

네트워크 반향제거기의 성능 향상

Performance Improvement of the Network Echo Cancellor.

유 재 하*
Jae Ha Yoo

ABSTRACT

In this paper, an improved network echo canceller is proposed. The proposed echo canceller is based on the LTJ(lattice transversal joint) adaptive filter which uses informations from the speech decoder. In the proposed implementation method of the network echo canceller, the filter coefficients of the transversal filter part in the LTJ adaptive filter is updated every other sample instead of every sample. So its complexity can be lower than that of the transversal filter. And the echo cancellation rate can be improved by residual echo cancellation using the lattice predictor whose order is less than 10.

Computational complexity of the proposed echo canceller is lower than that of the transversal filter but the convergence speed is faster than that of the transversal filter. The performance improvement of the proposed echo canceller was verified by the experiments using the real speech signal and speech coder.

Keywords: network echo canceller, adaptive filter, residual echo cancellation, lattice predictor

1. 서 론

음성통신의 궁극적인 목적은 통화자간의 음성신호를 정확하고 시간지연 없이 전송하는 것이다. 그러나 그림 1과 같이 통신선로 상에서 완전하지 못한 하이브리드회로의 임피던스 정합에 의해 네트워크 반향(network echo)이 발생한다[1].

하이브리드 회로는 PSTN(Public Switched Telephone Network) 2 선 가입자선로의 종단이 위치한 전화국에 위치하며 원단으로부터 전달된 4 선 선로에 존재하는 음성신호를 가입자 측에 전달하고, 가입자 측의 음성신호를 4 선 선로를 통하여 원단으로 전달하게 된다. 그러나 새로운 통화가 형성될 때마다 임피던스가 달라지기 때문에 임피던스 부정합이 발생하게 되어 전달된 음성신호가 하이브리드 회로를 통하여 피드백 되어 다시 원단의 화자에게로 전송되어 반향 신호가 된다. 이러한 반향은 원단화자가 말한 시점과 되돌아온 반향신호의 시간차이가 적은 경우에는 반향신호로 들

* 한경대학교 전자공학과

리지 않기 때문에 통화 품질을 떨어뜨리지 않지만 시간차이가 큰 경우에는 확연한 반향 신호로 인지되어 통화품질이 급격히 저하된다. 그러므로 일반적인 전화통화에서는 문제가 되지 않지만 이동통신, VoIP(Voice over IP) 등과 같이 사용되는 경우에는 긴 시간지연이 발생하게 되어 반향신호로 인해 통화품질이 심각하게 저하된다[1].

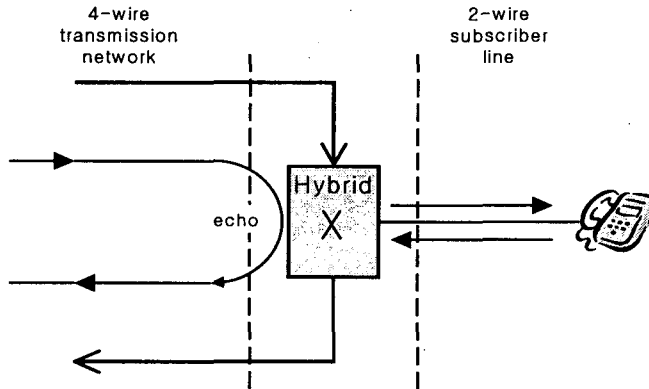


그림 1. 네트워크 반향

그림 2와 같이 하이브리드를 포함한 반향경로(echo path)를 반향 제거기를 설치하여 반향경로를 추정하고 추정된 반향을 빼줌으로서 반향을 제거할 수 있다. 반향제거기는 디지털 필터를 사용해 구현되며, 필터계수는 적응알고리즘을 사용하여 추정하게 되는데 이는 반향 경로가 고정된 값을 갖는 것이 아니라 시간에 따라 변화하기 때문이다. 반향 제거기의 출력 단에는 잔여 반향 신호(residual echo signal)가 남게 되는데 이는 완벽하지 못한 적응 알고리즘에 기인한다. 그림 2에서 $x(n)$ 은 원단화자(Far-End Talker) 신호, $y(n)$ 은 원단화자 신호가 하이브리드를 통과한 반향 신호, $d(n)$ 은 근단화자(Near-End Talker) 신호와 배경잡음이 합쳐진 신호이다. $z(n)$ 은 $y(n)$ 과 $d(n)$ 이 합쳐진 신호로서 적응필터의 주입력(primary input) 신호가 되며, $\hat{y}(n)$ 은 추정반향(estimated echo) 신호, $e(n)$ 은 오차신호(error signal)이다.

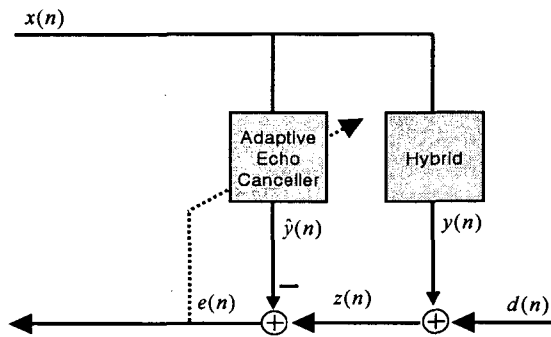


그림 2. 네트워크 반향제거기

LMS(least mean square) 알고리즘을 사용하는 트랜스버설(transversal) 필터는 구조가 단순하고 계산이 간단하여 실시간 구현이 용이하기 때문에 많은 분야에서 사용되어 왔으며 반향제거기에서도 많이 사용되어 왔으나 입력신호가 음성신호와 같이 상관도가 높은 신호인 경우에는 수렴속도가 느려지는 단점이 있어 적용에 문제가 발생한다. 이러한 수렴속도 문제를 개선하기 위하여 많은 고속 수렴 알고리즘이 제안되었으나 계산량이 많아 실시간 구현이 어렵거나[2] 지연이 발생하여 음성 통신 분야에 적용이 어렵다[3].

본 논문에서는 이와 같은 계산량 문제와 수렴속도 문제를 모두 다 개선할 수 있는 방법을 제안하고자 한다. 제안한 반향제거기는 디코더(decoder) 정보를 사용하는 LTJ 적응필터를 기본 구조로 하여 매 샘플마다 트랜스버설 필터 부분을 적용시키지 않고 D(D는 1보다 큰 정수) 샘플마다 적용시키는 구현 방법이다. 이러한 구현 방법은 트랜스버설 필터보다 계산량은 적으면서도 수렴속도는 빠르다. 또한, 적응필터에 의해 제거되지 않은 잔여반향을 10 차 이내의 격자예측기를 사용하여 필터링 함으로서 반향제거율을 더욱 향상시킨다.

2. 제안한 성능 향상 방법

2.1 제안한 적응필터

LTJ 적응필터는 그림 3과 같이 입력신호 $x(n)$ 을 격자예측기에 통과시켜 얻은 상관도가 많이 제거된 후방예측오차신호 $b(n)$ 을 트랜스버설 필터의 입력으로 사용하여 필터 출력신호 $\hat{y}(n)$ 을 얻는다. 여기서 $z(n)$ 과 $e(n)$ 은 주입력신호와 오차신호(error signal)를 나타내며, $c_m(n)$ 은 트랜스버설 필터의 계수를 나타낸다. $f_m(n)$ 과 $b_m(n)$ 은 각각 전방예측오차신호와 후방예측오차신호를 나타내며, $k_m(n)$ 은 반사계수(reflection coefficient)를 나타낸다. 격자예측기와 트랜스버설 필터의 각 신호 간 관계와 적응 방정식은 [4]에 잘 나타나 있다.

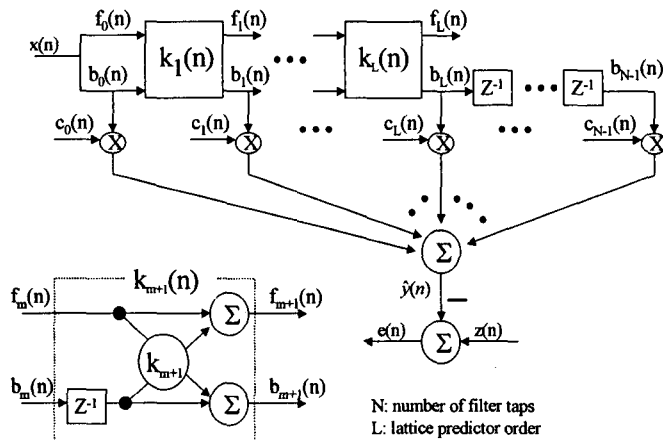


그림 3. LTJ 적응필터[4]

LTJ 적응필터는 정상상태 성능저하를 방지하기 위하여 필터계수보상이 필요한데 이를 위한 계산량은 $2LN[5]$ 으로서 트랜스버설 필터의 계산량 $3N$ 에 비하며 매우 크다. 여기서 N 은 적응필터의 탭수를 L 은 격자예측기의 차수를 나타낸다. 음성 코덱(CODEC:COder-DECOder)을 사용하는 응용 분야에서는 프레임 또는 서브프레임 단위로 LPC(Linear Prediction Coding) 계수가 변하기 때문에 매 샘플마다 필터계수의 보상이 필요하지 않으며, 디코더로부터 LPC 계수 정보를 얻어내어 반사계수로 전환할 수 있다는 사실을 사용함으로써 트랜스버설 필터에 비해 10% 정도의 계산량 증가만으로도 고속 수렴이 가능한 LTJ 적응필터가 제안되었다[4].

본 논문에서는 이에서 더 나아가 수렴특성은 트랜스버설 필터보다 우수하면서 계산량은 더 적은 LTJ 적응필터의 구현방법을 제안하여 네트워크 반향제거기에 적용하고자 한다. 그림 4에서 보는 바와 같이 LTJ 적응필터는 입력 음성신호의 상관관계를 제거하는 격자예측기 부분과 격자예측기 부분을 통과하여 상관관계가 감소한 후방예측 오차 신호를 트랜스버설 필터링 하는 트랜스버설 필터 부분으로 나눌 수 있다. 트랜스버설 필터부분의 계산량은 NLMS(Normalized LMS) 알고리즘을 사용하는 일반적인 트랜스버설 필터와 같이 $3N$ 이다. 또한 격자예측기를 통과한 후방예측 오차신호는 상관관계가 매우 줄어든 신호이기 때문에 트랜스버설 필터 부분을 매 샘플마다 계수 적용시키지 않고 $D(D>1, \text{정수})$ 샘플마다 한번만 적용시키더라도 트랜스버설 필터만을 사용하는 경우보다 빠른 수렴속도를 기대할 수 있으며 이를 실제 음성신호를 사용한 실험을 통하여 제안된 구현 방법이 가능함을 보이고자 한다. 표 1은 트랜스버설 필터와 제안한 적응필터와의 계산량을 비교하여 나타낸 것이다. 2 샘플마다 1번만 적용시키는 경우($D=2$) 트랜스버설필터에 비해 22% 이상 계산량을 감소시킬 수 있다. D 를 증가시킬수록 계산량을 더 감소시킬 수 있지만 수렴속도 또한 느려지므로 사용에 제한을 갖게 된다. D 값에 따른 성능은 모의실험을 통하여 고찰한다.

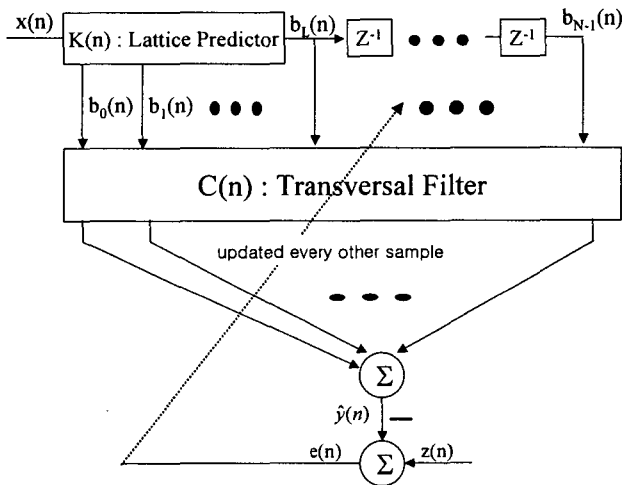


그림 4. 제안한 적응필터 구현 방법

표 1. 제안한 방법의 계산량 비교

	Multiplication
트랜스버설 필터	3N
제안한 적응필터	$(\frac{4}{3} + \frac{2}{D})N$

(N: 적응필터 차수, L: 격자예측기 차수, D: 트랜스버설 필터 부분 계수 적용 주기)

2.2 잔여반향 제거기

반향제거기는 적응필터를 사용하여 반향신호를 제거하게 되는데 적응필터의 부정확한 추정으로 인하여 오차신호에는 배경잡음뿐만 아니라 잔여반향 신호가 존재하게 되므로 동시통화 검출을 위한 음성검출기(VAD: Voice Activity Detector)의 성능을 저하시킬 수 있을 뿐만 아니라 반향제거기 후단에 위치할 비선형처리기(NLP: Non-Linear Processor)의 동작구간을 단축시킴으로서 전체 성능을 저하시킬 수 있다[6]. 그러므로 그림 5와 같이 오차신호를 입력신호로 하는 잔여반향제거기를 설치하여 전체적인 성능을 향상시킬 수 있다.

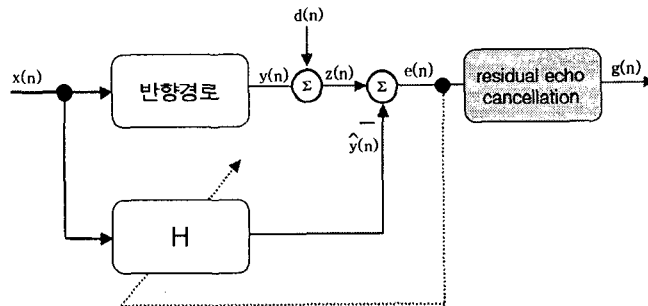


그림 5. 잔여반향제거기

[6]에서는 잔여반향 신호 e(n)을 AR(Auto Regressive) 모델링을 통하여 구한 AR 계수를 (1) 과 같이 사용하여 잔여반향을 제거한다. 이 때 AR 계수는 levinson-durbin recursion에 의한다[7].

$$g(n) = \sum_{i=0}^P a_i e(n-i), \text{ with } (a_0 = 1). \tag{1}$$

여기서 P는 AR 모델링 차수이다.

본 논문에서는 [6]과는 달리 격자예측기(lattice predictor)를 사용하여 잔여반향을 제거하며 격자 예측기의 반사계수는 블록단위가 아닌 샘플단위로 적응적으로 구한다. 격자예측기의 차수를 L이라 하면 음성신호의 경우 10 차 이내를 사용한다. L 번째 단의 전방예측 오차신호가 잔여반향이 제거 된 신호 g(n)이 된다.

잔여반향제거를 위한 계산량은 9 L이며, L은 10 이하이다[8]. 반면, 128 ms의 반향경로를 처리

하기 위해서 사용되는 적응필터의 탭수 N 은 1024이다. 이와 같은 점을 고려하면 제안한 적응필터의 계산량에 격자예측에 필요한 계산량을 더하더라도 트랜스버설 필터의 계산량보다 적게 되므로 제안한 방법은 LTJ 적응필터의 우수한 수렴특성과 잔여반향제거기를 통한 오차신호 감소라는 두 가지 성능 향상을 얻어 낼 수 있다.

3. 모의실험 및 결과 고찰

본 논문에서 제안한 적응필터와 잔여반향제거기의 성능을 검증하기 위하여 실제 음성 신호를 사용하는 네트워크 반향제거기에 적용한 모의실험을 수행하였다. 사용된 음성신호는 실제 음성신호를 8 kHz의 표본화율로 녹음하였으며, 음성 코덱은 ITU-T G.723.1을 사용하였다. 반향경로와 적응필터 탭수 N 은 모두 30이며, 반향신호 대 배경잡음비(ENR:Echo to Noise Ratio)는 40 dB이다. 필터의 성능은 ERLE(Echo Return Loss Enhancement)를 사용하여 표현하였다[4].

3.1 모의실험 : 적응필터

실험은 반향제거기의 수렴성능과 추적성능에 대해 평가하였다. LTJ 적응필터의 계수 보상은 서브프레임 단위로 수행되도록 하였으므로 60 샘플마다 계수 보상이 이루어지며 트랜스버설 필터 부분은 필터링은 매 샘플 마다 수행하지만 계수 적용은 2샘플에 한번씩만 수행한다.

그림 6은 수렴성능을 나타내고 있다. (a)는 반향제거기의 기준입력신호가 되는 음성신호이고 (b)는 트랜스버설 필터의 ERLE 곡선을, (c)는 제안한 적응필터의 ERLE 곡선을 나타낸다. (c)는 (b)에 비해 매우 빠른 수렴속도를 나타냄을 알 수 있다. 2 초 동안의 평균 ERLE는 (b)의 경우 30.8 dB인 반면 제안한 방법인 (c)의 경우 42.7 dB로서 약 12 dB 정도 성능향상이 되었다.

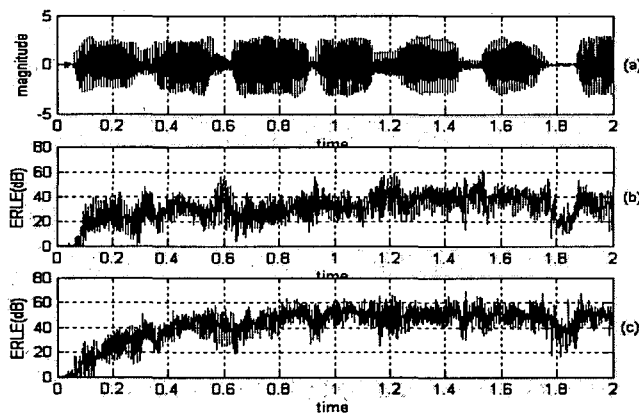


그림 6. 수렴 성능 비교

(a) 입력신호, (b) ERLE: 트랜스버설 필터, (c) ERLE: 제안한 적응필터

그림 7은 추적성능을 나타낸 결과이다. 1 초 경과한 후에 반향경로의 특성이 갑자기 바뀐 경우이다. 제안한 적응필터인 (c)는 새로이 바뀐 반향경로를 빠르게 추적해 나감을 볼 수 있다. 반면, (b)는 수렴특성에서의 경우와 마찬가지로 새로이 바뀐 반향경로를 빠르게 추적하지 못하고 있다. 2 초 동안의 평균 ERLE는 (b)의 경우 24.3 dB인 반면 제안한 방법인 (c)의 경우 30.6 dB로서 약 6 dB 정도 성능향상이 되었다.

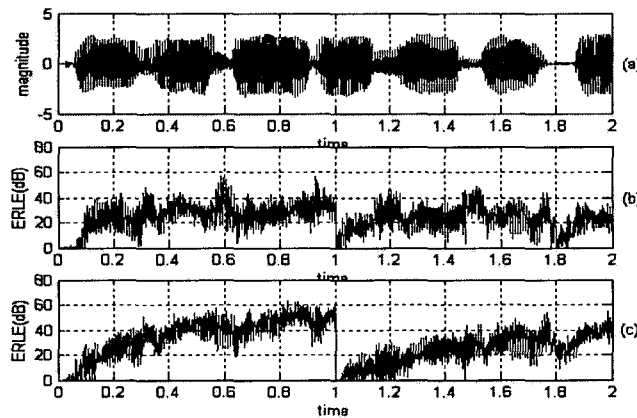


그림 7. 추적 성능 비교

(a) 입력신호, (b) ERLE: 트랜스버설 필터, (c) ERLE: 제안한 적응필터

3.2 모의실험: D 값에 따른 성능 비교

제안한 적응필터의 D 값을 2부터 차례로 증가시켜가며 수렴성능을 실험하였다. 표 2는 D 값에 따른 2 초 동안의 평균 ERLE를 나타낸다. D 값이 6이 되면 트랜스버설 필터의 평균 ERLE인 30.8 dB보다 성능이 저하됨을 볼 수 있다. D값을 1 증가시킬 때마다 성능은 약 3 dB 가량 저하되지만 그에 따른 계산량 감소 효과는 상대적으로 적으므로 D를 2로 하는 경우가 가장 타당한 값으로 판단된다.

표 2. D값에 따른 수렴 성능

D	2	3	4	5	6
평균 ERLE(dB)	42.7	40.2	35.1	32.4	29.1

3.3 모의실험: 잔여반향제거기

잔여반향제거기에 대한 실험은 적응필터에 대한 모의실험과 동일한 환경이며, 오차신호도 제한한 방법을 사용한 경우의 오차신호이다. 잔여반향제거기는 10 차의 격자예측기를 사용한 경우이다. 그림 8은 잔여반향제거 전후의 신호를 나타낸 것이다. 그림에서 볼 수 있는 바와 같이 잔여반향제거기를 통과한 오차신호인 (b)가 오차신호 (a)에 비해 상대적으로 많이 감소하였음을 알 수 있으며

수치적으로는 2 dB 이상 감소하였다.

그러므로 제안한 방법은 LTJ 적응필터를 사용하여 12 dB의 성능 향상과 잔여반향제거기를 사용하여 2 dB의 성능 향상을 얻어내었다.

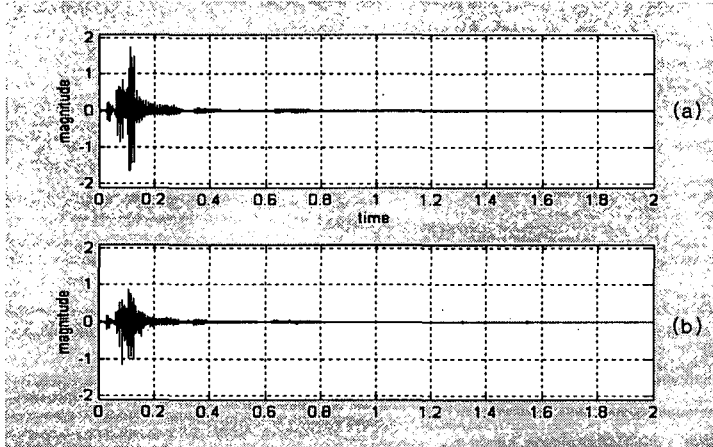


그림 8. 잔여반향제거 실험
(a) 오차신호, (b) 잔여반향제거기를 통과한 신호

5. 결 론

본 논문에서는 네트워크 반향제거기에 주로 사용되는 LMS 알고리즘을 사용하는 트랜스 버설 필터 보다 계산량은 적으면서도 반향제거 성능은 우수한 네트워크 반향제거기를 제안하였다. 제안된 반향제거기는 LTJ 적응필터의 트랜스버설 필터 부분을 2 샘플당 1 번만 적용시키는 방법을 사용하여 트랜스버설 필터에 비해 수렴속도도 빠를 뿐만 아니라 전체 계산량도 적은 LTJ 적응필터 구현 방법을 제안하였다. 또한, 적응필터에 의해 제거되지 않은 잔여반향은 10 차 이내의 격자예측기를 사용하여 필터링 함으로서 반향제거율을 향상시켰다. 제안된 반향제거기를 실제 음성신호와 음성부호화기를 사용한 실험을 통하여 제안된 반향제거기의 성능 향상을 검증하였다. 제안된 적응필터를 사용하여 12 dB의 성능 향상을 얻어냈으며, 격자예측기를 사용한 잔여반향 제거를 통하여 부가적으로 2 dB의 성능이 향상되었다.

참 고 문 헌

- [1] *Int. Telecommunication Union. ITU-T G.168*, 2002.
- [2] Farhang-Boroujeny, B. 1998. *Adaptive Filters Theory and Applications*, John Wiley & Sons.
- [3] Gilloire, A., Vetterli, M. 1992. "Adaptive filtering in subbands with critical sampling: analysis,

- experiments, and application to acoustic echo cancellation." *IEEE Transactions on Signal Processing*, Vol., 40, No. 8, pp. 1862-1875, Aug.
- [4] 유재하. 2004. "LTJ 적응필터의 실용적 구현과 적응반향제거기에 대한 적용." 한국음성과학회 논문지, 제11권, 제2호, pp. 227-235, 6월.
- [5] Naoki Tokui, Kenji Nakayama and Akihiro Hirano. 2001. "A Synchronized Learning Algorithm for Reflection Coefficients and Tap Weight In a Joint Lattice Predictor and Transversal Filter." *Proceedings of International Conference on Acoustics, Speech, and Signal Processing 2001*, pp. 3741-3744.
- [6] Park, Seon Joon, Lee, Chungyong, Youn, Dae Hae. 2002. "A Residual Echo Cancellation Scheme for Hands-Free Telephony." *IEEE Signal Processing Letters*, Vol.9, No.12, pp.397-399, Dec.
- [7] Thomas, F. 2002. Quatieri, *Discrete-Time Speech Signal Processing Principles and Practice*, Prentice-Hall.
- [8] Yoo, J-H., Cho, S-H., and Youn, D-H. Youn. 1998. "An Acoustic Echo Cancellation Based on the Adaptive Lattice-Transversal Joint (LTJ) Filter Structure." *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, **E81-A** (9), 1951-1954.

접수일자: 2004. 11. 01

게재결정: 2004. 11. 30

▲ 유재하

경기도 안성시 석정동 67 (우: 456-749)

한경대학교 전자공학과

Tel: +82-31-670-5196

E-mail: yjh@hknu.ac.kr