

새로운 학습 하이브리드 실내 충격 응답 모델

New Learning Hybrid Model for Room Impulse Response Functions

신민철*·왕세명†

Mincheol Shin and Semyung Wang

Key Words: Room Impulse Response (실내 충격 응답), Learning Hybrid Modeling Method (학습 하이브리드 모델링 방법)

ABSTRACT

Many trials have been used to model room impulse responses, all attempting to provide efficient representations of room acoustics. The traditional model designs for room impulse response seem to fail in accuracy, controllability, or computational efficiency. In time domain, a room impulse response is generally considered as the combination of three parts having different acoustic characteristics, initial time delay, early reflection, and late reverberation. This paper introduces new learning hybrid model for the room impulse response. In this proposed model, those three parts are modeled using different models with learning algorithms that determine the length or boundary of each model in the hybrid model. By the simulation with measured room impulse responses, it was examined that the performance of proposed model shows the best efficiency in views of both the parameter numbers and modeling error.

———— 기호설명 ————

$h[n]$: 실내 충격 응답 함수

$h'[n]$: 실내 충격 응답 함수의 위치 민감도

1. 서론

효과적인 실내 음향 특성의 표현을 위해 정확한 실내 충격 응답 (Room Impulse Response, RIR) 함수를 얻어내고자 많은 방법들이 시도되었으며 이는 실내 음향 연구 분야에 많이 이용되었다. 측정을 통해 얻어진 RIR 함수는 적당한 모델로 대체되어 음향 신호 처리 과정을 통해 3 차원 게임, 실시간 입체 음향 등의 분야에 응용되고 있다. 또한 음향 공간의 RIR 함수를 실시간으로 예측하는 것을 주요 기술로 사용하는 잔향 제거기, 음향 피드백 제거기와 같은 응용 분야에서는 측정된 RIR 함수를 어떤 모델로 대체하느냐가 기기의 성능을 좌우한다.

현재까지 대표적으로 쓰이고 있는 RIR 함수의 모델은 자기회귀 (Autoregressive, AR) 모델과 이동평균 (Moving Average, MA) 모델, 그리고 둘을 복합한 ARMA 모델을 대표적으로 사용해 왔다. 또한 통계적인 음향 모델 [1] 도 소개 되었고 협대역 필터를 이용한 방법(Subband Technique) [2], 부분 인공 재생 방법 (Partial Artificial Reproduction) [3] 등, 원래 길이의 RIR 함수를 신호처리가 용이

하도록 작은 신호 블록으로 나누어 처리하는 기술들이 개발 되었으며 최근에는 주파수 영역 확대 pole/ zero 모델 [4] 등 많은 연구들이 진행되어 왔다. 하지만 현재까지의 대표적인 모델링 기법들은 만족할 만한 정확도를 갖게 되면 이를 처리하는 계산량이 많아 실질적으로 적용하는데 한계가 있고 계산량이 적은 모델의 경우에는 정확도가 문제가 되어 왔다. 그러므로 RIR 함수를 정확하게 표현할 수 있으며 처리해야 할 계산량을 기존의 모델과 비교하여 줄일 수 있는 모델을 개발하는 것이 필요하다.

본 논문에서는 계산량을 줄이며 만족할 만한 정확도를 유지시키기 위한 방법으로 학습 하이브리드 모델을 제안하였다. 일반적으로 RIR 함수는 시간 영역의 음향학적인 관찰을 통해서 세 개의 구간으로 나눌 수 있으며 이는, 초기 지연 구간, Early Reflection (ER) 구간, Late Reverberation (LR) 구간으로 명명된다. 각각의 구간은 뚜렷이 구별되는 특성이 있으며 이를 효과적으로 표현하기 위해서 본 논문에서는 각 구간별 모델을 달리하는 하이브리드 모델을 제안하였다. 이때 실시간으로 변하는 각 구간의 경계는 학습 알고리즘을 이용하여 정의하였다. 이와 같이 제안된 모델은 기존의 전통적인 RIR 함수의 모델들과 비교한 결과, 정확도 유지와 계산량 감소 측면에서 동시에 향상된 결과를 얻을 수 있었다.

본 논문의 2 장에서는 실내 충격 응답의 특성에 대해 고찰하였고 3 장은 본 논문을 통해 제안된 학습 하이브리드 실내 충격 응답 모델에 대해 소개하였다. 4 장에서는 제안된 새로운 실내 충격 응답 함수 모델을 기존에 많이 쓰이던 전통적인 모

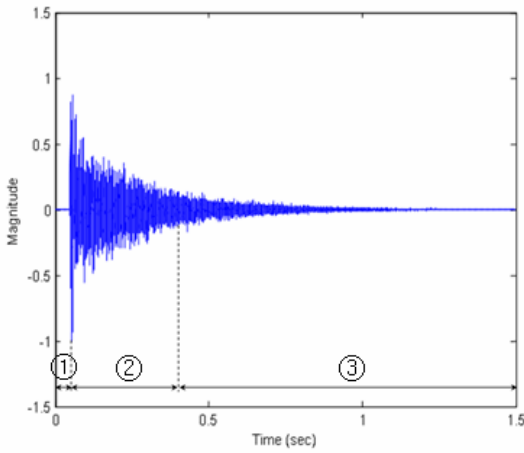
† 교신저자; 광주과학기술원, 기전공학과
E-mail : smwang@gist.ac.kr
Tel : (062) 970-2390, Fax : (062) 970-3263

* 광주과학기술원, 기전공학과

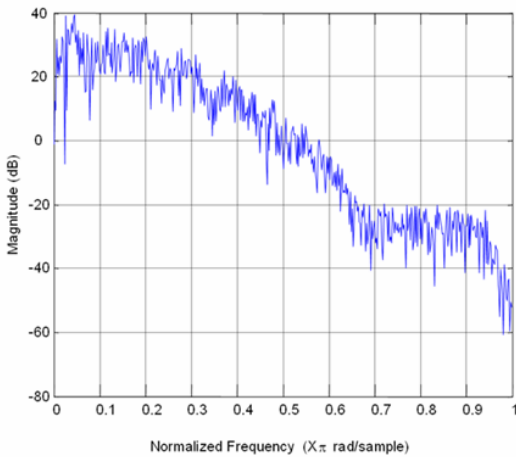
델과 비교하여 성능 평가를 실시하였다. 마지막으로 5 장을 통해 본 논문을 마무리 하였다.

2. 실내 충격 응답의 특성

본 장에서는 실내 충격 응답 함수의 시간 및 주파수 영역 특성을 한 예를 통하여 알아보려 한다. New Mexico State University 에서 48kHz 의 Sampling Rate 로 측정된 RIR 함수를 그림 1 에서 시간 영역과 주파수 영역으로 표현하고 있다[5].



(a) Time-domain representation



(b) Frequency-domain representation

그림 1. 실내 충격 응답 함수

그림 1(a)은 RIR 의 시간 영역에서의 표현을 나타내고 있다. 시간 영역 RIR 에서는 각각 현저히 다른 시간 영역 음향 특성과 함께 ① 초기 지연 구간, ② Early Reflection (ER) 구간, ③ Late Reverberation (LR) 구간으로 확연하게 구분된다. 반면에 (b)의 주파수 영역의 표현은 분리 기준이 모호하고 주파수 영역별로 음향학적 의미를 내포

하고 있지 못하다.

결과적으로 RIR 은 시간 영역 표현에서 확연하게 구분된 영역을 음향학적 특성을 고려하여 각기 다른 모델을 적용하여 모델링하는 것이 효과적이다.

3. 학습 하이브리드 실내 충격 응답 모델

본 장에서는 앞 장에서 음향학적 특성을 고려하여 초기 지연, ER, LR 의 세 개의 구간으로 나뉜 RIR 을 위한 효과적인 모델을 제안하였다. 각 구간의 특성을 잘 반영할 수 있는 각기 다른 모델을 적용하였고, 여기에 시간에 따라 변하는 각 구간의 경계를 정의하는 학습 알고리즘을 구현하였다. 그림 2 에 나타낸 바와 같이 초기 지연 구간에는 Delay Model, ER 구간에는 MA (All-zero) Model, 그리고 LR 구간에는 AR (All-pole) 혹은 ARMA (Pole-zero) Model 을 적용하고 이를 학습 하이브리드 모델이라 명명하였다.

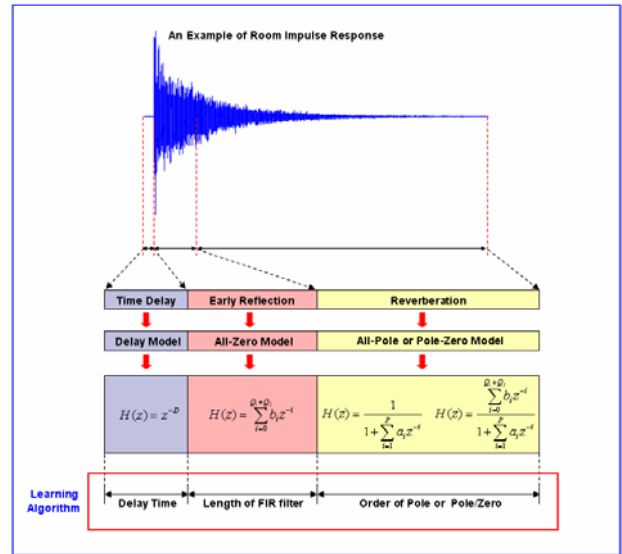


그림 2. 제안된 학습 하이브리드 모델의 개념도

3.1 하이브리드 실내 충격 응답 모델

언급한 바와 같이 RIR 의 세 부분의 음향학적 특성은 확연히 구분되어 있다. 초기 지연 구간은 음원과 수음점 사이의 거리를 나타내고 시간 지연 이외의 정보를 가지고 있지 않다. ER 의 경우 음원과 수음점의 위치 변화에 굉장히 민감한 부분이며 LR 의 경우는 공간 자체의 음향학적 특성을 잘 반영하며 위치의 변화에는 그리 민감하지 않다.

이와 같은 각 부분의 특성을 반영하여, 초기 지연 부분에 대해서는 다음의 식 (1)과 같은 시간 지연 모델을 적용하였다. 여기서 시간 지연은 음원과 수음점 사이의 거리에 비례한다.

$$H(z) = z^{-D}, \quad (1)$$

여기서 D 는 직접 경로의 시간 지연에 해당하는 샘플의 개수를 의미한다.

ER 은 음원에서 수음점까지의 직접 경로를 포함하는 간접 반사 경로로 구성되어 있다. 이러한 반사 경로들은 음원과 수음점의 위치 변화에 상당히 민감하다. 그러므로 이러한 음향학적 특성 때문에 이러한 변화를 민감하게 반영할 수 있는 식 (2)와 같은 MA (All-zero) Model 을 적용하였다.

$$H(z) = Cz^{-Q_1} \prod_{i=1}^{Q_2} (1 - q_i z^{-1}) = \sum_{i=0}^{Q_1+Q_2} b_i z^{-i}, \quad (2)$$

여기서 C 는 상수이며, Q_1 과 Q_2 는 All-zero 모델의 차수를 나타낸다. q_i 는 zeros 의 값을, b_i 는 MA 모델의 계수값을 나타낸다.

LR 구간은 ER 과 같이 반사 경로의 정보를 제공하기 보다는 전체적인 실내 음향 특성을 잘 나타낸다. 이러한 특성 때문에 LR 구간을 효과적으로 표현하기 위해서는 음원과 수음점의 위치에 민감한 모델보다 이러한 위치 변화에 강인하면서 전체적인 특징을 잘 나타내는 모델이 적당하다. 결과적으로 식 (3)으로 표현되는 AR (All-pole) 혹은 식 (4)의 ARMA (Pole-zero) Model 이 이러한 특성을 효과적으로 반영한다.

$$H(z) = \frac{C}{\prod_{i=1}^P (1 - p_i z^{-1})} = \frac{1}{1 + \sum_{i=1}^P a_i z^{-i}}, \quad (3)$$

여기서 P 는 All-pole 모델의 차수를 나타내며 p_i 는 pole 의 값을, a_i 는 AR 모델의 계수값을 나타낸다.

$$H(z) = Cz^{-Q_1} \frac{\prod_{i=1}^{Q_2} (1 - q_i z^{-1})}{\prod_{i=1}^P (1 - p_i z^{-1})} = \frac{\sum_{i=0}^{Q_1+Q_2} b_i z^{-i}}{1 + \sum_{i=1}^P a_i z^{-i}}, \quad (4)$$

C 는 상수이며, Q_1 과 Q_2 는 Zero 의 차수를 나타낸다. q_i 는 Zeros 의 값을, b_i 는 MA 의 계수값을 나타낸다. P 는 Pole 의 차수를 나타내며 p_i 는 Poles 의 값을, a_i 는 AR 의 계수값을 나타낸다. 본 논문에서는 ARMA Model 을 적용하였다.

3.2 Learning Algorithm

음원과 수음점의 움직임 또는 내부 구조물이나 사람의 움직임에 따라 RIR 함수 역시 변하게 된다. 이렇게 시간에 따른 환경 변화에 따라 변하는 RIR 을 앞서 언급한 하이브리드 실내 충격 응답 모델로 표현하기 위해서는 초기 지연과 ER 구간, ER 과 LR 구간의 경계에 대한 명확한 정의가 필요하다. 이를 위해 학습 알고리즘이 도입되었다.

(1) 초기 지연과 ER 의 경계

RIR 함수의 초기 지연 구간과 ER 구간 사이의 경계는 두 구간 사이를 지날 때 발생하는 RIR 함수값의 급격한 변화로 쉽게 찾아낼 수 있다. 따라서 적당한 문턱값 (Threshold, THD) 를 정의하여 그림 3 과 같은 RIR 함수값, $h[n]$ 과의 비교를 통한 학습 알고리즘으로 초기 지연과 ER 구간 사이의 경계를 용이하게 찾아낼 수 있다. 본 논문의 THD 는 0.005 로 하였다.

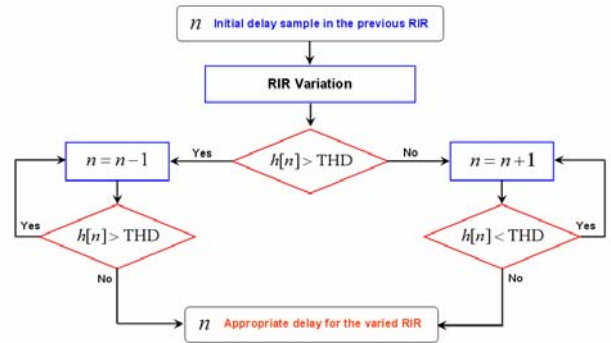


그림 3. 학습 알고리즘 I- 초기 지연과 ER 의 경계

(2) ER 과 LR 의 경계

ER 구간은 앞서 언급한 바와 같이 음원과 수음점의 위치 변화에 가장 민감한 부분이고 LR 구간은 전체적인 실내 음향 특성을 반영하며 음원과 수음점의 위치 변화에 덜 민감한 부분이다. 그러므로 위치 변화에 따른 민감도를 정의하면 두 구간의 경계를 찾는 데 효과적이다. 일반적인 구조 해석 분야에서 민감도는 다음 식 (5)와 같이 정의된다.

$$\psi' = \frac{\psi(r_0 + \Delta r) - \psi(r_0)}{\Delta r}, \quad (5)$$

여기서 ψ' 는 구조의 민감도, ψ 는 주어진 구조의 함수, r_0 와 Δr 은 각각 목적 변수의 초기 상태와 단위 변화량을 나타낸다. 이러한 개념을 목적 변수를 음원과 수음점의 위치 변화로 하여 실내 충격 응답 함수에 적용하기 위해 다음의 그림 4 와 같이 단위 위치의 변화량을 정의하였다.

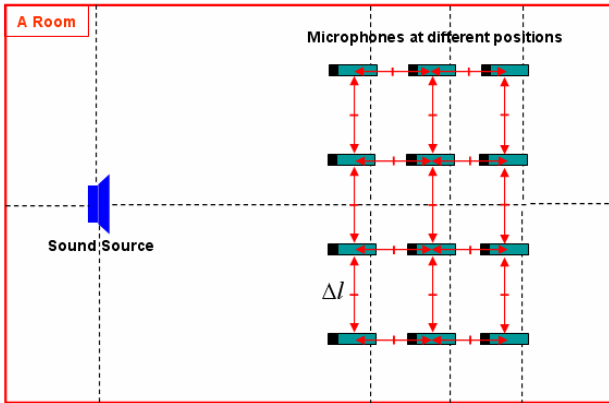


그림 4. 고정된 음원과 여러 위치의 수음점으로 정의된 단위 위치의 변화량 Δl

그림 4 와 같이 정의된 단위 위치 변화량, Δl 을 이용하여 RIR 의 위치 변화에 따른 민감도, h' 을 다음과 같이 정의한다.

$$h'[n] = \frac{|h_{l_0+\Delta l}[n] - h_{l_0}[n]|}{\Delta l}, \quad (6)$$

h_{l_0} 와 $h_{l_0+\Delta l}$ 은 각각 초기 위치의 RIR 과 단위 위치를 변화시킨 후 RIR 를 의미한다.

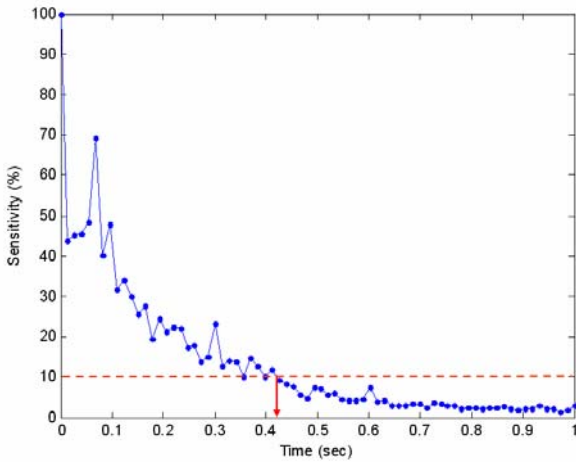


그림 5. 단위 위치 변화에 따른 RIR 의 민감도

위의 그림 5 는 단위 위치 변화, Δl 에 따른 RIR 의 민감도의 예를 보여주고 있다. 이 때 초기 지연 구간의 값들은 첫 번째 학습 알고리즘을 이용하여 제외하였다. 이 그림은 민감도가 가장 큰 값을 100%으로 하여 정규화시킨 결과를 나타낸 것이다. 빨간색 점선으로 나타낸 것은 민감도가 10%일 때를 표시한 것으로 대략 0.42 초 정도를 나타내었다. 본 논문에서는 ER 과 LR 구간의 경계를 10% 민감도를 기점으로 이보다 큰 곳을 ER 구간으로 이보다 작은 구간을 LR 로 정의하였다.

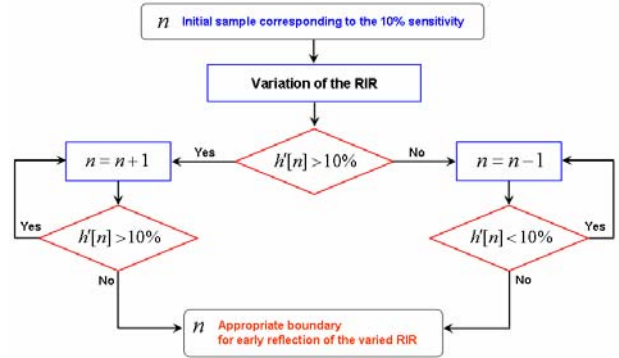


그림 6. 학습 알고리즘 II-ER 과 LR 의 경계

그림 6 은 ER 과 LR 구간의 경계를 정의하는 학습 알고리즘을 소개하였다. 이는 학습 알고리즘 I 과 비슷하지만 RIR 함수값, $h[n]$ 을 대신하여 RIR 의 위치에 따른 민감도, $h'[n]$ 을 미리 정의한 10%와 비교하여 시간에 따라 변하는 RIR 의 ER 과 LR 의 경계를 찾으려 하였다.

4. 제안된 모델의 성능 평가

본 논문에서 제안된 학습 하이브리드 RIR 모델의 성능을 평가하기 위해 다음 그림 7 과 같이 실제 RIR 을 광주과학기술원(GIST) 대강당에서 측정하였다.

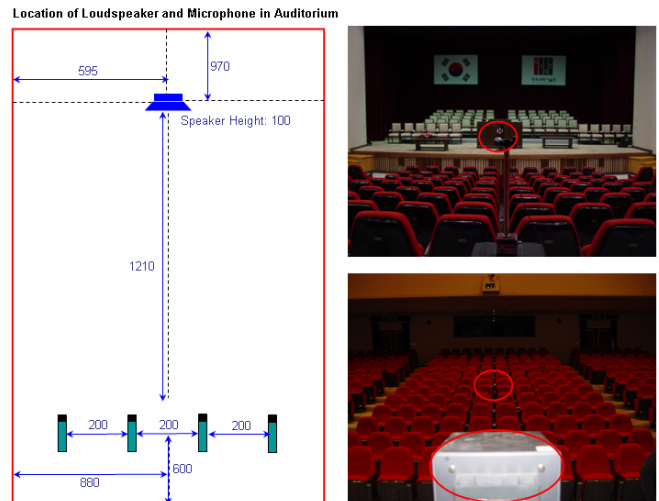


그림 7. GIST 대강당의 RIR 측정

그림에서 빨간색 동그라미는 음원과 마이크를 나타낸 것이며 단위 위치 변화, Δl 은 2 미터로 하여 측정을 하였다. 또한, 성능을 평가의 척도로서 실제 측정한 RIR 과 예측된 RIR 의 오차를 출력 신호간의 차이를 이용하여 검출해 내며, 모델의 정확도를 나타내는데 많이 사용되는 모델링 에러 (Modeling Error)를 선정하였다.

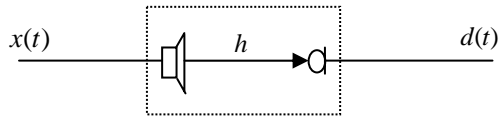


그림 8. 단일 채널 음향 시스템

위의 그림 8 과 같은 단일 채널 음향 시스템을 가정하면 모델링 에러 J (%)는 다음의 식 (7)과 같이 정의된다.

$$J(\%) = \frac{E[e^2(t)]}{E[d^2(t)]} \times 100, \quad (7)$$

입력신호를 $x(t)$ 라고 가정하면 $d(t) = h(t) * x(t)$, $e(t) = \hat{h}(t) * x(t) - d(t)$ 이며 $h(t)$ 는 측정된 RIR 을, $\hat{h}(t)$ 는 특정 모델을 이용하여 표현한 RIR 을 나타낸다. 여기서 * 는 컨벌루션 연산을 의미한다.

그림 7 의 측정을 통해 얻어진 RIR 을 기존의, MA, AR, ARMA 모델을 이용하여 얻은 모델링 에러와 이 때 사용된 가변 파라미터의 개수를 제안된 학습 하이브리드 모델 (Learning Hybrid, LH)과 비교한 결과를 표 1 을 통해 나타내었다.

표 1. 모델에 따른 가변 파라미터의 개수와 모델링 에러 (MA: MA coefficients 의 개수, AR: AR coefficients 의 개수, Delay: delay coefficient 의 개수, MA1: ER 구간의 MA coefficients 의 개수, MA2: LR 구간의 MA coefficients 의 개수)

RIR Model	Variable Parameters		Modeling Error (%)
	Total	Detail	
MA	4000	MA: 4000	0.1
AR	3000	AR: 3000	99.01
ARMA	3000	AR: 1500, MA: 1500	52.37
LH	2626	Delay: 1, MA1: 1745, AR: 430 MA2: 450	3.28

표 1 에 나타난 바와 같이 MA 모델의 경우 가장 작은 모델링 에러를 갖지만 또한 많은 가변 파라미터를 갖기 때문에 이를 처리하는 과정에서 많은 계산량을 야기시킨다. AR 과 ARMA 모델의 경우 3000 개의 상당한 가변 파라미터를 사용했음에도 불구하고 모델링 에러는 개선되지 않았다. 하지만 본 논문에서 제안된 Learning Hybrid (LH) 모델의 경우 가장 적은 개수의 가변 파라미터를 가지며 모델링 에러 또한 상당히 작았다. 그러므로 가변 파라미터와 모델링 에러를 동시에 고려하면 제안된 학습 하이브리드 모델이 가장 효과적인 알고리즘이라고 할 수 있다.

5. 결론

본 논문에서는 학습 알고리즘을 이용한 새로운 하이브리드 RIR 모델을 소개하였다. RIR 은 각각 다른 음향 특성으로 인해 초기 시간지연, ER, LR 의 세 구간으로 구성되어 있다. 각각의 구간은 서로 다른 모델을 이용하여 표현되었으며 시간에 따라 변하는 각 구간의 경계는 학습 알고리즘을 통해 정의되었다. 기존에 사용되던 MA, AR, ARMA 모델과 제안된 학습 하이브리드 모델의 계산량을 결정하는 가변 파라미터의 개수와 모델의 정확도를 나타내는 모델링 에러가 실제로 측정된 RIR 을 기초로 하여 비교되었다. 모델링 에러의 측면에서는 가장 많은 가변 파라미터의 개수에도 불구하고 MA 모델이 가장 정확한 모델로 판명되었다. AR 과 ARMA 모델의 경우 현저하게 큰 모델링 에러로 인해 RIR 전체를 표현하는데 적합하지 않은 모델로 판단되었다. 제안된 학습 하이브리드 모델은 가변 파라미터의 개수로 대변되는 처리 시의 계산량과 모델링 에러를 동시에 고려할 때 가장 효과적인 모델이었다. 또한 이 모델은 초기지연과 ER, ER 과 LR 의 경계 문턱값인 THD=0.005 와 민감도 10%를 조절하면 계산량과 정확도를 응용분야에 따라 적절히 조절할 수 있는 장점도 가지고 있다.

참고문헌

- [1] J. Martin, D. Van Maercke, J. P. Vian, 1993, "Binaural simulation of concert halls: A new approach for the binaural reverberation process," Journal of the Acoustical Society of America, vol. 94, no. 6, pp. 3255-3264.
- [2] J. Laroche, 1989, "A New analysis/synthesis system of musical signals using prony's method - application to heavily damped percussive sounds," in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing 89, pp. 2053-2056.
- [3] J. M. Jot, A. Chaigne, 1991, "Digital delay networks for designing artificial reverberators," in Proc. 90th Audio Engineering Society Convention, Paris, France, Preprint 3030.
- [4] M. Karjalainen, P.A.A. Esquef, P. Antsalo, A.M. Akivirta, V. Valimaki, 2002, "Frequency-Zooming ARMA modeling of resonant and reverberant systems," Journal of Audio Engineering Society, vol. 50, no. 12, pp. 1012-1029.
- [5] http://www.ece.nmsu.edu/~pdeleon/BSS/Impulse_Responses/README_Impulse_Responses.pdf.