DBpia

論 文

# 디지탈 음성방식의 성능비교에 대한 연구

正會員 裵 澈 秀*

# A Study on the Comparison of Digital Speech Coding Performance

Chul Soo Bai* *Regular Member*

"이 논문은 1990년도 교육부지원 한국 학술 진흥재단의 자유공모과제 학술연구조성비에 의하여 연구되었음."

要 約

본 논문은 음성 시스템과 통신망에서 이용되는 음성 품질 평가 모델의 구축을 위한 기본 연구로서, 음성 부호화 평가 방법중 주관적 평가에서 발생되는 여러 문제점을 해결하여 안정된 객관적 평가값을 얻기 위해서, 여러 객관적 평가량과 주관적 평가량을 상호 비교한 후, 주관적 평가값에 적합한 객관적 평가량을 검토하였다.

ABSTRACT

Resonable speech quality assessment methodologies are required for speech quality assessment model which is used at speech system and communication network.

There are objective measures and subjective measures and subjective measure has the various problems in speech quality assessment methodologies.

The objective of this study is to compare objective measures with subjective measure and obtain the objective measure as close as possible to subjective measure.

## I. 서 론

통신망의 디지탈(digital)화가 추진되고 있는 가운데, 음성 품질 평가 기술은 여러 통신 시스템에서 이용되는 음성 부호화 장치의 평가와 비교, 효과적인 통신망의 계획, 시스템의 알고리듬(algorithm)의 진단 평가등에서 이용되는 중요한 기술로서, 성능 평가법에는 주관적 방법과 객관적 방법이 있고, 두 방법 모두 중요한 연구 과제가 된다.[1]

이와 같은 음성 품질 평가 기술은, 음성 파형으로부터 계산된 신호대 잡음비(Signal to Noise Ratio), 스펙트럼 왜곡(Spectrum Distortion)등의 물리량을

*關東大學校 電子通信工學科
Dept. of Electronic Communication Engineering,
Kwan Dong University.
論文番號：92-88(接受1992. 5. 25)



www.dbpia.co.kr

추정하는 방법과, 실제로 신호를 청취하여 평가를 구한 MOS(Mean Opinion Score) 등의 주관적 판단이 있다.

현재 몇몇 주관적 방법에 의해 통신 기기들의 품질 평가가 수행되고 있지만, 앞으로 객관적 평가법이 확립되면 주관적 평가에 필요한 상당한 시간을 낭비하지 않고, 평가의 능률을 높여줄 것으로 생각된다. 그러나 주관적 평가에서 발생될 수 있는 여러 문제점을 해결하기 위한 객관적 평가법을 사용할 경우, 주요한 주관적 요소 각각을 확실히 추출하기 위하여, 다양한 음원을 제시할 필요가 있다.

또한 기존의 전화 선로를 이용한 효율적인 전송 방법에 관심이 고조되고 있는바, 이에따른 선결 과제로서 음성 신호 부호화 방법에 의한 성능 평가는 필수적인 과제이다.

그러므로 본 연구에서는 디지탈 부호화 방식의 가장 기본이 되는 PCM방식, 단구간의 음성 파워(power)를 이용한 ADPCM과 DM의 적응 양자화 방법인 CFDM과 CVSD등을 사용하여, 통신망의 이상적인 선로상에서와 비트 에러률(bit error rate)에 따른 각 부호화 방식의 성능 평가를 객관적 방법(SEGSNR, LPC-CD, FFT-SD)과 주관적 방법(MOS)를 통하여, 두 평가값의 결과를 상호 비교하였다.

## II. 디지탈 음성 부호화 방식

디지탈 음성 부호화기는 아날로그(anlaog) 음성 신호를 디지탈 신호로 변환하여 신호의 저장이 쉽고, 신호대 잡음비(SNR)가 향상되는 특징을 갖고 있다. [그림 2-1]에 여러 음성 부호화 방식의 전송 비트율(bit rate)과 음질 사이의 관계를 나타내었다.

[그림 2-1]에 나타낸 것과 같이 음성 부호화 방식은, 음성 신호 자체를 직접 양자화하여, 신호와 유사한 파형을 재생하는 파형 부호화(waveform coding) 방식과 음성 신호의 발생 모델을 기본으로하여, 파형을 재생하는 보코딩(vocoding) 방식으로 크게 구분할 수 있으며[1], 80년대 이후부터 위의 두 방법들의 장점들을 동시에 이용한 복합 부호화(hybrid codidng) 방식이 있다.

본 장에서는 24Kbps-64Kbps에서 음질이 매우 좋은 파형 부호화 방식중 PCM, ADPCM과 적응 DM 등에 대하여 기술한다.

### 1. PCM

PCM 방법은 [그림 2-2]에 나타낸 것과 같이, 아날로그 신호를 표본화하여 시간적으로 이산(discrete)시킨 후, 그 샘플(sample) 값을 양자화한다. 이와같이 양자화한 결과를 유한개의 2진 부호중 1개로 (B비트 부호화인 경우, $2^B$ 개중 1개) 대응시킨 결과의 펄스열로 구성하는 방법이다[1, 8].

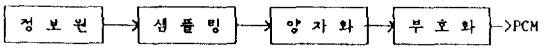



그림. 2-2 PCM 부호화 방식의 과정

한편, 이산 신호를 균일하게 양자화시킬 경우, 양자화로 인한 상대오차는 큰 진폭을 갖는 샘플에서 보다는 작은 진폭을 갖는 샘플에서 더 크게 된다. 그러므로 이러한 오차로부터 발생하는 작은 진폭의 신호 구간에서의 정보 손실을 줄이기 위하여, 입력 신호의 크기에 비례하여 양자화 레벨을 조절하는 방법이 제안되었으며, 그 한 방법이 μ-law compression law

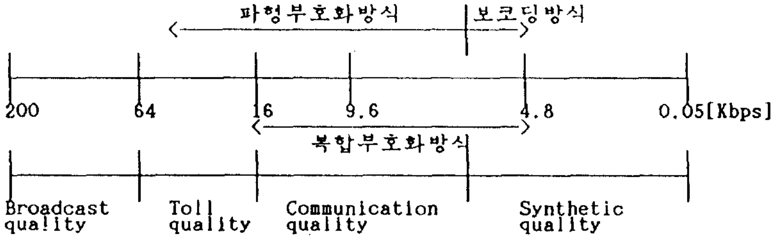



그림. 2-1 부호화 방식의 전송 비트율과 음질



www.dbpia.co.kr

이다. 이 규칙은 LOG 함수를 취하는 값의 "1"보다 큰 범위에 둠으로서, compression 된 값의 dynamic range를 입력 신호의 동작 범위와 같게 한다. 이 규칙의 수식은 식 (2-1)과 같다.

$$y(n)=F[x(n)]$$

$$=X_{max}\frac{\log\left[1+\mu\dfrac{X(n)}{X_{max}}\right]}{\log[1+\mu]}\,\mathrm{sign}[X(n)] \qquad (2\text{-}1)$$

$X(n)$ : 입력 샘플
$y(n)$ : compress 된 샘플
$X_{max}$ : 입력측의 최대값
$\mu$ : 압축률

단순한 LOG 양자화에서는 $X(n)=0$일 경우, 양자화된 값이 무한대인데 반하여, 식 (2-1)에서는 $y(n)=0$임을 알 수 있다.

식 (2-1)의 특성을 갖는 양자화를 위한 compressor /expandor 시스템의 블록도를 [그림 2-3]에 나타내었다.

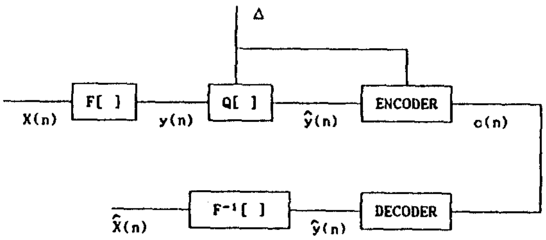



그림. 2-3 Compressor /Expandor 시스템

## 2. ADPCM

디지탈 음성 통신에 있어서의 전형적인 전송 방식은 64Kbit /s PCM 방식이다. 그러나 음성 파형을 분석해보면, 샘플간의 상관성 (correlation)이 상당히 크다는 것을 알 수 있기 때문에, 전송 음질을 유지하면서 전송 비트율을 낮출 수 있는 다른 부호화 기법이 가능하다. 이러한 특징을 갖는 부호화 기법중의 하나가 ADPCM으로서, 이 방법은 APCM(Adaptive PCM)과 DPCM(Differential PCM)의 두 방법을 결합시킨 것이다.[2]~[8].

기본적인 ADPCM의 블록도를 [그림 2-4]에 나타내었다.

[그림 2-4]에 나타낸 부호기는 64Kbit /s log PCM 신호를 입력으로 하여, 32Kbit /s ADPCM 신호를 출력하는 것을 보여주고 있다. 그리고 차분 신호 d(k)는 log PCM 신호를 균일 (linear) PCM 신호로 변환한 후, 다음식에 의해 구해진다.

$$d(k)=Sl(k)-Se(k) \qquad (2\text{-}2)$$

여기에서 Se(k)는 추정 신호, Sl(k)는 균일 PCM 신호를 나타낸다.

## 3. 적응 DM

Linear DM에서의 slope overload 문제를 해결하기 위하여 1960초 이래로 많은 적응 DM 방법들이 개발되었다[8], [9], [10].

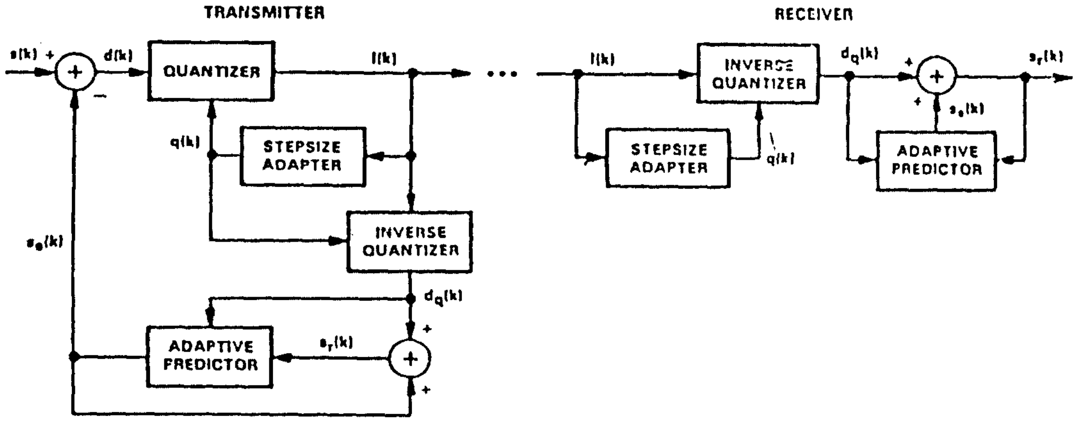



그림. 2-4 ADPCM의 블록도



www.dbpia.co.kr

예를들면 CVSD(Continuously Variable Slope DM), CFDM(first-order Constant Factor DM), HCDM(Hybrid Companding DM) 등이 있다.

CFDM은 출력 이진 신호 $\{b_n\}$의 현재값과 한 샘플 전의 값의 부호를 비교하여, 매 샘플 순간에 순간 적응(instantaneous adaption)을 한다. Jayant가 제안한 CFDM의 step 폭은 매 샘플 순간에 1.5나 0.66배만큼 변화하며, 수식적으로는 다음과 같다.

$$\Delta_n = 1.5\,\Delta_{n-1}, \text{ if } b_n = b_{n-1}$$
$$\quad 0.66\,\Delta_{n-1}, \text{ otherwise} \qquad (2\text{-}3)$$

여기에서 최대 및 최소 step 폭을 제한하여 바람직하지 못한 순간적인 효과를 피해야 한다.

CVSD는 출력인 2진 부호열의 형태로 부터 slope overload를 검출하며, 이때 syllabic 한 stepsize를 적응화해야 한다. 이렇게하여 변화된 stepsize를 일종의 leaky 적분기인 신호 추정 필터 (signal estimate filter)의 입력으로 사용해서 CVSD 부호화기의 입력을 예측한다.

CVSD의 step 폭은 다음과 같다.

$$\Delta n = (1-k)\,\Delta_{n-1} + k\,d_n \qquad (2\text{-}4)$$
$$b_n = b_{n-1} = b_{n-2},\ d_n = 1$$
$$\text{otherwise}\quad,\ d_n = 0$$

여기에서 $b_n$은 이진 신호 출력, k는 $2^{-8}+2^{-9}$의 값을 나타낸다.

## Ⅲ. 음성부호화 방식의 음질평가

### 1. 음질의 주관적 평가

음성이나 영상과 같은 통신 신호를 부호화하는데 있어서, 잡음 측정에 대한 결정적인 메카니즘은 인간의 인지 메카니즘이다. 그러므로 개념적인 주관적 평가가 모든 부호화기 구성과 평가에 적용되어야만한다. 비록 의미있는 객관적 평가법을 이용하여, 여러 부호화 방법에 적용한다 하더라도 객관적 데이터를 보충, 조정하기 위하여 어느 정도의 주관적 평가를 수행할 필요가 있다[4], [6], [7].

#### 1.1. MOS(Mean Opinion Score)

오피니언(Opinion) 등급이 음질 만족도를 평가하는데 적용된다. 이러한 등급을 5등급(아주 좋다, 좋다, 보통이다, 나쁘다, 아주 나쁘다)으로 나누어지며, 각 등급에서 4-0의 가중치를 부과하여 평균값을 구한다. 이 평균값을 MOS라 한다.

MOS 방법은 여러 피해 인자들(impairment factors)이 동시에 측정 가능하며, 훈련된 피험자가 필요없으므로 시험하기에 편리하다는 장점을 갖고 있다. 그러나 시험 조건의 선택과 다른 여러 요인들이 시험 결과에 영향을 미치는 문제점을 갖고 있다.

#### 1.2. Paired Comparison

MOS 방법으로 고품질 디지탈 통신 시스템 등을 평가할 경우에는, 대부분의 측정값들이 '매우 좋다'등급에 집중되어 정확한 차리를 구부할 수 없게 된다. 이러한 고품질 영역에서의 정확한 주관적 평가를 하기 위해서는 Paired Comparison이 필요하다. Paired Comparison은 피험자에게 두 신호를 제시하여, 더 나은 신호를 선택하게 하는 방법을 말하며, 표시 방법은 선택된 신호의 %를 나타내거나, 기준 신호에 값을 첨가하여 신호의 50% 수준까지의 preference level이 도달할때의 기준 신호의 신호대 잡음비로 나타내기도 한다. 후자의 평가치를 주관적 SNR이라고 한다.

Paired Comparison은 섬세하고 정확한 판단에 이용되는 방법이지만, 다른 피해 인자들로 구성된 두 신호를 비교할 경우에, 두 신호중 더 나은 신호를 판가름하기가 어렵게 된다는 문제점을 갖고 있다.

### 2. 음질의 객관적 평가

근본적으로 음질은 주관적 방법에 의해 평가되어야 하지만, 주관적 평가법은 매우 많은 비용과 시간이 소요되기 때문에, 주관적 평가 방법으로만 부호화 시스템과 통신망을 구성하기에는 불가능하다. 그러므로 주관적 평가법과 잘 일치되는 객관적인 방법에 의한 평가법이 바람직하다.

#### 2.1. 신호대 잡음비 (Signal-to-Noise Ratio)

SNR은 아날로그(analog) 시스템이나 파형 부호화 시스템에 적용되는 객관 측정 방법으로서, 입력과 출력 사이에 지연이 없는 왜곡 시스템에 적용 가능하다. 일반적인 SNR은 다음과 같다.



www.dbpia.co.kr

$$\mathrm{SNR}=10\log_{10}\frac{\sum_n X^2(n)}{\sum_n \{X(n)-Y(n)\}^2}\ [\mathrm{dB}] \qquad (3\text{-}1)$$

여기에서 X(n)은 입력 신호, Y(n)은 통신 시스템에 의해 재생된 음성 신호를 나타낸다.

이와같은 SNR에 시간에 따른 동적인 가중치를 가해주는 형식인 SEGSNR(Segmental SNR)은 다음과 같다.

$$\mathrm{SNRSEG}=E[\mathrm{SNR}(m)]\ [\mathrm{dB}] \qquad (3\text{-}2)$$

여기에서 SNR(m)은 m번째 단구간내의 SNR을, E는 시간에 대한 평균을 나타낸다. 이 방법은 묵음 구간이나, 음성의 에너지가 상당히 작은 구간을 포함할 경우, 전체 평가값에 영향을 주기 때문에, 이러한 구간들을 제거하는 작업이 필요하다.

### 2.2. FFT-SD(Fast Fourier Transform Spectral Distance)

주파수 영역에서 신호의 왜곡을 측정하는 척도로서 FFT-SD와 LPC-CD(Linear Predictive Coefficient Cepstral Distance를 이용한 방법이 주로 이용된다. FFT-SD 방법은 입·출력 신호의 스펙트럼 전체에 대한 차이로 왜곡을 측정하는 방법이다.[3],[4].

$$\mathrm{SD}=\sqrt{\frac{1}{W}\int\{Sx(f,t)-S\hat{x}(f,t)\,df} \qquad (3\text{-}3)$$

식 (3-4)에서 Sx(f,t), $S\hat{x}(f,t)$는 각각 입력 신호와 출력 신호의 Log Spectrum 이고, W는 주파수 영역에서의 적분 구간이다.

### 2.3. LPC-CD

스펙트럼 전체에 대한 왜곡을 측정하는 FFT-SD 방법과는 달리 LPC-CD 방법은 선형 예측 원리를 적용한 성도 모델링을 통하여, 입·출력 신호에 대한 포락 스펙트럼(envelope spectrum)을 구하고, 이들 간의 차이로써 왜곡의 정도를 추정한다.[3],[5].

$$\mathrm{CD}=10[\mathrm{Log}_e 10\sqrt{2\sum_{i=1}^{P}\{Cx(i)-C\hat{x}(i)\}]^{-1}} \qquad (3\text{-}4)$$

식 (3-5)에서 Cx(i), $C\hat{x}(i)$는 각각 입·출력 신호의 cepstrum 계수이고, P는 cepstrum 계수의 차수이다.

## Ⅳ. 실험 결과 및 분석

### 1. 데이터 베이스

본 연구에서 수행된 데이터는 훈련받지 않은 20대 남녀 각 1명이 5초 정도의 의미없는 문장을 100-3.4KHz 대역 필터 (bandpass filter)에 통과시키고, PCM과 ADPCM을 위해 8KHz, CFDM과 CVSD를 위해 32KHz로 샘플링 하여 16 비트로 A/D 변환시켰다.

### 2. 객관적 평가

본 연구에서 수행한 객관적 통화 품질 평가량으로서는 SEGNR, LPC-CD, FFT-SD를 사용하였으며, LPC-CD의 평가량을 구할시, 선형 예측 계수의 차수는 10차, FFT-SD 계산시에는 512point 퓨리에 변환을 수행하였다. SEGSNR을 구할 경우, 한 구간은 샘플링 주파수에 따라 변화시켰으며, LPC-CD의 계산식은 다음과 같은 두 식을 이용하였다.

$$D1=\frac{1}{N}\sqrt{2\sum_{i=1}^{10}\{Cx(i)-C\hat{x}(i)\}^2} \qquad (4\text{-}1)$$

$$D2=\frac{1}{N}\sqrt{\{Cx(1)-C\hat{x}(1)\}^2+2\sum_{i=2}^{10}\{Cx(i)-C\hat{x}(i)\}^2} \qquad (4\text{-}2)$$

여기에서 N은 한 프레임당 데이터의 수이고, Cx(i), $C\hat{x}(i)$는 각각 입·출력 신호의 cepstrum 계수이다.

또한 본 연구에서는 각 부호화 방법 수행시 발생할 수 있는 채널 비트에러(channel bit error)에 따른 객관적 통화 품질 평가량을 살펴보았다. [그림 4-1]에 각 부호화 방식에 의해 구한 스펙트럼을 나타내었으며, 각 부호화 방식의 객관적 평가량을 [표 4-1]과 [표 4-2]에 성별로 구분하여 각각 나타내었다.

[표 4-1]과 [표 4-2]에 나타난 것과 같이 32Kbps ADPCM 부호화 방식이 모든 객관적 평가량에서 다른 부호화 방식보다 에러가 적게되어 좋은 결과를 보여주고 있으며, 32Kbps, PCM, CFDM, CVSD의 순



www.dbpia.co.kr

으로 우수한 객관 평가량을 나타내고 있다.

그리고 채널 비트 에러에 따른 각 부호화 방식의 객관적 평가량을 살펴보면, $10^{-4}$비트당 랜덤(random)하게 에러 비트를 발생시킨 경우를 제외하고는 에러가 없는 경우와 비교하여 큰 차이의 특성 변화를 나타내고 있지 않다.

### 3. 주관적 평가

2절에서 구한 각 부호화 방식의 객관적 평가치와 주관적 평가치의 대응 관계를 살펴보기 위하여, 각 부호화 방식에 의해 구한 음성 품질을 오피니언(opinion) 시험을 통해 주관적 평가를 수행하였다.

오피니언 시험시 청취자수는 20대 남성5명, 여성 5명으로 수행하였으며, 입력 신호 레벨은 각 청취자가 각자 자신의 청각에 불쾌감을 주지 않은 신호 레벨로 조정하였다. 그리고 청취 기구는 헤드폰(headphone)을 사용하였으며, 대역 필터는 전송채널의 삽음의 영향을 포함시키기 위하여 광대역(100－9000Hz)으로 하였다. 녹음 청취 순서는 원음을 기준음으로 하여, 각 부호화 방식을 랜덤하게 배열하는 방법으로 편집하여 오피니언 시험을 수행하였다. 각 부호화 방식의 녹음 간격은 약 10초로 하여, 각 청취자가 판단할 시간을 부여하였다.

그리고 각 부호화 방식의 품질의 종합적인 만족도를 '아주 좋다'에서 '아주 나쁘다'까지 5단계로 구분하여 나타내도록 하였다.

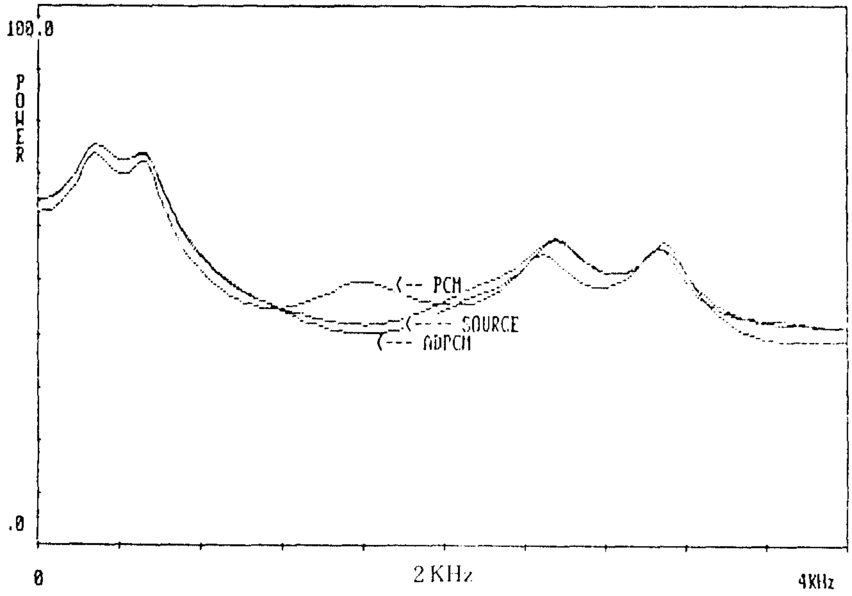
오피니언 실험을 한 평균 결과와 편차를 [표 4-3], [표 4-4]에 나타내었다.

그리고 이 오피니언 실험 결과를 객관 평가값과의 대응 관계를 검토하기 위하여, 객관 평가량과 주관 평가량의 관계를 [표 4-5]에 나타내었다.

[표 4-5]는 객관 평가량과 주관 평가량의 상관 계수와 근사식에 의한 편차를 나타낸 것이다. 이때의 근사식은 5차 다항식을 이용하였다. [표 4-5]에 나타난 것과 같이 SEGSNR 방법의 객관 평가치와 주관 평가치와의 상관 계수가 제일 큰 것으로 나타나는 것으로 보아, 음성 품질 평가에서 다른 객관 평가 방법보다 SEGSNR 방법이 주관적 평가값이 잘 일치하는 것으로 판단할 수 있으며, CEP-2, CEP-1, FFT의 방법의 순으로 객관 평가값과 주관 평가값의 상관 계수 크기를 보여주고 있다.

## V. 결 론

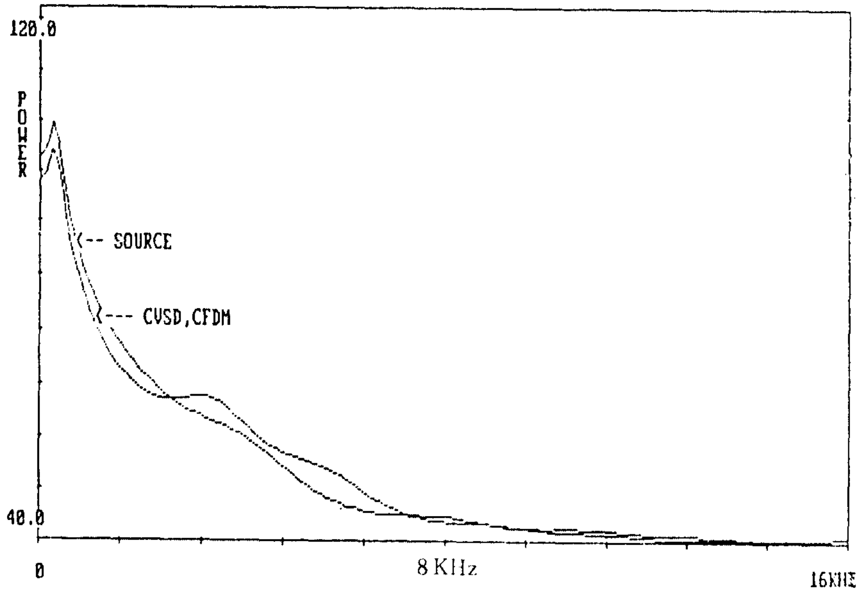
본 연구에서 각 부호화 방식의 평가량을 상호 비교하여 본 결과, 객관 평가에서는 32Kbps ADPCM 부




(a) 원음＋PCM＋ADPCM



www.dbpia.co.kr




(b) 원음+CFDM+CVSD

그림. 4-1 각 부호화 방식에 의한 출력 스펙트럼

[표 4-1] 각 부호화 방식의 객관적 평가(남성)

(A) SEGSNR

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 10.28 | 26.21 | 2.726 | 0.551 |
| 3 | 9.569 | 18.55 | 2.735 | 0.552 |
| 4 | 10.21 | 25.49 | 2.726 | 0.551 |
| 5 | 10.26 | 26.18 | 2.726 | 0.551 |

(B)CEPSTRUM-1 (식 4-1)이용

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 4.414 | 0.974 | 3.056 | 6.462 |
| 3 | 6.110 | 0.992 | 3.463 | 6.463 |
| 4 | 4.577 | 0.977 | 3.046 | 6.463 |
| 5 | 4.423 | 0.974 | 3.044 | 6.453 |

(C) CEPSTRUM-1 (식 4-2)이용

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 3.351 | 0.889 | 2.539 | 4.726 |
| 3 | 4.571 | 0.903 | 2.832 | 4.727 |
| 4 | 3.468 | 0.890 | 2.533 | 4.727 |
| 5 | 3.358 | 0.889 | 2.532 | 4.726 |

(D) FFT-SD

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 279.2 | 99.84 | 191.8 | 406.0 |
| 3 | 401.7 | 107.8 | 196.2 | 405.7 |
| 4 | 291.9 | 100.8 | 190.8 | 406.1 |
| 5 | 279.8 | 99.86 | 189.9 | 406.3 |



www.dbpia.co.kr

[표 4－2] 각 부호화 방식의 객관적 평가(여성)

(A) SEGSNR

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 12.29 | 31.01 | 2.806 | 0.618 |
| 3 | 11.35 | 20.74 | 2.794 | 0.621 |
| 4 | 12.24 | 29.76 | 2.804 | 0.618 |
| 5 | 12.28 | 30.83 | 2.807 | 0.615 |

(B)CEPSTRUM-1 (식 4-1)이용

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 3.893 | 0.738 | 2.987 | 6.051 |
| 3 | 5.090 | 0.804 | 3.041 | 6.058 |
| 4 | 4.116 | 0.742 | 2.989 | 6.052 |
| 5 | 4.024 | 0.744 | 2.987 | 6.052 |

(C) CEPSTRUM-1 (식 4-2)이용

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 2.964 | 0.645 | 2.465 | 4.427 |
| 3 | 3.827 | 0.697 | 2.501 | 4.429 |
| 4 | 3.224 | 0.648 | 2.466 | 4.428 |
| 5 | 3.057 | 0.650 | 2.463 | 4.427 |

(D) FFT-SD

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 223.2 | 76.39 | 239.4 | 388.8 |
| 3 | 316.1 | 90.76 | 246.4 | 389.8 |
| 4 | 240.5 | 79.62 | 239.5 | 388.8 |
| 5 | 230.9 | 77.24 | 238.5 | 388.7 |

호화 방식이 다른 부호화 방식보다 더 나은 평가 결과를 얻었으며, 32Kbps PCM, CFDM, CVSD 순으로 좋은 객관적 평가값을 구하였으며, 채널 비트 에러에 따른 객관 평가값들은 $10_{3}$ 비트당 랜덤하게 에러를 발생시킨 경우를 제외하고는 각 부호화 방식에 따라서 큰 차이를 얻지 못하였다.

그리고 오피니언 실험을 수행한 결과, 4가지 객관 평가방법 중에서 SEGSNR 객관 평가 방법의 결과가 주관 평가 결과와의 0.97106의 상관 계수를 얻어, 다른 객관 평가 주관 평가량과 잘 일치 대응되는 평가 방법임을 알 수 있었다. 한편 주파수 영역에서의 평가값인 LPC-CD와 FFT-SD는 각 성별에 따라 다른 결과를 보여주고 있지만, FFT-SD 보다 LPC-CD가 주관적 평가량에 더 접근된 결과를 보여주고 있다.

본 연구에서는 주관 평가량으로 MOS만을 이용하였으나, 앞으로 다른 주관 평가량과 MOS에 의해 구한 주관 평가량의 차이를 구하여볼 필요가 있으며, 또한 채널 비트 에러에 따른 영향외에, 백색 잡음이나 음성의 에너지에 따른 잡음의 영향도 연구할 과제로 남아있다.

[표 4－3] 오피니언 시험 결과(남성)

(A) 평균

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 1.4 | 2.4 | 1.2 | 1.4 |
| 3 | 1.6 | 2.2 | 0.8 | 1.0 |
| 4 | 1.8 | 3.4 | 1.5 | 1.7 |
| 5 | 1.4 | 3.5 | － | － |

(B) 편차

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 1.02 | 0.92 | 0.98 | 0.92 |
| 3 | 0.49 | 0.60 | 0.40 | 0.89 |
| 4 | 0.87 | 0.49 | 0.92 | 0.60 |
| 5 | 0.92 | 0.50 | － | － |



www.dbpia.co.kr

[표 4-4] 오피니언 시험 결과(여성)

(A) 평균

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 1.9 | 3.8 | 1.0 | 1.2 |
| 3 | 1.8 | 3.3 | 1.0 | 1.2 |
| 4 | 1.9 | 3.7 | 1.3 | 1.6 |
| 5 | 1.9 | 3.7 | – | – |

(B) 편차

| CODING / Channel bit error | PCM | ADPCM | CFDM | CVSD |
|---|---|---|---|---|
| 0 | 0.54 | 0.40 | 0.63 | 0.60 |
| 3 | 0.75 | 0.46 | 0.77 | 0.60 |
| 4 | 0.89 | 0.46 | 0.64 | 0.80 |
| 5 | 1.14 | 0.46 | – | – |

[표 4-5] 객관 평가와 주관 평가의 상관 관계

(A) 남 성

| | SEG | CEP-1 | CEP-2 | FFT |
|---|---|---|---|---|
| 상관편수 | .87767 | .71253 | .70590 | .60832 |
| 편 차 | .10663 | .18194 | .17311 | .21424 |

(B) 여 성

| | SEG | CEP-1 | CEP-2 | FFT |
|---|---|---|---|---|
| 상관편수 | .97106 | .78522 | ..82717 | .84497 |
| 편 차 | .05217 | .18974 | .15757 | .09506 |

## 참 고 문 헌


1. J.L. Flanagan, M.R.Schroeder, B.S.Atal, M.S. Jayant and J.M.Tribolet, "Speech Coding," IEEE Trans. Commun., Vol.27, No.4, April. 1979.
2. CCITT, "Recommendation G.721 : 32Kbit/s Adaptive Differential Puls Code Modulation (ADPCM)" Red Book, Vol. Ⅲ, Oct. 1984.
3. N.Kitawaki, K. Itoh, M.Honda and K.Kakehi, "Comparison of Objective Speech Quality Measures for Voiceband Codecs" Proc. IEEE ICASSP '82, Paris, France, pp.291-294, April. 1980.
4. N.Kitawaki, K.Itoh and M.Honda, "Speech Quality Assessment Methods for speech coding systems" IEEE Communication magazine, Vol. 22, No.10, pp.26-33, Oct. 1984.
5. 김원구, "스펙트럼 매칭 방법에 따른 소음 환경에서의 단독음인식" 석사학위 논문, 연세대학교 대학원, 8월. 1989.
6. N.Kitawaki and H.Nagabuchi, "Quality Assessment of Speech Coding and Speech Synthesis System," IEEE Commumnication magazine, pp.36-44, Oct. 1988.
7. 伊藤憲三, 北脇信彦, "高能率符號化 音聲の品質評價," 電子通信學會論文志, Vol.J68-A, No.9, pp. 183-190, 9. 1985.
8. N.S.Jayant, "Digital Coding of Speech Waveforms : PCM, DPCM, and DM quantizers," Proc. IEEE, Vol.62, pp.611-632, May. 1974.
9. N.S.Jayant, "Adaptive delta modulation with a one bit memory," Bell Syst. Tech.J., Vol.49, pp.321-342, Mar. 1970.
10. J.E.Abate, "Linear and adptive delta modulation," Proc. IEEE, Vol.55, pp.298-308, Mar. 1967.




www.dbpia.co.kr

**裵 澈 秀**(Chul Soo Bai) 正會員
1950년 3월 14일생
1979년 2월 : 명지대학교 전자공학과 졸업
1982년 2월 : 명지대학교 대학원 전자공학과 석사
1988년 8월 : 명지대학교대학원 전자공학과 박사
1981년 3월~1990년 11월 : 관동대학교 전자공학과 부교수
1990년 12월~현재 : 관동대학교 전자통신과 부교수
1989년 11월~현재 : 한국통신학회 강원지부장



www.dbpia.co.kr