

화성정보 추출을 이용한 음악 장르분류

손우람^{1)O}, 정민석²⁾, 안주영³⁾, 윤경로⁴⁾
건국대학교 컴퓨터공학부¹⁾⁴⁾, 송실대학교 정보통신전자공학부²⁾, 경희대학교 전자정보학부³⁾
{sonwr^{1)O}, yoonk⁴⁾}@konkuk.ac.kr, leader@ssm.samsung.co.kr²⁾, juneday@myjung.com³⁾

Automatic Genre Classification using Music Harmonic Detection

Wooram Son^{1)O}, Minseok Jung²⁾, JooYoung An³⁾, Kyoungro Yoon⁴⁾
Div. of CSE, Konkuk University¹⁾⁴⁾, Div. of EE, Soongsil University²⁾, Div. of EE, Kyunghee University³⁾

요 약

저장매체의 대용량화와 인터넷을 이용한 디지털 음원의 활성화로 개인이 소유하는 음원이 급속도로 증가하고 있다. 많은 양의 음원을 보유하고 있는 상황에서 사용자의 편의를 증가시키기 위하여 다양한 검색/분류 방법들이 개발되고 사용되고 있다. 본 논문에서는 음원에 사용된 표현방식이나 디렉토리 구조, 파일이름, 텍스트 태그 등에 독립적으로 적용될 수 있도록 디지털 신호처리 이론에 기반하여 파형데이터를 분석하고, 화성학 이론에 기반한 패턴매칭 기술을 응용하여 음악의 장르와 나아가 분위기를 기반으로 분류하는 방법을 제시한다.

1. 서 론

음악은 언어의 장벽을 뛰어넘는 대표적인 의사소통 도구의 하나이다. 이러한 특징에 의해 음악은 인터넷을 통해 디지털 음원의 형태로 배포되면서 결국 개인이 소유하는 음원의 양적 증가를 가져왔다. 또한 최근 휴대용 미디어 재생장치들도 대용량 저장장치를 제공하면서 사용자의 음악에 대한 선택권을 더욱 넓어지고 있는 상황이 되었다.

현재 대부분의 음악 재생 시스템에서의 음원 분류는 디렉터리, 파일의 이름을 기반으로 하거나 음원 파일 포맷의 헤더에 담긴 정보를 사용하고 있다. 후자의 방식을 취한 대표적 방법인 ID3는 MP3를 위한 태그 포맷으로 제목, 가수(음악가), 음반 이름, 곡 번호 등의 메타 데이터를 MP3 파일에 포함시킬 수 있도록 하여준다. 하지만 이런 방법은 같은 내용의 음원이라도 태그에 삽입된 내용에 따라 다르게 분류될 가능성이 있다.

오디오 신호 처리기술을 사용하여 음원의 특성들을 추출해내고 신경망 알고리즘을 사용하여 음원을 분류하는 방법이 많이 연구되고 있는 상황에서 [1][2][3] 본 논문은 화성학적인 특성을 고려한 특성추출 방법을 제시한다. 오디오 신호 처리기술을 사용하여 시간도메인에서의 파형신호를 주파수 도메인으로 변환하고 시간변화에 따른 주파

수 스펙트럼의 특징들을 추출하여 음악의 장르를 분류하고 음악이론에 의한 표현인 화음, 음계(스케일), 선법(모드) 등을 추출해내어 분류의 정확성을 높이고 추상적인 영역인 음악의 분위기를 분류해내도록 한다.

2. 본 론

본 시스템은 다양한 형태로 표현되고 있는 음원 파일을 복호화하여 PCM 형태의 데이터로 입력 받는 부분과 웨이브 파형을 분석하는 모듈 (Wave Analyzer), 신호처리 기법을 사용하여 음악적 피치를 추출해 내는 모듈 (Musical Pitch Estimator), 추출된 음악적 피치를 사용하여 화성학적인 정보를 추출하는 모듈 (Harmonic detection), 주파수 도메인으로 변환된 신호를 스펙트럼 형태로 보여주는 모듈(Spectrogram Analyzer) 그리고 스펙트럼형태의 특징들을 추출하는 모듈(Wave Feature Detector)로 구성되어 있다.

각 모듈들은 그림(1) 과 같이 오디오 시그널을 입력으로 받고 최종적으로 음악의 장르와 추출된 화성적 정보에 근거한 분류를 결과로 출력한다.

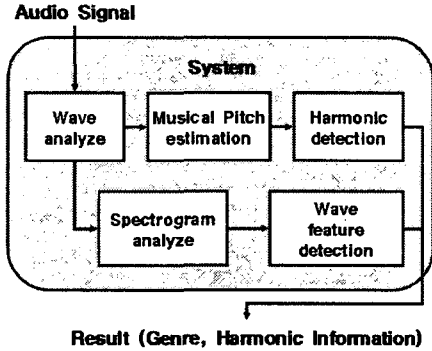


그림 (1). 시스템 다이어그램

2.1. 오디오 신호처리를 통한 특성 추출

2.1.1. 비트 추출 (Beat Detection)

비트를 추출하는 것은 음악의 특징인 빠르기를 알 수 있고 마디를 나눌 수 있는 요소로 작용을 하여 다른 화성학적인 특징을 추출할 수가 있는 중요한 작업이다. 대중적인 음악의 비트는 평균적으로 120Hz 부근에서 발생하므로 FIR필터의 한 종류인 Low-pass Filter를 사용하여 신호를 걸러낸다.[4] 필터를 통과한 샘플에서 순간에너지 (E_i) 와 지역평균에너지 (E_{LA}) 를 구하여 비트영역을 찾아낸다.

$$Beat := Instant E_i > Local Average E_{LA}$$

순간 에너지는 Instant E_i 로 아래 식(1)과 같이 구한다.

$$E_i = \sum_{k=i_0}^{i_0+n} a[k]^2 + b[k]^2 \quad \text{식 (1)}$$

샘플의 인덱스는 i , 윈도우의 샘플개수는 n , $a[n]$ 은 Left 채널의 샘플값, $b[n]$ 은 Right 채널의 샘플값을 의미한다.

$$E_{LA} = \frac{n}{N} \sum_{i=0}^N B_a[i]^2 + B_b[i]^2 \quad \text{식 (2)}$$

사람은 약 1초의 시간 동안의 에너지와 Instant E_i 와 상대적인 차이로 비트를 지각하므로 평균에너지는 위 식 (2)과 같이 구한다.

1초 동안의 샘플 개수는 n , B_a 는 Left 채널 버퍼, B_b 는 Right 채널 버퍼를 의미한다.

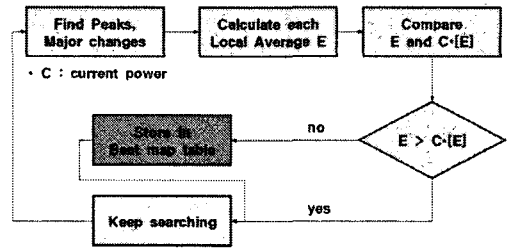


그림 (2). Beat Detection 모듈 다이어그램

2.1.2. 스펙트럼 분석 (Spectrogram analyze)

시간도메인에 따른 주파수 성분의 분석을 위해 STFT 를 사용하였다.

스펙트럼을 이용하여 추출할 수 있는 정보는 주파수의 분포를 나타내는 Bandwidth, 음악의 에너지의 총 크기, 음악의 피치가 있다.

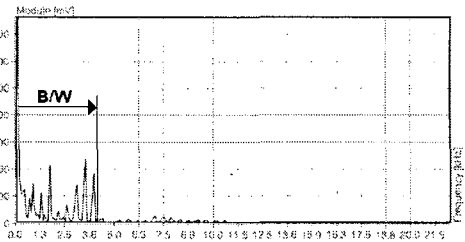


그림 (3). 스펙트럼으로 추출한 B/W - 펑크(Funk)

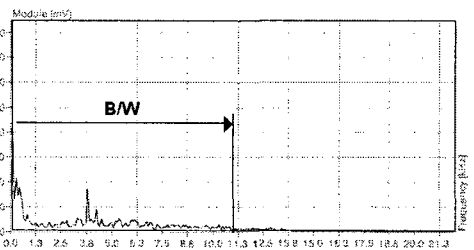


그림 (4). 스펙트럼으로 추출한 B/W - 재즈(Jazz)

위의 그림 (3)과 (4)는 장르별 스펙트럼의 Bandwidth 를 추출한 데이터이다. 장르별 Bandwidth와 에너지의 총 크기가 다르게 나타나고 있으므로 분류를 위한 하나의 특징벡터로 사용할 수가 있다.

$$E = \sum_{k=0}^n a[k]^2 + b[k]^2 \quad \text{식 (3)}$$

음악의 에너지의 총 크기(E)는 위 식(3)과 같이 구할 수 있다. n 은 음악의 마지막 샘플이다.

2.2. 음악 이론을 기반으로 한 통계적 정보 추출

2.2.1. 박자표(Time Signature) 추출

박자를 추출하기 위해선 Beat Detection으로 구한 Beat map table에서 박자표의 규칙성을 찾아낸다.

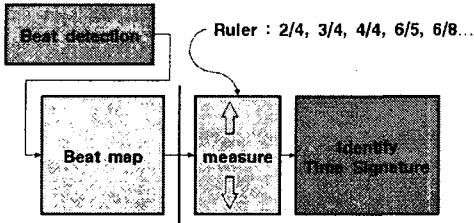


그림 (5). 박자표 추출 모듈

2.2.2. 음계(스케일), 선법(모드) 추출

음악의 스케일은 반음의 개수, 3음의 위치, 5음의 위치로 결정되게 된다. 모드의 결정요소는 반음의 개수, 2, 3, 6, 7음의 위치이다. 스케일과 모드를 추출한다면 음악의 분위기를 파악 할 수 있다.

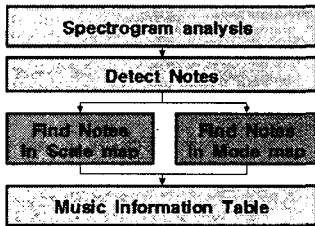


그림 (6). 스케일, 모드 추출 모듈

스펙트럼 분석기의 시간의 변화에 따른 주파수도메인 성분을 통해 음악의 피치[5], 화음을 추출하여 스케일을 추정한다. [6]

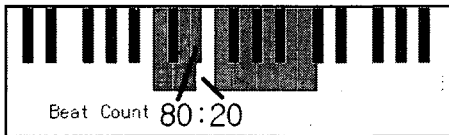


그림 (7). 스케일 추정 모듈 -.minor (Over Pentatonic)

음의 간격과 그 사이에 들어가는 반음과 온음의 개수, 빈도, 비율을 이용하여 스케일과 코드를 추정하였다. 현대 음악에서는 Passing Note라고 하여 Avoid Note도 연주하게 된다. 그러므로 위 그림 (7)과 같이 Avoid Note의 온음 혹은 반음과의 비율로 확률적으로 찾아내게 된다.

2.3. 음악 분류

장르별로 선정된 샘플에서 성분들을 추출하고 평균 값과 Offset 값을 구하여 필터를 만든다. 그리고 음원을 입력

받았을 때 성분분석 후 음악분류 필터를 통해 분류한다.

2.4. 실험결과

음악의 장조, 단조 스케일을 분류한 결과이다.

	Major	Minor
Result	51.7 %	78.5 %
False Positive	0.0 %	19.6%
False Negative	19.6%	14.2 %

표 (1). 장조, 단조 스케일 분류 결과

총 112개의 샘플을 테스트하였으며 56개의 장조 (Major)음악과 56개의 단조(minor)음악으로 실험하였다.

3. 결론

음원에서 정보를 추출하는 경우 개인의 취향을 고려한 음악 선곡시스템, 현재 듣고 있는 음악의 분위기와 유사한 음원을 검색해주는 시스템 등 다양한 응용이 있을 수 있다. 본 시스템의 개선해야 할 방안은 스케일 장조(Major)의 분류 정확도가 떨어지는 것으로 Passing Note의 영향으로 생각되며 관련 알고리즘의 성능향상이 필요하다. 음악분류 모듈 또한 신경망 알고리즘을 적용하여 성능을 향상시킬 수 있는 가능성이 존재하고 있다.

4. 참고문헌

[1] Tzanetakis G., and etc. " Automatic Musical Genre Classification of Audio Signals, Proc" . *ISMIR*, 2001.

[2] Shenoy A., Mohapatra and Wang Y. " Key Determination of Acoustic Musical Signals" , *International Conference on Multimedia and Expo*, 2004

[3] Klapuri A. P., " Musical meter estimation and music transcription" , *Cambridge Music Processing Colloquium*, Cambridge University UK, 2003

[4] E.D. Scheirer, " Tempo and Beat Analysis of Acoustic Musical Signals" , *Journal of the Acoustic Society of America*, 103 (1), pp. 588-601, 1998

[5] Yongwei Zhu., and etc. " Music Key Detection for Musical Audio" , *11th International Multimedia Modelling Conference(MMM' 05'*), 2005

[6] Anssi Klapuri, " Pitch Estimation using Multiple Independent Time-Frequency Windows " , *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New York, 1999.