

시퀀스 기반의 유사 음악 검색 기법

Sequence-based Similar Music Retrieval Scheme

전 상 훈*, 황 인 준**
Sanghoon Jun*, Eunjung Hwang***

Abstract

Music evokes human emotions or creates music moods through various low-level musical features. Typical music clip consists of one or more moods and this can be used as an important criteria for determining the similarity between music clips. In this paper, we propose a new music retrieval scheme based on the mood change patterns of music clips. For this, we first divide music clips into segments based on low level musical features. Then, we apply K-means clustering algorithm for grouping them into clusters with similar features. By assigning a unique mood symbol for each cluster, we can represent each music clip by a sequence of mood symbols. Finally, to estimate the similarity of music clips, we measure the similarity of their musical mood sequence using the Longest Common Subsequence (LCS) algorithm. To evaluate the performance of our scheme, we carried out various experiments and measured the user evaluation. We report some of the results.

요 약

음악은 다양한 하위 레벨 음악 특징을 통하여 인간의 감정을 유발시키거나 음악적 무드를 만들어낸다. 보통 음악은 하나 이상의 무드로 구성되며 이것은 음악간 유사도를 결정하는 데 주요한 단서로 사용된다. 본 논문에서는 음악의 무드 변화 패턴을 기반으로 하는 새로운 음악 검색 기법을 제안한다. 이를 위해서, 우선 모든 음악에 대해 유사한 하위 레벨 특징을 가지는 세그먼트로 나누고, K-means 군집화 알고리즘을 적용하여 유사한 특징을 가지는 클러스터로 그룹화한다. 각 클러스터에 대해 유일한 무드 심볼을 정의하고 나면, 각 음악의 무드 변화 패턴은 일련의 무드 심볼 시퀀스로 표현이 가능하다. 마지막으로 음악간 유사도를 측정하기 위해서 longest common subsequence (LCS) 알고리즘을 적용한다. 제안된 검색 기법의 성능을 측정하기 위해 다양한 실험과 사용자 만족도 조사를 수행하고 결과를 분석한다.

Key words : Similar music retrieval, music mood, longest common subsequence, feature extraction

1. 서론

디지털 미디어에 대한 수요가 급증함에 따라 내용 기반 검색 (content based retrieval) 은 멀티미디어 데이터의 활용에 있어 중요한 위치를 차지하게 되었

다. 내용기반 검색 기법은 검색 대상이 되는 미디어의 특성에 크게 좌우된다. 음악의 경우, 디지털 녹음된 음악이 가지는 다양한 하위 레벨(low level) 음향 특징 (acoustic feature)이 검색 조건으로 사용될 수 있다. [1-7].

최근 들어, 음악 검색을 위한 다양한 접근 방식이 새롭게 제안되었다. [8,9]에서는 음향의 특징보다 한 단계 더 추상화된 템포, 조성, 하모니 등과 같은 음악 특징에 중점을 두었다. 특히, [8]에서는 음악 무드의 변화를 기반으로 하여 무드 트래킹 및 탐지를 자동적으로 수행할 수 있는 방법을 제시하였다..

한편, 음악의 특징과 음악 무드 및 듣는 사람의 감정 사이의 상관관계를 정의하기 위해 많은 연구가 진

* 고려대학교 전자전기공학과 석사과정
** 고려대학교 전기전자전파공학과 부교수, 교신저자
★ 교신저자 (Corresponding author)
※ 감사의 글 (Acknowledgment)
이 논문은 2007년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2007-313-D00758).

행되어 왔다 [10]. 실제로 인간의 지각 관점에서 우리는 시끄러움의 정도, 템포 및 조성과 같은 음악의 특징을 통해 음악과 음악 무드를 구분할 수 있다. 이를 기반으로 본 연구에서는 음악의 검색과 분류를 위해 음악의 특징을 고려한다.

클래식과 같은 장르에서는 하나의 음악에 뚜렷하게 구분되는 두 가지 이상의 음악 무드가 나타난다. 예를 들어, 베토벤 5번 교향곡의 경우 음악 무드는 웅장함으로 시작하여 경쾌한 무드로 마무리된다. 실제로 이와 같은 무드 사이에는 음악 특징의 큰 차이가 나타난다. 우리는 지난 연구[1]에서는 음악의 분류와 검색을 위한 특징 추출에 있어서, 음악의 전역적(global) 특징만을 고려하였다. 하지만 음악이 다수의 무드로 구성된 경우 전역적 특징만으로는 정확한 검색이 어렵다. 따라서, 본 논문에서는 다양한 음악 특징을 기반으로 무드의 변화를 감지하고 이러한 변화 패턴으로 유사한 음악을 검색하는 기법을 제안한다. 우선 음악 특징에 따라 음악을 여러 개의 세그먼트로 분할하고, 유사한 특징을 갖는 세그먼트들을 같은 그룹으로 분류한다. 분류된 각 그룹마다 유일한 심볼을 정의하면 각 음악은 이러한 심볼의 시퀀스로 표현이 가능하다. 최종적으로 시퀀스 비교 알고리즘을 통하여 음악 간의 유사도를 측정한다.

II. 관련 연구

본 장에서는 먼저, 음악 무드 기반의 음악 검색과 관련된 기존 연구에 대하여 서술하고 시퀀스 유사도 측정을 위한 longest common sequence 알고리즘에 대하여 소개한다.

1. 음악 무드

우리는 이전 연구에서 음악의 시끄러움 정도, 하모니, 템포, 조성 등의 음악 특징과 음악 무드 사이에서 밀접한 관계가 있음을 보였다 [1]. 또한 이러한 결과를 바탕으로 음악 추천과 검색에 있어서 음악의 특징을 통해 음악 무드를 예측할 수 있음을 증명하였다.

Yazhong Feng 外[9]는 템포와 같은 음악 특징을 이용한 음악 무드 기반의 음악 검색 모델을 제안하였다. Lie Lu 外[8]는 강도(intensity), 음색 및 리듬 등을 이용하여 음악 무드 인식을 자동화한 계층적인 프레임워크를 제안하였다.

특히, [8]에서는 무드 인식에 있어서 음악에서 나타나는 하나 이상의 무드 변화에 기반한 접근을 시도하였다. 하나의 음악을 무드 변화에 따라 몇 가지의 독립적인 무드를 갖는 세그먼트로 나누는 작업을 자동화하여 수행하였다

립적인 무드를 갖는 세그먼트로 나누는 작업을 자동화하여 수행하였다

2. Longest common subsequence

Longest common subsequence는 두 시퀀스간의 가장 긴 공통되는 시퀀스를 찾아내는 알고리즘이다 [11]. 예를 들어, $S_1 = \text{AmERICA}$ 와 $S_2 = \text{Africa}$ 의 두 시퀀스 사이에는 RICA와 ARI 등의 공통된 문자를 갖는 다수의 공통 시퀀스가 존재한다. 그러한 공통 시퀀스 중에서 가장 긴 공통 시퀀스를 그들의 LCS(longest common sequence)라 하며, 위의 S_1 과 S_2 의 경우 ARICA가 두 시퀀스의 LCS가 된다 [12,13].

$S_1[1..i]$ 와 $S_2[1..j]$ 를 시퀀스 S_1 와 S_2 의 순차적 집합이라 정의하면, $LCS(S_1, S_2)$ 는 다음의 식에 따라 두 시퀀스의 조합인 (S_1, S_2) 의 가장 긴 공통 시퀀스의 길이, 즉 LCS의 길이를 반환한다.

$$LCS = R[i, j] \quad (1)$$

$$R[i, j] = \begin{cases} 0 & (i=0 \text{ or } j=0) \\ R[i-1, j-1] & (S_1[i] = S_2[j]) \\ \max\{R[i-1, j], R[i, j-1]\} & (S_1[i] \neq S_2[j]) \end{cases} \quad (2)$$

따라서, 이 수식에 따라 시퀀스 S_1 과 S_2 에 대한 $LCS(S_1, S_2)$ 는 가장 긴 공통 시퀀스인 ARICA의 길이, 즉 5를 반환한다.

LCS는 DNA 유사도 측정과 영상 시퀀스 측정 및 웹 브라우징 패턴인식 등 다양한 분야에서 널리 사용되고 있다. 또한 음악은 시간순의 특징 벡터를 갖기 때문에 관련 연구에서는 음악 무드 시퀀스의 유사도 측정에 LCS 알고리즘을 사용할 수 있다.

III. 유사음악검색

본 장에서는 음악 무드 시퀀스 기반의 음악 검색 기법을 제안한다. 검색을 위한 전체 데이터베이스 구축 절차는 그림 1의 (a)와 같다. 데이터베이스의 구축을 위해서는 음악의 intensity를 기반으로 세그먼트로 분할하고, 각 세그먼트의 특징에 따라 분류 및 시퀀스 매핑 과정을 통하여 각 음악을 음악 무드 시퀀스로 변환한다. 그림 1의 (b)는 주어진 질의 음악과 유사한 음악을 검색하기 위한 전체적인 흐름을 보여주고 있다.

주어진 음악의 무드 세그먼트는 그 음악에 존재하는 하나 이상의 음악 무드의 변화를 기준으로 나누는 음악의 구성 요소를 나타낸다. 즉 무드가 변하는 시점은 바로 무드 세그먼트간의 경계가 된다. 또한, 음악 무드 시퀀스는 무드가 변화하는 음악의 무드 세그

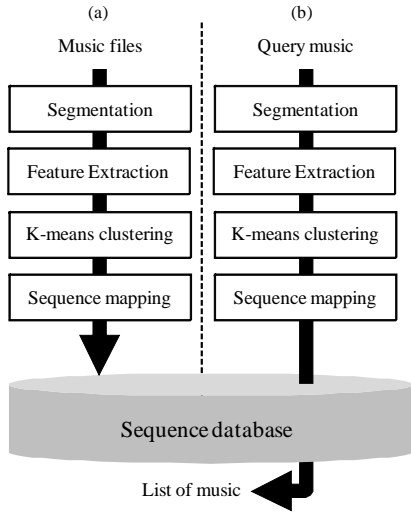


Fig. 1. Overall architecture of proposed retrieval scheme

그림 1. 제안된 검색 기법의 전체 구조

먼트들의 시간적인 시퀀스이다.

1. 무드에 따른 음악 분할

본 논문에서는 두 가지 이상의 무드를 갖는 음악을 대상으로 데이터 셋을 구성한다. 음악의 무드의 경계를 찾는 작업을 자동화하기 위하여 [8]에서 제안된 무드 경계 탐색 기법을 적용하였다.

음악 무드 변화의 경계를 찾기 위해 프레임 단위 신호의 특성을 나타내는 intensity의 변화를 이용한다. intensity는 소리 신호 윈도우의 표준편차를 계산함으로써 도출할 수 있다. intensity의 평균 μ 와 표준편차 σ 를 사용하여 다음의 수식과 같이 두 가지 임계치(threshold)를 정의할 수 있다.

$$Th1 = \mu - 0.5\sigma \tag{3}$$

$$Th2 = \mu + 0.5\sigma \tag{4}$$

변화하는 intensity가 두 임계값 중 하나를 가로지르는 시점을 잠재적인 경계로 정의하고, 가장 인접한 intensity의 계곡(valley)으로 조정을 수행한다. 구간이 짧은 경계를 제거하고, intensity의 차이가 크지 않은 경계를 병합함으로써 무드의 경계를 설정한다. 상기에 해당하는 과정은 그림 2에 나타나 있다. 설정된 경계를 기준으로 음악은 보통 3개에서 7개까지의 무드 세그먼트로 분할되었다.

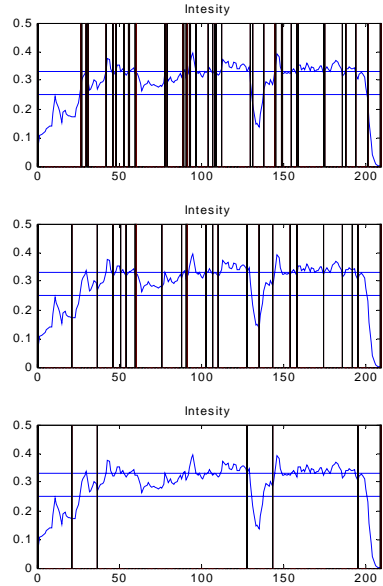


Fig. 2. Processing of mood boundary tracking

그림 2. 무드 경계 탐색의 처리 과정

2. 특징 추출

음악 무드의 세그먼트들을 비슷한 특징을 가지는 집단들로 분할하기 위해서 다양한 음악 특징들을 추출한다. 본 논문에서는 이러한 목적으로 템포, 조성, 멜로디, 시끄러움의 정도, 하모니와 같은 음악 특징과 spectral centroid, spectral rolloff, spectral flux, time domain zero crossing, mel-frequency cepstral coefficient와 같은 음색 특징을 고려하였다 [14].

가. 음악 특징

음악이 가지는 템포, 조성, 시끄러움의 정도와 같은 음악 특징은 하나의 스칼라 값으로 표현할 수 있다. 예를 들어, 어떤 음악은 평균적으로 120BPM(템포)과 C장조(조성) 및 20dBv의 시끄러움 정도를 가지는 것으로 표현할 수 있다. 하지만 하모니 분포와 같은 특징들은 하나의 스칼라의 값으로 표현하기 어려우며 대개 다차원의 벡터와 같은 표현 방식을 사용한다. 이러한 음악 특징의 추출에 대한 자세한 내용은 [14]에서 설명하고 있다.

나. 음색 특징

(1) Spectral centroid는 STFT(Short Time Fourier Transform)로 얻어진 스펙트럼 magnitude의 무게 중

심을 나타내는 스펙트럼 형태의 측정 방법 중 하나이다. 대개 높은 주파수를 갖는 '밝은' 음색의 경우 spectral centroid는 높은 값을 갖게 된다. Spectral centroid는 다음과 같은 수식으로 표현할 수 있다.

$$C_t = \frac{\sum_{n=1}^N M_t[n] * n}{\sum_{n=1}^N M_t[n]} \quad (5)$$

(2) Spectral rolloff는 가장 낮은 주파수부터의 순차적인 합이 전체 magnitude의 85%에 해당하는 값을 가질 때의 주파수 R_t 로 정의되는 스펙트럼 형태의 또 다른 측정 방법이다. 보통 낮은 R_t 를 가질 경우 저역의 주파수에 밀집되어 강한 저음의 음색을 갖는다. Spectral rolloff는 다음과 같은 수식으로 표현할 수 있다.

$$\sum_{n=1}^{R_t} M_t[n] = 0.85 * \sum_{n=1}^N M_t[n] \quad (6)$$

(3) Spectral flux는 연속적인 스펙트럼 분포의 정규화된 magnitude 차이에 제곱을 함으로서 도출되며 스펙트럼 변화의 측정에 사용된다. 특히, 음악에 사용되는 악기와 밀접한 관련이 있다. Spectral flux를 구하는 계산식은 다음과 같다..

$$F_t = \sum_{n=1}^N (N_t[n] - N_{t-1}[n])^2 \quad (7)$$

(4) Time Domain Zero Crossings는 시간 축에서 오디오 신호 파형의 위상이 중심축을 통과하는 회수를 고려한 신호의 잡음 정도(noisiness)를 나타내며 구하는 수식은 다음과 같다.

$$Z_t = \frac{1}{2} \sum_{n=1}^N |sign(x[n]) - sign(x[n-1])| \quad (8)$$

(5) Mel-Frequency Cepstral Coefficients(MFCC)는 인간의 청각에 전달되는 주파수의 관점으로 모델링된 mel-frequency를 기준으로 신호를 고려하는 방법이다. MFCC 계산의 결과로서 13가지의 계수를 얻을 수 있는데, 본 논문에서는 음악의 특징 표현에 6번째부터 13번째까지의 계수를 사용하여 보다 효과적인 음색 특징 추출이 되도록 한다. [15]에서는 MFCC를 계산하는 자세한 과정이 나와 있다

다. 군집화

K-means 알고리즘은 데이터 군집화 알고리즘으로서 데이터 분류 및 컴퓨터 영상 처리 분야에서 다양하게 사용되고 있다. K-means 알고리즘은 빠르고, 가벼우며 효율적이기 때문에, 본 논문에서는 수많은 무드 세그먼트의 분류를 위해 K-means 알고리즘을 사용하였다. 분류를 위하여 우선 900여개에 달하는 무드 세그먼트들을 대상으로 다양한 특징을 추출하였다. 실험적으로 효과적인 K값을 설정하고 K-means 군집화 알고리즘을 적용하였다. 무드의 세그먼트는 특징 정보에 따라 다섯 개의 개별 그룹으로 군집화되었다.

라. 시퀀스로의 변환

음악 무드 시퀀스를 보다 효율적으로 다루기 위하여 하나의 음악을 문자열로 이루어진 시퀀스로 변환하는 과정이 필요하다. 본 논문에서는 앞서 분류된 각 클래스에 독립적인 문자를 심볼로 부여하여 각 음악을 구성 세그먼트에 해당하는 심볼의 시퀀스로 변환하였다. 예를 들어 C_1 부터 C_N 까지 N개의 클래스로 이루어진 음악 데이터가 존재하고, m_1 부터 m_4 까지 4개의 세그먼트로 구성되어있는 음악 m 이 존재한다고 가정하자. K-means 군집화 알고리즘에 의하여 $m_1 \in C_1$, $m_2 \in C_5$, $m_3 \in C_3$, $m_4 \in C_6$ 의 조건이 성립한다면, 음악 m 은 $\{C_1, C_5, C_3, C_6\}$ 의 시퀀스로 표현된다.

마. 유사도 측정

두 시퀀스 사이의 유사도를 측정하기 위하여 본 논문에서는 [16]에서 제안된 비디오 시퀀스의 LCS 유사도 측정 방법을 사용하였으며 그 수식은 다음과 같다.

$$LCSsim(S_q, S_d) = \frac{LCS(S_q, S_d)}{|S_q|} \quad (9)$$

S_q 와 S_d 는 각각 질의 시퀀스와 데이터베이스의 시퀀스를 나타낸다. $LCS(S_q, S_d)$ 는 두 시퀀스 사이의 LCS의 길이를 계산한다. 이를 질의 시퀀스 S_q 의 길이로 나누면 두 시퀀스 사이의 유사도를 계산할 수 있다.

IV. 실험 및 결과

본 논문에서 제안하는 유사 음악 검색 기법의 효율성을 평가하기 위하여 다양한 실험을 수행하였다. 이장에서는 자세한 실험 내용 및 결과를 기술한다.

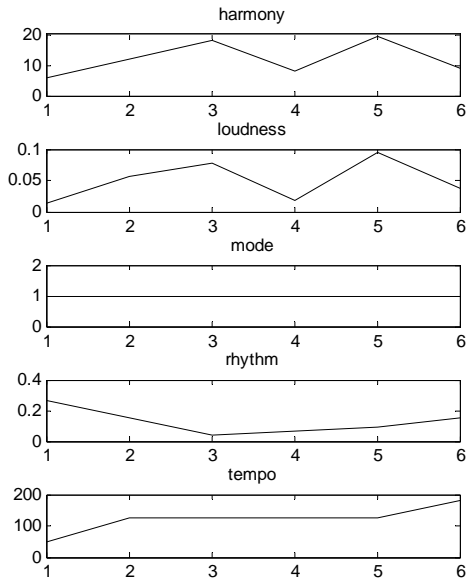


Fig. 3. Variation of musical features
 그림 3. 음악 특징의 변화

첫 번째 실험에서는 전역적인 음악 특징보다는 본 논문에서 제안하는 무드 변화 패턴을 이용하는 것이 보다 정확한 검색 결과를 얻을 수 있음을 보인다. 이러한 결과를 설명하기 위해서 다음 실험에서는 전역적인 음악 특징 추출의 경우 부분별 특징의 희석 위험성을 보이고, LCS 알고리즘을 이용한 세그먼트 기반 음악 특징 추출이 보다 적합함을 보인다. 세 번째 실험에서는 제안된 기법의 유효성을 확인하기 위해 원곡과 리메이크 곡으로 이루어진 두 음악 사이의 유사도를 측정하였다. 마지막으로 사용자의 평가를 통해 이전 방식보다 제안된 방식이 유사 음악검색에 효과적임을 보인다.

본 논문에서는 팝송 및 국내가요로 구성된 180개의 음악 데이터를 사용한다. 또한 모든 음악 데이터는 하나 이상의 뚜렷한 음악 무드의 차이를 보이는 음악들을 대상으로 수집하였다.

가. 무드 세그먼트별 특징 벡터 분석

제안된 검색 방식에 따라 나누어진 각 음악 세그먼트들을 대상으로 다양한 음악 특징을 추출하였다. 그림 3은 각 무드 세그먼트별 음악 특징의 변화를 나타낸다. 그림에서와 같이 같은 음악이라도 무드의 변화에 따라 특징 벡터의 변화에 큰 차이가 있음을 볼 수

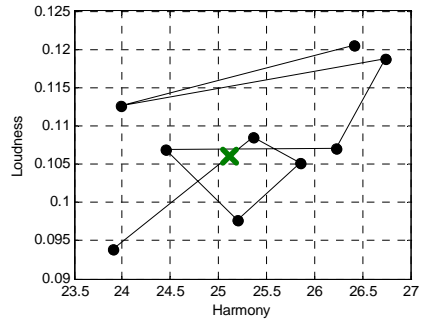


Fig. 4. Harmony-loudness variation
 그림 4. 음악의 harmony-loudness 변화

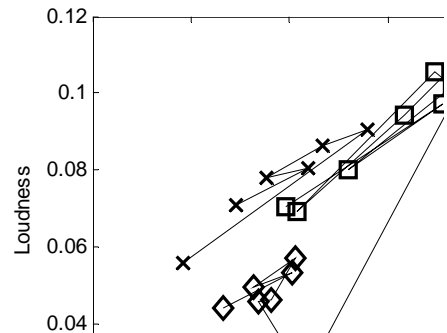


Fig. 5. Harmony-loudness variation of 3 songs
 그림 5. 3가지 음악의 harmony-loudness 변화

있다.

예를 들면, 그림 4는 하모니와 시끄러움의 정도를 나타내는 두 축으로 이루어진 특징 공간(feature space)에 한 음악의 음악 특징 변화 추이를 보여주고 있다. 그림에서 음악 무드별 세그먼트 분할 없이 추출된 전역 음악 특징을 'x'로 표시하였고, 세그먼트로 분할된 음악 특징의 변화는 '●'로 표시하고 있다. 결과적으로 여러 번의 음악 무드 변화를 가지는 음악에 대해 전역적 음악 특징을 사용할 경우 각 무드의 특징이 희석되어 사라질 수 있는 위험성을 보여준다.

그림 5에서는 서로 다른 세 음악에 대하여 하모니(harmony)와 시끄러움의 정도(loudness)로 이루어진 음악 특징 공간상의 특징 변화를 나타내고 있다. 각 음악의 무드 세그먼트들은 X, ◇, □의 도형으로 표현되어 있다. 그림으로 부터 각 음악은 별개의 무드별 특징 변화 패턴을 가지고 있음을 볼 수 있다.

나. 전역적 특징 추출방식의 LCS 유사도 측정

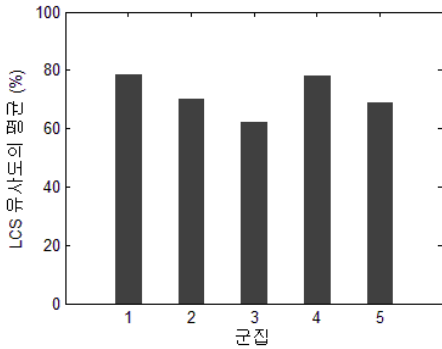


Fig. 6. Sequence similarity under global feature based clustering

그림 6. 전역 특징 추출을 이용한 군집화 결과의 LCS 유사도 측정

이 실험에서는 음악의 무드별 분할 없이 특징을 추출하는 전역 음악 특징 추출 방식의 효율성을 평가한다. 우선 전체의 음악에 대하여 전역 음악 특징을 추출하고, K-means 군집화 알고리즘을 통하여 음악을 분류하였다. 전역 특징 추출을 통해 같은 군집으로 분류된 음악들을 대상으로 무드 세그먼트로 분할한 뒤, 특징 추출 및 분류를 통하여 각각의 무드 시퀀스를 생성하였다. 최종적으로 같은 군집상의 음악에 대해 LCS 알고리즘을 사용하여 유사도를 측정한다. 전역 음악 특징으로 분류된 군집에서 가능한 모든 쌍을 대상으로 유사도를 측정한 뒤 그 평균을 구하면 그림 6과 같다. 실험 결과에 따르면 군집 1이 가장 높은 유사도 평균값인 78.6%을 나타내고 군집 3이 가장 낮은 유사도 평균값인 62.3%을 나타내고 있다. 군집들의 유사도 평균값의 전체 평균은 71.6% 이다. 실제로, 낮은 유사도 평균값을 갖는 군집에 소속된 음악의 경우 무드 시퀀스간의 상이한 차이를 가짐을 볼 수 있다. 이에 따라 전역적 음악 특징 추출은 음악의 특징을 정확히 나타내기에 한계가 있음을 알 수 있고, 음악 검색 성능에 부정적인 영향을 미칠 수 있음을 알 수 있다.

다. 유사한 음악의 유사도 측정

본 실험에서는 원곡과 리메이크 곡의 쌍으로 구성된 음악에 대하여 제안된 검색 기법의 성능을 평가한다. 보통 원곡과 리메이크 곡은 서로 매우 흡사한 음악의 특징을 갖고 대부분 동일한 무드 변화의 특징을 나타낸다[17]. 각각의 쌍에 대하여 제안된 검색 기법을 적용하여 LCS 알고리즘을 통한 유사도 측정을 수행한 결과는 그림 7과 같다. 또한 전역적 음악 특징

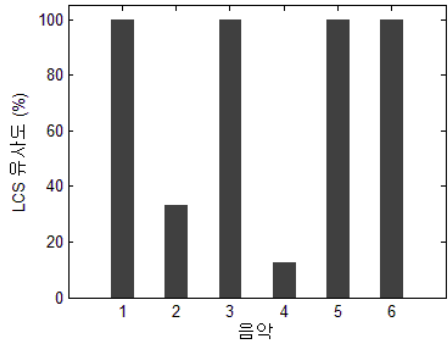


Fig. 7. LCS similarities between original and remake song

그림 7. 원곡과 리메이크곡의 LCS 유사도

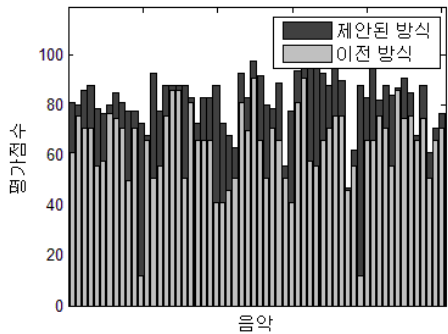


Fig. 8. User evaluation result

그림 8. 사용자 평가 결과

Table 1. Song clustering based on global feature

표 1. 원곡과 리메이크곡의 동일군집소속 여부

음악	동일군집소속 여부
1. Dont't stop me now	예
2. Surfin' USA	예
3. The power of love	예
4. How deep is your love	아니오
5. This love	예
6. Lemon tree	예

추출 방식을 적용하였을 때 같은 군집으로 분류되는지 여부를 실험한 결과는 표 1과 같다. 실험 결과 중 "Surfin' USA"의 경우 전역적 음악 특징 추출 방식에서는 동일한 군집에 속하지만 유사도 측정에 있어서는 상대적으로 낮은 음악 특징을 보인다. 실제로 해당 리메이크 곡은 리듬과 음색 부분에서 원곡의 음악 특징이나 노래 구성에 차이가 있음을 알 수 있다.

라. 사용자 평가

본 장에서는 유사 음악 검색 결과에 따른 사용자 만족도 평가 실험에 대하여 기술한다. 음악을 즐겨듣는 대학생으로 이루어진 20명의 피험자들을 대상으로 무작위로 선택된 각각 다른 10 곡의 대중적인 가요 음악을 들려주었다. 각 곡에 대해 제안하는 방식과 전역적 음악 특징 추출 방식에 따라 유사도가 가장 높은 음악을 제시하였을 때의 만족도를 0에서 100사이의 점수로 입력하도록 진행하였다. 여기서 만족도란 실험에서 두 곡의 음악 간의 유사도에 대한 피험자의 정량적 평가를 의미한다. 그림 8은 180여개의 음악 중 60개의 만족도 조사 결과를 나타낸다. 전체 180여개의 음악의 만족도 평균 점수는 기존 방식이 65.07, 제안하는 방식이 82.57로 제안하는 방식에 의한 결과에 더 높은 만족도를 보였다.

V. 결론

본 논문에서는 음악 무드의 변화를 기반으로 하는 새로운 음악 검색 기법을 제안하였다. 우선 무드의 변화별로 세그먼트를 분할하고, 각 세그먼트에 대하여 여섯 가지의 음악 특징을 추출한 뒤 K-means 군집화 알고리즘을 적용하여 세그먼트를 분류하였다. 각각의 군집 즉 클래스에 대하여 독립적인 심볼을 부여함으로써 각 음악을 심볼의 시퀀스로 표현하였다. 질의 음악 무드 시퀀스와 데이터베이스에 저장된 무드 시퀀스의 유사도를 측정하기 위하여 LCS 알고리즘을 사용하였다. 실험에서는 기존의 무드 세그먼트 분할 과정이 없는 전역적 음악 특징 추출 방식과 제안하는 방식의 검색 정확도 및 만족도 분석을 통하여 음악 무드별 분할의 우수성을 확인하였다.

참고문헌

[1] Sanghoon Jun, Seungmin Rho, Byeong-jun Han, Eunjung Hwang, "A Fuzzy Inference-based Music Emotion Recognition System," *Visual Information Engineering*, pp.673-677, Aug. 2008

[2] Hoashi, Zeitler, Inoue, "Implementation of relevance feedback for content-based music retrieval based on user preferences," *ACM SIGIR*, pp.385-286, 2002

[3] Hoashi, Matsumoto, Inoue, "Personalization of user profiles for content-based music retrieval based on relevance feedback," *ACM Multimedia*, pp.110-119, 2003

[4] Carol Krumhansl, *Cognitive Foundations of Musical Pitch*, Oxford University Press, 1990

[5] S. Rho, B. Han, E. Hwang and M. Kim, "MUSEMBLE: A Novel Music Retrieval System with Automatic Voice Query Transcription and Reformulation," *Journal of Systems and Software*, Vol. 81(7), pp. 1065-1080, 2008

[6] Hung-Chen Chen and Arbee L. P. Chen, "A Music Recommendation System Based on Music and User Grouping," *Journal of Intelligent Information Systems*, Vol. 24(3), pp. 113-132, 2005

[7] A. Ghias, et al., "Query by humming - musical information retrieval in an audio database," *Proceedings of ACM Multimedia 95*, pp.231-236, 1995

[8] Lie Lu, Dan Liu, Hong-Jiang Zhang, "Automatic Mood Detection and Tracking of Music Audio Signals," *IEEE Transactions on Audio, Speech and Audio Processing*, vol.14, no.1, pp. 5-18, Jan. 2006

[9] Yazhong F, Yueting Z, Yunhe P, "Music information retrieval by detecting mood via computational media aesthetics," *IEEE/WIC International Conference on*, pp.235-241, Oct. 2003

[10] P.N. Juslin, J.A. Sloboda, *Music and Emotion: Theory and research*, New York: Oxford University Press, 2001.

[11] L. Bergroth, H. Hakonen, T. Raita, "A Survey of Longest Common Subsequence Algorithms," in *Proceedings of the 7th International Symposium*, 2000.

[12] S. Daniel, Hirschberg, "Algorithms for the Longest Common Subsequence Problem," *Journal of ACM*, vol.24, no.4, Oct. 1977.

[13] R.A. Wagner and M.J. Fischer, "The String-to-String Correction Problem," *Journal of the ACM*, vol.21, no.1, pp 168-173, Jan. 1975

[14] Tzanetakis, G. and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Audio, Speech and Audio Processing*, 10(5): 293-302. 2002

[15] L.R. Rabiner, B.-H. Juang, *Fundamentals of Speech Recognition*, Prentice Hall, New Jersey, 1993.

[16] Young-tae Kim and Tat-Seng Chua, "Retrieval of News Video Using Video Sequence Matching," *Multimedia Modelling Conference, 2005. MMM 2005. Proceedings of the 11th International*, pp.

68-75, 12. Jan. 2005

[17] Daniel P.W. Ellis, Graham E. Poliner, "Identifying 'Cover Songs' with Chroma Features and Dynamic Programming Beat Tracking," *IEEE Conferences on Acoustics, Speech, and Signal Processing*, vol.IV, pp.1429-1432, 2007

저 자 소 개

전 상 훈 (학생회원)



2008년 : 고려대학교
전기전자전파공학부 졸업 (공학사)
2008년 ~ 현재: 고려대학교
대학원 전기전자공학과 (석사과정)
<주관심분야> 데이터베이스,
멀티미디어 검색, 내용기반
음악검색 및 음악추천시스템

ysbhjun@korea.ac.kr

황 인 준 (정회원)



1988년 : 서울대학교 컴퓨터공학과 졸업 (공학사)
1990년 : 서울대학교 컴퓨터공학과 (공학석사)
1998년 : Univ. of Maryland at College Park 전산학과 (공학박사)

1998~1999: Bowie State Univ., 조교수
1999~1999: Hughes Research Lab. 연구교수
1999~2003: 아주대학교 정보통신전문대학원 조교수
2003~2004: 아주대학교 정보통신전문대학원 부교수
2004~2007: 고려대학교 전기전자전파공학과 조교수
2008~현재: 고려대학교 전기전자전파공학과 부교수
<주관심분야> 데이터베이스, 멀티미디어 검색,
정보통합, 전자상거래, 영상처리, 유비쿼터스 컴퓨팅
ehwang04@korea.ac.kr