

입/출력 신호의 상관계수를 이용한 반향제거기

강명구, 유창동

한국통신 멀티미디어연구소 음성언어연구실

Echo Canceller using Cross-Correlation of Input and Output Signals

Myung-Goo Khang, Chang-Dong Yoo

Spoken Language Research Team, Multimedia Technology Research Laboratory, Korea Telecom

Email : {mgkang, cdyoo}@smm.kotel.co.kr

요약

전화망을 이용한 음성인식 시스템에서 출력신호의 반향에 의한 입력신호의 음질 열화현상을 막기위해서 적응 디지털 필터를 이용한 반향제거기가 필수적이다. 대표적인 적응 필터 알고리즘인 LMS 와 NLMS 들을 각각 이용한 적응 반향제거기들과 입/출력신호의 상관계수를 이용한 개선된 적응 필터 알고리즘의 성능을 비교하였다. 개선된 알고리즘의 경우 NLMS 알고리즘의 빠른 수렴특성을 가지면서도 더블톡(double talk)구간에서의 음질왜곡 현상을 LMS 보다 개선시켰다.

I. 서론

전화망을 이용한 서비스 중에서 음성인식 기술을 응용하여 인간과 기계간에 대화형식으로 정보를 제공하는 시스템의 경우에 시스템에서 제공하는 안내방송 신호와 사용자가 입력하는 음성신호가 서로 양방향으로 전달되게 된다. 전화선로에서는 단방향 4 선을 양방향 2 선 신호로 변환해주는 하이브리드(Hybrid) 회로와 전화라인간의 임피던스 불일치에 의해 원단(far end) 신호의 반향성분이 중첩되어 근단(near end) 신호의 음질을 저하시키게 된다. 이를 전자 반향신호(electrical echo)라하며, 스피커 전화기나 핸드프리(hand-free) 전화기의 스피커를 통해서 흘러나오는 안내방송 신호가 피드백(feedback)되어 사용자의 음성신호와 중첩되어 마이크 입력으로 흘러들어가게 되는데 이것을 음향 반향신호(Acoustic Echo)라 한다. 정보통신 서비스의 품질을 개선하기 위해서는 반향문제는 반드시 해결되어야 한다.

대화 방식으로 음성의 입/출력이 양방향으로 동시에 진행되는 음성인식 서비스 시스템에서 반향신호에 의한 음질열화 현상을 줄이고 음질저하로 인한 인식을 저하를 막기 위하여 출력음성(안내방송)의 반향성분을 자동적으로 제거하기 위한 적응 디지털 필터를 이용한 방법이 있다[4]. 이 방법은 이미 알고 있는 출력음성을 참조로 하여 적응 필터를 이용하여 입력 음성에서 반향 경로에 의한 출력 신호의 영향을 자동적으로 제거해 주는 방식이다. 본 연구에서는 널리 알려진 적응필터 알고리즘인 LMS(Least Mean Square) 와 NLMS(Normalized Least Mean Square) 알고리즘의 특성을 분석하고, 입/출력 신호가 더블톡 구간에서 상호 상관성이 떨어지는 점을 이용하여 NLMS 알고리즘의 빠른 수렴특성을 살리면서 더블톡 구간에서 왜곡현상을 줄이는 새로운 반향제거 알고리즘을 제안하고 비교분석 하였다. RLS(Recursive

Least Square) 알고리즘은 수렴속도가 빠르고 오차가 적지만 계산량이 많고 수치적 특성이 불안정하기 때문에 실시간에 인정되게 처리되어야 하는 반향제거기의 구현 특성상 본 연구에서는 비교에서 제외하였다[2].

2 장에서는 일반적인 반향제거기의 모델을 설명하고 3 장에서는 상관계수를 이용한 적응필터 알고리즘을 제안하며 4 장에서는 LMS, NLMS 알고리즘과 상관계수를 이용한 알고리즘을 비교실험한 결과를 분석하며 5 장에서는 마지막으로 본 연구결과를 종합 요약한다.

II. 반향이 생기는 시나리오

반향은 크게 전화선로 상에서 생기는 전자 반향과 공간 환경에 의해 생기는 음향 반향으로 구분할 수 있으며 반향신호는 고정된 경로를 유지하는 것이 아니라 실제 상황에서 항상 변화하기 때문에 반향제거를 위해서 적응 필터 방식을 사용하는게 바람직하다.

2.1 전자 반향

전화선로상에서 4 개 라인을 2 개 라인으로 줄이는 하이브리드(Hybrid)회로를 이용하는데, 하이브리드 회로의 라인 임피던스 불일치 때문에 원단(far end)에서 전송된 신호가 반전되어 원하지 않는 반향신호를 만들어 낸다. 이러한 반향신호를 제거하기 위하여 반향제거기가 필요하다. 반향제거기는 4 방향으로 연결된 그림 1 과 같이 표현될 수 있다. ITU-T G.165 에서 나타내어진 것과 같이 그림 1 에서 Rin 은 Receive-In Port , Rout 은 Receive-Out Port , Sin 은 Send-In Port , Sout 은 Send-Out Port 를 나타낸다.

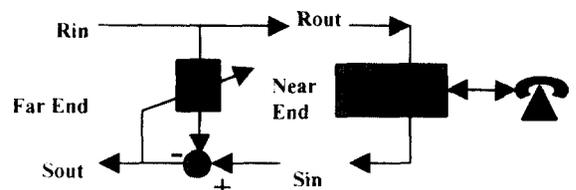


그림 1 전자 반향제거기

2.2 음향 반향

스피커 전화기나 핸드프리 셀룰라(ellular) 단말기등에서

스피커를 통해 나온 출력신호가 피드백되어 입력으로 중첩되어 들어가면서 원하지 않는 반향신호를 만들어 낸다. 이러한 반향신호를 제거하기 위하여 음향반향제거(Acoustic Echo Cancellation) 신호처리 기법을 이용하게 된다. 그림 2는 음향 반향반향제거기의 구조를 나타낸다. 음향 반향신호는 대체적으로 전자 반향신호에 비해 딜레이(delay)가 훨씬 크다.

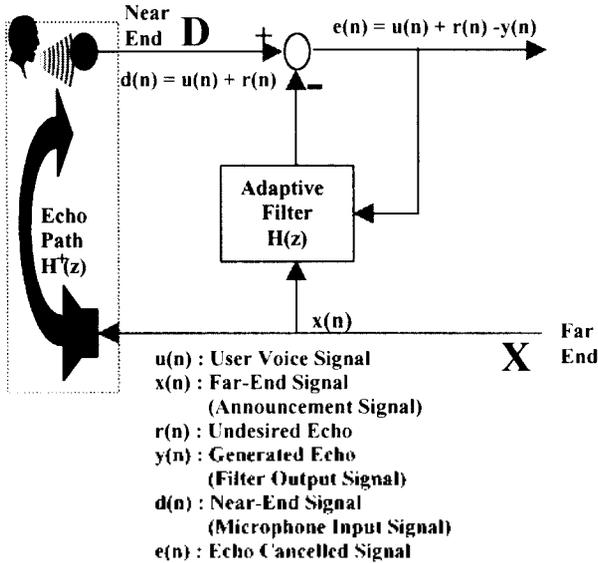


그림 2. 음향 반향제거기

2.3 적응필터를 이용한 반향제거기

적응필터를 이용하여 원 안내방송(출력음)을 참조로 하여 안내방송의 반향신호를 모델링 할 수 있는데 이렇게 모델링된 반향신호를 이용하여 반향신호와 중첩된 사용자 입력신호로부터 반향성분을 제거하고자 하는 방법이 적응필터를 이용한 반향제거 기법이다. 그림 3은 적응필터를 이용한 반향제거기의 구조를 나타낸다. 그림 3에서 $u(n)$ 은 사용자 음성신호, $r(n)$ 은 반향신호, $y(n)$ 은 필터 출력신호, $x(n)$ 은 안내방송 신호(출력신호), $d(n)$ 은 사용자 음성신호와 반향신호의 합인 시스템 입력신호, $e(n)$ 은 적응필터에 의해 반향이 제거된 신호를 나타낸다. 적응필터가 침적으로 작용할 경우에 $e(n)$ 은 $u(n)$ 과 같은 값을 가지게 된다.

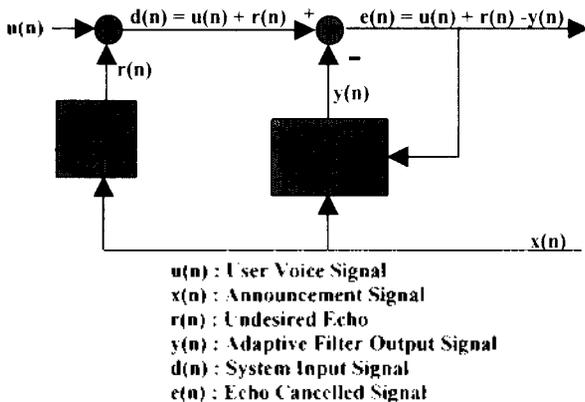


그림 3. 적응필터를 이용한 반향제거기 구조

III. 입/출력 신호의 상관계수를 이용한 적응필터 알고리즘

LMS 알고리즘은 적응 필터 방식 중 가장 기본이 되며 일반적으로 많이 쓰이는 알고리즘으로서 수치적 특성이 좋고 계산량이 적지만 수렴 속도가 느리며 필터의 성능을 좌우하는 스텝 크기를 정하는 것이 어렵고, 고정된 스텝 크기를 사용하기 때문에 필터 입력신호의 진폭 변화에 따라 적응 필터의 성능이 영향을 받는다[1,2]. 이런 문제를 해결하기 위하여 입력 신호를 입력 전력에 대해서 정규화 시키는 방법이 NLMS 알고리즘이다. NLMS 알고리즘의 경우에는 LMS 알고리즘에 비해 수렴속도가 빠르고 정규화된 스텝크기를 정하는 것이 LMS에 비해 매우 용이하다. 그러나 이러한 적응 필터들은 사용자 음성이 반향성분과 공존하는 실제 환경에서 사용자 음성의 영향으로 신호를 심하게 왜곡시킨다[5]. 사용자 음성이 없을 때 LMS보다 성능이 우수한 NLMS의 경우, 사용자 음성이 공존하는 실제 상황에서 사용자 음성에 의해서 LMS 보다 더 심한 왜곡이 생긴다. 실제 환경에서는 반향신호와 더불어 사용자 음성신호가 중첩되어 입력으로 들어가게 된다. 이 경우에 반향신호와 사용자 음성신호가 중첩되는 부분이 생기는데 이렇게 두개의 신호가 중첩되어 있는 구간을 더블톡(double talk) 구간이라 한다. 더블톡 구간에서 적응필터는 사용자 음성에 의한 오차에 영향을 받아 신호를 왜곡시키게 된다. 신호의 왜곡을 줄이기 위해서는 적응필터에 사용하는 스텝크기를 작게 하는게 유리하지만 스텝크기를 작게 할 경우 수렴속도가 느려져서 제대로 반향제거 기능을 하지 못하게 된다. 수렴속도에 영향을 주지 않으면서 이러한 왜곡현상을 없애기 위해서는 더블톡 구간을 감출하는 부분이 필요하다. 본 논문에서 제안하는 개선된 적응필터 알고리즘은 시스템 입력신호(사용자음성신호 + 반향신호)와 출력신호(안내방송:적응필터 입력신호)가 더블톡 구간에서는 서로 상관성이 떨어지는 점을 이용하여 시스템 입/출력 신호의 상호 상관계수를 구하여 NLMS 적응필터의 스텝크기에 공해중으로써 더블톡 구간에서의 왜곡현상을 막도록 하였다. 알고리즘은 아래와 같다.

$$y(n) = x_n h_n^t = \sum_{k=0}^{N-1} x(n-k) h_k^{(n)}$$

$$R_x(n) = \sum_{k=0}^{N-1} x(n-k)^2$$

$$R_d(n) = \sum_{k=0}^{N-1} d(n-k)^2$$

$$R_{cd}(n) = \sum_{k=0}^{N-1} x(n-k)d(n-k)$$

$$COR_{cd}(n) = \frac{R_{cd}(n)}{MAX(R_x(n), R_d(n))}$$

$$\mu(n) = \frac{COR_{cd}(n)}{1 + R_x(n)}$$

$$e(n) = d(n) - y(n)$$

$$h_{n,1} = h_n + \mu(n)e(n)x_n$$

$$\mathbf{x}_n = [x(n), x(n-1), \dots, x(n-N+1)]$$

: 필터 입력 벡터 (안내방송 신호)

$$\mathbf{h}_n = [h_0^{(n)}, h_1^{(n)}, \dots, h_{N-1}^{(n)}]$$

: 적응필터 가중치 벡터

$y(n)$: 적응필터 출력

$x(n)$: 필터 입력신호 (안내방송)

$r(n)$: 안내방송의 반향신호

$u(n)$: 사용자 음성 입력신호

$d(n)$: 시스템 입력 신호 ($d(n) = r(n) + u(n)$)

$e(n)$: 오차 신호

N : 적응필터 탭수

IV. 실험결과

더블톡 구간에서 기존의 LMS/NLMS 알고리즘과 입/출력 상관계수를 이용한 새로운 알고리즘의 성능을 비교하기 위하여 실제 전화라인을 통해서 양방향으로 입/출력신호가 공존하는 환경에서 실험하였다. 그림 4는 적응필터의 입력이 되는 시스템 출력신호(안내방송) $x(n)$ 과 반향성분을 포함한 시스템 입력신호 $d(n)$ 이다. 그림 4와 $d(n)$ 은 전화라인을 통해서 생기는 반향성분에 실제 음성 신호가 인가되어 6000~12000 사이의 구간에 더욱 짙은 구간이 형성되었다. 그림 5는 각각 LMS, NLMS 그리고 상관계수를 이용한 경우의 반향제거 처리결과 반향이 제거된 오차신호 $e(n)$ 을 나타낸다. 그림 6은 각 알고리즘의 처리결과 시스템 입력신호 대 오차신호의 평균전력비, ERLE(echo return loss enhancement) 값의 변화를 나타낸다.

$$ERLE = 10 \log \frac{P_d(n)}{P_e(n)}$$

$P_d(n)$: 시스템 입력신호 $d(n)$ 의 작은 구간에서의 파워 추정치

$P_e(n)$: 오차신호 $e(n)$ 의 작은 구간에서의 파워 추정치

그림 5를 보면 NLMS 알고리즘의 경우 사용자 음성이 없는 더블톡 이전 구간(6000 이하 구간)에서 LMS 알고리즘과 비교해서 잔존 오차가 적으나 더블톡 구간을 거치면서 신호를 심하게 왜곡시킴을 알 수 있다. 상관계수를 이용한 새로운 알고리즘의 경우 더블톡 이전구간에서 NLMS와 같이 좋은 성능을 보이면서 더블톡 구간에서도 신호의 왜곡현상이 현저히 개선됨을 알 수 있다. LMS 알고리즘의 경우 반향경로가 일정하지 않은 실제 전화라인의 환경에서 적응속도가 떨어지기 때문에 반향제거 성능이 NLMS 알고리즘에 비해서 떨어지지만 더블톡 구간에서의 왜곡현상은 심하지 않았다. 그림 6에서 NLMS 알고리즘의 경우 더블톡 구간을 거치고 난 뒤 신호왜곡에 의해 ERLE 값이 현저하게 떨어짐을 알 수 있다. 전체적인 ERLE 값의 분포에 있어서 LMS에 비해 상관계수를 이용한 방법이 높은 값을 유지하고 있으며, 그림 5의 결과와 같이 잔존 오차신호도 LMS에 비해 적었다.

시험환경
음성신호 : 8KHz 샘플링
탭수 N : 80

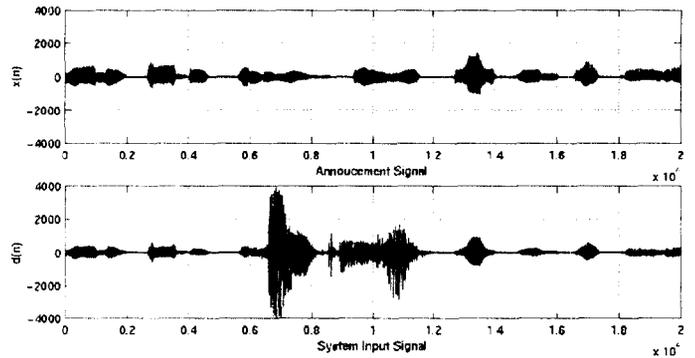


그림 4. 안내방송 신호 $x(n)$ 과 시스템 입력신호 $d(n)$

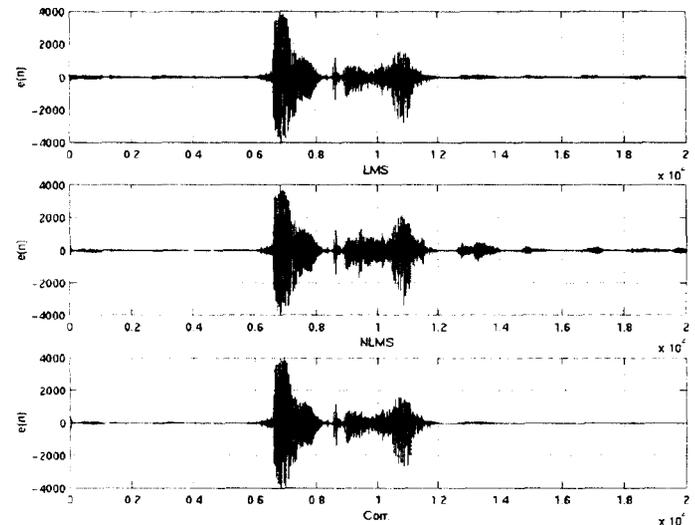


그림 5. 반향제거후 각 알고리즘별 오차신호 $e(n)$

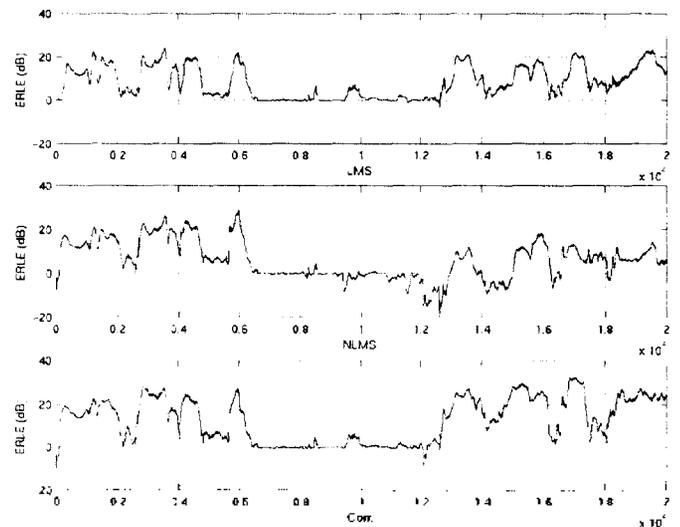


그림 6. 각 알고리즘별 입력신호대 오차신호의 평균전력비

V. 결론

본 논문에서는 출력신호(안내방송)의 영향으로 입력신호의 음질이 저하되는 현상을 막기 위한 자동적인 적응필터에 대해서 알아보고, 실험을 통하여 성능을 분석하였으며 새로운 적응필터 알고리즘을 제안하였다. LMS 알고리즘의 경우 고정된 스텝크기를 정하는 것이 어렵고 수렴속도가 느린 반면 NLMS의 경우에는 사용자의 음성신호가 없을 경우에는 수렴속도도 빠르고 매우 우수한 성능을 보인다. 그러나 사용자 음성이 반향성분과 공존하는 더블톡 구간에서는 NLMS 알고리즘은 매우 심한 왜곡현상을 보였다. 수렴속도를 떨어뜨리지 않으면서 더블톡 구간에서의 심한 왜곡현상을 없애기 위하여 입/출력 신호의 상관계수를 이용한 적응필터 알고리즘을 제안하였다. 새로운 방법은 NLMS와 같이 적응필터의 입력신호를 정규화 하면서 시스템의 입/출력 신호의 상관계수를 이용하여 더블톡 구간을 검출하기 때문에 수렴속도를 떨어뜨리지 않으면서도 더블톡 구간에서의 심한 왜곡현상을 막을 수 있었다. 따라서 새롭게 제안한 상관계수를 이용한 알고리즘을 적용하면 반향제거기의 성능을 높일 수 있을 것이다.

참고문헌

- [1] Simon Haykin, "Adaptive Filter Theory", Prentice Hall Inc., 1991
- [2] John R. Treichler, C. Richard Johnson and Micheal G. Larimore, "Theory and Design of Adaptive Filters, John Wiley & Sons, 1987.
- [3] Bernard Widrow and Samuel D. Stearns, "Adaptive Signal Processing", Prentice-Hall, 1985.
- [4] Hong Fan and W. Kenneth Jenkins, "An Investigation of an Adaptive IIR Echo Canceller : Advantages and Problems", IEEE Trans. on Acoust. Speech and Signal Processing, vol. 36, no. 12, pp. 1819-1833, Dec. 1988.
- [5] Hector Perez Meana and Luis Nino, "A Time Varying Step Size Normalized LMS Echo Canceller Algorithm", Proc. ICASSP'94, pp. II-249-II-252, 1994.