International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in
Vol.3., Issue.3, 2015
RESEARCH ARTICLE
ISSN: 2321-7758

SPEECH ENHANCEMENT ADAPTIVE FILTER AND MASKING PROPERTIES FOR MOBILE VOICE COMMUNICATION

RAVINDRA D. DHAGE[1], Prof. PRAVINKUMAR R. BADADAPURE[2]
[1]M.E Scholar, [2]M.E Teacher,
Imperial College of Engineering & Research,Pune

Article Received:01/06/2015 | Article Revised on:17/06/2015 | Article Accepted on:20/06/2015

RAVINDRA D. DHAGE

ABSTRACT

In this paper a single channel speech enhancement system based on Adaptive filtering and masking properties of the human auditory system is studied. Many speech enhancement implementations of today are either digital or analog. Digital solutions are often superior in time to market price per unit structured and powerful development tools, flexibility, high degree of reconfiguration, robustness, the ability to use a Digital Signal Processor (DSP) for many tasks and the possibility to handle high complexity algorithms, where the input signal is divided into a number of sub bands that are individually and adaptively weighted in time domain according to a short term SNR estimate in each sub band at every time an enhanced noise reduction method. The input signal is divided into a number of sub bands that are individually weighted in the time domain according to the short time signal to noise ratio estimate (SNR) in each sub band. Instead of focusing on suppression of the noise the method focuses on speech enhancement algorithms. To develop Adaptive filtering speech enhancement system, the noise variance estimate he modified as a function of masking threshold in each sub-band. By using such Adaptive filtering with masking method, we can achieve better results to sub-band Adaptive filtering, especially in the case of weak speech spectral components in noise.

I. INTRODUCTION

In this paper single channel speech enhancement is studied. Despite the availability of a wide variety of theoretical and relatively effective techniques, the problem *of* enhancing noisy speech still poses a challenge to many researchers. There are many publications which reported speech enhancement methods such as spectral subtraction, minimum mean square error (MMSE) method proposed by Ephraim & Malah [I], autoregressive (AR) model-based Adaptive filtering [2,3], auditory perceptual criteria based method [4], hidden Markov model (HMM) and subspace methods.

The objectives of speech enhancement are to either improve one or more perceptual aspects of speech such as overall quality and intelligibility or to reduce the amount of distortions to improve speech recognition rate. In hands-free car mobile communication systems, enhancement of the speech severely contaminated by additive colored noise in a running car can potentially improve the quality of voice communications.

The Adaptive gain equalizer (AGE) is a time domain speech enhancement algorithm in which the speech signal is amplified based on signal-to-noise (SNR) estimates in sub bands. A signal is divided into sub bands for calculation of a gain which is independent for each band. The algorithm has shown advantages over contemporary techniques because of its low complexity implementation no requirement of voice activity detector and has no presence of musical noise [1]. Different types of background noise corrupt the otherwise clean speech signals in everyday communication. A phone call can be disturbed by a variety of noises present nearby ranging from computer fan noise to factory noise.

There are a wide variety of context in which it is desired to enhance speech. The objective of enhancement is usually to improve the overall speech quality to increase intelligibility and to reduce listener's fatigue etc. The specific goal we attempt to attain is to increase output to input SNR gains which is defined as the ratio of the output SNR to the input SNR. A very important application for speech enhancement is in conjunction with speech compression system. Because of the increasing role of digital channels coupled with the need for encrypting of speech and increased emphasis on integrated voice data networks speech compression system based on speech production model is destined to play an increasing important role in speech communication system.

There are quite a number of publications which reported speech enhancement methods using Adaptive filtering. The advantages of Adaptive filtering as compared to spectral domain processing such as spectral subtraction and MMSE spectral suppression are that it could overcome the tonal noise problem, and achieve quite g o d speech quality by reducing the processing distortion introduced to the speech signal. Sub-band Adaptive filtering [2] is also reported for its good computation cost and improved performance as compared to the full hand case.

Applications of the properties of the human auditory system to speech enhancement have been reported in the past especially in spectral domain speech enhancement [4]. The main aim of this method is to find an optimal tradeoff between noise suppression, speech distortion and residual level of the tonal noise. We study how the masking properties could he applied to a time domain Adaptive filtering system involving subband analysis. Experiments indicate that our proposed estimator has better performance in terms of both objective and subjective evaluations as compared to many existing speech enhancement methods.

II. MASKING PROPERTIES INCORPORATING INTO SUBBAND FILTERING

Subband Filter:

A uniform and critical polyphase MDCT filter bank is used as analysis and synthesis filter banks in our system. The number of hands is set to *L.* The down sampling and up sampling ratios are

M : 1 and 1 : *M* respectively, and we set *M* = L.

Suppose the prototype filter of sub-band is *h(n),*

the pth polyphase filter *e(p,* n) is then denoted by,

$e(p,m) = R(mM + p)$, $p = 0,1,.., L - 1$ (1)

The k-th analysis filter *hk(n)* and k-th synthesis filter *gk(n)* can be proved to be given by the following equations,

$$h_k(n) = h_k(mM + P) = 2e(p,m)cos[\left(k - \frac{1}{2}\right)(p-1)\frac{\pi}{M} + \emptyset_k(m) + (-1)^{k-1}\frac{\pi}{4} \quad (2)$$

$$g_k(n) = g_k(mM + P) = 2e(p,m)cos[\left(k - \frac{1}{2}\right)(p-1)\frac{\pi}{M} + \emptyset_k(m) - (-1)^{k-1}\frac{\pi}{4} \quad (3)$$

where *k* , *p* = 0,1, ..., *L* - 1 and *&(m)* = *m7r*(*k* - 4).

III. ADAPTIVE FILTER

Each sub band specific gain function constitutes a quotient of a short term average and a noise floor level estimate. The noise floor level estimate should be set to track slow changes in the background noise and the short term average should track the bursts of speech. The proposed system used for the enhancement of noisy speech signal x (n). A K bands band-pass filter is used to divide the input speech signal x (n) into sub-bands according to:

$$x_k(n) = h_k(n) * x(n) \quad (4)$$

Where $h_k(n)$ impulse response of the k is sub band.

Natural signals such as speech can be represented by the corresponding high frequency and low frequency components. The final enhanced signal is obtained by adding all the modified sub bands according to the synthesis equation:

$$\hat{x}(n) = \sum_{k=1}^{k} \hat{x}_k(n) \quad (5)$$

The observed noisy modulator for sub-band k is given by $S_k(n)$ and where (p) is a short spectral estimation window. The center of gravity approach estimates the $w_k(n)$as the average frequency of instantaneous spectrum of x_k Center of Gravity (CoG) estimation $w_k(n)$ is given by:

$$m_k(n) = \sum_{j=1}^{p} a_k m_k(n-j) + w_k(n) \quad (6)$$

$$x_k(n) = m_k(n) + v_k(n)$$

$$H^T = [0,0 \dots 1]$$

At time instant n estimated sample is given by following relationship:

$$\hat{m}_k(n) = H^T \hat{m}_k(n) \quad (7)$$

B. Adaptive System

The AGE consists of a filter bank and each sub-band is weighted by a gain function which amplifies the signal when speech is present and keeps the noisy part of the signal where no speech is present to unity

$$x_k(n) = h_k(n) * x(n) \quad (8)$$

A filter bank of K band pass filters divides the input signal (n) into K sub-bands [7]. Here hk is the impulse response of the filter bank sub-band k and denotes the convolution. The output signal with the amplified speech signal is computed as

$$\hat{x}(n) = \sum_{k=1}^{k} G_k(n) x_k(n) \quad (9)$$

Where (n) is the AGE weighting function which amplifies the signal when speech is active and is given by

$$G_k(n) = \min\{(\frac{A_k(n)}{L_{opt} B_k(n)})^{p_k}, L_k \quad (10)$$

Where L_{opt} is the optimized suppression level for gain function and pk gain rise exponent constant, L_k is a limiting threshold limiting gain function value, Fast average (n) and slow average $B(n)$ of sub-band k calculated according to:

$$A_k(n) = a_k A_k(n-1) + (1-a_k)|x_k(n)|$$

Where $a_k = \frac{1}{f_s T_a}$forgetting factor constant and f_s is is sampling frequency.

$$B_k(n) = \quad A_k(n) \qquad \text{if } A_k(n-1) \leq B_k(n-1)$$

$$(1+B_k)(B_k(n-1)) \qquad \text{Otherwise}$$

$$\hat{m}_k(n) = m_k(n) G_k$$

Where $B_k = \frac{1}{f_s T_b}$is a positive constant control the noise level based on the above mentioned principle of AGE a speech signal modulator can also be enhanced by the equalizer Modulation domain separates each sub-band signal into a carrier and a modulator. While only modulators are considered here, the AGE is implemented on each modulator to enhance the speech. This system mathematics for AGE in the modulation domain is the same as for AGE in the sub-band domain the long term average and the short term average are calculated for each sub-band modulator instead of the sub-band itself. The gain function is multiplied with the modulator of the sub-band to yield a modified modulator which is then used with the carrier in the reconstruction stage of the modulation system.

IV. PERFORMANCE EVALUATION

The effectiveness of the enhancement algorithms is evaluated at the sampling rate of 8kHz which is down-sampled from 16 kHz utterances from the TIMIT database after pre-filtering. For our proposed algorithm with 8kHz sampling rate, the frame size used for masking (full band domain). It is expected since babble spectrum is not flat for all frequencies contrast to white noise and pitch is present (at SSR = -30 Db to -15 dB) as if there is speech. As a result, babble pitch can be seen initially. It implies that pitch (which is similar to original speech) exist when SNR + SSR _ -10 dB ignoring babble pitch. We may neglect babble pitch as it become smaller and insignificant compared to speech. In other words, speech dominance over noise is fulfilled provided that SNR + SSR _ -10 dB. Thus, babble pitch should be categories as false detection of speech dominance over noise.

By comparing spectrum from Figure 1, the spectrum is less obvious compared to white. With decreasing SSR, the effect is deteriorating since the noise begins to dominate over the speech. Notice that Figure 2 has weak speech recovery at SSR = - 30 dB to -15 dB (with dark red spectrum). Since there is presence of pitch but SNR at that frame is high, the decision maker in turns attempts to retain the original content but not as much as white noise speech. Compared to white noise, babble noise has more dominant effect on the speech signal, which means that it is more difficult to differentiate between speech and noise by listening.

International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in
Vol.3., Issue.3, 2015

Fig3.Compared to exhibits similar pattern where overall SNR is better for proposed algorithm compared to original SS. Improvement of proposed method is slightly worse for babble noise compared to white noise because there are more energy content of babble noise (speech also present) are concentrated at low frequency band whereas white noise is equally distributed across all frequency band.

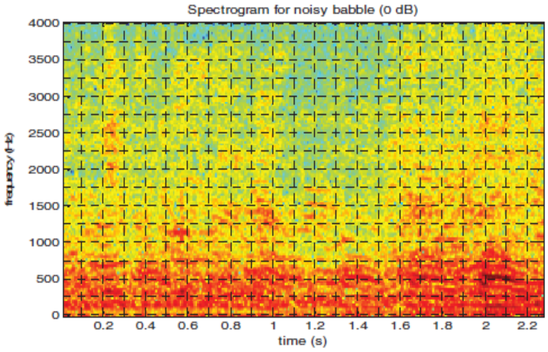

Fig1. Spectrogram Noise Speech

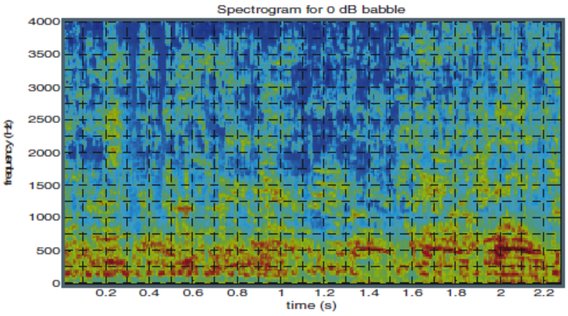

Fig2. Spectrogram Speech Enhancement

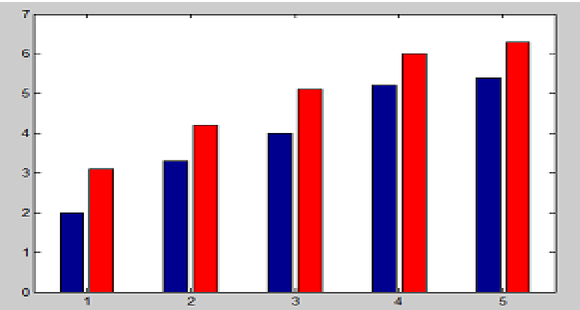

Fig3. Speech Improvement

CONCLUSION

In this paper Adaptive filter and masking properties using noise cancellation when only corrupted speech signal by an additive noise is available for processing. The observation sequence contains the speech component and the unwanted noise. An effective speech enhancement system for car mobile communication has been introduced. It employs Adaptive filtering and exploits the masking properties of the human auditory system. The improvement arising from the use of masking properties is significant. It follows an inherent improvement of Adaptive filtering in the sub-band domain that is characterized by low computational complexity and higher SNR over the full band case. Our system is proven a good performance for estimation of weak speech spectral components.

REFERANCES

[1]. S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE trans. Accout. Speech and Sig. Proc.*, vol. 27, no. 2, pp. 113–120, 1979.

[2]. T. Q. Nguyen, "Near-Perfect-Reconstruction Pseudo- QMF Bank," ***IEEE Trans. Signal Processing,*** vol. 42, pp.65-76, Jan. 1994.

[3]. W.R. **Wu** and P.C. Chen, "Subband Kalman Filtering for Speech Enhancement," ***Ieee Trans.*** *On* ***Circuits And*** *Sys***temsli:** ***Analog And Digital Signal Processing,*** Vol. 45, no. 8, pp. 1072-1083, Aug. 1998.

[4]. N. Virag, "Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System,"***IEEE Trans. Speech and Audio Processing,*** Vol. 7, No.2, pp, 126-137, Mar. 1999.

[5]. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean square error short-time spectral amplitude estimator," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 32, no. 6, pp. 1109 – 1121, dec 1984.

[6]. N. Westerlund, M. Dahl, and I. Claesson, "Speech enhancement for personal communication using an adaptive gain equalizer," *Elsevier Signal Processing.*, vol. 85, pp. 1089–1101, 2005.

[7]. B. S¨allberg, N. Grbic, and I. Claesson, "Implementation aspects of the adaptive gain equalizer," 2006.

[8]. M. Shahid, R. Ishaq, B. S¨allberg, N. Grbic, B. L¨ovstr¨om, and I. Claesson, "Modulation domain adaptive gain equalizer for speech enhancement," in *Signal and Image Processing Application 2011, by IASTED*, 2011.