

주입력신호의 LPC 필터 이득을 이용한 반향제거기의 수렴전 동시통화검출 성능 개선

Performance Improvement of Double Talk Detection before Convergence of the Echo Canceller by Using Linear Predictive Coding Filter Gain of the Primary Input Signal

유재하
Jae-Ha Yoo

한경대학교 전기전자제어공학과/ IT융합연구소

Dept. of Electrical, Electronic and Control Engineering, Hankyong National University

요 약

본 논문에서는 반향제거기가 수렴하기 전에 동작할 수 있는 기존 동시통화검출법의 성능을 개선하기 위한 방법을 제안하였다. 제안된 방법은 LPC 필터 계수를 주입력신호로부터 추정하게 된다. 동시통화검출을 위한 문턱치는 주입력신호의 크기별로 기대되는 LPC 필터 이득치를 사용하여 가변적으로 설정하게 된다. 제안한 방법은 기존의 방법이 갖고 있는 단일통화를 동시통화로 잘못 판단하는 오검출률을 개선할 수 있을 뿐만 아니라 동시통화 검출 지연 시간도 단축시킬 수 있다. 장시간의 음성신호를 사용한 모의실험을 수행하였다. 제안한 방법이 기존 방법의 오검출률을 감소시키고 동시통화 검출 지연시간을 단축시킴을 확인할 수 있었다.

키워드 : 동시통화검출, 문턱치, 반향제거기, LPC 필터 이득, 음성신호

Abstract

This paper proposes a performance improvement method of the conventional double talk detection method which can operate before convergence of the echo canceller. The proposed method estimates the coefficients of the linear predictive coding(LPC) filter by using the primary input signal. The time-varying threshold for double talk detection is determined based on the LPC filter gain of the primary input signal level. The proposed method can reduce not only false detection rate which means wrong detection of single talk as double talk but also double talk detection delay. Computer simulation was performed using a long-term real speech signals. It is shown that the proposed method improves the conventional method in terms of lowering the false detection rate and shortening the detection delay.

Key Words : Double Talk Detection, Threshold, Echo Canceller, LPC Filter Gain, Speech Signal

1. 서 론

동시통화는 음성통신에서 발생하는 반향을 제거하는데 사용되는 반향제거기에서 원단화자와 근단화자가 동시에 음성을 발생하는 구간을 의미한다[1]. 동시통화 구간에서도

반향제거기의 적응필터를 계속 적용시킬 경우에는 필터계수가 발산하게 되므로 동시통화를 빠르고 정확하게 검출하여 동시통화가 발생하였다고 판단된 경우에는 즉시 반향제거기의 필터계수 적용을 멈추도록 하여야 한다[2][3][4]. 반향제거기에 사용되는 동시통화검출기는 그림 1과 같이 나타낼 수 있다. 이 때, $x(n)$ 은 원단화자의 음성신호를 나타내며 반향제거기의 기준입력신호(reference input signal)가 된다. $f(n)$ 은 반향경로를 통과하여 발생한 반향신호를 나타내며, $w(n)$ 과 $d(n)$ 은 각각 배경잡음과 근단화자의 음성신호를 나타낸다. 마이크로폰 입력신호인 $y(n)$ 은 반향제거기의 주입력신호(primary input signal)가 된다. 동시통화검출기는 주입력신호에 근단화자의 음성신호가 포함되어 있는지를 검출하여 반향제거기의 적응을 멈추도록 제어신호(control signal)을 보내주어야 한다.

동시통화검출기는 동시통화의 시작구간을 빠르게 검출하여야 한다. 또한, 단일통화를 동시통화로 잘못 판단하지 않도록 정확하게 동작하여야 한다. 반향제거기가 수렴한 이후

접수일자: 2014년 10월 16일

심사(수정)일자: 2014년 11월 16일

게재확정일자: 2014년 11월 17일

† Corresponding author

본 연구는 2013년도 한경대학교 교비 해외 파견 연구비의 지원에 의한 것임.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

에 동시통화를 검출하는 방법에 대해서는 많은 연구가 진행되어 왔으며, 반향제거기의 오차신호를 사용하여 동시통화를 잘 검출할 수 있다[5][6]. 그러나, 반향제거기가 수렴하기 전에도 근단화자신호에 의해 반향제거기의 적응필터의 계수가 발산할 수 있기 때문에 수렴전에도 동시통화를 검출하는 것이 필요하다.

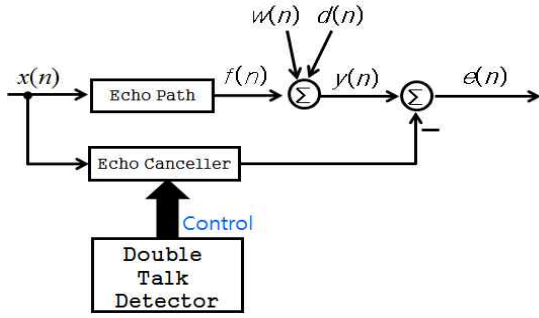


그림 1. 동시통화검출
Fig. 1. Double Talk Detection

반향제거기가 수렴하기 전에도 동시통화를 검출할 수 있는 방법은 Geigel 검출법[7], LPCCC(Linear Predictive Coding Cepstral Coefficients)[8] 등이 제안되었으며 최근에는 LPC 필터를 통과시킨 신호의 레벨을 사용한 동시통화 검출법이 상대적으로 우수한 방법으로 제안되었다[9]. 본 논문에서는 [9]의 동시통화검출법에서 발생할 수 있는 오검출 즉, 단일통화를 동시통화로 잘못 판단하는 문제점을 제시하고 이를 개선할 수 있는 방법을 제안하고자 한다. 또한, 주입력신호와 주입력신호를 LPC 필터에 통과시킨 신호의 상대적인 크기비와 관련 있는 가변적인 문턱치를 사용하여 동시통화 검출 지연시간을 줄일 수 있는 방법도 제안하고자 한다.

2. 기존의 동시통화 검출법

[9]에서 제안한 동시통화검출법은 그림 2와 같이 나타낼 수 있다. LPC 필터 계수를 기준입력신호 $x(n)$ 으로부터 추정한다. 주입력신호 $y(n)$ 을 이 LPC 필터에 통과 시켜 $z(n)$ 을 얻게 된다. 이 신호의 크기를 문턱치와 비교하여 문턱치 보다 큰 경우를 동시통화로 검출하는 방법이다. 문턱치는 통화초기 일정 시간동안 $z(n)$ 크기의 최대값을 측정하고 이 값에 일정 마진을 두어 설정한다.

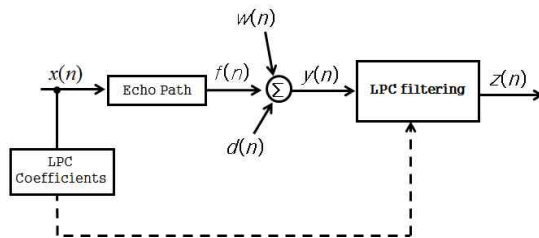


그림 2. 기존의 동시통화검출법
Fig. 2. Conventional Double Talk Detection

이 방법은 [8]에 비해 동시통화를 빠르게 검출할뿐만 아니라 반향경로의 변화를 동시통화로 오인하여 반향제거기의 필터계수적용을 멈추는 일이 발생하지 않는 장점이 있다.

그러나, 다음과 같은 문제점이 있다. 첫째, 문턱치를 설정하는 방법이다. LPC 필터링된 신호의 크기를 일정시간동안 관찰하여 최대값에 일정 마진을 두어 동시통화를 검출한다. 그러므로 처음 일정시간동안 동시통화 검출을 수행하지 못한다. 둘째, 민감하게 동시통화를 검출하기 위하여 문턱치의 마진을 작게 설정하는 경우 단일통화를 동시통화로 잘못 판단하는 오검출이 발생한다. 셋째, 문턱치의 마진을 크게 설정하는 경우 동시통화 검출 지연시간이 커져 반향제거기의 필터계수가 발산하는 경우가 발생한다.

그림 3은 단일 통화를 동시통화로 오검출하는 경우를 나타내고 있다. 모든 신호는 음성신호로서 샘플링 주파수는 8kHz 이다. 여기서 반향경로는 [10]에서 제시한 임펄스응답을 사용한 것이다. (a)는 주입력신호 $y(n)$ 을, (b)는 근단화자신호 $d(n)$ 을 나타낸다. 동시통화는 4초에서 6초 까지 2초 동안 진행된 경우이다. (c)는 $z(n)$ 을 각 프레임별로 신호의 크기 평균값을 나타낸다. 한 프레임은 10ms 이다.

통화의 시작부터 처음 3초 동안 각 프레임별로 평균 크기 값을 계산하고 최대평균치를 갖는 프레임의 값에 20%의 마진을 두었다. 즉, 1.2배로 문턱치를 설정하였다. 동시통화는 4초부터 시작되지만 3.6초 부근에서 이미 동시통화로 오검출하였다. 그러므로 실제로 단일통화구간에서 필터계수를 적용시켜야 함에도 불구하고 적용을 중지하여 수렴속도가 늦어진다.

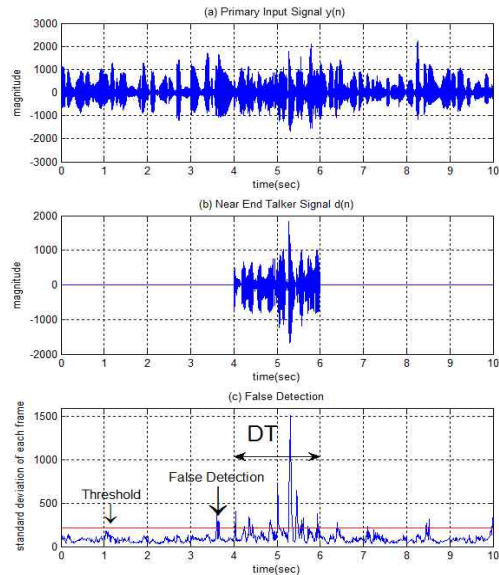


그림 3. 기존방법의 오검출 (a) 주입력신호 (b) 근단화자신호 (c) 오검출

Fig. 3. False Detection of the Conventional Method (a) Primary Signal (b) Near End Talker Signal (c) False Detection

3. 제안한 동시통화 검출법

기존의 방법에서 문턱치에 마진을 적게 설정하는 경우 단일통화를 동시통화로 잘못 판단할 확률이 높아진다. 반면, 마진을 크게 하여 문턱치를 설정하는 경우에는 오검출은 감소하지만 동시통화 검출 지연시간이 커지게 된다. 본 논문에서는 이를 개선하기 위하여 그림 4와 같은 동시통화 검출법을 제안한다.

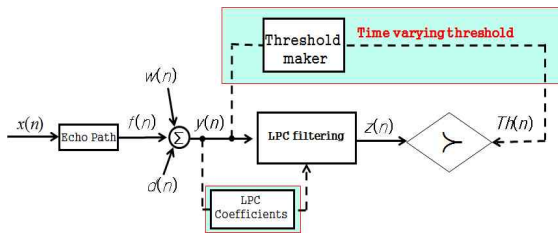


그림 4. 제안한 동시통화검출법

Fig. 4. Proposed Double Talk Detection Method

1. LPC 필터링

제안한 방법은 주입력신호 $y(n)$ 으로 부터 LPC 필터 계수를 추정한다. LPC 필터링은 다음과 같은 전달함수를 갖는다.

$$H(z) = 1 - A(z) = 1 - \sum_{j=1}^P a_j z^{-j} \quad (1)$$

시간영역에서 이를 표현하면 다음과 같다 .

$$z(n) = y(n) - \sum_{j=1}^P a_j y(n-j) \quad (2)$$

여기서 a_j 는 LPC 필터 계수를 나타내며, P는 필터 차수를 나타낸다[11][12].

그림 5는 기존의 방법으로 부터 구한 LPC 필터 계수를 사용하여 필터링한 경우의 필터링 전과 후의 신호레벨을 나타낸다. 이때 LPC 필터계수는 프레임 단위로 추정하게 된다. x축은 필터링 전의 프레임단위의 신호의 평균크기 레벨을 나타내며, y축은 필터링 후의 신호에 대한 값이다. 대상은 8kHz로 샘플링된 음성신호를 한 프레임은 10ms로 한 경우이다. 18,000 개의 프레임에 대한 신호레벨의 상관도표를 나타낸다. 붉은 점선으로 표시된 직선은 $y=x$ 형태의 직선으로서 직선 윗부분에 있는 점들은 필터링 후의 신호레벨이 필터링 전의 신호 레벨보다 커지게 되는 경우이다. 즉, LPC 필터링 이득이 0dB 보다 작게 되는 경우이다. 그림으로 부터 이러한 프레임이 많이 발생함을 알 수 있다. 이것은 기준입력신호로 부터 LPC 필터계수를 추정하는 것이 문제를 일으키고 있으며, 문턱치를 설정하는데 어려움을 주고 있는 원인으로 판단할 수 있다.

그림 6은 주입력신호 $y(n)$ 으로부터 LPC 필터 계수를 추정하여 필터링한 경우의 신호레벨 상관 도표이다. LPC 필터링 이득이 0dB 보다 작게 되는 경우가 주입력신호의 레벨이 100보다 큰 경우에는 발생하지 않고 100 이하에서만

매우 드물게 발생한다. 그러므로 LPC 필터계수 추정을 위하여 주입력신호를 사용하는 것이 훨씬 더 안정적으로 LPC 필터링 이득을 얻을 수 있어 문턱치 설정을 용이하게 할 수 있다.

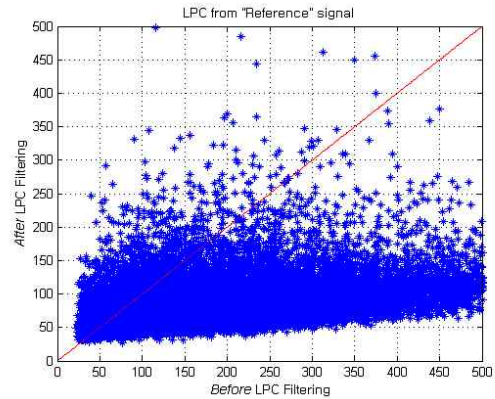


그림 5. 상관도표 : $x(n)$ 에 의한 필터계수

Fig. 5. Scatter Diagram : Filter Coefficients from $x(n)$

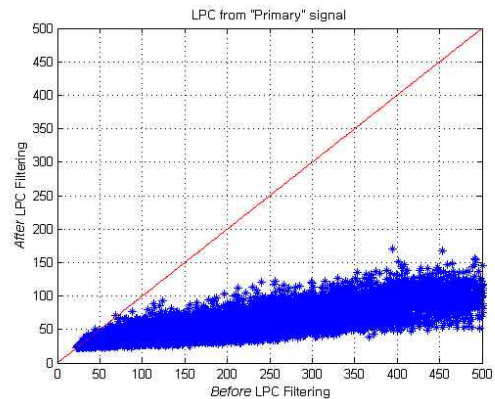


그림 6. 상관도표 : $y(n)$ 에 의한 필터계수

Fig. 6. Scatter Diagram : Filter Coefficients from $y(n)$

2. 문턱치

그림 6으로 부터 LPC 필터를 통과 하기 전의 프레임단위 신호레벨의 평균치에 따라 LPC 필터링 이득이 일정 범위 이하로 내려가지 않는다는 것을 알 수 있기 때문에 가변적인 문턱치를 설정할 수 있다. 그림 7은 본 논문에서 제안한 문턱치 설정 방법이다. 프레임 신호레벨이 100 보다 작은 경우와 큰 경우로 나누어 문턱치에 해당하는 일차함수를 달리한다. LPC 필터링 전의 신호인 주입력신호의 레벨이 100보다 작은 경우에는 (3), 100보다 큰 경우에는 (4)와 같은 문턱치 함수를 갖게 된다.

$$Th(i) = a \times (y(i) - 100) + 100 \quad (3)$$

$$Th(i) = b \times (y(i) - 100) + 100 \quad (4)$$

이때, $Th(i)$ 는 i 번째 프레임의 문턱치를 나타내며, $y(i)$ 는 주입력신호의 i 번째 프레임의 신호의 크기 평균 크기를 나타낸다. (3)과 (4)는 점(100,100)을 지나는 일차함수의 형태를 가진다. (3)과 (4)에서 기울기는 그림 7의 모든 점들이 각 일차함수 보다 아래에 있도록 정하게 된다. 실제 음성신호의 주입력신호로 부터 추정한 LPC 필터 계수를 사용한 필터링 이득의 최저값을 기준으로 문턱치를 설정하기 때문에 단일통화를 동시통화로 잘못 판단하는 오검출률을 매우 낮게 할 수 있다. 기존의 방법에서는 일정한 문턱치를 가지는 반면 제안한 방법은 각 프레임별로 다른 값을 갖는 가변 문턱치를 사용하게 되어 빠르게 동시통화를 검출할 수 있다.

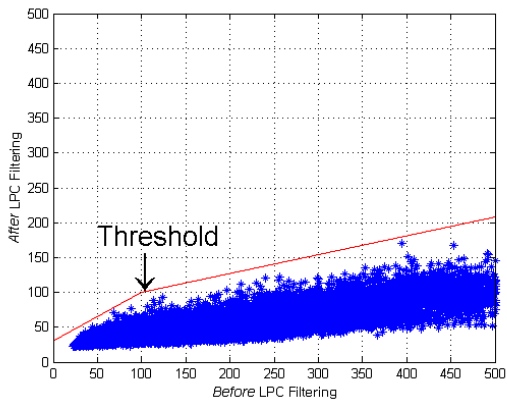


그림 7. 제안한 문턱치
Fig. 7. Proposed Threshold

4. 실험 결과

제안한 방법이 기존의 방법에 비해 성능이 우수함을 확인하기 위하여 모의실험을 수행하였다. 실험에 사용된 음성신호는 실제 음성신호로서 8 kHz로 샘플링된 신호이다[13]. 녹음된 음성신호 데이터베이스에서 그림 3(a)와 같이 일정 구간을 선택하여 기준입력신호로 사용하였다. 동시통화구간에 사용된 음성신호는 그림 3(b)와 같이 2초 동안의 근단화자 음성신호를 사용하였다. 반향경로는 [10]에서 권고하고 있는 표준 임펄스응답을 사용하였다. 반향신호 $f(n)$ 과 배경잡음 $w(n)$ 의 비율은 20dB이다[14][15]. 동시통화검출은 프레임 단위로 수행되며 한 프레임은 10ms로서 80 샘플에 해당하며 1초는 100 프레임으로 구성된다. 성능 평가는 단일통화를 동시통화로 잘못 판단하는 오검출률과 동시통화 검출지연 시간에 대해 비교하였다.

1. 오검출률

기존의 방법과 제안한 방법의 오검출률을 평가하였다. 기존의 방법은 처음 3초 동안은 문턱치 설정을 위한 구간으로 하고 이후의 10초 동안의 단일통화구간에서 동시통화로 오검출 하는 비율을 계산하였다. 음성신호의 시작부분을 다르게 하여 100번의 시뮬레이션을 수행하였다. 그러므로 각 문턱치 값마다 10만 프레임에 대해 오검출률을 실험하였다. 그림 8은 각 문턱치 별로 오검출률을 나타내는 그래프이다. 처음 3초 동안의 최대치에 0%의 마진을 부여한 문턱치부터

10%씩 증가하여 100% 마진을 부여한 경우까지 실험하였다. 문턱치를 0%로 하는 경우 오검출률이 0.6%로서 1.7초마다 오검출이 발생하여 실제로는 단일통화 구간이어서 적응필터의 계수적응이 필요한 구간임에도 불구하고 자주 계수적응을 멈추게 되어 반향제거기의 성능 저하를 가져오게 된다. 문턱치를 높게 가져갈수록 오검출률이 작아지게 된다. 100% 마진을 두는 경우 즉 문턱치를 3초동안의 최대치의 2배로 설정하는 경우에도 0.024%의 오검출률을 나타내고 있다. 즉 40초에 한번 비율로 단일통화를 동시통화로 잘못 판단하게 된다.

제안한 방법의 오검출률은 (3), (4)의 기울기를 어떻게 설정하는가에 따라 변화한다. a 는 작게 할수록, b 는 크게 할수록 오검출률이 적어지는 반면 동시통화검출 지연시간은 커지게 된다. 표 1은 기울기 a, b 에 따른 오검출률을 나타내고 있다. a 가 0.51 보다 같거나 적고 b 가 0.23 보다 같거나 크면 오검출이 전혀 발생하지 않는다. a 의 경우 0.75까지 증가시켜도 오검출률은 0.001% 수준이며, b 를 0.22까지 감소시켜도 오검출률은 0.001% 수준에 머무르게 된다.

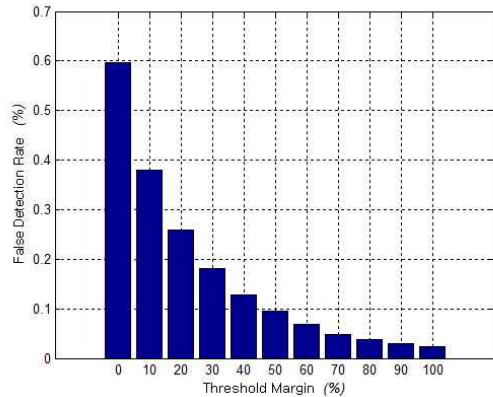


그림 8. 오검출률 : 기존의 방법
Fig. 8. False Detection Rate : Conventional Method

표 1. 제안한 방법의 오검출률

Table 1. False Detection Rate : Proposed Method

a	b	False Detection Rate(%)
0.51	0.25	0
0.52		0.001
0.75		0.001
0.51	0.24	0
	0.23	0
	0.22	0.001

2. 동시통화검출

그림 3과 같은 환경을 만들어 동시통화 검출 성능을 실험하였다. 동시통화는 2초 동안 존재하도록 하였다. 동시통화가 시작되는 시간을 다르게 하여 100번의 반복실험을 수

행하였다. 기존의 방법과 제안한 방법의 공정한 성능평가를 위하여 오검출률이 동일한 경우의 문턱치 파라미터를 먼저 조사하고 이를 사용하여 비교 평가하였다. 표 2는 문턱치 파라미터를 나타낸다.

표 3은 오검출률이 0.024%, 0.001%, 0% 인 경우의 동시통화 검출지연시간과 동시통화 검출 실패율을 비교 평가한 결과이다. 동시통화 검출에 실패한 경우는 검출지연시간을 2초로 계산하였다. 동시통화검출 지연시간의 경우 제안한 방법이 기존의 방법에 비해 작은 지연시간을 가지므로 기존의 방법의 성능을 개선할 수 있음을 알 수 있다. 오검출률이 0%인 경우 평균 1.36 초 정도 검출지연 시간을 단축시킬 수 있음을 볼 수 있다. 동시통화 검출 실패율의 경우도 제안한 방법이 기존의 방법에 비해 안정되게 낮은 값을 갖게 되어 성능 개선을 이루어냄을 알 수 있다.

표 2. 문턱치 파라미터
Table 2. The Threshold Parameters

Conventional Threshold Margin (%)	Proposed		False Detection Rate (%)
	a	b	
100	0.82	0.184	0.024
250	0.515	0.225	0.001
400	0.51	0.23	0

표 3. 동시통화 검출 : 지연시간과 실패율
Table 3. Double Talk Detection : Delay and Fail Rate

False Detection Rate (%)	Delay(sec)		Fail Rate(%)	
	Conventional	Proposed	Conventional	Proposed
0.024	0.93	0.29	9	1
0.001	1.51	0.37	44	1
0	1.75	0.39	69	1

5 결론

기존의 방법은 기준입력신호로부터 LPC 필터 계수를 추정하고 LPC 필터링된 주입력신호의 일정시간 동안 관측치의 최대값을 사용한 고정 문턱치를 사용하였다. 본 논문에서는 성능개선을 위하여 다음과 같은 개선 방안을 제시하였다. 첫째, 기준입력신호 대신에 주입력신호를 사용하여 LPC 필터 계수를 추정하는 것이다. 둘째, 주입력신호의 크기를 계산하고 그 크기 별로 주어지는 가변 문턱치를 사용하는 것이다. 주입력신호로부터 LPC 필터계수를 추정하는 방법이 기준입력신호로부터 추정하는 방법에 비해 더 큰 LPC 필터 이득을 제공하기 때문에 동시통화 검출을 위한 문턱치

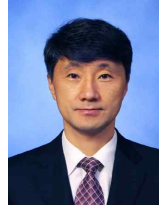
를 설정하는데 유리한 점을 제공해 준다. 주입력신호의 평균크기별로 가변적인 문턱치를 사용하기 때문에 보다 민감하게 동시통화를 검출할 뿐만 아니라 오검출률도 감소시킬 수 있다. 실제 음성신호를 사용한 실험결과로부터 오검출률 감소, 동시통화검출 실패율 감소, 동시통화 검출 지연시간 감소 측면에서 제안한 방법이 기존 방법의 성능을 개선함을 확인하였다.

References

- [1] K. Murano, S. Unagami and F. Amano, "Echo Cancellation and Applications", *IEEE Comm. Magazine*, pp. 49-55, Jan. 1990.
- [2] Seon Joon Park, Chom Gun Cho, Chungyong Lee and Dae Hee Youn, "Integrated Echo and Noise Canceller for Hands-Free Applications," *IEEE Trans. Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 49, no 36, pp. 188-195, March 2002.
- [3] Thomas Gänsler, Steven L. Gay, M. M. Sondhi, and Jacob Benesty, "Double-Talk Robust Fast Converging Algorithms for Network Echo Cancellation", *IEEE Trans. Speech Audio Processing*, vol. 8, no. 6, pp. 656-663, Nov. 2000.
- [4] Jiaquan Huo, Ka Fai Cedric Yu, Sven Nordholm, and Kok Lay Teo, "A robust transform domain echo canceller employing a parallel filter structure," *Signal Processing*, vol. 86, no.2, pp.3752-3760, 2006.
- [5] J. Benesty, D.R. Morgan, and J.H. Cho, "A new class of double-talk detectors based on cross-correlation", *IEEE Trans. Speech Audio Process*, vol.8, no.2, pp.168-172, Mar. 2000.
- [6] H. Ye and B.X. Wu, "A new double-talk detection algorithm based on the orthogonality theorem", *IEEE Trans. Commun*, 39, no. 11, pp. 1542-1545, Nov. 1991.
- [7] D. Duttweiler, "A twelve channel digital echo canceler," *IEEE Trans. Commun*, vol. com-26, no 5, pp. 647-653, May. 1978.
- [8] Jun-Eun Park, Yoon-Jae Lee, Ki-Hyeon Kim, Han-Seok Ko, "Speech Feature based Double-talk Detector for Acoustic Echo Cancellation," *Journal of Institute of Korean Electrical and Electronics Engineering*, vol. 2. pp. 132-139, 2009.
- [9] Jae Ha Yoo, Soo Chan Kim and Dong Yon Kim, "Double Talk Detection before the Convergence of Echo Canceller," *The Journal of The Institute of Internet, Broadcasting and Communication*, vol. 13, no 5, pp. 203-208, 2013.
- [10] ITU-T G.168 : Recommendations for Digital Network Echo Cancellers, International Telecommunication Union Telecommunication Standardization Sector, August, 2004.
- [11] A. M. Kondoz, *Digital Speech: Coding for Low Bit Rate Communication Systems*, 2nd Edition Wiley,

- 2004.
- [12] Thomas F. Quatieri, *Discrete-Time Speech Signal Processing Principles and Practice*, Prentice-Hall, 2002.
- [13] Kwee-Bo Sim, Chang-Hyun Park and Young-Hoon Joo, "Analyzing the Acoustic Elements and Emotion Recognition from Speech Signal Based on DRNN", *Journal of Korean Institute of Intelligent Systems*, vol.13, no. 1, pp. 45-50, 2003.
- [14] Sang-bae Jeong and Sung-doke Lee, "Multi-channel input-based non-stationary noise canceller for mobile devices", *Journal of Korean Institute of Intelligent Systems*, vol.17, no. 7, pp. 945-951, 2007.
- [15] Joung-Kook Seo and Hyung-Tai Cha, "Tone Quality Improvement Algorithm using Intelligent Estimation of Noise Pattern", *Journal of Korean Institute of Intelligent Systems*, vol.15, no. 2, pp. 230-235, 2005.
-

저 자 소 개



유재하(Jae-Ha Yoo)

1990년 : 연세대학교 전자공학과
공학사

1992년 : 연세대학교 전자공학과
공학석사

1996년 : 연세대학교 전자공학과
공학박사

1996년~2002년 : (주) LG전자
책임연구원

2002년~현재 : 국립 한경대학교 전기전자제어공학과 교수

관심분야 : 적응신호처리, 신호처리응용

E-mail : yjh@hknu.ac.kr