

## 음성전송지수를 이용한 확성전화기의 명료도 평가 방법

### A Study on the Speech Transmission Index Method for Estimating Articulation of Loudspeaking Telephony

장 대 영\*, 강 성 훈\*, 심 동 연\*\*, 김 천 덕\*\*  
(Daeyoung Jang, Seonghoon Kang, Dongyeon Sim, Chunduck Kim)

#### 요 약

전화기의 통화품질은 음량정적으로 규정하고 있으나, 이 방법은 핸드셋 전화기에만 국한되는 방법이다. 핸드 프리 전화기는 실내의 음장의 영향을 더 많이 받으므로 전송 특성뿐만이 아니라 주위 잡음, 에코, 잔향도 포함하여 평가하여야 한다. 따라서 핸드 프리 전화기의 품질을 평가할 수 있는 새로운 방법이 필요하다.

Steeneken은 음성 전송 지수(Speech Transmission Index; STI)를 계산하여 음성 전송 특성을 평가하는 객관적인 방법을 제안하였다. 본 논문에서는 STI를 핸드 프리 전화기의 통화품질 평가 방법에 적용 가능성을 고찰하고, 고속으로 STI를 계산할 수 있는 시스템을 구현하였다. 이 시스템을 이용하여 잔향 시간이 다른 세군데의 실내에서 핸드 프리 전화기의 STI를 측정된 결과, 실내의 잔향시간이 길어질수록 STI가 감소되는 것을 알았다. 이 결과는 STI를 음장 특성을 포함하는 명료도 평가 방법에도 응용할 수 있다는 것을 시사하고 있다.

#### ABSTRACT

The speech transmission quality in telephone is quantified in terms loudness rating, but this method has been validated only for the handset telephony. The transmission quality of loudspeaking telephony in any room must be evaluated not only with speech transmission but also with background noise, echo and reverberation since the effect of room acoustics is much stronger for loudspeaking telephony. Therefore, it requires a better approach to specify the quality of loudspeaking telephony.

By calculating the speech transmission index(STI), a physical method for measuring the quality of speech transmission was proposed by Steeneken. In this paper, the application of a STI method for estimating articulation of loudspeaking telephony was discussed. And the STI measurement system with high speed calculation was also implemented. By using this system, the STI measurements of loudspeaking telephony system were performed in three rooms, having different reverberation times. The results show that the STI decreases as the reverberation time of rooms increases. It suggests that speech transmission index method can be useful evaluating articulation of a loudspeaking telephony including the sound field characteristics.

\*한국전자통신연구소 음향정보처리연구실

\*\*부산수산대학교 전자공학과

접수일자: 1994년 3월 9일

## I. 서 론

화성 전화기의 통화품질은 핸드셋 전화기와는 달리 송·수화기와 사용자간의 거리가 멀기 때문에 핸드셋 전화기의 통화품질 열화요인이외에 주위의 소음과 실내음장의 영향을 많이 받게 되므로 이러한 것을 고려하여 평가하여야 한다. 일반적으로 음성통신계의 명료도를 평가하기 위한 방법으로는 사람을 대상으로 한 주관적 방법과 물리적 파라미터를 기초로 한 객관적 방법이 있다. 전자는 청취자의 청각에 의한 무의미 단음절을 이용한 단위 명료도 또는 임의의 문장에 대한 요해도등이 있으며, 후자에는 French등이 제시한 명료도 지수(AI : Articulation Index)법에 의한 평가방법이 있다[1]. AI방법은 실내음장에서 잔향이나 환경잡음의 영향을 평가할 수 없으므로, 1973년 Haugast등은 변조전송함수(MTF : Modulation Transfer Function)를 제안하였고[2]. 그후 1980년에 이 함수를 이용하여 음성명료도를 정량적으로 표현한 음성전송지수(STI : Speech Transmission Index)를 제안하였다[3][4].

그후 Hamada등은 디지털 신호처리 기술과 미니급 전자 계산기를 이용하여 Haugast등이 제안한 MTF-STI법을 실내 음장에 설치한 음향기기의 통화 품질 평가에 실제로 응용하였다[5][6][7]. 그들은 실내 음장의 임펄스 응답을 실험적으로 측정하여 변조전송함수를 이용한 음성전송지수를 환경잡음의 영향과 잔향을 분석하는데 이용하였다. 여기서 임펄스 응답을 측정하기 위하여 Hamada등은 음원신호로 백색잡음을 사용하고 PCM 레코더에 수신신호를 저장한 다음, 미니급 계산기에서 크로스 스펙트럼법으로써 임펄스 응답을 산출하였다. 그러나 이 방법은 잔향시간이 긴 경우 정확한 임펄스 응답을 구하기 위해서는 수분 이상의 평균을 필요로 하므로 PC급 컴퓨터로는 현장에서 실시간 처리가 불가능하다는 단점을 가지고 있다.

본 논문에서는 현장에서 실시간에 가까운 처리를 할 수 있는 방법을 제안하기 위하여, MLS(Maximum Length Sequence)신호를 음원신호로 사용하여 화성 전화기를 포함한 실내 음장의 임펄스 응답의 측정에 있어서 짧은 시간으로도 고정밀 측정이 가능하도록 한 측정법과, 음성명료도지수를 산출하기 위한 MTF-STI법의 소프트웨어적인 처리, 샘플링 주파수가 높

은 경우의 필터링 오차를 줄이기 위한 옥타브 대역통과필터와 서여통과필터의 구성 알고리즘의 개발에 대해서 기술한다.

## II. 변조전송함수의 이론적 배경

### 1. 변조전송함수의 정의

음장의 명료도를 평가하는데 있어서 일반적으로 명료도 지수법을 이용하고 있으나, 잔향 및 환경 잡음이 있는 상황에는 적용할 수 없다. 따라서 Haugast등은 왜곡이 없는 100% 정현파 변조한 음원신호의 포락선과 유상을 통과한 신호의 포락선을 비교하여, 그 신호들의 상대적 감쇠도를 나타내는 변조전송함수를 제안하여 잔향이나 환경잡음을 분석하는데 응용하였다[2].

100% 정현파 변조한 입력신호인 경우,  $m(F) = 1$ 을 나타내며 음장을 경유한 출력은 다음과 같다.

$$\begin{aligned} I_{ik}(t) &= \bar{I}_{ik} |1 + \cos(2\pi Ft)| \\ &= \bar{I}_{ik} |1 + \operatorname{Re}[\exp(j2\pi Ft)]| \end{aligned} \quad (1)$$

$$I_{ok}(t) = \bar{I}_{ok} |1 + m(F)\cos(2\pi Ft)| \quad (2)$$

윗식에서  $\bar{I}_{ik}$ ,  $\bar{I}_{ok}$ 는 입력신호와 출력신호의 옥타브당 평균 인텐시티(Octave Mean Intensity)이고,  $F$ 는 변조주파수이며 점자  $k$ 는  $k$ 번째 옥타브 대역을 나타낸다. 식(2)의 출력신호는 진폭변조파형이며, 실내 음장의 임펄스 응답의 사승과 입력신호의 콘볼루션으로 다음과 같이 나타낼 수 있다[7].

$$I_{ok}(t) = \int_0^t I_{ik}(t-\tau) \cdot h^2(\tau) d\tau \quad (3)$$

식 (3)에 식 (1)을 대입하면 다음과 같이 된다.

$$I_{ok}(t) = \int_0^t \bar{I}_{ik} |1 + \operatorname{Re}[\exp(j2\pi F(t-\tau))]| \cdot h^2(\tau) d\tau \quad (4)$$

식 (4)를 식 (2)와 같이 정리하면 식(5)와 같이 된다.

$$I_{ok}(t) = a \cdot \bar{I}_{ik} |1 + \frac{1}{a} \cdot \operatorname{Re}[b \cdot \exp(j2\pi Ft)]| \quad (5)$$

여기서 파라미터  $a$ ,  $b$ 는 다음과 같다.

$$a = \int_0^T h^2(\tau) d\tau \quad (6)$$

$$b = \int_0^T h^2(\tau) \exp(-j2\pi F\tau) d\tau \quad (7)$$

그러므로 변조전송함수는 음향기기를 포함한 실내 음장의 임펄스 응답으로 표현하면 다음과 같다.

$$m(F) = \frac{|b|}{a} = \frac{\left| \int_0^T h^2(\tau) \exp(-j2\pi F\tau) d\tau \right|}{\int_0^T h^2(\tau) d\tau} \quad (8)$$

## 2. 변조전송함수를 이용한 음성전송지수 산출

음성전송지수는 잔향이나 잡음이 있는 음장에서 음성명료도를 0~1범위에서 정량적으로 표현하는 물리적 지표이다. 산출과정은 명료도지수법을 기초로 하여 음성대역 500Hz~8kHz를 5개의 옥타브 밴드로 나누고, 0~40Hz 주파수에 대한 변조율의 포락선을 1/3 옥타브 밴드로 분석하여 계산한다[3].

음성전송지수를 구하는데 있어서 청각기관의 저주파수 성분에 의한 고주파수 성분의 매스킹 효과를 보상에 주어야 한다. 이를 고려한 신호의 세기는 다음 식과 같이 표현된다.

$$I_{AM,k} = \bar{I}_{ok-1} AMF \quad (9)$$

여기서  $I_{AM,k}$ 는 분석하고자 하는 옥타브 대역당 Apparent Masking Intensity이고,  $AMF$ 는 청각 매스킹 감쇠 인자로서 그 값은 0.0003이다[3].

식 (9)를 이용하여 청각보상인자  $ACF_k$ 를 구하면 식 (10)과 같이 된다.

$$ACF_k = \bar{I}_{ok} / (\bar{I}_{ok} + I_{AM,k}) \quad (10)$$

1절에서 구한 변조전송함수에 청각 매스킹 효과를 보상하면 다음과 같다.

$$m'_k(F) = ACF_k \cdot m_k(F) \quad (0 \leq m'_k(F) \leq 1) \quad (11)$$

또한, 등가 S/N비를 구하여 전송지수  $TI_k$ 로 변환하면 식 (13)과 같다.

$$SNR_k(F) = 10 \log \{ m'_k(F) / (1 - m'_k(F)) \} \quad (12)$$

$$TI_k = (SNR_k(F) - S) / R \quad (13)$$

여기서  $R$ 은 전송지수에 기여하는 S/N의 범위를 표시하고,  $S$ 는 그 지수에 대한 기여도가 0일때의 등가 S/N비이다. 옥타브 중심주파수  $k$ 의 기여율 변조전송지수 (Modulation Transfer Index)는 식 (14)와 같다.

$$MTI_k(F) = \frac{1}{N} \sum_{i=0}^{N-1} TI_k(F_i) \quad (14)$$

여기서  $N$ 은 분석하고자 하는 1/3 옥타브대역의 수이다. 옥타브당 음성대역 가중치  $W_k$ 를 곱함으로써 음성전송지수는 다음과 같이 구할 수 있다.

$$STI(F) = \sum_{k=1}^5 (W_k \cdot MTI_k(F)) \quad (15)$$

## III. MTF-STI 측정 시스템

### 1. MLS 신호를 이용한 임펄스 응답 측정 시스템 설계

변조전송함수는 고전적 방법인 정현파 변조에 의해 구할 수 있지만, 식 (8)에서 알 수 있듯이, 실내음장의 임펄스 응답을 구하게 되면 변조전송함수를 정확히 계산해 낼 수 있다. Hamada등[5]이 제시한 백색잡음을 이용한 크로스 스펙트럼법에 의한 임펄스 응답의 산출방법은 PC급 컴퓨터로 실시간 처리가 불가능하다.

본 논문에서는 시험신호로 단속적인 MLS신호를 사용함으로써 FFT 연산횟수를 대폭적으로 줄이는 임펄스 응답 측정 시스템을 제안하였다. 그리고 MLS신호를 이용하면, 외부장치에서 동기화하는 것에 의해 PC급 컴퓨터상에서 5회의 FFT만으로 고정도의 실내 음장의 응답특성을 구할 수 있다. MLS신호에 의한 크로스 스펙트럼에 의한 전달함수를 수식적으로 기술하면 다음 식과 같다.

$$\hat{H}_{xy}(k) = [W_{1y}(k) + W_{2y}(k)] / [W_{11}(k) + W_{22}(k)] \quad (16)$$

$W_{1y}$ 와  $W_{2y}$ 는 입력신호와 출력신호간의 크로스 스펙트럼이며,  $W_{11}$ 과  $W_{22}$ 는 2종류의 MLS신호의 파워 스펙트럼이다. 여기서 고려해야 할 점은 그림 1에 나타난 바와 같이, 두 종류의 MLS신호를 발생시키는

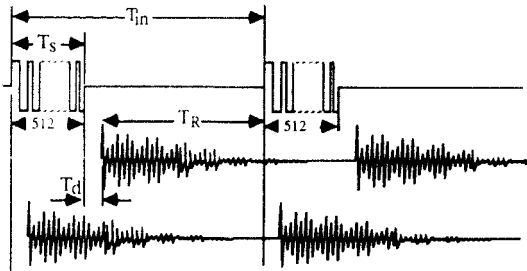


그림 1. 신호원과 그 응답과의 관계

데 있어서 실내 잔향시간과 입출력간의 군지연(Group delay) 시간 및 발생신호 시간과의 관계이다. MLS 신호의 주기는 다음과 같은 조건을 만족해야만 한다.

$$T_{in} \geq T_d + T_s + T_R \quad (17)$$

(17)식에서  $T_d$ 는 음원신호의 발생점에서 수신점간의 거리에 따른 군지연이고,  $T_s$ 는 MLS신호의 지속 시간이며,  $T_R$ 은 실내의 잔향시간이다.

본 논문에서는 무향실 및 잔향시간이 다른 두 장소에서 실험을 하였다. 음원신호는 10차의 MLS신호 ( $X^{10} \oplus X^9 \oplus X^6 \oplus X^5 \oplus X^4 \oplus X^3 \oplus 1$ )를 1023점 생성하여 마지막점에 한 점을 추가하여 512점씩 두 종류의 신호를 발생시켰다. 시스템에서 사용하는 A/D변환기의 샘플링 주파수는 48kHz이고, 입출력 신호를 차단주파수 22kHz의 저역필터를 통과시키므로 음원신호의 지속시간  $T_s$ 는 11ms(512점/48kHz)이다.

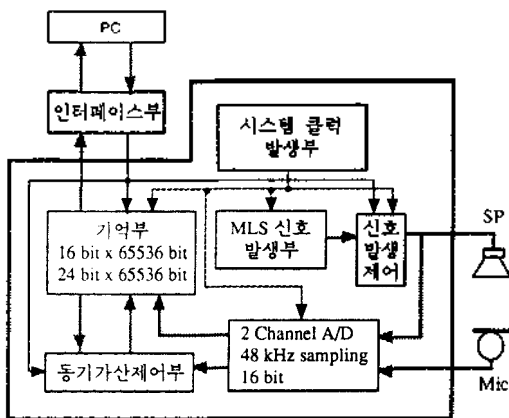


그림 2. MTF 측정시스템의 구성도

## 2. 시스템 구성도

본 논문에서 설계·제작한 시스템의 하드웨어 구성은 그림 2와 같다. 구성도에서 PC는 시스템 제어부로서 MLS신호의 발생, A/D변환, 동기 가산 및 수(16, 32, 64, 128회)를 결정하고, 음원신호의 구동주기의 제어 및 실시간에 동기 가산된 입출력 신호를 시스템 메모리로부터 하드 디스크에 저장한 후, 식(16)에 근거한 FFT연산에 의해 임펄스 응답을 계산한다.

## IV. 변조전송함수 및 음성전송지수의 산출

### 1. 변조전송함수 산출

변조전송함수를 산출하는 과정은 음원의 구동신호로서 10차의 MLS신호를 이용하여 확성 전화기를 포함한 실내음장의 임펄스 응답을 측정하고, 변조전송함수는 그림 3과 같이 컴퓨터상에서 소프트웨어적으로 계산한다.

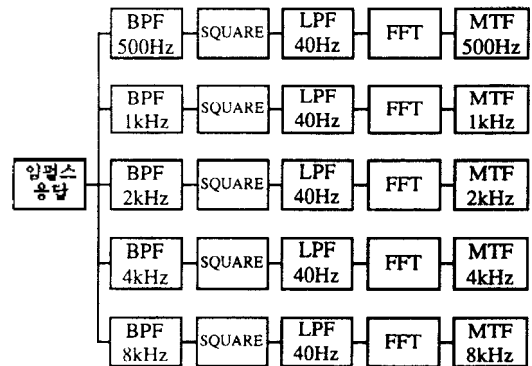


그림 3. MTF의 산출과정

여기서 이용한 대역통과필터와 저역통과필터는 현재 이용되고 있는 필터 알고리즘[4][9]으로는 샘플링 주파수(48kHz)가 너무 높기 때문에 정확한 필터링이 되지 않으므로, 주파수 영역에서 먼저 통과대역을 포함하는 구형파를 설정한 다음 역푸리에 변환하고, 시간상 함수를 부과하는 방법으로 필터를 실현하였다.

### 2. 음성전송지수 산출 알고리즘

컴퓨터상에서 계산된 MTF를 이용하여 식 (9)~식

(15)에 의하여 그림 1와 같은 절차에 의해 컴퓨터상에서 음성전송지수를 구하였다.

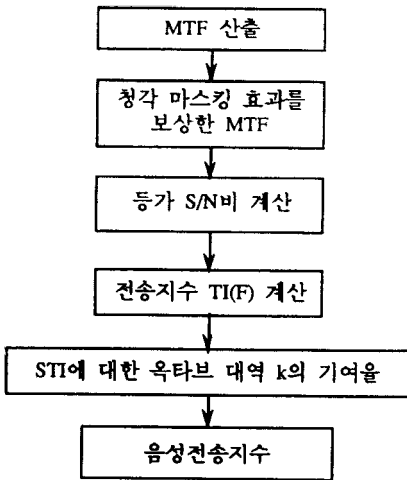


그림 4. 음성전송지수의 산출과정

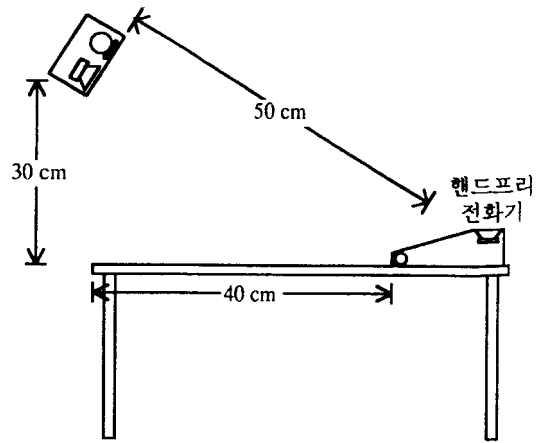
V. 실험

1. 실험 방법

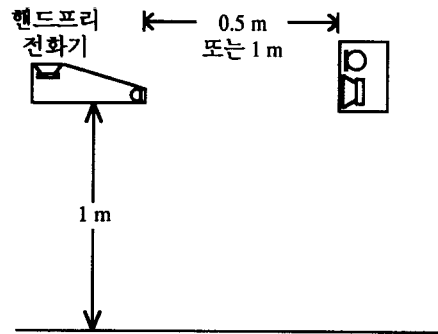
본 논문에서 제시한 실내음장의 임펄스응답 측정 방법과 변조전송함수 및 음성전송지수 산출 알고리즘의 타당성을 검토하기 위하여, 실내음장에 확성전화기를 설치하여 실험하였다. 측정장소는 실내의 잔향시간의 차이에 따른 MTF의 변화를 고찰하기 위하여 무향실(3.4m x 2.63m x 2.03m), 잔향시간이 0.3초인 연구실(7m x 4m x 2.8m), 잔향시간이 0.7초인 강의실(7m x 15m x 2.8m)로 선정하였다. 확성전화기의 설치는 그림 5와 같이 ITU-T 권고[9]와 0.5m 및 1m의 위치로 하였다.

2. 실험 결과 및 고찰

실험방법에서 제시한 3가지 조건중 ITU-T 권고에 따라 MTF를 무향실에서 측정된 결과를 그림 6에 나타낸다. 그림 6의 (a)는 임펄스 응답, (b)는 스펙트럼, (c)는 변조전송함수의 측정결과이며, 여기서 주파수는 옥타브 밴드의 중심 주파수를 의미한다. 중심 주파수 500Hz의 MTF는 무향실인데도 불구하고 감쇠가 크게 나타나는 이유는 무향실의 규격이 파장에 비해 적어 마이크로폰으로 들어오는 500Hz 이하의



(a) ITU-T 권고의 측정 조건



(b) 0.5 m, 1m 거리 조건

그림 5. 확성전화기의 STI 측정조건

신호성분이 감쇠되기 때문이다.

그림 7과 그림 8은 설치조건과 실내음장에 따른 음성전송지수의 변화를 나타낸다. 그림 7은 실내음장에 따른 확성전화기의 마이크로폰계와 스피커계의 음성전송지수를 비교한 것이다. 그림에서 실내의 잔향시간이 길어짐에 따라 마이크로폰계와 스피커계의 음성전송지수가 다같이 낮아지고 있는 것을 알 수 있다. 또한 그림 7의 (a)가 (b), (c)보다 값이 큰 이유는 탁자에서 1차 반사되는 음향파워가 공간음장을 경유한 잔향성분의 음향파워보다 크기 때문이다. 그림 8 (a)는 각 실내 음장에 대한 마이크로폰계의 음성전송지수를 나타내고, (b)는 스피커계의 음성전송지수를 나타낸다. 이들 그림에서 각 실험장소에서 잔향시간이 길어지면 음성전송지수가 낮아지는 것을 알 수 있다.

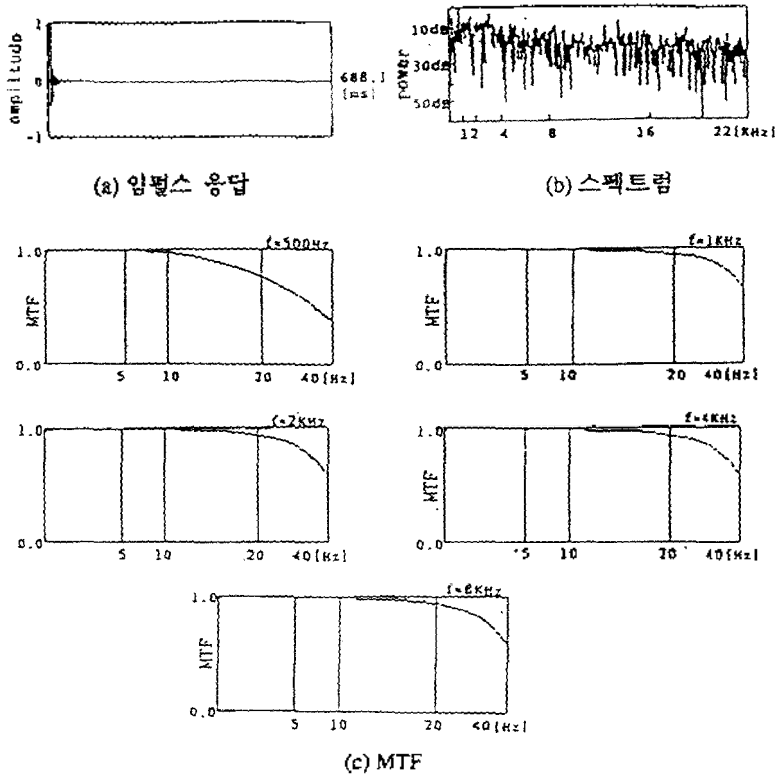


그림 6. 무향실 측정결과(TTU-T 권5)

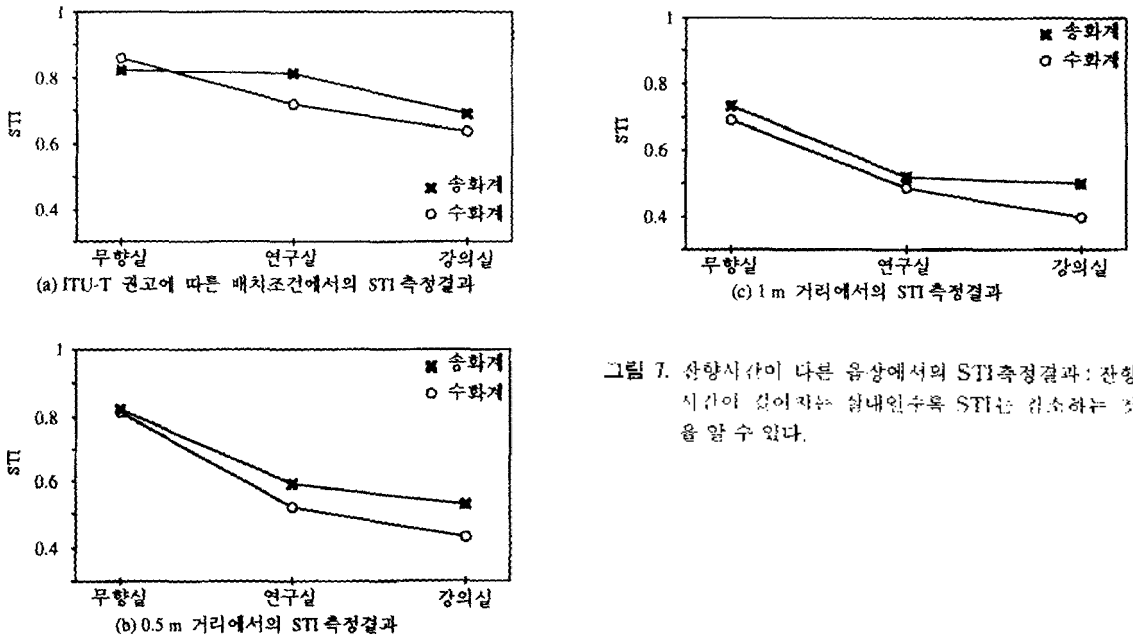
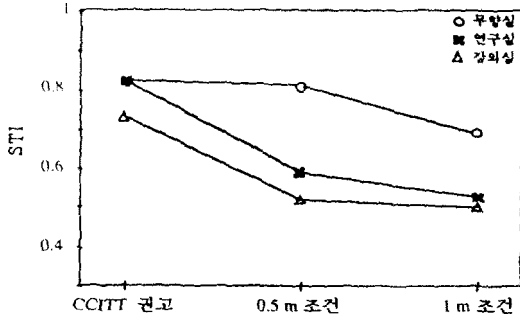
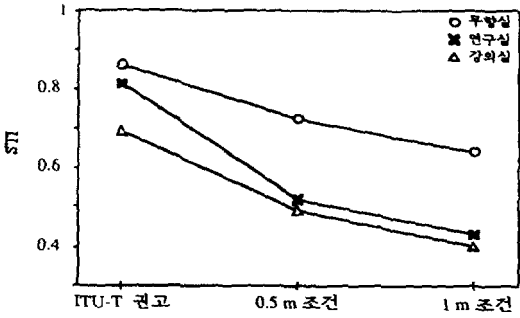


그림 7. 소향시간이 다른 음상에서의 STI 측정결과: 잔향 시간이 길어지는 실내일수록 STI는 감소하는 것을 알 수 있다.



(a) 송화 STI의 측정 결과



(b) 수화 STI의 측정 결과

그림 8. 확성 전화기의 송·수화 STI의 측정결과

VI. 결 론

본 논문에서는 실내음장을 경유하는 확성전화기의 객관적 음성전송 품질을 평가함에 있어, AI법에 대신하여 Haugast가 제안한 MTF-STI법을 이용하고 디지털 신호처리 기술을 응용하여 고정도, 고속으로 측정할 수 있는 시스템을 구현하였다. 측정시스템은 마이크로 컴퓨터에 외부장치를 접속시켰으며, 그 장치는 MLS 신호 발생부, 임펄스응답 동기가산부, 2채널 고속 변환부와 이들 각부들 제어하는 제어부로 구성되어 있다.

측정 시스템의 주요 기능은 확성전화기를 설치한 실내음장의 임펄스응답을 고속으로 정확하게 측정하는 것이다. 외부장치에서 임펄스응답 측정을 위한 MLS 신호를 발생시키고, 동기가산한 응답신호를 마이크로 컴퓨터에 입력해서 5회의 FFT 연산만으로 음향 전달계의 임펄스응답을 구하므로 데이터량 및 처리시간을 대폭적으로 단축시킬 수 있었다. 이 시스템의 성능을 평가하기 위해서 잔향시간이 다른 무향실, 연

구실, 강의실에서 확성전화기를 설치하여 ITU T 설치조건에 따라 음성 전송지수를 측정하였다. 음장에 대한 음성전송지수의 정량적평가는 설치방법을 동일한 조건으로 하고, 구동원인 스피커와 확성전화기의 마이크로폰 사이의 음성전송지수를 구한 결과, 잔향시간이 길어질수록 STI가 점점 낮아지는 것은 예상된 결과와 일치하는 것이다. 그러므로 이 결과는 STI를 음장특성을 포함하는 확성전화기의 명료도 평가 방법에도 응용할 수 있다는 것을 시사하고 있다.

참 고 문 헌

1. French N. R., Steinberg J. C., "Factors governing the intelligibility of speech sounds," J. Acoust. Soc. Amer. 19, 1947.
2. T. Hougast, H. J. M. Steeneken, "The modulation transfer function in room acoustics as a predictor of speech intelligibility," Acustica 28, 1973.
3. H. J. M. Steeneken, T. Hougast, "A Physical method for measuring speech transmission quality," J. Acoust. Soc. Am. 67, 1980.
4. T. Hougast, H. J. M. Steeneken, R. Plomp "Predicting speech intelligibility in rooms from the modulation transfer function, I, General room acoustics," Acustica 46, 1980.
5. Hitoshi Fujigaki, Hareo Hamada, Tanetoshi Miura, "Evaluation of the hands free telephones using MTF-STI methods," EA87-42, 1987.
6. Hitoshi Fujigaki, Hareo Hamada, Tanetoshi Miura, "Relation between the physical characteristics and the articulation of a hands free telephone," EA88-4, 1988.
7. Yasuo Ogura, Hareo Hamada, Tanetoshi Miura, "音響における音響伝送品質のための MTF と STI について," 日本音響學會, Vol. 40, 1984.
8. E. O. Brigham, *The Fast Fourier Transform*, Prentice-Hall Inc., New Jersey, 1974.
9. CCITT Recommendation, "Transmission characteristics of loudspeaker telephone," Rec. P.34, 1984.

▲장 대 영(Daeyoung Jang) 1966년 2월 24일생  
현재 : 한국전자동신연구소 연구원  
(1992년 제 11권 1호 참조)

▲강 성 훈(Seonghoon Kang) 1956년 6월 15일생  
현재 : 한국전자동신연구소 음향정보처리연구실장  
(1990년 제 9권 1호 참조)

▲김 천 덕(Chunduck Kim) 1946년 5월 23일생  
현재 : 부산수산대학교 전기공학과 교수  
(1988년 제 7권 2호 참조)

▲심 동 연(Dongyeon Sim) 1964년 7월 18일생



1989년 2월 : 부산수산대학교 전  
자동신공학과 졸업

1991년 2월 : 부산수산대학교 대  
학원 전자통신공학  
과 졸업(공학석사)

1991년 6월 ~ 1992년 : 일본 EAG-  
LE 주식회사 근무

1993년 4월 ~ 현재 : 일본 동북대  
학 기계전자공학과  
박사과정

※주관심분야 : 디지털 신호처리, 액츄에이터, 마이크  
로 머신