감성모델을 이용한 음악 탐색 인터페이스

Music Exploring Interface using Emotional Model

유민준, Min-Joon Yoo*, 김현주, Hyun-Ju Kim**, 이인권, In-Kwon Lee***

요약 본 논문에서는 감성 모델을 이용하여 음악을 정렬한 후, 이를 바탕으로 음악을 탐색하는 인터페이스를 제안한다. 먼저 다양한 곡들에 대한 Arousal-Valence 요소를 설문조사 한 후, 곡들의 다양한 audio feature 들과 Arousal-Valence 요소들간의 상관관계를 계산하여, AV모델을 수립한다. 그 후, 다양한 음악들을 수립된 AV모델에 대하여 정렬을 하여 음악을 배치한 후, 이를 마우스를 이용하여 탐색하는 인터페이스를 제공한다. 기존의 관련 인터페이스보다 더욱 직관적으로 원하는 곡을 선택할 수 있게 하기 위해서, 마우스의 위치에 따라서 여러 음악들이 페이드 인/아웃 되게 하였으며, 여러 가지 모드의 인터페이스를 제공하여, 사용자가 가장 편리한 인터페이스를 사용할 수 있게 하였다. 사용자는 본 음악 탐색 인터페이스를 이용하여, 더욱 감정적으로 원하는 음악을 쉽게 찾을 수 있게 된다.

Abstract In this paper, we introduce an interface for exploring music using emotional model. First, we survey arousal-valence factors of various music and calculate a correlation between audio fefatures of music and arousal-valence factors to build an AV model. Then, various music is aligned and arranged using the AV model and the user can explore music in this interface. To select the desired music more intuitively, we introduce new fade in/out function based on the location of the user's mouse point. We also offer several mode of selecting music so user can explore music using most suitable mode of interface. With our interface, the user can find the emotionally desired music more easily.

핵심어: music selection, emotional model, arousal-valence model, music exploring.

본 논문은 문화관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소(CT) 육성사업의 연구결과로 수행되었음.
*주저자 : 연세대학교 컴퓨터과학과 박사과정 e-mail: debussy@cs.yonsei.ac.kr
**공동저자 : 연세대학교 컴퓨터과학과 석사과정 e-mail: uyunamu@cs.yonsei.ac.kr
***교신저자 : 연세대학교 컴퓨터과학과 부교수; e-mail: iklee@yonsei.ac.kr

1. 서론

MP3 플레이어의 폭넓은 보급 및, 디지털 음악파일의 대중화로 인하여, 개인적으로 보유하고 있는 음악파일의 수는 나날이 크게 늘어나고 있다. 이를 효율적으로 관리하는 것은 점점 더 큰 문제가 되고 있으며, 기존의 사전식 정렬 수준을 벗어나 더욱 효과적으로 음악파일들을 관리하고자 하는 노력들이 이루어지고 있다. 본 논문에서는 음악에 대한 감성 모델을 이용하여 음악을 정렬하고, 이를 탐색하는 인터페이스를 제공한다.

최근에 음악의 메타 데이터 혹은 MFCC등의 오디오 시그널 데이터들을 이용하여 음악을 정렬하여 이를 바탕으로 음악을 탐색하고자 하는 연구들이 진행되고 있다[1]. 본 연구에서는 이전 연구들보다 더욱 높은 수준의 음악 탐색을 위하여, 음악에서 느껴지는 감정을 계산하는 모델을 수립하고, 이를 바탕으로 음악을 정렬하여, 음악을 탐색하는 방법을 제안한다.

우리는 먼저, 임의의 음악에 대하여 사람이 느끼는 감정을 추출하는 모델을 수립한다. 본 논문에서 사용하는 방법은 Thayer[2]의 감정 모델, 즉 Arousal-Valence 모델을 사용한다. 다양한 장르의 음악에 대하여 사람들의 설문조사를 통하여 각 곡에 대한 Arousal 값과 Valence 값을 추출한 후, 이를 각 곡에서 계산된 오디오 피쳐들과 연관지어, 음악에 대한 감성 모델을 수립한다. 이에 대한 관련 연구는 [3,4] 등에서 찾아볼 수 있다.

Thayer 모델에 따라 음악에서 느껴지는 감정을 분석하면, 임의의 음악을 Arousal과 Valence 두 축으로 나타낼 수 있다. 따라서 모든 음악을 2차원 상의 좌표로 나타낼 수 있으므로, 우리는 2차원 상에서 원하는 음악을 탐색할 수 있는 인터페이스를 제공한다. 적절한 페이드 인/아웃 기능을 추가하여, 음악 선택의 효율성을 증대시켰으며, 다양한 모드를 제공하여 사용자에가 가장 자신의 마음에 맞는 인터페이스를 선택하도록 한다.

본 논문에서 제시한 방법을 이용하면, 사용자가 느끼는 감정에 따라 음악을 고를 때, 더욱 빠르며, 효율적으로 선택할 수 있다. 이는 기존의 장르별 구분, 혹은 가수별 구분 보다 더욱 직관적으로 음악을 선택할 수 있는 방법으로 사용될 수 있다.

2. 음악에 대한 감정 모델 수립

2.1. Arousal-Valence 감정 모델

Thayer[2] 의 감정 모델에 따르면, 크게 감정은 Arousal와 Valence 의 두 가지 요소로 나누어질 수 있다. Arousal은 감정의 강도를 의미하며, 이 값이 커질 수 록 더욱 신나는 (exciting) 기분을 의미하며, 작을 수 록 더욱 조용한(calm) 한 기분을 의미한다. Valence는 감정의 긍정/부정의 정도를 나타내며, 이 값이 +로 더욱 커질 수 록 더욱 긍정적인 감정을 의미하게 된다 (그림 1).

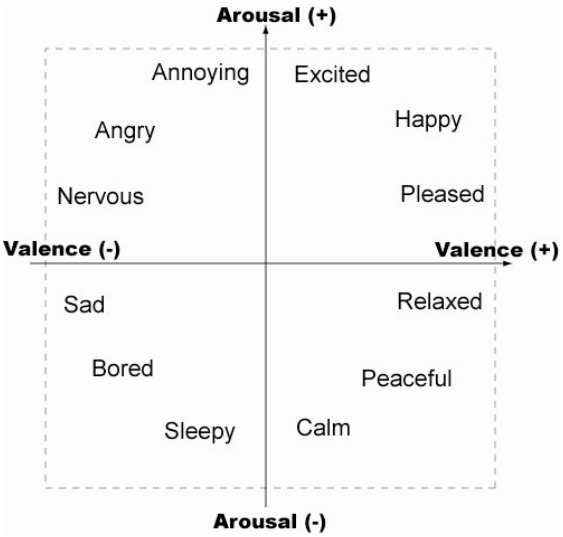

그림 1 Arousal-Valence 모델

2.2. 사용자 감정 조사

임의의 음악에 대한 감정의 정도를 측정하기 위한 모델을 수립하기 위하여, 우리는 먼저 총 135 곡의 음악에 대하여 설문조사를 통하여 각 곡에 대한 Arousal 및 Valence 수치를 조사하였다. 테스트 음악들은 한국, 일본 및 서양의 다양한 장르의 대중 음악에서 선택되었다.

실험방법은, Lang[5]의 방법을 약간 수정하여 수행되었다. 이 방법은 그림을 통하여 실험자가 자신의 감정상태를 더욱 직관적으로 표시할 수 있게 하는 방법으로, 이 방법에서 Arousal 및 Valence의 수치는 1~9 사이에서 선택되게 된다 (그림 2). 실험에 참가한 총 사용자는 모두 14명이다.

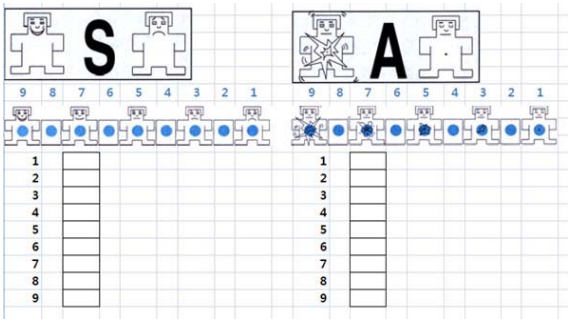

그림 2 본 논문에서 사용한 Arousal 및 Valence 조사 쉬트 (sheet)

2.3. 오디오 피쳐 추출

다음으로 테스트 음악에 대한 다양한 오디오 피쳐를 추출하였다. Psysound[6] 라이브러리를 이용하여, 총 55개의 오디오 피쳐들(FFT Spectrum, Cepstrum, Auto-Correlation, Loundess등)을 추출하였으며 이 중 선별작업을 통하여 의미가 적은 피쳐들이 제거되었다.

2.4. 음악에 대한 감정 모델

앞서 구한, 오디오 데이터에 대한 사용자의 Arousal-Valence값과 각 오디오의 피쳐들을 이용하여 음악에 대한 감정 모델을 수립할 수 있다. 본 논문에서는 두 가지 요소가 서로 선형적인 관계를 지니고 있다고 가정하여, 오디오데이터의 피쳐와 Arousal-Valence 값들간의 연관관계를 나타내는 선형 모델을 정의하였다. 이 때 오디오 데이터의 숫자와 사용된 오디오 피쳐의 수는 일치하지 않으므로, Pseudo inverse를 이용하게 된다. 여기서 구한 모델을 이용하여 임의의 음악이 들어오면, 이 음악의 오디오 피쳐들을 구한 후, Arousal-Valence 값들을 구할 수 있고, 이를 통하여 음악에서 느껴지는 감정을 계산할 수 있다.

3. 음악 탐색 인터페이스

음악에 대한 감정 모델은 Arousal과 Valance의 두 축으로 나타내어지므로, 임의의 곡들을 모두 2차원으로 표현할 수 있다. 즉, 기존의 연구에서 주로 사용된 SOM[7] 등을 이용한 차원축소 기법을 사용하지 않고 더욱 직관적으로 음악을 정렬 할 수 있다. 사용자는 2차원으로 정렬되어있는 음악에 접근하여 직접 원하는 음악을 선택할 수 있다.

3.1. 그레이 모델

여러 음악을 2차원 Arousal-Valence 평면의 점들로 표현하고, 각 점들마다 일정 높이를 갖는 원뿔의 부피를 갖는다고 가정하였다. 만일 여러 점들이 서로 근처에 있다면, 서로 원뿔의 부피는 중첩되도록 하였다. 이 방법을 이용하면, 비슷한 느낌의 음악의 수가 많은 부분과 적은 부분을 효과적으로 표현할 수 있는 인터페이스를 나타낼 수 있다. 그림 3(a)에 그레이 모델을 나타내었다.

3.2. 음악의 섬

Pampalk [8]이 제안한 섬모양의 인터페이스도 제공한다. 이 모델에서는 일정 값 이상의 데이터는 단계별로 색을 적용하여 섬과 유사한 형태의 색을 갖도록 한 인터페이스이다. 유사한 느낌의 음악이 적은 음악의 경우, 눈에 잘 안띈다는 단점이 있으나, 재미있는 의미와 모양을 갖는 인터페이스로 생각될 수 있다. 그림 3(b)에 음악의 섬 인터페이스를 나타내었다.

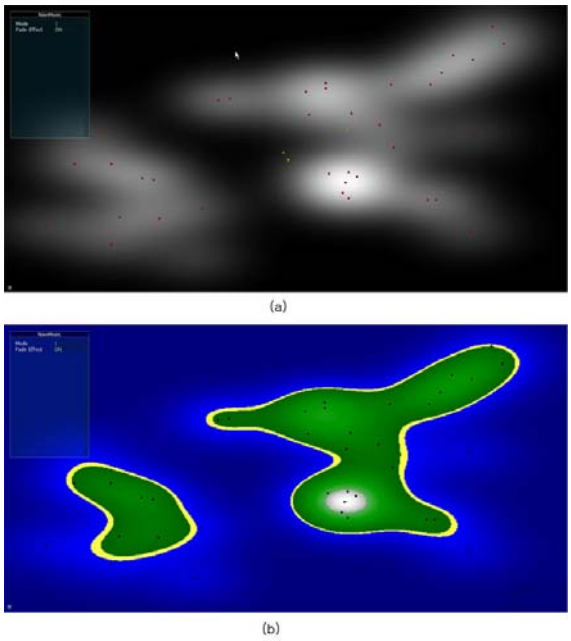

그림 3 (a) 그레이 모델, (b) 음악의 섬 모델

3.3. 멀티 음악 선택 기능

본 논문에서는 음악을 더욱 빠르게 선택하게 하기 위해서, 여러 음악을 동시에 들을 수 있는 기능을 제공한다. 사용자의 마우스 포인트의 위치에 따라서 사용자가 지점한 범위 안에 있는 음악이 모두 들리게 되며, 이때의 각 음악들의 볼륨은 마우스 포인트의 위치와 각 음악의 포인트의 위치를 고려하여 계산된다. i번째 음악의 볼륨을 V_i라고 한다면,

$$V_i = \alpha \exp(-\|P_{mouse} - P_i\|),$$

여기에서, P_{mouse}는 마우스의 위치, P_i는 i번째 음악의 위치. α는 사용자 정의 파라미터.

마우스의 위치 이동에 따라서 각 음악들의 볼륨은 자연스럽게 페이드 인/아웃 되는 효과가 발생하게 된다. 따라서 AV평면을 탐색하면서 여러 음악을 듣다가 더욱 마음에 드는 음악이 있을 때, 마우스를 그 음악 포인트로 가까이 움직이기만 해도 그 음악을 바로 들을 수 있게 된다. 따라서 원하는 음악을 선택할 수 있는 속도가 더욱 빨라지게 되며, 특히 모르는 음악 중에서도 자신에게 맞는 음악을 찾을 수 있는 가능성이 매우 커진다. 만일 이 기능이 제공되지 않을 때, 자신이 모르는 음악 중에 마음에 드는 음악을 선택하기 위해서는 모든 음악을 일일이 다 선택하여 들어봐야만 한다.

여러 음악의 볼륨에 따라 각 음악의 제목을 표기하는 텍스트의 투명도 또한 조정함으로써, 더욱 직관적으로 음악을 선택할 수 있게 하였다 (그림 4). 또한 만일 원하는 음악을 찾아서 그 음악만 듣고 싶을 경우에는, 그 음악의 포인트에 마우스를 정확히 위치하면 한 곡만 플레이되게 함으로써, 음악선택/다중음악선택 기능을 빠르게 바꿀 수 있게 하였다.

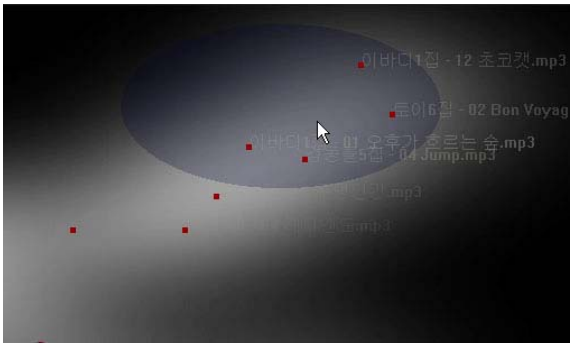

그림 4 멀티 선택 인터페이스. 보라색 타원 안에 들어오는 음악은 모두 재생이 되며, 볼륨은 타원의 바깥쪽으로 갈 수록 작아진다. 원 밖에 있는 음악의 제목도 희미하게 보이는 것도 확인할 수 있다.

3.4. 재생 음악 부분

각 음악을 재생할 때, 음악의 시작부터 재생하는 것보다는 후렴부분같은 더욱 반복되고 중요한 의미를 갖는 부분을 재생하는 것이 더욱 효과적임이 우리의 실험에서도 나타나었다. 특히 본 인터페이스처럼 순간적으로 각 음악들을 표현해야하는 시스템에서는 더욱 중요하다. 본 논문에서는 임의의 수작업을 통하여 이루어졌으나, 자동으로 음악을 분석하는 논문[9]등을 통하여 어느정도 자동화를 이루는 것도 가능하다.

4. 결론 및 향후 연구 계획

본 논문에서는 감정 모델에 따라 음악을 정렬하고 이를 탐색할 수 있는 인터페이스를 제안한다. 또한, 더욱 직관적인 음악 탐색을 위한 여러 가지 기능들을 소개한다. 본 인터페이스를 통하여, 사용자는 자신의 감정상태에 따라서 원하는 음악을 더욱 빠르고 정확하게 찾을 수 있다.

본 연구는 현재 터치 장치 및 모바일 장비를 이용한 인터페이스로 확장 중이며, 사용자 편의성을 더욱 높일 수 있는 여러 가지 기능을 고려중이다. 그레이 모델과 음악의 섬 모델 외에, 감정 모델을 더욱 직관적으로 적용시킬 수 있는 인터페이스 모델을 개발 중이다.

또한 더욱 높은 수준의 감정 모델을 수립하기 위한 연구가 진행 중이다. 음악의 감정을 추출하는 시스템 자체가 아직 완전한 단계가 아니기 때문에, 다른 분위기의 음악들이 근처에 있는 경우가 종종 있다. 더욱 엄밀한 오디오 피쳐의 선택과 계산을 통하여 본 시스템의 정확성을 더욱 높이는 것도 향후 개선 방향 중에 하나이다.

참고문헌

[1] P. Knees, M. Schedl, T. Pohle, G. Widmer, "Exploring Music Collections in Virtual Landscapes" IEEE multimedia Vol. 14, No. 3. pp. 46-54, 2007.

[2] R. E. Thayer, The Biopsychology of Mood and Arousal, New York, Oxford University Press, 1989.

[3] M. D. Korhonen, D. A. Clausi, M. E. Jernigan, "Modeling Emotional Content of Music Using System Identification," IEEE Transactions on Systems, Man, and Cybernetics - Part B : Cybernetics, Vol. 36, No.3, 2006.

[4] L. Lu, D. Liu, H-J, Zhang, "Automatic Mood Detection and Tracking of Music Audio Signals," IEEE Transactions on Audio, Speech, and Language Processing, Vol. 14 No.1, 2006.

[5] P. J. Lang, M. M, Bradley, B. N. Cuthbert. International Affective Picture System (IAPS) : Instruction Manual and Affective Ratings. NIMH Center for the Study of Emotion and Attention, 2007.

[6] Psysound. http://psysound.wikidot.com

[7] T. Kohonen. Self-Organizing Maps, Springer Series in Information Sciences, Springer, Berlin, 3rd Edition, Vol. 30. 2001.

[8] E. Pampalk, "Islands of Music: Analysis, Organization, and Visualization of Music Archives," master's thesis, Vienna Univ. of Technology, 2001.

[9] N. C. Maddage. Automatic Structure Detection for Popular Music. IEEE Multimedia 2006.