

# 웨이블렛 기반 바크 코히어런스 함수를 이용한 VoIP 음질평가

정회원 박 상 옥\*, 박 영 철\*\*, 윤 대 희\*

## Speech Quality Measure For VoIP Using Wavelet Based Bark Coherence Function

Sang-Wook Park\*, Young-Cheol Park\*\*, Dae-Hee Youn\* *Regular Members*

### 요 약

본 논문은 객관적 음질 평가법으로 웨이블렛 변환을 이용한 향상된 바크 코히어런스 함수 (Wavelet based Bark Coherence Function : WBCF)를 제안한다. 바크 코히어런스 함수 (Bark Coherence Function : BCF)는 심리 음향 영역에서 코히어런스 함수를 정의 함으로서 음성 통신 시스템의 아날로그 부분에 의하여 발생할 수 있는 선형 왜곡에 강한 객관적 음질 평가법이다. VoIP (Voice over Internet Protocol)와 같은 패킷 기반의 음성 전달 시스템은 가변 지연등이 발생 될 수 있는데, 이것은 원음과 왜곡음의 정확한 시간축 정렬을 불가능하게 하여 기존의 객관적 음질 평가법의 성능을 저하시킨다. 제안된 WBCF는 고주파 영역에서 시간 분해능이 높으며, 저주파 영역에서 주파수 분해능이 높은 웨이블렛 변환을 사용한 후 BCF를 계산하여 VoIP 시스템에서의 객관적 음질을 평가한다. 주/객관적 음질 평가 실험을 통하여 WBCF가 ITU-T 권고안인 Perceptual Speech Quality Measure (PSQM)에 비하여 높은 성능을 가짐을 확인하였다.

### ABSTRACT

The Bark Coherence Function (BCF) defines a coherence function within perceptual domain as a new cognition module, robust to linear distortions due to the analog interface of digital mobile system. Our previous experiments have shown the superiority of BCF over current measures. In this paper, a new BCF suitable for VoIP is developed. The Improved BCF is based on the wavelet series expansion that provides good frequency resolution while keeping good time locality. The proposed Wavelet based Bark Coherence Function (WBCF) is robust to variable delay often observed in packet-based telephony such as Voice over Internet Protocol (VoIP). We also show that the refinement of time synchronization after signal decomposition can improve the performance of the WBCF. The regression analysis was performed with VoIP speech data. The correlation coefficients and the standard error of estimates computed using the WBCF showed noticeable improvement over the Perceptual Speech Quality Measure (PSQM) that is recommended by ITU-T.

### I. 서 론

음질 평가는 주로 청취자들에 의해서 직접 평가 되는 주관적 음질 평가 방법과 수학적인 계산으로

언어지는 객관적 음질 평가법이 있는데, 이중 주관적 음질 평가 방법은 사람에 의해서 직접적으로 평가되므로 실제 사용자의 체감 음질을 가장 잘 표현할 수 있다. 그러나 주관적 음질 평가는 시간과 비용이 많이 소모되므로 객관적 음질 평가 척도를 사

\* 연세대학교 전기전자공학과 미디어 통신 신호처리 연구실(latest@mcsp.yonsei.ac.kr), \*\* 연세대학교 정보기술학부  
논문번호 : 010261-0926, 접수일자 : 2001년 9월 26일

용하여 주관적 음질을 예측하는 것이 보다 바람직하다<sup>[1]</sup>. 효과적인 객관적 음질 평가 척도는 다양한 왜곡환경과 음성 부호화 알고리즘에 대해서 주관적 음질과 상관 관계가 높아야 한다. 비록 객관적 음질 평가법이 주관적 음질평가를 완전하게 대체할 수 없으나 음성부호화기 개발, 음성 통신 장비 선택 및 최적화 그리고 네트워크 관리등에 사용될 수 있다.

근래에는 인간의 청각적 특성을 반영한 심리음학을 이용한 Bark Spectral Distortion (BSD)<sup>[3]</sup>, Perceptual Speech Quality Measure (PSQM)<sup>[4]</sup> 그리고 Measuring Normalizing Blocks (MNB)<sup>[5]</sup>등의 객관적 음질 평가법이 주로 연구되어 왔다<sup>[6-10]</sup>. 이러한 방법들은 임계 대역 분석, 등감 곡선 보정 그리고 주관적 세기 보정 등의 과정이 포함된다. 그러나 기존의 방법들은 음성 부호화기 음질 평가를 위하여 연구되어 왔기 때문에 실제 사용되는 음성 통신 시스템의 음질을 평가하기에는 부적절하다. 최신의 음성 통신 시스템은 디지털 기술을 사용하여 음성을 깨끗하고 선명하게 전달하나, 아날로그 접속단은 때때로 선형 왜곡을 발생시킨다. 선형 왜곡은 일반적으로 주관적 음질에 영향을 주지 않으며, 심지어 음질이 향상될 경우도 발생하나, 객관적 음질 평가 시스템의 성능을 저하시킨다<sup>[11][12]</sup>. 이러한 선형 왜곡 문제를 해결하기 위하여 [11][12]등의 방법이 제안되었다.

최근에 관심이 집중되고 있는 VoIP (Voice of Internet Protocol) 시스템은 음성을 패킷 단위로 전송하여 가변지연을 발생 시킨다<sup>[15]</sup>. 객관적 음질 평가에 앞서 원음과 왜곡음의 동기화는 필수적이며 일반적으로 상호 상관법을 주로 사용된다. 패킷 전송 시스템에서 발생한 가변 지연은 상호 상관법을 사용한 동기화 성능을 저하 시키고, 결과적으로 주관적 음질을 효과적으로 예측할 수 없게 된다<sup>[11]</sup>. 본 논문에서는 기존의 바크 코히어런스 함수 (Bark Coherence Function : BCF)<sup>[12][22]</sup>를 기본으로 하여 VoIP 시스템 음질 평가에 적합한 WBCF를 제안한다. WBCF는 음성 신호를 고주파 영역에서 시간 분해능이 높으며, 저주파 영역에서 주파수 분해능이 높은 웨이블렛 변환을 사용한 후 BCF를 계산하여 VoIP 시스템에서의 객관적 음질을 평가한다.

본 논문의 구성은 다음과 같다. 2장에서 가변지연 현상에 대하여 설명한 후 3장에서 새로운 객관적 음질 평가법인 WBCF에 대하여 설명한다. 4장에서 기존 객관적 음질 평가법과 WBCF의 성능을 비교 평가하며 5장에서 결론을 맺는다.

## II. 가변지연

객관적 음질평가에 앞서, 원음과 왜곡음은 시간축 상에서 정확하게 동기화가 이루어져야 한다. 시간축 동기화는 일반적으로 상호상관 함수를 사용하여 이루어지는데, 상호상관 함수는 시분별 지연 시스템에 매우 효과적이다.

VoIP<sup>[14]</sup>와 같은 통신 시스템은 음성을 부호화 한 후 그 비트 스트림을 패킷 단위로 전송한다. 전송되는 패킷은 일반적으로 UDP(User Datagram Protocol)로 전송되며 각 패킷은 네트워크 상태에 따라 각각 다른 도달 시간을 갖는다. VoIP 시스템은 지연 버퍼를 이용하여 실시간 복호화를 가능하게 하나, 버퍼의 사용은 전체 시스템의 지연 시간을 증가시킨다. 그러므로 실용 시스템에서는 지연 버퍼의 크기를 네트워크 상황에 따라 조절 하는데 이 때 복호화된 음성에 가변지연이 발생된다. 지연 버퍼의 크기 조절은 대개 목음 구간에서 이루어지며 주관적 음질에 영향을 거의 미치지 않으나, 원음과 왜곡음의 시간축 정확한 동기화를 불가능하게 하여 객관적 음질 평가 척도의 성능을 저하시킨다. 목음 구간에서 가변지연을 보상하기 위하여 일반적으로 음절(utterances)을 분리하여 각각 시간 축 동기화를 시키는 방법이 사용되기도 한다<sup>[11]</sup>.

그림 1은 가변지연의 예로서 첫 번째 문장과 두 번째 문장의 목음사이에서 가변지연으로 인하여 두 문장사이의 정확한 시간축 정렬이 불가능하다. 일반적인 상호 상관법을 이용하여 시간축 정렬을 시행할 때에 앞문장을 기준으로 시간축 정렬이 이루어지면 뒷문장에서는, 예를 들어 원음의 “길동”이라는 음성 스펙트럼과 왜곡음의 “홍길”이라는 음성 스펙트럼을 비교하게 되어 올바른 객관적 음질평가를 수행할 수 없다. 이러한 목음 사이의 가변지연은 주로 지연 버퍼의 크기를 변화할 때 발생하며 주관적 음질에는 거의 영향을 미치지 않으나 기존의 입력-출력 기반 객관적 음질 평가법의 성능을 저하시킨다<sup>[11]</sup>.

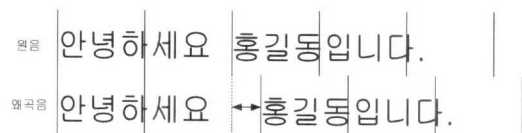


그림 1. 가변지연의 예. 목음 구간에 가변지연이 발생되어 원음과 왜곡음의 정확한 동기화가 불가능하다.

### III. 웨이블릿 변환 기반 바크 코히어런스 함수

본 논문에서는 선형왜곡과 가변지연에 강인한 웨이블릿 변환 기반 바크 코히어런스 함수 (Wavelet based Bark Coherence Function : WBCF) 를 제안한다. WBCF는 BCF를 기반으로 하여 VoIP의 객관적 음질 평가에 보다 적합하게 웨이블릿 변환을 사용한다.

BCF는 주파수 영역에서 두 신호의 상관관계를 나타내는 코히어런스 함수(magnitude coherence function : MSC)<sup>[18]</sup>을 인간의 심리 음향 반영한 바크(Bark) 영역에서 적용시킨 것이다. MSC는 선형 주파수 영역에서 다음과 같이 정의 된다.

$$\gamma_{xy}^2(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)} \quad (1)$$

여기서  $\gamma_{xy}^2(f)$ 는 입력 신호  $x(t)$ 와 출력 신호  $y(t)$ 의 상호 파워 스펙트럼이고,  $S_{xx}(f)$ 는  $x(t)$ 의 파워 스펙트럼, 그리고  $S_{yy}(f)$ 는  $y(t)$ 의 파워 스펙트럼이다. MSC는 원음과 왜곡음 사이의 주파수 영역에서 상관관계를 나타낸다. 다음으로 선형 시스템을 통과한  $x(t)$ 의 스펙트럼인 코히어런스 출력 파워 스펙트럼 (Coherent Power : CP)와 잡음 성분에 의한 출력 잡음 스펙트럼 (Non Coherent Power : NCP)를 식 (2)로 나타낼 수 있다.

$$\begin{aligned} CP(f) &= \gamma_{xy}^2(f) S_{yy}(f) \\ NCP(f) &= [1 - \gamma_{xy}^2(f)] S_{yy}(f) \end{aligned} \quad (2)$$

NCP와 CP의 비인 잡음 대 신호 비 (Distortion-to-Signal Ratio : DSR)는 다음과 같이 표현된다.

$$DSR(f) = \frac{NCP(f)}{CP(f)} = \frac{1 - \gamma_{xy}^2(f)}{\gamma_{xy}^2(f)} \quad (3)$$

잡음이 없는 이상적인 경우 MSC는 1의 값을 갖고 이때, DSR는 0이 된다. 또 비선형 잡음만이 존재할 경우 MSC는 0이 되어 DSR은 무한대의 값을 갖게 된다. 일반적인 경우 통신 시스템은 비선형 잡음이 존재하므로 DSR은 0과 무한대 사이의 값을 갖는다.

선형 주파수 영역에서 MSC의 정의와 더불어, 심리 음향 영역에서 정의된 MSC인 BCF는 다음과 같이 정의된다<sup>[12][22]</sup>.

$$BCF(b) = \frac{|L_{xy}(b)|^2}{L_{xx}(b)L_{yy}(b)} \quad (4)$$

여기서  $b$ 는 바크 주파수,  $L_{xx}(b)$ 와  $L_{yy}(b)$ 는 각각  $x(t)$ 와  $y(t)$ 의 자기 바크 스펙트럼이고,  $L_{xy}(b)$ 는  $x(t)$ 와  $y(t)$ 의 상호 바크 스펙트럼이다. 바크 스펙트럼을 구하기 위하여, 원음과 왜곡음을 푸리에 변환을 통하여 주파수 영역으로 천이한 후, 임계 대역 분석을 한다. 인간의 귀는 저주파 영역에서는 분해능이 세밀하고 고주파 영역에는 분해능이 떨어지며, 주파수 별로 자극에 대하여 다르게 민감하다. 예를 들어 100Hz의 톤이 1000Hz의 톤과 동일한 크기로 들리려면 35dB 더 큰 세기를 가져야 한다. 이러한 청각적 특징을 고려하여 각각의 주파수에 대하여서도 사람에게 의해 지각되는 소리의 크기가 동일하도록 보정할 필요가 있는데, 이를 표현한 것이 등감 곡선이다. 주관적 세기 보정은 인간이 소리의 세기 변화에 따라 비선형적으로 반응하는 특징을 반영한 것이 3가지 과정을 적용하여 계산한 결과가 바크 스펙트럼이다.

웨이블릿 변환의 기본 개념은 임의의 함수를 시간-주파수 공간에서 동시에 지역성을 갖는 웨이블릿 기저함수의 선형 결합으로 표현함으로 신호를 다 해상도로 해석하는 기법이다<sup>[17]</sup>. 고주파 대역은 시간 분해능을 높이게 되며, 저주파 대역은 주파수 분해능을 높이는 옥타브 대역 분할을 통해 시간과 주파수에 대해서 국부성(locality)를 가진다. WBCF는 이러한 특성을 이용하여 가변지연에 강인한 특성을 지닌다. 그림 2는 WBCF의 블록도이다.

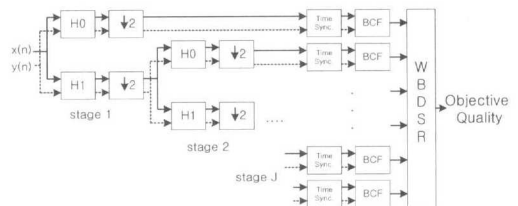


그림 2. WBCF의 블록도

WBCF는 웨이블릿 변환, 시간축 정렬과 BCF등의 3부분으로 나뉜다. 먼저 원음과 왜곡음은 이산 웨이블릿 급수 변환을 사용하여 옥타브 밴드 필터뱅크를 통과 한다. 각 필터를 통과한 신호들은 고주파 영역에서 시간 분해능이 높으며, 저주파 영역에서 주파수 분해능이 높은 특성을 가진다. 웨이블릿

변환된 원음과 왜곡음에 대하여 상호 상관법을 이용하여 시간축 정렬을 한다. 상호 상관법을 이용한 시간축 정렬은 비교적 정확한 결과를 나타내나 가변 지연이 발생시 신호의 에너지가 큰 부분을 기준으로 정렬이 이루어진다. 제안된 방법에서는 이산 웨이블릿 급수 변환 후 각 단계에서 시간축 정렬을 행하므로 각 옥타브 밴드 필터 출력 별로 보다 정확한 시간축 정렬을 행한다. 마지막으로, 시간축 정렬이 이루어진 신호들로 임계 대역 분석, 등감 곡선 보정, 주관적 세기 보정등의 심리 음향<sup>21,19)</sup>을 이용하여 BCF를 구한다. 웨이블릿을 이용한 심리 음향 영역에서 원음과 왜곡음의 코히어런스함수<sup>18)</sup>로 표현되는, WBCF는 다음과 같이 정의된다.

$$WBCF(b)^j = \frac{|L_{xy}(b)^j|^2}{L_{xx}(b)^j L_{yy}(b)^j} \quad (5)$$

여기서  $b$ 는 바크 주파수,  $L_{xx}(b)^j$ 와  $L_{yy}(b)^j$ 는 각각  $j$ 번째 단의  $x(t)$ 와  $y(t)$ 의 자기 바크 스펙트럼이고,  $L_{xy}(b)$ 는  $x(t)$ 와  $y(t)$ 의 상호 바크 스펙트럼이다. WBCF는 선형 왜곡에 영향을 받지 않으므로 보다 정확하게 통신 시스템의 음질을 평가할 수 있다.

WBCF를 계산한 후 최종적으로 왜곡 정도를 정량하기 위하여 다음과 같이 웨이블릿 기반 바크 신호대 잡음비(Wavelet based Bark Distortion-to-Speech Ratio : WBDSR)를 정의하여 심리 음향 영역에서 측정된 음성 통신 시스템의 음질을 의미한다.

$$WBDSR = \sum_{b=0}^N w_b \frac{1 - WBCF(b)^j}{WBCF(b)^j} \quad (6)$$

여기서  $w_b$ 는 각 바크 밴드별 가중치이며 주관적 음질과 상관관계가 높도록 최적화된다.

#### IV. 실험 및 성능 평가

본 논문에서 제안한 WBCF의 성능을 평가하기 위하여 여러가지 음성 데이터에 대하여 주관적 음질평가의 결과와 회귀분석<sup>20)</sup>을 하였다. 본 논문에서는 1과 5사이를 점근선으로 갖는 다음과 같은 회귀 함수를 사용하였으며, 주관적 음질평가로 MOS 평가<sup>13)</sup>를 실시 하였다.

$$\widehat{MOS} = 1 + \frac{4}{1 + \exp(ax + b)} \quad (7)$$

여기서  $a$ 와  $b$ 는 회귀 분석 계수,  $x$ 는 객관적 음질 평가 척도,  $\widehat{MOS}$ 는 예측된 주관적 음질이다. WBCF의 성능을 나타내는 지수로 상관 계수와 표준 추정 오차(SEE)를 사용하였다. 상관 계수는 객관적 음질 평가값과 MOS와의 선형성을 나타내며, 0에 가까울수록 낮은 상관 정도를, 1에 가까울수록 강한 상관 정도를 보여 준다. 표준 추정 오차는 실제 MOS와 예측된 주관적 음질인 예측된 MOS와의 차이를 통계적으로 나타낸다.

주관적 음질을 예측하기 위하여 회귀 분석을 한 후, 예측된 MOS와 실제 MOS와의 “Per-File Correlation” 상관 관계를 구하였다. “Per-File Correlation” 분석은 각 조건에 대한 평균 MOS값과 평균 객관적 척도값을 이용하는 “Per-Condition Correlation”과는 달리 하나의 파일에 대한 MOS값과 객관적 음질 척도 값들로 분석한다<sup>5)</sup>.

제안된 알고리즘의 성능을 평가하기 위하여 VoIP 시스템을 통과한 음성 데이터에 대하여 회귀분석을 실시하였다. 실제 인터넷망은 연구에 필요한 다양한 네트워크 상황을 재현하지 못하므로 실시간 네트워크 에뮬레이터인 NIST Net(National Institute Standards and Technology Network Emulation Tool)<sup>12)</sup>을 사용하여 패킷 손실률과 가변지연을 제어하였다. 설정된 평균 지연 시간은 150ms, 가변지연의 표준 편차는 0-60ms, 패킷 손실률은 0-25%로 설정하여 다양한 네트워크 환경을 반영하였다. 사용된 VoIP 시스템은 H.323기반의 상용화 시스템으로 PC-PC 그리고 PC-PSTN을 모두 지원한다. 녹음에 사용된 음성은 NTT CD-ROM에서 선택하였으며 모두 남성 화자 2명, 여성 화자 2명의 음성이 사용되었다. 전체적인 녹음과정은 그림 3과 같다. H.323<sup>14)</sup> 시스템 A에서 부호화된 음성은 패킷 형태로 변환된 후 NIST Net에 의하여 지정된 평균 지연시간, 가변지연 그리고 패킷 손실률을 가지고 LAN환경의 시스템 B로 전달된다. 녹음에는 G.723.1 6.3kbps 음성 부호화기가 사용되었다

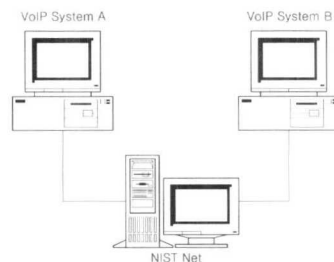


그림 3. VoIP 음성 데이터 녹음 과정

표 1. 가변지연 발생 음성 데이터에 대한 상관 계수와 표준추정 오차.

Test	Correlation Coefficient/SEE				
	PSQM I	PSQM II	BCF I	BCF II	WBCF
I (VoIP)	0.832/0.600	0.849/0.571	0.812/0.631	0.822/0.615	0.891/0.492
II (DCS)	0.657/0.701	0.692/0.671	0.704/0.660	0.840/0.505	0.887/0.429

표 1에 VoIP 음성 데이터에 대한 PSQM과 WBCF의 상관계수와 표준 추정 오차를 정리하였다. PSQM은 가변 지연에 대한 고려를 하지 않으므로 2가지 형태의 실험을 실시하였다. 표에서 PSQM I은 원음과 왜곡음의 동기화를 위하여 상호상관법을 이용하여 동기화를 한 후 결과 정렬이며 PSQM II는 PSQM I과 마찬가지로 동기화를 행한 후 앞 뒤 128ms 내에서 8ms 마다 PSQM값을 계산 한 후 최소값을 택하였다<sup>[16]</sup>. BCF I과 BCF II는 PSQM의 경우와 같은 동기화 기법을 사용하여 실험을 반복하였다. PSQM II는 PSQM I에 비하여 좀 더 정확한 시간축 정렬이 이루어졌으므로 주관적 음질과 높은 상관관계를 나타낸다. 그러나 이러한 방법은 반복적인 많은 계산량을 요구하므로 실제 사용에 적합하지 않다. 기존의 연구 내용인 BCF의 경우도 PSQM과 같이 가변지연에 낮은 성능을 나타낸다. 제안된 방법인 WBCF는 웨이블렛 기반의 바크 코히어런스 함수와 각 옥타브 밴드별 시간축 정렬을 함으로써 주관적 음질을 효과적으로 예측한다.

테스트 II는 CDMA 디지털 셀룰러 시스템(Digital Cellular System : DCS)으로 이 경우에도 테스트 I의 음성 데이터와 마찬가지로 강한 가변 지연이 관찰되었다. PSQM과 BCF는 주관적 음질과 낮은 상관도를 나타내며, WBCF는 VoIP경우와 마찬가지로 상관계수 값이 매우 높다.

### V. 결론

음성 통신 시스템에서 선형 왜곡과 가변지연등이 발생할 수 있는데 선형 왜곡은 통신 시스템의 아날로그 부분에 의하여 쉽게 발생되며 주관적 음질에 영향을 미치지 않은 객관적 왜곡을 발생 시킨다.

가변 지연은 주로 VoIP와 같은 패킷 기반의 음성 전달 시스템에서 발생되며 CDMA 시스템에서도 관찰되었다. 이는 선형 왜곡과 마찬가지로 기존의 객관적 음질 평가법의 성능 저하의 큰 요인이 된다.

본 연구는 웨이블렛 기반의 바크 코히어런스 함

수를 정의하고 이를 이용하여 음성 통신 시스템의 객관적 음질 평가를 수행하였다. 실험을 통하여 WBCF는 기존 ITU-T 권고안인 PSQM에 비하여 주관적 음질과 높은 상관 관계를 가짐이 입증하였다.

### 참고 문헌

- [1] S. Quackenbush, T. Barnwell and M. Clements, *Objective measures of speech quality*, Prentice Hall, 1988.
- [2] Nynek Hermansky, "Perceptual Linear Predictive (PLP) analysis of speech", *J. Acoust. Soc. Am.*, vol. 87, pp1738-1752, April 1990.
- [3] Shihua Wang, et al, "An Objective measures for predicting subjective quality of speech", *IEEE J. Select. Areas Commun.*, vol 10, No5, pp819-829, June 1992.
- [4] J. G. Beerends and J. A. Stemrindk, "A perceptual speech-quality measured based on a psychoacoustic sound representation", *J. Audio Eng. Soc.*, vol 42, No 3., pp115-123, March, 1994.
- [5] S. Voran, "Objective estimation of perceived speech quality, Part I : Development of the measuring normalizing block technique", *IEEE Trans. on Speech and Audio Processing*, vol. 7, No. 4, pp371-382, July 1999.
- [6] M. Hansen and B. Kollmeier, "Using a quantitative psychoacoustical signal representation for objective speech quality measurement", in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process*, pp 1387-1390, 1997.
- [7] R. F. Kubichek, "Mel-cepstral distance measure for objective speech quality assessment", in *Proc. IEEE Pacific Rim Conf. Communications, Computer, and Signal Processing*, pp 125-128.

- 1993.
- [8] W. Yang, M Dixon, and R. Yantorno, "A modified bark spectral distortion measure which uses noise masking threshold", in *Proc. IEEE Speech Coding Workshop*, pp. 55-56, 1997.
- [9] Markus Hauenstein, "Application of meddis' inner hair-cell model to the prediction of subjective speech-quality", in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process*, pp 545-548, 1998.
- [10] D. S. Kim, O. Ghitza and P. Kroon, "A computational model for mos prediction", in *Proc. IEEE Speech Coding Workshop*, pp141-143. 1999.
- [11] A. Rix and M. Hollier, "The perceptual analysis measurement system for robust end-to-end speech quality assessment," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process*, pp.1515-1518. 2000.
- [12] Sang-Wook Park, Seung-Kyun Ryu, Young-Cheol Park, and Dae-Hee Youn, "A Bark Coherence Function For Perceived Speech Quality Estimation," *Proc. of Intl. Conf. Spoken Language Processing 2000*, Vol 2, pp218-221. 2000.
- [13] ITU-T Rec. P.800, "Method for subjective determination of transmission quality", 1996.
- [14] ITU-T Rec. H.323, "Packet based multimedia system", 1998.
- [15] R Babbage, I Moffat, A O'Neill and S Sivaraj, "Internet Phone-change the telephony paradigm?" *BT Technol J*, Vol. 15 No.2, pp145-157, 1997.
- [16] ITU-T Rec. P.861, "Objective quality measurement of telephone-band speech codecs", 1998.
- [17] Martin Vetterli, Jelena Kovacevic, *Wavelets and subband coding*, Prentice Hall, 1995.
- [18] Julius S. Bendat and Allan G. Piersol, *Engineering Applications of Correlations and Spectral Analysis*, John Wiley & Sons, 1980.
- [19] E. Zwicker and H. Fastl, *Psychoacoustics Facts and Models*, Springer-Verlag, 1990.
- [20] N.R. Draper, H. Smith, *Applied Regression Analysis*, John Wiley & Sons, New York, 1981.
- [21] <http://www.antd.nist.gov/itg/nistnet/>
- [22] 박상욱, 류승균, 박영철, 윤대회, "바크 코히어런스 함수를 이용한 이동 전화 음질 평가," *한국통신학회 논문지*, 제 26권 제4B호, pp437-446, 2001.

박 상 욱(Sang-Wook Park) 정회원  
 한국통신학회논문지 제25권 제12B호, 2000년 12월  
 참조

박 영 철(Young-Cheol Park) 정회원  
 한국통신학회논문지 제25권 제6B호, 2000년 6월 참  
 조  
 2002년 3월 ~ 현재: 연세대학교 정보기술학부 조교수

윤 대 희(Dae-Hee Youn) 정회원  
 한국통신학회논문지 제25권 제6B호, 2000년 6월 참  
 조