A Study on the Design of Integrated Speech Enhancement System for Hands-Free Mobile Radiotelephony in a Car

*Kyu-Sik Park and, *Sang-Hun Oh

※ The Authors wish to acknowledge the financial support of the Korea Research Foundation made in the program year of 1997.

Abstract

This paper presents the integrated speech enhancement system for hands-free mobile communication. The proposed integrated system incorporates both acoustic echo cancellation and engine noise reduction device to provide signal enhancement of desired speech signal from the echoed plus noisy environments. To implement the system, a delayless subband adaptive structure is used for acoustic echo cancellation operation. The NLMS based adaptive noise canceller then applied to the residual echo removed noisy signal to achieve the selective engine noise attenuation in dominant frequency component. Two sets of computer simulations are conducted to demonstrate the effectiveness of the system; one for the fixed acoustical environment condition, the other for the robustness of the system in which, more realistic situation, the acoustic transmission environment change. Simulation results confirm the system performance of 20~25dB ERLE in acoustic echo cancellation and 9~19 dB engine noise attenuation in dominant frequency component for both cases.

I. Introduction

The noisy environment in a car is known to severely degrade performance of hands-free mobile communication and the enhancement of the desired speech signal from the surrounding noise is an important problem. Two major impairments in a car communication comes from the background noise levels due to the engine noise, road condition, airflow, etc, and the acoustic echo due to the coupling between loudspeaker and the microphone. These combined echo plus noise disturbances corrupt the speech signal, thereby making the serious annoyance to the far-end listener in a hands-free mobile communication situation. This calls for the method of front-end integrated system that simultaneously take account of the acoustic echo cancellation and the noise attenuation device to provide the sufficient speech quality for the hands-free mobile communication.

Over the last decade, the underlying theory of the acoustic echo cancellation and noise reduction has been extensively studied and widely utilized on various hands-free communication such as mobile telephony in an independent manner. Only a few works consider the problem of simultaneous cancellation of acoustic echo and background noise in the hands-free mobile radiotelephony either by using single microphone or multiple microphone. Pascal, et al[1] showed the analysis of the combined structure for acoustic echo cancellation and noise reduction with a single microphone. They show that the use of noise reduction device before the acoustic echo cancellation causes the nonlinear distortion to the identification process of acoustic echo canceller and thereby, the acoustic echo canceller with affine projection algorithm(APA) is applied before the MMSE based noise reduction device. In this manner, they reported the average ERLE of 11.46dB. On the other hands, the microphone array techniques[2][3][4] takes the advantage of the bearing informations on the array of microphones and it is known to have a little better performance than that of single microphone.

In this paper, our objective is to achieve simultaneous reduction of acoustic echo and the background noise, especially on the engine noise, in the hands-free mobile situation. Our approach is based on the two microphone technique to take the advantage of direct reference

* Division of Computer, Information and Telecommunication, Sangmyung University
Manuscript Received : May 11, 1999

informations on the far-end signal and the engine noise signal. Figure 1 shows the proposed integrate structure that incorporates both adaptive echo cancellation(AEC) and adaptive noise cancellation(ANC) device. A delayless subband adaptive technique[5] is used for acoustic echo cancellation and conventional NLMS based adaptive noise cancellation[6][7] method is then applied for the engine noise attenuation, focusing on the dominant noise frequency component, and thereby achieving the overall cancellation of the primary echoed plus noise input signal.

This paper is organized as follows. Section II briefly describes the delayless subband adaptive structure for acoustic echo cancellation. Section III discuss the engine noise characteristics and the corresponding adaptive noise cancellation method to achieve selective attenuation in dominant frequency is explained. We then investigate the integrated structure of AEC and ANC with some simulation results in section IV. Two sets of computer simulations are conducted depending on the acoustic environmental conditions. Finally, section V concludes this paper.

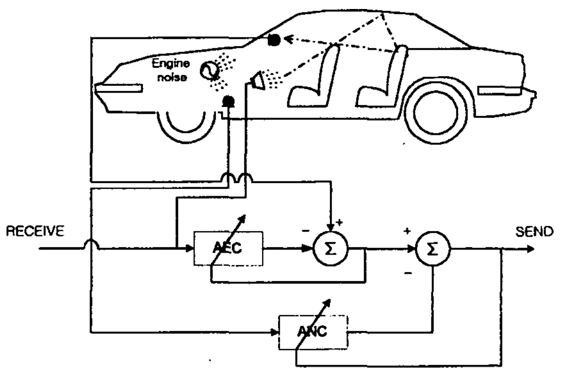


Figure 1. Block diagram of the integrated speech enhancement system.

II. A Delayless Subband Adaptive Echo Cancellation

An acoustic echo in a car arise from the coupling between the loudspeaker and microphone such that a delayed and distorted echo version of received signal is reflected back to the far-end speaker. The basic principle of echo cancellation is to first estimate the characteristics of the echo path, then generates a replica of the echo. This echo is then subtracted from the received signal. Adaptive digital filtering is required to obtain a good echo cancellation performance, since the echo path is usually unknown and time-varying.

Recently, Morgan[5] proposed delayless subband adaptive filter structure that overcomes the inherit nature of signal path delay and aliasing effects presented in subband structure while maintaining the convergence speed and computational complexity advantages. Figure 2 shows the adaptive echo cancellation(AEC) device that implements Morgran's delayless subband adaptive structure. From the figure, we first note that the structure does not include the synthesis filter bank so that no echo delay path is introduced in the overall structure. The far-end reference $x(n)$ and echoed version of signal $d(n)$ p ass through the two sets of FIR analysis filter bank to generate decorrelated complex subband signals, then downsampled by factor of D.

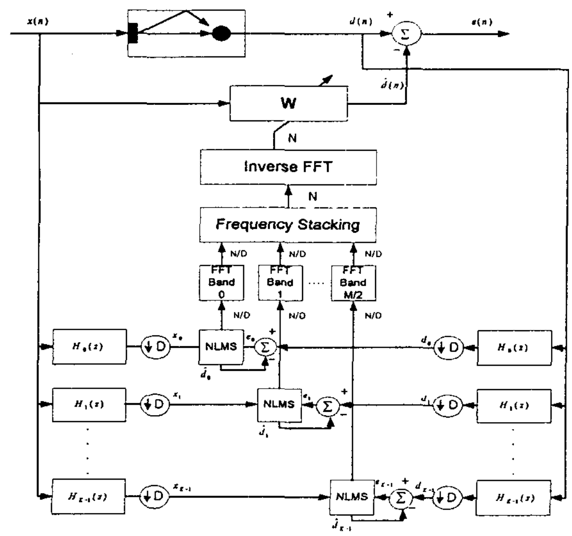


Figure 2. A delayless subband adaptive echo canceller.

For the fast implementation of the analysis subband filter bank, polyphase FFT filter bank can be used. This method is to first design the prototype low-pass filter and then, for the rest of the filter bank, it uses the equally frequency shifted of prototype filter over the entire frequency range. In this way, we can implement the FFT type of the filter bank and save the computation time. Figure 3 shows the basic procedure for generation of the k th subband channel by making use of prototype low pass filter $h(n)$ together with modulating factor W_K^{-nk}. Based on the block diagram of single channel DFT filter bank, the k th subband output can be written as following FFT expression[8]

$$X_k(m) = FFT*\{y_p(m)\}W_K^{-mMk} = FFT*\{y_{p-mM}(m)\} \quad (1)$$

where

$$y_p(m) = \sum_{r=-\infty}^{\infty}\overline{P_p}(r)x_p(m-2r),$$

$$\overline{P_p} = h(mK+p), \qquad x_p = x(mM-p) \quad (2)$$

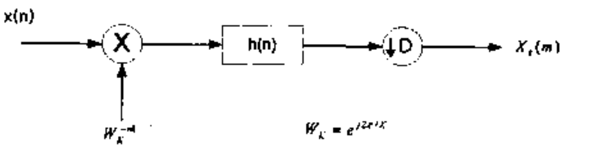


Figure 3. kth subband signal procedure for polyphase FFT operation.

Figure 4 shows the typical design of prototype low pass filter $h(n)$ based on hamming windowed linear-phase FIR design with 128 tap of cutoff frequency f_s /32[9].

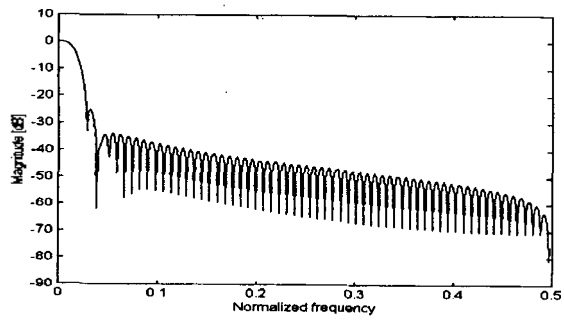


Figure 4. Prototype low pass filter design with 128 tap.

After the analysis subband filter bank, the downsampled subband signal is used for the narrowband adaptive filter weights update. To update the narrowband adaptive filter weights on each subband, we use the following complex NLMS algorithm

$$w_m(n+D) = w_m(n) + \mu x_{m^*} e_m(n)$$

$$e_m(n) = d_m(n) - d_m(n), \quad \mu = \frac{\alpha}{x_m^T(n)\, x_m(n)} \quad (3)$$

where $w_m(n)$, $x_m(n)$, $e_m(n)$ are narrowband adaptive filter weight vector, the subbanded reference signal vector, and the residual error signal on the m th subband respectively. The wideband adaptive filter is then constructed by collectively transforming the narrowband adaptive filter weights computed in each subband through the frequency stacking process. It is noteworthy that the subband structure is only used to update narrowband adaptive weights and the residual error signals are purely computed in the wideband so that the overall structure can avoid the aliasing effects.

III. Car engine noise characteristics and adaptive noise cancellation

The noise in a car mainly consists of the periodic type noise such as engine noise, and the random type of noise due to the road condition, airflow, and tire, etc. These noises corrupt the speech signal, thereby making the transmission of near-end speaker bad in a hands-free mobile telephone system. The previous studies[10][11] has shown that the car noise is dominated by engine noise with the most of noise energy concentrated below 500Hz when the speed of the car is low. Increasing the car speed introduces more random environmental noise in the car and it increases the noise energy, especially at higher frequencies.

Figure 5 shows the magnitude spectra of the engine noise at four different RPM conditions, 3000, 3500, 4000, 4500. They are recorded by a single microphone placed on the floor beneath the driver' s seat.

The engine noise is basically a periodic noise that consists of narrow band components at the fundamental frequency and harmonics which can be expressed as[12]

$$r(n) = \sum_{k=1}^{K} C_k \sin(2\pi k f_0 n) + v(n) \quad (4)$$

where f_0 is the fundamental frequency, K is the total number of harmonics, C_k is the amplitude of the corresponding harmonic, and $v(n)$ is white noise.

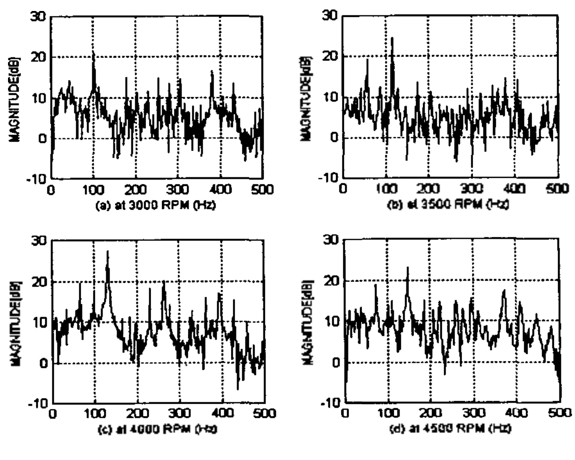


Figure 5. Noise spectra under four different RPM conditions.

From the magnitude spectrum of the car engine noise in figure 5, we see that the highest noise peak resides at the second harmonics of the spectrum(C2) and they are at f_2 = 102, 116, 132, 148 Hz corresponding to 3000, 3500, 4000, 4500 RPM respectively.

In the hands-free mobile communication situation, it is well known that the most annoyance of the car engine noise arise from the highest noise peak component(C_2), therefore noise cancellation method with good attenuation capability at the C_2 harmonic is preferred. For this purpose, conventional NLMS based adaptive noise cancellation(ANC) technique shown in figure 6 is used to reduce C_2 component of noise spectra. In general, the NLMS based ANC is known to have a problem of misadjustment in adaptive process when the speech like colored input is presented in the system. However, in ref [7], they showed that the NLMS adaptive filter can be approximated 2nd order notch filter, thereby it has a good selective cancellation nature in the dominant frequency components.

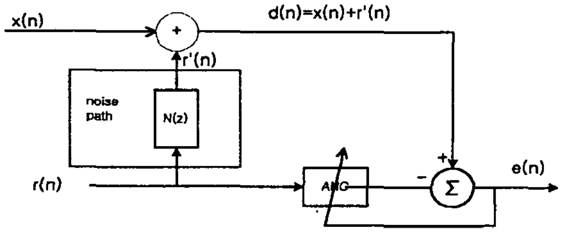


Figure 6. Adaptive noise cancellation structure.

Here, the primary input to the ANC is the desired signal x(n) plus noise r' (n) and the reference signal is the engine noise. The goal of ANC is to have the adaptive filter that models the transfer function N(z) between the engine noise source and the primary microphone so that the noise subtracted at the output cancels the noise in the primary signal.

The FIR adaptive filter weights in ANC are updated by using the NLMS algorithm

$$w(n+1) = w(n) + \mu e(n)\, r(n) \tag{5}$$

where $w(n)$, $r(n)$ are the adaptive weight vector and input noise vector defined as following

$$w(n) = [w(n), w(n-1), \cdots, w(n-N+1)]^T$$
$$r(n) = [r(n), r(n-1), \cdots, r(n-N+1)]^T \tag{6}$$

and $e(n)$ is the error signal between the primary input and the stimate of the adaptive filter. Here μ is the normalized convergence ratio given as

$$\mu = \frac{\alpha}{r^T(n)\, r(n)},\ 0 < \alpha < 1 \tag{7}$$

where the α is the step size for adaptation.

In this study, the described NLMS adaptive noise canceller will be used as a post processor to attenuate the C_2 component of the residual engine noise after the acoustic echo cancellation.

IV. Integrated Speech Enhancement System and Computer Simulations

Figure 7 shows the proposed integrated speech enhancement system that combine the acoustic echo canceller and noise canceller with two reference microphones. Here x(n), r(n) are two primary reference for far-end speaker signal and engine noise signal. After the acoustic propagation path in a car, echoed plus noise signal d(n) arrives to the receiving microphone, it pass through the acoustic echo canceller and noise canceller sequentially. We note that the system does not take any advantage of SNR improvement of noise canceller before the AEC, but it can at least avoid the distortions caused by the noise cancellation operation as described in [1].

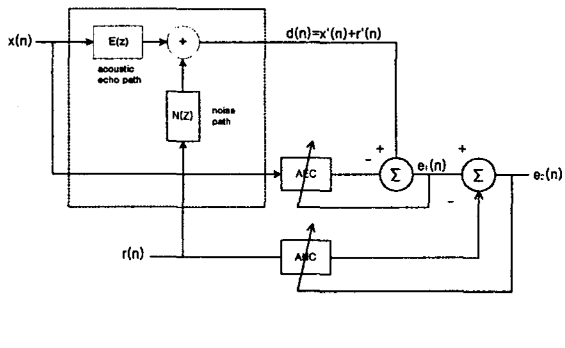


Figure 7. Proposed integrated speech enhancement system.

For the computer simulations of the integrate system, internally generated colored gaussian AR(6) signal and the actual recorded engine noise source as in figure 5 are used as the far-end signal reference x(n) and the noise signal reference r(n). For the acoustic echo cancellation operation, a delayless subband adaptive structure described in section 2 is used with narrowband adaptive filter size 32 tap with μ =0.99 so that the broadband filter has 512 taps. The number of subband implementation is 32 and the decimation factor of 16 is used so that the signal is 2X oversampled. The filter design shown in figure 4 is used as aprototype subband low-pass filter. On the other hands, the FIR adaptive filter size 256 tap is implemented with convergence ratio μ =0.001 for ANC operation.

To measure the acoustic and noise path, separate impulse response measuring experiments has been made between internally generated white noise source from the loudspeaker and receiving microphone signal, and measured 256 tap impulse responses are used to simulate receiving signals to the primary microphone[13].

Two sets of simulation are made to evaluate the system performance of proposed structure: one for the fixed acoustical environment condition, the other for the robustness of the system in which, more realistic situation, the acoustical transmission environment change. In order to evaluate integrated system performance, two kind of quantitative measures are used.

1) $$ERLE(n) = 10\log_{10}\frac{E[x'^2(n)]}{E[e_2^2(n)]}$$

2) $$ATT_{db} = 10\log_{10}\frac{E[r'^2(n)]}{E[e_2^2(n)]}, \quad ATT_{dblc2}$$

ERLE(Echo Return Loss Enhancement) measures the effectiveness of the echo canceller. The second parameter ATT_{db} measure the attenuation in overall noise spectra while ATT_{dblc2} measure the attenuation in highest noise peak at C2 component frequency which has more important information about the engine noise. Note that the described measures are calculated based on primary input and the residual error output $e_2(n)$of integrated structure to provide the information on the overall system performance.

A. Simulation 1: Fixed acoustic environment condition

The first set of simulation is conducted under the fixed acoustical environment condition. Figure 8, 9, 10 shows the simulation results for the first set of experiments. Considering figure 8 which show the ERLE curves when the four different sources of engine noise is added to the far-end echoed signal, we can notice that the echo cancellation operation has quite good convergence characteristics for all type of engine noise source. They all shows over 20dB ERLE convergence after first 3,000 iterations while, among them, the one at 3500 RPM has best convergence ERLE up to 25 dB.

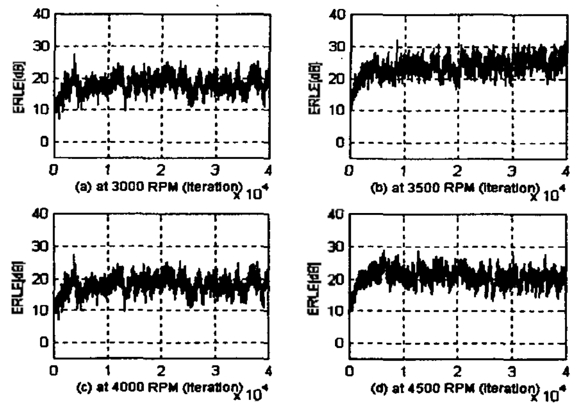


Figure 8. ERLE [dB] under four different RPM conditions.

The noise attenuation capability of the integrated system at 4000 RPM is shown in figure 9. Here the dashed line represents the residual noise spectra at the overall system output. As we see the figure, the overall attenuation of noise spectra, ATT_{db}= 14.12 dB can be observed at the corresponding RPM. Regarding to the C2 component frequency, the attenuation of ATT_{dblc2} =13.34 dB can be achieved. We note that the main contribution of attenuation occurs at the peaks of the noise spectra which demonstrate the selective attenuation feature of NLMS based ANC in dominant frequency component. Figure 10 summarize the improvement of noise attenuation capability at all engine RPMs. We note that the performance of ATT_{dblc2} , the noise peak attenuation, tends to decrease as the engine RPM increase while the overall noise spectra attenuation, ATT_{dblc2} , shows reverse characteristics.

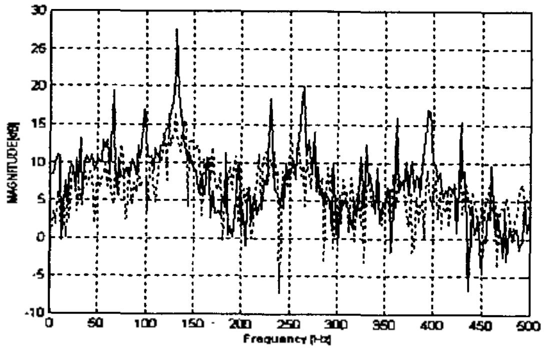


Figure 9. Noise spectra before(-) and after(--) the system at 4000 RPM.

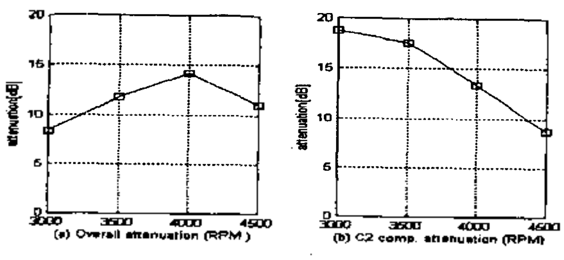


Figure 10. Engine noise attenuation under four different RPM conditions.

B. Simulation 2 : Variable acoustical environmental condition

The second set of simulation is done for the case of which the environmental acoustic transmission path is changed. For this purpose, we intentionally altered the acoustic echo path transfer function E(z) in figure 7 at sample instant n= 20,000 with new transfer function to reflect the transmission path change into the integrated system. The rest of simulation parameters are still applied as in simulation 1.

The simulation results for the variable acoustical environment of transmission path is shown in figure 11, 12, 13. Figure 11 shows the ERLE curves for the corresponding engine RPM. They all shows quite good tracking of the environmental change with fast convergence of ERLE over 20 dB. This brings the evidence of robustness in echo cancellation performance of the integrated system.

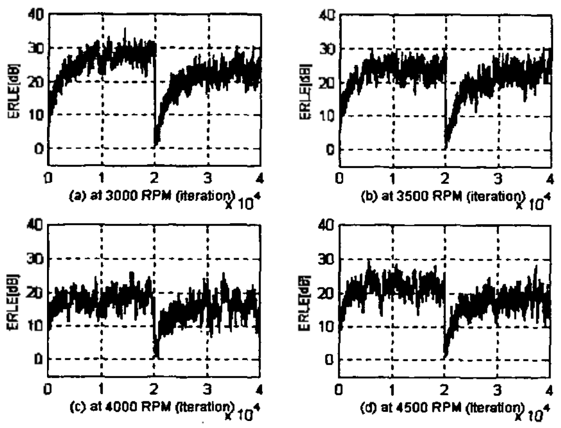


Figure 11. ERLE [dB] under four different RPM conditions.

Figure 12 shows the noise spectra before and after the integrated system operation at 4000 RPM. The dashed line again represents the residual noise spectra at the system output. From the figure, we see the efficiency of noise cancellation operation both for over the entire frequency range and for the C2 frequency component. The overall attenuation of ATT_{dB}= 11.82 dB is achieved and it is about 3 dB less than that of fixed environment condition. We note that this degradation comes from the additional error introduced in short tracking period for the adaptive echo canceller due to the environmental change. On the other hands, at C2 frequency, NLMS based adaptive noise canceller still plays a good role in the attenuation of C2 component noise and the attenuation of $ATT_{dB|C2}$=17.1 dB can be achieved. Figure 13 summarizes the noise attenuation performance for both overall frequency range and the C2 frequency component under all RPM conditions. Comparing to figure 10 of simulation 1, we notice that the noise attenuation performance tends to increase as the engine RPM increase under the environmental transmission change condition.

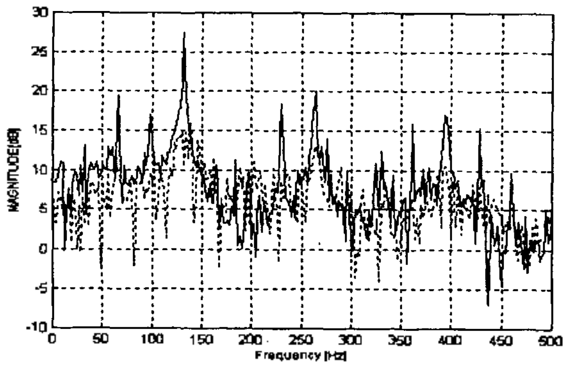


Figure 12. Noise spectra before(-) and after(--) the system at 4000 RPM.

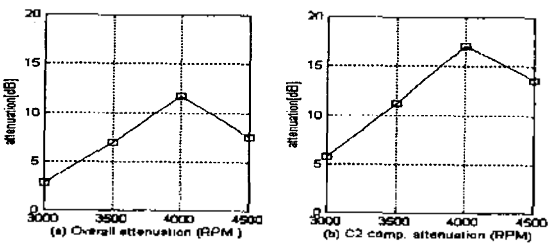


Figure 13. Engine noise attenuation under four different RPM conditions.

V. Conclusion

We have presented the integrated speech enhancement system that simultaneously cancel the acoustic echo and the engine noise which are the main sources of annoyance in hands-free mobile communication. In order to cancel the acoustic echo due to the reverberation effect of a car cabin, a delayless subband adaptive structure is used while for the mitigation of the engine noise at C2 component frequency, conventional NLMS based adaptive noise cancellation technique is applied.

Two sets of the computer simulation with the primary echoed plus the noise input are conducted under the fixed and variable acoustic environment condition. For the fixed environment condition, above 20 dB ERLE is observed with quite good convergence behavior, while for the noise attenuation in overall noise spectra and C2 component frequency, up to 8~15 dB and 9~19 dB is observed respectively. In order to test for the robustness of the system, we change the acoustic transmission environment and the results showed quite good tracking of the environmental change with fast convergence of ERLE over 20 dB. On the other hands, the noise attenuation in overall noise spectra and C2 component, the improvement of 3~11 dB and 9~14 dB are achieved. We note that, for the case of variable acoustic environment condition, the overall noise attenuation is about 5 dB less than that of the fixed condition and this degradation comes from the additional error introduced in short tracking period of the adaptive echo canceller due to the environmental change. These simulation results confirm the robust nature of the proposed system to the variable acoustic environment condition and the effectiveness of selective engine noise attenuation capability in the dominant frequency.

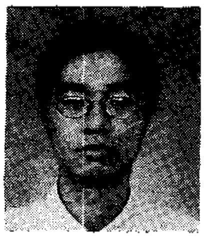
References

1. Yuan Guelou, Abdelkrim Benemar and Pascal Scalart, "Analysis of two structures for combined acoustic echo cancellation and noise reduction," *IEEE ICASSP'96*, pp. 637-640, 1996
2. Martie M. Goulding and John Bird, "Speech Enhancement for Mobile Telephony," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 4, November 1990
3. Mattias Dahl, Ingvar Glaesson and Sven Nordebo, "Simultaneous echo cancellation and car noise suppression employing a microphone array," *IEEE ICASSP'97*, pp. 239-242, 1997
4. Joerg Meyer and Klaus Uwe Simmer, "Multi-channel speech enhancement in a car environment using weiner filtering and spectral subtraction," *IEEE ICASSP'97*, pp. 1167-1170, 1997
5. Dennis R. Morgan, and James C. Thi, "A delayless subband adaptive filter architecture," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, August 1995
6. Simon Haykin, Adaptive Filter Theory, Processing, Prentice Hall Inc, Englewood Cliffs, 1991
7. Sen M. Kuo and Jyothsna Kunduru, "Multiple reference subband adaptive noise canceller for hands-free cellular phone application," *Journal of Signal Processing*, Elsevier Science, vol. 333, no. 5 pp.669-686, 1996[8] Ronald E. Crochiere and Lawrence R. Rabiner, Multirate Digital Signal Processing, Prentice Hall Inc, Englewood Cliffs, 1983
9. Alan V. Oppenheim, and Ronald W. Schafer, Discrete-Time Signal Processing, Prentice Hall Inc, Englewood Cliffs, 1989
10. R. A Gourban and H. M. Hafez, "Background noise reduction in mobile telephony," *Proc. IEEE Veh. Tech. Conf.* pp. 72-75, 1986
11. S. F. Boll and D. C. Pulsipher, "Suppression of acoustic noise in speech using two microphone acoustic noise cancellation," *IEEE Trans. on ASSP*, vol. 28, no. 6, pp. 752-763, Dec. 1980
12. Sen. M. Kuo, Holmes Chuang and Padma Mallela, "Integrated hands-free cellular phone, active noise control and audio system," *IEEE Trans.* Consumer Electronics, vol. 39, pp. 522-532, August 1993
13. J. Borish and J. B. Angell, "An efficient algorithm for measuring the impulse response using pseudorandom noise," *J. Audio Eng. Soc.*, vol. 31, no. 7, pp. 478-488, 1983

▲Kyu-Sik Park

Kyusik Park received B.S, M.S, and Ph.D degrees in 1986, 1988, and 1994, respectively, all from the Department of Electrical Engineering of Polytechnic University, Brooklyn, NY USA. In 1994, he joined the Semiconductor Division of Samsung Electronics as a staff engineer. He is currently an assistant professor with the Department of Information and Telecommunication in Sangmyung University, Chnungchongnam-Do, Korea. His research interests are digital signal, speech, and audio processing, and digital communication.

▲Sang-Hun Oh



Sanghun Oh received B.S degree in 1998 from the Department of Radio Sciences and Engineering, Korea Maritime University. From 1998, he has studied toward master degree in digital signal processing major in Sangmyung University. His research interests are digital signal processing, specially on speech and audio applications.