http://dx.doi.org/10.7236/JIIBC.2013.13.5.203
JIIBC 2013-5-25

반향제거기의 수렴전 동시통화검출

Double Talk Detection before the Convergence of Echo Canceller

유재하*, 김수찬**, 김동연**

Jae-Ha Yoo, Soo-Chan Kim, Dong-Yon Kim

요 약 본 논문에서는 반향제거기가 수렴하기 전에 동작할 수 있는 동시통화검출기의 성능 향상에 대한 방법을 제안하였다. 마이크로폰에 입력된 신호를 선형예측필터에 통과시키고, 이 필터링 된 신호를 사용하여 검출을 수행하게 된다. 선형예측필터 계수는 원단화자신호의 음성으로 부터 얻어진다. 단일통화 시에는 반향신호가 원단화자의 신호와 유사한 음성특성을 가지므로 필터링 된 신호가 작은 값을 갖게 된다. 그러나 동시통화가 발생한 경우에는 마이크로폰 신호에 다른 특성을 갖는 신호가 포함되어 있어서 파워가 줄어들지 않게 된다. 이를 사용하여 동시통화를 검출하게 된다. 음성신호를 사용한 모의 실험을 통하여 제안한 방법이 기존의 방법에 비해 동시통화를 빠르게 검출함으로서 성능이 우수함을 확인할 수 있었다.

Abstract In this paper, we proposed a performance improvement method of the double talk detector which can operate before the echo canceller converges. Microphone input signal is filtered by the linear prediction filter and this filtered signal is used for detection. The coefficients of the linear prediction filter are given by the far-end talker signal. During single talk, filtered signal has low power since the characteristics of the echo signal is similar with those of the far-end talker signal. But, during double talk, the filtered signal does not have low power because the signal of different characteristics is included in the microphone signal. Double talk is detected by this difference. Simulations using real speech signals verified that the proposed method outperformed the conventional methods.

Key Words : Double Talk Detection, Echo Canceller, Convergence, LPC, Adaptive Filter

Ⅰ. 서 론

반향은 음성통신에서 발신자의 음성이 일정시간 후에 되돌아오는 현상을 말하며 이러한 반향을 제거하는 장치를 반향제거기라 한다[1]. 동시통화검출기는 반향제거기에서 필요로 하는 기능적 요소를 말하며 자체적으로는 사용되지 않는다. 동시통화를 검출한 경우 반향경로를 모델링하는 적응필터의 계수적응을 중지하도록 명령한다. 동시통화 중에도 계수적응을 계속하는 경우 적응필터가 발산하여 반향제거 기능이 마비될 수 있기 때문에 동시통화검출기는 반향제거기에서 핵심적인 역할을 수행하는 부분이다[1].

반향을 제거하려면 그림 1과 같이 적응필터를 사용하여 반향을 제거하는 장치인 반향제거기가 필요하다. 여기서, $x(n)$ 과 $s(n)$은 원단화자의 음성신호와 근단화자의 음성신호를 나타낸다. $y(n)$은 원단화자의 음성신

*정회원, 한경대학교 전기전자제어공학과/IT융합연구소 (교신저자)
**정회원, 한경대학교 전기전자제어공학과/IT융합연구소
접수일자 : 2013년 9월 25일, 수정완료 2013년 10월 10일
게재확정일자 : 2013년 10월 11일

Received: 25 Sepetmber, 2013 / Revised: 10 October, 2013 / Accepted: 11 October, 2013
*Corresponding Author: yjh@hknu.ac.kr
Dept. of Electrical, Electronic and Control Engineering, Hankyong National University, Korea

호가 반향경로를 거쳐 생성되는 반향을 나타내며, $v(n)$은 마이크폰에 유입되는 배경잡음을 나타낸다. 적응필터에서 추정한 반향경로의 임펄스응답 $\hat{h}$을 사용하여 추정 반향신호 $\hat{y}(n)$ 을 만들고 이를 마이크로폰 입력신호 $m(n)$으로 부터 빼줌으로서 근단(near-end)에 있는 화자의 음성신호인 $s(n)$만을 원단(far-end)으로 보냄으로써 원활한 통화가 이루어지도록 하는 것이다.

반향신호를 추정하기 위하여 적응 필터는 반향경로를 추정하게 되는데 원단에 있는 상대방 통화자의 음성신호인 $x(n)$과 오차신호 $e(n)$을 사용하여 오차신호 속에 포함된 반향신호의 크기가 작아지는 방향으로 적응필터의 계수를 조정하게 된다. 그런데 오차신호에 근단화자의 신호 $s(n)$이 존재하는 통화구간에서 적응필터의 계수를 조정하는 경우에는 반향경로와 무관한 방향으로 필터 계수가 발산하게 되어 반향제거기로서의 기능을 수행하지 못하게 된다.

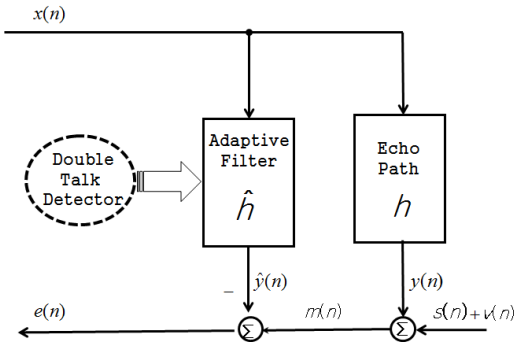

그림 1. 반향제거기와 동시통화검출기
Fig. 1. Echo canceller and double talk detector

이러한 적응필터의 발산을 막기 위해서는 원단화자와 근단화자가 동시에 존재하는 동시통화구간을 검출하여 적응필터의 계수적응을 중지하도록 하는 동시통화검출기가 필요하다.

동시통화검출기는 빠르게 동시통화를 검출하여야 할 뿐만 아니라 반향경로의 변화를 동시통화로 잘못 판단하지 않아야 하는데 이는 반향경로가 변화한 경우에는 필터계수적응이 필요한 경우이기 때문이다.

동시통화 검출을 위해서 크기 비교법[2], 상호상관계수 사용법[3][4], 음성신호의 특성을 이용한 방법[5][6] 등 많은 방법이 제시되어 왔으며, 상호상관계수를 사용하는 방법이 상대적으로 검출능력이 우수하며 계산량 측면에서도 실용적이어서 대부분의 연구가 이에 관한 것이다. 이 방법은 적응필터가 수렴한 후라는 가정을 하고 오차신호를 사용하여 민감하게 동시통화를 검출하게 된다. 그러나 반향제거기가 수렴하기 전에는 이 가정을 사용할 수 없는 문제점이 있다.

본 논문에서는 반향제거기가 수렴하기 전에 동작할 수 있는 동시통화검출 방법을 연구하고자 한다. 즉, 오차신호를 사용할 수 없고 원단신호 $x(n)$과 마이크로폰 신호 $m(n)$이 사용 가능한 환경에서 성능이 우수한 동시통화검출법을 제안하고자 한다.

II. 동시통화 검출기

다음의 식 (1)과 같이 마이크로폰 입력신호에서 추정된 반향신호를 빼주어 오차신호만을 전송함으로써 반향제거가 이루어지게 된다.

$$e(n) = m(n) - \hat{y}(n) = m(n) - \hat{h}^T(n)X(n) \quad (1)$$

적응필터의 계수 $\hat{h}(n)$을 구하는 방법은 많은 방법들이 연구되어 왔다[1]. 본 논문에서는 가장 널리 사용되는 NLMS(Normalized Least Mean Square) 알고리즘을 사용하기로 하며 계수 적응식은 다음과 같다[1].

$$h_{n+1} = h_n + \frac{\mu}{X_n^T X_n + \delta} X_n e(n) \quad (2)$$

동시통화검출기는 원단신호 $x(n)$, 근단신호 $m(n)$, 오차신호 $e(n)$을 사용하여 동시통화가 일어난 것을 빠르고 정확하게 검출하여 적응필터의 계수적응을 멈추도록 하는 것이 목적이다. 단일통화(single talk) 시와 동시통화 시에 차이가 나는 특정 값을 사용하여 문턱치(threshold)를 설정하고 단일통화시의 값에서 문턱치를 넘어 동시통화 시의 값으로 변화된 경우 동시통화로 판단한다[2].

Geigel 검출법은 신호의 크기를 비교하는 방법중 가장 최초의 검출법으로서 식 (3)과 같이 적응필터의 기준 입력신호벡터에 있는 값 중 가장 큰 절대값의 1/2 보다 마이크로폰 입력신호의 절대값이 큰 경우를 동시통화가 발생한 것으로 판단하는 방법이다.

$$|m(n)| > \frac{1}{2}\max\{|x(n)|, |x(n-1)|, \cdots, |x(n-N+1)|\} \quad (3)$$

여기서 N은 적응필터의 탭수를 나타낸다. 이 방법은 반향경로를 통하여 6dB 이상 신호가 감쇄한다는 것을 가정하여 만든 매우 구현이 단순한 방법이다. 그러나, 반향경로 마다 반향신호의 감쇄율이 다름을 반영하지 못하는 문제점이 있다. 이의 검출 능력을 향상시키기 위하여 적응필터가 수렴한 것을 가정하여 오차신호를 사용한 크기 비교법이 제안되었다[5].

상호상관 계수를 사용하는 방법은 [3]에서 처음 발표된 후로 많은 방법들이 제안되어 왔으며 [4]에서는 다음과 같은 방법으로 동시통화를 검출한다.

$$\xi = 1 - \frac{r_{me}}{\sigma_m^2} > Th \quad (4)$$

여기서 r_{me}은 마이크로폰신호와 오차신호와의 상관도를, σ_m^2 은 마이크로폰 신호의 분산을 나타낸다. Th는 동시통화를 검출하는 문턱치를 나타낸다. 그러나 이와 같이 적응필터의 수렴을 가정한 방법은 적응필터가 수렴하기 전에는 사용할 수 없는 문제점이 있다.

음성신호의 특징 파라메터를 사용하는 방법으로는 근단화자의 존재 유무에 따라 마이크로폰 입력신호의 반사계수의 차이가 생기는 것을 사용하는 방법[5], LPCCC(Linear Predictive Coding Cepstral Coefficients)를 사용하는 방법 등이 제안되었다[6]. 원단화자의 신호가 반향경로를 통하여 마이크로폰에 유입되기 때문에 원단신호와 마이크로폰 신호 간에는 유사한 음성특성을 갖는다는 것을 사용하는 방법이다. 다음은 LPCCC 검출법이다.

$$\xi = \sum_{i=1}^{P} (c_x(i) - c_m(i))^2 > Th \quad (5)$$

여기서 $c_x(i)$와 $c_m(i)$는 각각 $x(n)$과 $m(n)$의 i 번째 LPCCC를 나타낸다. P는 음성신호의 LPC 차수를 의미한다. 동시통화검출은 샘플마다 수행되지 않고 프레임단위로 수행된다. 단일통화시에는 $c_x(i)$와 $c_m(i)$의 차이 값인 ξ가 작은 값을 갖는 반면, 동시통화가 일어난 경우에는 $c_m(i)$에 원단신호와 무관한 신호가 포함되게 되어 $c_x(i)$와는 다른 값을 갖게 되어 ξ이 커지게 된다. 이 때 문턱치는 단일통화시의 ξ값에 따라 실험적으로 설정하게 된다. 이 방법은 오차신호를 사용하지 않아 수렴전 동시통화 검출법으로 사용될 수 있으나 잡음이 있는 경우에 대한 결과를 제시하지 않고 있다[6].

Ⅲ. 제안한 동시통화검출법

Geigel 계열의 크기 비교법은 동시통화가 발생하면 근단신호에 근단화자의 음성신호가 더해져 크기가 커지는 것을 사용하는 방법이다. 그러나 근단화자신호가 반향신호에 비해 확실하게 크지 않은 경우 이를 구별할 수 없는 문제점이 있다. 본 연구에서는 이를 개선하기 위하여 음성신호의 특성을 나타내는 LPC 계수를 사용한 선형예측필터를 통과한 신호를 사용한 검출법을 제안한다. 동시통화검출법은 그림 2와 같다.

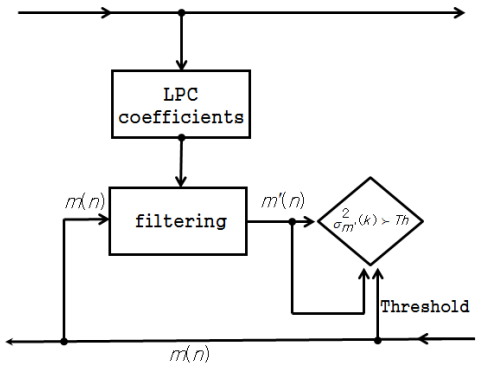

그림 2. 제안한 동시통화검출기
Fig. 2. Proposed double talk detector

여기서 $m'(n)$은 $m(n)$이 LPC 필터를 통과한 신호이다. LPC 계수는 수신된 원단화자신호로 부터 구하게 되며 반향경로를 통한 지연시간을 고려하여 이전 프레임으로 부터 계산된 것을 사용한다. 디지털 음성통신의 경우 음성디코더 단에서 계수가 주어지기 때문에 LPC계수 추정을 위한 별도의 계산량이 필요하지 않다. 문턱치는 이전 프레임에 계산된 최대치에 마진을 두어 정하게 된다.

그림 3은 마이크로폰신호 $m(n)$에 4초부터 5초 까지 1초 동안 동시통화가 발생한 경우를 나타내고 있다. 근단화자신호의 레벨은 반향신호의 레벨과 동일한 값을 갖는

다. (a)는 $m(n)$의 파워를, (b)는 LPC 필터를 통과한 신호인 $m'(n)$의 파워를 나타낸다.

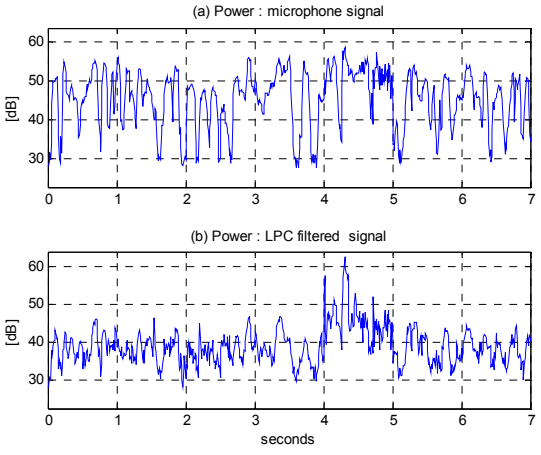

그림 3. 파워 (a) 마이크로폰신호 (b) 필터링된 신호
Fig. 3. Power (a) microphone signal (b) filtered signal

마이크로폰 신호 $m(n)$의 파워는 동시통화가 발생하여도 단일통화시와의 차이가 매우 적어 적당한 문턱치를 설정하기가 매우 어렵다. 그러나 LPC 필터링된 신호는 단일통화시에는 원단신호로 부터 얻어진 LPC 계수를 사용하여 필터링 됨으로서 그 파워가 많이 감소하게 된다. 그러나, 동시통화가 발생한 구간에서는 감소량이 적어지게 되어 단일통화시와 매우 다른 특성을 나타낸다. 그러므로 단일통화시의 최대값에 마진을 더한 문턱치를 설정하고 이보다 큰 값이 발생한 경우를 동시통화로 판단하면 된다.

Ⅳ. 실험 결과

제안한 동시통화검출법의 성능을 평가하기 위하여 모의실험을 수행하였다. 사용된 신호는 실제 음성신호이며 8 kHz로 샘플링 된 신호이다. 반향경로의 임펄스응답은 국제표준 권고안에서 제시한 모델을 사용하였으며 길이는 16ms이다[7]. 반향신호 대 잡음비는 20dB이다. 비교 대상으로 삼은 방법들은 반향제거기가 수렴전에도 동작할 수 있는 방법으로서 Geigel 검출법과 LPCCC 검출법이다. 음성신호의 한 프레임은 10ms로 하였다. 동시통화검출기는 동시통화를 빠르고 정확하게 검출하여야 할 뿐 아니라 반향경로의 변화를 동시통화로 오인하지 않아야 한다. 그러므로 모의실험은 각 방법에 대해 반향경로 변화시의 성능과 동시통화 검출 성능을 비교하였다.

1. 반향경로 변화시 성능

Geigel 방법은 반향경로의 형태가 변화하여도 영향을 받지 않는 방법이며, LPCCC 방법과 제안된 방법에 대한 실험 결과는 그림 4와 같다. 반향경로의 변화는 2초에서 발생하였다. LPCCC 방법과 제안한 방법 모두 반향경로의 변화에 따라 검출값이 변하지 않는 것을 관찰할 수 있다. 그러므로 세가지 방법 모두 반향경로의 변화를 동시통화로 잘못 판단하는 경우는 없다고 평가된다.

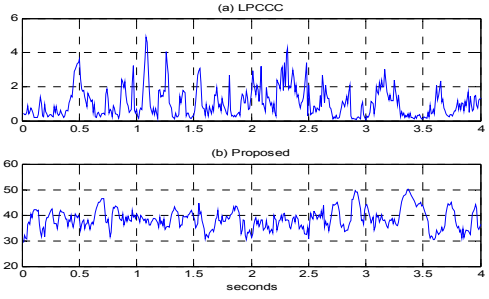

그림 4. 반향경로 변화 (a) LPCCC (b) 제안한 방법
Fig. 4. Echo path change (a) LPCCC (b) proposed method

2. 동시통화 검출 성능

동시통화 검출 성능을 비교하기 위하여 그림 5와 같은 신호를 사용하였다. 3초에서 4초까지 1초 동안 근단화자 신호가 존재한다. 문턱치는 단일통화시의 최대값에 마진을 두어 설정하였으며 LPCCC 방법은 6, 제안한 방법은 51dB 이다.

성능평가는 실제 반향경로의 임펄스응답 h와 적응필터의 임펄스응답 $\hat{h}$과의 차이 값인 misalignment로 표현하였다[8][9]. 실제 임펄스응답에 가까울수록 작은 값을 갖는다. 동시통화가 검출된 경우 일정기간 동안 적응필터의 계수적응을 중단하게 되는데 이를 hangover[2]라 하며 본 논문에서는 1초로 설정하였다.

그림 6은 근단화자신호 $s(n)$과 반향신호 $y(n)$의 비가 0dB인 경우의 성능을 나타낸다. Geigel 방법과 LPCCC 방법은 제안한 방법에 비해 늦게 동시통화를 검

출하여 계수적응을 중단시켰기 때문에 필터계수가 근단화자신호에 의해 많이 오염되어 misalignment가 커져버렸다. 특히, LPCCC의 경우 매우 늦게 동시통화를 검출하여 적응필터의 임펄스응답이 매우 왜곡되었다.

그림 7은 LPCCC 방법과 제안한 방법의 검출치를 나타낸다. 각 방법에서 붉은색으로 표현된 것이 문턱치를 나타낸다. LPCCC의 경우 3.8 초 부근에서나 동시통화를 검출함을 알 수 있다. 이는 배경잡음에 의해 동시통화구간에서도 검출값이 크게 증가하지 않았기 때문으로 판단된다. 반면, 제안한 방법은 동시통화가 발생한 이후 빠르게 동시통화를 검출하여 계수적응을 멈추어 기존 방법에 비해 우수한 성능을 가질 수 있었다. 동시통화 구간에서 제안한 방법은 Geigel 검출법 대비 4dB 성능향상을 나타냄을 볼 수 있다.

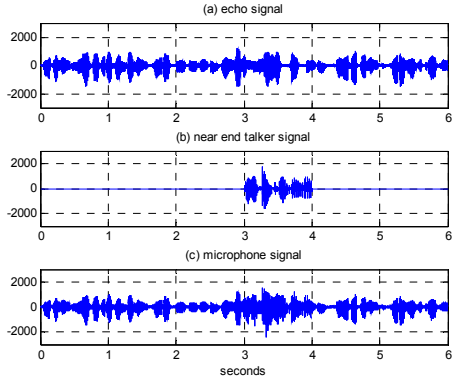

그림 5. 신호 (a) 반향 (b) 근단화자 (c) 마이크로폰
Fig. 5. Signals (a)echo (b) near-end talker (c) microphone

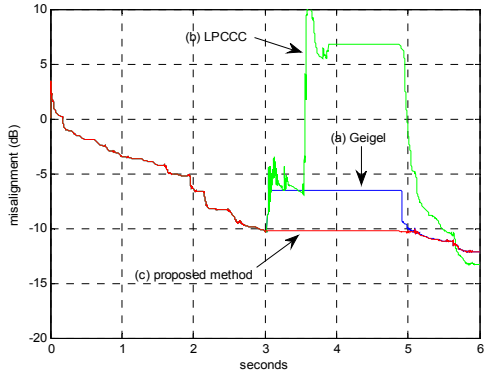

그림 6. 성능 비교
(a) Geigel (b) LPCCC (c) 제안한 방법
Fig. 6. Performance comparison
(a) Geigel (b) LPCCC (c) proposed method

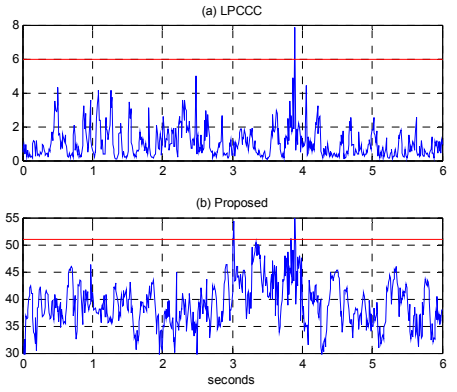

그림 7. 동시통화 검출 값
(a) LPCCC (b) 제안한 방법
Fig. 7. Double talk detection value
(a) LPCCC (b) proposed method

V. 결론

본 논문에서는 반향제거기가 수렴하기 전에 동작할 수 있는 성능이 향상된 동시통화 검출법을 제안하였다. 제안한 방법은 마이크로폰 신호를 LPC 필터에 통과 시킨 신호의 단일통화 시와 동시통화 시의 파워 변화를 이용하여 동시통화를 검출하는 방법이다. LPC 필터계수는 이전 프레임의 원단화자 신호로 부터 얻어진다. 음성신호를 사용한 모의실험을 통하여 제안한 방법이 기존의 Geigel 검출법과 LPCCC 검출법보다 우수한 동시통화 검출 성능을 나타냄을 확인하였다.

References

[1] J. Benesty, T. Gänsler, D.R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*, Springer, Inc. New York, 2001.

[2] D. Duttweiler, "A twelve channel digital echo cancer, IEEE Trans. Commun, vol. com-26, no 5, pp. 647-653, May. 1978.

[3] 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.

[4] M. Iqbal, J. Stokes and S. Grant, "Normalized

double-talk detection based on microphone and AEC error cross-correlation", Proc. IEEE Int, Conf. Multimedia Expo, Jul. 2007. pp 360-363.

[5] Jae-Ha Yoo and Sung-Ho Cho, "A New Double Talk Detector Using The Lattice Predictors For An Acoustic Echo Canceller," Proc, IEEE TENCON'97 Speech and Image Technologies for Computing and Telecommunication, vol. 2. pp. 1542-1545, Nov, 1997.

[6] 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.

[7] ITU-T G.168 : Recommendations for Digital Network Echo Cancellers, International Telecommunication Union Telecommunication Standardization Sector, August, 2004.

[8] 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.

[9] Jae Ha Yoo, "Tracking Performance Improvement of the Double-Talk Robust Algorithm for Network Echo Canceller," The Journal of th Institute of Webcasting, Internet and Telecommunication, vol. 12. No. 1, pp. 195-200, Feb, 2012.

Acknowledgement

"본 연구는 한경대학교 2012년도 학술연구조성비의 지원에 의한 것임"

저자 소개

유 재 하(정회원)

- 1990년 : 연세대학교 전자공학과 (공학사)
- 1992년 : 연세대학교 대학원 전자공학과 (공학 석사)
- 1996년 : 연세대학교 대학원 전자공학과 (공학 박사)
- 2002년 ~ 현재 : 한경대학교 전기전자제어공학과 교수

<주관심분야 : 신호처리, 잡음제거 등>

김 동 연(정회원)

- 1986년 : 연세대학교 전자공학과 (공학사)
- 1988년 : 연세대학교 대학원 전자공학과 (공학 석사)
- 1995년 : 연세대학교 대학원 전자공학과 (공학 박사)
- 1996년 ~ 현재 : 한경대학교 전기전자제어공학과 교수

<주관심분야 : 통신, 네트웍>

김 수 찬(정회원)

- 1994년 : 인제대학교 의용공학 (공학사)
- 1998년 : 연세대학교 대학원 생체공학 (공학 석사)
- 2003년 : 연세대학교 대학원 생체공학 (공학 박사)
- 2004년 ~ 2011 : 한경대학교 바이오융합기술대학원 교수
- 2012년 ~ 현재 : 한경대학교 전기전자제어공학과 교수

<주관심분야 : 의공학, 생체신호처리>