

# 인식된 방문객을 위한 지향성 오디오 시스템

김민준, 김민기, 윤도현, 홍준표, 김시현

수원대학교 정보통신공학과

{minjun9713, chkm7847, hiho0718, hjp0232, seehyun}@suwon.ac.kr

## Personalized Directional Audio System for Identified Visitors

Min-Jun Kim, Min-Ki Kim, Do-Hyeon Yoon, Joon-Pyo Hong, Seehyun Kim

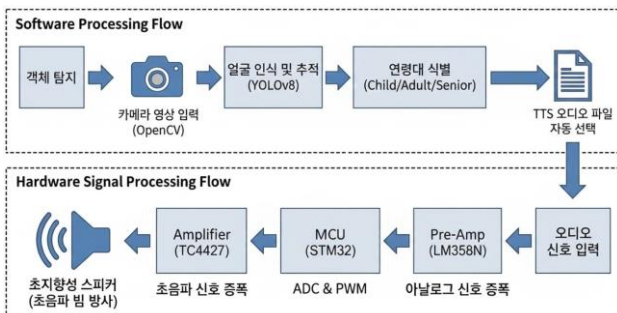
Dept. of Information and Communication Engineering, The University of Suwon

### 요약

본 논문은 인공지능의 객체 인식 기술과 초음파 기반의 지향성 음향이라는 두 가지 기술을 융합한 지능형 맞춤 도슨트 시스템을 제안한다. 이 시스템은 객체 인식 모델을 통해 관람객의 위치와 연령대를 실시간으로 식별하고, 이를 변조기 및 트랜스듀서 어레이와 연동하여 연령별 맞춤 해설을 지향성 오디오 신호로 전달한다. 성능평가 실험을 통해 정면 신호는 좌우 30도 측면 대비 약 30dB 높은 음압 이득을 보여 우수한 지향성을 검증하였으며, 기계 학습 기반 객체 인식률은 83.27%의 정확도를 달성하였다. 본 연구는 개방된 공간에서도 관람객에게 독립적인 음향을 제공할 수 있음을 시사하며, 향후 다양한 공공장소에서의 개인화된 정보 전달 기술로 활용될 가능성을 제시한다.

### I. 서론

현대의 박물관 및 미술관은 단순한 전시 관람을 넘어 능동적인 상호작용을 지향하고 있으나[1], 기존 도슨트의 획일적인 정보 전달 방식[2]과 개방된 공간에서 발생하는 음향 간섭 현상[3]은 관람객의 몰입을 저해하는 주요 한계로 지적된다. 이에 본 논문은 인공지능 기반의 객체 인식 기술과 초음파 변조 기술을 융합함으로써 개별 관람객에게 맞춤형 서비스를 제공하는 새로운 도슨트 시스템을 제안한다. 본 시스템은 ‘인지’, ‘판단’, ‘전달’의 세 단계 순으로 동작한다. 먼저 ‘인지’ 단계에서는 카메라에 입력되는 영상을 기반으로 YOLOv8과 EfficientNet-B0 모델을 적용하여 관람객의 위치와 연령대를 실시간으로 식별하고, ‘판단’ 단계에서는 식별된 연령대에 부합하는 맞춤형 해설 콘텐츠를 자동으로 선별한다. 마지막으로 ‘전달’ 단계에서는 STM32를 통해 40kHz 초음파 캐리어에 오디오 신호를 변조하고, 이를 4개의 트랜스듀서 어레이로 방사하여 공기의 비선형성을 이용한 자가 복조를 유도함으로써 목표 대상에게만 명확한 소리를 전달한다. 본 연구는 이러한 통합 시스템을 통해 관람객의 눈높이에 맞는 안내 서비스를 구현하는 데 그 목적이 있다.

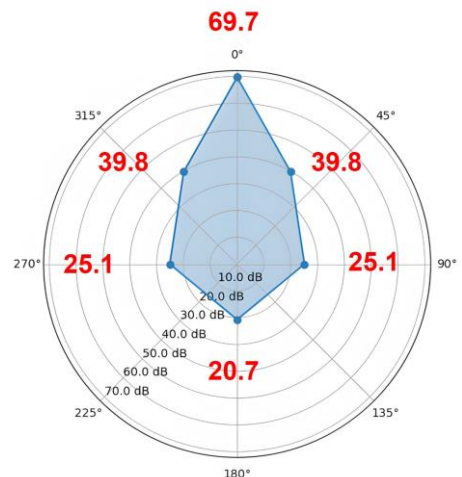


[그림 1] 시스템 전체 흐름도

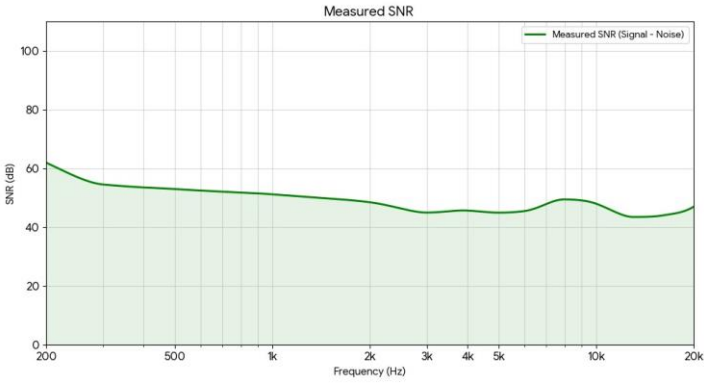
### II. 본론

#### 1. H/W 검증 결과

스피커 구동 결과, [그림 2]와 같이 정면(0°) 방향에서 최대 음압 레벨 69.7dB가 측정되었으며,  $\pm 30^\circ$  이상의 측면 영역에서는 39.8dB 이하로 급격히 감소하는 뚜렷한 지향 특성을 보였다. 이는 정면 대비 약 30dB 감소한 수치로, 소리가 측면으로 거의 확산하지 않음을 의미한다. 후면(180°) 방향으로 갈수록 음압이 현저하게 감소하는 이러한 결과는 본 시스템이 일반 스피커와 달리 특정 방향으로만 음향 에너지를 집중시킬 수 있음을 검증함과 동시에, 공기의 비선형 복조가 제한된 방향에서만 효율적으로 발생함을 보여주는 데이터이다.



[그림 2] 각도별 음압 수준 (SPL, 단위: dB)



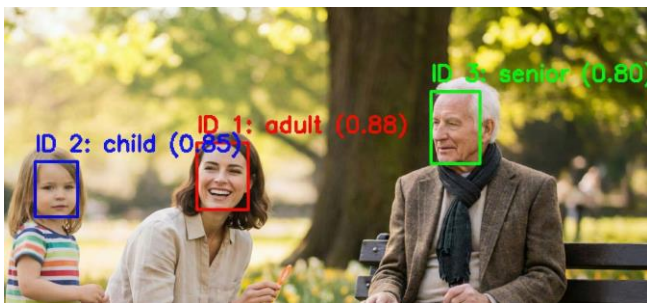
[그림 3] 주파수 대역별 SNR 측정 결과

[그림 3]은 200Hz부터 20kHz에 이르는 가청 주파수 대역에서 신호 대 잡음비를 측정한 결과이다. 전체 구간에 걸쳐 최소 43dB 이상의 우수한 성능이 유지됨을 확인하였다. 본 검증에 사용된 지표인 SNR은 출력된 신호 전력과 잡음 전력의 비율을 로그 스케일로 나타낸 것으로, 수치가 높을수록 잡음 대비 유효 신호의 세기가 강하여 정보 전달의 명료도가 높음을 의미한다. 측정 결과에 따르면 저주파 대역인 200Hz 인근에서 약 62dB의 최대 SNR을 기록한 이후 주파수가 증가함에 따라 완만한 하향 곡선을 그리지만, 인간의 청각이 민감하게 반응하는 2kHz~4kHz 대역과 아울러 고주파 대역에서도 43dB 수준을 견고하게 상회하며 급격한 성능 저하가 없는 평탄한 특성을 보였다. 이러한 결과는 개발된 스피커가 내부 회로 노이즈나 외부 간섭에 대해 높은 내성을 지니고 있음을 객관적으로 지표화하며, 결과적으로 타겟 객체에 왜곡없는 고품질의 사운드를 전달할 수 있는 안정적인 동작 성능과 하드웨어적 신뢰성을 확보했음을 입증한다.

## 2. S/W 검증 결과

--- Classification Report ---				
	precision	recall	f1-score	support
child	0.88	0.96	0.92	730
adult	0.91	0.82	0.86	3013
senior	0.63	0.78	0.70	998
accuracy			0.83	4741
macro avg	0.81	0.85	0.83	4741
weighted avg	0.85	0.83	0.84	4741

[그림 4] 객체 인식 성능 지표



[그림 5] 본 시스템의 객체 탐지 및 식별 결과

개발 완료된 객체 인식 모델의 실효성을 검증하기 위해 4,741건의 테스트 데이터셋을 기반으로 한 성능 평가와 [그림 5]와 같이 다수의 사람을 대상으로 한 식별 테스트를 병행한 결과, 시스템이 타겟 대상을 안정적으로 식별함을 확인하였다. [그림 4]가 그러한 사실을

뒷받침하는 근거이다. 우선 정량적 지표인 분류 리포트를 분석한 결과, 전체 정확도는 83%를 기록하였으며 데이터 불균형을 고려한 가중 평균 F1-score는 0.84로 산출되어 모델의 견고함이 입증되었다. 여기서 정밀도는 탐지된 결과가 실제 타겟과 일치하는 정도를, 재현율은 대상 객체를 놓치지 않고 찾아내는 민감도를 의미하며, 이 둘의 조화 평균인 F1-score는 모델의 종합적인 신뢰 척도를 나타낸다. 검증 차원에서 시스템이 실제 인식한 [그림 5]의 다수 객체에 표시된 바운딩 박스는 Confidence Score로, 각 연령대의 사람에 대해 우수한 신뢰도를 반영하고 있다. 결론적으로, 통계적 수치 지표와 실제 환경에서의 성공적인 식별 결과의 일치성은 본 시스템이 연령별 객체 특징을 명확히 구분하여 초지향성 사운드를 송출하기 위한 전처리 단계로서 충분한 판별 성능과 기술적 완성도를 확보하고 있음을 객관적으로 시사한다.

## III. 결 론

본 연구에서는 방문객을 인식하여 맞춤형 해설을 선택하고 이를 초음파 기반의 지향성 스피커를 통해 대상 고객에게만 선명한 오디오 해설을 전달하는 개인화된 도슨트 시스템을 개발하였다. 지향 각도별 SPL 및 SNR 데이터에 근거하여 하드웨어의 지향 음향 성능을 입증하였고, 객체 인식 모델의 성능 지표를 통해 개인화 성능을 확인하였다. 이러한 기술적 성과는 전시장 뿐만 아니라 공공 안전 시스템, 개인 엔터테인먼트 등 비접촉 개인화 서비스로의 확장 가능성을 보여주며[4][5], 기존 연구에서 보고된 파라메트릭 배열 기반 디지털 빔 스티어링 성과[6]를 바탕으로 향후 이동 객체 추적 기술로의 발전 방향을 제시한다. 결론적으로 본 연구는 개인화와 공간 음향 제어를 실증적으로 구현하여 차세대 미디어의 기술적 토대를 확보했다는 점에서 의의가 있다.

## 참 고 문 헌

- [1] 나서현·변영철·이동철, 「증강현실 기술을 활용한 사용자 중심의 체험형 박물관 전시콘텐츠 구현」, 『한국정보기술학회논문지』 제21권 제8호, 한국정보기술학회, 2023, 211-218쪽.
- [2] 어일선, 「박물관 전시에서의 디지털 매체 적용에 관한 고찰 : 서울시립대학교 기획전시를 중심으로」, 『한국콘텐츠학회논문지』 제10권 제9호, 한국콘텐츠학회, 2010, 212-219쪽.
- [3] 서정석·김재수, 「대형전시공간의 건축음향설계」, 『대한설비공학회 학술발표대회논문집』, 대한설비공학회, 2004, 115쪽.
- [4] 전계현·김태환·강신욱, 「지향성스피커를 활용한 방재성능 향상에 관한 연구」, 『한국재난정보학회 논문집』 제12권 제3호, 한국재난정보학회, 2016, 235-241쪽.
- [5] 장성택 외 3인, 「지향성 스피커를 활용한 컴퓨터 비전 기반 이용자 추적 음향 시스템 : 사운드 영역과 사용자 행동유도 연구를 위한 시스템구축」, 『한국HCI학회 학술대회』, 한국HCI학회, 2013, 786-788쪽.
- [6] 이재일, 「파라메트릭 배열 기반 초음파 시스템 개발 및 응용」, 제주대학교 박사학위논문, 2015.