

# 블루투스 네트워크 환경에서의 효율적인 음성전송 시스템 구현<sup>1)</sup>

\*김명중 \*\*박지훈 \*\*김홍국

\*한국산업기술대학교 전자공학과

\*\*광주과학기술원 정보통신공학과

\*oopsbbada@kpu.ac.kr \*\*{jh\_park, hongkook}@gist.ac.kr

## Implementation of an Efficient Voice Transmission System in Bluetooth Network Environments

\*Myung Jong Kim \*\*Ji Hun Park \*\*Hong Kook Kim

\*Dept. of Electronics Engineering, Korea Polytechnic University

\*\*Dept. of Information and Communications, Gwangju Institute of Science and Technology

### 요약

IPTV의 상용화에 맞추어 사용자와 TV간의 정보 교환에 의한 대화형 서비스들이 제공되고 있으며, 특히 음성인식 기술은 이러한 서비스를 실현하기 위한 중요한 기술 중의 하나로 대두되고 있다. TV에서의 음성인식 수행을 위해서는 가정환경과 같은 제한된 공간에서 효율적으로 사용자의 음성을 TV에 전송할 수 있는 근거리 무선통신 수단이 필요하게 된다. 특히, 리모트 컨트롤러와 같은 저전력 시스템 환경에서 구현이 가능해야 한다. 따라서 이러한 제한된 조건에서 최적의 성능을 갖는 음성전송 시스템 개발이 요구되고 있다.

본 논문에서는 블루투스 환경 하에서 음성인식을 위해 필요한 음성전송 시스템을 실시간 구현한다. 효율적인 음성전송을 위해 G.711을 기본 코덱으로 사용하며, 음성전송 시 발생하는 패킷손실에 따른 음성 품질 저하를 줄이기 위해 G.711 패킷손실 은닉 알고리즘을 음성전송 시스템에 적용한다. 특히 G.711 패킷 손실 은닉 알고리즘 수행을 위해 블루투스 프로토콜 스택 application layer에 RTP 프로토콜을 적용하여 패킷 손실 여부를 확인하고, 패킷 손실 발생 시 패킷손실 은닉 알고리즘을 통해 음성의 품질 저하를 줄인다. 구현된 시스템의 성능을 평가한 결과, G.711 패킷 손실 알고리즘을 적용하여 2~10%의 패킷손실 환경에서 14.7%의 음질개선을 얻을 수 있었다.

### 1. 서론

IPTV의 상용화에 맞추어 사용자와 TV간의 정보를 교환하는 대화형 서비스를 근간으로 하는 콘텐츠들이 제공되고 있으며, 특히 음성인식기술은 이러한 서비스를 실현하기 위한 중요한 기술 중의 하나로 대두되고 있다. TV에서의 음성인식을 수행하기 위해서는 가정환경과 같은 제한된 공간에서 효율적으로 사용자의 음성을 TV에 전송할 수 있는 근거리 무선통신 수단이 필요하게 된다. 이와 같은 무선통신 수단으로는 ZigBee (IEEE 802.15.4), WirelessLAN (IEEE 802.11b), HomeRF, 블루투스 (IEEE 802.15.1), RFID 등을 들 수 있다. 이 중에서도 블루투스(Bluetooth)는 HomeRF, WirelessLAN에 비하여 상대적으로 가격이 저렴하며, 전력소모가 적다는 장점이 있으며, 또한 ZigBee나 RFID 등에 비해서는 전송률 측면에서 음성 전송 서비스를 제공할 수 있는 장점이 있다 [1][2]. 따라서 리모트 컨트롤러 (remote controller)와 같은 저전력 시스템 환경에서 블루투스 네트워크를 이용하여 효율적으로 음성을 TV에 전송할 수 있다.

본 논문에서는 블루투스 환경 하에서의 음성인식을 위해 필요한 음성전송 시스템을 실시간 구현한다. 효율적인 음성전송을 위해 G.711 [3]을 기본 코덱으로 사용하며, 음성전송 시 발생하는 패킷 손실

에 따른 음성 품질 저하를 줄이기 위해 G.711 은닉 알고리즘 [4]을 음성전송 시스템에 적용한다. 특히 G.711 패킷 손실 은닉 알고리즘 (PLC: packet loss concealment) 수행을 위해 블루투스 프로토콜 스택 상위단에 RTP (real-time transport protocol) 프로토콜 [5]을 적용하여 패킷 손실 여부를 확인한다. 구현된 시스템은 실시간 동작여부 확인 및 패킷 손실 은닉 알고리즘 적용 전후에 대한 음질 평가를 통해 성능을 검증한다.

서론에 이어 2장에서는 블루투스 피코넷 기반 음성전송 시스템 및 적용하고자 하는 음성코덱과 패킷 손실 은닉 알고리즘에 대해 설명한다. 3장에서는 이들에 대한 구현에 있어서 문제점 및 그 해결책을 제시한다. 이어 4장에서는 성능 평가 결과에 대해 기술하고 마지막 5장에서 결론을 맺는다.

### 2. 블루투스 피코넷 기반 음성전송 시스템

그림 1은 두 peer 간에 일대일로 이루어지는 음성전송 시스템의 구성도를 나타낸다. 먼저 음성 리모트 컨트롤러에 해당하는 송신단 (client)은 사운드 카드로부터 음성신호를 입력받아 음성 인코더에 의

1) 본 논문은 제 1저자가 광주과학기술원 정보통신공학과 2007년도 겨울학기 인턴과정 수행중에 진행한 연구 결과임.

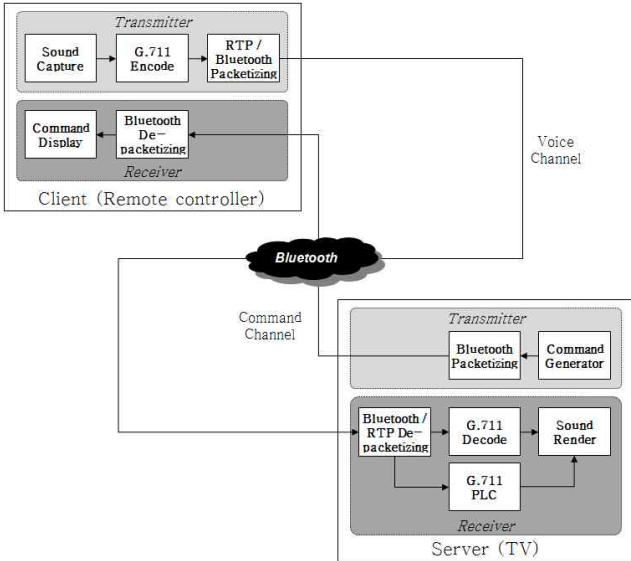


그림 1. 블루투스 피코넷 기반 음성전송 시스템 구성도

해 압축된 음성 비트스트림을 생성하며, 이를 위해 G.711을 기본코덱으로 사용한다. 생성된 음성 프레임은 패킷 순서 (sequence number) 및 전송시간 (timestamp), 참여자 식별정보 (SSRC identifier) 등을 포함하는 RTP 패킷화 과정을 거친다. 이 패킷으로부터 블루투스 header format과 payload format으로 구성된 부가적인 패킷을 생성한 후 블루투스 네트워크를 통하여 전송된다. RTP 프로토콜은 수신단에서 패킷 손실에 따른 음질 저하를 줄이기 위하여 G.711 패킷 손실 은닉 알고리즘 수행을 위해 블루투스 프로토콜 스택에서 application layer에 추가된다. 한편, TV 수신기에 해당하는 수신단(server)에서는 수신된 블루투스 패킷의 header format과 payload format, 그리고 RTP 패킷을 분석하여 패킷 순서를 통해 패킷 손실 정보를 얻는다. 이어 G.711 복호화를 수행한다. 복호화 과정에서 패킷 손실이 발생하였을 경우 G.711 패킷 손실 은닉 알고리즘을 통해 손실된 패킷을 복원한다. 한편, 양방향 통신을 위해 수신단에서는 역으로 필요한 정보를 블루투스 header format과 payload format으로 패킷을 구성한 후 블루투스 네트워크를 통하여 송신단으로 전송한다.

### 가. G.711 음성 코덱

G.711은 VoIP (voice over Internet protocol)에서 가장 기본적으로 사용되는 음성코덱으로 A/μ-law PCM (pulse coded modulation) 기반의 부호화를 수행하며 8 kHz의 표본화율(sampling rate)에서 64 kbps의 비트율(bitrate)을 가진다. 또한 G.711은 샘플단위의 코덱으로서 샘플당 8 bit PCM을 정의하고 있다. 즉, A-law 방식의 경우 sign bit와 마지막 3 bit를 제외한 나머지 13 bit 값을, μ-law의 경우 sign bit와 마지막 2 bit를 제외한 14 bit 값을 8 bit 로그 샘플로 변환한다.

### 나. 블루투스 프로토콜 스택

그림 2는 블루투스 프로토콜 스택 [6]을 보여준다. 프로토콜 스택이란 그림 2에서 보는 바와 같이 하위 계층에서부터 상위 계층까지 쌓아올린 프로토콜의 집합을 말한다. 블루투스 프로토콜 스택은 HCI (host controller interface) 프로토콜을 기준으로 호스트 컨트롤러와

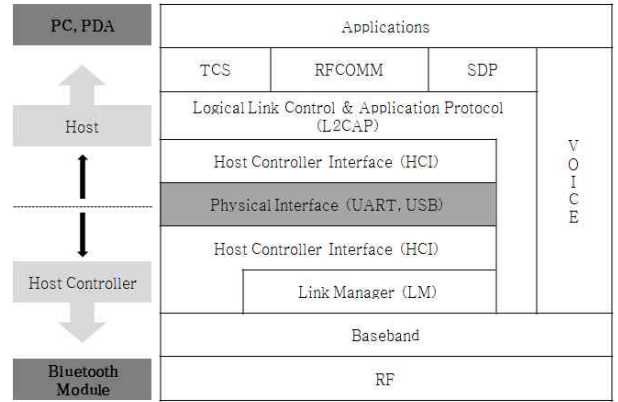


그림 2. 블루투스 프로토콜 스택

호스트로 나뉘게 되며, UART (universal asynchronous receiver/transmitter), USB (universal serial bus), PCMCIA (personal computer memory card international association) 등의 물리적 인터페이스로 연결된다.

호스트 컨트롤러는 블루투스 모듈에 해당하며, 각 모듈간의 물리적 채널을 구성한다. 블루투스 규격에서는 음성 및 오디오 서비스를 위해 동기식 SCO 채널(synchronized connection oriented link)을, 패킷단위의 데이터 전송을 위해 비동기식 ACL 채널(asynchronous connectionless link)을 사용하는 것을 명시하고 있다. 그러나 SCO 채널을 통한 음성전송의 경우 정해진 음성코덱만을 이용해야 하며, 음성의 부호화 과정이 블루투스 모듈상에 구현되어 있다. 따라서 하드웨어에 의한 음성처리가 민감하여 음질에 영향을 미치며, 소프트웨어적 음성처리 구현이 어려운 단점이 있다 [7][8]. 따라서 본 논문에서는 소프트웨어 환경에서 음성전송 시스템을 구현하기 위해 ACL 채널을 이용하여 패킷 단위로 음성신호를 전송한다.

호스트 프로토콜은 블루투스 모듈과 연결되어 블루투스 모듈을 제어하고 어플리케이션을 수행하는 곳으로 보통 PC, PDA 등이 이에 해당한다. 특히 본 논문에서는 G.711 패킷 손실 은닉 알고리즘을 수행하기 위해 application layer에 RTP 프로토콜을 적용하며, RTP 패킷의 header format을 분석하여 패킷 손실 여부를 확인하도록 한다.

### 다. G.711 패킷 손실 은닉 알고리즘

패킷 손실 은닉 알고리즘은 수신된 비트스트림 내의 데이터 손실을 은닉하기 위해 합성된 음성 신호를 만들어내는 것이다. 이상적으로 합성된 신호는 손실된 신호와 동일한 음색과 스펙트럼 특성들을 가지며, 또한 부자연스러운 인공효과를 발생시키지 않도록 한다.

그림 3은 ITU-T G.711 Appendix I에 규정된 패킷 손실 은닉 알고리즘을 보여준다. G.711 패킷 손실 은닉 알고리즘은 히스토리 버퍼(history buffer)에 저장되어 있는 48.75 ms의 이전 음성신호로부터 피치(pitch) 구간을 계산한 후, 계산된 피치 구간을 반복해서 삽입하는 피치구간삽입방식이다. 패킷 손실 은닉을 위해 복호화된 출력 신호는 히스토리 버퍼에 저장되며, 히스토리 버퍼는 손실된 구간에서의 피치를 계산하고 손실구간 동안의 신호를 추출하는데 사용된다 [9].

패킷 손실이 발생하게 되면, 히스토리 버퍼의 내용은 피치 버퍼로 복사되어 피치 버퍼의 음성으로부터 피치를 계산하는 데 사용된다. 이

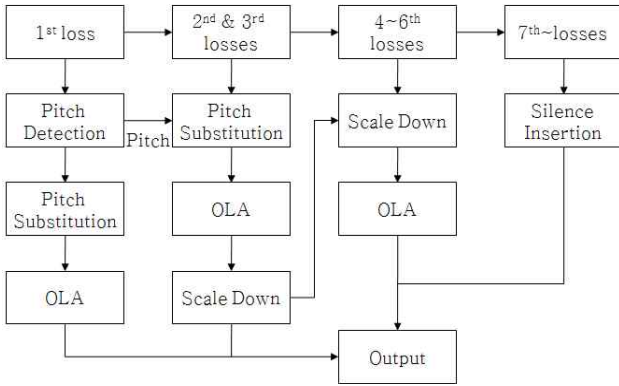


그림 3. G.711 패킷 손실 은닉 알고리즘

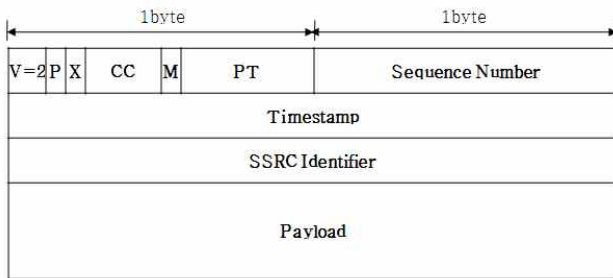


그림 4. RTP 패킷 구조

때 부드럽게 음성을 재생하기 위해 계산된 피치의 1/4 지점을 삼각 윈도우를 이용하여 overlap and add (OLA) 방법을 사용하여 피치 구간을 반복 삽입한다.

### 3. 음성전송 시스템 구현

블루투스 피코넷 기반의 음성전송 시스템을 구현하기 위해 G.711 공개 C-소스를 이용하여 부호화 및 복호화 API (application program interface) 를 구현하였으며, G.711 패킷 손실 은닉 알고리즘 또한 표준안에 의거하여 정상 프레임은 히스토리 버퍼에 저장하고, 손실 프레임에 대해서는 패킷 손실 은닉을 수행하는 API를 구현하였다. 또한 IETF 표준에서 제공하는 RTP 프로토콜의 패킷 구조를 기반으로 RTP 패킷화 과정 및 패킷 분석 과정을 구현하였으며, 마지막으로 구현된 API들과 RTP 패킷화 및 패킷분석 소스를 블루투스 네트워크에 추가함으로써 최종적인 음성전송 시스템을 구현하였다.

#### 가. 블루투스 피코넷 구현

블루투스 피코넷은 Microsoft 사에서 제공하는 Windows XP Platform SDK 기반의 Bluetooth API [10]를 이용하여 구현되었다. 제공된 API는 소켓을 통해 블루투스 RFCOMM 프로토콜과 논리적으로 연결되어 있다. 즉, G.711 공개 C-소스를 이용하여 부호화된 음성 비트스트림에 RTP 패킷화 과정을 수행하여 생성된 RTP 패킷을 RFCOMM payload에 저장하고 이를 블루투스 하위 프로토콜로 전송한다. 이때 음성의 한프레임은 8 kHz 표본화율에서 10 ms 단위로 설정하였다. 그림 4는 RFCOMM 프로토콜로 전송된 RTP 패킷 구조를

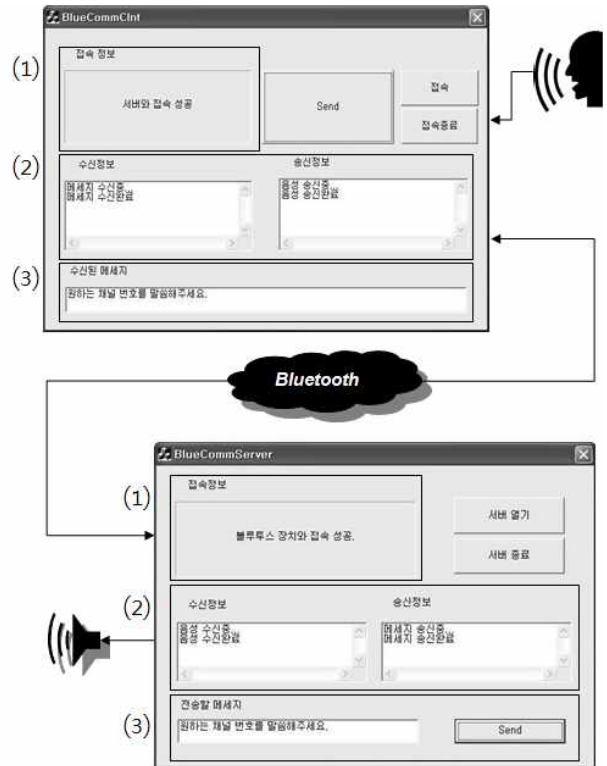


그림 5. 블루투스 피코넷 기반 음성전송 시스템 시뮬레이션 환경

보여준다. 블루투스 baseband에서는 두개의 블루투스 동글간에 생성된 물리적 링크 즉, ACL 채널을 통해 음성을 전송한다.

블루투스 피코넷 구현을 위한 하드웨어 플랫폼으로는 CSR사의 칩셋이 탑재된 USB 블루투스 동글을 사용하였다. 사용된 adapter는 블루투스 2.0 및 1.1을 지원하며 최대 2.1 Mbps의 전송속도를 지원한다. 또한 100 mW의 최대 출력파워를 가지며 최대 100 m의 전송영역을 지원한다.

#### 나. G.711 패킷 손실 은닉 알고리즘 구현

음성전송 시스템의 수신단에서는 패킷 손실로 인한 음질 저하를 줄이기 위해 G.711 패킷 손실 은닉 알고리즘을 구현하였다. 즉, 현재 수신된 데이터의 RTP 패킷을 분석하여, 현재 패킷의 패킷 순서가 이전 패킷의 패킷 순서와 연속적일 경우 정상 프레임으로 판정하고 정상적인 G.711 복호화를 수행한다. 반면, 비 연속적일 경우 손실 프레임으로 판정하여 G.711 패킷 손실 은닉 알고리즘을 수행하고 이를 통해 손실된 음성프레임을 복원할 수 있게 된다.

#### 다. 음성전송 시스템

그림 5는 음성전송 시스템의 시뮬레이션 환경을 보여준다. TV 수신기에 해당하는 수신단과 음성 리모트 컨트롤러에 해당하는 송신단의 물리적 링크 형성여부가 그림 5의 (1)과 같이 송신단과 수신단의 접속정보창에 표시된다. 두 단말간의 접속이 완료되면 송신단에서 수신단으로 음성전송이 가능하며, 수신단에서는 전송된 음성을 실시간으로 재생한다. 데이터 전송과 관련한 송수신 정보는 그림5의 (2)와 같은 상태표시창을 통해 표시된다. 한편, 수신단에서 송신단으로 필요한 메시

표 1. 구현된 G.711 패킷 손실 은닉 알고리즘의 음질 평가 (MOS).

손실률(%)	적용 전	적용 후	증가율(%)
2	3.322	3.676	10.7
5	3.165	3.570	12.8
7	2.851	3.260	14.3
10	2.626	3.179	21.1

지를 전송할 수 있다. 그림 5의 (3)은 메시지 전송과 관련된 인터페이스를 보여 준다. 수신단에서 송신단으로 메시지를 전송하기 위해 수신단에서는 전송할 메시지를 입력할 수 있는 메시지전송창이 있으며, ACL 채널을 통해 해당 메시지를 송신단에 전송한다. 그리고 송신단에서는 수신된 메시지를 출력하게 되며, 이는 IPTV 환경에서 메시지를 통한 사용자 인터페이스에 활용될 수 있다.

#### 4. 성능평가

성능 평가를 위해 G.711 코덱을 A-law로 설정하고, 패킷 손실률은 ITU-T 표준 G.191에 정의된 Gilbert-Elliot [11] 모델을 이용하여 2, 5, 7, 10%로 각각 생성하였다. 실험음원은 NTT-AT [12] 데이터베이스에 있는 각 8초 길이에 해당하는 8 kHz의 한국어 여성음성, 영어 여성음원, 한국어 남성음성, 영어 남성음원 4 샘플을 사용하였다. 즉, 송신단에서는 Gilbert-Elliot 모델을 통해 생성된 패킷 손실 패턴을 이용하여 정상패킷의 경우 수신단으로 데이터를 전송하며, 손실프레임의 경우 해당 패킷을 전송하지 않는다. 수신단에서는 10 ms 단위로 수신된 패킷을 분석하여 패킷 손실이 발생하였을 경우 패킷 손실 은닉 과정을 수행하였다.

음질 평가는 ITU-T의 PESQ (Perceptual Evaluation of Speech Quality) [13]를 사용하였고, G.711 패킷 손실 은닉 알고리즘을 적용하기 전과 후의 음질을 비교하였다. 표 1은 각 패킷 손실률에 대한 PESQ 음질 측정 평균치를 보여준다. 표에서 보는 바와 같이, G.711 패킷 손실 은닉 알고리즘은 이를 사용하지 않는 경우에 비해 평균적으로 약 14.7%로 음질을 개선함을 확인할 수 있었다.

#### 5. 결론

본 논문에서는 블루투스 환경 하에서의 음성인식을 위해 필요한 음성전송 시스템을 실시간으로 구현하였다. 효율적인 음성전송을 위해 G.711을 기본 코덱으로 사용하였으며, 블루투스 통신 환경에서의 음성 전송 시 발생하는 패킷손실에 따른 음성품질 저하를 줄이기 위해 G.711 은닉 알고리즘을 음성전송 시스템에 적용하였다. 특히 G.711 패킷 손실 은닉 알고리즘 수행을 위해 블루투스 프로토콜 스택 상위단에 RTP 프로토콜을 적용하여 RTP 패킷의 header format을 분석하여 패킷 손실 여부를 확인하였으며, 패킷 손실 발생 시 패킷 손실 은닉을 통해 음성 품질을 개선하였다. 구현한 음성전송 시스템의 성능을 평가한 결과, G.711 패킷 손실 은닉 알고리즘 적용함으로써 2~10%의 패킷 손실률을 갖는 환경에서 14.7%에 상당한 음질개선을 이룬 것을 확인할 수 있었다.

#### 감사의 글

본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (IITA-2008-(C1090-0801-0017))

#### 참고문헌

- [1] 박성수, "무선 홈 네트워킹 기술 - 블루투스 및 관련기술," *대한전자공학회 텔레콤*, 제 16권, 제 2호, pp. 47-62, 2000년 12월.
- [2] 이정환, 한역수, "블루투스를 이용한 이동통신 서비스의 동향과 전망," *IITA 주간기술동향*, 제1159호, 2004년 8월.
- [3] ITU-T Recommendation G.711, *Pulse code modulation (PCM) of voice frequencies*, Nov. 1998.
- [4] ITU-T Recommendation G.711 Appendix I: *A high quality low-complexity algorithm for packet loss concealment with G.711*, Sept. 1999.
- [5] RFC Recommendation 1889 RTP, *A Transport Protocol for Real-Time Application*, Jan. 1996.
- [6] 최진섭, 정중수, 정태운, 이용수, "블루투스 프로토콜 설계 및 구현을 통한 교육용 소프트웨어 개발," *한국정보과학회 2002년도 가을학술발표논문집* 제29권, 제2호, pp. 196-198, 2002년 10월.
- [7] 양영배, *VoIP 통신을 위한 블루투스 피코넷에서의 다중 태널 검출 및 LAN Access 시스템 구현*, 석사학위 논문, 제주대학교, 2003년
- [8] 김명규, 양일식, 손지연, 박준석, "블루투스 스테레오 오디오 공유 시스템의 설계 및 구현," *한국정보기술학회논문지*, 제3권, 제1호, pp. 7-11, 2005년 3월.
- [9] 한승호, 김진술, 이현우, 류원, 한민수, "음성 특성을 이용한 G.711 패킷 손실 은닉 알고리즘의 성능개선," *말소리*, no. 57, pp. 175-189, 2006년 3월.
- [10] <http://msdn2.microsoft.com/en-us/library/aa362932.aspx>
- [11] NTT-AT, *Multi-lingual speech database for tele-phonometry*, 1994.
- [12] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*, Feb. 2001.
- [13] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*, Feb. 2001.