

## 청각계 스펙트럼을 이용한 음색 지각공간

## Perceptual Space for Timbre using Auditory Spectrum

이 채 봉\*, 차 경 환\*, 曾根 敏夫\*\*

(Chai-Bong Lee\*, Kyung-Hwan Cha\*, Toshio Sone\*\*)

## 요 약

기본 주파수 뿐만 아니라 스펙트럼의 형태를 여러 가지 변화시킨 복합음을 이용하여 음색의 유사도에 관한 청취실험을 하였다. 그 결과에서 다차원 척도구성법에 의해 3차원 음색 지각공간을 얻었다. 기본 주파수의 차이는 복합음의 음색을 변화시키는 요인의 하나라는 것을 알 수 있었다.

또한 청각계 스펙트럼을 이용하여 음색의 대략적인 지각을 어느 정도 표현할 수 있는가에 대하여 검토하였다. 청각계 스펙트럼 행렬의 작은 특이치를 제외하여 구한 주파수 가중치행렬을 이용하여 재합성한 결과는 실험결과와 거의 일치하였다.

## ABSTRACT

We study the similarities in timbre by varying both the form of the spectrum and the fundamental frequencies. A three dimensional configuration was drawn as a result of multidimensional scaling analysis. This configuration suggests that the difference in fundamental frequencies is one of the factors that change the timbre of complex tones. Also we considered the possibility of expressing the broad perception of timbre using auditory spectrum. The experimental result is well agreed with the calculation using the frequency weight matrix excluding the small singular value of auditory spectrum.

## I. 서 론

인간은 들을 수 있는 모든 음을 청각계로 처리하고 있다. 여기서 처리의 결과로써 여러 가지 물리적 특징을 가진 공기의 소밀파는 음으로 지각되어 여러 가지 인상을 가진다. 음에 대한 인상으로서의 크기, 높이, 음색이 알려져 있다. 음의 크기는 라우드네스(loudness)라고 불리며 물리량인 음파의 진폭과 밀접한 관계가 있으며 sone이라는 1차원 척도로 표현하는 것이 가능하다. 또, 높이는 음파의 주파수와 밀접한 관계가 있으며 mel이라는 1차원 척도로서 표현하고 있다[1, 2]. 높이와 크기가 1차원 척도로써 거의 기술되고 있는 반면 음색은 매우 복잡하다. 음색은 파형과 밀접한 관계가 있으며, 파형은 여러 가지 물리량에 관계하여 복잡하게 변화하고 이것은 음색이 높이, 크기와는 달리 다차원적인 인상에 대응하고 있다.

음색에 관한 과거의 연구는 음의 물리적인 스펙트럼 구조와 음색을 대응시키는 것이 대부분이다. 曾根[3]는 8종의 악기음을 재생하여 많은 음색 표현어에서 도출한

사용빈도가 높은 21개의 표현어에 대하여 7단계의 평가 척도를 이용하여 실험을 하여 얻은 데이터에 대하여 각 척도간의 상관관계를 구해 인자분석을 한 결과, 음색의 평가를 다음의 4가지 인자로 나타내었다. 제 1인자는 미적·서정적 인자, 제 2인자는 양적·공간적 인자, 제 3인자는 밝음 인자, 제 4인자는 부드러움 인자로 표현하였다. Bismark[4]는 크기와 높이, 연속시간이 같은 스펙트럼 형태가 다른 30종류의 음에 대하여 서로 다른 음색 표현어를 7단계로 평가하였다. 이 결과 dull-sharp, compact-scattered, full-empty, colorful-colorless의 4개 인자를 도출하였다. 이와 같은 연구는 음색지각의 주요인자를 도출하여 음색의 주관척도를 구성하였으며 단순히 다차원적이라고 생각되어져 온 음색이 몇 개의 인자에 의해 다차원성이 설명되었다.

청각계에는 가청 주파수 전체에 걸쳐 어떤 대역폭의 밴드 패스 필터가 연결되어 이 대역폭에서 주파수 분석이 이루어진다고 생각하면 여러 가지 청각현상을 설명하는데 합리적이다. 물리스펙트럼에서 청각계 내에 있는 스펙트럼 표현으로의 변환은 가상적 필터(청각필터)에 의한 주파수 분석이 크게 기여한다. 이러한 청각필터는 Patterson[5, 6], Moore[7, 8]가 통과대역의 면적과 같은 면적을 가진 구형 필터의 통과 대역폭(ERB: Equivalent Re-

\* 동서대학교 정보통신공학과

\*\* 동북대학 전기통신연구소

접수일자: 1996년 8월 1일

ctangular Bandwith)으로 나타내었다.

본 연구에서는 이와 같은 절대적인 음색의 지각이 청각계 스펙트럼을 이용함으로써 어느 정도까지 표현되는가를 고찰하였다. 이를 위해 기본 주파수 뿐만 아니라 스펙트럼 형태도 다른 자극음을 이용하여 청취실험을 함으로써 다차원 음색 지각공간을 도출하였다. 얻어진 음색 지각공간으로부터 음색이 어떻게 지각되는가를 고찰하였다. 또 다차원 지각공간이 자극음의 청각계 스펙트럼의 어느 특징이 음색지각으로 형성되고 있는가를 고찰하기 위해 청각계 스펙트럼과 음색 지각공간의 대응관계에 대하여 검토하였다.

## II. 자극음의 음색 지각공간

### 2.1 실험방법

자극음은 표 I에 나타낸 바와 같이 기본 주파수와 스펙트럼 구조가 다른 36종류의 광대역 복합음이다. 자극음 1에서 12까지는 기본 주파수 100Hz로 성분수를 20, 30, 40으로 하고 성분 주파수 스펙트럼 경사는 0dB/oct, -3dB/oct, -6dB/oct, -9dB/oct로 하였다. 자극음 13에서 24까지는 기본 주파수 200Hz, 자극음 25에서 36까지는 기본 주파수를 400Hz로 하고, 위와 같은 스펙트럼 경사를 각각 두었다. 자극음은 컴퓨터로 합성하고 16bit D/A로 변환하여 출력하였으며 표본화 주파수는 50kHz로 하였다. 전폭의 측정은 인공귀(B&K 4153)를 이용하였다.

자극음은 헤드폰(STAX SR-Lambda Sprii)을 통해 피실험자의 양귀에 제시하였으며 라우드네스가 피실험자의 판단에 영향을 주는 것을 방지하기 위해 Zwicker의 라우드네스 산출법(ISO532B)을 이용하여 자극음의 라우드네스가 약 10sone이 되도록 조절하였다. 자극음의 제시는 랜덤으로 얻은 자극음 2개씩을 피실험자에게 제공하고,

높이를 포함한 넓은 의미의 음색을 7단계로 그 유사도를 판단하도록 하였다. 피실험자는 자극음에 대하여 8회의 반복 판단을 하였으며, 1명당 총 판단 횟수는  $36C_2 \times 8 = 5040$ 회가 된다. 피실험자는 청력이 정상인 성인 남성 5명으로 하였다.

### 2.2 음색 지각공간의 표현 및 고찰

피실험자가 8회 반복 판단하여 얻은 결과를 합으로 하여 각각의 자극음에 대한 음색의 비유사도를 데이터로 하였다. 이 비유사도 데이터를 주관적 거리로 보고, 다차원 척도구성법(동북대학 대형계산기센터 통계패키지 SPSSx의 서브프로그램 ALSICAL, 모델:INDSCAL)을 적용하여 자극음 공간에 분포하였다. 공간은 해석하기 쉽게 3차원으로 나타내었다.

다차원 척도구성법에 의해 얻은 음색 지각공간을 그림 1에 나타내었다. 각 자극음의 분포를 II-I축 및 II-III축에 투영하였으며 ○안에 있는 숫자는 자극음의 번호를 나타내고 있다. 다차원 척도구성법에 의해 얻어진 음색 지각공간을 보면 기본 주파수가 같은 자극음마다 그룹을 형성하고 있다는 것을 알 수 있다. 이것은 기본 주파수가 다름에 따라 주관적 거리가 크게 되지 않는다는 것을 알 수 있다. 즉, 기본 주파수가 같고 스펙트럼의 형태가 다른 자극음의 거리와 기본 주파수만 다른 스펙트럼의 형태가 같은 자극음 사이를 비교하면 양자는 차이가 없다. 이것은 기본 주파수가 다른 자극음에 대해서도 음색을 비교하는 것이 가능하다. 즉, 높이에 의존하지 않는 음의 인상을 비교하는 것이 가능하다는 것을 시사하고 있다. 그러나 높이가 다른 자극음은 음색 지각공간에서 그룹을 형성하고 있기 때문에 음색이 같고 높이가 다른 음은 존재하지 않는다는 것을 시사하고 있다. 따라서 기본 주파수는 스펙트럼 형태와 같은 정도로 음색을 변화시키는 요

표 1. 자극음과 스펙트럼의 구조

자극음 번호	기본주파수 [Hz]	성분수	기본주파수 전폭[dB], 경사[dB/oct]	자극음 번호	기본주파수 [Hz]	성분수	기본주파수 전폭[dB], 경사[dB/oct]
1	100	20	47.3, 0	19	200	20	58.5, 6
2	100	30	43.0, 0	20	200	30	57.6, 6
3	100	40	40.1, 0	21	200	40	57.2, 6
4	100	20	56.1, 3	22	200	20	64.5, 9
5	100	30	53.0, 3	23	200	30	64.0, 9
6	100	40	51.1, 3	24	200	40	63.8, 9
7	100	20	63.5, 6	25	400	20	43.4, 0
8	100	30	61.6, 6	26	400	30	42.7, 0
9	100	40	60.4, 6	27	400	40	42.6, 0
10	100	20	69.7, 9	28	400	20	50.6, 3
11	100	30	68.6, 9	29	400	30	50.3, 3
12	100	40	68.0, 9	30	400	40	50.2, 3
13	200	20	43.1, 0	31	400	20	56.8, 6
14	200	30	40.7, 0	32	400	30	56.6, 6
15	200	40	39.6, 0	33	400	40	56.6, 6
16	200	20	51.4, 3	34	400	20	61.8, 9
17	200	30	49.9, 3	35	400	30	61.7, 9
18	200	40	49.1, 3	36	400	40	61.7, 9

인의 하나라는 것을 알 수 있다. II-I 축을 보면 3개의 그룹이 발급형으로 분포되어 있다. 이것은 기본 주파수가 단일의 자극음에 대해서 2차원의 음색 지각공간[9~12]이 기본 주파수에 대응하는 제 III의 축에 따라 단순히 나열되어 있지 않다는 것을 알 수 있다. 또, I 축은 스펙트럼의 미세한 구조 차이에 대응하고 있음을 알 수 있다. II 축은 sharpness[13]에 관계된 축이며 III 축은 "확장감이 있는"이라는 청감에 대응하는 축[11]이라고 해석되고 있다.

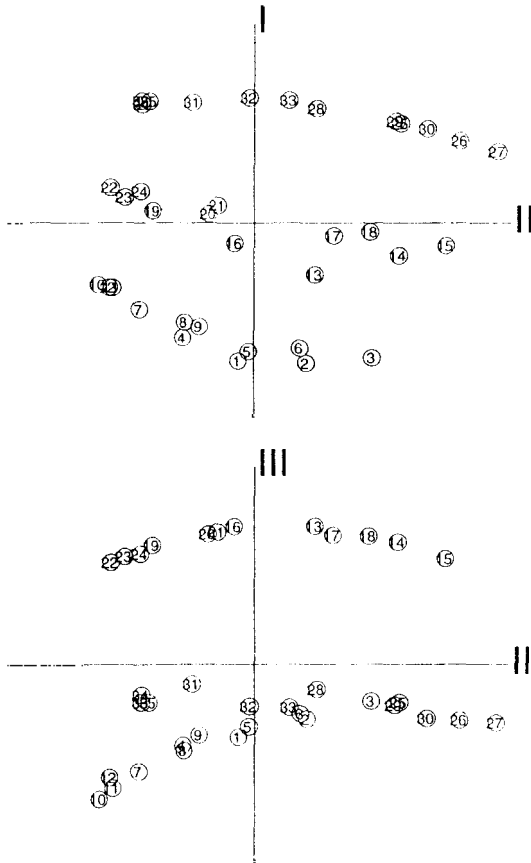


그림 1. 다차원 척도법에 의한 음색 지각공간

III. 청각계 스펙트럼과 음색 지각공간과의 대응

3.1 주파수 가중치를 이용한 음색 지각공간의 표현

r 차원의 음색 지각공간은 자극음을 n개라고 하면 n×r의 좌표행렬 T로 기술된다. 한편 자극음이 m개의 주파수 성분으로 되어 있다면 자극음의 주파수 스펙트럼은 n×m의 행렬 S로 표현된다. 즉, T는 아래의 식으로 나타낼 수 있다[14].

$$T = S \cdot G \tag{1}$$

여기서, 행렬 G는 자극음의 주파수 스펙트럼에서 음색 지각공간을 도출하기 위한 교환행렬이다. 음색 지각공간 T에 있어서 자극 i의 j축 좌표에 대응하는 요소 t<sub>ij</sub>는

$$t_{ij} = \sum_k S_{ik} g_{kj} = \sum_k g_{kj} S_{ik} \tag{2}$$

로 나타낼 수 있다. 이 경우 행렬 S의 요소 S<sub>ik</sub>는 자극음 i의 주파수 k에 대한 라우드네스이다. 행렬 G의 요소 g<sub>kj</sub>는 자극음의 주파수 스펙트럼에서 음색 지각공간 j축을 구하기 위한 자극음 스펙트럼의 k성분 가중치이며 행렬 G는 자극음이 주파수 교환행렬(주파수 가중치행렬)이다. 결국 자극음 공간과 음색 지각공간과의 대응을 생각하는 것은 주파수 가중치행렬 G를 구하는 것이 된다.

3.2 주파수 가중치행렬을 이용한 음색 지각공간의 합성

표 1의 36개 자극음에 대한 3차원 음색 지각공간으로 표현하므로 T는 36×3의 행렬이 된다. 청각계 스펙트럼으로는 3ERB에서 39ERB까지 361개의 필터 출력을 이용하였기 때문에 S는 36×361의 행렬이 된다. 주파수 가중치행렬 G는 미지수 361, 방정식 36개이므로 방정식계의 해는 부정이 된다. 이 해를 풀기위해 S의 의사(擬似) 역행렬을 생각하였다. 임의의 m×n의 행렬 A의 의사 역행렬 A<sup>+</sup>는 A를 특이치 분해하는 것에 의해 아래의 식과 같이 구하는 것이 된다.

$$A^+ = Q_2 \Sigma^+ Q_1^T \tag{3}$$

여기서 Q<sub>1</sub>, Q<sub>2</sub>, Σ는 식 (4)를 만족한다.

$$A = Q_1 \Sigma Q_2^T \tag{4}$$

또, Σ, Σ<sup>+</sup>는 식 (5)와 같이 표현되고, 행렬 A의 특이치 μ<sub>r</sub>를 이용한 행렬이며, Σ는 m×n, Σ<sup>+</sup>는 n×m의 행렬이 된다.

$$\Sigma = \begin{bmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \dots & & \\ \dots & & \mu_r & \\ & & & \\ 0 & \dots & \dots & 0 \end{bmatrix}, \quad \Sigma^+ = \begin{bmatrix} \mu_1^{-1} & 0 & \dots & \dots & 0 \\ 0 & \dots & & & \\ & & \mu_r^{-1} & & \\ 0 & \dots & \dots & & 0 \end{bmatrix} \tag{5}$$

여기서 r은 행렬의 계수이다.

또, S의 작은 특이치를 0으로 두고 의사 역행렬을 구하였다. 이것은 작은 고유치에 대해서는 고유벡터의 방향에 세울 공간을 생각하였을 때, 고유벡터의 방향에 공간의 퍼짐이 적은 것으로 생각하여 고유벡터 방향에 공간의 퍼짐을 고려하지 않은 것에 상당한다. 그림 2는 S의 특이치를 큰 순으로 나타낸 것으로 전후보다 큰 값으로 감소하고 있는 4번째까지의 특이치를 이용하여 주파수 가중치행렬 G를 구하였다. 구한 G의 각 요소를 그림 3에 나타내었다. 그림의 종축은 각 자극음에 대한 청각계 스펙트럼에서 좌표를 구하기 위한 가중치를 나타내며, 횡축은 주파수를 나타낸다.

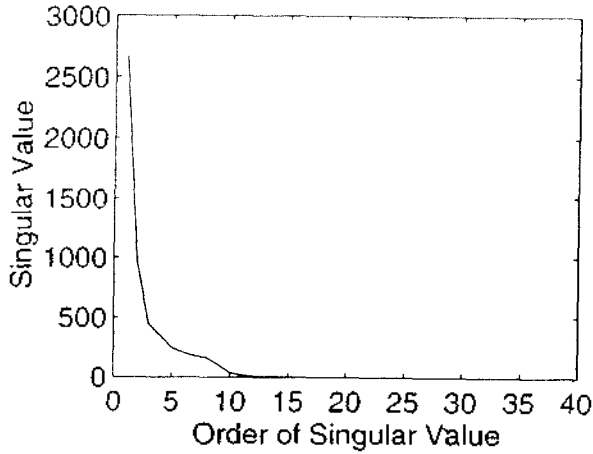


그림 2. 청각계 스펙트럼의 특이치

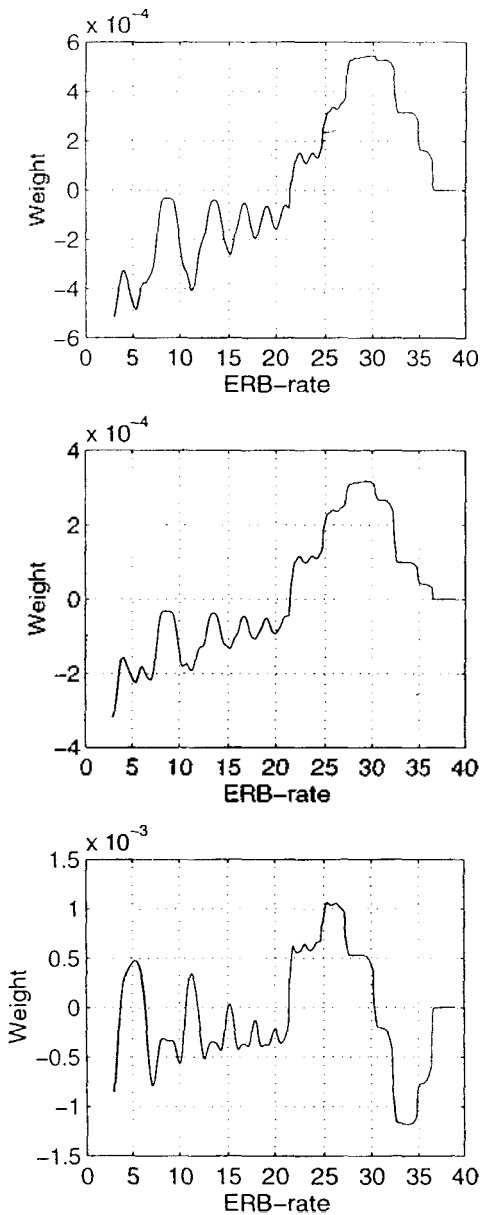


그림 3. 작은 특이치들 0으로 하여 구한 주파수 가중치행렬

3.3 결과 및 고찰

그림 3에 나타낸  $G$ 의 각 요소를 보면 낮은 주파수에서는 툰니 모양으로 증가하고 있음을 알 수 있다. 이것은 청각계 스펙트럼의 진폭이 작은 주파수이며 작은 특이치를 제외한 의사 역행렬을 구하기 위해 조그마한 차이에 가중치를 걸지 못하였기 때문이라 생각되어 진다. 청각계 스펙트럼에서 II축에 대응하는 좌표를 구하기 위한 주파수 가중치는 대략적으로 ERB-rate축에 대해서는 위로 올라가는 형태로 되어 있다. 이것은 sharpness[13]를 추정하기 위한 주파수 가중치함수와 유사하다. I축과 III에 대한 가중치의 값은 정 부로 변화하고 있으며 이것은 청각계 스펙트럼의 미세한 구조와 I축의 좌표가 대응하고 있음을 나타내고 있다. 또, III축에 대한 가중치는 저역과 고역은 정이고 그 사이의 대역은 부로 되어 있을 경우, 이것은 청각계 스펙트럼이 확장되는 음에 대해서는 III축의 좌표가 크게 되며 "확장감이 있는"이라는 청감을 반영하는 것이 되나 여기서는 확실히 나타나고 있지 않다.

그림 3의 가중치를 이용하여 재합성한 자극음의 분포를 그림 4에 나타냈다. 청취실험의 결과와의 비유사도는 0.01이며 거의 일치하는 분포를 얻고 있다.

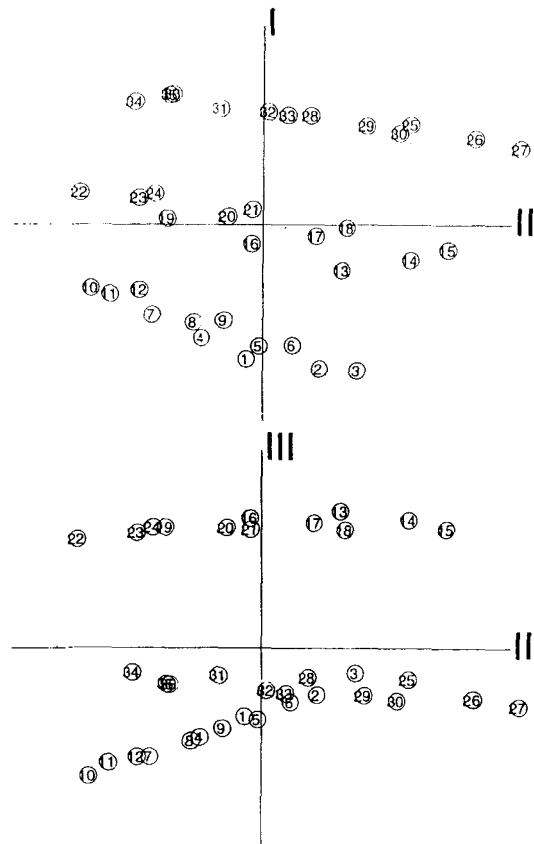


그림 4. 재합성한 자극음의 분포

IV. 결 론

스펙트럼의 형태 뿐만 아니라 기본 주파수도 다른 자극음을 이용하여 청취실험을 하였다. 그 결과 얻은 음색 지각공간에 대하여 고찰하였다. 기본 주파수가 같은 자극음이 그음을 형성함에 따라 기본 주파수의 차이는 복합음의 음색을 변화시키는 요인의 하나라는 것을 알 수 있다. 또 기본 주파수의 차이는 음색의 유사도 판단을 막지 않는다는 것을 알 수 있다.

청각계 스펙트럼 행렬에서 작은 특이치를 제외하고 주파수 가중치를 적용하여 재합성한 결과는 실험결과와 완전히 일치하지 않았으나 sharpness를 추정하기 위한 주파수 가중치합수와 유사하다는 것 등이 얻어졌으며, 실험 결과의 해석에 대응한 특징이 얻어졌다. 이것은 주파수 가중치행렬의 타당성을 나타낸 것이라고 생각되어 진다.

[謝辭]

본 연구는 일본 학술진흥회 외국인 특별연구원으로서 이루어진 연구의 일부분이며 송교수 및 송근연구실 음색 group 대학원생들에게 감사의 뜻을 표한다.

참 고 문 헌

1. 境久雄, "聽覺と音響心理," コロナ社(1992).
2. 城戸健一, "基礎音響工學," コロナ社(1990).
3. 會根敏夫, 城戸健一, 二村忠元, "音の評価に使われることばの分析," 日本音響學會誌, Vol.18, No.6, pp.320-326(1962).
4. G. von Bismark, "Timbre of sound: A factorial investigation of its verbal attributes," Acustica, Vol.30, pp.146-159(1974).
5. Patterson, R., "Auditory filter sharpes derived with noise stimuli," J.Acoust.Soc.Am. 59, pp.640-655(1976).
6. Patterson, R., "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," J.Acoust.Soc.Am. 72, pp.1788-1803(1992).
7. Moore, B.C.J. and Glasberg, B.R., "Suggested formulae for calculating auditory-filter band-widths and excitation patterns," J.Acoust.Soc.Am. 74, pp.750-753(1983).
8. Moore, B.C.J. and Glasberg, B.R., "Formulae describing frequency selectivity as a function of frequency and level. their use in calculating excitation patterns," Hear. Res. 28, pp.209-225(1987).
9. 小澤賢司, "聽覺系における音色知覺過程に関する研究," 東北大學博士學位論文(1993).
10. A. de Bruijn, "Timbre-classification of complex tones," Acustica, Vol.40, pp.108-114(1978).
11. 末木 隆, "廣帶域複合音の音色知覺過程に関する基礎的研究," 東北大學碩士學位論文(1990).
12. 末木 隆, 小澤賢司, 會根敏夫, "廣帶域複合音のスペクトル包絡と音色の知覺過程に関する考察," 聽覺研究會資料, H-90-21S, pp.1-8(1990).

13. G. von Bismark, "Sharpness as an attribute of the timbre of steady sounds", Acustica, Vol.30, pp.159-172(1974).
14. 小澤賢司, 森谷 彰, 鈴木 陽一, 會根敏夫, "廣帶域音の周波數スペクトルと音色知覺空間の對應關係について," 日本音響學會講演論文集, pp.455-456(1995).

▲이 채 봉(Chai-Bong Lee)

1960년 2월 2일생



1985년 2월: 동야대학교 전자공학과 (공학사)  
 1988년 3월: 동북대 대학원 전기 및 통신(공학석사)  
 1992년 3월: 동북대 대학원 전기 및 통신(공학박사)  
 1993년~현재: 동서대학교 정보통신공학과 조교수

▲차 경 환(Kyung-Hwan Cha)

1963년 4월 16일생



1985년 2월: 부산 수산대학교 전자통신공학과(공학사)  
 1990년 2월: 부산 수산대학교 전자통신공학과(공학석사)  
 1996년 8월: 부경대학교 전자통신공학과(공학박사)  
 1990년 1월~1995년 2월: LG전자(주) 생활시스템 연구소 주임 연구원

1995년 4월~1996년 3월: 한국전자통신 연구소 위촉 연구원  
 1995년 3월~현재: 동서대학교 정보통신공학과 전임강사

▲송 근 敏夫(Toshio Sone)

1935년 5월 14일생

현재: 동북대학 전자통신연구소 교수, 동북대학 대형계산기 센터장  
 (1993년 제 12권 2E호 참조)