

음악 무드의 변화 기반 유사 음악 검색 기법

전상훈, 한병준, 황인준
고려대학교 전기전자공학과
e-mail : {ysbhjun, hbj1147, ehwang04}@korea.ac.kr

A Music Retrieval Scheme based on Variation of Musical Mood

Sanghoon Jun, Byeong-jun Han, Eenjun Hwang
School of Electrical Engineering, Korea University

요 약

음악에서는 다양한 감정의 표현을 시간에 따른 음악 무드의 전이로 표현한다. 본 연구에서는 Longest Common Subsequence (LCS) 알고리즘 및 k-Means 알고리즘에 기반한 유사 음악 검색 기법을 제안한다. 우선, 음악 무드의 흐름을 무드 세그먼트 단위로 나누고, 이를 추출된 다양한 음악 특성을 k-Means 알고리즘으로 분류하여 무드 시퀀스로 변환한다. 또한, 유사한 무드의 흐름을 가지는 음악을 검색하기 위해 LCS 알고리즘에 기반한 무드 시퀀스의 유사도를 정의한다. 본 논문은 제안된 내용을 바탕으로 실험과 설문 조사를 통해, 기존의 전역적 특성 검색 방식보다 시퀀스를 이용한 검색 방식이 좀 더 효율적임을 증명하였다.

1. 서론

최근 멀티미디어 정보 검색을 위한 다양한 기술이 제안 및 응용되고 있다. 특히 근 20 여년 간 주석 기반(annotation-based) 정보 검색 연구로부터 내용 기반(content-based) 정보 검색 연구의 필요성이 증대되어 왔다. 멀티미디어 정보 검색의 다양한 세부 분야 중 하나인 음악 정보 검색(MIR; Music Information Retrieval)의 경우, 음악 콘텐츠 자체로부터 추출 가능한 특성 정보에 기반한 검색 및 제반 시스템에 대한 다양한 연구의 필요가 증가하고 있다.

본 논문에서는 음악 정보 검색의 세부 분야 중 유사 음악 검색을 위한 새로운 기법을 제안한다. 우선, 음악 검색을 위해 하나의 음악이 하나 이상의 무드를 가짐에 착안하였다. 이를 시간 흐름에 따른 음악 무드의 변화를 무드 세그먼트로 분할하고, 각각의 세그먼트에 k-Means 알고리즘을 적용하여 무드를 분류한 뒤, 이를 무드 시퀀스로 표현한다. 또한, Longest Common Subsequence (LCS) 알고리즘을 사용하여 다양한 음악으로부터 추출된 무드 시퀀스 간 유사도를 계산한다. 마지막으로, 실험을 통해 제안된 검색 기법이 기존의 기법보다 우수함을 보인다.

2. 음악 무드의 표현과 비교

2.1. 음악 무드의 특성과 무드 시퀀스의 표현

기존 연구[1,2]에서 소리의 세기(loudness), 하모니(harmony), 템포(tempo), 조성(key) 등의 다양한 음악 특성을 통하여 음악 특성과 음악 무드 사이에 밀접한 연관관계가 있음을 밝혔다. 이는 어떤 음악 특성으로부터 음악 무드를 유추하는 것이 가능함을 의미한다.

따라서, 음악 검색 및 추천 시스템에서 이러한 특성을 한 음악의 무드가 검색을 위한 중요한 단서로 활용할 수 있다.

한편, 기존 연구는 한 음악이 단일의 무드를 가짐을 전제로 수행되었다는 점에서 한계가 있다. 왜냐하면 한 음악은 하나 이상의 무드를 가질 수 있고, 이들은 시간 흐름에 따라 변화할 수 있기 때문이다. 따라서 보다 정확한 검색을 위해서 음악 무드 변화에 따라 나뉜 무드 세그먼트 별 음악 특성이 추출되어 활용되어야 한다.

이 때, 한 음악의 무드 변화를 표현하는 무드 시퀀스 S 를 정의하면 수식(1)과 같다.

$$S = \{s[1], s[2], \dots, s[n]\} \quad (1)$$

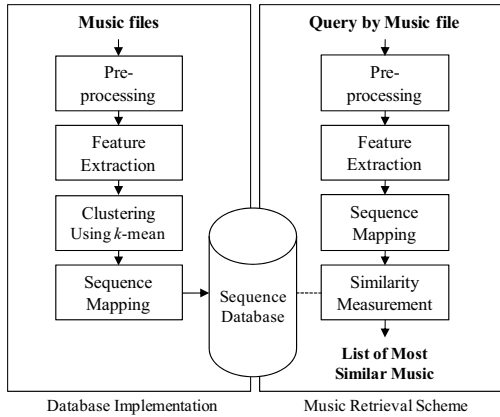
여기서 S 의 각 원소는 시간의 흐름에 따른 각 세그먼트의 무드를 순차적으로 표현한다. 또한, i 번째 세그먼트가 가지는 무드를 $s[i]$ 라 하고, $s[i+1]$ 는 $s[i]$ 보다 나중의 시간 순서를 가지는 무드이다.

2.2. Longest Common Subsequence (LCS) 알고리즘

Longest Common Subsequence (LCS)는 두 시퀀스의 공통되는 가장 긴 서브 시퀀스로 정의된다. 예를 들어, 두 시퀀스 $S_1 = \{T, U, e, S, D, A, Y\}$, $S_2 = \{T, h, U, r, S, D, A, Y\}$ 에 대하여 LCS 알고리즘은 S_1, S_2 에 포함되는 수많은 서브 시퀀스 중 공통이 되는 가장 긴 길이를 갖는 서브 시퀀스, 즉 $\{T, U, S, D, A, Y\}$ 라는 시퀀스를 계산하는 알고리즘이다[3,5].

2.3. 시퀀스 간 유사도 계산

임의의 두 시퀀스 간 유사도 계산을 위해 질의 시



(그림 1) 시스템의 다이어그램

퀀스와 목표 시퀀스 간 LCS 를 구하고, 다음 수식(2)으로 유사도를 계산한다.

$$Sim(S_q, S_d) = \frac{LCS(S_q, S_d)}{|S_q|} \quad (2)$$

이 때, S_q 는 질의 시퀀스, S_d 는 목표 시퀀스이다. $LCS(S_q, S_d)$ 는 S_q 와 S_d 의 LCS 의 길이를 반환한다. 마지막으로, 유사도 Sim 은 질의 시퀀스의 길이로 나누는 표준화 과정을 거쳐 구한다.

3. 유사 음악 검색 시스템 구조

효과적인 유사 음악 검색 기법으로써 무드 시퀀스 사상 기법에 기반한 검색 방식을 제안한다. 제안하는 시스템은 무드 순차 데이터베이스와 유사 음악 검색기로 구성된다. 검색 시스템 구조는 그림 1 과 같다.

3.1. 무드 순차 데이터베이스

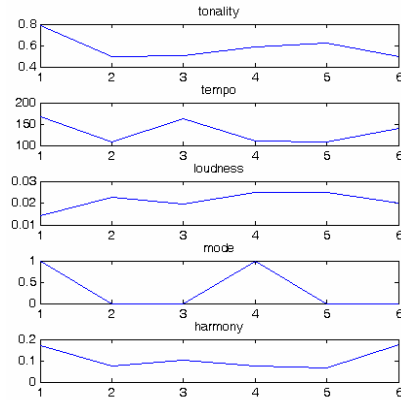
먼저 데이터베이스 설계를 위해 하나 이상의 무드를 갖는 다수의 음악에 대해 전처리 과정을 수행한다. 전처리 과정에서는 무드 세그먼트 분할 및 각 세그먼트로부터의 음악 특성 추출이 포함된다.

이후 추출된 음악 특성을 사용해 k-Means 알고리즘 기반으로 각 무드 세그먼트를 분류하여 무드를 분류한다. 분류된 무드는 비슷한 무드끼리 군집된 대표 클래스의 인덱스로 사상(mapping)하여 표현한다. 최종적으로, 분류된 각 무드 세그먼트들을 연결하여 하나의 무드 시퀀스로 표현된다.

예를 들어, 현재 무드 클래스 $C_1 \sim C_N$ 이 존재하며, 어떤 음악이 $\{m_1, m_2, m_3, m_4\}$ 의 무드 시퀀스를 가진다고 가정한다. 이후, k-Means 알고리즘에 의해 $m_1 \in C_1, m_2 \in C_5, m_3 \in C_3, m_4 \in C_6$ 의 사상 관계가 형성될 경우, 무드 시퀀스는 $\{C_1C_5C_3C_6\}$ 로 표현된다.

3.2. 무드 유사 음악 검색기

질의 시퀀스를 생성하기 위해 데이터베이스 설계과



(그림 2) 5 가지 음악 특성의 무드기반 변화

정과 동일한 전처리 및 음악 특성 추출 과정을 수행한다. 이후, 무드 순차 데이터베이스 설계 과정에 의해 생성된 클래스 중 각 무드 세그먼트가 포함되는 클래스를 검색하여 질의 시퀀스를 생성한다.

생성된 질의 시퀀스에 대하여 우선 시퀀스 데이터베이스에서 정확히 일치하는 시퀀스를 검색한다. 여기서 정확히 일치하는 시퀀스는 검색 결과의 최상위에 표시된다. 이후, 일치하지 않는 데이터베이스 내 시퀀스 각각에 대해 수식 (2)의 유사도를 계산하고, 유사도 순으로 검색 결과를 출력한다.

4. 실험 및 결과

본 장에서는 2.1 절에서 가정한 무드에 따른 음악 특성의 변화를 음악 특성 평면에 보인다. 또한 제안한 시스템이 기존 방식[1]에 비해 개선되었음을 보이기 위해 사용자 성능평가를 수행하였다.

4.1. 무드 변화에 따른 음악 특성 비교

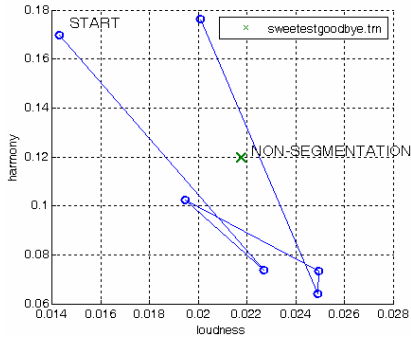
5 가지 음악 특성(템포, 조성, 음계, 소리의 세기, 하모니)을 각 무드 세그먼트에서 추출하였다. 그림 2 는 음악 특성이 무드 전환에 따라 달라짐을 보인다.

하모니-소리 세기 평면 상에서 각 무드 세그먼트 별 추출된 음악 특성(O)과 전역 음악 특성(X)의 분포 차이는 그림 3 과 같다. 이를 통해, 특성 변화가 무드 세그먼트 단위 음악 특성에서는 발견되는 반면, 전역 음악 특성을 통해서는 관찰할 수 없음을 알 수 있다.

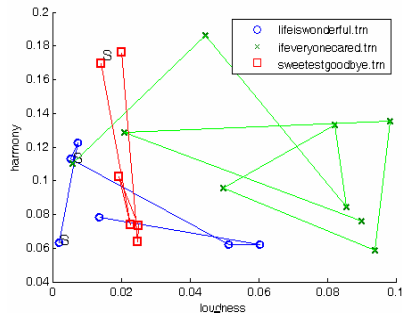
또한 그림 4 는 3 곡의 음악으로부터 추출된 무드 세그먼트 별 음악 특성들을 하모니-소리 세기 평면에 표현하고 있다. 이 때 몇몇 특성들은 서로 인접함을 관찰할 수 있다. 이는, 무드 세그먼트 분할 및 무드 세그먼트 별 특성 추출을 통해, 무드의 변화와 음악적 구조가 서로 다른 음악일지라도 부분적으로는 공통점을 찾을 수 있음을 의미한다.

4.2. 전역 특성 검색과 시퀀스 특성 검색의 비교

무드 시퀀스를 사용한 유사 음악 검색과 기존 단일



(그림 3) 하모니와 소리세기의 무드기반 변화



(그림 4) 3 가지 음악의 음악특성 변화

무드에 의한 검색 성능 비교를 위해 30 곡의 팝 음악으로 실험 데이터베이스를 구성하였다.

그림 5 는 기존의 전역 특성 추출 방식으로 검출된 무드로 데이터를 군집화하고, 군집화된 데이터와 각 음악의 무드 시퀀스 유사도를 계산한 후 평균한 결과이다. 이 결과는 기존 방식과 제안된 방식 간의 결과 차이가 큰 경우 0 에 가까운 결과를 보인다. 전체 30 개의 노래 중 50% 이하의 결과가 14 개(47%)로, 제안된 방식이 기존 방식보다 정확도가 높음을 알 수 있다.

4.3. 사용자 평가

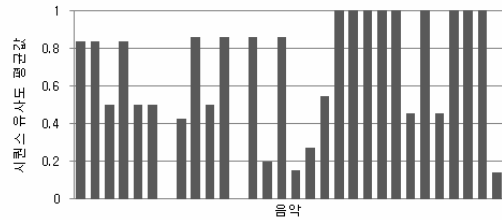
제안된 방식과 기존 방식 간 만족도 비교를 위해 대학교 내에서 모집된 다양한 연령대의 5 명의 피험자를 대상으로 다음 절차에 따라 평가를 실시하였다.

1) 피험자가 선택한 음악 하나에 대해 전처리 및 특성 추출을 한 뒤, 2) 기존 방법 및 제안된 방법에 의한 검색 결과를 출력한다. 피험자는 각 결과에 대해 0~100 점 사이의 값으로 만족도를 평가한다. 4) 이후, 1~3 번 과정을 피험자마다 30 회 반복한다.

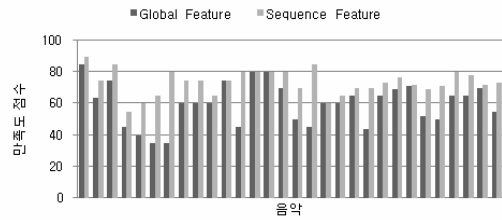
만족도 조사 결과는 그림 6 와 같다. 실험 결과 사용자 만족도의 평균은 기존 방법에서 59.8, 제안된 방법에서 73.5 였다. 따라서, 기존 방식보다 제안된 방식이 사용자를 좀더 만족시킴을 알 수 있었다.

5. 결론 및 향후 연구

본 논문에서는 하나의 음악에 존재하는 다수의 무



(그림 5) 유사 전역 특성을 갖는 음악에 대한 시퀀스 유사도



(그림 6) 시스템의 사용자 평가

드에 대해 무드 세그먼트를 정의하고, 각 무드 세그먼트에서 특성 추출 및 분류 과정을 통해 무드 시퀀스를 생성해냈다. 또, 유사한 무드 시퀀스를 가지는 음악에 대해 효과적으로 검색하는 시스템을 제안했다. 이를 통해 음악의 무드 변화에 대해 능동적으로 대응할 수 있는 검색 시스템을 제안하였다. 실험을 통해 기존 방식보다 제안한 방식이 좀더 뛰어난 결과를 보였으며, 사용자의 만족도 또한 증가함을 보였다.

향후 현재 사용된 음악 특성 외에 다양한 음악 특성을 고려하고, 좀더 발전된 시퀀스 생성 및 사상 기법을 적용하여 더 효과적인 음악 검색 방식의 연구를 진행할 것이다.

Acknowledgements

이 논문은 2006 년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2007-313-D00758).

참고문헌

- [1] 전상훈, 노승민, 황인준, “음악 무드와 감정의 퍼지 추론을 기반한 음악 검색 기법,” 정보처리학회논문지, vol.15, no.1, 2008.
- [2] Sanghoon Jun, Seungmin Rho, Byeong-jun Han, Eunjung Hwang, “A Fuzzy Inference-based Music Emotion Recognition System,” Visual Information Engineering , 2008.
- [3] L.Bergroth, H.Hakonen, T.Raita, “A Survey of Longest Common Subsequence Algorithms,” Proceedings of the Seventh International Symposium, 2000.
- [4] Daniel S. Hirschberg, “Algorithms for the Longest Common Subsequence Problem,” Journal of ACM (JACM), vol.24, no.4, Oct. 1977.