Proceedings of KFIS Fall Conference, 2003. 9. 27

개인 취향에 맞는 음악 장르 추천 시스템

The Adaptable Music Genre Recommendation System to The Individual Taste

강성춘, 이고은, 박정근, 손영선
동명정보대학교 정보통신공학과

Sung-Chun Kang, Go-Eun Lee, Jung-Kun Park and Young-Sun Sohn,
Dept. of Information & Communications Engineering
TONGMYONG University of Information Technology
E-mail : yssohn@tmic.tit.ac.kr

ABSTRACT

본 논문에서는 사용자가 음악을 직접 선곡하지 않고 락, 트로트, 댄스, 힙합, 발라드 등 5가지의 장르 중 사용자가 선호하는 음악의 장르를 추천하는 시스템을 구현하였다. 실시간으로 연주되는 음악에서 Bass Drum 신호를 추출 · 분석하여, 기본적으로 한 마디에 소요되는 시간, 주법, 진폭 등 세가지 파라메터를 이용하여 5가지 장르로 분류하였다 선택 곡 수와 들은 시간으로 퍼지 추론을 통해 각 장르에 대한 사용자 만족도를 평가한다. 평가된 만족도에 의해 사용자가 선호하는 장르의 음악을 제공하는 시스템을 제안한다.

Key words : 음악, 장르, 퍼지 추론

I. 서 론

음악 세계에는 다양한 종류의 장르가 존재한다. 장르는 음악가들의 새로운 시도에 의해 만들어 지기도 하고 상업적인 목적에 의해 만들어 지기도 한다. 이처럼 음악의 장르는 특정한 기준에 의해 만들어진 것이 아니기 때문에, 어떤 특정 기준으로 나누기가 힘들다.[1]

음악 장르 구분 연구로는, 시간 지연 신경 망을 이용한 분류 시스템이 있으나 마디 추출 방법에서 다양한 리듬과 빠르기의 음악을 수용할 수 없어 60%의 인식 율을 가진다.[2] 최근에는 에이젼트를 이용한 개인화 시스템 연구[4,5] 등 개인 취향을 고려하는 연구가 행하여지고 있다.

본 논문에서는 사용자가 듣는 음악의 특정 악기의 신호를 추출·분석하여 장르를 분류하고 퍼지 추론을 통한 장르별 사용자 만족도로 개인 취향에 맞는 장르의 음악을 제공하는 시스템을 구현 하였다.

II. 시스템 알고리즘

본 논문의 시스템 알고리즘은 그림 1과 같다. 사용자가 듣는 음악을 5초간 녹음하고 베이스 드럼의 신호를 추출·분석하여 주법, 한마디 시간, 진폭 등을 이용하여 기본적으로 5가지 장르로 구분한다. 구분된 각 장르별 선택 곡 수와 재생 시간 등의 사용자 데이터 통해 만족도를 추론 하여 높은 만족도를 나타내는 장르의 음악을 추천해준다.

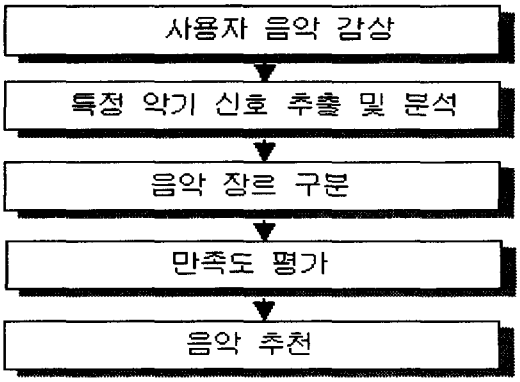


그림 1. 전체 시스템 구성도

III. 음악 장르 분류 알고리즘

3.1 음악 장르 분류

본 논문에서 분류하고자 하는 음악 장르는 Ballad, Dance, Hip-Hop, Rock, Trot등 5개의 장르이다.

5초간의 데이터를 녹음하고 녹음된 wave 데이터를 FFT(Fast Fourier Transform)를 취해 주파수 영역으로 변화시킨다. 주파수 영역의 신호를 BPF(Band-Pass-Filter)를 통해 20~80Hz의 신호만 통과 시키고 나머지 신호들은 차단시켜 다시 시간영역으로 바꾸기 위해 IFFT(Inverse Fast Fourier Transform)를 취한다. 추출 된 베이스드럼의 신호를 분석하여 주법과 마디시간 진폭 등을 기본 파라메타로 하여 장르 구분을 하였다. 음악 장르 분류 시스템 구성은 그림 2와 같다.

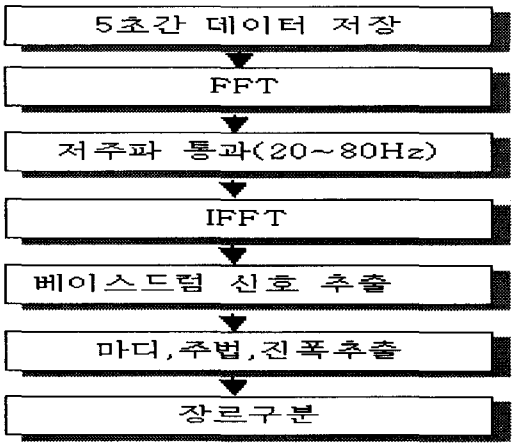


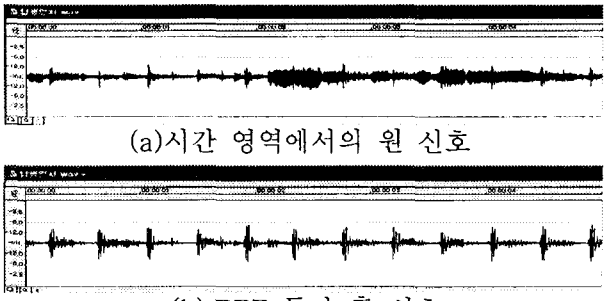
그림 2. 음악 장르 분류 시스템 구성도

3.2 베이스드럼 신호 추출

대중 음악은 멜로디(가락)를 담당하는 부분과 리듬을 담당하는 부분으로 구성되어 있으며 리듬을 구성하는 부분에서 규칙성을 찾을 수 있다. 리듬은 대부분 드럼과 베이스 기타 또는 이와 유사한 기능을 하는 악기를 사용하여 연주되는 특성을 가지고 있다. 본 논문에서는 규칙성을 가진 리듬 악기 중 리듬을 주도하면서 다른 악기와 주파수 영역이 겹치지 않는 베이스 드럼을 선정하였다.

일반적으로 곡들은 음악가 들이 의도한 곡의 분위기를 만들기 위하여 도입부분을 가지고 있다. 이러한 도입 부분들은 대부분 특별한 음악 장르의 특성을 갖지 않는다. 실험 결과 도입부분이 40초 미만이므로 데이터를 저장할 시간은 노래 재생후 40초 이후로 하였고 한 마디의 시간이 4.8초 이하이므로 불필요한 계산량을 줄이기 위해 5초간 녹음 하도록 하였다.

베이스 드럼의 신호를 찾기 위해 원 신호에 FFT를 취하여 주파수 영역의 신호로 변환하였다. 주파수 분포가 저주파(20~80Hz)에서 두드러지게 나타나므로 BPF를 사용하여 저주파만 통과 시킨 후, 주파수 영역의 신호를 시간영역으로 고치 기 위해 IFFT를 취하여 베이스 드럼의 신호를 추출해낼 수 있었다. 그림 3은 음악의 원 신호와 BPF를 통해 걸러진 베이스 드럼 신호를 나타내고 있다.



(a)시간 영역에서의 원 신호

(b) BPF 통과 후 신호

그림 3. 원 신호와 BPF를 통해 걸러진 베이스 드럼 신호

3.3 장르 구분 알고리즘

본 논문에서 제시한 5가지 장르 구분의 기준 파라메타로는 베이스 드럼의 주법, 주법을 통해 나온 한마디의 소요 시간, 베이스드럼 신호의 진폭 등 3가지 이다.

베이스드럼의 기본 주법으로는 그림 4에 보여지듯이 9가지가 있고, 음악장르에 따라 베이스 드럼의 주법이 다음과 같이 다르다. 그림 4의 (1),(2),(3)은 Ballad, Trot주법, (4),(5),(6)은 Dance, Hip-Hop, Trot주법, (7),(8)은 Dance, Rock 주법, (9)은 Hip-Hop주법이므로, 이와 같은 주법을 이용하여 1차적으로 장르를 분류한다.

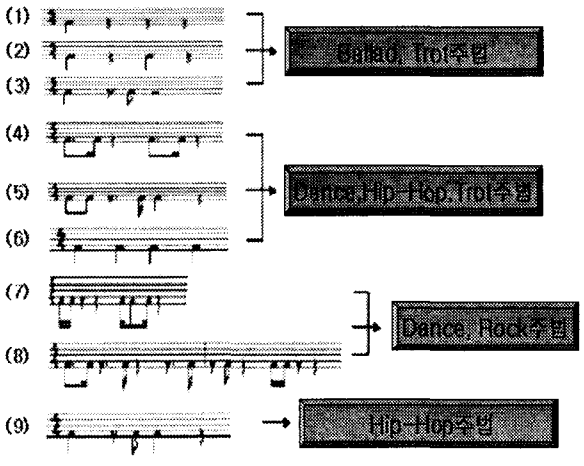


그림 4. 장르를 나누는 9가지 주법

음악은 장르에 따라 한 마디에 소요되는 시간이 다르다. Dance는 1.2~2초, Hip-Hop은 1.8~2.8초, Rock과 Ballad는 2.8~4.8초, Trot는 1.2~4.2초 등 차이가 나므로 이것을 이용

하여 2차적으로 장르를 분류한다. Trot는 시간과 주법이 Ballad와 중복되는 느린 Trot와 Dance와 중복되는 빠른 Trot가 있기 때문에 2차적인 장르분류까지도 구분되어 지지 않는다. 따라서, 음악의 특성들로 다시 나누면 Trot는 반주보다 사람의 음성을 위주로 된 음악, Dance는 음성보다 반주의 효과가 큰 음악이므로 진폭을 통해 진폭 값이 크면 Dance, 작으면 Trot로 구분하여 3차적으로 장르를 분류한다. 또한 느린 Trot와 Ballad는 모두 반주보다는 사람의 음성을 위주로 하는 음악이므로 3차적인 장르 분류 후에도 분류가 되지 않는다. 따라서, 반주에 들어가는 악기들로 구분을 두는데 Trot는 21~48인조의 오케스트라로 구성되고 Ballad는 6~7인조의 반주로 구성된다. 이러한 특징들 때문에 필터를 거친 신호의 peak치 개수 와 peak치들의 합한 값을 가지고 제4차적인 장르 구분을 통해 최종 5가지 장르로 분류하였다.

3.4 장르 구분 결과

본 시스템을 통해 장르별 200곡씩을 무작위로 선택하여 테스트한 결과 표 1과 같다.

표1. 장르 구분 결과

장르 \ 장르구분	정인식율	오인식율
Rock	77%	23%
Trot	68.5%	31.5%
Dance	71.5%	28.5%
Ballad	89%	11%
Hip-Hap	75%	25%

장르 구분 요소에는 가사가 담고 있는 의미나 시대의 문화적인 요인, 음악가의 음악적 특성 등이 있다. 이런 요소들을 다 반영하지 못한 결과로써 장르 오인식이 발생한다. 그러나 다양한 리듬과 빠르기를 수용함으로써 인식 율이 개선되는 결과를 얻을 수 있었다.

IV. 사용자 평가 알고리즘

4.1 만족도 추론

사용자의 주관적인 만족도를 얻기 위해 선택 곡 수, 청취 시간 등 2가지 요소를 이용하여 그림 5과 같이 퍼지 추론한다. 최종 추론 된 결과는 평가 요소에 대한 만족도를 나타낸다.

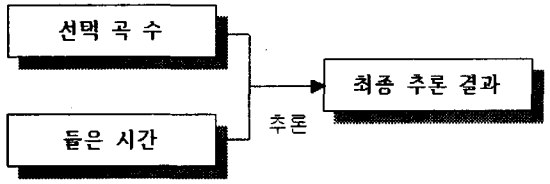
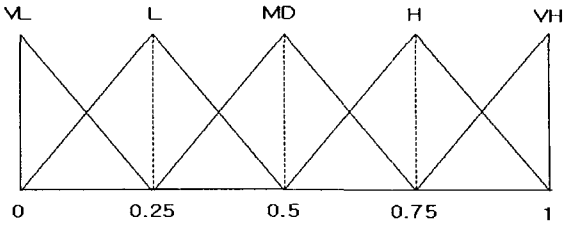


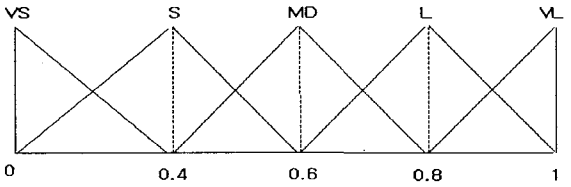
그림 5. 만족도 추론의 블록도

그림 6은 선택 곡 수에 대한 멤버십 함수, 그림 7은 들은 시간에 대한 멤버십 함수 이다.



VL(Very Low) : 선택 곡 수의 빈도가 매우 낮다.
L(Low) : 선택한 곡 수의 빈도가 낮다.
MD(MiDdle) : 선택한 곡 수의 빈도가 중간이다.
H(High) : 선택한 곡 수의 빈도가 높다.
VH(Very High) : 선택한 곡 수의 빈도가 매우 높다.

그림 .6 선택 곡 수에 대한 멤버십



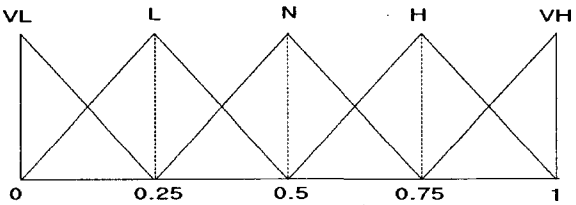
VS(Very Short) : 음악을 들은 시간이 매우 짧다.
S(Short) : 음악을 들은 시간이 짧다.
MD(MiDdle) : 음악을 들은 시간이 보통이다.
L(Long) : 음악을 들은 시간이 길다.
VL(Very Long) : 음악을 들은 시간이 매우 길다.

그림 7. 음악 들은 시간에 대한 멤버십 함수

선택 곡 수, 들은 시간에 의한 만족도를 표 3과 같이 구성하였고, 그림 8의 멤버십 함수로 나타내어진다.

표 3 퍼지 추론 규칙

들은 시간 \ 선택 곡 수	VL	L	MD	H	VH
VS	VL	VL	L	L	N
S	VL	L	L	N	H
MD	L	L	N	N	H
L	N	N	H	VH	VH
VL	H	H	H	VH	VH



VL(Very Low) : 만족도가 아주 낮다.
L(Low) : 만족도가 낮다.
N(Normal) : 만족도가 보통이다.
H(High) : 만족도가 높다.
VH(Very High) : 만족도가 매우 높다.

그림 8. 실제적인 만족도 멤버십 함수

퍼지 추론 결과를 바탕으로 각 장르별 사용자 만족도를 비교하여 사용자가 제일 크게 만족하는 장르의 음악을 추천 한다.

V. 실험 및 결과

5.1 실험

본 실험은 동명정보대학교 남,여 대학생 8명을 대상으로 하였고, 피험자에게 음악을 감상할 수 있는 인터페이스를 그림 9와 같이 제공하였다. 인터페이스를 사용한 피험자는 실험이 끝남과 동시에 추천된 곡에 대한 사용자 만족도 설문 조사를 행하였다.

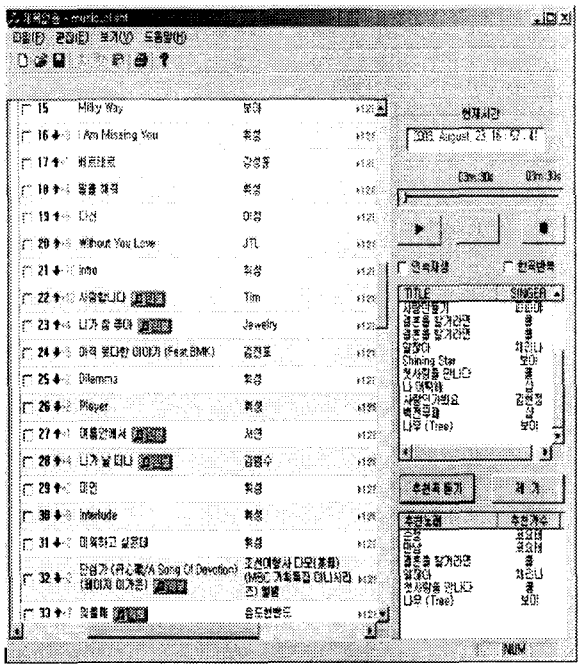


그림 9. 사용자 인터페이스 화면

5.2 결과

본 논문에서 제안한 방법에 의해 추론 된 만족도와 실제 설문으로 얻어진 만족도를 표3에 보여지고, 비교 결과가 유사함을 알 수가 있었다.

표 3. 추론 된 만족도와 실제 설문 비교

ID	추천 장르	만족도 추론값	설문 결과
park5697	댄스	0.85	매우 만족
tlstjsgp1	댄스	0.7	만족
radpooh	발라드	0.75	만족
toy_eterity	발라드	0.51	만족
sooki75	락	0.74	만족
uni350	발라드	0.7	보통
hjp2000	락	0.51	불만족
yasi0929	발라드	0.71	보통

IV. 결론 및 향후 과제

본 논문에서는 특정악기 신호를 추출·분석하여 기본적으로 한마디에 소요되는 시간, 주법, 진폭 등 세가지 파라메터를 이용하여 5가지 장르로 분류하였다. 기존의 장르 분류 시스템 연구에서는 다양한 리듬과 빠르기를 수용할 수 없어 60%의 인식률을 보였으나 본 시스템은 이를 수용함으로써 평균 76.5%의 상당히 개선된 결과를 얻을 수 있었다. 구분된 각 장르별 선택 곡 수와 재생 시간 등의 사용자 데이터로 만족도를 추론 하여 높은 만족도를 나타내는 장르의 음악을 추천 해주었다.

향후 과제로는 다른 많은 음악의 장르를 구별하기 위해 가사가 담고있는 의미나 시대의 문화적 요인, 음악가의 음악적 특성 등을 반영할 수 있는 효과적인 특징 추출에 관한 연구와 카메라나 마이크를 통한 사용자의 표정이나, 음성을 분석 사용자 감정을 인식하여 보다 정확한 사용자 만족도를 반영하여 더 정확한 취향 분석을 통한 음악 제공을 고려할 수 있다.

참고 문헌

[1] 신현준, 록 음악의 아홉 가지 갈래들, 문학과지성사, ISBN 89-320-0969-4, 1997.

[2] 이재원. 조찬윤. 김상균."시간 지연 신경망 을 이용한 음악 장르 분류". 멀티미디어 학회 논문지, Vol.4, No.5, 2001

[3] 손영선, "Evaluation of Human Interface using Fuzzy Measures and Fuzzy Integrals", 한국 퍼지 및 지능시스템학회, Vol.8, No.2, PP.31-36, 1998.

[4] 고경희, 오경환, "사용자의 동적인 관심변화를 학습하는 개인화된 뉴스 에이전트", 한국정보과학회, Vol. 28, No.2, PP.82~84, 2001.

[5] 한선미, 우진운, "지능형 에이젼트를 이용한 개인화된 유무선 뉴스 검색 시스템", 정보처리학회논문지B, Vol.8-B, No.6, PP.609-616, 2001.

[6]한학용."Multimedia Sound Programming".영진닷컴. 2003

[7]本多中二,大理有生, "퍼지 공학 입문",응보출판사, 2002.

[8]최영삼,손영선. "독자에게 적응하는 신문기사 추천 시스템". 한국퍼지 및 지능 시스템학회. 추계학술대회 학술 발표 논문집. pp.431-434. 2002.

[9]조용성. 강은영. 손영선. "개인 취향에 맞는 음악 선곡 시스템". 한국퍼지 및 지능 시스템학회. 추계학술대회 학술 발표 논문집. pp.435-438. 2002.