Volume 5, Issue 6, June 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Voice Recognition using Different Filtering Methods

Gaurav R. Patel*
P.G. Student, E&TC Department
P.S.G.V.P.Mandals, D.N.Patel College of Engineering, Shahada, India

Prof. N. C. Patil
E&TC Department
P.S.G.V.P.Mandals, D.N.Patel College of Engineering, Shahada, India

Abstract— Digital signal processing (DSP) has been a major field in the current advancements which includes noise filtering, system identification, and voice prediction. Some of the standard DSP techniques are not enough to solve those problems vastly and obtain desirable results. To promote accurate solutions, Adaptive filtering techniques must be implemented and it is timely convergence to that solution. Speech is most primary human communication for that reason there exists a big trend to increase & improve telecommunication. However, the background noise is an important handicap. If it is joined with other distortion, it can seriously damage the service quality. Two criteria are often used to measure performance: quality and intelligibility. It is very hard to satisfy both at time. Speech enhancement is an area of speech processing used to improve the intelligibility and pleasantness of a speech signal. In speech enhancement the noise removal is the most common approach where estimation of noise characteristics can cancel noise components and retain only the clean speech signal.

Keywords— Adaptive filtering, Digital signal processing (DSP), LMS, NLMS.

I. INTRODUCTION

This continuously environment discourse signs are ruined by a few types of clamor, for example, for example, contending speakers, foundation commotion, auto commotion, furthermore they are liable to bending created by correspondence channels; cases are room resonation, low-quality receivers, and so on. In every single such circumstance extraction of high determination signs is a key undertaking. In this perspective separating come into the photo. Fundamentally sifting systems are comprehensively named non-versatile and versatile separating procedures. In down to earth cases the measurable way of all discourse signs is non-stationary; thus non-versatile sifting may not be suitable. Discourse improvement enhances the sign quality by concealment of commotion and decrease of mutilation. Discourse improvement has numerous applications; for instance, versatile interchanges, powerful discourse acknowledgment, low-quality sound gadgets, and portable amplifiers.

Numerous methodologies have been accounted for in the writing to address discourse improvement. Lately, versatile separating has turn into one of the compelling and famous methodologies for the discourse upgrade. Versatile channels license to identify time shifting possibilities and to track the dynamic varieties of the signs. Furthermore, they change their conduct as indicated by the info signal. In this way, they can recognize shape varieties in the group and accordingly they can get a superior sign estimation. The principal versatile commotion wiping out framework at Stanford University was outlined and constructed in 1965 by two understudies. Their work was embraced as a feature of a research paper venture for a course in versatile frameworks given by the Electrical Engineering Department. Since 1965, versatile clamor crossing out has been effectively connected to various applications. A few strategies have been accounted for so far in the writing to upgrade the execution of discourse handling frameworks; probably the most essential ones are: Wiener separating, LMS sifting [1], otherworldly subtraction [2]-[3], thresholding [4]-[5]. On the other side, LMS-based versatile channels have been broadly utilized for discourse improvement [6]–[8]. In a late study, notwithstanding, a relentless state joining investigation for the LMS calculation with deterministic reference inputs demonstrated that the consistent state weight vector is one-sided, and in this manner, the versatile assessment does not approach the Wiener arrangement. To handle this disadvantage another methodology was considered for evaluating the coefficients of the straight extension, to be specific, the piece LMS (BLMS) calculation [9], in which the coefficient vector is upgraded just once every event taking into account a square inclination estimation. A noteworthy point of interest of the square, or the change area LMS calculation is that the data signs are more or less uncorrelated. Discourse coding in voice keeping money, another system for voicing identification and pitch estimation. This system is in light of the unearthly examination of the discourse multi-scale item.

Practically speaking, LMS is supplanted with its Normalized variant, NLMS. In viable uses of LMS separating, a key parameter is the stride size. On the off chance that the stride size is expansive, the meeting rate of the LMS calculation will be fast, yet the relentless state mean square mistake (MSE) will increment. Then again, if the stride size is little, the consistent state MSE will be little, however the merging rate will be moderate. Therefore, the stride size gives a tradeoff between the joining rate and the enduring state MSE of the LMS calculation. The execution of the LMS calculation may be enhanced by making the stride size variable instead of settled. [10] Rahman et.al displayed discourse sifting utilizing variable step measure minimum mean fourth based treatment and fair and standardized versatile separating systems. The vicinity of foundation clamor in discourse altogether diminishes the understandability of discourse. Clamor lessening or discourse upgrade calculations are utilized to stifle such foundation commotion and enhance the perceptual quality and comprehensibility of discourse. Uprooting different sorts of commotion is troublesome because of the arbitrary way of the clamor and the inalienable complexities of the discourse. Clamor decrease systems typically have an exchange off between the measure of commotion evacuation and discourse contortions acquainted due with preparing of the discourse signal. A few strategies have been proposed for this reason in the territory of discourse improvement, as otherworldly subtraction approach, wiener channel and kalman channel. The exhibitions of these procedures rely on upon the quality and clarity of the handled discourse signal. The change in the discourse sign to clamor proportion is the objective of most procedure.

II. LITERATURE SURVEY

The B. Widrow, J. R. Glover, J. M. McCool [1] clarified idea of discourse upgrade in a hypothetical methodology, utilizing diverse discourse improvement calculations. The discourse upgrade systems went for smothering the foundation clamor are in view of restricted or the other on the estimation of the foundation commotion. In the event that the clamor is developing more gradually than the discourse, it is anything but difficult to gauge the commotion amid the stops in the discourse. In the event that the clamor is changing quickly then estimation is more troublesome. This paper addresses the issue of decrease of added substance foundation clamor in discourse.

Radhika Chinaboina, D.S. Ramkiran [7] considered the versatile separating constitutes one of the center advances in advanced sign handling and finds various application regions in science and additionally in industry. The sign obstruction brought about by acoustic reverberation is diverting to clients and reasons a decrease in the nature of the correspondence. This paper concentrates on the utilization of LMS and NLMS calculations to lessen this undesirable reverberation, consequently expanding correspondence quality.

Sambur M. ITT, Nutley N. J. [2] portrayed that a novel compelled steadiness minimum mean-squares calculation for separating discourse sounds is proposed in the versatile commotion dropping issue. It is in view of the minimization of the squared Euclidean standard of the weight vector change under a strength imperative over the a posteriori estimation blunders. To this reason, the Lagrangian system has been utilized as a part of request to propose a nonlinear adjustment as far as the result of differential info and mistake.

Longbiao WANG, Norihide KITAOKA [5] considered a visually impaired dereverberation strategy in view of unearthly subtraction utilizing a multi-channel slightest mean squares (MCLMS) calculation for removed talking discourse acknowledgment. In a removed talking environment, the channel drive reaction is longer than the transient otherworldly examination window. By regarding the late resonation as added substance clamor, a commotion decrease method in view of unearthly subtraction was proposed to gauge the force range of the clean discourse utilizing force spectra of the misshaped discourse and the obscure reactions.

Sayed A. Hadei, M. lotfizad [4] examined in numerous utilization of clamor wiping out, the adjustments in sign qualities could be quick. This requires the use of versatile calculations, which join quickly. Slightest Mean Squares (LMS) and Normalized Least Mean Squares (NLMS) versatile channels have been utilized as a part of an extensive variety of sign handling application in view of its straightforwardness in processing and execution. The Recursive Least Squares (RLS) calculation has built up itself as a definitive versatile separating calculation as in it is the versatile channel showing the best meeting conduct

III. PROBLEM DEFINATION

A. Weiner Filtering

Title and The channel has its source in a Kalman's report (1960) where it is portrayed as a recursive answer for the direct sifting issue for discrete information. The exploration was in a wide connection of state-space models, where the fact is the estimation through the recursive minimum squares. Since that minute, because of the improvement of advanced count, Kalman channel has been explored and connected, especially in self and helped route, rockets inquiry and economy. The investigation of Kalman channel is in view of Wiener channel. The objective of the Wiener channel is to sift through commotion that has adulterated a sign. It is in light of a factual methodology. Regular channels are intended for a fancied recurrence reaction. On the other hand, the configuration of wiener channel takes an alternate methodology. One is accepted to have learning of the ghostly properties of the first flag and the commotion, and one looks for the direct time invariant channel whose yield would verge on the first flag as could be allowed. Weiner channels are described by the accompanying:

- Requirement: the channel must be physically feasible/causal (this prerequisite can be dropped bringing about a non-causal arrangement).
- Performance rule: least mean square lapse (MMSE).

B. Kalman Filter

The Kalman channel utilizes a framework's dynamic model (i.e., physical laws of movement), known control inputs to that framework, and estimations, (for example, from sensors) to shape an evaluation of the framework's differing amounts that is superior to the assessment acquired by utilizing any one estimation alone. All things considered, it is a typical sensor combination calculation. The Kalman channel midpoints a forecast of a framework's state with another estimation utilizing a weighted normal. The motivation behind the weights is that values with better assessed instability are "trusted" more. The weights are figured from the covariance, a measure of the assessed vulnerability of the forecast of the framework's state. The consequence of the weighted normal is another state assess that lies in the middle of the anticipated and measured state, and has a superior evaluated vulnerability than either alone. This procedure is rehashed each time venture, with new gauge and its covariance educating the forecast utilized as a part of the accompanying cycle. This implies that the Kalman channel lives up to expectations recursively and requires just the keep going best figure not the whole history of a framework's state to compute another state. At the point when performing genuine figurings for the channel, the state appraisal and covariance are coded into frameworks to handle the various measurements included in a solitary arrangement of computations. This takes into consideration representation of direct relationship between diverse state variables, for example, position, speed, and quickening in any of the move models or covariance. The utilization of Kalman channel for discourse upgrade was initially presented by Paliwal (1987). This system however is best suitable for lessening of background noise consent to Kalman presumption. In inferring Kalman mathematical statements it is regularly expected that the procedure clamor is uncorrelated and has an ordinary circulation. This suspicion prompts whiteness character of this clamor. It is accepted that discourse sign is stationary amid every casing that is the AR model of discourse continues as before over the portion. Kalman channel is a versatile slightest square mistake channel that gives a productive computational recursive answer for assessing a sign in vicinity of Gaussian clamors. Kalman channel hypothesis is in view of a state-space approach in which a state mathematical statement models the flow of the sign era process and a perception comparison models the loud and bended perception signal.

IV. PROPOSED SOLUTION

A. Basic Adaptive Filter Structure

The Figure 1 shows an adaptive filter with a primary input that is noisy speech signal s1 with additive noise n1. While the reference input is noise n2, which is correlated in some wa with n1. If the filter output is y and the filter error e= (s1+n1)-y, then

$ee2 = (s1 + n1)2 – 2y (s1 + n1) + y2$

$= (n1 – y)^2 + s1^2 + 2\, s1\, n1 – 2y\, s1$…….. (1)

Since the signal and noise are uncorrelated, the mean-squared error (MSE) is

$E[e2]=E[(n1 – y)2]+E[s12]$……………. (2)

Minimizing the MSE results in a filter error output that is the best least-squares estimate of the signal s1. The adaptive filter extracts the signal, or eliminates the noise, by iteratively minimizing the MSE between the primary and the reference inputs. Minimizing the MSE results in a filter error output y that is the best least-squares estimate of the signals.

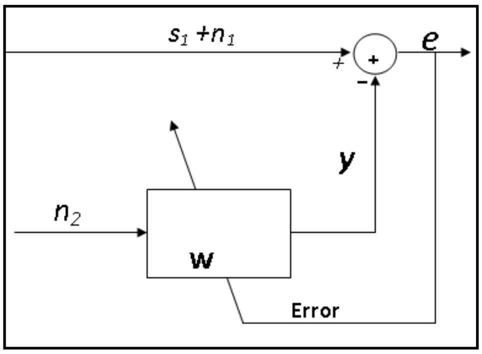

Fig.1. Magnetization Adaptive Filter Structure

B. LMS Algorithm

The LMS algorithm is a method to estimate gradient vector with instantaneous value. It changes the filter tap weights so that e (n) is minimized in the mean-square sense. The conventional LMS algorithm is a stochastic implementation of the steepest descent algorithm. It simply replaces the cost function ξ (n) = E [e2 (n)] by its instantaneous coarse estimate.

The error estimation e(n) is

e (n) = d(n) – w(n) X(n) . . . . (3)

Coefficient updating equation is

w (n+1) = w(n) + μ x(n) e(n), ... (4)

Where μ is an appropriate step size to be chosen as $0 < \mu < 0.2$ for the convergence of the algorithm. The larger step sizes make the coefficients to fluctuate wildly and eventually become unstable.

V. CONCLUSION

In this paper the issue of commotion expulsion from discourse signs utilizing Variable Step Size based versatile separating is exhibited. For this, the same arrangements for speaking to the information and also the channel coefficients as utilized for the LMS calculation were picked. Therefore, the strides identified with the sifting stay unaltered. The proposed treatment, however abuses the alterations in the weight redesign equation for all classifications further bolstering its good fortune and therefore pushes up the rate over the particular LMS-based acknowledge. Our reproductions, on the other hand, affirm that the capacity of MRVSSLMS and RVSSLMS calculations is superior to traditional LMS and Kong's VSSLMS calculations as far as SNR change and merging rate. Consequently this calculation is satisfactory for every reasonable purpose.

A versatile LMS FIR channel approach for discourse upgrade is proposed in this paper. This methodology relies on upon the adjustment of the channel exchange capacity from test to test in view of the discourse signal insights (mean and fluctuation). The outcomes delights that the proposed methodology gives the minimum discourse signal contortion The better execution is because of the ideal choice of the channel request and step size for the calculations a format to get ready paper so as to submit it for gathering.

VI. FUTURE SCOPE

There are numerous conceivable outcomes of extending this venture specifically examination of its usage utilizing whole number-crunching prompting potential outcomes for proficient equipment execution. Further examination of the versatile equalizer execution could likewise be done. Examination of the computational funds and the advantages acquired by utilizing a versatile channel as a part of the recurrence area instead of the time space in a true application where the motivation reaction is generally long would unquestionably be justified regardless of some examination.

In future the individual which driving auto, car vehicle can get the approaching call and make the correspondence of discourse. We likewise utilize the diverse channels with assortment of separating calculations use for discourse upgrade. It is lamented that because of time imperatives full examination in the range of further applications, for example, hands free telephony was not finished, however there was time to research versatile evening out and it was found that the recurrence area versatile channel would in reality be pertinent here if the drive reaction was adequately long. Unfortunately, there was no time for examination of the calculation utilizing whole number execution either.

REFERENCES

[1] B. Widrow, J. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H.Hearn, J. R. Zeidler, E. Dong, and R. Goodlin,“Adaptive noise cancelling: Principles and applications ”, Proc. IEEE, vol. 63, pp.1692-1716, Dec. 1975.

[2] Sambur M. ITT, Nutley N. J, “A parametric formulation of the generalized spectral subtraction method,” IEEE Trans. On Speech and Audio Processing, vol. 6, pp. 328-337, 1998.

[3] I. Y. Soon, S. N. Koh, and C. K. Yeo, “Noisy speech enhancement using discrete cosine transform,” Speech Communication, vol. 24, pp. 249-257, 1998.

[4] Sayed A. Hadei, M. lotfizad, “An improved wavelet-based speech enhancement system,” Proc. of the Eurospeech, 2001.

[5] Longbiao WANG, Norihide KITAOKA “Soft thresholding for DCT speech enhancement,” Electron. Letters, vol. 38,

[6] V. J. Mathews and Z. Xie, “A stochastic gradient adaptive filter with gradient adaptive step-size,” IEEE Transactions on Signal Processing, vol. 41, no. 6, pp. 2075–2087, 1993.

[7] Radhika Chinaboina, D.S. Ramkiran, “A robust variable step-size LMStype algorithm: analysis and simulations,” IEEE Transactions on Signal Processing, vol. 45, no. 3, pp. 631–639, 1997.

[8] G.V.S. Karthik, M.Ajay Kumar, Md.Zia Ur Rahman, “Speech Enhancement Using Gradient Based Variable Step Size Adaptive Filtering Techniques”, International Journal of Computer Science & Emerging Technologies, UK, (E ISSN: 2044-6004), Volume 2, Issue 1, February 2011, pp. 168-177.

[9] Md. Zia Ur Rahman, K.Murali Krishna, G.V.S. Karthik, M. John Joseph and M.Ajay Kumar, “ Non Stationary Noise Cancellation in Speech Signals using an Efficient Variable step size higher order filter”, International Journal of Research and Reviews in Computer Science, UK, Vol. 2, No. 1, 2011.

[10] Md Zia Ur Rahman et al., “Filtering Non Stationary noise in Speech signals using computationally efficient unbiased and Normalized Algorithm”, International Journal on Computer Science anD Engineering, ISSN: 0975-3397 Vol. 3 No. 3 Mar 2011, pp. 1106- 1113.

[11] Francesco Beritelli, Salvatore, Casale, and Salvatore Serrano, “Adaptive Robust Speech Processing Based on Acoustic Noise Estimation and Classification”, IEEE International Symposium on Signal Processing and Information Technology, 2005, p.p. 773-777.

[12] Jocob Benesty, Jingdong Chen, Yiteng Huang, Microphone array signal processing, 2008.

[13] Paul W.Shields and Dougals R. Campbell, “Multi-microphone sub-band adaptive signal processing for improvement of hearing aid performance preliminary results using normal hearing voluntreers”, Proc. of the IEEE International Conference on Acoustics, speech, and Signal Processing, 1997, p.p. 415-418.

[14] L.R.Rabiner/ R.W.Schafer, Digital processing of speech signals.

[15] Patrik Sorqvist, Peter Handel, and Bjorn Ottersten, “Kalman filtering for low distortion speech enhancement in mobile communication”, Proc. of the IE.G. Eason, B. Noble, and I. N. Sneddon, “On certain i ntegrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–5 51, April 1955. *(references)*

