

이종 기기 간 음성통신을 위한 자동전환장치의 구현

류창국* · 이배호**

Implementation of the automatic switching device for the voice communications between heterogeneous devices

Chang-Guk Lew* · Bae-Ho Lee**

요 약

무전기의 음성통신은 PTT(: Push To Talk)를 이용한 반이중(half-duplex) 방식으로, 송신 시 단일 통화선로를 점유한다. 전화와 무전기간, UHF와 VHF 간의 인터페이스와 같이 서로 다른 이종 장치 간 음성통신을 위해서는 두 장치간의 자동전환장치가 요구되고, 이 장치는 입력 신호로부터 전송해야 할 음성을 검출하는 음성전환장치의 성능에 따라 전송되는 음성신호의 손실여부에 많은 영향을 받는다. 기존방식은 단순 입력신호의 크기 즉, 에너지 레벨을 통해 기준을 정함으로써 잡음에도 반응하는 문제점을 지니고 있다. 본 논문에서는 음성신호처리기법을 이용하여 입력된 신호가 음성임을 판별함으로써, 이종 기기 사이의 음성을 자동으로 전달하는 장치를 구현하였다. 이를 통해 음성 자동전환장치의 성능향상을 확인하였고, 이종 기기 간 음성 손실 없는 전송을 수행할 수 있었다.

ABSTRACT

A radio is a half-duplex voice communication method using the PTT(: Push To Talk), occupy a single line calls during transmission. As an interface between the telephone and the radio, UHF and VHF, for voice communication between the different heterogeneous devices, A device automatically switches between the two devices is required. Therefore, in accordance with the performance of the voice switching apparatus for detecting a voice to be transmitted from an input signal, loss of the audio signal to be transmitted is subjected to Significant influence. Conventional method has the problem responding to noise by setting the level through simple means of amplitude of input signal, in other words, the energy level of the input signal. This paper, by using the audio signal processing techniques, this discriminated what the voice is among the input signal and substantiated a device for the automatic voice transmission between heterogeneous devices. With this proposal, I was confirmed of improvement of performance in the automatic voice switching device, could perform loss-less transmission of voice between heterogeneous devices.

키워드

VOX, VAD, PTT, Heterogeneous RADIO
음성 동작, 음성 검출, 푸시 투 토크, 이종 무전기

1. 서 론

유선 및 이동통신전화 등의 일반적인 음성통신장치

와 VHF(: Very High Frequency), UHF(: Ultra High Frequency) 무전기와 같은 이종 기기 간의 통신을 하거나 또는 VHF와 UHF 무전기와와의 서로 다른 주파

* 전남대학교 전자컴퓨터공학과(ds1gor@naver.com)

** 교신저자 : 전남대학교 전자컴퓨터공학부

• 접수 일 : 2015. 10. 15
• 수정완료일 : 2015. 12. 13
• 게재확정일 : 2015. 12. 24

• Received : Oct 15, 2015, Revised : Dec 13, 2015, Accepted : Dec 24, 2015

• Corresponding Author : Bae-Ho Lee

Dept. Chonnam National University, Electronics & Computer Engineering
Email : bhlee@chonnam.ac.kr

수 대역을 사용하는 장치들 간에 음성통신을 하기 위해서는 두 무전기 사이에 음성을 자동전환 하기 위한 장치가 필요하다.

유선 및 이동통신전화 등의 전이중(full-duplex) 음성통신 방식은 음성신호가 양방향에서 동시에 이루어지고 있다. 하지만, 이종의 통신장치와 무전기간 음성통신을 하기 위해서는 무전기의 고유기능 중에 하나인 PTT(Push To Talk) 스위치 조작을 통해 음성을 전달해야만 한다. 전송할 음성신호를 전달하기 이전에 반드시 PTT가 ON상태를 유지해야만 무전기의 송신단이 출력준비가 완료되고, 이때 입력된 음성신호만 정상적으로 전송할 수 있다. 만약 음성입력 전에 PTT가 지연되면 음성이 전달되지 않거나 음성신호의 앞부분이 손실된 상태의 일부만 전달되어 수신 측에서는 손실된 음성을 수신하게 된다. 따라서 무전기 음성통신에서 전송음성의 성능을 개선하기 위해서는 PTT의 제어시점을 결정해야하고, 이는 수신측에 전송된 음성품질 성능의 큰 영향을 미친다.

일반적으로 입력된 음성신호를 이용하여 자동으로 PTT신호를 만드는 것을 VOX(Voice Operating Transmission)[1]라고 한다.

VOX 에서 정확한 음성전송 결과를 얻기 위해서는 입력된 신호로부터 음성구간결정과 음성판별 처리는 시스템성능의 중요한 역할을 차지하고 있다. 음성 구간 검출(VAD)은 음성 부호화기, 잡음제거, 음성인식기 등의 성능에 직접적인 영향을 미치는 원인으로 인해 정확한 음성구간검출[2-3]을 위한 연구는 지속적으로 이루어지고 있다. 조용한 환경에서 음성검출은 그 성능을 보장할 수 있지만, 다양한 잡음에 노출되어 있는 실(real)환경에서는 그 성능이 현격하게 저하된다. 이러한 요인으로 음성통신에 있어 음성 인터페이스를 활용하기 위해서는 다양한 잡음 개입에 대한 잡음제거 알고리즘들이 필요하다[3-4].

본 논문에서는 잡음상황에서의 음성신호 존재여부를 정확하게 감지하여 이를 전화기와 무전기의 인터페이스 또는 UHF와 VHF로 주파수가 서로 다른 두 개의 무전기 간 교신 가능한 이종 기기 간 음성통신 자동전환장치를 구현하고, 이를 위한 기존 시스템의 성능저해 요소를 도출 한 후, 이의 개선을 위한 시스템을 구현하고자 한다.

II. 관련연구

2.1 기존 아날로그 방식의 VOX

기존 아날로그 방식의 VOX는 그림1과 같이 음성신호가 입력되면, 그 신호의 크기만을 가지고 음성의 시작여부를 비교기를 이용하여 판단한다. 입력신호 대비 임계(Threshold)값과 상대 비교하여 크면 신호가 입력된 것으로 간주하고, 작으면 잠음으로 처리함으로써 주변 환경에 따라 오동작을 할 수 있다.

또한, 임계값 보다 큰 노이즈는 음성신호로 처리되기도 하고, 임계값 보다 작은 음성신호가 잠음으로 처리되기도 하는 에러를 발생하기도 한다. 이는 아마추어 무선장치의 헬즈프리 등에서는 사용할 수 있지만, 정확도를 요구하는 장치에서는 사용이 불가능하다.

시스템 설계에 있어 임계값은 하드웨어적으로 고정되어 있으면 변경이 어렵고[5] 변경을 한다 해도 이를 수동으로 변경해야 하기 때문에 다양한 곳에 적용하기에는 현실적으로 어려움이 많이 존재한다.

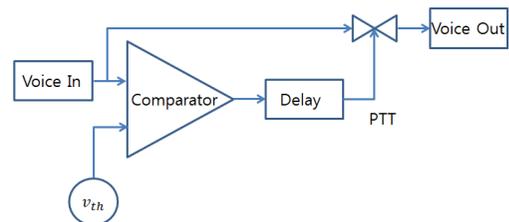


그림 1. 아날로그 VOX구현 예
Fig. 1 Example of analog VOX

2.2 AGC로 개선한 VOX

기존 아날로그 VOX의 성능의 일부를 개선하기 위해 그림2와 같이 AGC(Auto Gain Control)을 활용하기도 한다. 송신 무전기마다 음성 입력 크기가 다르고 사람마다 음성의 크기가 달라 정확한 음성을 감지하기 어렵다. 이를 위해 수신한 음성의 크기가 작을 경우 음량증폭을 위한 별도의 장치가 필요하다. 그러나, 큰 음량의 음성이 입력되는 경우 무전기는 수신측에 그대로 전달하면 음량 포화(Saturation) 현상이 발생한다. 따라서, 음성통신을 하는 무전기로부터 입력되는 서로 다른 소리의 음량크기를 일정한 크기로 전달할 필요가 있고 이를 위해 무전기 자동음량 조정 알고리즘(AGC)을 구현하기도 하였다[6].

이 방법은 음성신호가 입력이 되면, 음성신호의 크기를 항상 일정한 크기로 조정하여 이를 기준신호 크기와 비교하여 이를 음성신호의 입력인지 여부를 판단하도록 구성하는 방식이다. 이는 입력되는 음성신호의 크기가 일정하지 않아 임계값을 수동으로 변경하던 것을 개선하는 효과가 있다. 그러나, 입력된 신호를 단순히 증폭 또는 감쇄시키게 되면 동시에 입력된 신호와 분리가 불가능한 배경잡음 및 통신잡음과 같이 원하지 않은 잡음도 같이 증폭 또는 감쇄되기 때문에 PTT의 정확한 동작을 보장할 수 없다. 이는 입력되는 신호가 사람의 음성인지, 주변잡음인지, 통신잡음인지를 판단할 수 없어 단순히 기준신호 크기와 비교를 통해 입력된 신호와의 비교를 하는 것은 성능적인 한계를 가지고 있어서 활용이 제한적이다.

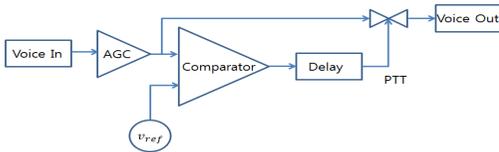


그림 2. AGC VOX구현 예
Fig. 2 Example of AGC VOX

2.3 입력 음성신호의 초기 손실

무전기는 출력장치가 안정화 되어 실제로 음성신호를 송출하기까지 일정한 시간이 필요하다. 아날로그 VOX나 AGC로 개선된 VOX는 입력된 음성신호를 기초로 그림3과 같이 PTT의 동작여부를 판단하고, 일정한 크기의 신호라고 판단이 되면 PTT 제어가 되고 이를 통해 음성 송신준비를 하게 된다. 그러나, 이때 이용된 음성신호는 저장되지 않았고, 무전기 출력이 안정화되기 전까지의 음성은 전달되지 못한 채 소멸된다. 수신 측에서는 송신한 음성신호의 초기의 일부를 전달하지 못하는 기본적인 한계를 가지고 있다.

이러한 한계를 극복하기 위해 음성손실의 최소화를 위해 불필요한 신호나 휘파람으로 PTT를 동작시킨 후에 전송하고자 하는 음성을 전달한다. 이러한 시스템 운영은 운영방법을 알지 못하는 일반사용자가 사용하기에는 한계가 있고, 전달속도를 지연시키는 것은 물론 많은 오동작을 발생시킬 여지가 있다. 또한 음성신호의 지속 정도에 따라 PTT의 비정상 제어로 무전기의 고장을 유발하는 원인이 될 수 있다.

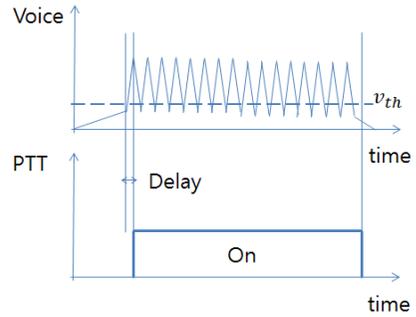


그림 3. 기존 VOX의 타이밍
Fig. 3 Timing of existing methods VOX

III. 제안한 개선방법

3.1 음성 신호처리를 통한 VOX

본 논문에서 제안한 자동전환장치의 구현을 위하여 VAD(Voice Activation Detection)[7-9] 방법을 이용하여 음성구간을 검출한다. 이를 위해 단 구간의 음성 프레임의 초기 음성구간이 아닌 구간에서 신호에너지의 평균값 M_x 을 구해 그 값이 일정한 기준을 초과 또는 미달하였는지를 판단한다[3]. $X_{noise(i)}$ 는 입력된 신호를 n 개의 입력된 샘플의 수로 나누어 이를 프레임 단위로 평균값 M_x 을 구해 이를 기준값과 비교한다.

$$M_x = \sum_{i=0}^{n-1} \frac{X_{noise(i)}}{n} \quad (1)$$

이 방법은 기존의 아날로그 방식에서 동작하는 수준으로 주변의 잡음환경에서는 사용하기가 부적당하다.

VAD에서 중요한 특징은 새로운 임계치인 th_{est} 를 추정해 나가는 것으로, 이는 전 평균치의 데이터인 M_x 와 새로 입력되는 데이터에 가중치를 가산하여 식 (2)와 같이 재추정(Re-estimation)을 수행한다.

$$th_{est} = M_x + 2^* \sum_{i=0}^{n-1} \frac{X_{noise(i)}}{n} \quad (2)$$

VOX를 구현하기 전에 그림4와 그림5과 같이 시뮬레이션을 통하여 알고리즘의 유효성을 확인하였다. 그

림4에서 같이 시간영역 신호의 비음성 잡음입력과 스펙트로그램의 주파수 대역별 잡음신호를 분석하였다 [10]. 만약 이러한 신호가 입력되었을 때는 음성이 아닌 잡음이 입력된 것으로 간주하여 PTT의 억제와 음성출력을 제한한다.

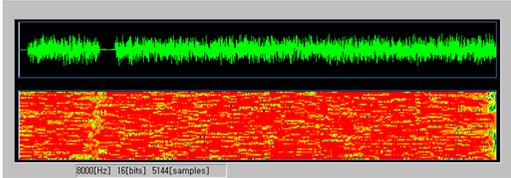


그림 4. 비 음성신호와 Spectrogram 분석
Fig. 4 Non-voice signal and Spectrogram analysis

그림5는 입력된 음성신호가 각 음절에 따라 짧게 단절되는 것을 보여주고 있고, 스펙트로그램을 통해 입력된 음성신호의 주파수대역을 보여주고 있다. 이는 입력된 신호가 잡음이 아닌 음성임을 알 수 있고, 음성 구간이 입력되었음을 알 수 있으며, 입력된 음성에 반응하여 PTT제어를 수행하고, 이후 저장된 음성을 출력한다.

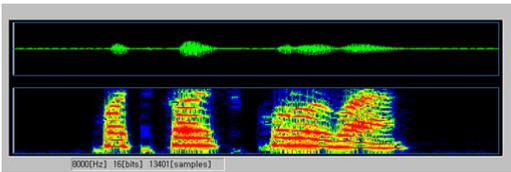


그림 5. 음성신호와 Spectrogram 분석
Fig. 5 Voice signal and Spectrogram analysis

3.2 DSP를 통한 실시간신호처리

아날로그의 실시간 음성을 입력 받아, 이를 주파수 대역으로 신호처리하기 위해서는 고속의 디지털 신호 변환장치인 CODEC(: Coder and Decoder), 저장장치와 처리를 위한 고성능의 프로세서가 필요하다.

음성 신호처리의 실시간 처리를 위해서는 일반적인 CISC(: Complex instruction set computing) 프로세서는 구현의 편의성은 있지만 고가의 장치가 요구되고, 소형화할 수 없는 문제로 인해 고려대상에서 제외하기로 한다. 본 논문에서는 신호처리 전용의 RISC(: Reduced Instruction Set Computer) 프로세서인

DSP(: Digital Signal Processor)를 사용하였다. 또한 음성신호의 입·출력을 위해 2채널 32bit CODEC과 신호 저장을 위한 메모리 등으로 구성되어 있다.

IV. 구현 및 실험

4.1 하드웨어 구현

제한한 시스템의 실시간 음성신호처리 구현을 위해 TI사의 TMS320VC5505 eZdsp USB STICK[9-10]을 사용하였다. 또한, TI의 CCS(: Code Composer Studio) 를 사용하여 신호처리 SW를 구현하였다. 표1은 음성 자동전환장치의 실험에 사용한 신호의 규격과 분석구간 시간을 나타내고 있다.

표 1. 실험용 음성규격

Table 1. Experimental sound specification

Sampling Rate	16bit
Sampling Freq.	8kHz
Frame Interval	25msec
Frame Shift	10msec
Frame Samples	400ea

그림6와 같이 TMS320VC5505 eZdsp USB STICK을 사용하여 VOX 알고리즘을 처리하고, 전화 또는 무전기와의 임피던스 매칭을 위한 회로와 무전기의 PTT와 음성신호를 출력하기 위한 회로제작을 통하여 이중 기기 간 음성통신을 위한 자동전환장치를 구현하였다.



그림 6. 실험용 보드
Fig. 6 Evaluation boards

4.2 소프트웨어 구현

구현시스템의 처리절차는 그림7과 같이 신호가 입력되면 이를 CODEC을 이용해 메모리에 25msec 단위의 프레임으로 저장하여, PTT 제어신호 발생 전에 손실되는 신호가 없도록 한다. 또한 저장된 프레임의 신호가 음성이라고 판단되면 VAD 프로파일을 생성하고, PTT 제어신호를 약 500msec 먼저 출력 후에 저장된 음성을 출력하도록 한다. 또한, 음성신호의 출력이 종료된 후에도 PTT를 유지시켜 저장된 음성신호의 마지막까지 모두 출력하고, 일정 시간이상 유지 후에 PTT 제어신호를 OFF한다.

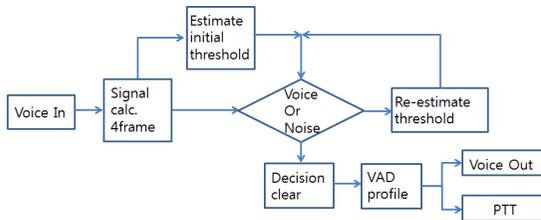


그림 7. 음성여부판단과 임계치 재추정

Fig. 7 Voice determination and re-estimating threshold

그림8과 같이 중간에 음절이 잠시 단절되어도 PTT 제어신호를 일정하게 유지하도록 한다. 그렇지 않으면 음성신호의 짧은 단절에도 민감하게 반응하여 수신측에 정상적인 음성신호를 전달할 수 없다.

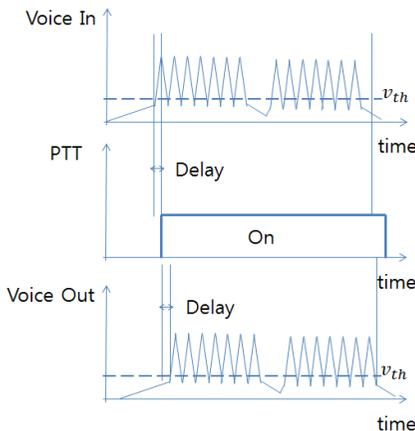


그림 8. 음성 신호처리를 이용한 VOX

Fig. 8 VOX with the audio signal processing

짧은 PTT의 제어는 음성신호 단절의 원인이 되고, 비정상적으로 PTT가 오래 지속되면, 반이중 통화로를 너무 오래 점유하여 양방향 통신을 정상적으로 할 수 없게 된다.

4.3 실험결과

그림9는 입력된 음성파형과 음성신호를 판단하여 발생된 PTT 신호와 지연 출력되는 음성신호를 나타내고 있다. 이는 입력된 신호의 초기 부분이 손실되지 않고, 모두 메모리에 저장된 후 식(2)를 이용해 음성으로 판단되면 순차적으로 음성을 출력하고, 반대로 비음성이나 잡음이 입력되면 PTT의 발생을 저지하여, 입력된 비음성이나 잡음은 출력되지 않는다.

기존 아날로그 방식의 VOX, AGC로 개선한 VOX, 입력 음성의 초기 손실의 문제가 본 실험에서 주변잡음에 적응적으로 동작하고 안정적인 PTT의 발생과 손실 없는 음성을 출력하는 것으로 확인되었다. 음성의 짧은 단절에도 PTT신호를 안정적으로 유지함으로써 수신측에서 끊김 없는 안정적인 음성을 전달 받는 것을 확인할 수 있었다.

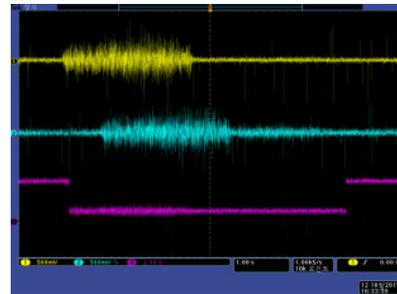


그림 9. 실험결과

Fig. 9 Result of experimental

입력된 음성신호에 정확하게 반응하는 PTT도 음성 전달의 품질면에서 중요하지만, 오동작으로 인한 PTT가 반복적으로 발생한다면 반이중 통화로의 일방적인 점유로 수신할 기회를 잃어 통신 본래의 기능을 원활하게 할 수 없기에 불필요한 PTT의 억제도 이에 못지않게 중요하다.

기존 아날로그 VOX와 개선된 음성전환 장치인 VOX와의 주요특징을 비교하면 표2와 같이 개선되었다.

표 2. 실험결과

Table 2. Experiment result

	existing VOX	Improving VOX
First word loss	X	○
Sustain PTT	△	○
Stability PTT	△	○
Voice loss	△	○

V. 결 론

유선과 이동통신 음성전화와 무전장치들을 이중 기간의 음성통신을 위해서는 음성전환 장치가 필요하다. 기존 장치 성능개선을 위하여 실시간 DSP를 이용하여 신호를 분석하고, 이중 기기 간 음성통신을 위한 자동전환장치인 VOX를 구현하였다. 잡음이 없는 순수한 음성과 이 음성에 각 5dB, 10dB, 15dB, 30dB의 잡음을 혼합[11]하여, 실험용 음성파일 5종을 준비한다. 이 파일을 실험용 보드를 이용하여 실험한 결과, 전달음성신호의 처음 부분에 대한 훼손방지, 전달음성의 짧은 단절에 대한 PTT 제어신호 지속시켰으며, PTT 제어신호의 안정적인 동작으로 고장발생요인을 낮춤으로써 송신측의 음성신호를 수신측에 정확하고 손실 없이 전달할 수 있도록 하였다.

향후 이중 기기 간 음성통신 자동전환장치를 이용하여, 다양한 무전기의 핸드프리, UHF와 VHF간 상호 연동, 전화를 통한 무전기 연동 등을 통해 비상 음성통신, 경찰 및 군작전용 장치에 적용함으로써 손실 없는 음성통신을 수행할 것이라 판단된다.

References

[1] Y. Liu and N. Na, "A Voice-Operated Transmit(VOX) Algorithm for Input Sinal with Low SNR," *Proc. of 2011 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference*, vol. 2, Harbin, Heilongjiang, China, July, 2011, pp.1243-1246.

[2] J. Song and S. Lee, "Voice Activity Detection Using Modified Power Spectral Deviation Based on Teager Energy," *J. of Rehabilitation*

Welfare Engineering & Assistive Technology, vol. 8, no. 1, 2014, pp. 41-46.

[3] J. Hong and M. Hahn, "Robust Feature Extraction for Voice Activity Detection in NonStationary Noisy Envrionments," *J. of the Korean Association of Speech Sciences*, vol. 5, no. 1, Mar, 2013, pp. 11-16.

[4] T. Choi and S. Kim, "A Speech Segregation Technique for Speech Quality Enhancement on Noise Enviroment," *J. of Korean Institute of Information Technology*, vol. 10, no. 10, 2012, pp. 47-53.

[5] C. Lew, "A Study on VOX Realization using Teager Energy," *Master's thesis, Graduate School Korea Aerospace University*, 2001,

[6] H. Park, "Heterogeneous Radio Adapter for Disaster Voice Communication," *J. of Korean Institute of Information Technology*, vol. 11, no. 12, 2013, pp. 143-150.

[7] Y. Park and S. Lee, "Voice Activity Detection Using Global Speech Absence Probability Based on Teager Energy in Noisy Environments," *J. of the Institute of Electronics Engineers of Korea*, vol. 49, no. 1, Jan, 2012, pp. 97-103.

[8] Y. Chung, "Speech Recognition based on Environment Adaptation using SNR Mapping," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 9, no. 5, 2014, pp. 543-548.

[9] G. Jang, "An Efficient Voice Activity Detection Method using Bi-Level HMM," *J. of The Korea Institute of Electronic Communication Sciences*, vol. 10, no. 8, 2015, pp. 901-906.

[10] C. Lee, "Adaptive Noise Reduction of Speech Using Wavelet Transform," *J. of the Korea Institute of Electronic Communication Sciences*, vol. 4, no. 3, 2009, pp. 190-196.

[11] T. Choi and S. Kim, "Gamma-tone Feature Extraction Acoustic Modeling for Improving Speech Recognition Performance," *J. of Korean Institute of Information Technology*, vol. 10, no. 11, 2012, pp. 155-160.

저자 소개



류창국(Chang-Guk Lew)

1991년 청주대학교 전자공학과 졸업(공학사)

2001년 한국항공대학교 항공산업대학원 정보통신공학과 졸업(공학석사)

2014년 ~ 현재 전남대학교 대학원 전자컴퓨터공학과 박사과정

※ 관심분야 : 음성신호처리, 영상신호처리, ITS, CCTV, u-City



이배호(Bae-Ho Lee)

1978년 한양대학교 전자학과 졸업(공학사)

1980년 한국과학기술원 대학원 전기 및 전자공학과 졸업(공학석사)

1993년 University of Missouri, Columbia 졸업(공학박사)

1980년 ~ 1983년 국방과학연구소

1993년 ~ 현재 전남대학교 전자컴퓨터공학부 교수

※ 관심분야 : 영상처리, 컴퓨터비전, 인공지능, 멀티미디어 등

