

비선형 특징추출 기법에 의한 머리전달함수(HRTF)의 저차원 모델링 및 합성

서 상 원[†]·김 기 홍^{††}·김 현 석[†]·김 현 빈^{†††}·이 의 택^{††††}

요 약

3D 음상정위(Sound Localization) 시스템의 바이노럴(Binaural) 필터링을 위해 사용되는 머리전달함수(HRTF) 필터는 그 차수가 높고 모든 방향에서의 값을 가지고 있어야 히브로 상당한 기억 공간을 요구하게 된다. 이에 대해, 많은 차수의 HRTF 데이터에 PCA 등 다차원 데이터의 저차원 특징추출 기법을 적용하여 정보의 큰 손실 없이 모든 방향에 대한 HRTF를 저차로 표현하고 이로부터 원래 HRTF를 효율적으로 합성하는 방법에 관한 연구가 수행되어 왔다. 이들 연구에서는 보편적인 선형 PCA를 주파수 영역 HRTF 데이터에 적용하여 적은 수의 주성분으로써 원래의 HRTF를 근사 합성하였다. 본 논문에서는 이 HRTF 저차원 모델링을 위해 신경망 기반 비선형 PCA 모델인 NLPCA와 비선형 PLS 회귀 모델인 NLPLS를 적용하여 선형 PCA에 의한 결과와 비교 분석하였다. 비선형 곡면 투영을 수행하는 NLPCA가 선형 PCA보다 더 효율적인 HRTF 특징추출 성능을 보였으며 HRTF의 방향정보가 고려되지 않는 PCA에 비해 방향정보를 포함시켜 특징추출 하는 NLPLS 회귀 모델이 학습되지 않은 일반적인 HRTF의 합성에 있어서 더 안정적인 결과를 나타냈다.

Low Dimensional Modeling and Synthesis of Head-Related Transfer Function (HRTF) Using Nonlinear Feature Extraction Methods

Sang-Won Suh[†] · Ki-Hong Kim^{††} · Hyun-Suk Kim[†] · Hyun-Bin Kim^{†††} · Ee-Taek Lee^{††††}

ABSTRACT

For the implementation of 3D Sound Localization system, the binaural filtering by HRTFs is generally employed. But the HRTF filter is of high order and its coefficients for all directions have to be stored, which imposes a rather large memory requirement. To cope with this, research works have centered on obtaining low dimensional HRTF representations without significant loss of information and synthesizing the original HRTF efficiently, by means of feature extraction methods for multivariate data including PCA. In these researches, conventional linear PCA was applied to the frequency domain HRTF data and using relatively small number of principal components the original HRTFs could be synthesized in approximation. In this paper we applied neural network based nonlinear PCA model (NLPCA) and the nonlinear PLS regression model (NLPLS) for this low dimensional HRTF modeling and analyze the results in comparison with the PCA. The NLPCA that performs projection of data onto the nonlinear surfaces showed the capability of more efficient HRTF feature extraction than linear PCA and the NLPLS regression model that incorporates the directional information in feature extraction yielded more stable results in synthesizing general HRTFs not included in the model training.

† 정 회 원 : 한국전자통신연구원 가상현실연구개발센터 선임연구원

††† 정 회 원 : 한국전자통신연구원 책임연구원(가상현실연구개발센터장)

†† 정 회 원 : 한국전자통신연구원 가상현실연구개발센터 연구원

논문접수 1999년 10월 23일, 심사완료 2000년 2월 11일

††† 정 회 원 : 한국전자통신연구원 책임연구원(인지정보연구팀장)

1. 서 론

머리전달함수(Head-Related Transfer Function, HRTF)란 자유음장에서 소리가 청취자의 외이에 이르기까지의 경로를 나타내는 전달함수를 말하며 음원이 청취자의 고막에 이르기까지 겪게되는, 주로 외이(outer ear)에 의한 필터링을 의미한다[1, 2]. 수평면에서의 소리의 방향 지각은 두 귀에 들어오는 음의 크기 차이와 시간 지연 차이가 주된 단서가 되며 음원의 고도각에 대한 방향 지각은 귓바퀴와 어깨 그리고 머리로부터의 반사 형태가 주된 단서를 이루는데, HRTF는 이러한 방향 지각 요소들에 대한 정보를 모두 포함한다. 모노 음원에 대하여 청취자의 왼쪽과 오른쪽 두 귀에 대한 HRTF를 각각 필터링하여 2채널 바이노럴(binaural) 신호를 합성함으로써 일정한 방향성을 갖는 음상(sound image)을 시뮬레이션할 수 있는데 이것이 현재 가상현실 응용시스템 등에서 널리 사용되고 있는 2채널 입체음향 생성 기술의 기본 원리이다[1, 10]

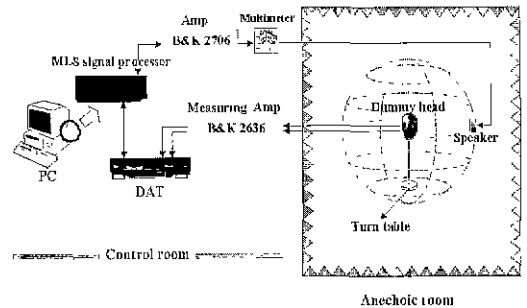
그런데 임의의 방향에 대한 음상정위를 하기 위해서는 3차원 공간 상의 모든 방향에 대한 HRTF 데이터를 갖고 있는 것이 필요하며 이는 HRTF 필터의 많은 차수와 아울러 상당한 양의 메모리를 요구하게 된다. 이 문제에 대해서 Kistler와 Wightman[3], Chen과 Van Veen[4] 등은 HRTF 데이터 집합을 다차원 공간 상의 점집합으로 보고 Principal Components Analysis(PCA)와 같은 통계적 특징 추출 기법을 적용하여 적은 차원의 성분으로 HRTF를 나타내는 방법을 제안하였다

본 논문에서는 이 HRTF의 저차원 모델링에 있어서 보편적인 선형 PCA 대신 신경망 기반 비선형 PCA 기법인 NLPCA와 비선형 회귀 모델인 NLPLS 모델을 적용한 결과를 선형 PCA의 결과와 비교 분석한다.

2. HRTF 데이터

HRTF는 보통 FIR(Finite Impulse Response) 필터의 형태로 표현되며 그 필터 계수들인 임펄스 응답을 시간 영역 HRTF 데이터로서 측정한다. 임펄스 응답의 측정에는 보통 더미헤드(dummy head)를 사용하여 무향실(anechoic chamber)에서 (그림 1)과 같이 측정 장비들을 구성하여 수행하게 된다. 더미헤드로부터 일정 거리에 세팅된 스피커 시스템으로부터 MLS(Maximum Length Sequence) 신호와 같은 입력 신호를 방사하면

더미헤드의 두 귀에 장착된 마이크로폰에 입력되어 출력 전압을 발생시키게 되는데 이들 동시에 측정하고 적절한 신호분석 및 증폭을 통해 좌우측 귀에서의 임펄스 응답 값을 산정해 저장하게 된다



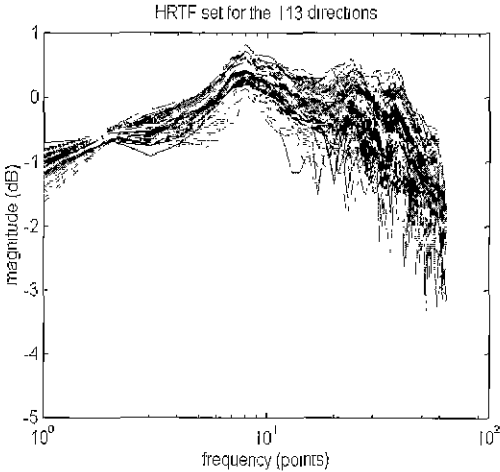
(그림 1) 무향실에서의 HRTF 측정과 장비의 구성

공간 상에서 음원의 방향은 두 방향 파라미터 방위각(azimuth)과 고도각(elevation)에 의해 지정되는데, 방위각은 더미헤드가 장착된 회전 테이블을 돌려서 그리고 고도각은 스피커가 고정된 빔을 회전시켜 변화시키면서 여러 방향에 따른 HRTF 데이터를 측정한다

본 논문에서 사용된 HRTF 데이터는 미국 MIT의 Media Lab이 KEMAR 더미헤드를 사용해 측정하여 공개한 것으로서 고도각 -40도~90도, 방위각 0도~360도 범위의 총 710개 방향에서 측정된 HRTF 임펄스 응답의 집합이다[5] 본 논문의 특성 추출 실험에서는 신경망 학습에 필요한 계산 양의 부담 관계로 이들 중 약 15% 정도인 113개 위치의 HRTF들을 샘플링하여 학습 데이터로 사용하였다 한편, 각 방향의 임펄스 응답 데이터에 대해서는 처음 128 포인트를 취했는데 이 시점은 임펄스 응답에서 주요한 부분인 머리응답(head response)은 대략 다 나타나는 시점이다 128 포인트 임펄스 응답을 이산푸리에변환(DFT)[15]하고 그 크기를 취하여 얻어지는 113개의 64차원 Log 크기응답(magnitude response)들을 HRTF 모델링을 위한 학습 데이터로 구성하였다. (그림 2)는 구해진 113개 HRTF의 주파수 영역 크기응답 데이터를 함께 보여준다.

3. 저차원 특징 추출의 일반 이론

이 장에서는 본 연구에서 적용된 다차원 데이터의 저차원 모델링에 대하여 그 일반적 모델과 이론을 고찰한다.



(그림 2) 총 113개 방향의 HRTF 크기응답

차원 p 의 다차원 데이터를 그 보다 적은 차원 r 의 초평면(hyperplane)이나 곡면으로 근사하여 원래 데이터를 중요 정보의 손실 없이 줄어든 차원으로 표현하고자 하는 것이 특징추출 기법 또는 차원축소 기법의 목적이다. 이에 대한 정식화는 다음 식 (1)과 같이 원래의 점들과 저차원 곡면 상의 투영(projection) 점들 간의 거리 제곱 합을 최소화하는 목적 함수를 설정하고 이 목적 함수를 최적화하는 근사 곡면과 투영 index를 구하는 것이다.

$$\min_{f, s_f} \sum_{i=1}^n \|x_i - f(s_f(x_i))\|^2 \quad (1)$$

여기서 $s_f: R^p \rightarrow R^r$ 는 투영 index를 나타내고 $f: R^r \rightarrow R^p$ 는 다차원 공간상의 점들 가운데를 적합하도록 모수화되고 추정되는 저차원 곡면을 나타낸다. f 가 선형 함수이면 보편적인 PCA의 목적함수가 되고 이때의 f 곡면은 바로 PCA의 근사하는 초평면이 된다. 즉, 선형 PCA는 다음과 같은 목적함수를 취한다.

$$\min_A \|X - \text{proj}_A X\|^2 \quad (2)$$

여기서 X 는 $n \times p$ 데이터 행렬을 그리고 A (rank $A = r < p$)는 저차원 근사 평면의 직교 basis. 그리고 $\text{proj}_A X$ 는 A 에 의해 생성되는 평면으로의 데이터의 투영을 나타낸다. $\text{proj}_A X = XAA^T$ 가 됨으로 (2)의 목적함수는 동등하게 $\text{trace}[(XA)^T(XA)]$ 의 최대화가 됨을 보일 수 있다 즉 PCA 목적함수를 달리 이야기하면 근사 평면

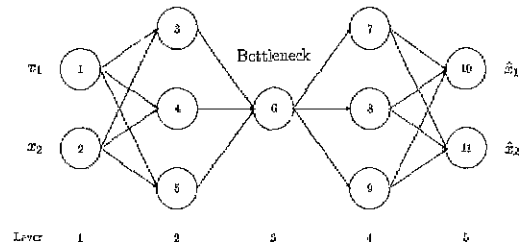
의 각 basis 축으로 투영된 투영점들의 분산의 합을 최대화하는 것이 된다.

최적 \hat{A} 는 X 가 평균 센터링 되었을 때, $X^T X$ 의 eigen vector들을 대응하는 eigen value 크기 순으로 배열하여 얻을 수 있다[13]. 이때 근사 오차들의 크기 제곱 합은 $p-r$ 개의 제일 작은 eigen value들 $(\lambda_{r+1} \dots \lambda_p)$ 의 합과 같으며,

$$\|X - X\hat{A}\hat{A}^T\|^2 = \sum_{j=r+1}^p \lambda_j$$

처음 r 개의 주성분에 의한 데이터 분산 설명율은 $\sum_{j=1}^r \lambda_j / \sum_{j=1}^p \lambda_j$ 로서 얻어진다.

근사 곡면에 대하여 선형성 가정이 없는 경우는 일반적인 비선형 특징 추출의 문제가 되는데 이에 대하여 여러 접근 방법이 연구되고 있다[12]. 본 연구에서는 이 비선형 특징추출 기법을 HRTF 저차원 모델링에 새롭게 적용하였으며 사용한 모델은 일반적 목적 함수 (1)에 대한 비선형 최적화를 신경망 구조를 통하여 수행하는 NLPCA(Non-Linear PCA)이다[11]. NLPCA는 신경망 은닉층의 비선형 transfer 함수를 이용하여 데이터의 가운데를 지나는 r 차원 비선형 곡면 f 를 적합한다. (그림 3)은 NLPCA의 수행을 위한 auto-associative 신경망 구조를 간략화하여 나타낸다. 입력된 p 차원 데이터에 대응하는 p 차원 근사치론 출력층에서 재생하도록 하는 구조이다.



(그림 3) NLPCA를 위한 5-layer feed-forward 신경망 구조

5개의 층으로 구성되며 2번째 층과 4번째 층의 은닉 노드들에 대한 활성화 함수는 비선형 함수를 사용하여 임의의 연속 함수를 근사하도록 한다. 많이 사용되는 비선형 활성화 함수의 형태는 logistic $h(s) = (1 + e^{-s})^{-1}$, scaled logistic $h(s) = (1 - e^{-s}) / (1 + e^{-s})$, hyperbolic tangent $h(s) = \tanh(s) = (e^s - e^{-s}) / (e^s + e^{-s})$ 등이다.

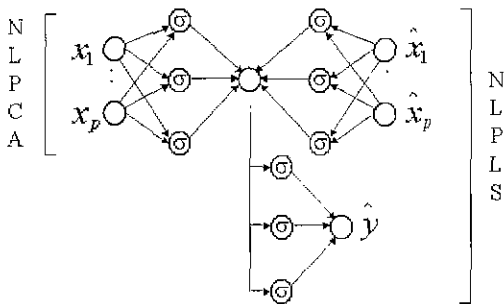
3번째 층과 출력층의 활성화 함수는 보통 선형 $h(s) = s$ 로 한다. 3번째 층은 입력층보다 적은 수의 노드들로 구성되어 병목(bottleneck)을 이루는데 이 층에서 데이터의 차원 압축이 일어나도록 설계한 것이며 병목층의 r 개 노드들의 활성화들은 r 차원 근사 곡면으로의 투영 index를 나타내고 이는 원래의 입력 데이터에 대해 추출된 저차의 특징 값들로 볼 수 있다.

은닉층의 활성화 함수들이 모두 선형일 경우 이 신경망은 선형 PCA의 목적 함수를 최적화하는 것이 되어서 PCA의 신경망 구현이 된다.

Nonlinear Partial Least Squares(NLPLS) 회귀 모델 [12]은 NLPCA모형을 기반으로 하며 예측변수(predictor variable)들에 대한 저차원 투영은 NLPCA를 사용하고 이에 확장하여 투영된 특징 값들로부터 반응변수(response variable)들을 적합하는 비선형 회귀 모델을 또한 포함한다. 목적함수는 다음과 같이 입력 데이터의 저차원 근사에 의한 오차 제곱합과 다시 저차원 특징들로부터 반응변수들에 대한 회귀오차 제곱합의 합으로 설정된다

$$\min_{f, s_r, h} \sum_{i=1}^N [\|x_i - f(s_r(x_i))\|^2 + \|y_i - h(s_r(x_i))\|^2]$$

첫째 항은 NLPCA의 목적함수와 같으며 둘째 항의 h 는 특징값들로서 반응변수 y_i 에 회귀한다. (그림 4)는 NLPLS 모델을 구현하는 신경망 구조를 도해한다



(그림 4) NLPLS 구현을 위한 신경망 구조

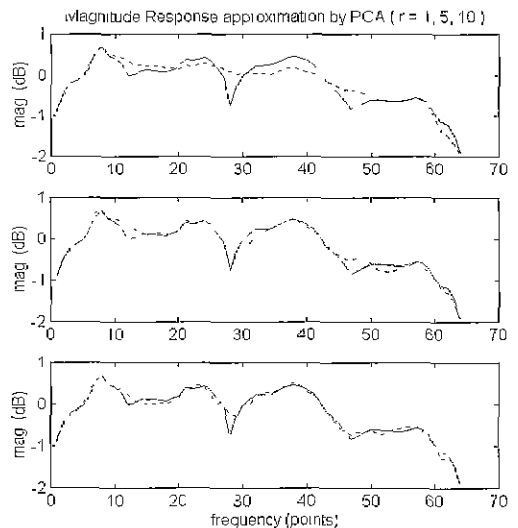
4. PCA와 NLPCA에 의한 HRTF 특징추출 및 합성

2장에서 설명한 HRTF 로그 크기응답 데이터를 대상으로 선형 PCA와 NLPCA를 적용하여 특징 추출 및 원래 크기응답의 근사를 실험했다.

Matlab 함수들을 사용하여 113개의 64 포인트 로그 크기 응답들의 평균 백터, 공분산(covariance) 행렬을 구하고 그 eigen value, eigen vector 쌍을 얻어 PCA를 수행하였다. 처음 r 개의 주성분(principal components)에 의해 설명되는 분산 비율의 결과는 <표 1>에 정리한다. 또한 3장에서 설명된 대로 원래 크기 응답을 PCA에 의해 r 차원 평면에 투영하여 근사시킨 결과물을 볼 수 있는데, (그림 5)는 PCA 결과 구해진 처음 1개, 5개, 10개의 주성분들을 각각 사용하여 원래의 크기응답(실선)을 근사한 결과(점선)의 예이다; 고도각 0도 방위각 270도의 경우

<표 1> 주성분에 의한 크기응답 데이터의 분산 설명률 비교

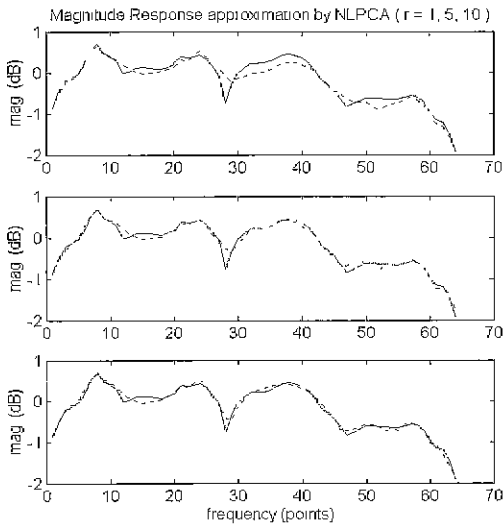
차원 r	Fraction of Variance Explained (%)		차원 r	Fraction of Variance Explained (%)	
	PCA	NLPCA		PCA	NLPCA
1	73.8	80.3	11	95.4	97.3
2	80.1	87.8	12	95.8	97.7
3	84.0	91.6	13	96.2	97.9
4	87.5	93.8	14	96.5	98.2
5	89.9	94.8	15	96.8	98.3
6	91.4	95.4	16	97.1	98.4
7	92.7	95.1	17	97.4	98.6
8	93.6	96.4	18	97.6	98.7
9	94.2	96.7	19	97.8	98.9
10	94.9	97.0	20	98.0	99.0



(그림 5) 선형 PCA에 의한 HRTF 크기응답의 근사 (고도각 0, 방위각 270)

다음으로 앞에 설명된 NLPCA를 신경망 simulator를 사용하여 실험 데이터에 적용하였다. NLPCA 목적 함수의 최적화를 위한 신경망 weight들의 학습은 LBFGS 비선형 최적화 알고리즘[14]의 반복을 통하여 수행했다. 병목(bottleneck) 층의 노드 수, 즉 축소된 차원 r 과 은닉층의 노드 수, 그리고 비선형 활성화 함수의 형태 등을 정해 네트워크를 구성하고 HRTF 데이터를 입력하여 반복 학습한다. NLPCA 신경망 학습은 각각의 차원 r 에 대하여 필요하다. 은닉층의 비선형 전달함수 형태는 scaled logistic으로 하였다.

(그림 6)은 학습된 NLPCA의 처음 1개, 5개, 10개의 주성분들을 사용한 근사 결과를 마찬가지로 보여준다. 스펙트럼 중간중간의 골(notch)들은 HRTF의 중요한 스펙트럼 단서들 가운데 하나로 여겨지는데[1], (그림 5)와 비교할 때 이 골들을 비선형 특징추출 기법인 NLPCA가 같은 주성분 수에 대하여 더 정확하게 근사하고 있다. <표 1>은 데이터의 총 분산 중 PCA와 NLPCA의 r 차원 근사에 의해 설명되는 분산의 비를 각각 비교한다. 1~4차원까지는 NLPCA의 설명율이 PCA보다 7% 정도 높은 결과이다.



(그림 6) NLPCA에 의한 HRTF 크기응답의 근사 (고도각 0, 방위각 270)

실험에서 구성된 HRTF 데이터에 대하여 95% 이상의 분산 설명율을 달성하는데 PCA는 11개 성분들을 필요로 하는데 대하여 신경망 기반 NLPCA는 6개 성

분들로 이를 달성하였는데, 이를 기준으로 축소 차원을 결정한다면 HRTF 크기응답 합성을 위해 저장에 필요한 주성분 데이터의 크기도 반 정도 더 줄일 수 있는 것이다.

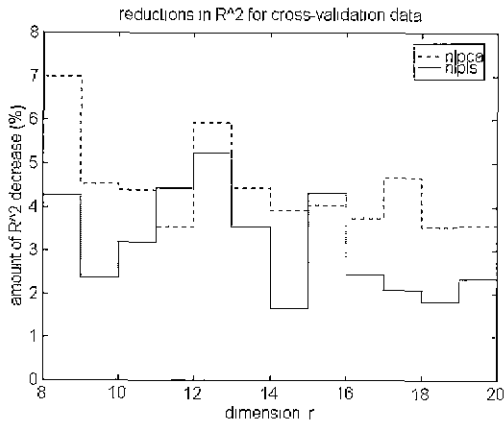
5. 방향정보를 포함한 HRTF의 특징추출 및 합성 : NLPLS의 적용

각 HRTF 데이터에는 방향정보가 연관되어 있다. 즉, 그 데이터가 측정된 위치의 방위각과 고도각이 알려지 있다. 그런데 위에서의 선형 및 비선형 주성분 분석 기법에는 그 목적함수의 최적화에 있어서 HRTF에 대한 방향정보는 전혀 고려되지 않는다. 그런데 방향정보를 반응변수(response variable)로 포함시켜 HRTF의 저차원 투영 시 추출되는 특징 값들이 또한 그 방향정보를 잘 적합할 수 있도록 최적화하는 것을 생각할 수 있는데 PLS(Partial Least Squares) 회귀(regression) 모델을 여기에 자연스럽게 적용할 수 있다.

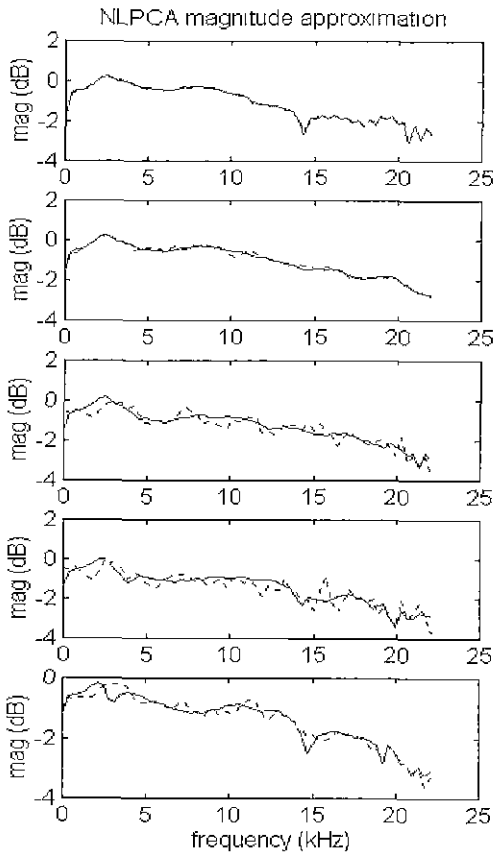
본 논문에서 적용한 것은 NLPCA 신경망 구조에다 투영된 특징값들로서 반응변수를 예측하는 회귀모델 부분이 첨가된 비선형 PLS인 NLPLS 모델이다[12]. NLPLS 모델링에서는 추출되는 저차원 특징들이 방향정보 또한 잘 적합하도록 최적화한다. 따라서 학습데이터로 사용되지 않은 HRTF들에 대해 예측하여 저차원 합성할 경우 방향정보가 고려되지 않은 PCA 방법보다 더 정확한 근사를 제공할 것으로 기대했는데 이를 뒷받침할만한 적용 결과를 얻었다.

(그림 7)이 나타내는 것은 학습에 사용 안 된 HRTF들을 포함하는 원래의 전체 710개 HRTF 집합에 대해 NLPCA와 NLPLS를 적용하여 근사 예측한 결과를 비교한 것이다. NLPCA는 방향정보의 사용 없이 학습되었으며 NLPLS 모델은 반응변수에 해당하는 방향정보로서 방위각을 포함시켜 학습한 것이다.

R^2 값은 얻어진 저차원의 특징들에 의해 원래의 HRTF 크기응답들이 근사되는 정도를 전체적으로 측정하는데, 그림에서는 두 방법을 학습된 HRTF 집합과 710개 전체 HRTF 집합에 대해서 적용했을 때 그 R^2 값이 감소한 크기를 보여준다. 실선으로 나타난 NLPLS 경우가 대부분의 축소 차원에 대하여 NLPCA 경우보다 그 감소의 정도가 낮은 것을 볼 수 있다. 즉, 학습되지 않은 새로운 HRTF들에 대해 근사 정확도가 떨어지는 정도가 NLPCA에 비해 작다는 것을 의미한다.



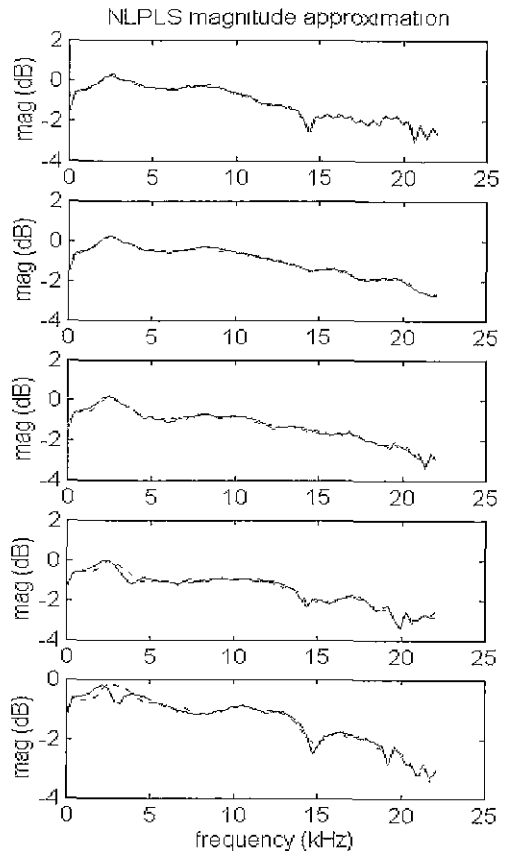
(그림 7) 학습 안된 HRTF들에 대한 근사 정확도의 감소. NLPCA vs. NLPLS



(그림 8) 크기응답의 NLPCA 예측 근사

(그림 8)과 (그림 9)는 두 방법을 적용하여 얻어진 크

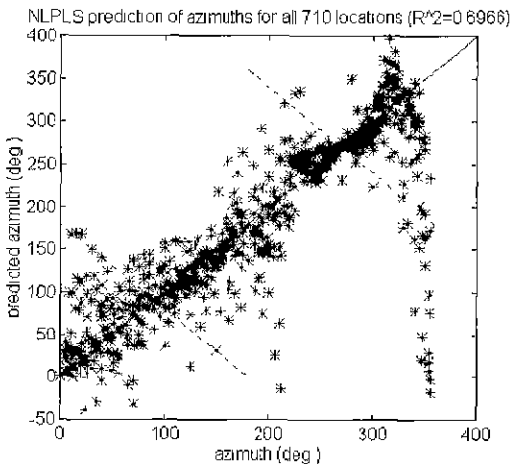
기응답 데이터의 근사 결과를 비교한다. 두 그림의 첫째 열의 크기응답은 학습에 포함된 데이터이고(고도각 10도, 방위각 90도인 경우) 나머지 열들의 크기응답은 학습에 포함되지 않았던 것들이며(고도각 10도, 방위각 각각 95, 100, 105, 110도인 경우) 이들을 축소된 차원 $r=20$ 에 대해서 NLPCA, NLPLS 두 방법에 의해 각각 예측 근사한 결과이다. 학습되지 않은 HRTF들에 있어 NLPLS에 의한 합성은 안정되어 있는 반면 NLPCA의 경우 합성이 불안정하게 된 것을 볼 수 있다. 전체 결과에서 이런 현상을 가끔씩 볼 수 있었다.



(그림 9) 크기응답의 NLPLS 예측 근사

(그림 10)은 방위각을 반응변수로 포함하여 학습된 NLPLS 모델이 주어진 710개 크기응답 테스트 데이터에 대해서 방위각을 예측한 결과를 보여준다. 테스트 데이터가 측정된 실제 방위각(가로축)에 대하여 학습된 회귀모델에 의해 예측된 방위각(세로축)을 나타낸

다. 그림에서 점선으로 표시된 2개의 선분 부근은 청취자의 음상정위에 있어 잘 알려져 있는 문제인 앞-뒤 혼돈(reversals) 현상(즉, 앞에 있는 음원을 뒤쪽에 보는 뒤에 있는 음원을 앞쪽에 있는 것으로 지각하는 현상)[1]을 나타내는데 이 앞뒤 혼돈 현상에 NLPLS 모델의 방위각 추정 결과들이 나타나는 것은 흥미롭다.



(그림 10) NLPLS 모델의 HRTF 방위각 예측

6. 논 의

본 논문에서 수행한 저차원 모델링 실험은 방향에 따른 HRTF 스펙트럼 특성을 나타내는, 핵심적인 크기 응답만을 대상으로 하였다. 일반적으로는 위상정보도 포함한 HRTF를 구하여 시간영역에서의 합성된 임펄스 응답을 얻고 이를 필터링에 사용하게 된다. 합성된 크기 응답으로부터 시간영역 데이터를 얻는 방법은 HRTF의 최소위상(minimum phase)특성을 가정하여 크기 응답과 위상의 힐베르트 변환 관계로부터 위상특성을 구해서 얻는 방법이 있다[15]. 위상특성의 가정 없이 직접 복소 HRTF(임펄스 응답의 complex Fourier Transform)의 실수부와 허수부 각각에 대하여 특징추출을 적용하여 합성한 후 역변환하여 임펄스 응답을 얻는 방법도 연구되었는데[4, 7] 이 경우에는 위상정보에 대한 특징추출도 같이 하게 되어서 크기응답만을 합성하는 경우보다 더 많은 수의 주성분을 필요로 하게 된다.

7. 결 론

본 논문에서는 HRTF의 저차원 특징 추출 및 근사적 합성을 위해서 기존에 제안된 선형 PCA 방식에서 나아가 일반적인 곡면 근사를 수행하는 비선형 PCA 방법과 회귀 모델을 사용해서 특징 추출시 방향 정보(방위각 및 고도각)를 포함시키는 방법을 제시하였다.

NLPCA의 적용 결과, 크기응답 데이터에 대하여 선형 투영에 의한 PCA 보다 좀 더 우수한 특징추출 성능을 보였다. 이는 다차원 공간 상에서 방향에 따른 HRTF 데이터의 분포가 비선형 곡면으로 모델링 하기에 더 유리하다는 것을 나타내는 것으로서 앞으로 HRTF 데이터에 대한 비선형 특징 추출의 연구가 요청된다고 하겠다.

또한, 기존의 PCA 특징 추출에서는 HRTF들이 다차원 공간상의 점집합으로 간주될 뿐 각 HRTF에 대한 방향정보는 고려되지 않았다. NLPLS 회귀 모델을 적용하여 특징추출한 결과, 모델 학습 데이터로 사용되지 않은 일반적 HRTF에 대해서 주성분 분석만 하는 경우보다 더 robust한 근사 합성의 결과를 얻었으며 방향정보도 포함시켜 최적화하는 것이 효과적임을 볼 수 있었다.

HRTF의 저차원 특징추출 방법의 유용한 응용으로서 불측정된 위치의 HRTF에 대한 보간(interpolation)을 생각할 수 있다. 원래의 모든 차수의 HRTF 대신 특징추출 기법으로 얻어진 압축된 특징 값들만을 보간에 사용함으로써 보간의 효율성을 높이는 것이다.

본 연구에 사용된 비선형 특징추출 기법을 향후 이 분야에 응용하기 위해서는 추출된 특징 값들에 대한 해석 및 근사 곡면의 유용한 모수화(parameterization)에 대한 보다 심도있는 연구가 필요할 것이다.

참 고 문 헌

[1] D. R. Begault, '3-D Sound for Virtual Reality and Multimedia.' Academic Press Inc., 1994.
 [2] H. Moller, M. F. Sorensen, D. Hammershoi, C. B. Jensen, "Head-Related Transfer Functions of Human Subjects," J Audio Eng. Soc., 43(5), pp 300-321, 1995.

[3] D. J. Kistler, F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction." J. Acoust. Soc. Am. 91(3), pp 1637-1647, 1992.

[4] J. Chen, B. D. Van Veen, "A spatial feature extraction and regularization model for the head-related transfer function," J. Acoust. Soc. Am. 97(1), pp.439-452, 1995.

[5] B Gardner, K. Martin. "HRTF Measurements of a KEMAR Dummy-Head Microphone," MIT Media Lab Perceptual Computing - Technical Report #280, 1994

[6] 김봉수, 안철용, 방희석, 성평모, "최소위상특성을 이용한 HRTF의 합성", 한국음향학회 학술발표대회 논문집 제15권 1(s)호, pp.259-262, 1996.

[7] C-Y Ahn, H-S Bang, K-M. Sung, "Model of HRTF based on complex-valued PCA considering group delay," Proc. of ASVA '97, pp.365-372, 1997.

[8] 서상원, 김기홍, 김진옥, 김현빈, "머리전달함수의 효율적인 합성을 위한 데이터의 차원 축소 기법에 관한 연구", 한국음향학회 학술발표대회 논문집, 제16권 1(s)호, pp.157-162, 1997.

[9] S.-W. Suh, K.-H Kim, H.-C. Lee, H.-B. Kim, "Neural Network Based Nonlinear Feature Extraction of Head-Related Transfer Function (HRTF) Data in 3-D Sound Processing," Proc. of NC '98, 1998.

[10] H. Myung, S.-W. Suh, et.al. "The Development of 3D Sound Signal Editor *SoriWave* for Multimedia Contents," Proc. of the 106th AES Convention, 1999

[11] E. Malthouse, R. Mah, and A. Tamhane. "Some theoretical results on nonlinear principal components analysis." Proc. of the American Control Conference, 1995.

[12] E. Malthouse, "Nonlinear Partial Least Squares." Ph.D. Dissertation, Northwestern University, 1995.

[13] R. A. Johnson, D. W. Wichern, 'Applied Multivariate Statistical Analysis.' Prentice Hall Inc., 1982.

[14] D. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," Mathe-

matical Programming. Vol.45, pp.503-528, 1989.

[15] A. V. Oppenheim, R. W. Schaffer, 'Discrete-Time Signal Processing,' Prentice Hall Inc., 1989.



서 상 원

e-mail Sangwcon@etri.re.kr

1989년 서울대학교 계산통계학과 (학사)

1990년 영국 Univ. of Oxford 통계학과(석사)

1993년~1998년 (구)SERI 인공지능연구부 연구원

1998년~현재 한국전자통신연구원 가상현실연구개발센터 선임연구원

관심분야 : 음성인식, 입체음향, 신호처리, 패턴인식



김 기 홍

e-mail : ghkim@seri.re.kr

1994년 경북대학교 전자공학과 (학사)

1996년 경북대학교 전자공학과 (석사)

1995년~1998년 (구)SERI 연구원

1998년~현재 한국전자통신연구원 가상현실연구개발센터 연구원

관심분야 : 음성신호처리, 입체음향처리, 생체신호처리



김 현 석

e-mail : kimmhs@etri.re.kr

1991년 한국과학기술원 기계공학과 (학사)

1993년 한국과학기술원 기계공학과 (석사)

1998년 한국과학기술원 기계공학과 (박사)

1998년~현재 한국전자통신연구원 가상현실연구개발센터 선임연구원

관심분야 : 신호처리, 적응신호처리 및 제어, 입체음향, 진동 및 소음제어



김 현 빈

e-mail : hbkim@etri.re.kr

1985년 중앙대학교 응용통계학과
(학사)

1988년 중앙대학교 응용통계학과
(석사)

1996년 일본 오카야마대학 대학
원 자연과학연구과(박사)

1984년~1998년 (구)SERI 책임연구원

1998년~현재 한국전자통신연구원 책임연구원
(창각정보연구팀장)

관심분야 : 입체음향, 생체신호처리, 인공지능, 패턴인식



이 의 택

e-mail : etlee@etri.re.kr

1978년 서울대학교 공과대학
공업교육학과(학사)

1980년 서울대학교 대학원 전자
공학과(석사)

1996년 한국과학기술원 전기전자
공학과(박사)

1980년~현재 한국전자통신연구원 책임연구원
(가상현실연구개발센터장)

관심분야 : 가상현실, 그래픽스, 애니메이션, 영상처리