

---

# VoIP 음질의 주관적 평가에 관한 연구

강영도\* · 강진석\* · 최연성\*\* · 김장형\*

## A Study of Subjective Speech Quality Measurement in VoIP

Young-do Kang\* · Jin-suk Kang\* · Yeon-sung Choi\*\* · Jang-hyung Kim\*

### 요 약

본 논문은 초고속 인터넷을 이용한 음성전송 방식인 VoIP(Voice over IP) 서비스에서의 음질의 주관적 평가에 관한 연구이다. PSNR이나 지터 같은 멀티미디어 서비스의 객관적 파라미터는 명료히 정의되어 있고 쉽게 측정되나, 그것이 이용자의 지각적 인지도와는 일치하지 않는다. 본 연구에서는 송화자 내용 발생과정에 있어서 어느 정도 완전히 표현되었는가를 나타내는 송화 품질과 음성이 전송계를 통해 수화자에게 전달되는 과정에서 왜곡이나 잡음 등의 방해요인에 의해 열화 되는 정도를 나타내는 전송 품질, 그리고 수화자가 청각에서 신호처리 과정을 거친 송화자의 내용을 어느 정도 이해할 수 있는지를 나타내는 수화 품질에 대한 주관적 평가법과 그 척도를 제시한다. 또한 제안된 방법으로 실험하여 그 유효성을 검증한다.

### ABSTRACT

In this paper, we discuss the scale of subjective speech quality measurement over VoIP(Voice over IP) network which is a component of broadband networks. Objective parameters of multimedia services like PSNR or jitter can easily measured and defined, but these factors are not easily meet the user's perceptual recognition. We suggest the speech quality measurement scale through the subjective measurement for end-to-end speech quality composed of sender-side quality, transmission quality, receiver-side quality, which provide the degree of correctness of representation of speaker, the degree of impairment caused by various factors, the degree of recognition of processed speech, respectively. Also, we examined the proposed method and verify it's availability.

### 키워드

VoIP, QoS, Subjective Speech Quality Evaluation, Delay, Jitter

---

\* 제주대학교 공과대학 정보공학과

\*\* 군산대학교 공과대학 전자정보공학부

접수일자 : 2001년 6월 9일

## 1. 서론

최근 인터넷 사용의 폭발적인 증가로 인하여 인터넷을 통한 다양한 서비스들이 제공되고 있으며, 혹은 준비중에 있다. 이러한 오늘날의 인터넷 서비스가 상업화됨에 따라 서비스 품질(Quality of Services)에 대한 관심이 고조되고 있으며, 이에 따라 QoS를 제공하기 위한 메커니즘의 필요성 여부도 활발하게 이루어지고 있다. 일각에서는 광섬유와 WDM (Wave length Division Multiplexing) 기술에 의해 QoS가 자동으로 제공될 수 있을 정도의 충분한 대역폭이 보장되기 때문에 QoS 제공을 위한 새로운 기술이 필요하지 않다고 주장하고 있다. 또 다른 일각에서는 비록 네트워크가 아무리 많은 대역폭을 제공하더라도 이를 고갈시킬 새로운 응용들이 개발되어질 것이므로 QoS를 제공하기 위한 메커니즘이 더욱 필요하다고 주장하고 있다.

본 연구에서는 VoIP(Voice over IP) 망의 서비스 품질 평가에 대한 지침을 제시하고자 한다. VoIP 서비스는 기존 공중 전화망에 비해 매우 저렴한 가격으로 이용할 수 있는 새로운 서비스이며 응용 프로그램에 따라 매우 다양한 서비스를 제공할 수 있는 고부가가치의 통신매체로 기존의 PSTN(Public Switched Telephone Network) 상에서 음성 서비스를 이용하여 왔던 사용자 입장에서 볼 때 인터넷상에서의 음성서비스는 서비스 품질 측면에서는 아직 부족한 부분이 많은 실정이다.

본 논문에서는 VoIP망에서의 통화 품질을 평가하는 방법으로 이용자가 품질 요인 영향을 주관적으로 측정하는(Subjective Measurement) 주관적 평가법을 품질에 영향을 주는 물리량을 측정하여 평가하는 객관적 평가법과 비교하여 고찰한다.

## II. 음질 평가 방법 및 모델

음질(Speech quality)은 인간의 지각 과정 속에서 심리적-청각과 관련된 복잡한 요소와 관련되어 주관적인 성격을 갖고 있다. 사람마다 음질을 느끼는 방법이 틀리고 심지어 동일인이라 할지라도

분위기, 관심, 기대감에 따라서 음질을 인지하는 수준이 다르게 된다. 그러나 음질을 정량화하고 측정할 때에는 이와 같은 개인의 의견은 배제되어야 하므로 음질은 보통 MOS(Mean Opinion Score)로 표현한다. 즉, 음질에 대해서 느끼는 여러 사람의 의견을 평균한 것으로 여러 대표 수화자 들로부터 이 값을 얻게 된다. MOS값은 보다 더 정확한 음질을 측정하는데 많이 사용될 수 있으며 이는 음질의 주관적 품질평가의 중요 요소가 된다. <그림 1>은 음성 통신에서의 음질에 대한 여러 면을 시각적으로 보여주는 것으로 단-대-단 음질 혹은 Mouth-To-Ear-Quality는 원격의 단-대-단 터미널에서 수화자가 지각하는 음질의 정도를 나타내는 것으로 음질은 이용자에 의해서 인지되는 전반적인 통신 품질의 요소 중 하나의 요소에 불과한 것이다[2].

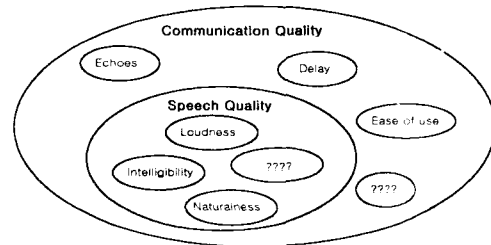


그림 1. 음질 품질의 관점  
Fig. 1 Aspects of speech quality

일반적으로, 음질 측정을 자동적으로 측정하기 위한 알고리즘이 많이 연구되어왔는데 객관적인 방법의 원리는 참조 음성 표본(Reference Speech Sample)과 손상된 표본(Impaired Sample)을 비교하는 것이며, 가능한 실제 청취 시도에서 얻을 수 있는 MOS수치에 대응하는 값을 얻는 것을 목표로 하고 있다. 일반적으로 음질을 측정하는 이유는 다음과 같다.

- ① User Perception : 이용자는 자신들이 느끼는 음질에 의해서 서비스 품질을 전적으로 판단한다. 즉, 음질은 고객 만족도를 나타내는 지표가 된다.
- ② End-to-End Measurement for any Impairment : 음질은 단-대-단에서 측정되므로 전체 전

송 링크의 비용을 압축해서 나타낼 수 있으며, 더욱이 이러한 음질 평가는 손상의 종류와 손상을 유발시키는 망 설비와 무관하게 총체적인 결과를 얻을 수 있는 “블랙 박스”로 볼 수 있다.

③ **Reliable Automated measurement** : 음질 측정 알고리즘의 기술 수준이 실제 사람이 듣는 수준과 근접한 결과를 산출해 낼 수 있기 때문에 자동화된 음질 평가 작업을 수행할 수 있으며, 비용절감은 물론 관리자와 서비스 제공자를 위한 새로운 어플리케이션을 가능하게 한다.

### 1. VoIP QoS

전통적으로 통신회사들은 음성과 데이터 망을 분리한 채 관리해 왔다. 이러한 주 이유는 기존의 망 기술이 음성과 데이터에 따라 다르게 요구하는 성능을 충족시킬 수 없었기 때문이다. VoIP 기술은 IP 주소를 기반으로 한 종단간의 채널 설정을 통하여 음성 신호를 압축하여 IP 패킷화한 음성데이터를 전달하는 기술이다. 압축 기술은 기존 전화망에서의 채널 사전 점유에 의한 64Kbps 급 대역폭 이용상의 문제점을 극복하고 매우 적은 대역폭만을 이용하여 음성 서비스의 제공을 가능하게 하였다. 또한, 인터넷의 QoS 기술의 발전에 따라 패킷화된 음성 데이터를 실시간으로 패킷의 손실 없이 종단간에 전달할 수 있는 기술이 발전하게 되고, 인터넷이 기존 망의 백본망으로서 위치할 가능성이 높아지면서 집중적으로 음성 서비스망인 PSTN/N-ISDN을 종단하고 인터넷을 통하여 전달하기 위한 인터넷 폰 기술에 대한 관심이 증대되며 신속한 호처리, 망의 효율적인 운용 및 지능망 서비스를 위한 No.7망을 돕으로써 호와 연결을 분리하여 처리할 수 있도록 구성되어 있다. VoIP와 같은 패킷 기반의 멀티미디어 통신을 위한 글로벌 표준인 H.323은 공중 교환 전화망과 비교할 수 있는 전화 기능성(Telephone Functionality)을 지원한다. 그러나 성공적인 VoIP 통신이 되기 위해 꼭 필요한 다른 요건은 QoS이다. 음성 통신은 저 지연, 저 지터, 최소의 패킷 손실을 갖는 망을 필요로 한다. 전화 이용자들은 이미 PSTN과 사설 PBX 기반의 망에서 지원되는 QoS에 익숙해져 있기 때문에 매우 높은 기대감

을 가지고 있다. 연결 중심형인 회선 교환망은 각각의 사용자에게 허용된 모든 대역폭을 각 호 동안 제공한다. 이러한 결과는 매우 낮은 지연과 지터 그리고 연결상의 잡음에 의해 발생하는 두절을 최소화로 만든다. 짧은 대기(Low Latency)는 이용자가 자연스러운 대화를 나눌 수 있게 해준다. 이용자들은 각자가 지연에 대한 인내하는 정도가 다르다. 그러나 보편적인 좋은 방법 중 하나는 한 방향의 지연을 약 150ms 이내로 제한하는 것이다. 이 지연은 종단 시스템에 의해 생성되는 프로세싱 지연과 망의 지연을 합친 값이다.

### 2. 객관적 음질 평가

객관적 평가는 품질에 대해 객관적이고 과학적인 방법을 통해 표준 검사한 것을 대상으로 사용자가 이 요소에 대한 비교를 하는 것이다. 표준검사는 품질과 관련된 많은 제품의 특성을 측정할 수 있다는 장점을 가지고 있으나 신뢰성 있고 타당성 있는 방법의 개발은 많은 요인들을 고려해야 하는 어려움이 있다.

객관적 음질 평가 척도에는 SNR, SegSNR, LPC-CD, BSD 등 여러 가지 방법이 있다. SNR 척도는 원 음성과 왜곡된 음성 파형간의 자승평균오차를 계산한다. SNR은 단지, 원 음성과 왜곡된 음성 파형의 신호 파형만의 비교이므로, 아날로그 이동 전화시스템이나 낮은 전송률을 갖는 음성부호화기의 주관적 음질을 예측하는 데에는 적합하지 않다. 반면에 SegSNR은 SNR 보다 주관적 음질과 상관 관계가 높다. 또, 인간의 청각 특성을 이용한 심리적 가중치(Perceptually Weighted) SNR인 Frequency-Variant SNR (FVSNR) 등도 있다.

LPC-CD는 원 음성과 왜곡된 음성의 선형예측성도 모델에 기본을 두고 있다. 파라미터는 LPC 자체가 직접 사용될 수도 있고, PARCOR, Log Area Ratio(LAR), Cepstral 계수 또는 LSP(Line Spectrum Pairs) 등과 같이 변환될 수도 있다. BSD는 인간의 청각 임계 대역(Auditory Critical Bandwidth)에 기본을 두고 인간의 귀를 모델링하는 필터 बैं크를 사용한다. BSD는 Bark 영역에서 스펙트럼의 차이를 계산한 것이다.

### 3. 주관적 음질 평가

주관적 평가는 사용자의 주관성에 초점을 둔 방법이다. 즉 이상적 선택에 대한 자신의 인식과 관련해서 멀티미디어 콘텐츠의 특성에 대한 개인의 인식을 기초로 많은 수의 실제 청취자가 실험에 참여한다. 이러한 방법은 콘텐츠 사용에 있어서 개인의 기호도 및 특정 상황에 따른 개인차이와 많은 노력과 비용의 필요성 및 개인의 주관적 이상에 대한 측정 도구를 만들어야 하는 어려움이 있지만 음질 측정에 대한 실질적이고 신뢰성 있는 평가를 얻을 수 있다는 장점을 갖는다. 채널을 통과한 음성의 주관적 평가를 하는 방법에는 여러 가지가 있다. 주로 사용되는 방법은 Degradation Category Rating(DCR), Comparison Category Rating(CCR), Absolute Category Rating(ACR) 등이 있는데 본 논문에서는 음질의 전체적인 평가에 주로 사용되는 ACR 방법으로 주관적 음질 평가를 수행하였다. ACR에서는 음질평가의 레벨로 주로 MOS를 사용한다. MOS는 음질 손상의 척도로 5 등급의 손상척도를 제시한다. 음질이 매우 좋을 때는 5점에 해당되고, 음질이 매우 나쁠 때는 1점이 된다. <그림 2>는 MOS 5등급을 나타낸 것이다.

Quality of the speech	Score
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

그림 2. MOS 5 등급  
Fig. 2 Rating scales for MOS

### 4. 음질품질 평가 모델

본 연구에서는 VoIP 망에서의 음성 품질의 주관적 평가 분석을 위한 평가 모델을 정의한 다음 주관적 평가를 수행해 본다.

첫째, VoIP 음성은 궁극적으로 이용자에게 도달되어 인간의 인지적인 판단에 의존해서 결정되는 특성을 갖고 있으므로 객관적인 실측만으로는 품질 평가가 부족하며, 주관적 방법을 통한 이용

자의 전화서비스 품질에 대한 불만족도를 정량적으로 파악함으로써 서비스 품질에 대한 종합적인 평가와 방향 설정을 할 수 있다.

둘째, 긍정적인 영향은 동기화 정확도 명세(SAS : Synchronization Accuracy Specification)에 긍정적인 영향을 미친다.

셋째, 프리젠테이션의 품질은 동기화 정확도 명세(SAS)에 의해 주관적 품질 모델을 제시할 것이다. 이를 모델로 도식화하면 <그림 3>과 같다.

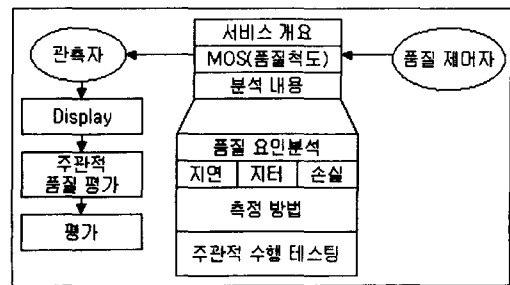


그림 3. VoIP의 주관적 품질 평가 모델 제안  
Fig. 3 Proposed model for subjective quality evaluation of VoIP speech

<그림 3>의 모델을 토대로 음성 품질의 여러 중속 변수를 대상으로 실질적인 품질 측정의 평가 모델을 실증적으로 분석해 보고자 한다.

## III. 음질 평가 실험 및 결과

### 1. 음질(MOS) 측정

측정시 음성 코딩 방식은 G.729(8K, CS-ACELP)을 공통으로 사용했다. 5초 길이의 샘플링된 남녀음성원음을 측정하였다. 먼저 측정장소에서 기준값 추출을 위해 VoIP G/W를 거치지 않은 상태에서 L to M 음성품질을 측정하는데 이동망에 의한 왜곡을 최소화하기 위해 정지 상태에서 시행했다.

측정결과 이동망의 상태는 상당히 양호한 편으로 나타났다. 일반 데이터 무부하 상태(음성호만 전송)에서 VoIP G/W를 통과한 음성품질은 기준값에 비해 PMOS는 남성의 경우 0.35, 여성의 경우 0.29정도 낮은 수치로 측정되어 VoIP G/W에

의한 품질저하는 거의 없는 것으로 나타났다. 800 바이트의 Ping datagram을 계속 전송하고 있는 부하상태에서 음성에 대한 별도의 QoS를 적용하지 않은 경우 PMOS가 매우 떨어짐을 알 수 있다. 실제 음성품질을 들을 경우 음성품질이 상당히 떨어짐을 알 수 있으나, PMOS 값이 매우 떨어진 주 요인은 Q-매니저의 측정요소가 끊김 현상을 매우 민감하게 평가하는 것도 주요한 원인이 되었다. 그리고 음성 데이터를 여러 번 코딩하고 디코딩하는 것으로 인한 신호왜곡(transcoding)으로 인한 품질저하도 포함되어 있다. 통상적으로 G.729인 경우 신호왜곡을 세 번 이상하면 정상 이하의 품질로 저하된다. 그런데 IP Precedence 6에 의한 Weighted Fair Queuing을 QoS로 적용해서 라우팅 우선 순위를 높게 부여했다. 부하상태에서도 무부하상태의 PMOS에 상당히 근접한 품질을 얻을 수 있었다.

이 모델은 일상적인 전화 음성대역(300Hz-3400 Hz)내에서 음성통화품질에 대한 척도를 제공한다. 이 모델에서는 <그림 4>에서와 같이 송화자(S) 측의 입에서 수화자 측(R)의 귀까지 인지된 음성통화품질을 추정한다. 이 모델에서 근본적인 원칙은 "심리적인 척도(Psychological scale)를 가지는 심리적인 요인들(factors)은 가감(additive)이 가능하다"는 개념에 기초하고 있다.

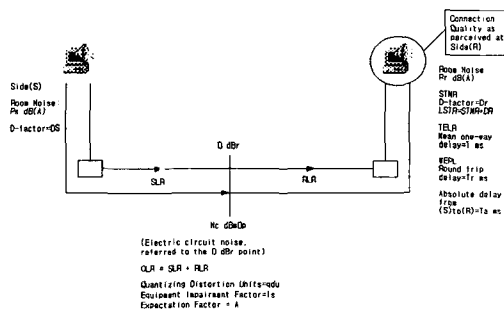


그림 4. E-모델 다이어그램  
Fig. 4 E-Model Reference Diagram

이 모델링은 다양한 전송품질요소의 효과를 결합하여 하나의 평가율 R로 제공함으로써 접속호에 대한 양호(Good or Better) 확률이나 불만족(Poor or Worse) 확률, 평균 의견 점수로의 환산 점수 등과 같은 사용자 반응의 예측을 가능하게

한다. 평가율 R은 아래와 같은 항으로 구성되었다.

$$R = R_0 - I_s - I_d - I_e + A \quad (\text{식-1})$$

여기서  $R_0$ 는 주로 여러 가지 잡음과 손실 등에 의한 항을 나타내고,  $I_s$ 는 너무 소리가 큰 연결(tool loud connection), 큰 측음, PCM에 의한 양자화 잡음과 같이 음성신호에서 동시에 발생하는 열화(Impairment)에 의한 항을 나타낸다.  $I_d$ 는 송화자 수화자 반향이나 너무 긴 지연 등과 같이 지연으로 인한 열화에 의한 항을 나타내고,  $I_e$ 는 저속 모뎀이나 디지털 회선 다중화 장치(DCME)와 같은 특정 장치로 인한 전송 열화에 의한 항을 나타내며, A는 접속 우월성으로 인한 사용자의 기대 요인(Expectation factor)을 나타낸다.

평가율 R을 접속호에 대한 양호 확률이나 불만족 확률로 환산하기 위해서는 (식-2)의 오류함수를 사용하여 (식-3)으로부터 얻을 수 있다.

$$Erf(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt \quad (\text{식-2})$$

$$GOB = 100 \cdot Erf\left(\frac{R-60}{16}\right)\% \quad (\text{식-3})$$

$$POW = 100 \cdot Erf\left(\frac{45-R}{16}\right)\%$$

평균 의견 점수로의 환산한 점수는 (식-4)로부터 얻을 수 있다.

$$MOS = 1 + 0.0035 * R + (R - 60) \quad (\text{식-4})$$

$$(100 - R) * 7 * 10^{-6}, 0 < R < 100$$

$$(MOS = 1, R < 0 \text{ 그리고 } MOS = 4.5, R > 100)$$

(식-1)을 구성하는 각 항의 구체적인 수식은 오랜 기간동안 ITU-T와 여러 사업자들에 의해 연구되어 왔다.

망을 계획하고 설계하는 데는 반드시 전송계획을 포함한다. 경제적·기술적인 실행가능성(Feasibility)과 바람직한 망 성능은 망 계획시 주된 고려요인이다. 통화품질에 대한 평가관리는 완전히 새로운 망을 설계할 때뿐만 아니라 기존 망을 확장하거나 성능을 향상시키는 데에도 사용될 수 있다. 체계적인 방법으로 열화요인들의 복합효과를 평가하는 수단으로서 특정 접속호에 대한 통화품질을 추정할 수 있을 뿐만 아니라 사용자

에 의해 인지되는 음성 통화품질의 통계적인 분포에 대한 추정도 할 수 있다. 또한 만족할만한 성능을 유지하기 위해 필요한 개별 열화요인의 기여량과 열화요인들 사이의 상충관계(Trade-off)를 찾을 수 있으며 망에서 나타날 수 있는 다양한 변화에 따른 음성 통화품질의 효과를 추정할 수 있다. 이것은 가입자 회선의 손실변화나 증가된 지연, 새로운 전화기의 사용 등으로 인한 통화품질 평가를 가능하게 하며 무선서비스나 인터넷 폰 등의 새로운 서비스의 음성 통화품질에 대한 효과를 추정할 수 있다.

## 2. 음질 평가 결과

주관적 음질 평가를 위해서 다음과 같은 남녀 화자가 발음한 두 개의 문장을 사용하였다. 사용된 내용은 다음과 같다.

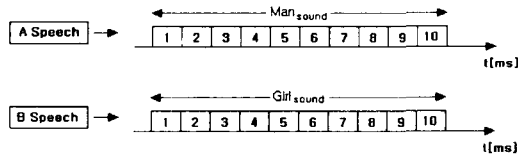


그림 5. 음질의 테스트 재료  
Fig. 5 Test material of speech quality

### 가. 지연(Delay) 측정 결과

본 연구를 수행하는데 있어 음성서비스의 품질에 영향을 미치게 되는 지연은 VoIP 터미널에서 코덱이 음성신호를 압축할 때 다음 3가지 지연을 발생한다.

- 프로세싱 지연 또는 알고리즘 지연
- Look ahead Delay
- 프레임 지연

ITU-T 표준코덱과 그 코덱이 발생시키는 지연 량

- G.711 : 압축되지 않은 64Kbps 음성으로 거의 지연이 없다.
- G.729 : 8Kbps의 음성을 인코딩. 약 25ms의 단방향 지연을 가진다.
- G.723.1 : 6.4Kbps 또는 5.3Kbps의 음성을 인코딩. 약 67.5ms의 단방향 지연

코덱에서 압축 수준이 높을 수록 많은 지연이 발생하며, 좋은 음성 품질을 유지하기 위해서는 낮은 망 지연(Network Latency)이 필요하게 된다. 또한 신뢰성 있는 VoIP 트래픽을 위해 다음과 같은 조건을 지원해야 한다.

- 패킷 전송 지연 : VoIP 대화의 한계수준 (Tolerable Level)을 초과해서는 안된다.
- 패킷 전송 지터 : VoIP 세션을 유지할 수 있는 한계 수준을 초과해서는 안된다.
- 망의 과부하 동안에도 VoIP 세션을 유지할 수 있는 대역폭과 용량을 보장한다.

이러한 내용을 토대로 본 연구는 VoIP의 실험적 테스트를 위해서는 H.323 표준에 의해서 정의되는 전화 기능(Telephone Function)과 PSTN 사용자가 경험해 왔던 수준에 상당하는 QoS를 제공할 수 있는 IP 망(IP Network)과 사설 PBX 기반 망(Private PBX-based Network)이 필요하다.

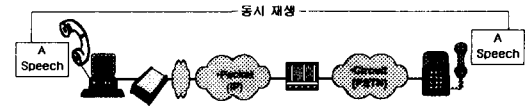


그림 6. 단-대-단 지연측정 방법  
Fig. 6 Method of assessing end-to-end delay

음성 망에서 점-대-점 지연이 높은 경우에 발생할 수 있는 두 가지 문제점이 반향(echo)과 혼선(Talker Overlap)이다. 반향은 Round-Trip 지연이 50ms 이상이 될 때 문제가 된다. 예코는 중요한 품질 문제로 인식되는 것이므로 VoIP 시스템은 반향 제어의 필요성을 제시해야 하며 반향을 제거하는 수단을 구현하여야 한다.

본 연구에서 사용한 이러한 지연의 측정 방법은 <그림 6>과 같다. 송신측과 수신측간의 녹음된 동일 음성을 동일 시간대에 재생시킴으로써 단 대 단 측정을 수행하였다. 지연으로 야기된 지각적 혼란으로 송수신시 메시지의 의미를 파악하는 능력을 실험한 결과는 <그림 7>과 같다. 여기서 보다시피 단방향 지연이 250ms 이상이 될 때 문제가 된다. 따라서 점-대-점 지연은 패킷망에서의 지연을 줄이기 위한 요구를 결정짓는 주요 제약 조건이라고 볼 수 있다.

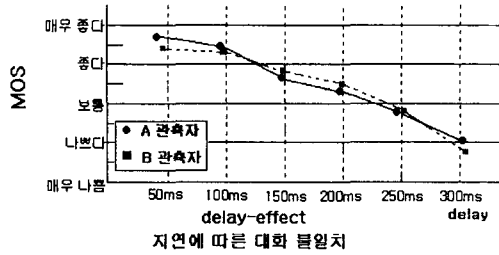


그림 7. 지연에 따른 파일럿 테스트  
Fig. 7 pilot test for impacts of delay

나. 음질 품질 측정 결과

본 절에서의 음질 품질 측정은 <그림 8>과 같이 통합 음성을 평가하기 위해서 문장 세트를 통해 음질의 품질 테스트를 수행하였다.

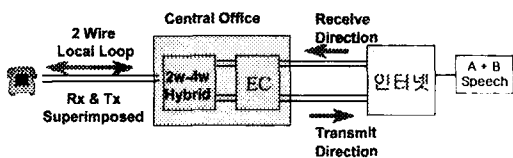


그림 8. 음질 측정 방법  
Fig. 8 method of assessing speech quality

방법 1 : 단어 명료성 검사

문장의 문맥 안에서 단어의 명료성을 검사하기 위한 방법으로서 해당 문장이 실제로 사용되는 빈도에 따라 단어 부분 음운들이 잘 표현될 수 있도록 문장을 선택한다. 테스트는 간단하게 각 문장 주제에 대한 연습은 불필요하며 점수를 매기는 것도 간단하게 사용될 수 있다. 본 실험에서는 A+B Speech의 내용 주제는 들어보지 않은 상태의 테스트 재료의 문장을 재생하였다.

방법 2 : 음성의 이해도 검사

문장 수준 혹은 단어 수준에서 음성의 명료도 (이해도)를 검사하기 위한 방법으로 일상 대화의 변칙적인 특성을 제공하였으며 이것은 손실된 항목이 의미 있는 문장에서와 같이 쉽게 포함될 수 없다는 것을 의미한다. 이 방법은 신뢰성 있는 결과를 얻기 위해서 테스트 주제를 단 한번 적용하여야 하며, 이 테스트 재료의 문장을 보면 다음과

같이 실생활에서 접하기 힘든 문장들이다.

- 문장 1. "다양한 왜곡의 음성 데이터를 얻기 위해서 음향 음성 실험을 해야 한다."
- 문장 2. "대학가라는 말을 붙일 수 있는 거리는, 어느 시중의 거리와 차별될 수 있는 풍속도를 지니고 있어야 한다."

방법 3 : 의미-예측도 검사

이 방법은 문장 수준에서 명료성(이해성)을 검사하기 위한 방법으로 테스트에 사용되는 단어들은 기 정의되어 있는 발음으로써 단어 대부분은 단절음으로 구성하였으며, 일정한 문장 구문 문법에 랜덤하게 명사를 대입하여 테스트하였다.

- 문장 1. "사람이 장독과 보았던 산에 올라 집이다."
- 문장 2. "쿠기에 넘친 카치초 들녘에 있는 바다 소리와 달리다."

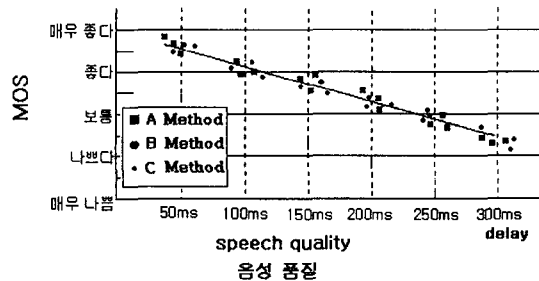


그림 9. 음질 파일럿 테스트  
Fig. 9 pilot test for speech quality

3. 결과 및 고찰

VoIP 음성은 궁극적으로 이용자의 인지 정도에 따라 품질이 결정된다는 가정 하에 주관적 음질 평가 모델을 설정하여 지연에 따른 대화의 불일치와 음질에 대한 파일럿 테스트를 수행하였다. 먼저, 송신측과 수신측간의 녹음된 동일 음성을 동시간대에 재생시킴과 동시에 서로가 청취케 하여 단-대-단에서의 MOS 점수를 얻었다. 청취자가 대체로 100ms 이내에서는 정상적인 대화가 가능하지만 200ms 이상의 지연에서는 사실상 정상적인 대화가 어렵다고 불평하였다.

<그림 9>는 앞에서 제시된 3가지 방법으로 지

연시의 만족도를 측정 한 결과이다. 여기서 알 수 있듯이 VoIP망 내에서 음질의 테스트 문장을 통하여 사람이 인지할 수 있는 척도를 마련한 결과, 지연범위에서는 100ms 범위에서 양호하게 측정되었으며 200ms에서 250ms 이상의 지연범위에서 매우 낮은 점수를 얻을 수 있었다. 따라서 사용자의 주관적 관점을 기준으로 한 품질 척도 설정의 타당성을 살펴 볼 수 있었다.

#### IV. 결론

본 논문에서는 멀티미디어 서비스 요소 중의 하나인 VoIP망에서의 음성 품질에 대한 음질 평가방법과 모델을 제시하였다. 그리고 멀티미디어 서비스 특성에 기인한 주관적 평가의 중요성을 제시하여 지연에 따른 대화 불일치와 음질에 대한 파일럿 테스트를 수행해 본 결과, 지연 요인과 사용자가 느끼는 음성 품질간의 추세를 통한 상관관계를 발견하여 주관적 품질 척도를 제안하였다.

본 연구를 통하여 얻어진 품질측정 척도 지침으로 VoIP 관련 사업의 품질 기준을 제정할 수 있으며, 공공서비스에 대한 측정 도구로 활용되어 개발 업체의 기술 발전은 물론 고객에 대한 양질의 서비스를 도모하는데 기여할 수 있을 것이다.

#### 참 고 문 헌

[1] Scott Keagy, 2000, Integrating Voice and Data Networks, Cisco Press, Indianapolis USA, p.145~p.164.

[2] Wu, Raymond, "Objective Assessment of the Perceived Quality in VoIP Systems", ETSI EP TIPHON 10, 1998.

[3] ETSI EG 201 377-1 v1.1.1: "STQ; Specification and measurement of speech transmission quality; part 1: Introduction to objective comparison measurement methods for one-way speech quality across networks", April, 1999.

[4] TechBrief Extreme Networks: "Combining Voice over IP with Policy-Based Quality of

Service"

[5] Sami Lemmetty, "Review of Speech Synthesis Technology; 10. Speech Quality and Evaluation", June, 1999

[6] 김장형, "멀티미디어 콘텐츠 서비스의 주관적 품질평가에 관한 연구", 정보통신부, 2001

[7] 이인섭, "인터넷 전화 서비스품질" TTA 저널, 68호, 2000

[8] Anna Watson, "Multimedia Conferencing Via Multicast: Determining the Quality of Service Required by the End User", 1999.

[9] Wu, Raymond, "Objective Assessment of the Perceived Quality in VoIP Systems", October, 1998



강 영 도(Young-do Kang)

1989년 2월 제주대학교 정보공학과(공학사)  
1998년 8월 조선대학교 교육대학원 전자계산전공(교육학석사)  
2001년 2월~현재 제주대학교

정보공학과 대학원 박사과정

\* 관심분야: 멀티미디어 시스템, 영상처리

\* e-Mail : jongcom@hanmail.net



강 진 석(Jin-Suk Kang)

1999년 2월 제주대학교 정보공학과(공학사)  
2001년 2월 제주대학교 대학원 정보공학과(공학석사)  
2001년 2월~현재 제주대학교

정보공학과 대학원 박사과정

\* 관심분야: 멀티미디어 시스템, 영상처리

\* e-Mail : jskang01@cheju1.cheju.ac.kr





**최연성(Yeon-Sung Choi)**

1982년 2월 중앙대학교 전자공학과 졸업(공학사)

1984년 2월 중앙대학교 대학원 전자공학과(공학석사)

1990년 2월 중앙대학교

전자공학과 대학원 졸업(공학박사)

1988년 3월-1991년 2월 제주대학교 정보공학과 조교수

1991년 6월-현재 군산대학교 전자정보공학부 부교수

1995년 - 1996년 군산대학교 전자계산소장

\* 관심분야 : 영상처리, 멀티미디어 시스템, 신호처리 시스템

\* e-Mail : yschoi@kunsan.ac.kr



**김장형(Jang-Hyung Kim)**

1981년 2월 홍익대학교 정밀기계공학과(공학사)

1983년 2월 연세대학교 대학원 기계공학과 (공학석사)

1990년 8월 홍익대학교 대학원 기계공학과 (공학박사)

1995년 2월~현재 제주대학교 통신컴퓨터공학부 교수

1998년 3월~2000년 5월 제주대학교 전자계산소장

1999년~현재 해양정보통신학회 제주지부장

2001년 2월~현재 정보과학회 이사

\* 관심분야 : CAD/CAM, 멀티미디어, 인공지능

\* e-Mail : janghkh@cheju.cheju.ac.kr