

음악의 패턴 분류 및 정보 검색을 위한 클러스터링 기반의 음악적 단어 설계 연구

*김현우, 김병희, 장병탁
서울대학교 전기컴퓨터공학부
e-mail : hwkim, bhkim, btzhang @ bi.snu.ac.kr

Design of Musical Words Based on Clustering for Pattern Classification and Information Retrieval of Music

*Hyun-Woo Kim, Byoung-Hee Kim, Byoung-Tak Zhang
School of Computer Science and Engineering
Seoul National University

Abstract

Typical content-based methods for music classification and retrieval mainly deal with global statistics or features of pre-divided songs. However, focusing on local, heterogeneous fragments and features is mandatory for flexible analysis. We propose a musical word scheme based on clustering as the base for local pattern extraction, robust classification and flexible music retrieval.

I. 서론

근래의 인터넷 기술의 발전과 멀티미디어 데이터의 압축 기술의 발달, 그리고 다양한 모바일 장치의 개발로 인해 MP3를 포함한 디지털 음원은 현대인들이 흔히 접할 수 있는 데이터의 형태가 되었다. 이에 따라 자동화된 음원의 분류 및 검색기술의 필요성이 커지고 있다. 상용화된 MP3 및 음원 검색 기술은 노래 가사, 노래 제목, 가수, 장르 등의 메타정보 이용에 국한되어 있으며, 음악 자체의 다양한 특성을 반영하지 못하여, 사용자들의 잠재적인 필요를 충족시키지 못하고 있다.

이를 해결하기 위해 근래에는 음원 자체의 내용을 기반으로 검색 및 추천, 분류하는 연구가 활발히 이루어지고 있다. 대표적인 예로 음원을 작은 조각으로 나누어서 여러 가지 특성값을 추출하고 이동평균과 분산 등을 이용해 분류한 연구결과가 있다 [1]. 그러나 현재의 기법들은 다양한 길이의 음원을 다루고, 음원의 부분적 특성을 분석하고 처리하기에는 제한적이다. 실제로 사람들은 음원의 부분만을 기억하거나 듣게 되며, 광고에서나 영화에 삽입 되는 형태 또한 음원의 일부분이다.

다양한 길이의 음원을 자유롭게 일부분만을 검색하고, 여러 음원의 부분들을 특징들을 조합한 질의를 다

루는 기술개발을 위한 기반 연구로서 본 논문에서는 음악적 단어를 정의하고, 이 기법의 유용성을 확인하기 위해 음악적 단어를 기반으로 음원을 가수별로 분류해보았다.

II. 관련 연구 및 음악적 단어 정의

2.1 관련 연구

음원 내용 기반 검색에 관한 연구는 근래에 활발히 이루어지고 있다. 널리 사용되는 음원은 MIDI 파일과 같은 symbolic 형태의 데이터와 Wave, MP3와 같은 signal 형태의 데이터가 있다. MP3 등의 signal 형태의 음원이 보편화됨에 따라, DSP(digital signal processing)을 통한 음원 분석과 내용 기반(content-based)의 음원 분류 및 검색기법 연구가 활발하게 이루어지고 있다. 대표적인 예로는, 내용 기반의 유사 음원 검색과 분위기 추출 연구[1] 및 음원 내용 기반 장르 분류 연구가 있다[2].

2.2 음악적 단어 정의

음악적 단어는 표 1에 설명된 과정을 통해 생성한다. 기본 아이디어는 음악을 작은 조각으로 분할하여 유사한 특성을 가지는 조각들의 집합을 단어로 정의하는 것이다. 1개의 조각을 1개의 단어에 매칭시킨 후 각 곡은 단어의 조합으로 다시 표현된다.

MP3음원을 분할하고 특성값(feature)을 추출하는 도구로는 jAudio를 사용하였다(<http://jaudio.sourceforge.net>). 특성값은 음성인식과 음악 정보 처리에 널리 쓰이는 MFCC(Mel-frequency cepstral coefficients)를 사용한다[3]. 클러스터링 과정에서 특성값 벡터로 표현되는 각 조각의 비교시 척도의 영향을 줄이기 위해 척도를 통일하는 표준화 과정을 추가한다.

표 1. 음악적 단어 정의 절차

단계	과정 명칭	세부 설정
1	음원 분할	1 구간 = 16kHz * 4092 samples
2	특성값 추출	각 구간에 MFCC를 13등급으로 구분하여 추출
3	표준화	각 MFCC 등급의 분산을 1로 표준화
4	군집화	표준화된 조각 데이터 집합에 k-means를 적용하여 200개의 클러스터로 군집화. 유사도 측정 기준은 Euclidean distance 적용
5	음악적 단어 정의	각 클러스터를 독립된 단어로 정의하며, 클러스터의 평균점을 단어의 특성값으로 지정.
6	음악적 단어로 음원표현	음원 조각에 최근접 이웃 기법을 적용하여 단어를 지정한다. 한 음원은 연속된 음악적 단어의 집합으로 표현된다

2.3 음악적 단어 기반의 응용

음악적 단어의 응용 가능성은 다양하다. 음원의 분류 또는 검색시 특정 장르 또는 가수에게 나타나는 단어들의 조합(패턴)을 기반으로 보다 간결하고 빠르게 작업을 수행할 수 있으며, 음악을 부분적으로 처리하기에 유용하다. 그리고, 단어의 집합으로 각 음원을 표현함으로써, 정보검색(information retrieval on text) 분야에서 축적된 다양한 패턴 분석 및 검색 기법을 적용할 수 있다.

III. 실험 및 분석

설계한 음악적 단어의 유용성을 확인하고자 음악적 단어의 분포 패턴을 기반으로 음악을 분류하는 실험을 수행하였다. 음악의 분류 문제로는 장르 구분 및 무드 구분 문제가 많이 다루어져왔다. 음악 장르 구분 문제에서는 장르가 음악 내용과 무관하게 지역, 또는 시대적으로 구분되는 경우가 있다[4]. 이러한 모호성을 피하기 위해, 장르 대신 가수를 분류하는 실험을 설계하였다. 본 논문에서는 각 곡을 bag of words 방식으로 표현하여 음악적 단어의 빈도 정보만으로 곡을 분류한다.

데이터는 특성이 잘 구분되는 유명가수 5명 (Megadeth, Maroon5, Boyz2men, Britney Spears, Eminem)의 MP3 곡들을 음악적 단어의 집합으로 변환한 자료를 사용하였다. 각 가수별로 20곡씩 총 100곡을 기반으로 표 1의 절차에 따라 음악적 단어 200개를 정의하고 해당 곡에 대한 분류 성능을 cross-validation 기법으로 평가한 결과는 표 2와 같다. 100곡을 학습 데이터로 사용하고 별도의 15개 곡을 테스트 데이터로 구성하여 분류 성능을 측정한 결과는 표 3과 같다.

분류 실험 결과 음악적 단어를 기반으로 한 가수 분류의 정확도가 높음을 확인할 수 있었다. 한 가지 특기할 만한 사항은, 음원의 특성값으로 MFCC만을 적용하였음에도 불구하고 성능이 높았다는 점이다.

표 2. 음악 단어로 표현한 5명의 가수 100곡을 다양한 분류 알고리즘으로 분류한 결과. 10-fold cross validation을 5회 반복 수행하여 측정한 분류 정확도의 평균 및 표준편차를 표기하였다.

(SVM: support vector machine, ANN: multi-layer perceptron, NB: naive Bayes, NB multi: NB multinomial, BNC: Bayesian network classifier 중 TAN을 적용, kNN: k-nearest-neighbor)

	SVM	ANN	NB	NB multi	BNC (TAN)	kNN
평균(%)	89.60	92.40	85.80	85.80	89.60	87.20
표준편차	9.68	8.22	11.65	11.26	10.68	9.04

표 3. 테스트 데이터에 대한 분류성능 평가 결과

	SVM	ANN	NB	NB multi	BNC (TAN)	kNN
정답비율	12/15	12/15	11/15	12/15	13/15	13/15

이는 기존의 내용기반의 장르 분류 연구에서 여러 가지 특성값을 이용했고, MFCC만을 사용했을 경우 매우 낮은 성능(5-class 문제에서 60% 전후의 정확도)을 보였던 결과와[5] 비교된다.

IV. 결론 및 향후 연구 방향

본 논문에서는 클러스터링 기반의 음악적 단어를 정의하고 음악적 단어가 음원 분류 성능 향상에 기여할 수 있음을 확인하였으며, 음악 정보 검색에의 적용 가능성을 점검하였다.

향후 연구과제로 단어의 순서를 고려한 패턴 분류 및 정보 검색 기법을 개발하고자 한다. 그리고 정보 검색 성능의 향상을 위한 방법으로서, 분류에 기여하는 단어의 선별, 지정된 클래스를 대표하는 단어의 연속 패턴(모티프) 탐색 기법을 연구하고자 한다.

감사의 글

본 연구는 교육인적자원부 BK21-IT 과제에 의하여 일부 지원되었다.

참고문헌

- [1] T. Li, M. Ogihara, Content-based music similarity search and emotion detection, *Proc. ICASSP*, Vol. 5, pp. 705-708, 2004.
- [2] G. Tzanetakis, P. Cook, Musical genre classification of audio signals, *IEEE Transactions on Speech and Audio Processing*, Vol. 10, No. 5, pp. 293-302, 2002.
- [3] M.J. Carey, E.S. Parris, H. Lloyd-Thomas, A comparison of features for speech, music discrimination, *Proc. ICASSP*, Vol. 1, pp. 149-152, 1999.
- [4] N. Scaringella, G. Zoia, D. Mlynek, Automatic genre classification of music content: a survey, *IEEE Signal Processing Magazine*, Vol. 23, No. 2, pp. 133-141, 2006.
- [5] T. Li, M. Ogihara, Q. Li, A comparative study on content-based music genre classification, *Proc. ACM SIGIR*, pp. 282-289, 2003.