases} \sum_{n=0}^{L-l-1} h(n)h(n+l) & l \geq 0 \\ r_h(-l) & l < 0 \end{cases} \quad (9)$$

Therefore, in terms of the aurocorrelation parameters, the double shift orthogonality condition in (7) can be expressed as

$$r_h(2k) = \delta(k), \ \text{for } k = 0, 1, \ldots, \left\lfloor \frac{L-1}{2} \right\rfloor \quad (10)$$

and the necessary condition in (8) as

$$\sum_{m=1}^{L-1} r_h(m) = 0.5 \quad (11)$$

by exploiting the orthogonality condition in (7) and the symmetry property in (9). Correspondingly, using (6) and (9) in (2) and (3) the average power of the impulse noise and speech in the finest scale are given by

$$\begin{aligned} \sigma_i^2 &\approx \sum_n \left[r_h(0) + 2\sum_{l=1}^{L-1} (-1)^l r_h(l) \cos(\omega_n l) \right] P_i(\omega_n) \\ &= \mathbf{1}^T \mathbf{C}_i \mathbf{A} \mathbf{r} \quad (12) \\ \sigma_s^2 &\approx \sum_n \left[r_h(0) + 2\sum_{l=1}^{L-1} (-1)^l r_h(l) \cos(\omega_n l) \right] P_s(\omega_n) \\ &= \mathbf{1}^T \mathbf{C}_s \mathbf{A} \mathbf{r} \quad (13) \end{aligned}$$

where

$$\begin{aligned} \mathbf{r} &= [r_h(0) \cdots r_h(L-1)]^T \quad (14) \\ \mathbf{A} &= \begin{bmatrix} a_{00} & \cdots & a_{0(L-1)} \\ \vdots & \vdots & \vdots \\ a_{(N-1)0} & \cdots & a_{(N-1)(L-1)} \end{bmatrix} \quad (15) \\ \mathbf{C}_i &= \text{diag}(c_0^{(i)}, \ldots, c_{(N-1)}^{(i)}) \quad (16) \\ \mathbf{C}_s &= \text{diag}(c_0^{(s)}, \ldots, c_{(N-1)}^{(s)}) \quad (17) \\ a_{nl} &= 2(-1)^l \cos(\omega_n l) \quad (18) \\ c_n^{(i)} &= P_i(\omega_n), \ \omega_n \in [-\pi, \pi] \quad (19) \\ c_n^{(s)} &= P_s(\omega_n), \ \omega_n \in [-\pi, \pi] \quad (20) \end{aligned}$$

and N is the number of samples. The optimization is formulated as the minimization of σ_s^2 while keeping σ_i^2 constant so that R_i in

4. EXPERIMENTAL RESULTS

In this segment we perform analyses to think about the drive detection execution of wavelets planned utilizing the proposed technique with other surely understood wavelets. To produce the motivation clamor signals for completing the experiments we utilize a drive commotion age display [14] that has been observed to be a decent portrayal for discourse signals corrupted by clicks. The model, recreated in Fig. 1, utilizes two commotion age forms. The first is a paired clamor age process, i(n), that controls a switch. The switch is associated when i(n) = 1, in this way

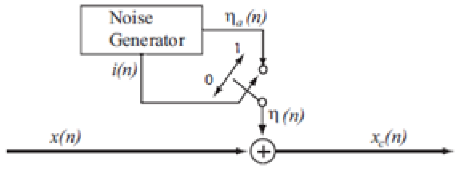


Fig. 1. Impulse noise generation model.

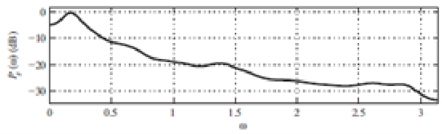


Fig. 2. Normalized average power spectrum of speech. The sam-pling frequency is 16 kHz.

Enabling a second noise process, $\eta_a(n)$ to be added to the speech signal x(n). As can be seen, the noise produced by such a system occurs in bursts, where its value is precisely zero for at least some of the time. A typical audio signal degraded with impulse noise can have an average impulse width of around 1 ms while the fraction of the signal that is contaminated is usually less than 20 percent [11]. If α is the fraction of signal samples contaminated by impulse noise the average signal to impulse noise ratio is given by [16]

$$SINR = \frac{P_s}{\alpha P_i} \quad (23)$$

where P_s is the power of the speech signal and P_i is the power of the impulse. For our experiments, we set the contamination level to 5 percent, which is a typical level for audio degraded by impulse noise [12]. The binary noise generation process for i(n) is implemented using a two-state Markov chain where the transition probabilities can be appropriately adjusted to have the desired average impulse width and contamination level. The second noise process, $\eta_a(n)$, is generated using a normal distribution. To evaluate the detection performance of the wavelets, we com-pare the discriminatory capability of the impulse-detection features of the wavelet by using the separability criterion J in (4). To compute J, the detection features need to be first classified into either class ω_1 or class ω_2: Class ω_1 if the features correspond to an impulse, and ω_2 otherwise.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

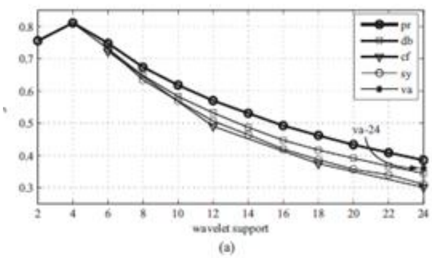
After the features have been classified, we then use (4) to obtain J. The signal from the first level, which corresponds to the finest scale, is the one that is used to detect the impulses. To carry out the classification of the detection features in ω_1 and ω_2, the discrete wavelet transform of the clean speech signal and the impulse noise are taken separately. If $x_f^{(s)}(n)$and $x_f^{(i)}(n)$ are the wavelet coefficients of the clean speech and impulse noise in the finest scale, respectively, the classification of the features in the two classes is given by

$$\mathcal{F}(n) \in \begin{cases} \omega_1 & \text{if } |x_f^{(i)}(n)| > 0 \\ \omega_2 & \text{otherwise} \end{cases} \tag{24}$$

where

$$\mathcal{F}(n) = |x_f^{(s)}(n) + x_f^{(i)}(n)| \tag{25}$$

The speech signal used in the experiments is clean near-microphone speech taken from the ATIS corpus database [13], with a sampling frequency of 16 kHz. The total duration of the signal used for computing J is about 5 minutes long with a total of 3 male and 3 female speakers. In Fig. 2, the average power spectrum of the speech signal, $P_s(\omega)$, is shown. The optimal wavelet filter coefficients are designed as in Section III by solving the optimization problem in (22) to obtain the optimal autocorrelation values and then performing spectral factorization with appropriate scaling to derive the wavelet low pass filter coefficients. For the optimization, we use the speech power spectrum shown in Fig. 2 to compute C_s in (17). Since the generated impulse noise has an average spectrum that is flat we normalize.



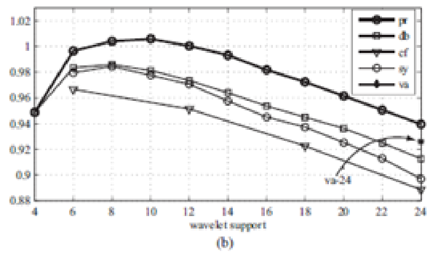


Fig. 3. Comparison plots of J versus support size when the SINR is 10 dB for the cases when (a) the average impulse width = 1 ms (b)The average impulse width = 15 ms. Note that the 'va' wavelet is only a single point with a support size of 24.

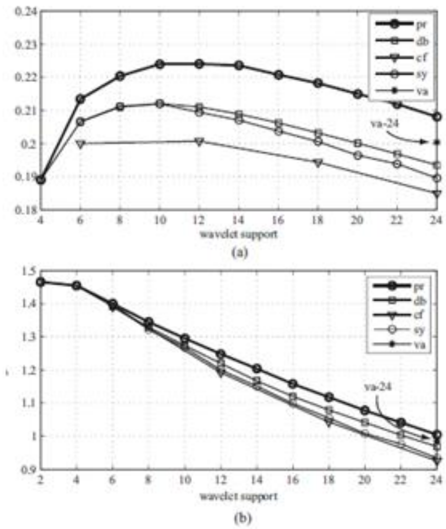


Fig. 4. Comparison plots of J versus support size when the average impulse width is 5 ms for the cases when (a) the SINR is 20 dB (b) the SINR is 0 dB. Note that the 'va' wavelet is only a single point with a support size of 24.

$P_i(\omega)$ = 1and, as a result,$C_i$in (16) simplifies to an identity ma-trix. For our experiments, twelve wavelets ranging from orders 2 to 24 were designed and their corresponding low-pass filter coefficients have been made available online [15]. In the figures, the wavelets designed using the proposed approach are denoted as 'pr'.

For the comparison, we consider various wavelets taken from either the WAVELAB toolbox [19, 20] or the MATLAB Wavelet Toolbox: Daubechies ('db') orders 2-24, Coiflet ('cf') orders 6-24, Symmlet ('sy') orders 6-24, and Vaidyanathan ('va') order 24. Two experiments are carried out to compare the wavelet impulse-detection performance. In the first experiment, we compare the detection performance using impulse noise with two different average widths while keeping the SINR constant. In the second experiment, we compare the detection performance for impulse noises with different SINR levels but having the same average widths.

4.1. Experiment 1:

In this trial, we consider two drive commotions that have the same SINR however extraordinary normal widths and utilize them to analyze the location execution of the wavelets for various help sizes. The main drive commotion has a normal motivation width of 1 ms while the second has a width of 15 ms. The SINR is set to 10 dB in the two cases. In Figs. 3(a) and (b), the detachability parameter, J, is com-pared for various wavelet bolster sizes. As can be seen from the figures, the execution of wavelets planned utilizing the proposed technique is equivalent to or superior to the majority of the contending wavelets. We additionally watch that this execution change has a tendency to improve with respect to alternate wavelets as the help estimate builds; this is on the grounds that the expansion in wavelet bolster relates to an expansion in the quantity of wavelet channel coefficients, in this way permitting more degrees of flexibility in the streamlining. Moreover, looking at the plots between Figs. 3(a) and (b) we watch that the ideal wavelet bolster measure is bigger for the drive commotion that has bigger normal motivation width. This is as per the conclusions attracted our past work [9].

4.2. Experiment 2:

In this trial, we consider two drive clamors that have a similar motivation width however extraordinary SINRs and utilize them to think about the identification execution of the wavelets with various help sizes. The main drive commotion has a SINR of 0 dB while the second has a SINR of 20 dB. The normal motivation width is set to 5 ms in the two cases. In Figs. 4(a) and (b), bends of the distinguishableness parameter, J, versus the wavelet bolster measure are plotted for the different wavelets. As can be seen, the bend comparing to the wavelets planned utilizing the proposed strategy demonstrate the most astounding distinctness at all of the wavelet bolster sizes. Furthermore, as in Experiment 1, the change over the contending wavelets has a tendency to show signs of improvement as the help estimate increments. Contrasting the plots between Figs. 4(a) and (b) we ob-serve that the ideal wavelet bolster estimate is bigger for the drive commotion with bigger SINR, as per the outcomes in our previous work [9].

5. CONCLUSION:

Another strategy for outlining orthogonal wavelets that are streamlined for recognizing motivation commotion in discourse has been depicted. In the strategy, the attributes of the motivation commotion and the basic discourse flag are considered and an arched enhancement problem was planned for determining the ideal wavelet for a given sup-port size. Execution examination with other surely understood wavelets demonstrated that the wavelets planned utilizing the proposed technique have prevalent drive identification properties.

REFFERENCES:

[1] P. Sprechmann, A. Bronstein, J.-M. Morel, and G. Sapiro, "Audio restoration from multiple copies," Proceedings of ICASSP 2013, pp. 878-882.

[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).

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[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 2005, vol. 1, pp. 1093-1096.

[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 2007, vol. 4, pp. 5-8.

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

[7] 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).

[8] R. C. Nongpiur, “Impulse noise removal in speech using wavelets,” Proceedings of ICASSP 2008, pp. 1593-1597.

[9] 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).

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

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

[12] P. Moulin and M. K. Mihcak, “Theory and design of signaladapted FIR paraunitary filter banks”, IEEE Trans. Signal Processing, 46(4), 920-929 (1998).

[13] A. H. Sayed and T. Kailath, “A survey of spectral factorization methods”, Numer. Linear Algebra Appl., 8, 467-496 (2001).

[14] S. J. GodSill and P. J. W. Rayner, “A Bayesian approach to the restoration of degraded audio signals”, IEEE Trans. Speech AudioProcess., 3, 267-278 (1995).