



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0012221
 (43) 공개일자 2014년02월03일

(51) 국제특허분류(Int. Cl.)
G10L 19/00 (2006.01) **G10L 25/00** (2014.01)
 (21) 출원번호 10-2012-0072577
 (22) 출원일자 2012년07월04일
 심사청구일자 2012년07월04일

(71) 출원인
금오공과대학교 산학협력단
 경상북도 구미시 대학로 61 (양호동, 금오공과대학교)
 (72) 발명자
김병만
 경상북도 구미시 대학로 61 (양호동 디지털관 335호)
문창배
 경상북도 구미시 대학로 61 (양호동 디지털관 336호)
 (뒷면에 계속)
 (74) 대리인
김순용

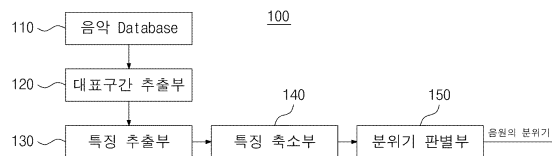
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 **표준편차 기반의 음원의 특징 축소 방법을 이용하는 음악 분위기 자동 판별 시스템 및 방법**

(57) 요약

본 발명은 표준편차를 기반으로 하는 음원의 특징을 축소하는 방법을 이용한 음악 분위기 자동 판별 시스템 및 방법에 관한 것이다. 본 발명에 따른 음악 분위기 자동 판별 시스템은 많은 음원의 특징들 중 일부를 선택하여 특징을 축소함으로써, 특징들 사이의 잡음을 제거하여, 보다 효과적으로 많은 음악파일들의 분위기를 정확하게 판별할 수 있다. 따라서, 본 발명에 따른 음악 분위기 자동 판별 시스템은 분위기 자동 판정을 통한 내용 기반 필터링, 추천 시스템 및 정보 검색 분야 등 많은 분야에서 유용하게 사용될 수 있다.

대표도



(72) 발명자
김현수
경상북도 구미시 대학로 61 (양호동 디지털관 33
6호)

송민균
경상북도 구미시 대학로 61 (양호동 디지털관 33
6호)

이 발명을 지원한 국가연구개발사업
과제고유번호 2010-0021097
부처명 교육과학기술부
연구사업명 일반연구자지원사업
연구과제명 다국어 태그를 지원하는 음악 분위기 폭소노미 기반의 자동태깅 시스템
기여율 1/1
주관기관 금오공과대학교 산학협력단
연구기간 2010.09.01 ~ 2013.08.30

특허청구의 범위

청구항 1

특정값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템에 있어서,

음악의 대표 구간을 추출하는 대표 구간 추출부;

상기 대표 구간 추출부에 의하여 추출된 대표 구간의 특징을 추출하는 특징 추출부;

상기 특징 추출부에 의하여 추출된 특징을 축소하는 특징 축소부; 및

상기 특징 축소부에서 축소된 특징을 가지고 음악의 분위기를 판별하는 음악 분위기 판별부;를 포함하되,

상기 특징 축소부는 음악의 특징을 정규화하고, 각 특징별로 기준을 정의한 후, 각 기준의 표준편차를 구하고, 임계치를 적용하여 특징을 선별하는 알고리즘을 이용하는 것을 특징으로 하는, 특정값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템.

청구항 2

제 1항에 있어서, 상기 대표 구간은 음악의 구조 정보를 이용하여 음악의 유사 구간을 획득한 후, 상기 유사 구간을 세그먼트화하여 음원의 단위로 나눈 것들 중에서 선택된 음원임을 특징으로 하는, 특정값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템.

청구항 3

제 2항에 있어서, 상기 선택된 음원은 대표 구간 추출부의 음악 세그먼트화를 통하여 유사 구간이 세그먼트화되어 음원의 단위로 나뉜 음원 중 도입부, 종결부 및 세그먼트화된 음원들 중 에너지가 가장 높은 음원으로 이루어진 것을 특징으로 하는, 특정값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템.

청구항 4

제 1항에 있어서, 상기 대표 구간 추출부는, 상태열 기반의 유사 구간 클러스터링을 이용하여 음악의 구조를 분석하여 음악의 대표 구간을 추출하는 것을 특징으로 하는, 특정값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템.

청구항 5

특정값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법에 있어서,

1) 대표 구간 추출부를 통하여 음악의 대표 구간을 추출하는 단계;

2) 특징 추출부를 통하여 상기 1)단계에서 추출된 대표 구간의 특징을 추출하는 단계;

3) 특징 축소부를 통하여 상기 2)단계에서 추출된 특징들 중 일부를 선택하여 특징의 갯수를 축소하는 단계; 및

4) 분위기 판별부를 통하여 상기 3)단계에서 축소된 특징을 이용하여 음악의 분위기를 판별하는 단계를 포함하되, 상기 3)단계의 특징의 갯수를 축소하기 위하여, 음원의 특징을 정규화하고, 각 특징별로 기준을 정의한 후, 각 기준의 표준편차를 구하고, 임계치를 적용하여 특징의 갯수를 축소하는 것을 특징으로 하는, 특정값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법.

청구항 6

제 5항에 있어서, 상기 1)단계의 대표 구간은 음악의 구조 정보를 이용하여 음악의 유사 구간을 획득한 후, 상기 유사 구간을 세그먼트화하여 음원의 단위로 나눈 것들 중에서 선택된 음원임을 특징으로 하는, 특정값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법.

청구항 7

제 6항에 있어서, 상기 선택된 음원은 대표 구간 추출부의 음악 세그먼트화를 통하여 유사 구간이 세그먼트화되

어 음원의 단위로 나뉜 음원 중 도입부, 종결부 및 세그먼트화된 음원들 중 에너지가 가장 높은 음원으로 이루어진 것을 특징으로 하는, 특징값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법.

청구항 8

제 5항에 있어서, 상기 2)단계의 음악의 대표 구간을 추출하기 위하여, 상태열 기반의 유사 구간 클러스터링을 이용하여 음악의 구조를 분석하여 음악의 대표 구간을 추출하는 것을 특징으로 하는 특징값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법.

명세서

기술분야

[0001] 본 발명은 표준편차를 기반으로 하는 음원의 특징을 축소하는 방법을 이용한 음악 분위기 자동 판별 시스템 및 방법에 관한 것이다.

배경기술

[0002] 분위기 추출에 관한 초창기의 연구들은 일반적인 기계 학습/판별 방법을 사용하였으나, 이러한 방법은 음악을 하나의 분위기로 판단하기 때문에 정확성이 떨어지는 문제가 있고, 또한 개인의 주관적인 느낌과 이질감을 반영하지 못하는 문제가 있었다(T. Li and M. Ogiwara, "Detection of Emotion in Music," Proc. of the International Symposium on Music Information Retrieval (ISMIR), pp.239-240, Washington D.C., USA, 2003; L. Lu, D. Liu and H. Zhang, "Automatic Mood Detection and Tracking of Music Audio Signals," IEEE Trans. on Audio, Speech, and Language Processing, Vol. 14, pp5-18, 2006; Y.Feng, Y.Zhang and Y.Pan, "Popular Music Retrieval by Detecting Mood," Proc. of ACM SIGIR 2003, pp 375-376, 2003). 단일 분위기로 판정하기 때문에 발생하는 불확실성을 해결하기 위하여, 퍼지 기반의 학습/분류 방법을 사용하였으나, 이 또한 개인에게 느껴지는 음악의 분위기에 대한 주관적 성향을 해결하기에는 한계가 있었다(Y.H. Yang, C.C. Liu and H.H. Chen, "Music Emotion Classification: a Fuzzy Approach," Proc. of ACM Multimedia 2006(ACM MM'06), pp.81-84, Santa Barbara, CA, USA, 2006).

[0003] 또한, 하나의 음악은 전체 내용이 동일한 분위기 특성을 유지하기보다는 중간 중간에 다른 분위기로 변화한다. 따라서 음악의 분위기를 탐지하기 위해서는 전체 음악을 의미있는 몇 개의 부분으로 나누고 각 부분들에 대하여 독립적인 분위기를 탐지하는 기법이 필요하다. 하지만, 기존 연구들에서는 이러한 특성을 고려하지 않고 각 연구의 필요성에 따라 음악의 일부분을 전문가의 수작업을 통해 잘라내어 사용하거나 임의로 설정된 구간(예를 들어 음악의 시작 후 30초 구간부터 30초 길이의 구간)을 사용하였다. 이러한 방법들은 새로이 출판되는 음악에 적용시키기에 무리가 있으며 변화가 많은 음악의 특성상 정확도가 떨어지는 단점이 있다.

[0004] 이러한 문제점을 해결하기 위하여, 기존 연구에서는 수동이 아닌 자동으로 음악 자체의 내용을 바탕으로 한 구조 분석기법을 통하여 음악을 의미 있는 구간들로 나누고, 각 구간들의 독립적인 분위기를 탐지하는 방법을 제안하였다(Jong In Lee, Dong-Gyu Yeo, Byeong Man kim, Hae-Yeoun Lee, "Automatic Music Mood Detection through Musical Structure Analysis", International Conference on Computer Science and its Application CSA 2009, pp. 510-515, 2009, 문창배, 김현수, 송민균, 김병만, "AV 모델을 이용한 음악 분위기 자동 분류", 한국정보처리학회, 제35회 한국정보처리학회 춘계학술발표대회 논문집 제18권 제1호, 2011). 하지만 기존문헌 (Jong In Lee, Dong-Gyu Yeo, Byeong Man kim, Hae-Yeoun Lee, "Automatic Music Mood Detection through Musical Structure Analysis", International Conference on Computer Science and its Application CSA 2009, pp. 510-515, 2009)은 피실험자의 수가 적어 대중적인 성향보다는 개인적 성향에 대한 비중이 높고, 다른 기존문헌(문창배, 김현수, 송민균, 김병만, "AV 모델을 이용한 음악 분위기 자동 분류", 한국정보처리학회, 제 35회 한국정보처리학회 춘계학술발표대회 논문집 제18권 제1호, 2011)에서는 실험자의 수를 증가시켜 개인적 성향의 문제를 해결하였다. 또한 상기 기존문헌에서는 MIRtoolbox(Olivier Lartillot, "MIRtoolbox 1.2.4", Finnish Centre of Excellence in Interdisciplinary Music Research, March, 18th, 2010)를 사용하여 391개의 특징을 추출하고, 특징 축소 방법으로 "김종완, 김희재, 김병만, "퍼지추론과 신경망을 사용한 유즈넷 뉴스 그룹 결정", 한국퍼지 및 지능시스템학회 2004년도 춘계학술대회 학술발표논문집 제14권 제1호, 2004"의 방법을

적용하여 50개의 특징을 사용하였지만 분위기 판별에서는 정확도가 떨어지는 단점이 있었다.

발명의 내용

해결하려는 과제

[0005] 상기와 같은 문제점을 해결하기 위하여, 본 발명의 목적은 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템에 있어서, 특징 축소를 위하여, 음악의 특징을 정규화하고, 각 특징별로 기준을 정의한 후, 각 기준의 표준편차를 구하고, 임계치를 적용하여 특징을 선별하는 알고리즘을 이용하는 것을 특징으로 하는, 특징값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템 및 방법을 제공함에 있다.

과제의 해결 수단

- [0006] 상기와 같은 목적을 달성하기 위하여, 본 발명은 특징값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템에 있어서,
- [0007] 음악의 대표 구간을 추출하는 대표 구간 추출부;
- [0008] 상기 대표 구간 추출부에 의하여 추출된 대표 구간의 특징을 추출하는 특징 추출부;
- [0009] 상기 특징 추출부에 의하여 추출된 특징을 축소하는 특징 축소부; 및
- [0010] 상기 특징 축소부에서 축소된 특징을 가지고 음악의 분위기를 판별하는 음악 분위기 판별부;를 포함하는 특징값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템을 제공한다.
- [0011] 또한, 본 발명은 특징값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법에 있어서,
- [0012] 1)대표 구간 추출부를 통하여 음악의 대표 구간을 추출하는 단계;
- [0013] 2)특징 추출부를 통하여 상기 1)단계에서 추출된 대표 구간의 특징을 추출하는 단계;
- [0014] 3)특징 축소부를 통하여 상기 2)단계에서 추출된 특징들 중 일부를 선택하여 특징의 갯수를 축소하는 단계; 및
- [0015] 4)분위기 판별부를 통하여 상기 3)단계에서 축소된 특징을 이용하여 음악의 분위기를 판별하는 단계를 포함하는, 특징값의 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법을 제공한다.

발명의 효과

[0016] 본 발명은 표준 편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 시스템은 많은 음원의 특징들 중 일부를 선택하여 특징을 축소함으로써, 특징들 사이의 잡음을 제거하여, 보다 효과적으로 많은 음악파일들의 분위기를 정확하게 판별할 수 있다. 따라서, 본 발명에 따른 음악 분위기 자동 판별 시스템은 분위기 자동 판정을 통한 내용 기반 필터링, 추천 시스템 및 정보 검색 분야 등 많은 분야에서 유용하게 사용될 수 있다.

도면의 간단한 설명

- [0017] 도 1은 본 발명에 따른 음악 분위기 자동 판별 시스템의 구성을 대략적으로 나타낸 도이다.
- 도 2는 본 발명의 자동 판별 시스템의 대표 구간 추출부의 구성을 대략적으로 나타낸 도이다.
- 도 3은 본 발명에 따른 특징값의 정규화 결과를 나타낸 도이다.
- 도 4는 본 발명에 따른 각 특징에 속한 각 클래스의 평균값을 계산을 통하여 기준을 구한 결과를 나타낸 도이다.
- 도 5는 본 발명에 따른 특징 선택 알고리즘을 나타낸 도이다.
- 도 6은 본 발명에 따른 신경망 구조를 나타낸 도이다.
- 도 7은 본 발명에 따른 분위기 수집 표를 나타낸 도이다.

도 8은 본 발명의 판별 성능 결과를 나타낸 도이다.

도 9는 본 발명의 음악 분위기 자동 판별 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 이하, 본 발명의 바람직한 실시예를 첨부된 도면을 참조하여 상세히 설명하면 다음과 같다. 그러나 이들 도면은 예시적인 목적일 뿐, 본 발명이 이에 한정되는 것은 아니다.
- [0019] 도 1은 본 발명에 따른 음악 분위기 자동 판별 시스템의 구성을 대략적으로 나타낸 도이다.
- [0020] 도 1을 참조하면, 본 발명의 음악 분위기 자동 판별 시스템은 대표 구간 추출부(120), 특징 추출부(130), 특징 축소부(140) 및 분위기 판별부(150)를 포함한다. 바람직하게는 음악 데이터베이스(110)를 추가적으로 포함할 수 있다.
- [0021] 대표 구간 추출부(120)는 음악 데이터베이스(110)로부터 전달받은 음악의 분위기 판별을 위한 대표 구간을 추출하는 기능을 수행한다. 보다 구체적으로, 본 발명의 대표 구간 추출부(120)는 음악의 구조 정보를 분석하고, 분석된 음악의 구조 정보를 이용하여 음악 세그먼트화하여 음원의 단위로 나누고, 상기 음원들 중 대표 구간을 추출하는 기능을 수행한다.
- [0022] 상기 음악의 구조 정보를 분석하는 방법은 음색의 변화 구간 탐색 방법, 음표 리듬 패턴을 통한 구조 분석 방법 및 특징 벡터간의 유사도 매트릭스의 클러스터링을 이용할 수 있으나, 상태열 기반의 유사 구간 클러스터링을 이용하는 것이 바람직하다.
- [0023] 도 2는 본 발명의 자동 판별 시스템의 대표 구간 추출부의 구성을 계략적으로 나타낸 도이다.
- [0024] 도 2를 참조하면, 상태열 기반의 유사 구간 클러스터링을 수행하기 위하여, 대표 구간 추출부(120)는 음악 특징 추출부(121), 음색-타입 시퀀스 추출부(122), soft k-mean 클러스터링부(123), 음악구조정보 생성부(124), 음악 세그먼트화부(125) 및 대표 음원 선택부(126)를 포함할 수 있다.
- [0025] 음악 특징 추출부(121)는 음악 신호의 1차 음향 특징을 추출하고, 음색 타입 시퀀스 추출부(122)는 상기 1차 음향 특징에서 음색-타입(timbre-type)시퀀스를 추출한다. 소프트 K-민즈(soft k-means) 클러스터링부(123)는 추출된 음색-타입 시퀀스를 이용하여 클러스터 정보를 획득하고 음악 구조 정보 생성부(124)는 클러스터 정보를 분석하여 음악 구조 정보를 생성시킬 수 있다.
- [0026] 음악 특징 추출부(121)를 통하여 상기 1차 음향 특징을 추출하기 위한 특징 추출 프레임 길이 결정 방법으로 비트 탐색 알고리즘을 사용할 수 있고, 상기 비트 탐색 알고리즘을 통하여 8개의 비트에 해당하는 길이를 프레임 윈도우의 홉사이즈로 사용할 수 있으나, 1.2S의 길이와 300ms의 홉사이즈를 가진 고정된 프레임을 사용하는 것이 바람직하다. 상기 1차 음향 특징 추출은 각 프레임 별로 MPEG-7의 1/8 옥타브의 해상도를 갖는 ASE(Audio Spectrum Envelop)를 추출한 후, PCA(Principal component Analysis)알고리즘을 통하여 상위 20 프로젝트션(ASP:Audio Spectrum Projection)을 추출할 수 있으나, PCA 알고리즘에 의해 정규화된 ASP 값을 사용할 경우 각 프레임의 에너지 차이에 대한 정보가 사라지기 때문에 각 프레임별 파워 스펙트럼 값의 L2-Norm을 구하여 총 21차의 멜로디와 에너지 모두를 표현하는 음향 특징 벡터를 추출하는 것이 바람직하다.
- [0027] 상기 추출된 음향 특징 벡터를 음색-타입 시퀀스로 맵핑하기 위하여, 음색-타입 시퀀스 추출부(122)는 HMM(hidden markov model)을 이용할 수 있다. HMM을 적용하기 위하여, 초기 HMM 파라미터를 생성하는 k-means 클러스터링부(122-1)와 HMM 파라미터를 생성하는 바움-웰치 파라미터 추정부(122-2)를 포함할 수 있다. 상기 바움-웰치 파라미터 추정부(122-2)를 통하여 HMM 파라미터를 얻은 후 비터비 경로 추정부(122-3)을 통하여 음색-타입 시퀀스를 추출할 수 있다.
- [0028] 소프트 k-민즈(soft k-mean) 클러스터링부(123)는 히스토그램 기반 소프트 k-민즈(soft k-means) 클러스터링 방법을 이용하여 클러스터 정보를 생성할 수 있다. 즉, M개의 세그먼트로 나누기 위하여 우선 가상코드를 W의 크기를 가지는 윈도우를 한 스텝씩 이동시키며 윈도우 내의 시퀀스를 사용한 데이터 히스토그램을 생성하고, 각 데이터 히스토그램이 속해있는 특징적 구간의 라벨을 할당하여 특징적 구간의 시퀀스를 생성할 수 있다.
- [0029] 음악 구조 정보 생성부(124)는 상기 생성된 클러스터 정보를 이용하여 음악의 구조 정보를 생성한다.
- [0030] 음악 세그먼트화부(125)를 통하여, 상기 생성된 음악의 구조 정보를 이용하여 음악의 유사 구간을 획득한 후, 세그먼트화하여 음원의 단위로 나눈다. 즉, 유사구간을 시작부부터 12초 단위로 음악을 분리시킬 수 있다.

[0031] 대표 음원 선택부(126)를 통하여, 상기 분리된 12초 단위의 음원 중, 음악의 도입부에서 1개와 종결부에서 1개를 선택하고, 음원들의 에너지를 계산하여 에너지가 가장 큰 샘플을 1개 선택할 수 있다(상기 선택된 샘플을 대표 음원이라 함). 즉, 음악당 최대 3개의 대표 음원이 선택되지만 도입부나 종결부에 에너지가 최대인 경우 음악당 2개의 대표 음원이 선택될 수 있다(상기 대표 음원을 이하 대표 구간이라고 함).

[0032] 특징 추출부(130)는 상기 대표 구간 추출부(120)에서 추출된 대표 구간(대표 음원)의 특징을 MIRtoolbox를 이용하여 추출할 수 있다. 상기 특징의 대분류로 다이내믹스(dynamics), 유동(Fluctuation), 리듬(Rhythm), 스펙트랄(Spectral), 음색(Timbre), 톤(Tonal) 이고, 중분류로 MFCC, 템포(Tempo), 크로마그램(Chromagram), 롤오프(Rolloff)등을 포함한 28개 일 수 있고, 상기 28개의 특징들 각각에 대해 민즈(Mean), Std, 슬롭(Slope), PeriodFreq, PeriodAmp, PeriodEntropy등의 특징들을 갖을 수 있고, 바람직하게는 391의 특징을 갖을 수 있다.

[0033] 도 3은 본 발명에 따른 특징값의 정규화 결과를 나타낸 도이고, 도 4는 본 발명에 따른 각 특징에 속한 각 클래스의 평균값을 계산을 통하여 기준을 구한 결과를 나타낸 도이고, 도 5는 본 발명에 따른 특징 선택 알고리즘을 나타낸 도이다.

[0034] 특징 축소부(140)은 상기 추출된 특징들 중 분별력이 좋은 특징을 선별하여 특징을 축소하는 기능을 한다. 본 발명에서 특징 축소부(140)은 표준편차를 이용한 특징 축소 알고리즘을 이용한다. 보다 구체적으로, 표준편차를 이용한 특징축소 알고리즘은 총 세 단계로 진행된다. 첫 번째 단계는 도 3에 나타난 바와 같이 정규화 단계이다.

[0035] 두 번째 단계는 도 4에 나타난 바와 같이, 각 특징별로 기준을 정의하는 단계이다.

[0036] 마지막 단계는 도 5에 나타난 바와 같이, 각 기준의 표준편차를 구하고, 임계치를 적용하는 단계이다. 정규화 과정은 각 특징들은 서로 다른 최대값을 가지기 때문에 도 3의 점선 사각형과 같이, 특징들을 0~1의 사이 값으로 변환하는 과정이다. 정규화 과정은 최초 각 특징들에 대하여 최대값을 구하고, 각 특징에 대하여 최대값으로 나누는 방법으로 하기의 수학적 식 1을 이용하여 정규화한다.

[0037] [수학적 식 1]

[0038]
$$Nf_{c,n}^k = (fd_{c,n}^k - Y_{c,n}) / (X_{c,n} - Y_{c,n})$$

[0039]
$$X_{c,n} = \max(fd_{c,n}^1, fd_{c,n}^2, \dots, fd_{c,n}^{k-1}, fd_{c,n}^k)$$

[0040]
$$Y_{c,n} = \min(fd_{c,n}^1, fd_{c,n}^2, \dots, fd_{c,n}^{k-1}, fd_{c,n}^k)$$

[0041] 여기서, n은 특징수이고, K는 데이터수이다. $fd_{c,n}^k$ 는 클래스 c에 속한 k 번째 데이터의 n번째 특징값을 의미하고, $Nf_{c,n}^k$ 는 $fd_{c,n}^k$ 를 정규화한 값이다.

[0042] 상기 정규화 후, 특징별로 각 클래스의 기준을 구한다. 상기 기준은 클래스에 속한 데이터의 해당 특징값들을 평균함으로써 구할 수 있다. 상기 기준은 하기의 수학적 식 2에 의하여 계산된다. 도 4에 나타난 바와 같이, 특징별 클래스의 막대 위에 표기된 점이 기준들이다.

[0043] [수학적 식 2]

[0044]
$$\mu_n^c = \frac{1}{k} \sum_{i=1}^k Nf_{on}^i$$

[0044]

[0045]

[0046] 여기서, k 는 클래스에 속한 데이터의 개수이며 u_n^c 는 n 번째 특징에 대한 c 번째 클래스의 평균을 의미한다.

[0047] 상기 기준을 구한 후, 특징선택은 특징별로 기준을 이용하여 표준편차를 계산하고, 표준편차를 이용하여 분별력이 좋은 특징을 선택할 수 있다. 즉, 클래스의 개수를 m 개, 특징의 개수를 n 개라 하면 j 번째 특징에 대한 c 번째 클래스($1 \leq c \leq n$)의 평균 u_j^c 을 구한 후 이들의 표준편차 σ_j^c 를 하기의 수학적 식 3을 이용하여 구할 수 있다. 모든 특징에 대해 동일하게 표준편차를 구하였으면 최종적으로 임계치를 적용하여 특징을 선택할 수 있다.

[0048] [수학적 식 3]

$$\sigma_j = \sqrt{\frac{1}{m} \sum_{c=1}^m (\mu_j^c - \bar{\mu}_j)^2}$$

[0049]

[0050] 도 5에 나타난 바와 같이, 랭크(rank)는 표준편차가 큰 값부터 작은 값까지 순위를 계산한 것이고, 임계치를 적용한 후 특정 값의 표준편차 이상 또는 이하를 갖는 특징을 선택할 수 있다.

[0051] 본 발명의 분위기 판별부(150)는 도 6에 나타난 바와 같은, 신경망 구조를 사용할 수 있다. 상기 신경망 구조의 각 층의 구성은 입력층 50개, 출력층 4개가 바람직하나, 이에 한정되는 것은 아니고 상황에 따라 가변적으로 구성할 수 있다. 또한, 은닉층을 가변적으로 구성시킬 수 있다.

[0052] 상기 입력층의 입력은 특징 축소부(140)에서 획득한 특징 데이터를 사용하고, 출력층은 4개로 구성하여 좌측부터 AV모델의 1사분면, 2사분면, 3사분면, 4사분면을 대응시킬 수 있다. 즉 01, 02, 03, 04의 값이 (1, 0, 0, 0)은 AV모델의 1사분면을 의미하고, (0, 1, 0, 0)은 2사분면, (0, 0, 1, 0)은 3사분면, (0, 0, 0, 1)은 4사분면을 의미한다. 도 7과 나타난 바와 같이, 각 분면에 해당하는 분위기를 나타낼 수 있다.

[0053] 분위기 판별을 위하여, 상기 신경망 구조를 선행적으로 학습을 시킬 수 있다. 보다 구체적으로, 특정 음악에 대하여 앞서 기술한 방법으로 대표 구간을 추출한 후, 설문 조사 등을 통하여 상기 대표 구간에 대한 사용자들의 분위기를 수집할 수 있다. 대표 구간의 특징과 설문 조사를 통하여 얻은 분위기를 매칭하여 대표 분위기를 정의한 후, 상기 신경망 구조를 학습시킬 수 있다.

[0054] 상기 대표 분위기를 정의하는 방법은 하기와 같은 방법을 사용할 수 있다.

[0055] 특정 음악에 대한 선택된 대표 구간에 대하여 여러 사용자가 다른 분위기를 지정할 수 있기 때문에 대표 분위기를 파악할 필요가 있다. 이를 위하여, 특정 음악 또는 음악의 대표 음원의 분위기에 대한 사용자들의 평가치의 합을 하기의 수학적 식 4와 같이 계산할 수 있다.

[0056] [수학적 식 4]

$$ed_i^s = \sum_{u=1}^n data_{ui}^s, \begin{cases} i = 1, 2, \dots, 12 \\ n: \text{피실험자 수} \end{cases}$$

[0057]

[0058] 여기서, ed_i^s 는 음원 s 에 대한 i 번째 분위기에 대한 피실험자들의 평가치 합이며, i 는 분위기 인덱스로 1~12까지의 범위를 가지며, 1번부터 반시계 방향으로 Pleased, happy, excited...relaxed에 대응할 수 있다(도 8 참조). $data_{ui}^s$ 는 음원 s 에 대한 피실험자 u 의 i 번째 분위기에 대한 평가치를 의미한다.

[0059] 모든 분위기에 대한 ed_i^s 를 구한 후, 이 값을 AV 공간 상의 좌표 값으로 변환할 수 있다. 변환은 각 ed_i^s 에 대해

원점(0,0)을 기점으로 하며 거리 ed_i^s 만큼의 위치에 표기한 후 그 점의 x,y 좌표값을 구할 수 있다. 즉, 수학적 5를 이용하여 x 값과 y 값을 구할 수 있다.

[0060] [수학적 5]

[0061] $x_i = \sin(f\theta_i)ed_i \quad i=1, 2, 3, \dots, 11, 12$

[0062] $y_i = \cos(f\theta_i)ed_i \quad i=1, 2, 3, \dots, 11, 12$

[0063] 여기서, $f\theta_i = f\theta_{i-1} + 30$, $2 \leq i \leq 12$ 이며, $f\theta_1$ 은 15도이다(즉, 15도마다 분위기를 나타내는 형용사에 대응한다).

[0064] 12개의 분위기에 대한 ed_i^s 의 값을 AV 좌표로 변환 후, 이들의 중심 좌표를 구하고 상기 중심좌표와 각 분위기 축과의 각도를 구하여 가장 근접한 분위기를 대표 분위기로 결정한다. 상기 해당 음원 또는 대표 구간에 대한 분위기의 대표 분위기를 정의한 후, 상기 음악의 대표 구간의 특징 및 상기 음악의 대표 구간에 대하여 정의된 분위기를 가지고 신경망 구조에 학습할 수 있다. 따라서, 상기 학습된 신경망 구조를 통하여 새로운 음악 또는 음악의 대표 구간의 특징으로부터 분위기를 판별할 수 있다.

[0065] [실시예 1] 음악 분위기 자동 판별 실험 및 확인

[0066] 신경망 학습을 위한 분위기 판별 학습 데이터를 얻기 위한 분위기 수집 시 사용한 모델은 도 7과 같은 Theyer의 2차원 모델을 사용하였다. 본 발명에서 구축한 데이터는 총 211명의 피 실험자가 참여하였고, 분위기 수집시 제공된 음원 샘플은 음악 101곡에서 추출된 총 281개로 189명에게는 41개의 음원 샘플을 랜덤하게 제공하고, 13명에게는 전체 음원샘플을 제공하였다.

[0067] 본 발명에서는 리브-원-아웃(Leave-one-out) 교차 검증방법을 이용하여 성능을 실험하였다. 리브-원-아웃(Leave-one-out) 교차 검증은 하나의 데이터를 테스트로 사용하고, 나머지 데이터를 학습으로 사용하는 방법으로 본 실시예에서 각 은닉층 갯수에 대하여 164번의 실험을 실시하였다. 즉, 2개의 은닉층부터 25개의 은닉층까지 각 164번의 실험을 실시하여 총 4,100번의 실험을 실시하였다. 또한 신경망 학습 수를 3,000번과 4,000번을 실시하였다.

[0068] 또한 본 발명에서 제안한 표준편차를 이용한 특징축소 방법의 성능을 비교하기 위해 기존문헌(문창배, 김현수, 송민균, 김병만, "AV 모델을 이용한 음악 분위기 자동 분류", 한국정보처리학회, 제35회 한국정보처리학회 춘계 학술발표대회 논문집 제18권 제1호, 2011)에서 사용한 결정계수를 이용한 특징축소 방법과 비교하였다. 실험환경, 실험방법, 실험에 참여한 실험자 수 등 모든 조건은 기존문헌의 방법과 동일하고, 특징 축소방법만 다르다. 상기 기존문헌에서는 특징축소 방법으로 결정계수를 사용하였다. 이의 결과를 도 8에 나타내었다.

[0069] 도 8에 나타난 바와 같이, 기존문헌의 방법에서 사용한 특징축소 방법의 경우 3,000번 학습, 히든층 17개에서 66.46%를 보이는 반면 본 발명의 특징 축소 방법에서는 3,000번 학습, 히든층 22개에서 70.12%의 판별성능을 보임을 확인하였다. 따라서, 본 발명에 따른 발명의 효과가 기존문헌의 방법보다 우수함을 확인하였다.

[0070] 도 9는 본 발명의 음악 분위기 자동 판별 방법을 설명하기 위한 흐름도이다.

[0071] 도 9를 참조하여 특징값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법에 대하여 설명한다.

[0072] 본 발명의 특징값의 표준편차를 기반으로 하여 음원의 특징 축소를 이용하는 음악 분위기 자동 판별 방법은 대표 구간 추출부(120)를 통하여 음악의 대표 구간을 추출하고(S110), 특징 추출부(130)를 통하여 대표 구간의 특징과 특징데이터를 추출하고(S120), 특징 축소부(140)를 통하여 상기 추출된 특징들 중 일부를 선택하여 특징의 갯수를 축소하기 위하여, 특징 데이터들을 정규화(S130)하고, 상기 특징 데이터들의 기준을 정의하고(S140), 상기 기준에 대한 표준편차를 이용하고, 임계치를 적용하여 특징을 선택하여 특징을 축소하고,(S150), 축소된 특징을 획득하고(S160), 분위기 판별부(150)를 통하여 상기 축소된 특징을 이용하여 음악의 분위기를 판별한다.

[0073] 본 발명은 특징 축소부(140)을 통하여 많은 음원의 특징들 중 일부를 선택하여 특징을 축소함으로써, 특징들 사이의 잡음을 제거하여, 보다 효과적으로 많은 음악파일들의 분위기를 정확하게 관별할 수 있다. 따라서, 분위기 자동 판정을 통한 내용 기반 필터링, 추천 시스템 및 정보 검색 분야 등 많은 분야에서 유용하게 사용될 수 있다.

[0074] 이상, 전술한 본 발명의 바람직한 실시예는, 예시의 목적을 위해 개시된 것으로, 당업자라면, 이하 첨부된 특허 청구범위에 개시된 본 발명의 기술적 사상과 그 기술적 범위 내에서, 또 다른 다양한 실시예들을 개량, 변경, 대체 또는 부가 등이 가능할 것이다.

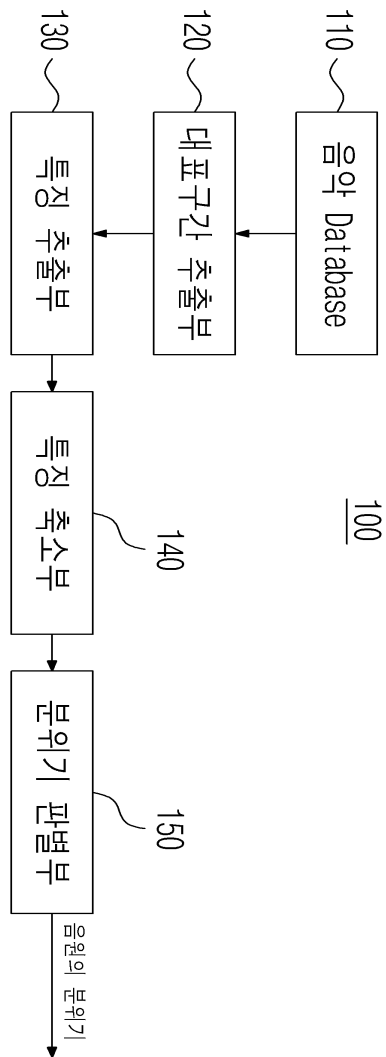
[0075]

부호의 설명

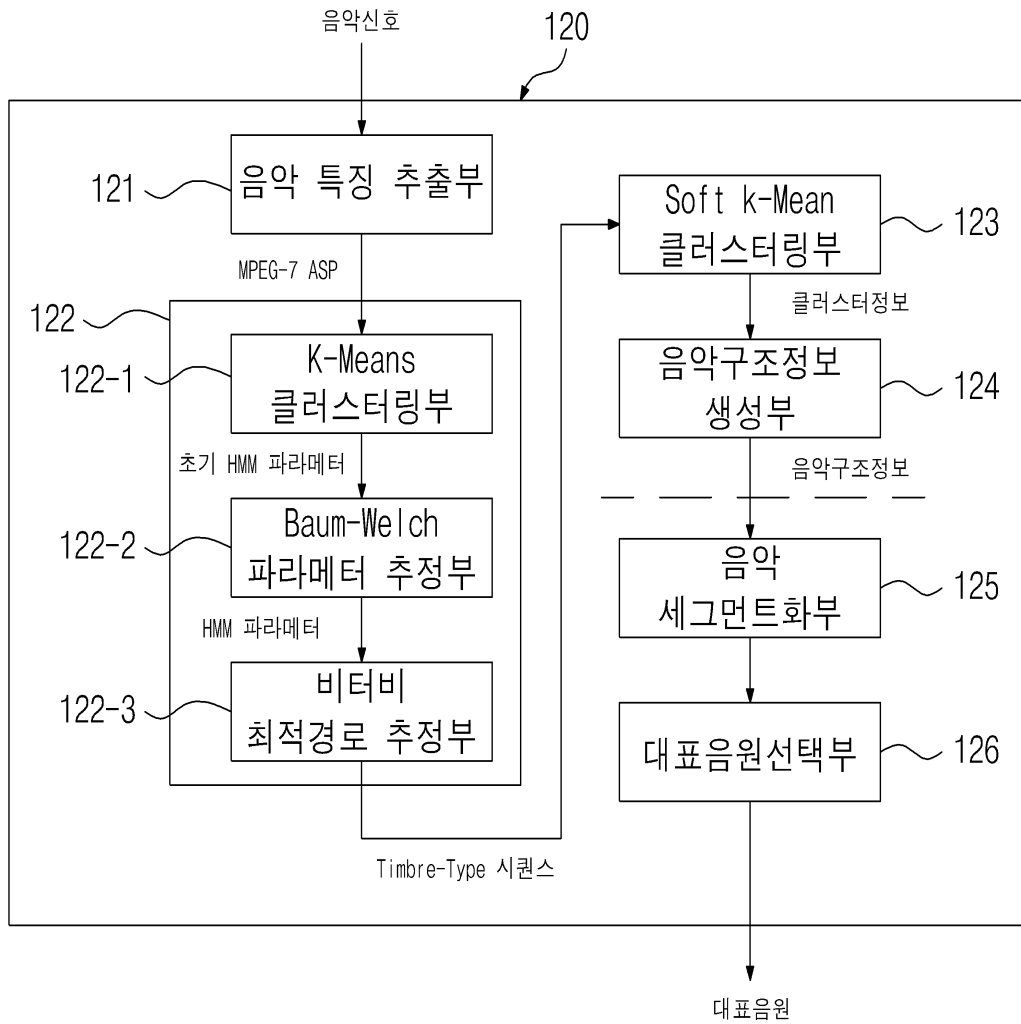
[0076]

- 110; 음악 데이터베이스
- 120; 대표 구간 추출부
- 121; 음악 특징 추출부
- 122; 음향-타입 시퀀스 추출부
- 122-1; K-민즈(K-means) 클러스터링 부
- 122-2; 바움-웰치 파라미터 추정부
- 122-3; 비터비 최적경로 추정부
- 123; 소프트 k-민즈(soft k-Means) 클러스터링부
- 124; 음악 구조 정보 생성부
- 125; 음악 세그먼트화부
- 126; 대표 음원 선택부
- 130; 특징 추출부
- 140; 특징 축소부
- 150; 분위기 관별부

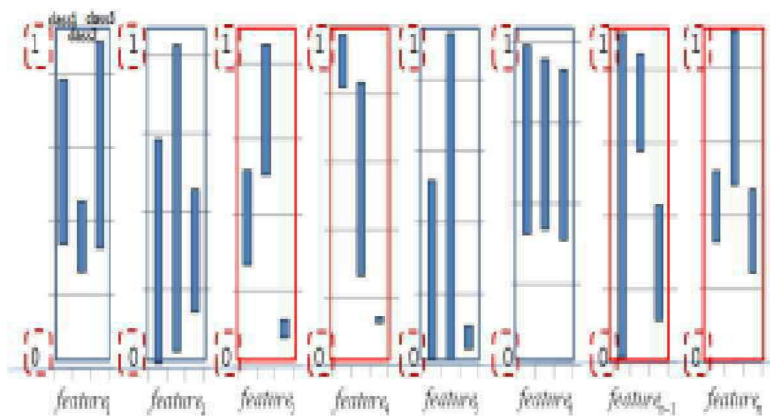
도면
도면1



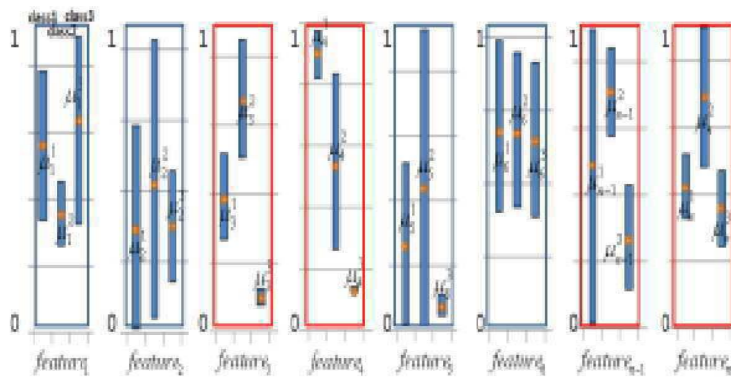
도면2



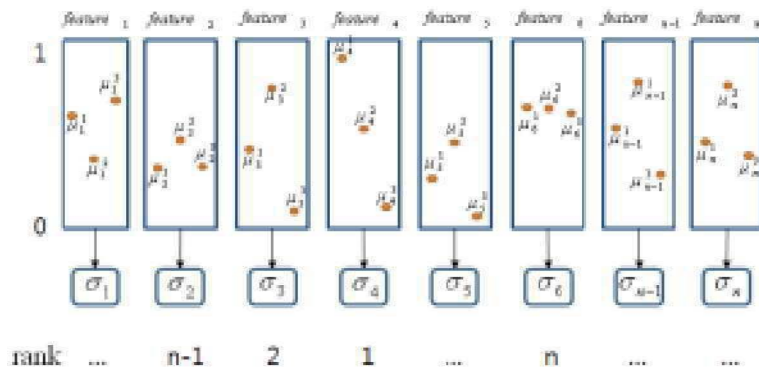
도면3



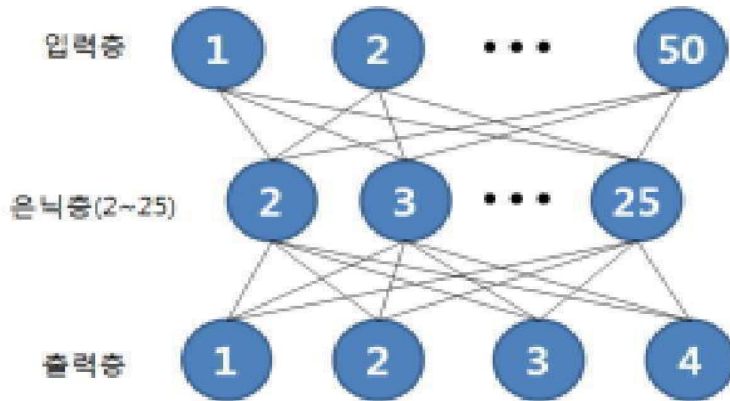
도면4



도면5



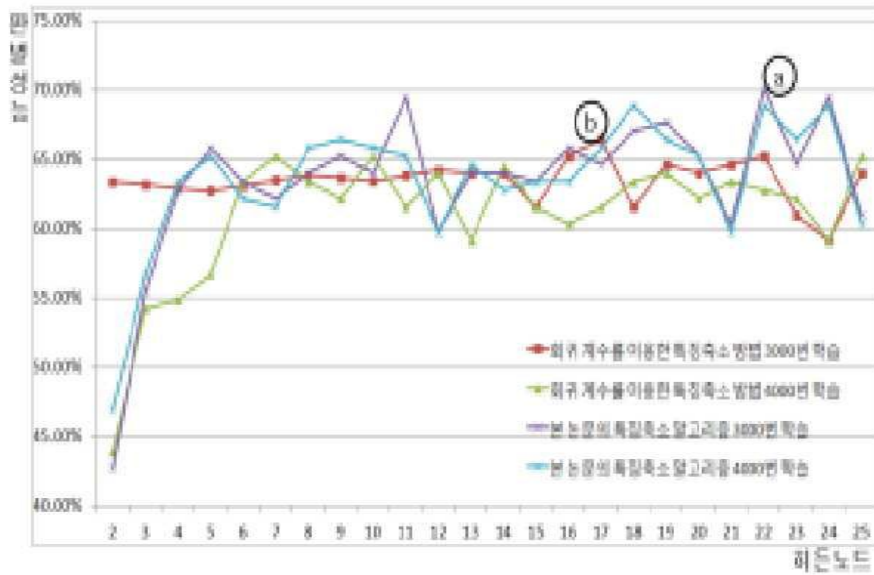
도면6



도면7



도면8



도면9

