

# CDMA 이동전화 통화품질평가를 위한 객관적 음질평가척도별 성능 비교

이 준 희\*\*, 김 광 수\*, 윤 정 오\*

\*경운대학교 컴퓨터 전자정보공학부

## Performance Comparison of Objective Measures for Speech Quality for Evaluation in CDMA Mobile Telephone

\*\* , Kwang-Soo Kim\* , Jung-Oh Yun\*

\*School of Computer, Electronics and Information Eng., Kyungwoon University

### 요 약

본 논문에서는 디지털 이동전화(CDMA) 채널환경을 통과한 왜곡된 전화음성에 대해 객관적 음질평가 척도의 개발을 위한 기초 연구로서 기존의 CD(Cepstral Distance), MSD(Mel Spectral Distance), BSD(Bark Spectral Distance), Modified BSD, PSQM(Perceptual Speech Quality Measure)를 대상으로 객관적 알고리즘을 성능평가 하였다. 이 척도들은 실제 이동전화 환경에서 수집된 PCS 음성데이터에 대해서 적용하였으며 이 결과치를 주관적 음질평가 방법인 MOS와 상관성을 비교 조사하였다. 실험 결과, BSD와 MBSD, 그리고 PSQM 척도의 상관성이 각각 0.80, 0.85, 0.84로 나타났으며 CD, MSD보다 성능이 상대적으로 더 우수함을 보였다.

가 기술은 이동 전화의 통화 품질분야에서 가장 중요한 문제중의 하나이다. 이 평가 작업의 주된 내용은 전파환경내의 여러 잡음 조건과 채널 손상 등 아래에서 수신된 이동 전화음성에 대해 사용자의 주관적 음질을 정확하게 예측하는 객관적 척도를 찾는 것이다.

본 연구의 목적은 현재 알려져 있는 여러 척도들 중에서 CD, MSD, BSD, MBSD, PSQM을 대상으로 실제 전파환경에서 수집한 이동 전화 음성에 대해 청취자의 주관적 평가(MOS)와 어느 정도 상관성을 가지는지를 조사하고 이들 척도들간의 성능을 비교 분석하는데 있다. 본 논문은 PCS 이동 통신 채널을 통하여 수집된 왜곡된 음성data를 대상으로 척도별 알고리즘 성능비교를 수행하였다.

### I. 서 론

일반적으로 이동 전화 사용자가 체감하는 음성의 통화품질을 알기 위해서는 다양한 전파채널 환경 내에서 반복 청취 실험에 의한 주관적인 평가를 실시해야 하는데 이 방법은 시간과 비용이 많이 소요되므로 신속히 음질 평가를 하기에는 많은 제한이 따르며 또한 평가 결과의 일관성이 부족하다는 단점이 있다.[1][2].

이동 전화와 같은 무선 전송채널 환경에서 발생하는 왜곡 원인으로는 채널의 대역 통과 특성과 부가되는 잡음, 그리고 복수개의 전파 경로를 통해 수신측에서 불규칙적으로 합성되어 발생하는 다중 경로 페이딩 현상이 있다.

이러한 다양한 상황을 반영한 이동 전화의 주관적 음질을 간접적으로 예측할 수 있는 음질의 자동 평

### II. 음질 평가 방법

#### 2.1 주관적 음질 척도

표1. 주관적 음질 평가 척도

명료성 시험 (articulation)	- 명료도 - AEN (명료도 듣기감쇄량) - 요해도, 등			
자연성 시험 (naturalness)	- loudness(음량감)			
opinion 평가 시험	- opinion 평가 (loudness로 평가되지 않는 잡음과 왜곡 등을 음성신호와 종합적으로 평가하고자 할 때 사용)	- Listening-only Test (청취 시험)	Absolute Category Rating (ACR)	MOS
		- Conversat.on Test (회화 시험) - Talking Test (송화 시험)	Degradation Category Rating (DCR)	DMOS

전화망 음질 평가에는 보통 표1.의 MOS(Mean Opinion Score)평가법을 자주 이용한다. MOS는 청취자의 반응 평가시 5단계 scale에 따라 5~1점의 점수를 주어 청취 시험에 참가한 다수 청취자에 의한 반응 의견에 대해서 가중 평균치를 구하는 척도이다.

## 2.2 객관적 음질 척도

음질평가 척도의 성능은 주관적인 청취 시험 결과인 MOS값과의 상관성 정도에 의해 평가되는 것이 일반적이다. 원 음성과 왜곡된 음성간의 왜곡량 계산시 사용되고 있는 측정 파라메타는 보통 표2.와 같이 4가지 형태가 이용되고 있다.[3]

표2. 객관적 음질 평가 척도

파형 왜곡 척도	음성의 파형 특성	SNR-based measures
스펙트럼 왜곡 척도	스펙트럼 전체적 특성을 고려	-SNR, Segmental SNR
스펙트럼 포락선 왜곡 척도	스펙트럼 포락선 특성을 고려	Spectral Distance-based measures -Spectral Distance, CD
지각적 왜곡 척도	귀내 음압 스펙트럼 레벨의 특성(청각특성)을 고려	LPC-based measures -LPC, Log LPC
		Perceptual based-measures -MSD, BSD, MBSD, PSQM

최근에는 BSD와 PSQM 등과 같이 청각의 심리 음향 모델을 반영한 지각적인 왜곡척도가 통화품질 연구에 많이 적용되어 왔다.[4][5]

특히 이러한 척도들은 상대적으로 높은 상관성을 나타내는 것으로 알려져 있어 본 연구에서도 BSD, PSQM척도의 알고리즘을 중심으로 성능을 비교 분석하였다.

### 2.2.1 CD(Cepstral Distance) 척도

CD척도는 식(1)과 같이 원 음성과 왜곡된 음성 사이에서 cepstral 계수간 차이를 계산한다.

$$CD = \frac{10}{\log 10} \sqrt{2 \sum_{i=1}^K [(C_x(i) - C_y(i))]^2} \quad (1)$$

$K$  : 프레임 수,

$C_x(i), C_y(i)$  : 원 음성, 왜곡 음성 cepstral 계수

### 2.2.2 BSD(Bark Spectral Distance) 척도

사람의 귀 내 기저막의 소리에 대한 임계대역 특성은 Bark scale로 나타내는데 이 과정에서는 지각적 모델의 3 단계를 거쳐 소리의 power 스펙트럼  $P(f)$ 가 Bark 스펙트럼  $B(z)$ 으로 변환(mapping)된다. 라우드니스를 계산하기 위해서는 그림1.의 3가지 단계를

거치게 된다.

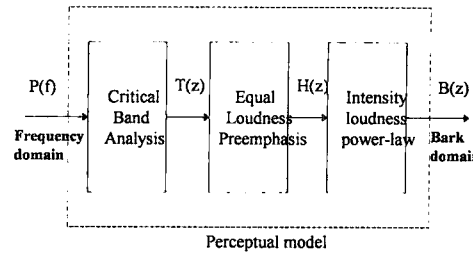


그림1. Perceptual model

BSD척도는 원 음성과 왜곡된 음성의 Bark 스펙트럼(청각 power 스펙트럼) 차이를 계산하며 임계 대역당  $BSDB(i)$ 는 식(2)와 같이 구해진다[6].

$$BSDB(i) = \frac{\sum_{l=1}^N dis\{B'_x(i), B'_y(i)\}}{\sum_{l=1}^N \sum_{i=1}^b [B'_x(i)]^2} \quad (2)$$

$B'_x(i), B'_y(i)$  : 입,출력프레임  $l$ 의 Bark 스펙트럼

$N$  : 음성 프레임의 수,  $b$  : 임계 대역 수

BSD척도의 계산과정은 그림2.와 같다.

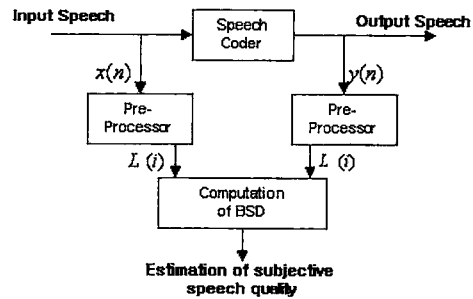


그림2. BSD척도 계산과정

### 2.2.3 MBSD(Modified BSD) 척도

이 척도는 BSD 계산시 noise masking threshold 개념을 도입하였다. 사람의 청각 기능은 음성신호에서 masking된 부분은 지각하지 못하므로 척도계산시에 제외하는 것이 바람직하다. 따라서, 왜곡량 계산시 지각된 audible distortion만을 계산하도록 하기 위해 noise masking threshold 개념을 도입하였다.

BSD와 MBSD의 차이점은 BSD는 loudness의 평균 Euclidean 거리로서 왜곡량을 계산하는 반면에, MBSD는 loudness의 평균 차이로서 왜곡량을 계산하고 audible distortion을 결정하기 위해 noise

masking threshold를 사용하는 것이다. MBSD의 계산과정은 그림3과 같다.

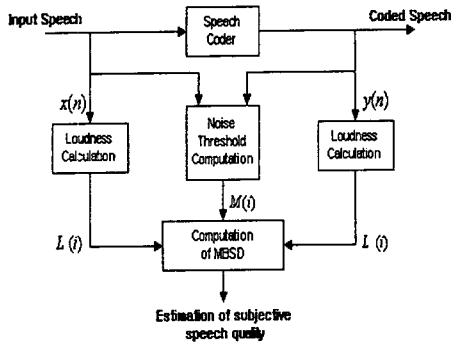


그림3. MBSD 계산과정

### 2.2.4 MSD(Mel Spectral Distance) 척도

Mel scale은 소리 tone의 크기를 청각적인 감각에 상응하도록 변환시키며 주파수와 Mel scale과의 관계는 식(3)으로 근사화된다.

$$\text{Mel} = 1000 \log_2(1 + f) \quad (3)$$

### 2.2.5 PSQM(Perceptual Speech Quality Measure)

ITU-T 권고안 P.861에서는 전화망내 음성 신호의 전송품질 평가시 PSQM 알고리즘으로 품질 평가하도록 권고하고 있다[4].

이 척도도 음성 왜곡량에 대한 심리 음향적 표현의 차이로서 나타내며 이 차이를 두 음성간의 평균 noise disturbance로 표현하고 있다. 그림4는 PSQM의 계산과정을 나타낸 것이다. PSQM 알고리즘에서는 성능의 향상을 위하여 지각적 모델외에 인지(cognitive)모델을 사용하고 있다[4].

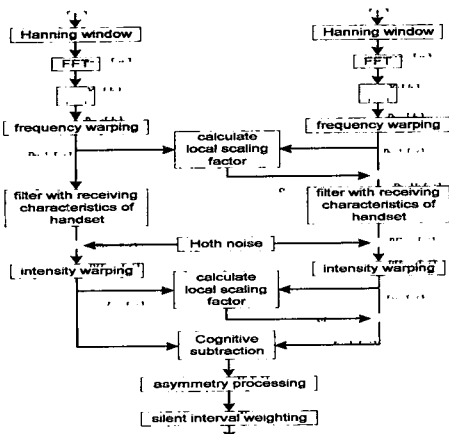


그림4. PSQM 계산과정

### 2.3 MOS와 객관적 척도와의 상관성

객관적 음질척도로부터 주관적 음질(MOS)을 예측하는데는 식(4)의 이차 회귀 함수가 널리 사용되고 있으며 우수한 성능을 나타낸다[1].

$$\hat{y} = ax^2 + bx + c \quad (4)$$

$x$  : 객관적 척도(CD, BSD, MSD, PSQM)

$\hat{y}$  : 예측된 MOS값

청취 test에 의한 MOS값과 예측된 MOS값 사이의 상관계수( $\rho$ )는 식(5)으로 구해진다.

$$\rho = \sqrt{\frac{\sum(\hat{y}_i - m_y)^2}{\sum(y_i - m_y)^2}} \quad (5)$$

$m_y$  : 평균 MOS값

$\hat{y}_i, y_i$  : 예측 및 측정된 MOS값

$\rho = 1$ 일 경우, MOS 예측 오차가 없음을 의미한다.

## III. 음질 평가 실험

본 연구에서 객관적 척도의 음질 평가 성능을 비교 분석하기 위해 수행한 과정은 다음과 같다[3].

- 1) speech database 구성
- 2) PCS phone을 통한 왜곡된 음성 수집
- 3) 수집된 음성(왜곡된 음성) data의 MOS 측정
- 4) 객관적 음질척도 계산
- 5) MOS와 객관적 척도와의 상관성 조사

### 3.1 speech database 구성

원 음성은 연속음성 데이터베이스에서 각각 8초 정도의 발성길이를 가진 5문장을 선택하였다.

### 3.2 PCS phone을 통한 왜곡된 음성 수집

PCS 이동전화 채널을 통과한 실제의 음성데이터를 수집하기 위한 과정은 그림5와 같으며 다양한 전파 채널 환경에 따른 왜곡된 음성을 주·야간 각3회에 걸쳐 수집하였다.

원 음성을 휴대형 DAT에 녹음하고 이것을 PCS 단말기에 접속시켜 차량을 정지 또는 정속 운행(60km 정도)하면서 PCS 전화기를 통하여 송출하였다. 연구실에서는 일반 전화기를 통하여 수신된 음성 신호를 desktop DAT에 저장하여 database화하였다. 원 음성 1문장당 80개의 왜곡된 음성을 수집하였다. 표3은 이동 차량을 통해 채록한 PCS음성데이터의 수집환경을 나타낸 것이다.

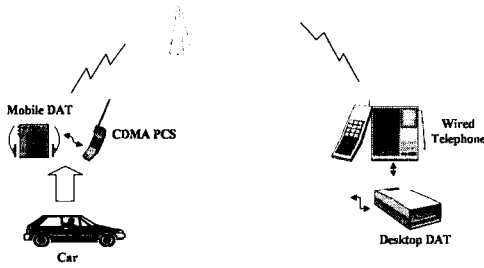


그림5. 왜곡된 음성data 수집 과정

표3. 음성 데이터 수집환경

1. PCS phone	송신: 1751.25~1778.75MHz 수신: 1841.25~1868.75MHz
2. 송신 출력	0.2+0.1 W
3. 음성 데이터량	1개 문장당 80개의 음성데이터 (전체 400개의 문장을 수록)
4. 차량 속도	정지 또는 주행(평균 60km/s)

### 3.3 왜곡된 음성에 대한 MOS test

MOS평가는 CCITT 권고안 562-3을 근거로 왜곡된 음성에 대해 청취자들 각자의 품질 반응 의견 (opinion)을 평가용지에 기입하도록 하였다.

MOS test 후 청취자가 기입한 평가 반응지에 나타난 시험결과는 문장당 5단계로 구분하여 각 단계에 5~1점의 가중치를 부여하고 MOS값을 계산한다.

### 3.4 객관적 음질척도 계산

저장된 왜곡된 음성은 원 음성과의 파형 동기화를 수행한 후 두 음성 신호간의 왜곡량 차이를 척도 알고리즘별로 계산하여 성능을 비교 분석한다.

## IV. 성능 비교 분석

PCS채널 환경을 통해 수집된 실제 음성에 대해 척도별 얻어진 상관계수는 그림6과 같이 나타났으며 상관계수의 평균값은 표4와 같다.

그림6의 가로축은 MOS평가시 사용된 왜곡된 문장 수이며 세로축은 그 문장에 대해 구한 상관계수이다.

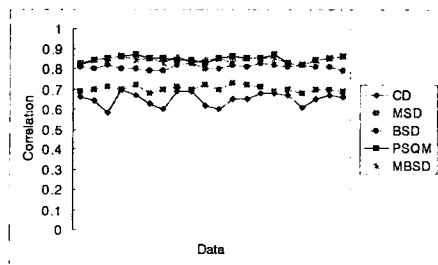


그림6. 척도별 상관계수

	평균 상관 계수
CD	0.65
MSD	0.70
BSD	0.80
MBSD	0.85
PSQM	0.84

표4. 척도별 평균 상관계수

왜곡된 음성에 대해 MOS와의 상관성을 비교 조사한 결과 BSD, MBSD와 PSQM척도는 CD, MSD척도에 비하여 10% 이상 상관성이 높게 나타나 주관적 음질을 더 잘 예측할 수 있음을 알 수가 있다.

또한 CD척도는 상대적으로 상관계수의 변동이 크게 나타나서 PCS 이동전화 환경의 음질 평가척도로는 적합하지 않음을 알 수가 있다.

## V. 결 론

본 연구에서는 디지털 이동전화의 통화음성에 대해 사용자의 음질 평가를 잘 예측하는 객관적 척도를 개발하기 위한 기초단계의 연구를 수행하였다.

실제 전파환경에서 PCS채널 음성데이터를 수집하여 MOS와의 상관성 비교를 통해 CD, MSD, BSD, MBSD, PSQM 척도의 성능을 비교 분석하였다. 실험 결과, BSD, MBSD 척도와 PSQM 척도가 각각 0.81, 0.85, 0.84정도로 상대적으로 더 높게 나타났음을 확인하였다. 이것은 PCS이동 전화음성에 대한 통화품질 평가척도로서 유효함을 나타낸다.

향후, 본 연구에서는 이러한 결과를 근거로 하여 향후 PCS 이동전화 음성에 대해 심리 음향적 모델을 기반으로 한 여러 척도 알고리즘 연구와 다양한 환경요인에 따른 성능비교 분석을 통해 보다 여러 지역의 전파 채널 환경에서도 더 높은 상관성을 나타내는 강건한 음질 척도 연구가 진행 중이다.

또한, 이동 통신 채널환경에서 다중 경로 페이딩 현상은 통신시스템의 전송 품질왜곡을 일으키는 주요 요인중의 하나로 취급되고 있다. 따라서 통화품질 척도의 개발 측면에서 이것이 평가척도의 성능에 어느 정도 영향을 미치는지를 확인해 보고자 한다.

## 참 고 문 헌

- [1] K.H. Lam, O.C. Au, C.C. Chan, K.F. Hui, S.F. Lau, "Objective Speech Measures for Chinese in Wireless Environment", ICASSP, vol.1, pp277-280, 1995.

- [2] 박상욱 외 4인, "아날로그 셀룰라 시스템을 위한 자동 음질 평가기 개발", 한국음향학회지, 제17권, 제7호, pp-28-35, 1998.
- [3] K.H. Lam, O.C. Au, C.C. Chan, K.F. Hui, S.F. Lau, "Objective Speech Quality Measure for Cellular Phone", ICASSP, vol.1, pp.487-490, 1996.
- [4] ITU-T Rec. P.861, "Objective Quality Measurement of Telephone-based Speech Codecs", Geneva, 1996.
- [5] S. Wang, A. Sekey, A. Gersho, "An Objective Measure for Predicting Subjective Quality of Speech Coders", IEEE J. on Select Areas in Comm., vol.10, No.5, pp819-829, 1992.
- [6] Mohamed M. Meko, and N. Saadawi, " A Perceptual-Based Objective Measure for Speech Codecs Using Abductive Network", ICASPP, vol.1, pp479-482, 1996.