

음원 내 보컬 주파수 대역 분석에 기반한 음향기기 추천시스템⁺

(A system for recommending audio devices based on frequency band analysis of vocal component in sound source)

김정현¹⁾, 석철민¹⁾, 김민주¹⁾, 김수연²⁾*

(Jeong-Hyun Kim, Cheol-Min Seok, Min-Ju Kim, and Su-Yeon Kim)

요약 음원 스트리밍 서비스와 Hi-Fi 시장이 성장함에 따라 다양한 음향기기들이 출시되고 있다. 이로 인해 소비자들의 제품 선택에 대한 폭은 넓어졌지만 자신의 음악적 취향과 일치하는 제품을 찾기는 더욱 어려워졌다. 본 연구에서는 사용자가 선호하는 음원으로부터 보컬 성분을 추출하고 이를 토대로 사용자에게 가장 적합한 음향기기를 추천하는 시스템을 제안하였다. 이를 위해 먼저 원본 음원을 Python의 Spleeter Library를 통해 분리하여 보컬 음원을 추출하고 제조사의 음향기기의 주파수 대역 데이터를 수집한 결과를 각각 격자 그래프로 나타내었다. 추출한 보컬 음원의 주파수 대역과 음향기기의 주파수 대역 측정치 데이터를 비교하기 위한 지표로서 Matching Gap Index(MGI)를 제안하였다. 산출된 MGI 값을 토대로 사용자 선호도와 유사도가 가장 높은 음향기기를 추천한다. 추천 결과는 음향 전문업체에서 제공하는 장르별 Equalizer 데이터를 이용하여 검증하였다.

핵심주제어: 추천시스템, 보컬 추출, 주파수 대역 분석, 음향기기 추천, 음원 이퀄라이저

Abstract As the music streaming service and the Hi-Fi market grow, various audio devices are being released. As a result, consumers have a wider range of product choices, but it has become more difficult to find products that match their musical tastes. In this study, we proposed a system that extracts the vocal component from the user's preferred sound source and recommends the most suitable audio device to the user based on this information. To achieve this, first, the original sound source was separated using Python's Spleeter Library, the vocal sound source was extracted, and the result of collecting frequency band data of manufacturers' audio devices was shown in a grid graph. The Matching Gap Index (MGI) was proposed as an indicator for comparing the frequency band of the extracted vocal sound source and the measurement data of the frequency band of the audio devices. Based on the calculated MGI value, the audio device with the highest similarity with the user's preference is recommended. The recommendation results were verified using equalizer data for each genre provided by sound professional companies.

Keywords: Recommendation System, Vocal Extraction, Frequency Band Analysis, Audio Device Recommendation, Music Equalizer

* Corresponding Author: sykim@daegu.ac.kr

+ 이 논문은 2021학년도 대구대학교 학문후속세대 연구과제로 수행되었음.

Manuscript received December 11, 2022 / revised

December 24, 2022 / accepted December 26, 2022

1) 대구대학교 대학원 IT융합공학과

2) 대구대학교 컴퓨터정보공학부, 교신전자

1. 서론

음향기기 시장은 지속적으로 성장하고 있다. 음향기기 제조사들은 매년 새로운 제품을 출시하고 있으며 삼성전자와 같은 스마트폰 제조업체 또한 하만카돈을 인수하여 음향 시스템의 품질을 높이는 등 음향 기술에 대한 관심을 기울이고 있다. 이처럼 새로운 음향기기 제품들이 지속적으로 출시되고 있지만 음향기기에 대한 전문적인 지식이 없는 일반 소비자는 자신에게 적합한 음향기기를 판별하기가 쉽지 않다. 음향기기 제조사들은 제품의 특성을 객관적으로 나타내기 위해 주파수 대역 측정치를 제공하고 있으나 전문지식이 없는 일반 사용자의 입장에서는 쉽게 정보를 파악하기 어렵다. 이처럼 일반 구매자가 음향기기의 특성을 제대로 이해하지 못하고 제품을 구매한 뒤 자신의 청음 스타일과 다를 경우 구매 불만족으로 이어지는 경우가 발생한다.

본 연구는 사용자가 선호하는 음원으로부터 보컬 성분을 추출하여 주파수 대역을 분석하고 음향기기와의 유사도를 산출함으로써 최적의 음향기기를 추천하기 위한 목적으로 수행되었다. 음향기기는 통상적으로 음악을 듣기 위한 기기와 음향을 이용하여 계측이나 탐사를 하는 기기를 모두 포함하지만, 본 연구에서는 이어폰, 헤드폰 등과 같이 음악을 듣기 위한 리시버 형태의 개인용 기기로 한정하여 사용한다.

연구의 범위는 연주곡이 아닌 보컬 중심의 음원을 대상으로 하였으며, 음원으로부터 아티스트를 특정할 수 있는 가장 뚜렷한 특성인 보컬 성분을 분리하여 주파수 비교 실험에 활용하였다. Spleeter Library를 활용하여 보컬 성분을 추출하였으며 이를 음향기기 주파수 대역별 응답치와 비교하였다. 보컬 성분과 주파수 대역 응답치 그래프는 5x5 격자 그래프의 형태로 나타내었으며 두 그래프를 비교한 결과의 차이를 MGI(matching gap index)라는 적합도 지수로 나타내었다. 따라서 MGI가 낮을수록 사용자의 선호에 가장 적합한 음악 스타일을 제공하는 음향기기인 것으로 설정하였으며, 결과를 음향 전문업체의 Equalizer를 통해 검증하였다.

2. 관련 연구

2.1 추천시스템

추천시스템은 다양한 분야에서 활용되고 있다. 유튜브, 넷플릭스와 같은 영상 스트리밍 업체들은 시청 기록을 통해 사용자가 선호할 만한 영상을 선별하여 추천한다. 쿠팡, 아마존과 같은 소셜커머스 업체들은 사용자의 구매 상품을 분석하여 잠재적으로 구매할 만한 상품을 상단에 노출하여 구매를 유도한다. 이러한 추천시스템에서 주로 사용되는 방법론은 Fig. 1과 같이 콘텐츠 기반 필터링(content based filtering)과 협업 필터링(collaborative filtering), 그리고 하이브리드 필터링(hybrid filtering)으로 분류된다.

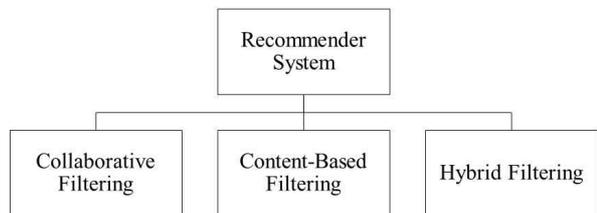


Fig. 1 Recommender System

2.1.1 콘텐츠 기반 필터링

콘텐츠 기반 필터링은 사용자의 제품이나 서비스 정보를 제공받아 콘텐츠의 속성과 유사한 정보를 파악하여 또 다른 콘텐츠를 추천하는 기법이다(Wu and Chen, 2000). 콘텐츠 기반 필터링은 사용자 프로파일과 콘텐츠에 대한 정보를 통해 추천하기 때문에 새로운 콘텐츠의 정보나 사용자의 반응을 추천에 이용할 수 있다. Lee et al.(2020)의 연구에서는 형태소 추출을 통한 영상 추천 방법을 제시하였고, Yoo and Jeong(2015)은 유튜브 상에서 사용자의 소셜 카테고리 추출을 이용한 추천 알고리즘을 제안하였다. Park et al.(2020)의 연구에서는 SNS 사용자가 게시한 사진 속성을 통해 사용자가 선호하는 장소를 추천하는 알고리즘을 도출하였다. Lee(2020)은 사용자가 미평가한 항목들의 평가등급을 예측하는 시스템을 제시하였고 Jeon and Ahn(2015)은 사용자 리뷰를 분석하여 정성적 데이터를 활용하는 추천시스템을 제안하였다. Hwang et al.(2022)는 취미 물품 대여 플랫폼에 추천 시스템을 반영하여 개인화 서비스를 제공하는 시스템

템을 구축하였다. Hong et al.(2022)는 영화 리뷰를 대상으로 LDA 기법을 이용하여 상품 및 고객 속성을 총평점 데이터와 통합한 지능형 추천시스템을 제안하였으며, Lee and Kim(2013)은 스마트 TV 상의 개인화된 콘텐츠를 추천하는 지능형 시스템을 설계 및 구현하였다.

넷플릭스가 사용자의 시청 기록을 통해 같은 시리즈나 비슷한 장르의 작품을 추천해주는 것도 콘텐츠 기반 필터링을 활용한 것이다. Fig. 2는 넷플릭스의 콘텐츠 기반 필터링 예시이다.



Fig. 2 Content-Based Filtering Example (netflix.com)

2.1.2 협업 필터링

협업 필터링은 사용자와 유사한 성향을 가진 타 사용자들이 선호했던 콘텐츠를 사용자에게 추천해주는 시스템이다(Sharma and Gera, 2013). Fig. 3은 협업 필터링에 대한 개념을 보여주고 있다.

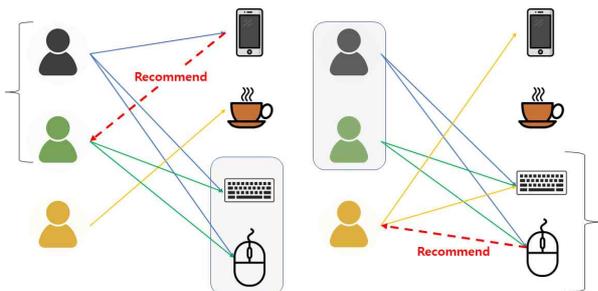


Fig. 3 Collaborative Filtering

협업 필터링은 사용자의 평가나 선호도 지표를 통해 누적된 콘텐츠 간의 유사도를 기반으로 추천하기 때문에 음원 서비스, 온라인 마켓 등에서 주로 사용된다. Kim et al.(2011)의 연구에서는 영화의 장르 유사도와 선호 장르를 이용한 추천시스템을 제안하였고, Park et al.(2020)의 연구에서는 사용자 연령과 국

내 거주 기간 정보를 통해 사용자 학습 레벨에 맞는 회화 스크립트를 추천하였다. Yun and Youn(2020)는 상품의 리뷰를 분석하고 이를 가중치로 사용한 협업 필터링 추천 기법을 제안하였다. 대표적인 협업 필터링 추천 서비스로 Amazon을 들 수 있다. Amazon은 아이템 중심의 방법을 이용하여 사용자에게 상품을 추천한다.

아이템 중심 방법은 아이템들 간 상관관계를 도출하여 사용자의 최근 데이터를 토대로 사용자 선호도를 예측하는 방법이다. 한 사용자가 콘텐츠를 구매하면, 타 사용자가 해당 콘텐츠와 같이 구매했던 콘텐츠를 추천하며, 유사한 특성을 갖는 사용자들이 구매했던 비슷한 콘텐츠를 추천한다. Fig. 4는 아마존의 아이템 중심 방법을 적용한 추천 결과 예시이다.



Fig. 4 Amazon Recommendation Example (amazon.com)

2.2 음향 분야 추천시스템

추천시스템은 음향 관련 분야에서도 널리 활용되고 있다. Oh and Lee(2019)의 연구에서는 연관규칙 (association rule)을 적용하여 사용자의 현재 위치 정보와 시간 정보에 대한 관계성 및 규칙들을 기반으로 실시간 상황에서 적합한 음악을 추천하는 시스템을 제안하였다. Chung and Kim(2018)은 음악 데이터와 이용자의 선호도를 분석하여 개인 맞춤형 콘텐츠 기반 음악추천 모델을 제안하였고, Choi and Hwang (2012)의 연구는 사용자의 SNS 텍스트를 통해 감정을 분석하고 분석한 감정에 적합한 음악을 추천하는 기법을 제안하였다. Kim et al.(2013)는 사용자의 청취 습관과 태그 데이터를 토대로 사용자 프로파일을 생성한 다음 협업 필터링 기법을 적용하여 음악을 추천하는 하이브리드 필터링 음악 추천시스템을 개발

하였다. Kim and Lim(2015)의 연구에서는 특정 상황에서 사용자가 선택한 음악으로부터 감정을 자극하는 상황 요인과 감정 상태들 간의 상관관계를 분석하여 감정 상태에 적합한 음악을 추천하는 방법을 제안하였다.

추천시스템을 활용하는 음원 서비스로는 'Youtube Music'을 예로 들 수 있다. Youtube Music은 사용자에게 선호하는 아티스트를 선택하도록 한다. Fig. 5는 Youtube Music 선호 아티스트 선택 화면 예시이다.



Fig. 5 Youtube Artist Selection Example (youtube.com)

사용자가 아티스트를 선택하고 나면 아티스트의 대표곡들을 우선적으로 추천하고 해당 아티스트와 유사한 음향적 특성을 가진 아티스트와 해당 아티스트를 좋아하는 타 사용자가 선호하는 음악을 추천한다.

이상에서 살펴본 바와 같이 사용자에게 적합한 음원을 추천하는 연구는 수행된 바 있으나 음원을 재생하는 음향기기를 추천하는 분야의 연구는 아직 이루어지지 않고 있다. 이에 본 연구에서 사용자에게 가장 적합한 음향기기를 추천하는 시스템을 제안하고자 한다.

3. 음향기기 추천시스템

본 연구에서는 사용자의 청취 데이터에 기반하여 가장 적합한 음향기기를 매칭하고 추천하는 방법을 제안한다. 이를 위해 먼저 사용자가 선호하는 아티스트의 특성을 정의한다. 사용자 선호 아티스트의 특성은 해당 아티스트의 고유한 속성값으로 설정해야 하며 본 연구에서는 아티스트의 보컬로 설정하였다. 특정 아티스트의 보컬 스타일은 다른 아티스트와 뚜렷이 구분될 수 있다. 아티스트 보컬의 음역대별 음량

을 5점 척도로 수치화하여 아티스트의 Feature로 정의한다. 또한 음향기기의 음역대별 음량 측정치를 마찬가지로 수치화하여 음향기기의 Feature로 정의한 후 최종적으로 아티스트의 Feature와 음향기기의 Feature를 비교하여 적합도를 산출함으로써 사용자가 가장 선호할만한 음향기기를 도출한다. 제안된 추천시스템은 Fig. 6과 같은 작업 절차로 이루어진다.

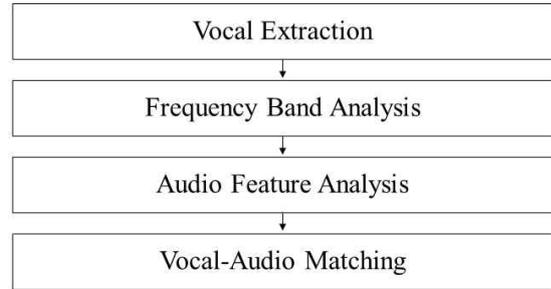


Fig. 6 Recommendation Procedure

3.1 음원 내 보컬 추출

일반적으로 음원은 여러 성분들이 동시에 녹음되어 하나의 파일로 저장된다. 기타, 베이스, 드럼, 피아노 등의 악기들과 아티스트의 보컬이 하나의 파일 안에 묻혀져 있다. 이 중에서 아티스트의 보컬은 고유한 주파수 대역을 가지게 되어 해당 아티스트를 특정할 수 있는 요소로 활용될 수 있다. 이러한 보컬 특성을 명확히 파악하기 위해 아티스트의 대표 음원에서 보컬 특성을 분리하고 악기 연주로 생성되는 배경음 성분을 제거하여 추출하는 작업이 요구되며 추출된 보컬 음원을 수치화하여 아티스트의 Feature로 정의할 수 있다.

음원 내 보컬 영역을 추출하기 위해 음원 분리 작업을 진행한다. 음원 분리는 여러 음원이 섞여 있는 오디오 신호로부터 개별 음원을 추출하는 기술이다. 동영상에서 사람의 목소리만 추출하거나 음악 파일에서 한 악기의 소리만을 추출하는 것을 예로 들 수 있다. 음원 분리를 진행하기 위해 프랑스 온라인 음악 스트리밍 서비스 'Deezer'에서 제공하는 Python 라이브러리 'Spleeter Library'를 활용하였다. Fig. 7은 Spleeter의 제작사 'Deezer'에서 소개되는 Spleeter의 개념도이다. Spleeter는 Tensorflow를 통한 사전 훈련 모델을 보유하고 있다(Hennequin et al., 2020).

Spleeter Library 활용 작업은 음원 파일을 Google Colab 환경에 마운트하여 진행하였다.



Fig. 7 Spleeter Library (deezer.io)

Spleeter Library는 Table 1과 같이 Stem 수를 각각 2개, 4개, 5개씩으로 하는 3단계 분리 메커니즘을 제공한다.

Table 1 Spleeter Library Stems: 3-Level Separation

Number of Stems	Components of Music Source
2-Stems	Vocal, Accompaniment
4-Stems	Vocal, Drum, Base, Guitar
5-Stems	Vocal, Drum, Piano, Base, Guitar

본 연구에서 제안하고자 하는 방법은 음원의 보컬을 아티스트의 특성으로 제한하였기 때문에 보컬 성분과 반주 성분으로 분리하는 2-Stems을 이용하였다. 실험을 위해 여성 보컬 밴드 'Fit for Rivals'의 'freakmachine.mp3' 파일을 사용하였다. Fig. 8은 음원이 분리된 파일 목록이다.



Fig. 8 Music Source Separation

분리된 음원 중 'vocals.wav' 파일이 음원의 보컬 성분을 추출한 것이며 이를 주파수 대역 분석 단계에서 활용한다.

3.2 주파수 대역 분석

추출한 보컬 성분의 주파수 대역을 추출하고 분석한다. 주파수 대역 분석을 위해 Librosa Library를 활용하였다(McFee et al., 2022). Librosa Library는 오디오 분석에 주로 활용되는 라이브러리로, Waveform을 시각화하고 변형하는 기능을 제공한다. 추가적으로 Numpy, Matplotlib, IPython 등의 라이브러리를 활용하였다.

보컬 음원 주파수 대역 분석에 앞서 보컬 분리가 제대로 이루어졌는지 검증하기 위해 원본 음원과 보컬 성분의 진폭(amplitude) 데이터를 비교하였다. 시간대에 따른 각각의 진폭 추이를 표시하기 위해 Matplot Library를 사용하여 데이터를 시각화하는 코드를 작성하였다. 실험을 위해 아래와 같이 원본 음원인 'freakmachine.mp3' 파일과 보컬을 추출한 'vocals.wav' 파일을 구분하여 각각 입력으로 사용하였다.

```

/* original music source */
wav,sr = librosa.load('/MyDrive/freakmachine.mp3')
/* vocal component */
wav,sr = librosa.load('/MyDrive/vocals.wav')
    
```

Fig. 9와 Fig. 10은 각각 원본 음원의 시간대별 진폭과 보컬 성분을 추출한 음원의 시간대별 진폭을 나타낸 그래프이다. 두 그래프에서 x축은 음원의 재생 시간대, y축은 진폭을 의미한다.

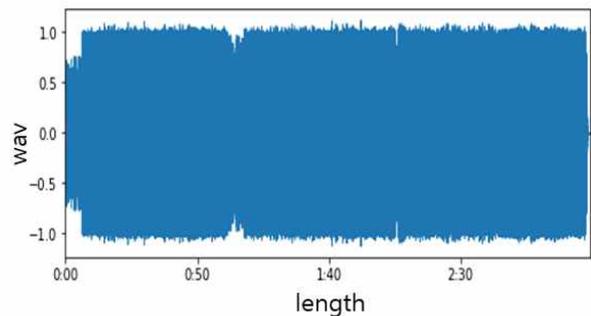


Fig. 9 Amplitude by Time of Original Music Source

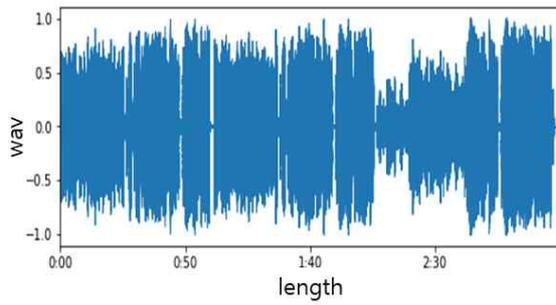


Fig. 10 Amplitude by Time of Vocal Component

두 그래프를 비교해 보았을 때 원본 음원이 진 구간에서 진폭이 일정한 패턴으로 나타난 것에 비해 보컬 음원에서는 반주만 연주되는 간주 구간에서 진폭이 작게 나타난 것을 확인할 수 있다.

다음으로 보컬 성분만을 포함하는 음원의 강조된 주파수 대역을 파악하였다. 이를 위해 보컬 음원의 각 주파수 대역에서의 진폭을 스펙트로그램으로 표시해주는 코드를 작성하였다. x축과 y축을 각각 시간대와 주파수 대역으로 설정하고 색상을 이용하여 진폭을 시각화하였다.

```
librosa.display.specshow(log_spectrogram, sr=sr, hop_length=hop_length, x_axis='time', y_axis='log')
```

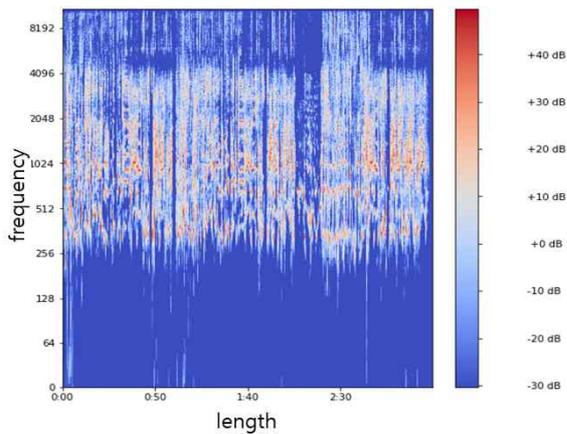


Fig. 11 Frequency-Amplitude Spectrogram of Vocal Component

Fig. 11은 보컬 성분을 추출한 음원의 시간대와 주파수 대역별 진폭을 시각화한 것이다. 색상

이 청색에 가까울수록 진폭이 낮음을 의미하고, 적색에 가까울수록 진폭이 높음을 의미한다. 해당 보컬 음원에서는 1000Hz 전후의 주파수 영역에서 높은 진폭이 출력되었다.

실험의 다음 단계에서 해당 정보를 활용하기 위해 색상과 시간대 축을 제외한 2차원 그래프로 표현하는 코드를 Fast Fourier Transform(FFT) 알고리즘을 이용하여 작성하였다. x축을 주파수 대역, y축을 진폭으로 설정하고 비교를 위해 원본 음원과 보컬 음원에 대해 각각 코드를 적용하여 시각화하였다.

```
fft = np.fft.fft(wav)
m = np.abs(fft)
f = np.linspace(0, sr, len(m))
lf = f[:int(len(f)/2)]
lm = m[:int(len(m)/2)]
```

원본 음원과 보컬 성분을 추출한 음원에 대해 위 코드를 실행한 결과는 각각 Fig. 12, Fig. 13과 같다. 보컬 추출 음원에서는 해당 아티스트의 가성 주파수 미만의 주파수 대역에서 낮은 진폭을 보였다. 가성 주파수는 사람이 낼 수 있는 범위의 주파수 대역을 의미하며 100Hz~6KHz의 범위를 갖는다. 대조적으로 보컬의 주요 음역대인 1000Hz 주변의 주파수 영역은 가장 강조되어 나타났다. 실험에 사용된 음원이 여성 보컬의 곡임과 동시에 Rock 장르의 곡이기 때문에 비교적 높은 음역대가 강조된 그래프가 형성된 것으로 보인다.

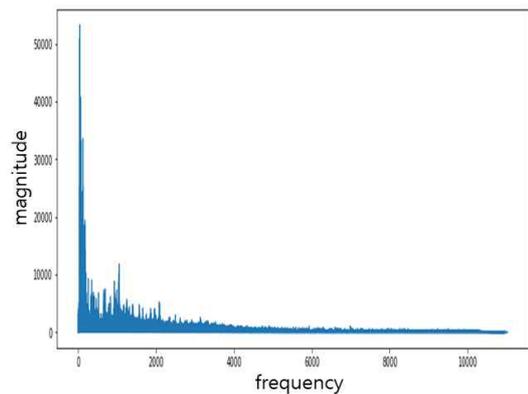


Fig. 12 Frequency-Amplitude Graph of Original Music Source

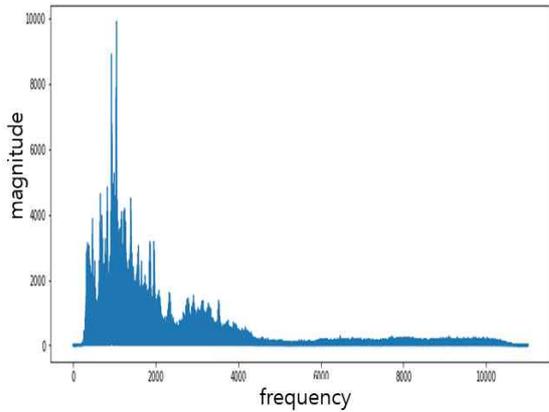


Fig. 13 Frequency-Amplitude Graph of Vocal Component

상기와 같은 과정을 통해 임의로 선정된 아티스트 A, B의 Feature를 5x5 격자 그래프 형태로 표현하였고 그 결과는 Fig. 14와 같다. 전체 범위가 아닌 보컬 성분이 포함되는 가성 주파수 대역을 극저음, 저음, 중음, 고음, 초고음의 5단계로 구분하였고 각 음역대는 그래프의 x축에 나타나 있다. 격자의 범위는 5점 척도를 사용하며 칸의 개수가 적을수록 진폭이 낮고 칸의 개수가 많을수록 진폭이 높음을 의미한다. 즉 한 칸은 최소 진폭, 다섯 칸은 최대 진폭을 나타낸다.

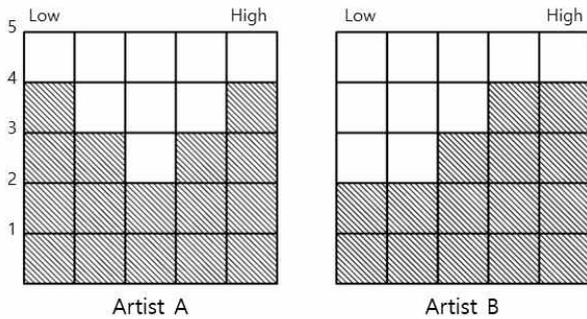


Fig. 14 Frequency-Amplitude Grid Graph by Artists

아티스트 A의 Feature는 극저음 4, 저음 3, 중음 2, 고음 3, 초고음 4 척도로 산출되었고, 아티스트 B의 Feature는 극저음 2, 저음 2, 중음 3, 고음 4, 초고음 4 척도로 산출되었다. 이 단계에서 도출된 Feature들을 보컬-음향기기 매칭 적합도 산출에 활용한다.

3.3 음향기기 Feature 분석

본 단계에서는 음향기기들의 Feature를 파악하여 수치화한다. 음향기기는 제작사의 설계에 따라 주파수마다 출력되는 음량이 각각 다르게 제작되며 저음이 감소하고 고음이 강조된 음향기기, 중음을 강조한 음향기기 등과 같이 다양한 유형의 음향기기로 구분된다. Fig. 15는 대표적인 몇 가지 음향기기의 주파수 대역별 응답치 그래프이다. 그래프에서 x축은 주파수 대역(Hz)을 의미하며, y축은 응답치(dB)를 나타낸다. 녹색 곡선은 음원의 출력값 자체를 나타내며, 청색 곡선은 외부상황을 고려해 보정된 값을 의미한다. 본 연구에서는 음원의 오리지널 출력값 자체를 실험에 사용하였다.

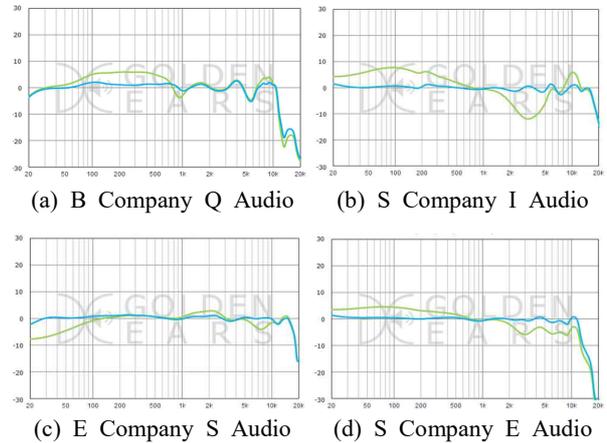


Fig. 15 Frequency Response Graph of Audio Devices (goldenears.com)

음향기기마다 출력되는 주파수 대역별 음량이 상이하므로 같은 음원을 재생하더라도 음향기기에 따라 출력이 달라지며 어떤 음향기기가 해당 음원과 가장 잘 부합하는지를 구분할 수 있다.

실험을 위해 음향기기의 주파수 대역 응답치를 아티스트 Feature와 동일한 방법으로 극저음, 저음, 중음, 고음, 초고음 5단계로 구분하여 5x5 격자 그래프로 매핑한다. 음역대의 구분에 따른 정확한 범위는 기관이나 주체에 따라 다소 차이가 있지만 대부분 5단계 분류를 사용한다. 본 연구에서는 극저음 0~62Hz, 저음 62~250Hz, 중음 250~1KHz, 고음 1~4KHz, 초고음 4KHz 이상으로 설정하였으며, 모든

음원에 대해 같은 기준을 적용하였다. Fig. 16은 서로 다른 6종의 음향기기의 주파수 대역별 응답치를 5x5 격자 그래프 형태로 나타낸 것이다.

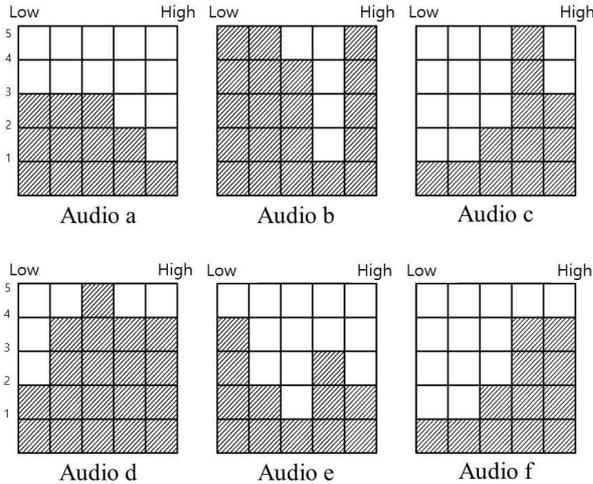


Fig. 16 Frequency-Amplitude Grid Graph by Audio Devices

3.4 보컬-음향기기 Feature 매칭 적합도 산출

산출된 아티스트 보컬 Feature와 음향기기 Feature 간 적합도를 산출한다. 본 연구에서는 보컬과 음향기기 매칭 적합도를 산정하기 위하여 MGI(matching gap index) 지표를 제안하였다. MGI는 아티스트의 보컬 주파수 대역과 음향기기의 주파수 응답치 격자 그래프를 비교하여 차이를 보이는 영역의 수를 합친 지표로서 구체적으로는 정수화된 아티스트 Feature와 음향기기 Feature의 차이를 합산한 값이다. 즉, 아티스트 Feature와 음향기기 Feature를 비교하여 차이가 발생한 영역마다 1점을 부여하고 차이가 발생한 영역의 모든 값을 합산하여 산출한다. 산출된 MGI 값이 작을수록 아티스트와 음향기기 Feature 간 차이가 적다는 의미이므로 해당 음향기기의 추천 순위가 높은 것으로 판단한다. Fig. 17은 아티스트 A와 음향기기 a의 격자 그래프를 이용한 매칭 예시이다. 아티스트 A와 음향기기 a의 MGI 값은 6점으로 산출되었다.

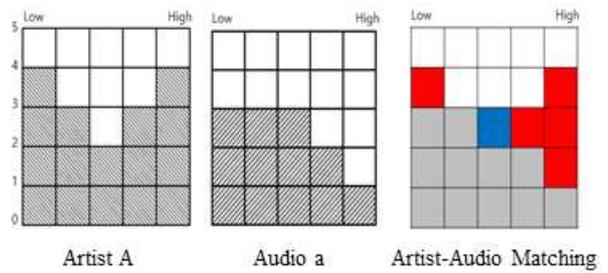


Fig. 17 Artist-Audio Matching Example

Table 2는 아티스트 A, 아티스트 B와 6종의 음향기기들 간 MGI 값을 산출한 결과이다.

Table 2 Artists-Audio Device MGI

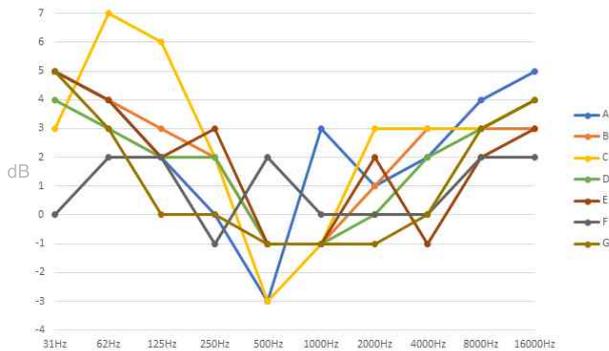
	Dev a	Dev b	Dev c	Dev d	Dev e	Dev f
Artist A	6	8	8	7	4	6
Artist B	7	11	5	4	7	3

아티스트 A와 음향기기의 MGI 산출 결과 음향기기 b와 음향기기 c의 MGI가 8점으로 가장 높아 아티스트 A와의 적합도가 가장 낮은 것으로 나타났고, 음향기기 e의 MGI는 4점으로 아티스트 A에 가장 적합한 음향기기로 선정되었다. 아티스트 B와 음향기기의 MGI 산출 결과 음향기기 b의 MGI가 11점으로 적합도가 가장 낮은 것으로 나타났고, 음향기기 f의 MGI가 3점으로 아티스트 A에 가장 적합한 음향기기로 선정되었다.

4. 추천 결과 검증

음향기기 추천이 적절하게 이루어졌는지 검증하기 위하여 입력으로 사용된 음원의 장르와 추천된 음향기기가 매칭되는지 확인하였다. 음원 장르의 특성을 파악하기 위해 S사에서 제공하는 Equalizer Preset을 사용하였다. S사는 음향기기 제조 및 음원 스트리밍 서비스를 운영하는 글로벌 기업으로 신뢰도 높은 음향 서비스를 제공한다. 이퀄라이저의 범위는 +7dB부터 -4dB까지 설정되었다. 해당 Preset에는 Rock,

R&B, Latin 등 여러 장르의 이퀄라이저가 포함되어 있다. 본 연구에서는 Preset에서 제공하는 정보를 토대로 기준 음역대를 설정하였다. Fig. 18은 장르별 이퀄라이저를 나타낸 그래프이다.



* A: Electronica, B: Rock, C: R&B, D: Jazz, E: Hip-Hop, F: Folk, G: Latin

Fig. 18 Equalizer Graph by Genre

사용자에게 추천된 음향기기의 주파수 대역별 응답치 그래프와 사용자가 선택한 음원 장르의 이퀄라이저의 차이를 비교하여 해당 음향기기가 적절하게 추천되었는지 검증한다. Fig. 19는 아티스트 A를 선호하는 사용자에게 음향기기 e가 매칭되었을 경우, 추천시스템에 입력된 음원의 장르(Rock) 이퀄라이저와 타 장르(Latin) 이퀄라이저, 그리고 본 실험에서 추천된 음향기기 e의 주파수 응답치 그래프를 비교한 예시이다.

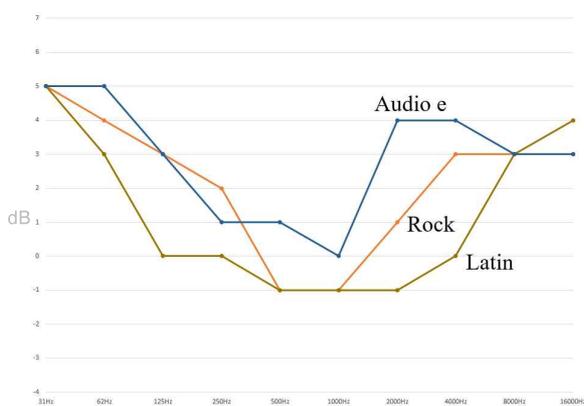


Fig. 19 Recommended Device-Genre Comparison

추천 음향기기의 주파수 응답치 그래프와 사용자가 선호하는 장르의 이퀄라이저, 타 장르 이퀄라이저를 주파수 대역에 따라 비교해 보았을 때 사용자의 선호 장르인 “Rock”과의 차이가 9점, 타 장르인 “Latin”과의 차이는 19점으로 산출되었다. 두 그래프 간 오차가 적을수록 음향기기가 더 적절하게 추천되었다고 볼 수 있으므로 본 연구의 추천 결과가 타당한 것으로 나타났다.

Table 3 Differences between Audio e and Equalizers by Genre

Genre	Difference from Audio e
A: Electronica	19
B: Rock	9
C: R&B	16
D: Jazz	15
E: Hip-Hop	15
F: Folk	22
G: Latin	19

Table 3은 음향기기 e와 각 음악 장르별 이퀄라이저와의 차이를 보여준다. 본 연구에서 제안한 방법을 통해 추천된 음향기기 e에 대하여 장르별 이퀄라이저와의 차이를 비교한 결과 “Rock” 이퀄라이저와의 차이가 가장 작았으며, “Folk” 이퀄라이저와의 차이가 22점으로 가장 큰 것으로 나타났다. 즉 추천 음향기기 e는 비교 대상이 되는 7개 장르 중 사용자가 선호하는 Rock 장르와 가장 유사한 것으로 판별되어 추천 결과의 타당성이 검증되었다.

5. 결론

본 연구에서는 사용자가 선호하는 음원의 보컬 성분을 추출하고 그에 맞는 최적의 음향기기를 추천하는 방법을 제시하였다. 사용자가 선호하는 음원의 보컬을 음원 데이터로부터 분리·추출하고 분석하여 주파수 대역 그래프에 기반한 격자 그래프의 형태로 나타내었다. 분석된 보컬 성분을 각 음향기기의 주파수 대역 응답치와 비교하여 가장 일치율이 높은 음

향기기를 추천하였다. 이 과정은 MGI(matching gap index)를 통해 수치화하여 표현할 수 있으며 추천의 근거를 사용자에게 제공할 수 있다.

본 연구의 결과는 다음과 같이 활용될 수 있다. 추천시스템의 사용자는 적합한 음향기기를 추천받음으로써 전문적인 지식이나 추가적인 조사과정 없이 자신에게 가장 적합한 형태의 청취 환경을 구성할 수 있다. 기존에는 특별한 기준 없이 가격에 기반하거나 타 사용자의 주관적인 리뷰에 근거하여 제품을 구매하였으나 본 연구에서 제안한 MGI 지표를 활용하여 선호에 맞는 제품을 구매함으로써 만족도를 높일 수 있다. 음원 공급업체는 사용자의 음악 취향을 파악하여 높은 품질로 재생할 수 있는 기기를 추천하는 부가 서비스를 제공함으로써 전체 서비스의 가치를 향상시킬 수 있다. 음향기기 업체는 제품의 출시 의도에 맞는 목표 고객들을 적절히 Targeting 함으로써 마케팅 전략에 활용할 수 있다.

향후 연구에서는 음원의 보컬 성분 외에도 음원에 포함되는 다양한 사운드 구성요소를 분석하여 장르별 특성, 악기별 특성 등을 종합적으로 고려한 다차원적인 추천 방법을 제시하여 보컬 성분이 없는 클래식 등의 장르에도 적용될 수 있도록 확장하고자 한다. 또한 다수 인원이나 남녀 혼성 그룹이 참여한 음원의 경우 본 연구에서 제안한 추천의 결과가 유의하게 나타나는지에 대한 검증이 필요하다. 부가적으로 음원의 구성요소들과 음향기기의 Feature 점수를 도출할 때 각 주파수 대역의 특성을 고려하여 구간별로 가중치를 조정하는 방안에 대한 연구가 필요할 것으로 보인다.

References

- Choi, Hong-gu and Hwang, Eunjung (2012). Emotion-based Music Recommendation System based on Twitter Document Analysis. *KIISE Transactions on Computing Practices*, 18(11), 762-76.
- Chung, Ji Yun and Kim, Myoung-Jun (2018). A Study on Personalized Music Recommendation Model through Analysis on Users' Music Preference Factors. *Journal of Digital Contents Society*, 19(11), 2041-2047.
- Hennequin, Romain, Anis Khelif, Felix Voituret and Manue Moussallam (2020) Spleeter: a fast and efficient music source separation tool with pre-trained models. *Journal of Open Source Software*, 5(50), 2154, <https://doi.org/10.21105/joss.02154>
- Hong, Taeho, Hong, Junwoo, Kim, Eunmi and Kim, Minsu (2022). An Intelligent Recommendation System by Integrating the Attributes of Product and Customer in the Movie Reviews. *Journal of Intelligence and Information Systems*, 28(2), 1-18.
- Hwang, Se-Won, Kim, Da-yeong, Kim, Eun-Bae, Lee, Ye-Eun, Eom, Ye-Sol and Lee, Ju-Young (2022). Recommender System for the Hobby-Activity-Item Rental Platform using User-based Collaborative Filtering. *Journal of Digital Contents Society*, 23(6), 985-993.
- Jeon, ByeoungKug and Ahn, Hyunchul (2015). A Collaborative Filtering System Combined with Users' Review Mining: Application to the Recommendation of Smartphone Apps. *Journal of Intelligence and Information Systems*, 21(2), 1-18.
- Kim, Dong Joo and Lim, Kwon Mook (2015). A Music Recommendation Method Using Emotional States by Contextual Information. *Journal of The Korea Society of Computer and Information*, 20(10), 69-76.
- Kim, Hyon Hee, Kim, Dong-Geon and Cho, Jin Nam (2013). A Hybrid Music Recommendation System Combining Listening Habits and Tag Information. *Journal of The Korea Society of Computer and Information*, 18(2), 107-116.
- Kim, Kyung-Rog, Byeon, Jaehee and Moon, Nammee (2011). Collaborative Filtering Design Using Genre Similarity and Preferred Genre. *Journal of The Korea Society of Computer and Information*, 16(4), 159-168.

- Lee, Hyun Sup, Kim, Minyoung, Lee, Ji-Hoon and Kim, Jin Deog (2020). Implementation of User Recommendation System based on Video Contents Story Analysis and Viewing Pattern Analysis. *Journal of the Korea Institute Of Information and Communication Engineering*, 24(12), 1567-1573.
- Lee, Sang Hoon and Kim, Su-Yeon (2013). Design and Implementation of an Intelligent System for Personalized Contents Recommendation on Smart TVs. *Journal of the Korea Society Industrial Information System*, 18(4), 73-79.
- Lee, Soojung (2020). Collaborative Filtering based Recommender System using Restricted Boltzmann Machines. *Journal of The Korea Society of Computer and Information*, 25(9), 101-108.
- McFee, Brian, Alexandros Metsai, Matt McVicar, Stefan Balke, Carl Thomé, Colin Raffel, Frank Zalkow, Ayoub Malek, Dana, Kyungyun Lee, Oriol Nieto, Dan Ellis, Jack Mason, Eric Battenberg, Scott Seyfarth, Ryuichi Yamamoto, viktorandreevichmorozov, Keunwoo Choi, Josh Moore, Rachel Bittner, Shunsuke Hidaka, Ziyao Wei, nullmightybofo, Adam Weiss, Darío Hereñú, Fabian-Robert Stöter, Lorenz Nickel, Pius Friesch, Matt Vollrath and Taewoon Kim (2022). librosa/librosa: 0.9.1 (0.9.1). Zenodo. <https://doi.org/10.5281/zenodo.6097378>
- Oh, Jae-Taek and Lee, Sang Yong (2019). A Music Recommendation System based on Context-awareness using Association Rules. *Journal of Digital Convergence*, 17(9), 375-381.
- Park, Se-Eun, Yoon, Jae-Hee, Choi, Se-Young, Ha, Su-Gyoung and Lee, Kyoung-Mi (2020). User-friendly Korean conversation learning application using face swap and personalized recommendation. *Journal of Digital Contents Society*, 21(12), 2125-2133.
- Park, So Yeon, Kim, Ji Yeon, Kang, Youngok, Cho, Nahye and Yoon, Ji Young (2020). Travel Sites Recommendation Considering User's Preferences Shown in User's SNS Photos. *Journal of Korean Society for Geospatial Information Science*, 28(4), 127-136.
- Sharma, L. and Gera, A. (2013). A Survey of Recommendation System: Research Challenges. *International Journal of Engineering Trends and Technology*, 55(5), 1989-1992.
- Wu, Y. H. and Chen, A. L. (2000). Index structures of user profiles for efficient web page filtering services. In *2012 IEEE 32nd International Conference on Distributed Computing Systems*, 644-644.
- Yoo, SoYeop and Jeong, OkRan (2015). The YouTube Video Recommendation Algorithm using Users' Social Category. *Journal of KIISE*, 42(5), 664-670.
- Yun, So-Young and Youn, Sung Dae (2020). Item-Based Collaborative Filtering Recommendation Technique Using Product Review Sentiment Analysis. *Journal of the Korea Institute Of Information and Communication Engineering*, 24(8), 970-977.



김 정 현 (Jeong-Hyun Kim)

- 대구대학교 컴퓨터정보공학부 공학사
- 대구대학교 경영학과 경영학사
- 대구대학교 IT융합공학과 공학 석사

- 관심분야: 추천시스템, 블록체인, NFT, Web 3.0



석 철 민 (Cheol-Min Seok)

- 대구대학교 컴퓨터정보공학부 공학사
- 대구대학교 경영학과 경영학사
- 대구대학교 IT융합공학과 공학 석사

- 관심분야: 퍼스널모빌리티, 3D프린터, 소방관제



김민주 (Min-Ju Kim)

- 대구대학교 컴퓨터정보공학부
공학사
- 대구대학교 경영학과 경영학사
- 대구대학교 IT융합공학과 석사
과정

• 관심분야: IT기획, 추천시스템, 증강현실



김수연 (Su-Yeon Kim)

- 종신회원
- 포항공과대학교 수학과 이학사
- 숭실대학교 정보산업학과 이학
석사
- 포항공과대학교 산업경영공학과
공학박사

• 대구대학교 정보통신대학 컴퓨터정보공학부
교수

• 관심분야: 지능형 시스템, 기술경영, 지식재산권