

---

# 음악정보와 음악적 성향 분석 및 협업 필터링을 이용한 음악추천시스템

공민서\* · 홍진주\*\* · 최재현\*\*\*

\*숭실대학교 소프트웨어특성화대학원

Music information and musical propensity analysis, and music recommendation  
system using collaborative filtering

Minseo Gong\* · Jinju Hong\*\* · Jaehyun Choi\*\*\*

\*Graduate School of Software, Soongsil University

E-mail : gutssoull@gmail.com\*, hellojinju\_@hanmail.net\*\*

## 요 약

모바일 음악 시장이 점차 커지고 있다. 하지만 현재 적용되는 서비스는 사용자가 선호할 만한 음악을 추천하기에는 정확도가 떨어진다. 본 연구에서는 음악 정보와 사용자의 음악적 성향을 분석해 협업 필터링기법으로 사용자가 보다 선호하는 음악을 자동으로 추천해주는 음악 추천 시스템을 제안한다. 본 시스템은 음원의 메타데이터에서 장르 데이터를 추출해서 장르별로 구분하고, STFT기법의 ZCR, Spectral roll-off, Spectral flux의 요소 벡터값을 추출하여 유사한 음원끼리 군집화를 한 후, TF-IDF기법으로 각 음원 가사의 무드를 분류한 다음, 이 요소들로 협업 필터링기법을 이용해 유사한 취향의 사용자를 발견해 자동 음악 추천을 하는 시스템을 제안한다.

## ABSTRACT

Mobile music market is growing. However, services what are applied recently are inaccurate to recommend music that a user is worth to prefer. So, this paper suggests music recommend system. This system recommend music that users prefer analyzing music information and user's musical propensity and using collaborative filtering. This system classify genre and extract factors what can be get using STFT's ZCR, Spectral roll-off, Spectral flux. So similar musics are clustered by these factors. And then, after divide mood of music's lyric, it finally recommend music automatically using collaborative filtering.

## 키워드

음악 추천, STFT, 클러스터링, 협업 필터링, TF-IDF

## 1. 서 론

많은 사람들이 스트리밍 서비스를 이용하고, 웹사이트를 통해 쉽게 음악을 다운로드 받고 들을수 있게 되면서 음악을 즐기는 장벽이 낮아졌다. 이에 따라 서비스를 제공하는 사이트나 업체는 개인화된 추천 서비스를 제공하긴 하지만 이 추천 서비스의 정확도가 높다고 보긴 어렵다. 사용자 개인의 성향을 반영하는 서비스를 지원하기 위해서는 문제점이 따른다.

첫 번째로, 사용자의 성향을 분석하는 방법에 어려움이 따른다. 두 번째는 분석된 데이터를 토대로 음악 자료들과의 비교를 통해 추천할 수 있어야 한다. 따라서, 두가지 문제점에 대하여 다음과 같은 해법을 제시한다. 사용자의 성향 분석 및 음악의 성향을 유추해내기 위해 음원의 메타데이터에서 장르 태그를 추출한 다음, 음과의 파형을 분석한다. 그리고 TF-IDF를 통해 음원 가사의 무드를 분류한다. 이를 통해 음원의 성질을 수치화하여 나타낼 수 있다. 수치화된 데이터를 토대로

군집화를 적용한다. 군집화를 하기 위해 고려한 방법에는 대표적인 군집화 알고리즘인 K-Means와 EM Clustering을 고려하였다. K-Means와 EM은 반복적인 모델 정제 과정을 통해 최적의 군집을 찾는다라는 면에서 유사하다. 그러나 k-means 알고리즘의 경우, 생성 시켜줄 군집을 정해줘야 하는 단점이 있고 데이터 간의 거리를 계산할 때 유클리디언(Euclidean) 거리 계산 방법을 사용한다. 이와는 달리 EM Clustering은 log-likelihood 함수를 사용하여 모델의 적합성으로 평가하고 통계적인 방법을 사용하므로 더욱 효과적인 기능을 제공한다.[1] 이에 추구하려는 목표에 적합한 EM 군집화(Clustering) 알고리즘을 가지고 음원 데이터를 분석한다.

일반적으로 추천시스템은 크게 협업 필터링(collaborative filtering)과 내용기반(content-based) 추천 방식의 두가지 방식을 사용한다. 내용 기반(content-based) 추천 방식은 사용자가 과거에 선호했던 내용을 분석하여 사용자의 취향을 판단하는 방식이며, 협업(collaborative)추천 방식은 비슷한 취향을 가지는 사용자를 그룹화 하여 추천하는 방식이다. 협업 필터링 기법은 이 중에서도 현업에서의 실제 적용[2] 한 사례에서도 속도와 질적인 측면에서 그 효용성이 입증된 추천 기법 중 하나이다. 그 동안의 음악 추천에 대한 논문들을 살펴보면, 사용자의 시간 상황 정보를 고려하고 협업 필터링을 이용해 음악을 추천[3] 해주는 연구 사례가 있었고, 음원의 음파를 분석하여 벡터 유사도(Vector similarity)를 통해 유사한 음악을 정렬시켜 사용자가 선호하는 음악을 추천[4] 하는 연구가 있다. 본 논문의 구성은 다음과 같다. 제 2장에서는 관련연구를 소개하고, 제 3장에서는 전체 시스템의 구조를 기술하고 제 4장에서는 결론 및 향후 과제에 대해서 기술한다.

## II. 관련 연구

### 2.1 음악 특징 추출

음악이란 파형으로 이루어져 있다. 파형의 특징 종류로는 음들의 높이(pitch), 소리의 세고여림(loudness), 음의 길이(duration), 음질(timbre) 등이 있고 좀더 상세적인 특징으로는 STFT(Short Time Fourier Transform)에 기반을 둔 ZCR(zero crossing-rate), Spectral centroid, Spectral roll-off, Spectral flux와 지각적인 특징을 표현하는 MFCC(Mel-Frequency Cepstral Coefficients), LPC(Linear Prediction reflection Coefficients)등이 있다[5]. 본 논문에서는 ZCR, Spectral roll-off, Spectral flux를 조합하여 사용한다. 이들 조합은 일정 구간 내에서 음악을 찾는 알고리즘에서 높은 정확도를 보여준 바 있으므로 채택하게 되었다[5].

#### 2.1.1 ZCR(Zero Crossing-Rate)

ZCR은 주어진 구간 내에 음성 신호가 기준선인 0을 통과하는 횟수를 측정하는 것이다. 즉, 이산 신호에서 연속 샘플링 값이 서로 다른 부호일 때 발생하는데, 이는 음성의 분할, 분석, 인식에 매우 유용하게 쓰인다. 음성의 신호의 크기를  $s$ 라 하고 샘플  $n$ 에 대한  $i$ 차 ZCR인  $Z_i$ 는 다음과 같이 정의된다[5].

$$Z_i = \sum_{n=1}^N \frac{|sgn|s_i(n)| - sgn|s_i(n-1)|}{2}$$

단,  $sgn|s_i(n)| = \begin{cases} 1 & s_i > 0(n) \\ -1 & s_i < 0 \end{cases}$   
 $s$  = 신호의 크기,  $n$  = 샘플,  $i$  = 차수

그림 1. ZCR의 수식

#### 2.1.2 Spectral roll-off

Spectral roll-off는 음성 구간과 무음성 구간 사이를 구분하는 특징으로써 centroid와 함께 스펙트럼 형태의 또 다른 측정 방법이다. roll-off의 정의는 크기(magnitude) 분포의 85%가 집중해 있는 주파수  $R_t$  이하를 말한다. 즉, 신호의 에너지 중 낮은 신호에 얼마나 많이 집중되어 있는가를 보여준다[5].

$$\sum_{n=1}^{R_t} M_t[n] = 0.85 * \sum_{n=1}^N M_t[n]$$

$M_t[n]$ 은 프레임  $t$ 와 주파수  $n$ 에서 Fourier Transform의 크기

그림 2. Spectral roll-off의 수식

#### 2.1.3 Spectral flux

Spectral flux는 스펙트럼의 변화율을 측정하는 방법으로 연속된 스펙트럼의 분포의 정규화된 크기들의 차를 제공한 것을 의미한다. Flux는 음악과 음성을 구분 짓는데 적합하다[5].

$$F_t = \sum_{n=1}^N (N_t[n] - N_{t-1}[n])^2$$

$N_t[n]$ ,  $N_{t-1}[n]$ 은 각각 현재 프레임  $t$ 와 이전의 프레임  $t-1$ 에서 Fourier transform의 정규화된 크기

그림 3. Spectral flux의 수식

### 2.2 음악 가사 무드 분류

이번에는 감정 범주와 음악 가사 무드 분류를 위한 관련 연구에 대해서 기술한다.

#### 2.2.1 감정 범주

사람의 감정을 수치적으로 표현하는 방법으로 Thayer[6]가 제안한 2차원 공간에 감정 상태를 나타내는 차원 접근법이 있다. 차원 접근법은 긍정성/부정성의 정도(Valence)와 감정 에너지의 정도

(Arousal)를 기준으로 감정을 분류한다. Thayer가 제안한 이차원 감정 모델은 간단하면서 강력한 모델로 이 모델을 이용하면, 그림 4와 같이 11개의 무드로 나타낼 수 있다. 본 논문에서는 이 중 angry, sad, happy, peaceful 4개의 무드를 사용한다.

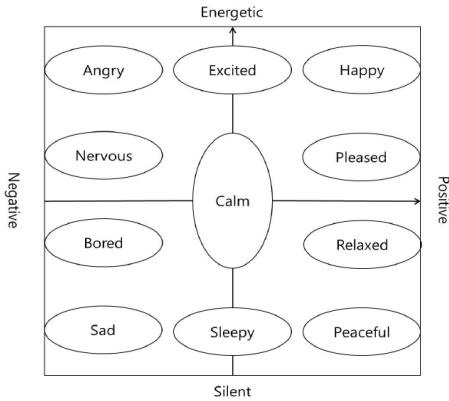


그림 4. Thayer's model

2.2.2 TF-IDF

TF-IDF[7] 가중치 모델은 정보 검색을 위해서 문서 내부의 단어간 상대적 중요도를 평가하기 위해 문서의 표현방식으로 고안된 것이다. TF-IDF 가중치로 표현된 문서는 정보 검색 엔진에서 주어진 질의어와 가장 유사한 문서들의 순위를 결정할 수 있게 할 뿐만 아니라, 유사 문서들의 그룹을 찾는 문서 군집화(document clustering)를 용이하게 한다. TF-IDF 값이 큰 단어는 그것이 속한 문서의 주제 또는 의미를 결정할 가능성이 크며, 따라서 이 측정치를 주요 키워드를 추출할 수 있는 척도로 활용할 수 있다.

2.3 EM(Expectation Maximization) 군집화

EM(Expectation Maximization) 알고리즘은 1958년 Hartley에 의해 처음 제안되었고 1977년에 Demster에 의해서 체계화된 군집(Clustering) 알고리즘이다. EM알고리즘은 K-means 알고리즘과 마찬가지로 초기 모델을 생성한 후 반복 정제 과정을 통하여 모델을 최적화된 모델로 만들어간다.[8] EM 알고리즘은 반복 정제 과정을 통하여 각 객체들이 혼합 모델(Mixture Model)에 속할 가능성(Probability)을 조정하여 최적의 모델을 생성해 간다. K-means 알고리즘이 유클리디언(Euclidean) 거리 함수를 사용해서 모델을 생성해 나가는 것과는 다르게 EM알고리즘은 log-likelihood 함수를 사용하여 모델의 적합성을 평가한다.[1] 즉, EM 알고리즘은 확률기반 군집(Probability-based clustering)이다.

III. 시스템 구조

이 장에서는 음악 추천을 위한 제안 시스템의 전체 구조를 설명한다.

3.1 전체 구조

음악을 추천하기 위한 전체적인 시스템은 그림 5와 같다. 제안 시스템의 입력은 사용자가 좋아하는 음악이고, 출력은 입력물의 정보를 바탕으로 한 유사한 음악이다. 이를 위해서는 충분한 양의 음악 데이터가 구축되어있어야 추천의 정밀도를 높일 수 있다.

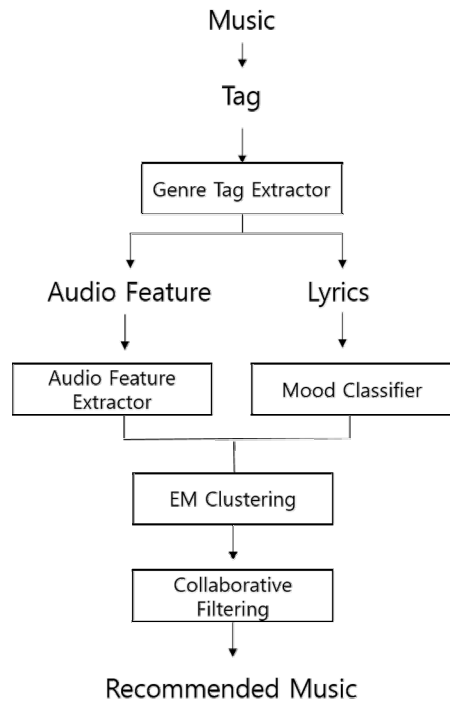


그림 5. 음악 추천 시스템 구조

IV. 결론 및 향후 과제

사용자들의 음악 청취 서비스를 이용하는 것이 점차 늘어나고 있고, 보다 사용자의 취향을 고려하는 추천 시스템이 필요하다. 본 논문에서는 음악을 먼저 장르별로 구분하고 음악을 파형으로 분석하여 데이터화한다. 그리고 음악의 가사를 분석해 음악 무드를 데이터화하고 파형 데이터와 무드 데이터를 군집화한 후, 데이터를 모아놓고 사용자가 선호하는 음악을 선택하면 위의 과정을 거쳐 같은 장르의 유사한 파형을 지닌 음악과 사용자와 비슷한 취향을 지닌다고 판단되는 다른 사용자의 선호 음악을 추천해주는 시스템을 제안하였다.

향후 본 연구에서는 제안된 시스템의 정확성과 유효성을 검증하기 위해 실증적인 연구를 수반할 것이다. 첫째로, 음원 분석을 수행해 방대한 음원 데이터DB를 구축할 것이다. 둘째로, 협업 필터링의 정확도를 높이기 위해서는 수많은 사용자들의 선호 음악 데이터가 필요하므로 사용자들의 데이터를 모을 것이다. 셋째로, 모여진 데이터를 기반으로 새로운 사용자가 제안된 추천 시스템을 사용했을 때의 정확도를 측정해 검증할 예정이다.

### 참고문헌

- [1] Ian Witten and Eibe Frank and Mark Hall, "데이터 마이닝 - 데이터 속 숨은 의미를 찾는 기계학습의 이론과 응용", 에이콘, 396-397, 2011
- [2] Greg Linden, Brent Smith and Jeremy York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering.", IEEE Internet Computing, Jan-Feb, 2003
- [3] 이동주, 이성근, 이상구, "시간 상황 정보를 고려한 협업 필터링을 이용한 음악 추천", 한국컴퓨터종합학술대회 논문집, Vol.36, 2009
- [4] 김동문, 이지형, "음파 분석을 이용한 사용자 적응형 음악 추천 시스템", 한국정보과학회 가을 학술발표논문집, Vol.33, 2006
- [5] G. Tzanetakis, "Manipulation, Analysis and Retrieval systems for audio signals", Ph.D dissertation, 32-33, 42-43, June 2002
- [6] R. E. Thayer, "The Biopsychology of Mood and Arousal", Oxford University Press, 1989
- [7] <http://en.wikipedia.org/wiki/Tf%E2%80%93idf>
- [8] 김현욱, 송하윤, "EM(Expectation Maximization) 군집화(Clustering)을 통한 인간의 이동 패턴 연구", 한국컴퓨터종합학술대회 논문집, Vol.38, 2011