特輯

技術解說

音場通信을 위한 음향신호처리 기술

A Review on the Acoustic Signal Processing for Sound Field Communication

강 성 훈
(Seonghoon Kang)
(한국전자통신연구소 신호처리연구실)

Ⅰ. 서 론

음성통신의 최종 목표는 멀리 떨어져 있는 송,수화자가 마치 같은 실내에 있는 것 처럼 자연스러운 환경에서 서로 회화통신을 할 수 있는 것일 것이다. 서로 떨어져 있는 회의실 상호간을 쌍방에 적절한 전기통신 수단으로 연결하여 서로 의사 등을 교환하는 회의 목적을 달성하는 형태를 遠隔會議(Tele-conference)라고 한다. 원격회의에는 음성통신만에 의한 전화회의(Audio-conference)나 팩시밀리 통신 또는 데이타 통신을 수반하는 것, 더우기, 정지화면(Still Video) 또는 동화면(Motion Video)을 수반하는 것 등 여러가지 형태가 있다. 쌍방의 통화자 전원이 마치 같은 실내에 있는 것과 같은 현장감있는 회의를 진행할 수 있는 시스템으로써 회의실 상호간을 영상과 음성통신으로 접속하는 비데오 회의 시스템(Video Conference system)이 있다.

최근에 사회활동의 활성화에 따라, 효율적인 기업활동을 수행하고, 회의 소집 시간 및 여행 경비를 줄이기 위하여 화상회의나 음성회의 등의 통신 서비스가 주목되기 시작하였고, 기업전략의 하나로서 원격회의 시스템을 구축하여 활용되고 있다[1]. 이러한 회의 시스템은 마이크로폰과 스피커를 이용한 핸드프리 통화를 실현하는 것이 필수적이고, 이를 위한 기술로서는 음성 스위치와 에코 캔슬러 방식이 이용되고 있다. 또, 원격회의 시스템으로 부터 한걸음 더 나아가서, 하나의 테이블에 마치 참석자가 앉아 있는 것처럼 송화자가 누구인가를 쉽게 확인할 수 있고, 한자리에 모여서 하는 통상의 회의에 가까운 통신 시스템을 실현하기 위한 音場通信 시스템이 연구되고 있다. 이러한 시스템을 구현하기 위해서는 傳送系 技術, 映像系 技術 및 音響系 技術등 복합적인 기술이외에 心理音響 및 人間行動學 등의 연구가 필요하다.

본고에서는 통신망에 있어서 신호처리기술 분야 및 공간음향에 대한 재고와 원격회의 시스템의 최종 목표인 음장통신의 구현에 있어서 필수적인 음향기술 Layer를 재정리하고, 이것을 기본으로 각종 음성회의 시스템의 특징 및 소요 음향신호처리 기술에 대하여 고찰한다.

Ⅱ. 음향기술 레이어

통신망을 구성하는 주된 기능과 기술을 분류하면 그림 1과 같다. 신호처리 기술에서 새롭게 등장한 분야는 話者檢出, 話者認識, 假象 共有空間 生成등을 위한 상태정보를 처리하는 기술로써, 미디어 신호만의 정보패스와 별도의 채널 또는 같은 채널에 제어정보를 주고 받음으로써 자연성이 풍부하고 편리성이 높은 통신을 지향하는 것이다[2].

기 능 / 기 술	통신 미디어 처리	신 호 변 환	신 호 전 달 처 리
정보처리	음성, 화상 및 문자 인식합성, 음성 및 화상 축적 등		
상태처리	음상정위, 음장공간재생, 화자위치 검출, 화자인식, 화자상태인식 등		
신호처리 (협의의 의미)	A/D변환, 각종 부호/복호화, 대역압축 등 Filter(고정형, 적응형)		
전달처리	파형전달(타이밍추출, 위상/주파수동기, 변복조, 전송부호) 다중화(주파수 분할/시분할, ATM), 원격전송품질측정 등		

그림 1. 통신 신호처리 기술의 종류

음의 정보는 크게 나누어 時間情報(Temporal Information)와 空間情報(Spatial Information)로 나눌 수 있다[1]. 그림2는 음향기술 레이어를 나타내고 있다. 레이어 1은 시간정보만을 전송하는 음성통신의 기본이 되는 기술, 레이어2와 Layer 3은 시간정보 이외에 공간정보까지 전송하기 위한 상태처리 기술이다. 레이어2는 발성자 또는 송화자(음원)가 어디에 있는가 하는 방향성 지각을 위한 처리기술이고, 레이어 3은 하나의 가상 공유공간(Virtual reality)을 생성 또는 재생하기 위한 배경음의 처리 기술이다. 이들 音像定立(Localization)나 배경음 처리는 이 공간구조인 상태정보의 한가지 처리기술이다. 음의 통신정보로서 필요한 것은 단순히 레이어1뿐만이 아니고, 기술 레이어가 높아짐에 따라 통신 정보로부터 생성하는 것 보다 청취측에서의 일체감을 생성하는 상태처리 기술이 중요한 과제가 될 것이다.

Ⅲ. 空間音響(Spatial Hearing)

음향기술의 진보로 인류의 꿈이었던 시간과 장소에 관계없이 음악을 즐길수 있게 된것은 약 100여년이 지났다. 그 사이의 진보에서 음악을 신호로서 전송하거나 기록하는 것은 비약적으로 진보한데 비해, 공간음향 신호의 전송이나 기록은 처음 시도된 것은 오래되었으나 별로 진전되지 않았다. 그 커다란 이유는 한점의 음향신호의 전송으로는 空間音響情報를

음향기술 Layer		송 화 계		전 송 계	수 화 계	
공간 정보 처리	Layer 3 배경음 처 리	잔 향 제 어	잔향제어 처리 · 송화계의 실내 잔향특성 · 명료성	· 심리음향 · 음향 공유공간 생성 처리	확 산 감 제 어	확산감 처리(배경음) · 수화계의 실내 잔향감 · 음원 일체화(복수 화자시)
	Layer 2 정 위 처 리	음원 위치 추정	수음처리(방향성) · 음원위치의 명확한 · 다수 음원의 분리화	· 음원위치의 부호화 · 음상정위의 부호화 · 실내공간 음향 특성.	음상정위 제 어	음상정위처리(방향감) · 음상생성 위치의 명확함 · 다수 음원 위치의 분리 · 원근감 제어 · 명료도
시간 정보 처리	Layer 1 통신로 처 리	· 음부호 · 음향전기 · 지향성 제어 · 잡음제거	수음처리(S/W) · 입력 다이나믹 레인지 · 실내잡음	· 전달 용량 · 전송 대역 · 전달지연 · 에코 · 하울링	· 음부호 · 전기음향	음의 박력 · 음량, 온밀, S/N · 주파수 특성 · 스페프럼, 대역, 비직선 · 왜곡과도 특성 · 통화 용담성

그림 2. 음향기술 레이어 및 각 요소기술

전송할 수 없기 때문이다.

심리음향의 진보로 사람의 청각 지각 현상이 밝혀지고, 兩耳效果(Binaural Effect)가 공간음향지각에 아주 중요하다는 점을 알았고, 그 응용예로서 1881년 파리 만국 박람회에서는 Binaural 전화가 출현하여 화제가 되었다. Binaural 전화 이후에 고감도 마이크로폰의 등장에 의해 마네킹의 귀에 마이크로폰을 설치한 더미헤드(Dummy Head)가 탄생되어, 두 귀에서의 음압 파형을 충실하게 수음하고, 헤드폰으로 청취함으로써 충실한 現場感을 재생할 수 있었다. 1931년에는 3개의 마이크로폰과 3개의 스피커를 이용한 음장 재현 실험이 미국의 Bell 연구소의 Fletcher에 의해 이루어졌다. 현재 일반적으로 사용되고 있는 2채널 스테레오는 이러한 방식중에서는 가장 단순한 2점 사이의 음장재생이다.

Binaural 전화와 비교하여 '더미헤드 녹음-헤드폰 재생' 방식의 현장감이 도약적으로 개선된 이유는 무엇인가? 사람이 자연상태에서 듣고 있는 신호자체를 수음하고, 그대로 재생하는 것이 충실한 현장감 재생이라는 것은 사실이다. 그러면 우리들은 어떻게 하여 현장감의 기본이 되는 음상이나 음색을 지각하는 것일까? 최근에는 그 이유가 명백해졌다.

청취자의 두 귀에 도달하는 음향신호는, 음원으로부터 나온 음이 공간에서 반사된 반사음을 포함하여 공간을 전반하는 동안에, 또 음원과 청취자의 위치관계에 의해 결정되는 머리와 컷바퀴의 영향에 따라 변한다. 이와같이 음이 전반되는 공간과 청취자에 대한 음원방향이 고유적으로 결정되는 음향신호의 변화를 만들어 내는 것을 音響傳達函數라고 한다[4][5]. 따라서 두 귀에 입사하는 음향신호의 음압은, 음원 자체의 음향신호의 음압, 음이 전반되는 공간의 공간전달함수, 그리고 음원과 청취자의 위치관계로 결정되는 청취자 고유의 머리 전달함수(Head-related Transfer Function)의 3가지 요인으로 결정된다[4]. 이 공간 전달함수 및 머리 전달함수를 전송하는 기술이 그림 2의 Layer2와 Layer3에 해당된다.

Ⅳ. 음성회의 시스템

음성회의 시스템은 마이크로폰과 스피커를 이용한 핸드-프리 통화를 실현하는 것이 필수적이고, 이를 위한 기술로서는, 음성 스위치와 에코 캔슬러 방식이 이용되고 있다. 이들을 이용한 구체적인 시스템예에 대해서 기술한다[6].

음성 스위치를 이용한 음성회의 시스템의 구성예를 그림3에 나타낸다. 다수의 회의 참가자를 위하여 복수의 마이크로폰 및 스피커가 이용되고 있으며, 그 결과 하울링이 생기기 쉽고 에코도 들리기 쉽다. 이것을 전지역에 걸쳐서 똑같이 억압하려고 하면, 어두.어미의 절단이 커진다. 이에 대처하기 위해서 대역 분할형 음성 스위치가 사용되고 있다. 대역 분할

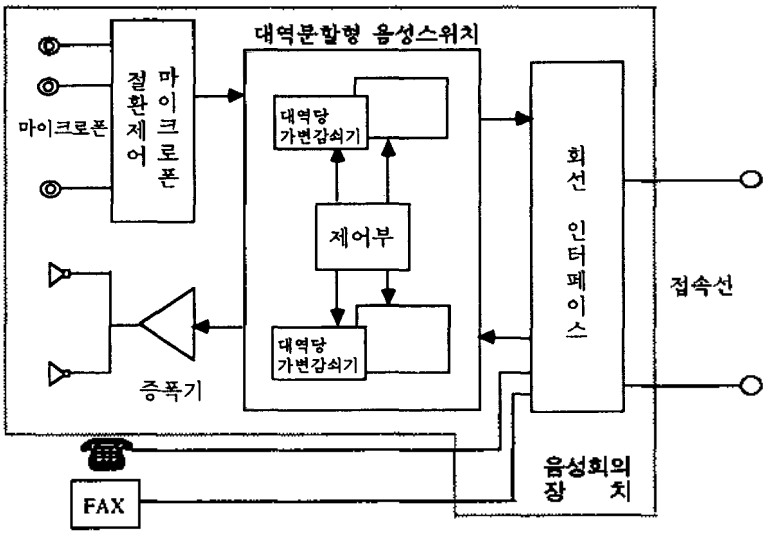

그림 3. 음성회의 시스템의 구성

형 음성 스위치는, 마이크로폰의 유효 갯수를 하나 또는 두개로 필요 최소한으로 제한하기 위한 마이크로폰 교환제어방식을 채용하고 있고, 하울링과 에코가 발생하기 쉬운 주파수 부근에만 손실을 주어 S/N비를 확보하고 동시에 어두.어미의 절단을 적게 하고 있다.

음성 스위치 대신에 에코 캔슬러를 이용한 회의 시스템도 개발되고 있다. 에코 캔슬러는 처리가 복잡하기 때문에 값이 비싸다는 결점이 있지만, 어두.어미의 절단이 생기지 않기 때문에 상호 통화시의 자연스러운 핸드-프리 통화가 가능하다. 이 시스템은 그림3에서 대역 분할형 음성 스위치를 에코 캔슬러로 바꿔 놓은 구성이다. 여러개의 마이크로폰과 스피커의 교환을 제어하기 위하여 여러개의 반향로를 가지게 되고, 단순한 에코 캔슬러의 채용으로는 안정된 성능을 얻을수 없기 때문에 복수 반향로 에코 캔슬러가 개발되고 있다. 복수 반향로 에코 캔슬러는 2개의 마이크로폰이 ON인 경우에 임펄스 응답이 각각의 마이크로폰 마다의 임펄스 응답의 선형 결합으로 표현되는 것을 이용하여, 각 마이크로폰 마다의 임펄스 응답을 기억하여 보다 빨리 적응을 실현하는 것이다.

이 시스템은 그림2의 음향기술 레이어에서 단순히 시간 정보만을 전달하는 Layer1의 통신로 처리에 해당된다.

V. 현장감 통신 실현방식

4장에서 기술한바와 같이 음성회의에 있어서 자연스러운 핸드-프리 통화의 실현이 서비스의 제1단계라고 하면, 그 다음 단계는 현장감 있는 회의통신 시스템을 실현하는 것이다. 또한 대향음성회의(point to point teleconferencing) 시스템을 복수의 통화자가 다른 장소에서 회의에 참가하는 다지점 음성 통신(multiconferencing)에 적용할 경우에는 다지점으로부터 복수인의 음성이 단지 하나의 수화기나 스피커로 재생되기 때문에, 다른 장소에서 발신한 송화자가 마치 같은 장소에서 발신한 것 처럼 들린다. 따라서 어느 지점의 누가 발신하였는지가 불명확하고, 상대의 음성에 특징이 있는 사람 이외는 분간하여 알아들을 수 없다. 이러한 경우에 송화자를 명확히 알 수 있도록 하기 위해서는 회의 참가자가 이름이나 소속을 먼저 이야기하고 나서 발신을 해야 하므로, 보통 회의의 분위기와는 다를뿐만 아니라 회의를 원만하게 진행할 수 없게 된다. 또 지점 수가 증가할수록 음성회의에 있어서 지점의 식별은 한층 더 어려워진다.

또한 분위기의 전송, 즉 어떻게 하면 현장감이 풍부한 음성 회의를 실현할수 있을까 하는 것이다. Yoshida는 "立體音의 본질은 무엇인가?"를 파악하기 위하여 품질의 인자를 분석한 결과, 方向性 인자와 殘響性 인자(배경인자)의 두가지가 음의 입체감에 크게 기여하고 있다고 하였고, 또 음을 충실하게 전송하고 재생하는 시간정보와 방향성.잔향감을 재생하는 공간정보의 전송이 필요하다는 점을 지적하였다[7].

음향신호 처리기술을 통신에 응용하는 목적은 송화자의 목소리와 송화자를 둘러싼 공간정보를 가능한한 충실하게 전송하는 기술을 실현하는 것이다. 따라서 대향 통신에서 상대측의 통화자가 착석하고 있는 테이블 위치에 음상을 정위(Localization : 음성이 들리는 위치의 제어) 시켜 통화자의 위치 식별이 가능한 시스템이나, 스테레오 회의 시스템 등이 연구되고 있다.

여기에서는 현재 연구되고 있는 음상제어 방식, Stereo재생 방식, 3차원 음장재생 방식에 관한 특징을 기술하고, 그 장단점을 비교·고찰해 본다.

5.1 음상제어 방식

音像制御(Sound Image Control) 방식은 청취측에서 복수의 발생 음원을 준비하여 송화자의 음성과 그 복수의 발생 음원과 대응시켜 송화자 식별신호에 의해 音像定位를 제어하는 방식이다[8]. 이 방식은 다지점 통신에 있어서 통화 상대의 음상 위치를 미리 통화자의 희망 위치에 설정하여, 그 설정된 음상위치로 부터 상대 통화자를 식별한다. 이 기술은 앞으로 복수지점을 동시에 접속할때 가상 공유공간 생성에 있어서 중요한 과제가 된다. 이 송화자 식별 신호는 2장에서 기술한 상태정보 신호에 포함된다. 이 방식의 개요는 다음과 같다.

(1)회의 참가자의 지역 번호를 미리 설정하여 두고, 통화자가 송화자가 된 지점으로 부터 그외의 다지점의 청취자에 대해, 지역 식별 번호를 송신하는 지역 식별 번호를 부호화한다.

(2)다지점 통화를 처리하는 MCU(Multipoint

Conferencing Unit)에서는 지금 어느 지역이 송화자인가를 지역 식별 번호를 처리하여 수화자가 복수 지역 식별번호를 전달하는 복수 지역 식별을 처리한다.

(3)청취측에서는 MCU로 부터의 복수 송화자의 지역 식별 번호를 수신하여, 그 부호에 따라 2개 이상의 스피커로 음상을 생성하여, 생성된 음상위치로부터 송신지역을 식별하는 음상생성 처리를 한다.

이 방식을 통신회의에 적용할 경우의 특징은 다음과 같다.

- 다지점 통신 및 대향 스테레오 통신에도 쉽게 확장할 수 있다.
- 수신정보 신호는 모노(Monaural)이어도 좋고, 공간 구조 파라미터(상대정보신호)를 별도로 전달하여 머리 움직 변화를 정보를 얻을 수 있기 때문에 에코제어 설계가 용이하다.

이 방식은 음향기술 Layer2의 정위처리중 음원 위치처리 기술에 해당된다.

5.2 Stereo 재생 방식

Stereo Audio-conference System은 Hi-Fi Audio System의 개념을 그대로 통신에 적용한 것이다. 즉, 2개의 마이크로폰으로 음을 픽업하여 두개의 신호를 전송하고, 상대방의 회의실에서 2개의 스피커로 재생하는 것이다[9]. 이렇게 함으로써 모노로 전송한 경우보다 명료도, S/N비 및 정위감이 향상된다.

그러나, 2채널 스테레오 재생에서는 재생 스피커의 중앙 전방 이외에서는 충분한 효과를 얻을 수 없다고 하는 결점이 있다. 이에 대해 넓은 청취 범위에 걸쳐 입체감을 얻을 수 있는 스테레오 재생 방식이 제안되고 있다[10][11][12]. 시스템 구성 방식을 그림4에 나타낸

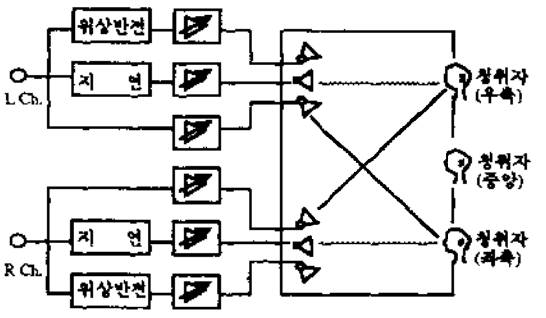

그림 4. 음성회의용 스테레오 재생계의 구성

다. 이 방식은 청각 특성과 스피커의 지향 특성을 이용하고 있는 점이 특징이다. 재생계는 정면 방향에서 지향 특성을 가진 종래의 스피커 이외에, 위상을 지연시킨 음성 신호를 방사하는 스피커와 위상을 반전시킨 음성신호를 방사하는 3개의 스피커로 구성된다. 주관평가 시험에 의하면 이 방식의 효과는 종래의 재생계와 비교하면 스피커의 중앙 전방의 청취 위치만이 아니라 좌우의 넓은 위치에 있는 청취자에 대해서도 양호한 정위감을 얻을 수 있다.

이러한 스테레오 통신의 음상재생 기술은 어디까지나 대향통신에 유효하고, 다지점 통신에는 적용하기 어렵다. 장래에는 보통 면담식 회의와 같이 복수와 원격지 통화자가 마치 하나의 테이블앞에 앉아 있는 것처럼, 음상정위에 의한 현장감 넘치는 다지점 통신이 기대된다. 이 방식은 그림2의 음향기술 Layer2에서 정위처리에 해당된다.

5.3 3차원 음장재생 방식

3차원 음장재생 방식은 일반적으로 Transaural이라고 명명되고, Atal이 제안한 것으로서 현재 많은 연구가 발표되고 있다[13][14][15]. 그림5에 이 방식의 구성도를 나타낸다. 이 방식은 3장의 공간음향 지각에서 기술한바와 같이, 음원의 장소에서 HATS(Head and Torso Simulator : 머리 몸통 모사기)로 수음한 음과 같은 음을 청취자의 두귀에 그대로 전송하는 것이고, 청취측에서는 두개의 스피커로부터 방사된 음이 청취자의 두 귀에 도달하는 음파를 원음장과 일치시키는 방식이며, 앞으로 회의통신에 있어서 기대되는 이른바 Binaural Communication System 방식이다.

이 방식을 통신 회의에 적용할 경우의 특징은 다음과 같다.

- 좌우 두 채널의 신호로를 사용하기 때문에 통상의 2채널 스테레오에 적용이 가능하다.
- 통화 상대의 음상 및 배경음 처리까지도 충실하게 재현할 수 있게 때문에 가장 자연스러운 통신이 가능하다.
- 확성계에서는 지연시간이 큰 적응 필터와 역필터의 Convolution 연산처리가 필요하다.

이 방식을 적용하면 음상정위 처리와 배경음 처리가 동시에 이루어지기 때문에 음향기술의 레이어2와

레이어3이 동시에 해결된다.

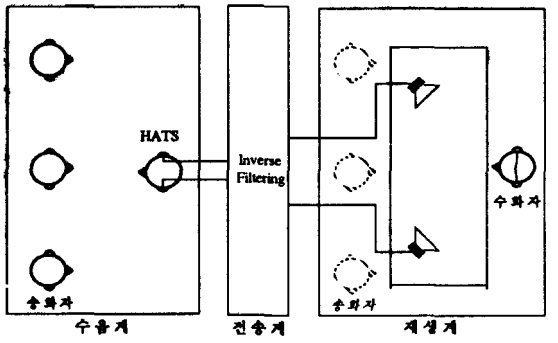

그림 5. Transaural System의 개요

Ⅵ. 음향신호처리 요소 기술

표1.에는 음성회의 통신 시스템을 실현하기 위한 각 요소와 음향 에코캔슬러에서 규정해야할 항목과의 관련을 나타낸다. 이들의 요소는 규정해야 할 항목과 서로 관련되어 있고, 임의적으로 정해서는 안된다. 따라서 하나의 규정항목에 대해서는 복수의 요소로 부터 결정하는 것이 중요하다[16].

표1. 음성회의 시스템을 실현하기 위한 외부조건과 규정해야할 항목

외부조건	규정해야할 항목
-심리적인 음향 요소	
· 음향레벨(송화, 수화)	-소요 주파수 대역
· 반향 에코 손실	
-실내 음향 요소	-음향에코 소거량
· 잔향시간과 임펄스 응답 길이	
· 실내 잡음	-임펄스 응답 길이
-심리적 전기 요소	
· 동시 통화시의 손실량 변화	-하울링 마진
-객관적 전기 요소	
· 전기 음향 변화기의 주파수 특성	-신호 대 잡음비
· 에코 경로의 비선형	
인간 행동학 요소	신호처리 지연
· 동시 통화시에서의 인간의 응답성	
· 화자의 자연동작에 대한 에코 경로 변경	-시험 측정 방법

6.1 음향용 에코 캔슬러

스피커와 마이크로폰에 의한 핸드 프리 통화의 경우에 스피커에서 발생되는 음은 마이크로폰에서 픽업되어 다시 상대측에게 전달되는 에코가 된다. 또 양단말에서 핸드-프리 통화를 하면 음향 결합이 강한 경우에 하울링(명음)이 생기는 경우도 있다. 따라서 에코 캔슬러를 이용하여 이들 현상을 제어하여야 한다.

전화회선용 에코 캔슬러에서는 대상으로한 에코 경로는 하이브리드 트랜스 등의 전기 소자로 구성되어 있기 때문에 30~60ms의 반향 신호를 소거하면 충분하다. 이것에 대해서 음향 결합에 의해 생기는 에코는 스피커와 마이크로폰이 설치된 실내의 잔향시간(수백ms이상)에 의해 영향받는다. 따라서 음향용 에코 캔슬러에서는 200~400ms에 달하는 임펄스 응답을 처리하지 않으면 안되고, 필요한 장치의 규모는 전화회선용 에코 캔슬러에 비하여 수배 이상이 된다. 따라서 연산의 고속화 알고리즘과 IIR형 에코 캔슬러 등의 검토가 한창이다[17][18][19].

6.2 실내의 잔향시간

일반적으로 같은 음향시스템이라도 통화 품질은 실내의 잔향특성에 따라 크게 달라진다. 즉, 잔향시간이 짧은 실내에서는 무향실에서 대화하는 것과 비슷하고, 잔향시간이 긴 회의실에서는 소위 "Rain-Barrel Effect"로 자연성의 회화 상태를 유지하기 어렵다. 회의통신에 최적한 잔향시간을 규정하여 통화 품질을 보증하여야 하며, 음향 에코를 소거할 수 있는 임펄스 응답 시간의 규정도 필요하다. 일반적인 회의통신용 회의실에 최적한 잔향시간에 대해서는

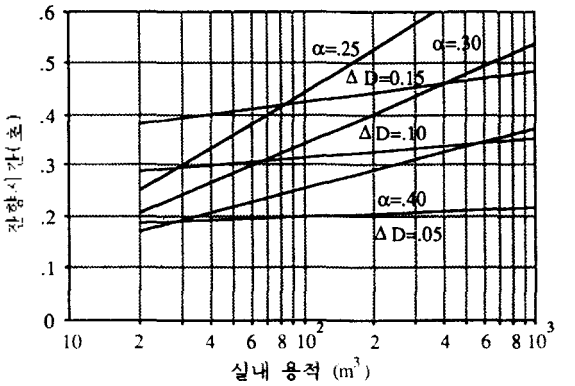

그림 6 리스닝룸의 용적에 따른 최적 잔향시간 :
α : 평균 흡음계수(최적치 : 0.25-0.40)
ΔD : 플래아백 음성에 있어서 명료도의 감쇠치
(최적치 : 0.05-0.15)

아직 충분히 검토되어 있지 않으므로, 현재는 방송용 스튜디오나 리스닝룸에 적합한 잔향시간을 참고로 하고 있다.

Hirata는 일반적인 거실에 있어서 재생음의 울림의 변화와 잔향시간과의 관계로부터 리스닝룸의 최적 잔향시간을 구하고 있다[20]. 이 결과에 의하면 명료도(D : Definition)의 감소분의 최적치가 0.1(+/−0.05)의 범위이고, 거실의 평균 흡음률은 0.25~0.40이므로 바람직한 잔향시간은 0.2~0.4초라고 보고하고 있으며(그림6), 또 CCITT에서도 통신 회의실의 잔향시간은 100~400ms가 바람직하다는 것을 지적하고 있다. 이러한 결과로 부터 잔향시간 400ms이하를 잠정치로 하고 있으며, 잔향시간 400ms 이상에 대해서는 음향 에코 제어장치의 기술적 실현성을 고려한 범위에서 금후의 과제이다.

6.3 마이크로폰의 지향성 제어

마이크로폰은 공간에서 동작하는 센서의 일종으로서 다른 센서에 비해 넓은 주파수 대역에 걸쳐 지향성을 비교적 자유롭게 제어할 수 있는 것이 특징이 있다. 녹음이나 방송을 위한 수음에서는 하울링 방지와 소음억제의 수단으로서의 지향성 마이크로폰을 사용하는 것이 일반적이다.

시판되고 있는 확성전화기에는 단순한 전지향성의 일렉트리트 콘덴서 마이크로폰을 조립한 것이 압도적으로 많다. 송화위치의 제한, 실장의 어려움, 가격이 비싸다 등의 이유로 지향성 마이크로폰을 그다지 사용하고 있지 않다. 음성 스위치 회로의 경우와 같이 이러한 불리함을 능가하는 이점이 인정되면 지향성 마이크로폰은 널리 이용되게 될 것이다. 최근 수요가 현저하게 증가된 음성회의 시스템용으로서는 비교적 대형이고 고가인 마이크로폰 시스템을 사용하고 있으며, 이러한 시스템에 디지탈 신호처리를 도입하여 고도의 환경적응 동작을 실시하는 시스템의 연구와 상품화가 이루어지고 있다.

6.4 기타

수화계에서는 스피커계의 신호처리 기술도 동시에 연구되고 있다. 예를 들면, 스피커 어레이를 이용하여 원하는 음장을 실현하기 위한 신호처리 이론과 스테레오 재생에 악영향을 미치는 실내음장 특성을 적응적으로 보정하는 기술등이 있다. 능동 잡음제어 기술(소음을 제어된 스피커의 음으로 상쇄시킨다)등도 소음하에서의 음장통신를 고품질화하기 위한 연구과제이다.

또한 핸드셋 전화기가 한쪽 귀로 청취(Monaural Listening or One-ear Listening)하는 반면, 회의 통신에서는 두 귀로 청취(Binaural Listening or Two-ears Listening)하기 때문에 수화음량 및 명료도를 평가할때 양이효과 (Binaural Effect)와 머리의 회절효과(Head Diffraction Effect)등 심리적인 면도 고려하여야 한다[16][21]. 따라서 HATS의 국제표준화도 동시에 연구되어야 하며, 이와같은 연구는 CCITT SGXII의 Question 2(Hand-Free Telephony)에서 주요 연구과제로 삼아 1989년부터 연구되어 1992년 Head and Torso Simulator for Telephonometry(Rec.P.58)로써 권고되었다.

화상회의에서는 같은 음향계라도 송화자의 얼굴을 음성과 동시에 전송하면, Audio/Video상승효과에 의해 명료도가 향상된다는 흥미있는 연구 보고도 있다[22].

Ⅶ. 결 론

통신의 신호처리에는 종래의 정보처리, 신호처리, 전달처리의 3가지 이외에 앞으로는 보다 자연성이 풍부하고 편리하게 정보를 전달하기 위한 상태처리 기술의 중요성, 그리고 음성회의 시스템에 있어서 신호처리 기술의 전개와 장래의 새로운 방향과 음향기술에 관한 기술 레이어에 관하여 총괄하였다.

종래의 대역.다이나믹 레인지의 확대 기술 개발을 중심으로한 시간정보 전송에서 공유 가상공간을 생성하기 위한 음상정위 처리 기술, 배경음 처리 기술 등을 대상으로한 공간 정보의 전송이 중요한 기술이라는 점을 지적함과 동시에 음의 공간정보 전송이 상태처리 기술의 일례에 대응한다는 것을 기술하였다. 또한 가상 공유공간을 생성하는 것과 같은 자연감 넘치는 통신의 실현 등 과제가 상태처리 기술이라는 점을 음향기술 관점에서 재정리하고, 이를 기본으로 각종회의 시스템 등의 특징에 대하여 기술하였다.

마지막으로 통신 미디어 서비스가 Visual(V), Intelligent(I), Personal(P)이라고 하는 VI&P서비스

로 바뀌어 가고 있다. 음의 통신 세계도 종래의 핸드셋 통화 또는 확성 통화의 음의 통신 패스 시대로부터 보다 자연성이 높은 서비스 시대가 도래해야 한다고 생각된다. 통신에 응용을 목적으로한 음향, 음성신호처리 연구가 현재 이상으로 활발히 진행되기를 기대한다.

참 고 문 헌

1. 한국전기통신공사, "영상회의 시스템 도입," 전기통신경영과 기술, 9.10월호, 79-83(1989).
2. 中島, "テレビ會議システム," テレビジョン學會誌, 482-489(1982).
3. 島田, "通信における信號處理技術の 展開," 日本音響學會誌, 47卷 7號, 491-497(1991).
4. Y. Ando, *Concert Hall Acoustics*, Springer-Verlag (1985).
5. 강성훈, 강경옥, 장대영, "Head and Torso Simulator의 음향연구에의 활용," 한국음향학회 학술논문 발표회 논문집, 81-94(1991.11).
6. 及川, "帶域分割型スイッチを用いた音聲會議裝置," 研究實用化報告 37, 183-190(1988).
7. 吉田, "立體音の本質はなにか," 日本音響學會誌, 16卷, 249-257(1960).
8. 島田, "多對地音聲會議通信システムの對地識別音像生成方式," 電子情報通信學會論文誌 J70-B, No.9, 1017-1023(1987).
9. R.Botoros, "Stereophonic Speech Teleconferencing,"ICASSP 86, Tokyo, 1321-1324(1986).
10. S. Aoki, "Stereo Reproduction with Good Localization over as Wide Listening Area," J.Audio Soc., 38, 433-438(1990).
11. 清木, "ステレオ受音,再生における音像定位制御," 日本音響學會論文誌表誌, 417-418(S. 63年 10月).
12. S. Aoki, "Stereo Reproduction Method for Teleconferening," International Symposim on Active Control of Sound and Vibration, Tokyo Japan, 497-502(1991).
13. B.S. Atal, "Apparent Sound Source Translator," U.S. Patent 3,236,949(1966, Feb.22).
14. M. Miyoshi, "New Transaural System for Teleconfering Services," International Symposim on Active Control of Sound and Vibration, Tokyo Japan 217-222(1991).
15. Hamada, "マルチィチャネル適應制御アルゴリズムとその音場再生システムの 應用," 日本音響學會論文誌表誌, 2-7-10, 431-432(1990. 3).
16. 강성훈, "고품질 확성 통화계를 위한 음향신호처리 기술 동향," ETRI 주간기술동향 88-19, 1-14(1989).
17. 中律, "音聲通信における音響,音聲信號處理," 日本音響學會誌, 45卷 6號, 466-470(1989).
18. 丸山, "音聲通信におけるエコ制御," 日本音響學會誌, 45卷 6號, 474-479(1989).
19. 島田, "音聲 制御裝置 標準化 動向," 日本音響學會 EA89-1(1989).
20. 平田, "リスニングルームの殘響時間と再生音の響きの質," 日本音響學會誌, 33卷 7號(1977).
21. 강성훈, "핸드-프리 전화의 음성품질 평가," (한국통신학회 발표 예정, 1992, 7월).
22. CCITT Contribution COM-41 : The Efeect on speech communication of head-and-shoulder picture presentation in a videophone terminal(1991).

• 부록 : 용어 정의

音原(Sound Source)과 音像(Sound Image)

음원의 존재와 상관 없이 어느 방향으로부터 음이 들린다고 느끼는 경우에, 지각적으로 그 방향에 음이 존재한다고 가정하고, 지각적으로 표현할 경우에 音原대신 音像이라는 용어를 사용한다.

定位(Localization)

음을 지각할때 그 음원의 방향을 특정할 수 있을때의 음의 방향이 정위되었다고 한다. 음원이나 음상에 대해서 똑같이 사용된다.

現場感, 臨場感(Presence, Ambience)

재생음을 들을 때 마치 연주자(또는 송화자) 앞에서 듣고 있는 것 같은, 또는 실제의 연주회장(또는 회의실)있는 것 같은 音響的인 印象.

Head and Torso Simulator(머리 몸통 모사기)

귀의 고막 위치에 마이크로폰을 설치하여 인공적으로 만든 인체 상반신 마네킹으로써, 음향 측정시에 사람의 머리, 외이도 및 몸통의 회절효과를 고려하기 위한 것으로써, 확성전화기, 헤드폰, 보청기등의 음향측정과 고충실도 음질재생을 목적으로 하는 Binaural 녹음에 이용된다. HATS의 머리 부분만을 의사 머리(Dummy Head, Artificial Head)라고도 한다.

▲강 성 훈(정회원)
현 한국전자통신연구소 신호처리연구실 선임연구원
9권 1호 참조