

음악 무드와 감정의 퍼지 추론을 기반한 음악 검색 기법

전상훈*, 노승민**, 황인준*
*고려대학교 전자전기공학과
**아주대학교 정보통신전문대학원
e-mail : *{ysbhjun,ehwang04}@korea.ac.kr, **anycall@ajou.ac.kr

A Music Retrieval Scheme based on Fuzzy Inference on Musical Mood and Emotion

Sanghoon Jun*, Seungmin Rho**, Eenjun Hwang*
* School of Electrical Engineering, Korea University
** Graduate School of Information and Communication, Ajou University

요 약

최근 오디오 압축 기술의 발전에 힘입은 디지털 음원과 웹 스트리밍의 보급으로, 사용자가 음악 정보에 손쉽게 접할 수 있게 되었다. 이에 따라 음악을 보다 쉽고 효율적인 방법으로 검색하는 방법뿐 아니라 사용자의 환경에 따라 적절한 음악을 검색할 수 있는 기능의 필요성이 증가하게 되었다. 본 논문에서는 음악의 특징에 따라 분류된 데이터베이스를 사용하고, 사용자의 감정을 분석하여 적절한 음악을 검색하는 시스템을 제안한다. 본 시스템은 사용자의 감정 입력을 효율적으로 처리하기 위한 방법으로 Thayer의 2D emotional space를 적용하여 Valence-Arousal model의 두 가지의 입력을 처리한다. 가장 적합한 음악의 정보를 얻기 위해 사용된 Fuzzy Inference System의 IF-THEN 규칙을 정의하기 위하여 언어적으로 정의된 기존의 음악 감정 연구 결과를 적용하였고, 도출된 결과와 가장 유사도가 깊은 음악을 우선적으로 검색하도록 설계하였다. 이와 같이 구현된 시스템의 타당성을 검증하기 위해 사용자 설문조사를 수행하였다.

1. 서론

월드 와이드 웹으로 인해 본격화된 멀티미디어 정보의 홍수 속에서 온라인 음원의 광대한 보급은 효율적이고 정확한 자동화된 음악 정보 처리와 검색의 필요성을 증가시켰다. 또한 음악의 분위기에 따른 분류와 검색의 필요는 최근 음악치료 (Music Therapy) 등에서 볼 수 있듯이 그 중요성이 증가하고 있다.

음악특징과 감정 사이의 상호관계를 정의하기 위하여 오랜 기간 동안의 연구가 지속되어왔다 [1, 2]. 이에 따라 감정모델과 음악특징을 수학적 상호관계로 연계하기 위한 연구도 진행되고 있다.

본 논문에서는 사용자 감정 분석과 퍼지추론에 기반한 음악 검색 시스템을 제안한다. Thayer [3]의 Valence-Arousal Model을 통하여 사용자의 감정을 입력 받고, IF-THEN 규칙을 적용하여 얻어진 적합한 음악특징을 토대로 유사도가 가장 큰 음악을 데이터베이스에서 선택한다.

2. 관련연구

사용자의 감정상태를 분석하기 위해서는 사람의 감정이 갖는 모호함에 대한 대안을 생각해야 한다. 감정을 표현하기 위해서는 수많은 형용사가 사용될 수

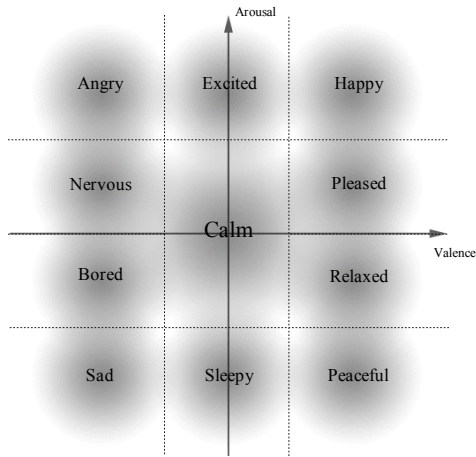
있지만, 감정상태를 선형적으로 나누기는 불가능하다. 이러한 감정의 모호함의 대안으로 Thayer의 Valence-Arousal Model을 사용하고 이를 처리하기 위한 방법으로 퍼지 추론 시스템을 적용한다.

2.1. Valence-Arousal 모델

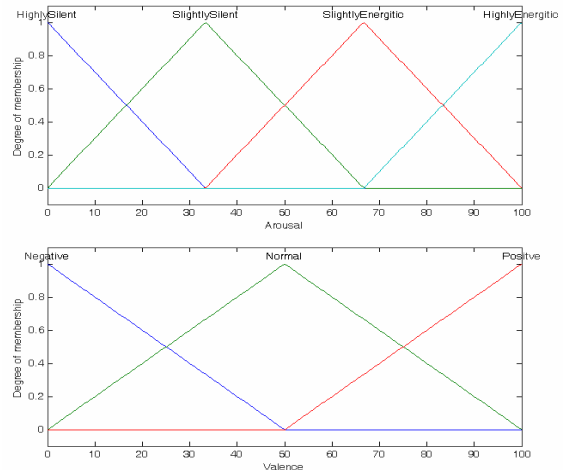
앞서 설명한 것처럼 본 시스템에서는 사용자의 감정을 효과적으로 인식하기 위하여 사용자 감정인식에서 보편적으로 많이 사용되는 모델인 Thayer의 Valence-Arousal 모델을 사용한다. 그림 1에서 보는 바와 같이 Valence-Arousal 모델은 2차원의 감정 공간(2DES: 2 Dimensional Emotion Space)으로 표현할 수 있다. 2DES의 Arousal에 대한 구분으로 제 1, 2사분면은 보다 능동적이고 활동적인 (Energetic) 감정을 표현하고, 제 3, 4사분면은 수동적이고 소극적인 (Silent) 감정을 나타낸다. Valence상의 제 1, 4사분면은 긍정적인 (Positive) 감정을, 제 2, 3사분면은 부정적인 (Negative) 감정을 나타낸다.

2.2 퍼지 추론 시스템

퍼지 추론 시스템이란 기존의 수학적 방법으로는 설명될 수 없는 시스템의 행동을 퍼지 이론에 기반한



(그림 1) Valence-Arousal Space 의 11 가지 감정



(그림 2) A/V 멤버십 함수의 정의

여 설명하기 위한 모델이다. 퍼지 시스템은 퍼지 규칙을 기본으로 구성된다.

퍼지 추론 시스템에서는 하나의 규칙을 나타내는 “IF A THEN B” 형식의 퍼지 조건문을 통해, 퍼지화된 입력 값을 퍼지 관계에 적용시키는 과정이 필요하다. 본 시스템에서는 Mamdani 형 퍼지 규칙을 통해 시스템을 구성한다.

퍼지 시스템에서 최종적으로 퍼지 출력 값을 얻기 위해서는 비퍼지화 (Defuzzification) 과정이 필요하다. 비퍼지화 과정을 통해 멤버십 함수의 확실적인 값에 해당하는 의도하는 퍼지 출력 값을 얻는다.

3. 감정 분석 기반 음악 검색 시스템

3.1 감정 분석기

본 시스템에서는 모호성을 갖는 감정의 상태를 표현하기 위하여 퍼지 추론 시스템을 도입한다. Valence-Arousal 의 두 가지의 백분율 입력을 받아 요구되는 음악의 분위기 (Mood)와 음악의 음조 (Pitch), 템포 (Bpm), 장·단조 (Tonality) 등을 출력한다.

Valence-Arousal 모델에서 사용자 감정에 대한 입력을 Valence 와 Arousal 의 두 가지의 입력을 받는다. 2DES 각각의 입력은 시스템으로 0 에서 100 사이의 백분율 값 (%)으로 입력된다.

가령, 사용자의 감정이 82%의 Valence 값과 78%의 Arousal 의 값을 갖는다면, 긍정적인고, 활동적인 “행복함”의 감정에 근접한 감정을 갖는다고 볼 수 있다. 하지만, 해당되는 값은 “행복함”의 감정에 대해 가까울 뿐, 명확한 정의를 내릴 수 없기 때문에 본 시스템에서는 퍼지집합을 도입한다.

본 시스템의 Valence 와 Arousal 의 입력은 그림 2 과 같은 퍼지 멤버십 함수 (Fuzzy Membership Function)를 사용하여 입력의 퍼지화를 수행한다.

3.2 Mamdani 형 퍼지 규칙에서의 추론

Valence, Arousal 두 가지의 입력 퍼지 집합을 정의한 뒤 이에 대한 규칙이 정의 되어야 한다. 객관성과 전문성을 높이기 위해 감정과 음악특징을 정의한 자

료를 토대로 퍼지 규칙을 정의하여 음악특징을 연결한다 [1, 2].

표 1. 퍼지 규칙의 예

| Input | valence | | arousal | |
|--------|---------|----------|----------|-----------------|
| | low | | high | |
| Output | tempo | tonality | loudness | Pitch variation |
| | fast | minor | high | small |

표 1 는 Anger 의 감정에 해당하는 퍼지 규칙으로서, 해당 감정에 대응되는 음악특징을 정의하고 있다. 음악 특징은 BPM (Beat Per Minute) 단위의 템포, 장·단조, 시끄러움의 정도, 높낮이의 변화율 등을 고려하여, 감정에 가장 적합한 음악을 검색할 수 있는 근거를 제공한다. 표 2 는 표 1 에서 정의한 규칙을 퍼지 조건문으로 표현한 것이다.

표 2. 퍼지 조건문의 예

| |
|--|
| IF (valence is low) and (arousal is high) THEN (tempo is fast) and (tonality is minor) and ... |
|--|

퍼지 조건문의 적용 후 비퍼지화를 통하여 음악특징의 결과를 얻는다. 추론된 값에 대하여 가장 근사한 값을 갖는 음악을 데이터베이스로부터 선택한다.

3.3 최소 (Min) 연산에 의한 대표 음악 감정 추론

본 시스템에서는 멤버십 값을 갖는 Valence 와 Arousal 에 대해서 적절한 음악특징뿐만 아니라 사용자가 대표적인 음악 감정에 대한 정보도 알 수 있도록 설계한다. Valence-Arousal Space 에서 나뉘어진 감정의 종류는 총 11 가지로서 그림 1 과 같다. 정의된 퍼지 집합은 그림 2 에서 표현된 바와 같이 Valence 가 3 가지, Arousal 은 4 가지로서 퍼지 멤버십 조합의 경우의 수는 총 12 가지가 정의된다. 12 가지의 감정 중 Arousal 과 Valence 의 강도가 적고 유사도가 깊은, 즉 원점에 가까운 두 감정은 동일한 대표 감정으로 판단한다.

각 감정에 대해 최소연산을 수행하고, 계산된 결과에 대하여 가장 높은 값을 갖는 음악 감정을 선택한다. 예를 들어 화남 (Angry) 감정에 해당하는 최소연산은 표 3 과 같이 표현할 수 있다. 11 개의 각 감정에 대하여 표 3 과 같은 연산을 수행하고, 가장 높은 멤버십 값을 갖는 대표 감정을 강도가 가장 높은 감정으로 판단하여 이를 선택한다.

<표 3> 음악 감정 추론의 최소(min)연산

| |
|--|
| Mood Selector : Angry |
| $\mu_{angry} = \mu_{arousal_high} \cap \mu_{valence_negative}$ |

4. 실험 및 결과

음악 검색 시스템에 대한 성능을 측정하기 위하여 사용자 성능 평가 방식을 도입하여 만족스런 결과를 얻는지 알아보았다.

4.1 사용자 평가

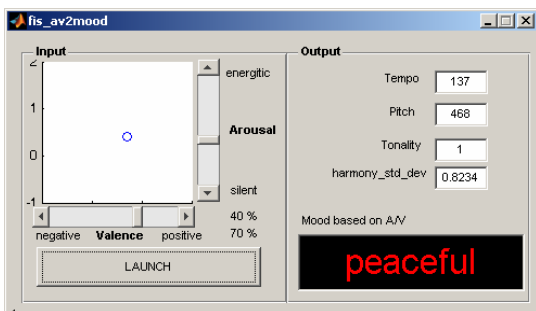
실험을 위해 음악특징이 정의가 되어있는 다양한 장르의 40 곡의 음악을 대상으로 데이터베이스를 사용하였다 [4]. Valence-Arousal 에 대한 적절한 지식을 가지고 있는 대학생 5 명의 사용자를 대상으로 직접 인터페이스 상의 입력을 통해 감정의 정도를 입력한다. 퍼지 추론 시스템을 통해 나온 음악특징에 대해 유사도 우선순위가 가장 높은 음악을 선택하여 기대했던 음악 감정을 느낄 수 있는지 만족도를 설문 조사하여 0~5 사이의 만족도 점수를 얻는다. 각 피험자는 10 번의 수행과정을 거친다.

4.2 사용자 인터페이스

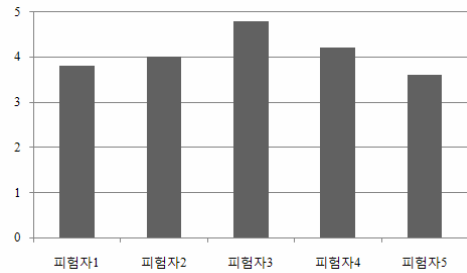
사용자의 Valence-Arousal 입력을 위해 보다 직관적인 인터페이스를 위해 그림 4 와 같이 Valence-Arousal 2 차원 공간입력을 사용한다 [5]. 사용자의 입력에 대하여 음악특징을 추론한 뒤 해당 Valence-Arousal 상태에 대한 음악특징과 근접한 대표적인 감정을 출력한다.

4.3 평가 결과

각 피험자에 대한 10 번의 수행 후 만족도를 조사한 결과, 피험자 각각 평균 3.8, 4, 4.8, 4.2, 3.6 의 점수



(그림 3) 시스템의 사용자 인터페이스



(그림 4) 피험자의 만족도 평균

로 응답하였다. 피험자 1 의 설문조사는 가장 낮은 평균 점수를 받았는데, 이는 10 번의 수행 중 한 번의 수행이 낮은 점수인 0 점을 받았기 때문이다. 이는 화남 (Angry)과 신나는 (Excited) 감정사이의 경계가 협소하기 때문에 분석되었고 더 다양한 음악특징의 추가가 요구되었다.

5. 결론 및 향후 연구

본 논문에서는 사용자의 감정을 분석하여 음악특징을 추출하고 가까운 특징을 갖는 음악을 검색하는 시스템을 제안했다. 사용자 감정상태의 효율적인 입력을 돕기 위해 Valence-Arousal 2 차원 공간입력을 적용하였고, 사용자에게 적절한 음악특징을 얻기 위하여 퍼지 추론 시스템을 설계하였다. 또한 사용자에게 현재 감정상태에 대한 대표적 감정을 표현하기 위해 퍼지 연산을 수행하였다. 하지만 검색의 근거가 되는 음악특징의 부족함과 음악감정에 대한 퍼지 집합의 조정이 필요함이 실험적으로 나타났다. 향후에는 다양한 음악특징 분석연구와 음악 감정 모델에 대한 보완이 필요하다. 또한 본 시스템을 확장하여 감정의 반전을 고려한 음악 추천 시스템 설계의 관련 연구를 진행할 것이다.

참고문헌

- [1] Juslin, P.N., Sloboda, J.A., "Music and Emotion: Theory and research", New York: Oxford University Press, 2001.
- [2] Russell, J. A., "A Circumplex Model of Affect," Journal of Personality and Social Psychology, Vol. 39, 1980.
- [3] R. E. Thayer, "The Biopsychology of Mood and Arousal," New York: Oxford University Press, 1989.
- [4] AllMusic. <http://www.allmusic.com>
- [5] Musicoverly. <http://www.musicoverly.com>
- [6] Yang, Y.H., et al., "Music emotion classification: a fuzzy approach," Proc. of ACM Multimedia 2006, pp. 81-84, 2006.
- [7] Hung-Chen Chen and Arbee L. P. Chen, "A Music Recommendation System Based on Music and User Grouping,"
- [8] Lie Lu, D. Liu, Hong-Jiang Zhang, "Automatic Mood Detection and Tracking of Music Audio Signals," IEEE Transactions on Audio, Speech & Language Processing, Vol. 14, No. 1, pp. 5-18, 2006.
- [9] Yang, Y.H., et al., "A Regression Approach to Music Emotion Recognition," Audio, Speech, and Language Processing, IEEE Transactions, Feb. 2008