ISSN 2319-8885,Volume01,Issue No. 03

International Journal of Scientific Engineering and Technology Research

www.semargroups.org

Jul-Dec 2012, P.P. 230-235

A New Delay-less Sub-band Adaptive Kalman Filtering Algorithm for Speech Enhancement on Active Noise Control Systems

M.SUDHEER[1], M.L.RAVI CHANDRA[2]

[1]M.Tech Student of NIET, Nalgonda, AP-India, surisudheer@gmail.com,

[2]HOD, ECE Dept, NIET, Nalgonda, AP-India.

Abstract: In this project, The speech enhancement is one of the effective techniques to solve speech degraded by noise. In this paper a fast speech enhancement method for noisy speech signals is presented, which is based on improved Subband adaptive kalman filtering techniques play a prominent role in designing active noise control (ANC) systems. They reduce the computational complexity of ANC algorithms, particularly, when the acoustic noise is a broadband signal and the system models have long impulse responses. In the commonly used uniform-discrete Fourier transform (DFT) -modulated (UDFTM) filter banks, increasing the number of subbands decreases the computational burden but can introduce excessive distortion, degrading performance of the ANC system. The conventional Kalman filter algorithm for speech enhancement needs to calculate the parameters of AR (auto-regressive) model, and perform a lot of matrix operations, which usually is non-adaptive. The speech enhancement algorithm proposed in this paper eliminates the matrix operations and reduces the calculating time by only constantly updating the first value of state vector X(n). We design a coefficient factor for adaptive filtering, to automatically amend the estimation of environmental noise by the observation data. Simulation results show that the fast adaptive algorithm using Kalman filtering is effective for speech enhancement.

Keywords: Speech enhancement, Adaptive kalman filter, Active noise control (ANC), discrete Fourier transform (DFT), filter bank.

I. INTRODUCTION

In the past several years, there were many applications in speech enhancement based on Kalman filtering algorithm. Those methods were proposed by [1]-[7]. Most of those methods need to estimate the parameters of AR model at first, and then perform the noise suppression using Kalman filtering algorithm. In this process, the calculations of LPC (linear prediction coding) coefficient and inverse matrix greatly increase the computational complexity of the filtering algorithm. Although these methods can achieve a good filtering efficiency, the noise suppressed signal may deteriorate the quality of the speech signal dependent on estimation accuracy of the parameters of the AR model. [2] and [3] have been given a simple Kalman filtering algorithm without calculating LPC coefficient in the AR model, but the algorithm still contains a large number of redundant data and matrix inverse operations. In addition, the algorithm is non-adaptive.

To overcome the drawback of conventional Kalman filtering for speech enhancement, we propose a fast adaptive algorithm of Kalman filtering. This algorithm only constantly updates the first value of state vector X(n), which eliminates the matrix operations and reduces the time complexity of the algorithm. Actually, it is difficult to know what environmental noise exactly is. And it affects the application of the Kalman filtering algorithm. So we need a real-time adaptive algorithm to estimate the ambient noise. We add the forgetting factor which has been mentioned by [4] and [5] to amend the estimation of environmental noise by the observation data automatically, so the algorithm can catch the real noise. Simulation results show that, compared with the conventional Kalman filtering algorithm, the fast adaptive algorithm of Kalman filtering is more effective. At the same time, it reduced its' running time without sacrificing quality of the speech signal. It also has good adaptability to improve the algorithm robustness.

Active noise control (ANC) is a method of canceling a noise signal in an acoustic cavity by generating an appropriate anti-noise signal via canceling loudspeakers. Due to recent advances in wireless technology, new applications of ANC have emerged to mitigate the environmental acoustic noise and therefore improve the speech and music quality. ANC as a real-time adaptive signal processing method should meet the some basic requirements like minimum computational complexity, stability and robustness to input noise dynamics, and maximum noise attenuation. Acoustical and electrical signal transmission path models such as those encountered in realistic ANC applications, usually have long impulse responses. Consequently, noise cancellation algorithms require long adaptive filters, resulting in significant computational burden. The computational complexity can be reduced by using frequency-domain filtering techniques based on decomposition, processing, and reconstructing the signals using filter banks such as sub-band adaptive filtering (SAF) and block adaptive filtering (BAF) techniques.

The delayless SAF scheme in an ANC system involves the decomposition of input noise (i.e., the reference signal) and error signals into subbands using analysis filter banks, and combining the sub-band weights into a full-band noise canceling filter by a synthesis filter bank called weight stacking. Typically, a linear-phase finite-impulse response (FIR) low-pass filter (i.e., prototype filter) is designed and modulated for the design of such filter banks. The filter must be designed so that the side-lobe effect and spectral leakage are minimized. The latter requires a high-order FIR filter, introducing a long delay, which increases with M as the bandwidth shrinks. The long delay and side-lobe interference introduced by the prototype filter degrade the performance of SAF algorithms for large M , limiting the computational saving that can be obtained by increasing the number of subbands. Improving the system performance and reducing the computational burden by increasing M has inspired the work presented herein.

The mainstay of the proposed model is improving the system performance and reducing the computational burden. In this paper, we first demonstrate that the increased delay degrades the system performance more than that of the spectral leakage (or side-lobe effects) in a uniform sub-band filtering method. It is shown how the spectral leakage can be reduced by choosing a proper decimation factor and weight stacking methodology. We then present a new SAF (Sub-Band Adaptive Filtering) algorithm that reduces computational complexity by increasing the number of subbands M without degrading the performance of the ANC (Active Noise Control) system. The performance of the proposed method is compared with those of MT (Moragan and Thi) and DFT-MDF(Discrete Fourier Transform and Multi-Delay Adaptive Filter) methods. The results show that the maximum noise attenuation level (NAL) of the proposed method is higher than that of MT and comparable to that of the DFT-MDF method.

II. ACTIVE NOISE CONTROL SYSTEM

Active noise control (ANC) (also known as noise cancellation, active noise reduction (ANR) or antinoise) is a method for reducing unwanted sound. Sound is a pressure wave, which consists of a compression phase and a rarefaction phase. A noise-cancellation speaker emits a sound wave with the same amplitude but with inverted phase (also known as antiphase) to the original sound. The waves combine to form a new wave, in a process called interference, and effectively cancel each other out - an effect which is called phase cancellation. Depending on the circumstances and the method used, the resulting sound wave may be so faint as to be inaudible to human ears.

A noise-cancellation speaker may be co-located with the sound source to be attenuated. In this case it must have the same audio power level as the source of the unwanted sound. Alternatively, the transducer emitting the cancellation signal may be located at the location where sound attenuation is wanted (e.g. the user's ear). This requires a much lower power level for cancellation but is effective only for a single user. Noise cancellation at other locations is more difficult as the three dimensional wavefronts of the unwanted sound and the cancellation signal could match and create alternating zones of constructive and destructive interference. In small enclosed spaces (e.g. the passenger compartment of a car) such global cancellation can be achieved via multiple speakers and feedback microphones, and measurement of the modal responses of the enclosure.

Modern active noise control is achieved through the use of a computer, which analyzes the waveform of the background aural or nonaural noise, then generates a signal reversed waveform to cancel it out by interference. This waveform has identical or directly proportional amplitude to the waveform of the original noise, but its signal is inverted. This creates the destructive interference that reduces the amplitude of the perceived noise.

The active methods (this) differs from passive noise control methods (soundproofing) in that a powered system is involved, rather than unpowered methods such as insulation, sound-absorbing ceiling tiles or muffler. The advantages of active noise control methods compared to passive ones are that they are generally:

- More effective at low frequencies.
- Less bulky.
- Able to block noise selectively.

Applications can be "1-dimensional" or 3-dimensional, depending on the type of zone to protect. Periodic sounds, even complex ones, are easier to cancel than random sounds due to the repetition in the wave form.

Protection of a "1-dimension zone" is easier and requires only one or two microphones and speakers to be effective. Several commercial applications have been successful: noise-cancelling headphones, active mufflers, and the control of noise in air conditioning ducts. The term "1-dimension" refers to a simple pistonic relationship between the noise and the active speaker (mechanical noise reduction) or between the active speaker and the listener (headphones).

Protection of a 3-dimension zone requires many microphones and speakers, making it less cost-effective. Each of the speakers tends to interfere with nearby speakers, reducing the system's overall performance. Noise reduction is more easily achieved with a single listener remaining stationary in a three-dimensional space but if there are multiple listeners or if the single listener moves throughout the space then the noise reduction challenge is made much more difficult. High frequency waves are difficult to reduce in three dimensions due to their relatively short audio wavelength in air.

Sinusoidal noise at approximately 1000 Hz is double the distance of the average person's left ear to the right ear; such a noise coming directly from the front will be easily reduced by an active system but coming from the side will tend to cancel at one ear while being reinforced at the other, making the noise louder, not softer. High frequency sounds above 1000 Hz tend to cancel and reinforce unpredictably from many directions. In sum, the most effective noise reduction in three dimensions involves low frequency sounds. Commercial applications of 3-D noise reduction include the protection of aircraft cabins and car interiors, but in these situations, protection is mainly limited to the cancellation of repetitive (or periodic) noise such as engine-, propeller- or rotor-induced noise.

Antinoise is used to reduce noise at the working environment with ear plugs. Bigger noise cancellation systems are used for ship engines or tunnels. An engine's cyclic nature makes FFT analysis and the noise canceling easier to apply.

The application of active noise reduction produced by engines has various benefits:

1.The operation of the engines is more convenient for personnel.

2.Noise reduction eliminates vibrations that cause material wearout and increased fuel consumption.

3.Quieting of submarines.

In the proposed SAF method, the analysis stage is a UDFTM filter bank. The decimation factor in the filter bank is M/4. The bandwidth of the linear-phase low-pass prototype FIR filter is 2π/M. The prototype filter has the length of introducing a delay of in the filter banks. This delay is less than the delay presented in the methods resulting in the improved performance in our algorithm. This description Using the UDFTM (Uniform DFT Modulated) structure, we propose the following low-pass prototype FIR filter for the filter banks,

$$H_0(z) = 1 + z^{-1} + \ldots + z^{-M+1} \tag{1}$$

The resulting filter bank is the simplest FIR perfect reconstruction filter bank which is made by,

$$H_k(z) = H_0(ze^{-j2\pi k/M}) \tag{2}$$

With a frequency response of the attenuation of its first side-lobe is about 13 dB relative to its main lobe. The first zero-crossing the decimation factor should not exceed M/2 in order to avoid spectral aliasing. For the proposed UDFTM filter banks, h(Z)is defined by

$$\mathbf{h}(z) = \frac{1}{M}\mathbf{F}^{*}\begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-M+1} \end{bmatrix} \tag{3}$$

Where **F** is the DFT matrix of order M. The central frequencies of the bandpass filters are located for 0<k<1. An important advantage of this UDFTM filter bank is that it can be realize using a tapped delay line of length M followed by an inverse FFT block.

The computational complexity of the proposed subband filtering technique is calculated by counting the number of real multiplications required for the following operations.

- Decimation by the factor.
- Calculation of the filter bank outputs. Only Sub bands are considered since is real.
- Weight update operation for calculating adaptive filters.

- Weight stacking and calculation of which includes FFT operations to compute s and an IFFT operation to compute

III. IMPROVED KALMAN FILTERING ALGORITHM

A. Conventional Kalman Filtering Method

Speech driven by white noise is All-pole linear output from the recursive process. Under the short-time stable supposition, a pure speech can establish L step AR model by

$$s(n) = \sum_{i=1}^{L} a_i(n) \times s(n-i) + \omega(n) \tag{4}$$

In (4), $a_i(n)$ is the LPC coefficient, $\omega(n)$ is the white Gaussian noise which the mean is zero and the variance is δ_u^2. In the real environment, the speech signal s(n) is degraded by an additive observation noise v(n) $f|f$nIts mean is zero, and its variance is δ_v^2. This noise isn¡¦t related to s(n). A noisy speech signal y(n) is given by

$$y(n) = s(n) + v(n) \tag{5}$$

In this paper, it is assumed that the variance δ_v^2 is known, but in practice we need to estimate it by the "silent segment" included in the y(n) . from (4) and (5) can be expressed as the state equation and the observation equation which are given by [State equation]

$$x(n) = F(n)x(n-1) + G\omega(n) \tag{6}$$

[Observation equation]

$$y(n) = Hx(n) + v(n) \tag{7}$$

F(n) is the L x L transition matrix expressed as A Fast Adaptive Kalman Filtering Algorithm for Speech Enhancement

$$F(n) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_L(n) & a_{L-1}(n) & a_{L-2}(n) & \cdots & a_1(n) \end{bmatrix} \tag{8}$$

G is the input vector and H is the observation vector. It is easy to see that the conventional Kalman filtering is using the LPC coefficient to estimate the observations of the speech signal. This part spends half the time of the whole algorithm. In [2] the transition matrix F and the observation matrix H are modified. They has defined as

$$F = H = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{9}$$

It also has defined the L„e1 state vector $X(n) = [s(n) \;\cdots\; s(n-L+1) \;\; s(n-L+2)]$, the L is input vector $Q(n)^T = [s(n) \;\; 0 \;\cdots\; 0]$, and the observation vector $R(n) = [1, v(n), \ldots v(n-L+2)]$. Finally, (6) and (7) can be rewritten into the matrix equations by [State equation]

$$X(n) = F \times X(n-1) + Q(n) \tag{10}$$

[Observation equation]

$$Y(n) = H \times X(n) + R(n) \tag{11}$$

This algorithm abrogates the computation of the LPC coefficient. The number of calls for the filtering equations is equal to the number of sampling point¡¦s n of the speech signals,

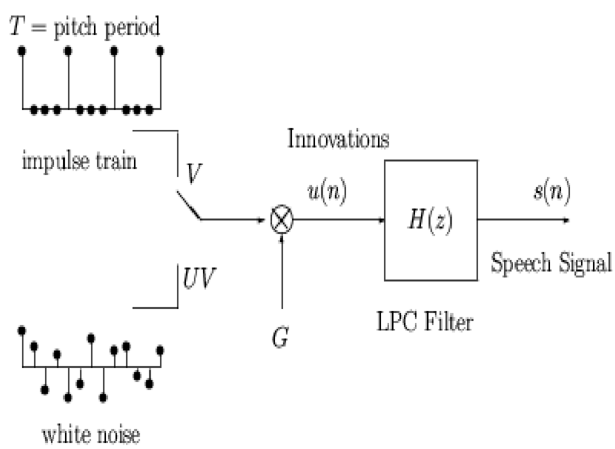

Figure1: LPC Mathematical Model

- The above model is often called the LPC Model.
- The model says that the digital speech signal is the output of a digital filter (called the LPC filter) whose input is either a train of impulses or a white noise sequence.
- The relationship between the physical and the mathematical models:

Vocal Tract	⟺	H(z) (LPC Filter)
Air	⟺	u(n)(Innovations)
Vocal Cord Vibration	⟺	V (voiced)
Vocal Cord Vibration Period	⟺	T (pitch period)
Fricatives and Plosives	⟺	UV(unvoiced)
Air Volume	⟺	G(gain)

- The LPC filter is given by:

$$H(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_{10} z^{-10}} \qquad (12)$$

this is equivalent to saying that the input-output relationship of the filter is given by the linear difference equation:

$$s(n) + \sum_{i=1}^{10} a_i s(n-i) = u(n) \qquad (13)$$

- The LPC model can be represented in vector form as:

$$\mathbf{A} = (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, G, V/UV, T) \qquad (14)$$

- **A** Changes every 20 msec or so. At a sampling rate of 8000 samples/sec, 20 msec is equivalent to 160 samples.
- The digital speech signal is divided into frames of size 20 msec. There are 50 frames/second.

The model says that

$$\mathbf{A} = (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, G, V/UV, T)$$

is equivalent to

$$\mathbf{S} = (s(0), s(1), \ldots, s(159)) \qquad (15)$$

Thus the 160 values of S are compactly represented by the 13 values of A.

- There's almost no perceptual difference in S if:

For Voiced Sounds (V): the impulse train is shifted (insensitive to phase change).

For Unvoiced Sounds (UV): a different white noise sequence is used.

LPC Synthesis: Given A, generate S (this is done using standard filtering techniques).

LPC Analysis: Given S, find the best S (this is described in the next section).

IV. SIMULATION RESULTS

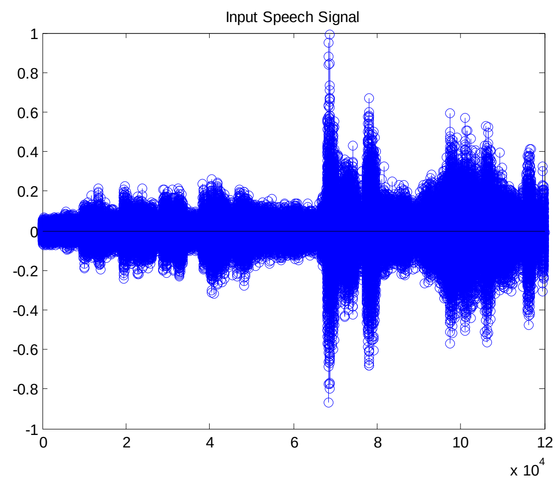

Figure2: Input Speech Signal.

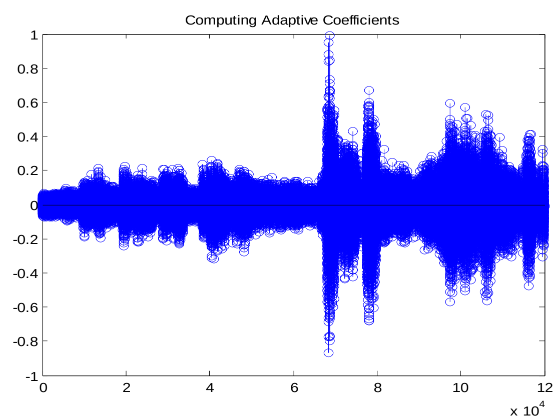

Figure3: Computing Adaptive kalman Filter Speech Co-efficients for input Speech Signal.

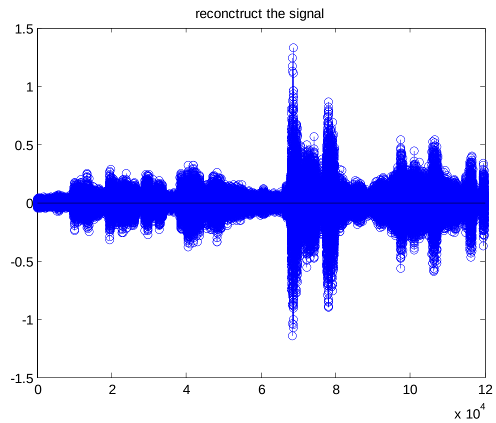

Figure4: Reconstruct output speech signal using Adaptive kalman Filter.

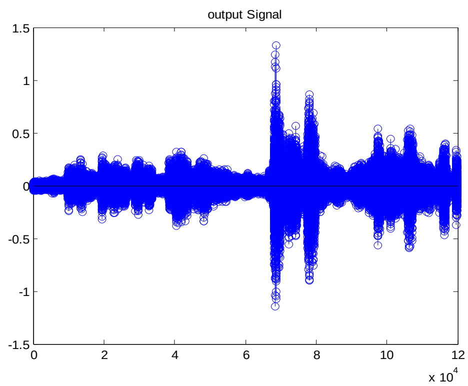

Figure5: Enhanced Speech Signal using Sub-band Adaptive kalman filter.

V. CONCLUSION

This paper has presented a sub-band adaptive Kalman filtering algorithm for speech enhancement by eliminating the matrix operation and designing a coefficient factor. It has been shown by numerical results and subjective evaluation results that the proposed algorithm is fairly effective. Especially, the proposed method contains two-step multiplications in each procedure so that it requires less running time, and the SNRout of this proposed method is higher when the speech signals are degraded by the colored noise. It is concluded that this proposed algorithm is simpler and can realize the good noise suppression despite the reduction of the computational complexity without sacrificing the quality of the speech signal.

VI. REFERENCES

[1] D. R. Morgan and J. C. Thi, “A delayless subband adaptive filter architecture,” IEEE Trans. Signal Process., vol. 43, no. 8, pp. 1819–1830,Aug. 1995.

[2] R. Merched and A. H. Sayed, “An embedding approach to frequencydomain and subband adaptive filtering,” IEEE Trans. Signal Process.,vol. 48, no. 9, pp. 2607–2619, Sep. 2000.

[3] P. Loizou, Speech Enhancement: Theory and Practice. Boca Raton,FL: CRC, 2007.

[4] V. R. Ramachandran, G. Kannan, A. A. Milani, and I. M. S. Panahi,“Speech enhancement in functional MRI environment using adaptive sub-band algorithms,” in Proc. ICASSP’07, 2007, pp. II-341–II-344.

[5] P. Loizou and K. Kokkinakis, “Subband-based blind signal processing for source separation in convolutive mixtures of speech,” in Proc. ICASSP’07, 2007, pp. IV-917–IV-920.

[6] N. Hirayamaa, H. Sakai, and S. Miyagi, “Delayless subband adaptive filtering using the hadamard transform,” IEEE Trans. Acoust., Speech, Signal Process., vol. 47, pp. 1731–1734, Jun. 1999.

[7] P. Eneroth, “Joint filterbanks for echo cancellation and audio coding,” IEEE Trans. Speech Audio Process., vol. 11, no. 4, pp. 342–354, Jul. 2003.

[8] D. Zhou and V. DeBrunner, “A new active noise control algorithm that requires no secondary path identification based on the SPR property,” IEEE Trans. Signal Process., vol. 55, no. 5, pp. 1719–1729, May 2007.

[9] J. Shynk, “Frequency-domain and multirate adaptive filtering,” IEEE Signal Process. Mag., vol. 9, no. 1, pp. 14–37, Jan. 1992.

[10] P. Vaidyanathan, Multirate Systems and Filter Banks. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[11] J. Huo, S. Nordholm, and Z. Zang, “New weight transform schemes for delayless subband adaptive filtering,” in Proc. GLOBECOM’01, New York, Nov. 2001, pp. 197–201.

[12] L. Larson, J. de Haan, and I. Claesson, “A new subband weight transform for delayless subband adaptive filtering structures,” in Proc. Digital Signal Process. Workshop, Aug. 2002, pp. 201–206.

[13] J. M. de Haan, N. Grbic, I. Claesson, and S. E. Nordholm, “Filter bank design for subband adaptive microphone arrays,” IEEE Trans. Speech Audio Process., vol. 11, no. 1, pp. 14–23, Jan. 2003.

[14] A. H. Sayed, Fundamentals of Adaptive Filtering. New York: Wiley, 2003.