

PCS 이동전화망에서의 객관적인 음질평가척도별 성능비교

김낙철**, 김광수*, 정호열*, 정현열*

*영남대학교 정보통신공학과
**대구기능대학 전자기술학과

Performance Comparison for Objective Measures of Speech Quality Evaluation in PCS Wireless Telephone Network

Nag-Cheol Kim**, Kwang-Soo Kim*, Ho-Youl Jung*, Hyun-Yeol Chung*

*Dept. of Information and Communication Eng., Yeungnam University

**Dept. of Electronic Technology, Taegu Polytechnic College

요약

본 연구에서는 PCS 이동전화의 객관적 통화품질 평가 척도개발을 위한 기초연구로 기존의 CD (Cepstral Distance), MSD (Mel Spectral Distance), BSD (Bark Spectral Distance), PSQM (Perceptual Speech Quality Measure) 척도를 적용하여 그 성능을 비교 분석하였다. 이 척도들을 실제환경에서 수집된 PCS 음성데이터에 대해서 적용하였고 이 결과치와 청취자들의 평가 반응에 의해 얻어진 MOS 결과치와의 상관성을 조사하였다. 실험 결과, BSD와 PSQM 척도의 상관성이 0.81, 0.84로 나타나 CD, MSD보다 성능이 더 우수함을 보였다.

척도를 찾는 것이다.

본 연구의 목적은 PCS 이동전화의 객관적 통화품질 평가 척도를 개발하기 위한 기초 연구로서 현재 알려져 있는 여러 척도들 중에서 CD, MSD, BSD, PSQM을 대상으로 실제 전파환경에서 수집한 이동전화 음성에 대해 청취자의 반응에 대해 어느 정도 상관성이 있는지를 조사하고 이들 척도들간의 성능을 비교 분석하는데 있다.

국내의 경우 CDMA채널의 PCS 전파환경을 통한 통화품질 평가 연구는 미흡하기 때문에 PCS 이동전화를 사용한 왜곡된 음성신호를 대상으로 척도별 비교 연구를 수행하였다.

II. 음질 평가 방법

I. 서론

이동 전화 사용자가 체감하는 통화 서비스 만족도는 대체로 통화 성공률과 통화품질로 평가된다고 할 때 통화 품질은 전화 사용자가 청각으로 느끼는 주관적 음질 평가에 의해 좌우된다. 이동전화의 통화 품질을 알기 위해서는 다양한 전파채널 환경내에서 사용자의 반복 청취 실험에 의한 주관적 평가를 실시해야 하는데 이 방법은 시간이 많이 소요되고 비용이 많이 들기 때문에 평가를 신속히 하기에는 많은 제한이 따른다[1][2].

주관적 음질을 간접적으로 예측할 수 있는 음질의 자동 평가는 이동 전화의 통화 품질분야에서 가장 중요한 문제중의 하나이다. 이 평가의 주된 내용 중 하나는 전파환경내의 변화하는 잡음 조건과 채널 손상 등 아래에서 수신된 음성 신호에 대해 사용자(청취자) 반응(acceptance)을 정확하게 예측하는 객관적

2.1 주관적 음질 척도

전화망 음질 평가에는 오피니언(opinion) 평가법이 주로 사용되고 있으며 그 중에 평가 scale이 excellent, good, fair, poor, bad의 5단계인 MOS(Mean Opinion Score)평가법을 많이 이용한다. 청취 반응 평가시 각 scale에 따라 5~1점의 점수를 주어 다수 청취자에 의한 반응 의견에 대해서 가중 평균치를 구하는 것이 MOS이다.

2.2 객관적 음질 척도

음질 평가 척도의 성능은 주관적인 청취 시험 결과인 MOS값과의 상관성 정도에 의해 판단될 수 있다. 즉, 상관성이 높을수록 유효성이 높다고 할 수 있다.

두 음성 신호간의 왜곡량 계산시 사용되고 있는 측정 파라메타는 표1과 같이 파형 왜곡척도, 스펙트

림 왜곡척도, 스펙트럼 포락선 왜곡 척도 그리고 지각적 왜곡 척도의 4가지 형태가 이용되고 있다[3].

표 1. 왜곡량에 대한 계산 파라메타

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

수 년 전부터 다양한 객관적 척도 연구분야에서는 BSD와 PSQM과 같은 청각의 심리 음향 모델을 반영한 지각적인 왜곡척도가 통화품질 연구에 많이 적용되어 왔다[4][5].

사람의 청각특성을 고려한 이러한 척도들은 상대적으로 높은 상관성을 나타내는 것으로 알려져 있으므로 본 연구에서는 CD와 MSD, BSD, PSQM척도의 알고리즘을 중심으로 성능을 비교 분석하였다.

2.2.1 CD 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 척도

Bark scale은 사람의 귀내 기저막의 소리에 대한 임계대역 특성에 대응하며 귀 내에서 지각적 모델의 3 단계를 거쳐 소리의 전력스펙트럼 $P(f)$ 가 Bark 스펙트럼 $B(z)$ 으로 변환(mapping)되어진다.

지각적 모델은 심리 음향적 항인 청각 감각의 크기로서 정의되는 음성 라우드니스(Loudness)에 직접 상관된다고 하는 가정에 근거하며 라우드니스를 계산하기 위해 그림 1의 3가지 단계를 거친다

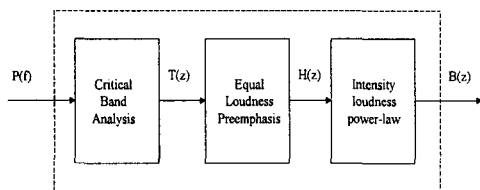


그림 1. Perceptual model

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

$$BSDB(i) = \frac{\sum_{l=1}^N dis\{B_x^l(i), B_y^l(i)\}}{\sum_{l=1}^N \sum_{j=1}^b [B_x^l(i)]^2} \quad (2)$$

$B_x^l(i), B_y^l(i)$: 입,출력프레임 l 의 Bark 스펙트럼

N : 음성 프레임의 수, b : 임계 대역 수

$BSDB(i)$ 의 계산과정은 그림 2와 같다.

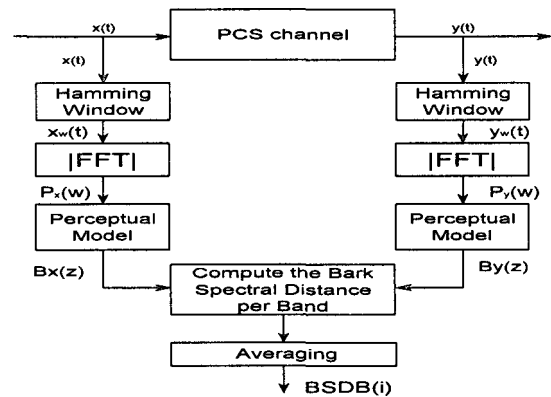


그림 2. BSDB(i) 계산과정

2.2.3 MSD 척도

BSD는 Bark영역에서 스펙트럼의 차이를 나타내는 반면에 MSD는 Mel scale 영역에서의 스펙트럼의 차이를 계산한다.

Mel scale은 소리 tone의 크기에 대한 청각적인 감각에 상응하도록 주파수를 변환시키며 주파수와 Mel scale과의 관계는 식(3)으로 근사화된다. Bark scale과 Mel scale은 1kHz이상의 주파수 대역에서 대수적인 주파수 스케일에 거의 비례한다.

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

2.2.4 PSQM 척도

전화망내 음성 신호의 전송품질 평가시 ITU-T 권고안 P.861에서는 PSQM 알고리즘으로 측정하도록 되어 있다[4].

PSQM 척도도 음성 왜곡량에 대한 심리 음향적 표현의 차이에 따라 결정되며 이 차이는 두 음성간의 noise disturbance로 주어진다.

이 평균 noise disturbance가 음질 특성을 나타내며 그림 3은 PSQM의 계산과정을 나타낸 것이다. 또한 PSQM 알고리즘에서는 주관적 평가와의 사이에 높은

상관성을 얻기 위해 지각적 모델외에 인지(cognitive) 모델링을 사용한다[4].

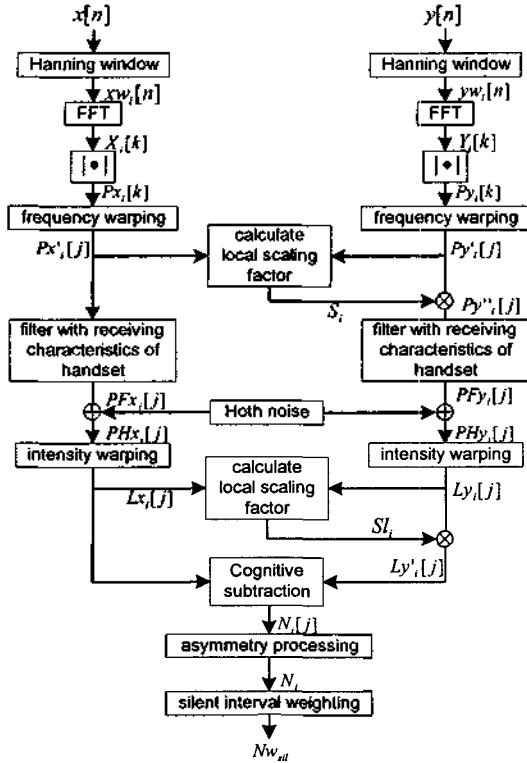


그림 3. PSQM 계산과정

2.3 MOS와 객관적 척도와의 상관성

계산된 객관적 척도로부터 MOS를 예측하는 방법에는 몇 가지 방법중 널리 이용되고 있는 식(4)의 이차 회귀 함수식을 사용하였다[1].

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

x : 객관적 척도(CD, BSD, MSD, PSQM)

\hat{y} : 예측된 MOS값

a, b, c : 최소자승회귀(Least Square Regression)에 의하여 구한 계수

청취 test에 의한 MOS값과 예측된 MOS값 사이의 상관관계수(ρ)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, 객관적 척도가 실제 환경의 주관적 음질을 완벽하게 예측함을 의미한다.

III. 음질 평가 실험

본 연구에서 4가지 객관적 척도의 음질 평가 성능을 비교 분석하기 위해 수행한 평가 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

본 연구에서는 국어공학연구소(KLE) 음성 데이터베이스에 수록된 문장 중에서 선정하여 사용하였다.

원 음성은 데이터베이스에서 무성음과 유성음이 균형있게 골고루 분포되어 있는 40초정도의 발성걸이를 가진 5문장(1문장 : 약 8초)을 선택하였다.

3.2 PCS phone을 통한 왜곡된 음성 수집

PCS 이동전화 채널을 통과한 실제의 음성데이터를 수집하기 위한 과정은 그림4와 같으며 다양한 전파 채널 환경에 따른 왜곡된 음성을 주, 야간에서 3회에 걸쳐 수집하였다.

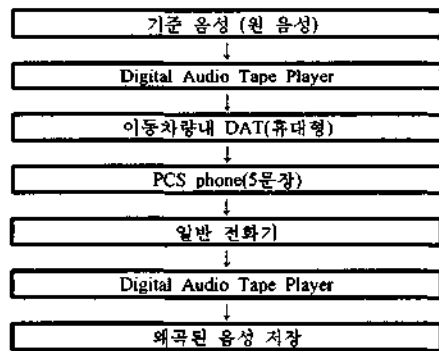


그림 4. 왜곡된 음성data 수집 과정

표 2는 이동 차량을 통해 채록한 PCS음성데이터의 녹음 환경을 나타낸 것이다

표 2. 수집된 음성 데이터 녹음환경

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 하여 청취된 40문장에서 각 왜곡된 문장에 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. 계산사 사용된 파라메타는 표 3과 같다.

표3. 척도별 계산 파라메타

CD	cepstral계수간의 평균 자승 Euclidian Distance
BSD	Bark 스펙트럼간의 평균자승 Euclidian Distance
MSD	Mel 스펙트럼간의 평균자승 Euclidian Distance
PSQM	평균 Noise Disturbance

대부분의 척도가 동기화의 정확도에 따라 민감한 영향을 받으므로 두 신호 파형의 시간축 정렬 작업의 경우 본 연구에서는 Flag 신호를 원 신호에 부가하여 두 음성 신호간의 파형 동기화시켰다.

IV. 성능 비교 분석

PCS채널 환경을 통해 수집된 실제 음성에 대해 척도별 얻어진 상관계수는 그림5와 같이 나타났으며 상관계수의 평균값은 표4와 같다.

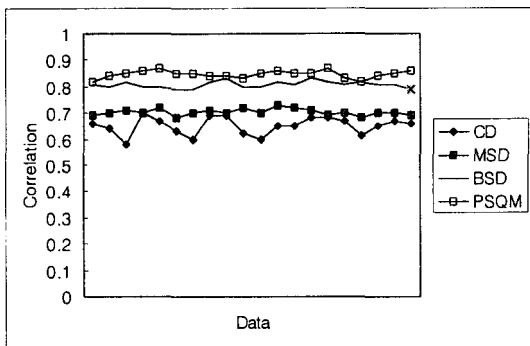


그림 5. 척도별 상관계수

그림 5의 가로축은 MOS 평가시 사용된 왜곡된 문장수이며 세로축은 그 문장에 대해 구한 상관계수이다. BSD와 PSQM척도는 왜곡된 음성에 대해 CD, MSD에 비하여 10% 이상 상관성이 높게 나타나 주

관적 음질을 더 잘 예측할 수 있음을 알 수 있다.

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

표 4. 척도별 평균 상관계수

척도	평균 상관 계수
CD	0.65
MSD	0.70
BSD	0.81
PSQM	0.84

V. 결 론

PCS 통화음성에 대해 사용자의 음질 평가 반응을 정확하게 예측하는 객관적 척도를 개발하기 위한 기초 연구로 음성데이터를 실제 수집하여 CD, BSD, MSD, PSQM 척도의 성능을 비교 분석하였다.

이들 척도들의 성능 평가치를 MOS test로 실시한 주관적 평가치와의 상관성을 비교 조사한 결과 BSD 척도와 PSQM 척도쪽이 각각 0.81, 0.84정도로 더 높게 나타났음을 확인하였다.

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

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, pp. 819-829, 1992.
- [6] Mohamed M. Meky, and N. Saadawi, " A Perceptual-Based Objective Measure for Speech Codecs Using Abductive Network", ICASSP, Vol.1, pp. 479-482, 1996.