

---

# 무선 채널 환경에서 디지털 이동통신용 음성 부호화기의 성능 평가

김형중\*, 윤병식\*, 최승인\*

## Performance Evaluation of Speech Coder for Digital Mobile Communication System in Radio Channel Environment

Hyung-Jung Kim\*, Byung-Sik Yoon\*, Song-In Choi\*

### 요약

본 논문에서는 현재 디지털 이동통신 시스템에서 운용되고 있는 QCELP(Qualcomm Code Excited Linear Predictor) 음성부호화 방식과 향후 IMT-2000 (International Mobile Telecommunications 2000) 등의 시스템에서 사용 예정인 CS-ACELP(Conjugate Structure Algebraic Code Excited Linear Prediction) 음성부호화 방식과의 성능을 비교한다. 특히 무선 채널을 사용하는 이동통신환경의 특징인 채널에러로 인한 음성부호화기의 성능을 비교함으로써 채널에러에 강인한 음성부호화 알고리즘 설계에 대한 고찰을 유도한다.1)

### Abstract

In this paper, we present a comparison between QCELP(Qualcomm Code Excited Linear Predictor) speech coder that is operating in digital mobile communication system and CS-ACELP(Conjugate Structure Algebraic Code Excited Linear Prediction) speech coder that is scheduled to use for IMT-2000 (International Mobile Telecommunications 2000) system. The performance comparison might give help to design of the speech coding algorithms so that the robustness of the algorithms to channel errors engaged by mobile communication system be optimized.

---

\* 한국전자통신연구원 신호처리연구실

## I. 서 론

1984년에 아날로그 이동통신 서비스인 AMPS (Advanced Mobile Phone System)가 도입된 이래 우리나라의 이동통신 시장은 급속도로 성장되어 왔다. 많은 수의 이동통신 가입자들이 제한된 스펙트럼을 동시에 사용하고자 함에 따라 서비스 품질 저하가 발생하게 되었다. 이로 인하여 큰 가입자 수용 능력을 가진 디지털 이동통신 시스템을 개발하게 되었다. 디지털 이동통신 시스템의 급속한 발전으로 인하여 기존의 아날로그 이동통신과 비교하여, 고품질의 다양한 서비스를 보다 많은 수의 가입자에게 제공하게 되었다.

디지털 이동통신 시스템에서 가장 중요한 요소 중 하나는 주파수 효율성의 향상과 양질의 음성 품질이다. 이런 요소는 디지털 이동통신에서 사용하는 음성부호화기와 밀접한 관계가 있다. 디지털 이동통신용 음성부호화기의 요구사항은 가능한 낮은 전송률에서 고품질의 음성 서비스가 가능하여야 하며, 이동통신 채널환경에 강인하여야 하며, 적절한 복잡도를 가져야 한다. 현재 TDMA 방식의 이동통신용 음성부호화 방식으로는 GSM(Global System for Mobile communication)에서 사용되는 13Kbps RPE-LTP(Regular Pulse Excited Long Term Prediction) 음성부호화기와, 북미와 일본에서 각각 표준안으로 채택된 8Kbps와 6.7Kbps로 동작하는 VSELP(Vector Sum Excited Linear Prediction) 음성부호화기가 있다[1][2]. CDMA 방식의 이동통신용 음성부호화 방식으로는 QCELP(Qualcomm Code Excited Linear Prediction) 음성부호화기가 있다[3]. 이 방식 역시 IS-96으로 표준화되었다[4]. 현재 활발히 연구 개발되고 있는 IMT-2000 시스템에 사용되어질 음성부호화 방식으로 8Kbps CS-ACELP(Conjugate Structure Algebraic CELP)[5], 13Kbps ACELP, 13Kbps QCELP 등이 제안되었다. 본 논문은 디지털 이동통신용 음성부호화기의 중요한 요구사항 중 하나인 채널에러에 강인한 음성부호화 방식이라는 관점에서, QCELP 음성부호화기와 차세대 이동통신용 음성부호화기로 주목받고 있는 CS-ACELP 음성 부호화기의 성능을 비교하였다. 다양한 채널에러 환경하에서 기존의 IS-96

QCELP 음성부호화기의 문제점이 발견되면서, IS-96A 잠정 표준안으로 변경된 QCELP 음성부호화기의 성능도 같이 비교하였다. 아울러 4bit CRC가 부여된 CS-ACELP 음성부호화기의 성능도 채널에러 환경하에서 CRC가 없는 경우와 비교하였다. 비교 방법으로는 객관적 음질평가 기준인 SNR (Signal to Noise Ratio)와 SEGSNR (Segmental Signal Noise Ratio)를 사용하였다.

본 논문은 다음과 같이 구성되어 있다. II장에서는 QCELP 음성부호화 방식의 구조와 특징에 대해, III장에서는 CS-ACELP 음성부호화 방식의 구조와 특징에 대해 간단히 설명한다. IV장에서는 채널에러 환경하에서 QCELP 음성부호화기와 CS-ACELP 음성부호화기의 성능을 비교하고, 마지막으로 V장에서 결론을 맺는다.

## II. QCELP 음성부호화 방식

QCELP 음성부호화 방식은 기본적으로 AT&T의 CELP 알고리즘을 기본으로 구성되어 있다. 분석/합성 (Analysis by Synthesis) 방법을 사용하여 음성 생성에 필요한 피치 이득(Pitch Gain), 피치 인덱스(Pitch Index), 코드북 이득(Codebook Gain), 코드북 인덱스(Codebook Index) 등을 먼저 결정하고, 여기 신호 (Excite Signal)를 벡터 양자화하기 위하여 코드북을 사용한다. 그러나 일반적인 CELP 구조와 구분되는 여러 특징을 지니고 있다. QCELP 음성부호화 방식은 음성신호의 에너지 양에 따라 전송률이 변하는 가변전송률 부호화 방식이고, 코드북 저장 메모리 양과 검색 계산량을 줄이기 위하여 128개의 Circular Stochastic 코드북을 사용하며, LPC(Linear Predictive Coefficient)의 양자화와 효과적인 보간(Interpolation) 및 적절한 안정성 검사 (Stability Check)를 위하여 LSP(Line Spectrum Pair)를 사용한다[6]. 특히 가변전송률 음성부호화 방식은 양방향 통화 시스템에서 묵음 시간이 전체 통화의 60-70%정도를 차지한다는 점을 이용하여 음성 품질의 저하없이 낮은 전송률로써 음성신호를 부호화할 수 있다는 장점이 있다. 각 전송률에 따른 전송프레임의 크기 및 비트 할당은 표 1에 있다.

표 1. QCELP 프레임 당 비트 할당

Table 1. Bits allocation of QCELP per frame

|                           | Rate 1 | Rate 1/2 |
|---------------------------|--------|----------|
| Bits per LPC update       | 40     | 20       |
| Bits per Pitch            | 10     | 10       |
| Number of Pitch update    | 2      | 2        |
| Bits per Codebook         | 8      | 4        |
| Number of Codebook update | 10     | 10       |
| Bits per Frame            | 160    | 80       |

|                           | Rate 1/4 | Rate 1/8 |
|---------------------------|----------|----------|
| Bits per LPC update       | 10       | 10       |
| Bits per Pitch            | 10       |          |
| Number of Pitch update    | 1        |          |
| Bits per Codebook         | 2        | 1        |
| Number of Codebook update | 10       | 10       |
| Bits per Frame            | 40       | 20       |

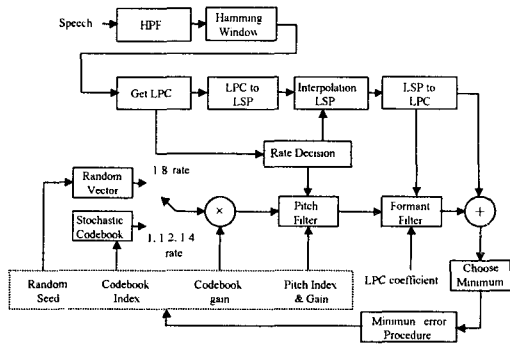


그림 1. QCELP 음성부호화기의 인코더  
Fig. 1. Encoder of QCELP speech coder

QCELP 음성부호화 방식의 인코더 구조는 그림 1과 같다. 음성신호를 8KHz 샘플링하여 20msec, 160개의 샘플로 이루어진 프레임 단위로 부호화가 시작된다. 샘플 데이터로 자기상관법을 사용하여 10차 LPC 계수를 추출하고 LPC 값을 수학적으로 동일한 LSP 계수로 변환한 후 각 전송률에 맞는 비트로 양자화한다. 양자화된 LSP값을 각 피치 및 코드북 부프레임(Subframe)에 맞게 보간한 다음 다시 LPC로 변환하여 피치 및 코드북 파라미터를 검색하는데 사용한다. 각 피치 및 코드북 파라미터들은 분석/합성 방법을 이용하여 입력 음성신호와

합성 음성신호를 비교하여, 오차가 최소화되도록 양자화된다.

### III. CS-ACELP 음성부호화 방식

CS-ACELP 음성부호화 방식은 1996년 ITU-T (International Telecommunications Union Telecommunications Standardization Sector)에서 G.729로 표준화되었다[7]. Univ. of Sherbrooke에서 개발된 ACELP 음성부호화기와 NTT사의 CS-CELP 음성부호화기와의 장점을 결합한 구조이며, PCS 및 IMT-2000 용으로 사용되어질 8Kbps급 음성 부호화기로 Toll-quality의 성능을 가진다고 알려져있다. CS-ACELP 음성 부호화기에서 한 프레임에 대한 비트 할당은 표 2에 나타내었다.

표 2. CS-ACELP 비트 할당

Table 2. Bits allocation of CS-ACELP

| Parameter      | Subframe1 | Subframe2 | Total |
|----------------|-----------|-----------|-------|
| LSP            |           |           | 18    |
| Pitch Delay    | 8         | 5         | 13    |
| Pitch Parity   | 1         |           | 1     |
| Codebook Index | 13        | 13        | 26    |
| Codebook Sign  | 4         | 4         | 8     |
| Codebook gain1 | 3         | 3         | 6     |
| Codebook gain2 | 4         | 4         | 8     |
| Total          |           |           | 80    |

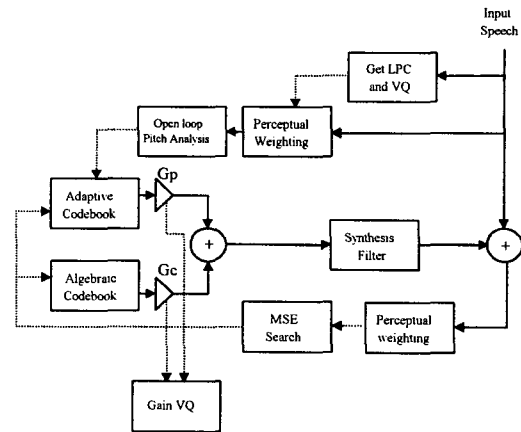


그림 2. CS-ACELP 음성부호화기의 인코더  
Fig. 2. Encoder of CS-ACELP speech coder

CS-ACELP 음성부호화기의 인코더 구조는 그림 2와 같다. 한 프레임은 10msec, 80 샘플로 구성된다. LPC 분석을 위하여 비대칭적인 Hamming windows를 사용하여 프레임 영역을 결정하고, LPC 파라미터를 추출한다. 이런 방식은 음성부호화기의 알고리즘 지연을 감소시키는 역할을 하게 된다. 알고리즘 지연은 IMT-2000 시스템에서 중요한 파라미터이다. LPC 계수가 결정되면, 양자화가 용이한 LSP로 변환한 후, 이전 4 프레임에 걸친 LSP 평균값을 구한 다음 현재 프레임의 LSP 값과의 차를 2-Stage 벡터 양자화를 수행한다. 첫번째 Stage는 128개의 7-dimensional 벡터 양자화를 수행하고, 두번째 Stage에서는 32개의 5-dimensional Split 벡터 양자화를 수행한다. 양자화된 LSP 값들은 각 피치 및 코드북 부프레임을 위하여 보관된다. 피치 검색은 먼저 개루프 검색이 수행되어진 후 얻어진 피치 지연값을 중심으로 아주 작은 영역 내에서 일반적인 정수형 피치 검색이 아닌 실수형 피치 검색이 이루어진다. 코드북은 ISPP (Interleaved Signle-Pulse Permutation) 구조에 따라 +1 혹은 -1로 구성된 4개의 펄스들로 구성된다. 이들은 Algebraic Structure로 구성되어 기존의 분석/합성 방법에서와 다른 간단한 연산을 통하여 코드북 인덱스 값을 구할 수 있으며 또한 많은 후보 코드북을 발생시키므로 음성품질의 향상을 가져왔다. 피치 이득과 코드북 이득은 2-stage Conjuagate 구조의 벡터 양자화기로 양자화된다.

IV. 채널에러하에서 성능 비교

이동통신 시스템은 무선채널을 사용하므로 페이딩, 잡음, 그리고 다른 사용자로부터의 간섭 등이 불가피하게 발생한다. 이런 다양한 에러환경들로 인하여 음성부호화기들의 성능평가는 이동통신 시스템에 있어 중요한 요소가 된다. 채널에러 환경하에서 음성부호화기의 성능은 파라미터의 양자화 방법에 따라 주로 그 성능이 좌우된다. 일반적으로 음성부호화기에서 많이 쓰이는 양자화 방법은 인접 프레임간의 상관성을 이용하여 적은 비트로 효율적으로 양자화하는 것이다. 그러나 이런 양자화 방법은 이전 프레임의 양자화값을 이용하여 예측

하는 구조이기 때문에 채널에러로 인한 품질저하가 비교적 크다. IS-96 QCELP 음성부호화기의 LSP 파라미터는 DPCM 방식 예측기를 사용하여 양자화되므로 채널에러에 의한 음성품질이 현격히 떨어지는 문제점을 드러내었다. 이를 보완하기 위하여 IS-96A QCELP 음성부호화기에서 제시한 LSP 파라미터 예측기는 기존 IS-96 예측기 식인 (식 1)에서 (식 2)로 변경 되었다.

$$P_w(z) = 0.90625Z^{-1} \quad \text{(식 1)}$$

$$P_w(z) = 0.0Z^{-1} \quad \text{(식 2)}$$

그러나 예측기를 사용하지 않으므로써 생기는 양자화 손실은 (식 3)에서와 같이 양자화기의 최대 및 최소 양자화 영역을 면밀히 조사하여 양자화기의 활성 영역을 최적화함으로써 양자화 성능저하를 막을 수 있었다.

$$Q_n(X) = \text{round}((2^N - 1) \frac{X - Q_w^{\min}}{Q_w^{\max} - Q_w^{\min}}) \quad \text{(식 3)}$$

여기서 N는 양자화 비트수이고,  $Q_w^{\max}$ 는 LSP 최대 양자화 영역이고,  $Q_w^{\min}$ 는 LSP 최소 양자화 영역을 나타낸다.

표 3. BER에 따른 IS-96과 IS-96A의 성능  
Table 3. Performance of IS-96 and IS-96A

| BER   | IS-96 |        | IS-96A |        |
|-------|-------|--------|--------|--------|
|       | SNR   | SEGSNR | SNR    | SEGSNR |
| 0%    | 13.96 | 11.66  | 13.97  | 11.44  |
| 0.01% | 12.73 | 10.91  | 13.07  | 11.10  |
| 0.05% | 9.89  | 9.55   | 11.22  | 10.18  |
| 0.1%  | 8.00  | 8.32   | 9.13   | 9.17   |
| 0.5%  | 2.08  | 3.32   | 3.94   | 5.27   |
| 1%    | 0.15  | 1.17   | 1.00   | 2.73   |

표 3은 BER 환경에서 QCELP 음성 부호화방식의 두가지 표준인 IS-96과 IS-96A와의 성능 평가 결과를 나타내었다. SNR과 SEGSNR을 사용하여 성능을 비교하였다. 표준안에 있는 에러보정능력, Rate 1 프레임인 경우 음성 파라미터 중 주요

한 28비트에 CRC 11비트를 사용하여 1비트 에러보정능력을 추가하여 시뮬레이션을 수행하였다. 채널에러가 없을 경우 IS-96 음성부호화기는 SEGSNR에서 약 0.2dB 높은 성능을 보이나 전송에러 발생시 성능저하는 더욱 심하게 나타나는것을 알 수 있다.

CS-ACELP 음성부호화기는 QCELP 음성부호화기와는 달리 LSP 파라미터 양자화 과정 및 피치, 코드북 이득의 양자화 과정에서 DPCM 방식의 예측기를 사용하지 않고 MA(Moving Average) 방식의 양자화 예측기를 사용하였다. 즉, 이전 4 프레임의 양자화된 LSP 값을 사용하여 LSP 값을 예측하고, 실제 LSP 값과 예측된 LSP 값의 차를 벡터 양자화한다. 또한 코드북 이득 또한 LSP 양자화 방법과 같이 이전 4개의 부프레임에서 양자화되었던 코드북 이득을 MA 방식의 양자화 예측기에 의하여 현재 코드북 이득 양자화에 사용하였다. 이러한 MA 방식 양자화 예측기를 사용하므로써 QCELP 음성부호화기에서 사용하였던 DPCM 방식의 양자화 예측기와 비교하면 현 프레임 에러 발생시 LPC 양자화로 인한 손실이 많이 줄어든다. 표 4는 BER환경에서 CS-ACELP 음성부호화기의 성능을 나타내었다. 이 결과는 QCELP 음성부호화기에서 11 비트의 에러보정 비트를 사용한 것과는 달리 에러보정 비트 없이 시뮬레이션하였다.

표 4. BER에 따른 CS-ACELP 성능  
Table 4. Performance of CS-ACELP per BER

| BER   | CS-ACELP |        | IS-96과 차이 |        |
|-------|----------|--------|-----------|--------|
|       | SNR      | SEGSNR | SNR       | SEGSNR |
| 0%    | 16.38    | 15.49  | 2.42      | 3.83   |
| 0.01% | 15.85    | 15.20  | 3.12      | 4.29   |
| 0.05% | 13.71    | 14.35  | 3.82      | 4.8    |
| 0.1%  | 11.81    | 13.33  | 3.81      | 5.01   |
| 0.5%  | 5.74     | 8.76   | 3.66      | 5.44   |
| 1%    | 3.13     | 5.44   | 2.98      | 4.27   |

에러보정 능력이 없는 CS-ACELP 음성부호화기가 에러보정 능력이 있는 IS-96 음성부호화기와 비교하여, 채널에러에 강인하다는 것을 표 4로 알 수 있다.

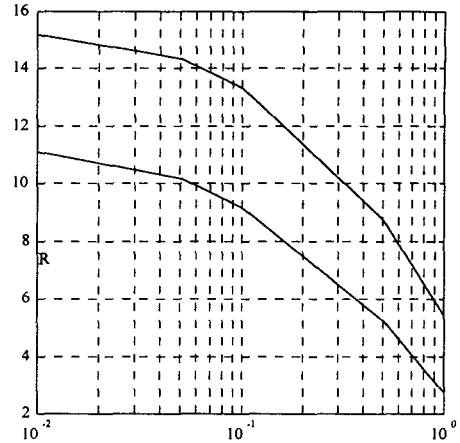


그림 3. BER에 따른 CS-ACELP와 IS-96A 성능  
Fig.3. Performance of CS-ACELP and IS-96A

채널에러에 강인하게 설계된 IS-96A 음성부호화기와 표 4를 통하여 채널에러에 강인하다고 생각되는 CS-ACELP 음성부호화기도 0.1% BER까지는 음성품질을 유지하나, 그 이상에서는 급격히 성능이 떨어짐을 그림 3을 통하여 알 수 있다. CS-ACELP 음성부호화기는 1% BER 경우 이전 BER과 비교하여 성능이 갑자기 저하되었다. 이 결과로 1% BER 이상의 채널에러하에서는 MA 예측기와 벡터 양자화기를 사용하더라도 큰 효과를 얻지 못할 것으로 추측된다.

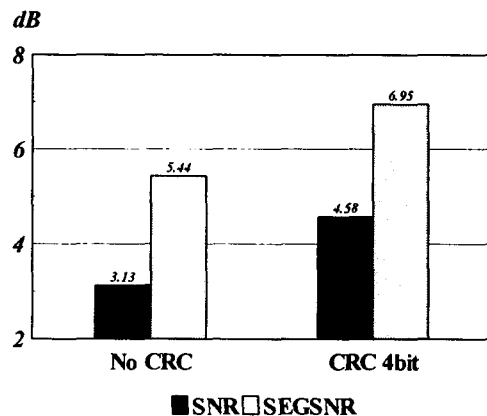


그림 4. 1% BER 시 4bit CRC CS-ACELP 성능  
Fig.4. Performance of CS-ACELP on 1% BER

그림 4는 CS-ACELP 음성부호화기에서 2-stage LSP 벡터 양자화 중 1-stage의 벡터 양자화된 7bit 와 피치 및 코드북 이득을 양자화하기 위한 2-stage 벡터 코드북의 MSB 4bit를 CRC 4비트로 1비트 에러보정 능력을 부여하여 시뮬레이션한 것이다. 1% BER 환경하에 SNR은 약 1.2dB SEGSNR은 약 1.5dB의 성능 향상을 가져옴을 알 수 있다.

### V. 결 론

본 논문은 디지털 이동통신 시스템용으로 개발된 QCELP 및 CS-ACELP 음성부호화기를 다양한 채널 환경하에서 성능을 평가하였다. 파라미터 양자화에 MA 예측기를 사용한 CS-ACELP 음성부호화기 경우 DPCM 방식의 예측기를 가진 IS-96 QCELP 음성부호화기와 비교하여 0.01% BER에서 0.5% BER까지 SNR 및 SEGSNR의 성능이 향상됨을 볼 수 있었다. 이는 파라미터 양자화 방법에서 MA 예측기를 사용한 CS-ACELP 음성부호화기가 채널에러 환경하에서 더욱 효과적임을 알 수 있다. 그러나 MA 예측기와 벡터 양자화를 사용하더라도 BER 1% 이상시에는 성능 향상을 기대할 수 없음을 볼 수 있다.

음성부호화기만으로 채널환경 하에서의 성능을 평가하였으며, 실제 시스템에서는 채널 코더와 결합되어 사용된다. 이를 프레임 당 4비트의 CRC를 추가한 CS-ACELP 음성부호화기의 경우 CRC를 추가 하지 않은 CS-ACELP 음성부호화와 비교하여 BER 1%시 약 1.5dB 이상의 SNR 및 SEGSNR 향상을 가져옴으로써 확인하였다. 음성부호화기가 에러보정 능력이 큰 채널코더와 결합된다면, 채널 에러환경 하에서도 품질저하가 감소될 것이다.

### 참 고 문 헌

[1] EIA/TIA/IS-54-B, "Cellular System Dual Mode Mobile Station Base Station Compatibility Standard", April 1992.

[2] M.H.Sunwoo, S. Park, "Real-time implementation of the VESLP on a 16bit DSP Chip", IEEE Trans. Consumer Electronics, pp.772-782, Vol.37, No.4 Nov. 1991.

[3] TIA/EIA IS-96 Interim Standard. "Speech service option standard for wideband spread spectrum digital cellular system", April 1993.

[4] TIA/EIA/IS-95, "Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System", July 1993.

[5] Redwan Salami, Claude Laflamme, Jean-Pierre Adoul, Dominique Massaloux, "A Toll Quality 8Kb/s Speech Codec for the Personal Communications System", IEEE Trans. On Vehicular Technology, Vol. 43, No. 3, Aug 1994.

[6] Noboru Sugamura, Nariman Farvardin, "Quantizer Design in LSP Speech Analysis-Synthesis", IEEE Comm. Vol. 6, No.2, Feb. 1988.

[7] ITU-T G.729, "Coding of speech at 8 kbits/s using conjugate-structure algebraic -code-excited linear-prediction", May. 1996.

[8] B.S.Yoon et al, "Fixed Point Implementation of the QCELP Speech Coder", ETRI Journal. Vol. 19. No. 3. Oct. 1997.



김 형 중(Hyung-Jung Kim)  
 1970년 9월 25일 생  
 1993년 2월 : 한양대학교 전자  
 공학과 졸업(공학학사)  
 1995년 8월 : 한양대학원 전자  
 공학과 졸업(공학석사)  
 1995년 8월~현재 : 한국전자통신연구원 신호처리  
 연구실 연구원  
 주관심분야 : 이동통신, 음성신호처리, 적응신호처리



윤 병 식(Byung-Sik Yoon)

1967년 10월 31일 생

1990년 2월 : 경북대학교 전자  
공학과 졸업(공학학사)

1992년 2월 : 경북대학원 전자  
공학과 졸업(공학석사)

1992년 1월~현재 : 한국전자통신연구원 신호처리  
연구실 연구원

주관심분야 : 이동통신, 음성신호처리, 영상신호처리



최 승 인(Song-In choi)

1957년 1월 26일 생

1982년 2월 : 광운대학교 응용  
전자공학과 졸업(공학학사)

1987년 2월 : 광운대학원 전자계  
산기공학과 졸업(공학석사)

1982년 7월~현재 : 한국전자통신연구원 신호처리  
연구실 책임 연구원

주관심분야 : 이동통신, 음성신호처리, 영상신호처리