

규칙 합성음의 객관적 품질평가에 관한 연구

홍진우, 유경완, 김순협

강원대학교 전자계산기공학과

A Study on Objective Quality Assessment of Synthesized Speech by Rule

J. W. Hong, K. W. You, S. H. Kim

Dept. of Computer Eng., Kwang Woon University

ABSTRACT

This paper evaluates the quality of synthesized speech by rule using the LPC CD in the objective measure and then compares the result with the subjective analysis. By evaluating the quality of synthesized speech by rule objectively. We have tried to resolve the problems (Evaluation time or size expansion, variables within the analysis results) that arise when the evaluation is done subjectively. Also by comparing intelligibility - the index for the subjective quality evaluation of synthesized speech by rule - with evaluation results obtained using MOS and the objective evaluation, we have proved the validity of the objective analysis and thus provides a guide that would be useful when R&D and marketing of synthesis by rule method is done.

1. 서론

최근 기계에 의한 음성의 합성에 대한 연구가 활발히 진행되고 있다. 또한, 음성 합성기술의 발전과 적용 범위가 넓어짐에 따라 다양한 방법에 의한 합성방식(녹음 편집 방식, 분석 합성 방식, 규칙 합성 방식)이 대두되고 있어 각 방식에 의한 합성 음성의 품질을 평가하고자 하는 연구의 필요성이 점차 확산되고 있다.

이러한 상황에서 규칙 합성음의 품질을 정량화할 품질 평가법을 확립할 필요가 있으며, 이에 기초한 평가 분석에 의해 양호한 품질의 합성음성을 서비스함을 들른 합성방식의 연구, 개발에 유요한 지침을 제공해줄 수 있다.

일반적으로 음성의 품질 평가는 주관적인 평가 방법과 객관적인 평가 방법으로 크게 분류할 수 있으며 각각 장, 단점을 가지고 있다. 주관적인 평가 방법은 품질을 평가하는 객체가 사람임을 중시하여 품질의 평가를 사람이 직접 결정하는 방법으로서 명도도, 만족도, 이해도등의 평가 척도가

사용된다. 그러나, 주관적인 평가 방법은 품질을 평가하기 위한 실험 규모가 커지고 실험에 소비되는 시간이 길며, 품질 평가 결과가 평가하는 사람의 심리적 환경에 매우 민감하게 작용하는 등의 문제점이 있다.

본 논문에서는 이러한 문제점을 감안하여 규칙 합성음의 품질을 객관적으로 평가하는 연구를 수행하였다. 본 논문에서 적용한 객관적인 평가 방법의 적용 알고리즘은 LPC CD (Linear Prediction Coefficient Cepstrum Distance)이며 평가 음성은 한국어의 속성 및 인도 분포를 고려한 단어 선택법에 의해 선정된 108개 단어를 반음절 규칙 합성 방식에 의해 만들어진 합성음이다.

제2장 음성 품질 평가법

11-1. 목적 및 필요성

인간은 주위의 환경으로 부터 다양한 형태의 자극을 받고 그것에 적절한 반응을 하면서 생활하고 있다. 자극에는 시각적 자극과 청각적 자극이 있으나 청각적 자극인 음성에 대해 그 품질을 평가하는 수행 방법과 평가 대상은 다양하게 적용될 수 있다. 반응을 평가하는 데는 그 대상과 목적에 가장 적절한 평가 체계를 찾는 것이 중요하다. 물체의 첫수를 잴때에 자를 사용하듯이 평가 체계에는 그 목적으로 부터 유도된 평가척도가 필요하다. 그러나, 하나의 평가 척도만으로 목적에 맞는 평가를 하는 경우는 거의 없고, 몇개의 평가 척도의 집합체가 필요하게 된다.

다수의 일반 상대자를 대상으로 하는 음성응답 시스템 (Audio Response System: ARS) 에서 음성 품질의 좋고 나쁨은 가장 중요하게 고려되어야 할 점의 하나이다. 많은 연구자들이 각종 음성품질 특성을 규명하고 품질기준을 설정하려는 것은 가장 경제적인 최상의 음성 품질 실현에 그 목적이 있다. 그러나 이러한 음성 품질의 좋은 정도 (goodness) 를 시스템 설계자가 독단적 판단으로 결정한다든지 할 경우에서 사용하고 있는 품질기준을 그대로 모방한다든지 할 경우 사회 문화적인 환경, 성별, 연령별 등 개인적 요인에 따라

많은 차이가 있기때문에 국내 실정에 맞는 음성품질의 평가를 기대 하기가 어렵다.

II-2. 품질 평가법의 종류

평가법은 크게 주관적인 방법과 객관적인 방법으로 나눌 수 있다. 주관적 방법에는 자연성, 이해성, 음량감과 만족도 등이 있으며 객관적 평가에는 LPC CD (Cep-strum Distance) 법과 FFT (Fourier Fast Transform) 분석에 의한 Coherence 함수법 등이 제시 되고 있다. 주관 평가법은 실용적 방법 (utilitarian method) 과 분석적 방법 (analytic method) 으로 나눌 수 있다.

그러나, 주관적인 평가 방법은 평가 실험을 위해 대규모의 모델 시스템이 필요하고 다수의 평가 인원이 요구되고, 평가 실험 시간이 많이 소요될 뿐만 아니라 평가 결과가 평가자의 심리적 요인이나 평가 환경에 많은 영향을 받는 문제점이 지적되고 있다.

본 논문에서는 이러한 문제점을 해결하기 위하여 객관적인 평가 방법을 도입하였다.

II-3. 객관적인 품질평가법

객관적인 품질측정은 기준 신호와 비교 신호와의 유사성 (likelihood) 신호 특성의 거리차, 또는 시스템을 경유한 입력과 출력 신호 사이의 왜곡 (distortion) 으로 정의되어 진다. 그림 1은 객관적인 평가 값을 결정하기 위한 측정방법의 블록도이다. 객관적인 품질 평가를 위해 시간 영역과 주파수 영역에서 다양한 객관적인 품질 측정 방법이 제안되고 있으며 이것들은 다음과 같다.

1. 파형 왜곡 척도 (Waveform Distortion Measures)

파형 왜곡 척도들은 입력과 출력 음성 파형 (speech waveform) 사이의 왜곡 (distortion) 에 의해서 정의되어 진다.

- (1) 신호대 잡음 비 (Signal to Noise Ratio : SNR) 측정법
- (2) 구분화 신호대 잡음비 (Segmental SNR: SNR_{seg}) 측정법

2. 스펙트럼의 왜곡 척도 (Spectral Distortion Measures)

주파수 영역에서 객관적인 품질측정은 입력과 출력 음성 스펙트럼 (speech spectra) 사이에 왜곡 (distortion) 으로 정의되어 진다.

- (1) 스펙트럼 왜곡 (Spectral Distortion ; SD) 측정법
- (2) 주파수 가중치된 스펙트럼 왜곡 (Frequency Weighted Spectral Distortion : WSD) 측정법

3. 스펙트럼 Envelope 왜곡 척도 (Spectral Envelope Distortion Measures)

스펙트럼의 envelope 왜곡에 기초한 측정 방법으로 FFT 기술과는 다른 선형 예측 부호화 (Linear Predictive

Coding: LPC) 에 의해서 계산되어 진다.

- (1) 유사도 비율(Likelihood Ratio : LR) 계산법
- (2) 가중치된 유사도 비율 (Weighted Likelihood Ratio : WLR) 계산법
- (3) LPC Cepstrum Distance Measure (LPC CD)
- (4) COSH Measure (COSH)

본 논문에서는 객관적인 품질 평가의 측정 방법중 음성 신호내의 가장 중요한 특징만을 추출하여 특징들간의 차를 거리값으로 이용하는 LPC Cepstrum 거리 (LPC Cepstrum Distance) 측정 방법을 사용하였다.

III. 객관적 품질 평가를 위한 LPC CD 계산

III-1. 선형 예측 계수 추출

음성의 특징을 추출하는 데에 다양한 방법이 있으나 LPC (Linear Prediction Coefficient) 를 이용해서 특징 파라미터를 추출하는 것이 매우 효과적이고, 합리적인 방식이라고 널리 알려져 있다. 음성 신호의 파형을 관찰하여 보면 음성 파형의 이웃한 샘플들은 상관 관계가 높음을 알 수 있다. 이와 같은 관계를 간단한 선형예측 형태로 표시하면 식 (1) 과 같다.

$$\tilde{x}_n = \alpha_1 x_{n-1} + \alpha_2 x_{n-2} + \dots + \alpha_p x_{n-p} \quad (1)$$

위 관계식은 음성파형의 한 샘플을 과거의 p개의 샘플들의 선형 결합으로 예측할 수 있다는 것을 가정하고 있는데 이때 각 샘플들에 곱하여 지는 가중치 ($\alpha_i, i = 1, 2, \dots, p$) 를 선형예측 계수라 한다. 선형예측 분석에서는, 예측오차 신호의 평균 지승치가 최소가 되도록 선형예측 계수들을 정한다.

입력 음성 신호 샘플 x_n 에 대한 예측치를 x_n 이라 하면

$$\tilde{x}_n = \alpha_1 x_{n-1} + \alpha_2 x_{n-2} + \dots + \alpha_p x_{n-p} \quad (2)$$

로 표시할 수 있다. 그러므로, 예측오차 e_n 은

$$e_n = x_n - \tilde{x}_n = x_n - \sum_{i=1}^p \alpha_i x_{n-i} \quad (3)$$

가 된다. 여기서 α_i 의 부호를 바꾸어 주면 식 (4)과 같이 나타낼 수 있다.

$$e_n = x_n + \sum_{i=1}^p \alpha_i x_{n-i} = \sum_{i=0}^p \alpha_i x_{n-i}, \quad \alpha_0=1 \quad (4)$$

예측오차 신호의 지승치는

$$e_n^2 = \left(\sum_{i=0}^p \alpha_i x_{n-i} \right)^2 = (x_n + \alpha_1 x_{n-1} + \dots + \alpha_p x_{n-p})^2 \quad (5)$$

$$E = \sum_n e_n^2 = \sum_n \left(x_n + \sum_{i=1}^p \alpha_i x_{n-i} \right)^2 \quad (6)$$

로 표시된다. 예측오차신호의 자승치를 최소화 하는 α_i 를 구하기 위해서 e_n^2 의 식을 α_i 에 대해서 편미분하면 다음과 같은 식을 얻게된다.

즉,

$$\frac{\partial E}{\partial \alpha} = 0, 1 \leq i \leq p \quad (7)$$

의 해를 구하면

$$R_{ij} \alpha = -r \quad (8)$$

여기에서 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_p]^T$, $r = [r_{01}, r_{02}, \dots, r_{0p}]^T$ 인 vector 양이고 R_{ij} 는 r_{ij} 를 갖는 $p \times p$ matrix로서

$$r_{ij} = \frac{1}{N} \sum_{n=0}^{N-1} x_{i+n} x_{j+n} \quad (9)$$

이다. 앞의 matrix식을 푸는 방법으로 자기상관 (autocorrelation) 방법을 사용한다.

자기상관 방법은 입력음성신호 $\{x_n\}$ 에 길이 N 인 window를 사용하여 그 window 밖의 샘플 값을 0으로 가정하고 자기상관 관계를 구하게 된다. 이 경우 자기상관 관계 matrix R_{ij} 는 positive definite한 Toeplitz 형태를 갖게 되어 Levinson-Durbin 알고리즘을 이용할 경우 (i)를 구하는데 계산량을 대폭적으로 줄일 수 있으면서도 안정된 계수들을 얻을 수 있다.

선형 예측은 음성의 스펙트럼의 기울기를 추적하는 한 방법으로 지금까지 연구된 여러가지 방법 중에서 정확하고 빠르게 계산 가능한 방법이다. 이것의 장점은

- 1) 근본적으로 시간 평면에서 분석하는 방법이다.
- 2) 상도 필터의 구조는 일련의 대역 필터 보다 융통성이 커서 스펙트럼 매칭도 정확히 할 수 있다.
- 3) 선형예측에 의한 power 스펙트럼은 FFT나 대역 여파기에 의한 것보다 훨씬 완만하다.

III-2. LPC 펠스트림 계수

LPC에 의한 스펙트럼 추정기 화자간의 고유한 성질을 완전히 제거하지 못하여 화자의 독립성을 고려하는 데는 장애 요소가 되어 왔다. 따라서 LP 계수에 대해 변형이 요구되는데 LP 계수를 각각 독립적으로 다룰 수가 없으므로 각 계수가 독립적이 되도록 변화시킨다.

본 논문에서는 이러한 특성을 갖는 LPC 펠스트림 계수를 거리값을 구하기 위한 특징 파라미터로 그림 2와 같이 추출하였다.

$$C_0 = \ln \sigma^2, C_i = -a_i - 1/i \sum_{k=1}^{i-1} k C_k a_i - k, 1 \leq i \leq p \quad (10)$$

III-3. LPC 펠스트림에 의한 거리 측정

선형예측 계수가 펠스트림 계수에 입각해서 거리 척도를 계산하기 위해서 사용되어 진다. 음성 파형으로 부터 직접 계산된 펠스트림과 다르게, 예측 계수로 부터 계산된 펠스트림은 평활화된 음성 스펙트럼의 평가를 제공한다. 이것은 식 (11) 과 같이 표현되어 진다.

$$\log \left(\frac{1}{A(z)} \right) = \sum_{k=1}^p c(k) z^{-k} \quad (11)$$

여기서, $A(z)$ 은 LPC 모델 다항식이고, $c(k)$ 는 펠스트림 계수이고, z 은 $e^{j\omega}$ 와 같게 설정되어 진다. 계수들은 다음에 따르는 수식을 사용해서 예측계수로 부터 재귀적으로 계산되어 진다.

$$nc(n) - na(n) = \sum_{k=1}^{n-1} (n-k) c(n-k) a(k) \text{ for } n=1, 2, \dots \quad (12)$$

여기서, $a(0) = 1$ 이고, $k > p$ 이면 $a(k) = 0$ 이다. 이 수식에서 $a(k)$ 는 LPC 예측계수이고, p 는 LPC 분석의 차수이다. 펠스트림 계수에 입각한 척도는 다음과 같이 계산되어 진다.

$$d(c\phi, c_d, 2, m)$$

$$= \{ [c\phi(0) - c_d(0)]^2 + 2 \sum_{k=1}^L [c\phi(k) - c_d(k)]^2 \}^{1/2} \quad (13)$$

여기서, d 는 프레임(frame) m 에 대한 L_2 거리이고, $c\phi(k)$ 는 자연 음성의 펠스트림 계수이고, $c_d(k)$ 는 규칙 합성 음성의 펠스트림 계수이다.

IV. 실험 및 결과 고찰

IV-1. 실험 환경

자연 음성과 규칙합성 음성의 LPC 펠스트림 거리를 측정하기 위한 대상어는 임의로 설정하지 않고 한국심리학회에서 조사한 "한국어 어휘 빈도 조사" 자료에 수록된 17,883 단어에 대해 속성 및 빈도 분포를 고려한 단어 선택법에 의해 선택된 108개 단어로 선정하였다. 선정된 대상어는 한국어 음성의 중요한 속성인 푼사, 단어의 길이, 중요도, 변칙 발음현상, 음운의 종류등이 고려된 단어들로써 실험 결과에서 이것들에 대한 속성별 결과도 기술하였다. 측정을 위해 사용된 음성용 자연 음성은 국내에 기준이 되는 음성이 규정되어 있지 않아 일반 성인 남성이 방음 환경에서 발생하여 녹음한 테이프를 사용하였다.

본 논문에서 LPC 펠스트림을 추출하고 펠스트림 거리를 구하기 위해 사용된 실험 시스템은 그림 3과 같다.

같은 단어에 대해 자연 음성과 규칙 합성음성의 LPC 펠스트림 계수가 추출되면 서로 다른 프레임의 크기를 정규화시키기 위해 프레임 압축/팽창에 의한 정규화 기법을 적용

한후 전체 프레임에 대해 LPC 쉼스트림 거리를 산출하는 방법에 의해 두 음성간의 거리값을 산출하여 출력시켰다.

한편, 규칙 합성음의 직관적인 평가와 주관적인 평가와의 상관관계 및 직관적 평가의 타당성을 입증하기 위해 주관적 평가를 수행하였다. 주관적 평가의 척도로는 이해도(intelligibility)와 MOS (Mean Opinion Score)에 의한 만족도를 선정하였다. 이해도는 언어로서 의미를 갖는 단어 또는 문장을 송화하여 이해율을 구한 것이고, MOS에 의한 만족도 시험은 규칙 합성음을 청취하고 난후 자연 음성과의 품질을 비교하여 1) 매우 좋다(4점) 2) 좋다(3점) 3) 보통이다(2점) 4) 나쁘다(1점) 5) 매우 나쁘다(0점)로 한정된 결과를 합산하여 평균한 값으로 결과를 유도하였다.

본 논문의 주관적 평가인 MOS 시험에 참가한 피험자는 성인 남성 12명이다.

IV-2. 결과 고찰

1. LPC CD와 이해도 및 MOS와의 관계

본 논문에서 수행한 종합적인 실험결과는 그림 4와 그림 5와 같다. 그림 4는 자연 음성과 규칙합성 음성간에 LPC 쉼스트림 거리 계산에 의해 산출된 거리 값(직관적 평가 결과)과 이해도 결과(주관적 평가 결과)와의 상관관계를 나타낸 것이고, 그림 5는 계산된 LPC 쉼스트림 거리 값과 본 논문에서 주관적 평가로 실시한 MOS 값과의 관계를 나타낸 것이다.

그림 4와 그림 5에서 표시된 점들은 108개 단어들중 같은 LPC CD 값을 갖는 단어들의 평균 LPC CD 값이며, 2차 곡선은 이 값들에 의해 구성되는 곡선 피크 모형(curvilinear regression model)을 표현하고 있으며 각각 다음과 같은 식으로 구성된다. 만일 가로축을 X, 세로축을 Y라 하면

$$Y = -4.359447 X^2 + 5.212056 X + 102.320076$$

$$Y = 0.141032 X^2 - 1.555205 X + 5.634372$$

그림 4와 그림 5로부터 LPC CD 값이 커지면 커질수록 즉, 규칙합성 음성의 품질이 자연 음성의 품질에서 멀어지면 질수록 이해도와 MOS 값이 하락되는 것을 관측할 수 있어 직관적 평가 결과와 주관적 평가 결과가 일치하고 있음을 알 수 있다. 이해도에 의한 주관적 평가(그림 4)에서 보통 80% 이상의 이해도가 확보되면 양호한 상태라고 할 수 있으며 이때 직관적 평가치인 LPC CD는 2.9dB가 된다. 반면에 MOS에 의한 주관적 평가(그림 5)에서 50% 이상의 피험자가 보통 이상이라고 응답하는 MOS 값이 2.5 이상이기 때문에 이 이상의 품질이 확보되어야만 실제의 응용에 사용될 수 있을 것으로 판단되며 이때의 직관적 평가치인 LPC CD는 2.6 ~ 2.7 dB가 된다. 따라서, 규칙합성 음성의 직관적 품질은 주관적 평가(이해도 및 MOS 평가)에 의해 양호한 품질이라고 판단되는 값의 범주에 대응할 수 있어야 하므로 자연 음성(기준 음성)과의 LPC CD 값이 약간의 차를 보이고 있는데 이것은 이해도가 음성 통신의 주 목적인 의미 전

달을 평가하는데 반해 MOS는 품질의 만족도를 평가하는 것에 따른 요인이라고 생각된다.

2. 단어의 속성별 LPC CD와 MOS와의 관계

규칙합성 음성의 직관적 평가를 위해 적용한 108개 단어에 대한 각 속성을 분석하여 각 단어 속성별 LPC CD 값과 MOS와의 관계를 살펴보고자 하며 그 결과는 그림 7과 같다. 여기서 세로축은 MOS 값을, 가로축은 LPC CD 값을 의미하고, 2차곡선은 각각의 피크 모형을 나타낸다.

그림 6의 a)는 품사유형별 MOS와 LPC CD와의 관계를 나타낸 것이다. 품사유종 명사의 LPC CD 값이 가장 다양하게 분포되어 있는데 이것은 108개 단어중 명사가 차지하는 비율이 매우 높기 때문이다. 이 그림에서 MOS 값의 기준을 2.5 이상으로 잡을때 명사가 가장 짧은 LPC CD 값을 나타냈고, 그 다음 부사, 동사의 순이었다.

부사의 경우 LPC CD 값의 작은 변화에 대해 MOS 값이 급변하는 현상을 나타내고 있어 규칙합성 음성을 생성할때 특히 주의를 기울여야 할 것으로 판단된다. 명용사의 경우는 데이터가 적어서 분포를 알 수 없어 생략하였다.

그림 6의 b)는 단어 길이별(2음절, 3음절, 4음절이상 단어별) MOS와 LPC CD와의 관계를 나타낸다. 이 그림으로부터 MOS값 2.5 이상에 대응하는 LPC CD 값은 단어 길이에 따른 영향이 거의 없음을 나타내고 있다.

그림 6의 c)는 단어의 중요도별(일상 생활에 많이 사용되는 순서에 따라 중요도 1, 중요도 2, 중요도 3으로 분류) MOS와 LPC CD와의 관계를 나타낸다. 중요도별로 약간의 LPC CD 값이 변화하고 있으나, 큰 영향을 미치지 않고 있음을 알 수 있다.

그림 6의 d)는 변칙 발음별(단어와 발음이 다른것과 다르지 않은 것을 구분) MOS와 LPC CD와의 관계를 나타내고 있다. 전반적으로 변칙발음 현상이 있는 단어가 변칙 발음이 없는 단어에 비해 LPC CD 값이 작아진 것을 알 수 있다. 이것은 변칙발음 현상(즉, 자음결핍, 구개음화, 경음화 등의 법칙)에 많은 고려사항을 부여하여 규칙합성 음성을 생성한 결과라고 생각된다.

그림 6의 e)는 단어의 제 1 음운별종((평음: 바, 다, 가), (격음: 포, 도, 쑤), (마찰음: 사, 하, 마), (복모음: 카, 키, 나, 니, 나, 내, 구, 네, 니)에 대한 MOS와 LPC CD와의 관계를 나타낸다.

그림으로부터 MOS값 2.5 이상에 대응하는 LPC CD와 값이 작은 순은 제 1 음절의 음운이 마찰음, 복모음, 평음, 격음의 순으로 나타나고 있음을 알 수 있다. 특히, 격음의 경우는 LPC CD 값이 작아져도 MOS값이 증가하지 않는 현상을 보이고 있는데 이러한 현상으로부터 격음에 해당되는 단어를 규칙합성 음성으로 생성할때 음의 억양이나 고조, 운율등에 많은 고려를 적용해야만 한다.

V. 결론

본 논문에서는 규칙합성 음성의 품질을 정성적으로 또는 정량적으로 평가하기 위한 방법중 LPC 펄스폭 거리를 적용한 객관적 품질 평가 방법 및 실험 결과에 대해 기술하였다. 또한, 객관적 품질 평가 결과와 주관적 품질 평가 결과와의 상관 관계를 살펴보기 위하여 주관적 품질평가인 이해도와 MOS실험을 수행하였으며 다음과 같은 결과를 얻었다.

1. 규칙 합성음성의 품질을 정확히 평가하기 위해서는 주관적 평가와 객관적 평가가 종합적으로 수행되어야 한다.
2. 본 논문에서 적용한 객관적 평가 시스템을 이용할 경우 양호한 품질의 규칙 합성음성에 대한 LPC CD 값은 2.6 ~ 2.9 dB 이하 이어야 한다.
3. 흡사류중 부사의 경우 LPC CD 값의 작은 변화에 MOS 값이 급변하는 현상이 있어 규칙 합성음성을 합성할때 주위가 필요하고, 격음의 경우 LPC CD 값이 작아져도 MOS 값이 증가하지 않는 현상이 있으므로 음의 역양이나 고조, 운율등에 세심한 고려를 필요로 한다.

앞으로 규칙합성 음성의 품질을 더욱 정확히 평가하기 위한 기준 음성의 규정 및 체계화된 측정 방법이 확립되어야 할 것이며 본 논문의 연구 결과가 국내의 규칙 합성음의 품질 향상을 위한 참고자료로 활용되기를 바란다.

참고문헌

- [1] 김영채, 한국어 어휘 인도조사, 한국심리학회지, Vol. 5, No. 3, pp216 - 285, 1986.
- [2] 한국전자통신연구소, 통화품질 평가법 및 표준화에 관한 연구, 연구보고서, 1989. 12.
- [3] N. Kitawaki, K. Itoh, and K. Kakehi, "Speech Quality Measurement Methods for Synthesized Speech," Review of ECL, Vol. 29, No. 9-0, Sep, 1981.
- [4] T. Watanabe, "Intelligibility Assessment of Synthesized Speech Using Word Intelligibility Score," NTT R&D, Vol. 38, No. 10, pp.1143 - pp.1152, 1989.
- [5] T. Watanabe, H. Nagabuchi, and N. Kitawaki, "A Study on Quality Evaluation Method of Speech Synthesized by Rule," NTT ECL, SP 86-11, 1988.
- [6] S. R. Quackenbush, T. P. Barnwell III, and M. A. Clements, Objective Measures of Speech Quality, Prentice Hall, 1988.
- [7] 김성한, 규칙 합성음의 이해성 평가를 위한 단어표 구성 및 실험법, 장문대학교 대학원 석사학위 논문, 1991. 8.
- [8] Thomas P. Barnwell III, "Objective measures for speech quality testing", J. Acoust. Soc. Am. 66(6), Dec. 1979.

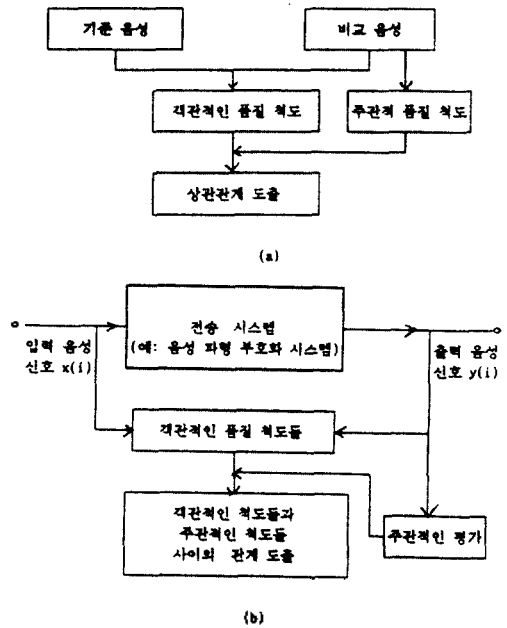


그림 1. 객관적인 품질측정 방법의 블록도
Fig. 1 Block Diagram of objective quality measures.

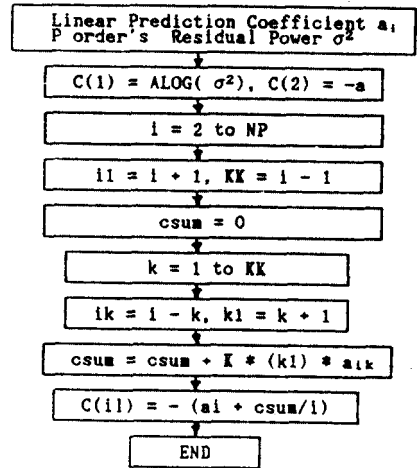


그림 2. LPC 펄스폭 계수 계산
Fig. 2 A calculation of LPC cepstrum coefficients

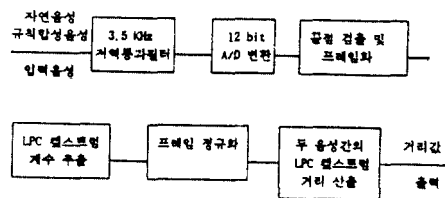


그림 3. LPC CD 계산 시스템의 구성도
Fig. 3 An organization diagram of LPC CD calculation system

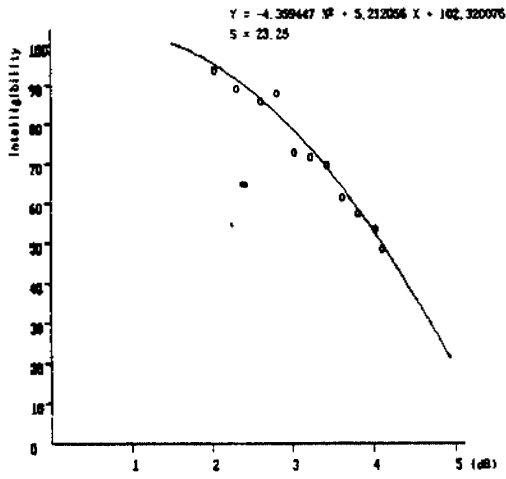


그림 4. 베타코프 LPC CD와의 관계
Fig. 4 Relation between intelligibility and LPC CD

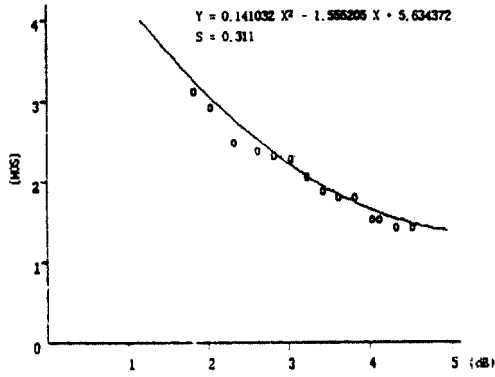
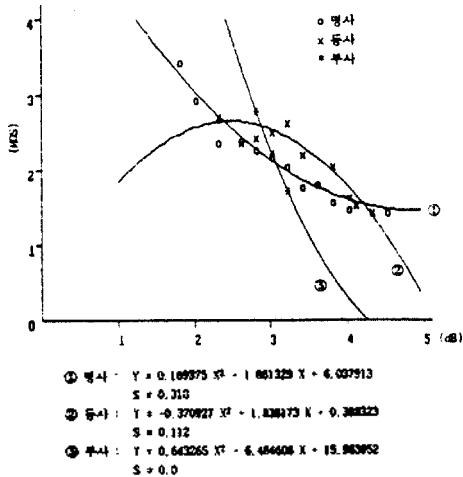
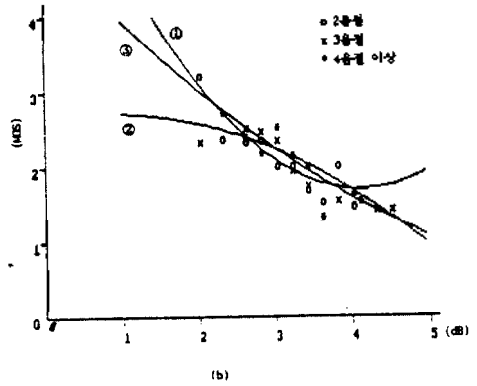


그림 5. MOS와 LPC CD와의 관계
Fig. 5 Relation between MOS and LPC CD

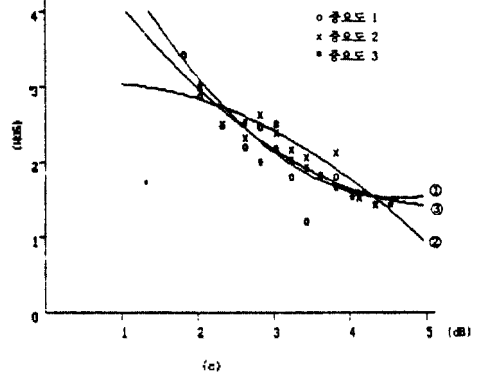


① 명사 : $Y = 0.169975 X^2 - 1.861329 X + 6.007913$
 $S = 0.310$
 ② 동사 : $Y = -0.370927 X^2 + 1.838473 X + 0.368323$
 $S = 0.112$
 ③ 부사 : $Y = 0.643265 X^2 - 6.484608 X + 15.963952$
 $S = 0.0$

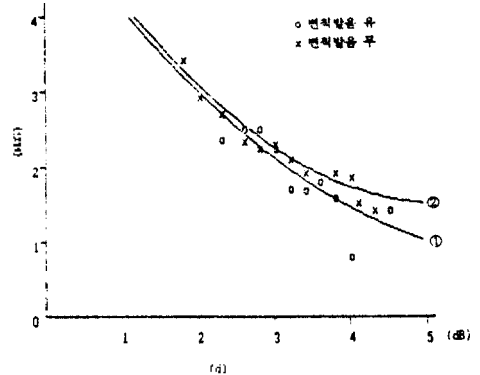
(a)



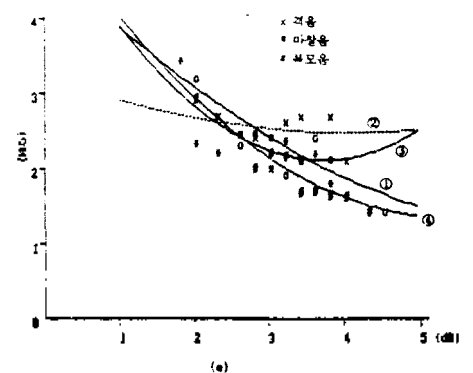
(b)



(c)



(d)



(e)

그림 6. 각 단어 속성별 MOS와 LPC CD 관계

Fig. 6 Relation between MOS and LPC CD for each word attributes