



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년06월30일

(11) 등록번호 10-1530481

(24) 등록일자 2015년06월15일

(51) 국제특허분류(Int. Cl.)

G10L 25/51 (2013.01) G10L 19/04 (2006.01)

(21) 출원번호 10-2014-0039279

(22) 출원일자 2014년04월02일

심사청구일자 2014년04월02일

(56) 선행기술조사문헌

변가람 외 1인, '다중 옥타브 밴드 기반 음악 장르 분류 시스템', 전자공학회 논문지, 제50권 제12호, pp.238~244, 2013년 2월.*

Karam Byun et al., 'Music genre/mood/composer classification: MIREX 2013 submissions', The 14th International conference on music information retrieval, November 2013.*

정원진 외 1인, 'Multi-frame AR model을 이용한 LPC 계수 양자화', 한국음향학회지, 제31권 제2호, pp.93~99, 2012년.*

Shin-Cheol Lim, et al., 'Music genre/mood classification using a feature-based modulation spectrum', in Proc. IEEE Int. Conf. Mobile IT Convergence, pp.133~136, 2011.*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

세종대학교산학협력단

서울특별시 광진구 능동로 209 (군자동, 세종대학교)

(72) 발명자

김무영

서울 강남구 선릉로 221, 210동 402호 (도곡동, 도곡테크아파트)

변가람

서울 서대문구 증가로 191, 111동 1301호 (남가좌동, 삼성아파트)

(74) 대리인

특허법인태백

전체 청구항 수 : 총 13 항

심사관 : 정성윤

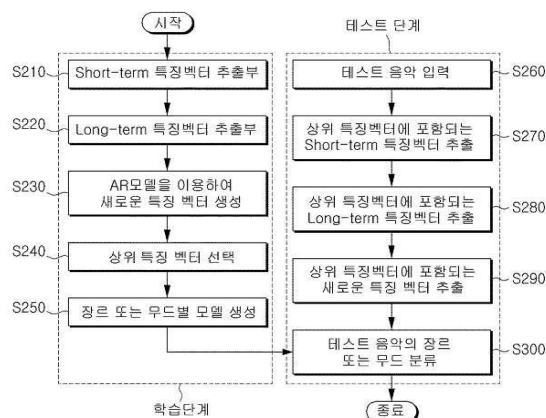
(54) 발명의 명칭 자기 회귀 모델을 이용한 음악 분류 장치 및 그 방법

(57) 요약

본 발명은 음악 분류 장치 및 그 방법에 관한 것이다. 본 발명의 음악 분류 장치는 음원을 입력받는 입력부; 상기 입력부에 입력된 음원의 음색 특징에 대한 Short-term 특징 벡터를 추출하는 Short-term 특징 추출부; 상기 Short-term 특징 벡터를 이용하여 Long-term 특징 벡터를 추출하는 Long-term 특징 추출부; 상기 추출된 Short-term

(뒷면에 계속)

대표도 - 도2



term 특징 벡터를 이용하여 선형예측계수(LPC)를 추출하고, 추출된 선형예측계수(LPC)를 자기 회귀 모델로 모델링하여 차수가 증가된 새로운 특징 벡터를 생성한 뒤, LSP 파라미터로 변환하는 AR 모델링부; 상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터 중에서 인식률이 높은 상위 특징 벡터를 선택하는 특징 선택부; 상기 선택된 상위 특징 벡터를 이용하여 음악의 분류 모델을 생성하는 모델 생성부; 및 상기 분류 모델을 기반으로 입력된 테스트 음악의 장르 또는 무드를 분류하는 음악 분류부를 포함한다.

이와 같이 본 발명에 따르면, 음악의 특징 벡터 추출 후의 상위 특징 벡터 선택 과정을 통해, 입력된 음악 데이터의 장르 분류에 요구되는 계산량을 줄일 수 있기 때문에 분류 결과를 얻는데 걸리는 시간을 단축 할 수 있으며, 높은 인식률을 얻을 수 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호	1345192901
부처명	미래창조과학부
연구관리전문기관	한국연구재단
연구사업명	일반 연구자 지원 사업
연구과제명	Interactive Music Application을 위한 객체기반 오디오 콘텐츠 코딩, 신호분리 및 잡음 제거 기법에 관한 기초 기술 연구
기 여 율	1/1
주관기관	세종대학교 산학협력단
연구기간	2013.05.01 ~ 2014.04.30

명세서

청구범위

청구항 1

음원을 입력받는 입력부;

상기 입력된 음원의 음색 특징에 대한 Short-term 특징 벡터를 추출하는 Short-term 특징 추출부;

상기 Short-term 특징 벡터를 이용하여 Long-term 특징 벡터를 추출하는 Long-term 특징 추출부;

상기 추출된 Short-term 특징 벡터를 이용하여 선형예측계수(LPC)를 추출하고, 추출된 선형예측계수(LPC)를 자기 회귀 모델로 모델링하여 차수가 증가된 상기 Short-term 특징 벡터에 대한 새로운 특징 벡터를 생성한 뒤, LSP 파라미터로 변환하는 AR 모델링부;

상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터 중에서 인식률이 높은 상위 특징 벡터를 선택하는 특징 선택부;

상기 선택된 상위 특징 벡터를 이용하여 음악의 분류 모델을 생성하는 모델 생성부; 및

상기 분류 모델을 기반으로 입력된 테스트 음악의 장르 또는 무드를 분류하는 음악 분류부를 포함하며,

상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터는,

MFCC(Mel-Frequency Cepstral Coefficient) 및 DFB(Decorrelated Filter Bank)를 포함하고,

상기 Short-term 특징 벡터는 Texture Window 기법을 이용하여 추출되는 음색에 대한 통계적 특징을 포함하되,

상기 통계적 특징은 평균값 및 분산값을 포함하는 음악 분류 장치.

청구항 2

삭제

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 Long-term 특징 벡터는,

FMS(Feature-based Modulation Spectrum)를 이용하여 추출되는 FMSFM(Feature-based Modulation Spectral Flatness Measures) 및 FMSCM(Feature-based Modulation Spectral Crest Measures)를 포함하며,

상기 FMSFM 및 FMSCM은 아래 수학식으로 정의되는 음악 분류 장치.

$$FMSFM(k) = \frac{M}{2} \frac{\sqrt{\prod_{m=0}^{M/2-1} FMS(k, m)}}{\sum_{m=0}^{M/2-1} FMS(k, m)}$$

$$FMSCM(k) = \frac{M}{2} \frac{\max_{n=0, \dots, M/2-1} (\overline{FMS}(k, m))}{\sum_{n=0}^{M/2-1} \overline{FMS}(k, m)}$$

청구항 5

제4항에 있어서,

상기 FMS 및 상기 FMS의 평균은,

아래 수학적식으로 정의되는 음악 분류 장치:

$$FMS_t(k, m) = \sum_{p \in P_t} X_t(k, p) \exp(-j2\pi mp/M)$$

$$\overline{FMS}(k, m) = \frac{1}{T} \sum_{t=0}^{T-1} FMS_t(k, m)$$

여기서, T는 총 Texture Window의 수이고, $X_t(k, p)$ 는 t번째 Texture Window내의 p번째 프레임의 Short-term 특징의 k번째 요소이고, P_t 는 t번째 Texture Window내에 속한 프레임의 총수이고, M은 Modulation Fourier-transform의 크기이다.

청구항 6

제1항에 있어서,

상기 Long-term 특징 벡터는 통계적 특징을 포함하되,

상기 통계적 특징은 평균값 및 분산값을 포함하는 음악 분류 장치.

청구항 7

제1항에 있어서,

상기 음악 분류부는,

분류 대상이 되는 상기 테스트 음악이 입력되면, 상기 테스트 음악에 대하여 상기 상위 특징 벡터에 해당하는 특징 벡터를 추출하고, 상기 추출된 특징 벡터를 기초로 상기 분류 모델과 비교하여 상기 입력된 테스트 음악의 장르 또는 무드를 분류하는 음악 분류 장치.

청구항 8

제1항에 있어서,

상기 특징 선택부는,

SVM(Support Vector Machine) ranker를 이용하여 상기 상위 특징 벡터를 선택하고,

상기 음악 분류부는,

One-against-one SVM(Support Vector Machine)를 이용하여 상기 테스트 음악의 장르 또는 무드를 분류하는 음악 분류 장치.

청구항 9

음악 분류 장치를 이용한 음악 분류 방법에 있어서,

입력된 음원의 음색 특징에 대한 Short-term 특징 벡터를 추출하는 단계;

상기 Short-term 특징 벡터를 이용하여 Long-term 특징 벡터를 추출하는 단계;

상기 추출된 Short-term 특징 벡터를 이용하여 선형예측계수(LPC)를 추출하는 단계;

추출된 선형예측계수(LPC)를 자기 회귀 모델로 모델링하여 차수가 증가된 상기 Short-term 특징 벡터에 대한 새로운 특징 벡터를 생성하여 LSP 파라미터로 변환하는 단계;

상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터 중에서 인식률이 높은 상위 특징 벡터를 선택하는 단계; 및

상기 선택된 상위 특징 벡터를 이용하여 음악의 분류 모델을 생성하는 단계;

상기 분류 모델을 기반으로 입력된 테스트 음악의 장르 또는 무드를 분류하는 단계를 포함하며,

상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터는,

MFCC(Mel-Frequency Cepstral Coefficient) 및 DFB(Decorrelated Filter Bank)를 포함하고,

상기 Short-term 특징 벡터는 Texture Window 기법을 이용하여 추출되는 음색에 대한 통계적 특징을 포함하되,

상기 통계적 특징은 평균값 및 분산값을 포함하는 음악 분류 방법.

청구항 10

삭제

청구항 11

삭제

청구항 12

제9항에 있어서,

상기 Long-term 특징 벡터는,

FMS(Feature-based Modulation Spectrum)를 이용하여 추출되는 FMSFM(Feature-based Modulation Spectral Flatness Measures) 및 FMSCM(Feature-based Modulation Spectral Crest Measures)를 포함하며,

상기 FMSFM 및 FMSCM은 아래 수학적식으로 정의되는 음악 분류 방법.

$$FMSFM(k) = \frac{M}{2} \frac{\sqrt{\prod_{m=0}^{M/2-1} \overline{FMS}(k, m)}}{\sum_{m=0}^{M/2-1} \overline{FMS}(k, m)}$$

$$FMSCM(k) = \frac{M}{2} \frac{\max_{n=0, \dots, M/2-1} (\overline{FMS}(k, m))}{\sum_{n=0}^{M/2-1} \overline{FMS}(k, m)}$$

청구항 13

제12항에 있어서,

상기 FMS 및 상기 FMS의 평균은,

아래 수학식으로 정의되는 음악 분류 방법:

$$FMS_t(k, m) = \sum_{p \in P_t} X_t(k, p) \exp(-j2\pi mp/M)$$

$$\overline{FMS}(k, m) = \frac{1}{T} \sum_{t=0}^{T-1} FMS_t(k, m)$$

여기서, T는 총 Texture Window의 수이고, $X_t(k, p)$ 는 t번째 Texture Window내의 p번째 프레임의 Short-term 특징의 k번째 요소이고, P_t 는 t번째 Texture Window내에 속한 프레임의 총수이고, M은 Modulation Fourier-transform의 크기이다.

청구항 14

제9항에 있어서,

상기 Long-term 특징 벡터는 통계적 특징을 포함하되,

상기 통계적 특징은 평균값 및 분산값을 포함하는 음악 분류 방법.

청구항 15

제9항에 있어서,

상기 분류 모델을 기반으로 입력된 음악의 장르 또는 무드를 분류하는 단계는,

분류 대상이 되는 상기 테스트 음악이 입력되면, 상기 테스트 음악에 대하여 상기 상위 특징 벡터에 해당하는 특징 벡터를 추출하는 단계; 및

상기 추출된 상위 특징 벡터를 기초로 상기 분류 모델과 비교하여 상기 테스트 음악의 장르 또는 무드를 분류하는 단계를 포함하는 음악 분류 방법.

청구항 16

제9항에 있어서,

상기 인식률이 기준값보다 높은 특징을 선택하는 단계에서,

SVM(Support Vector Machine) ranker를 이용하여 특징을 선택하고,

상기 분류 모델을 기반으로 입력된 음악의 장르 또는 무드를 분류하는 단계에서,

One-against-one SVM(Support Vector Machine)를 이용하여 상기 테스트 음악의 장르 또는 무드를 분류하는 음악 분류 방법.

청구항 17

제9항에 있어서,

상기 추출된 Short-term 특징 벡터를 이용하여 선형예측계수(LPC)를 추출하는 단계는;

레빈슨 알고리즘을 이용하여 상기 선형예측계수(LPC)를 추출하는 음악 분류 방법.

발명의 설명

기술 분야

[0001] 본 발명은 자기 회귀 모델을 이용한 음악 분류 장치 및 그 방법에 관한 것으로서, 연산량을 줄이고 입력된 음악의 장르 또는 무드를 정확하게 구별할 수 있는 자기 회귀 모델을 이용한 음악 분류 장치 및 그 방법에 관한 것이다.

배경 기술

- [0002] 최근 많은 대중 음악이 제작되고 있는 가운데 음악 데이터를 어떻게 하면 효과적으로 카테고리 별로 분류할 것인지, 특히 음악 데이터의 장르별 분류가 이슈가 되고 있다. 기존에는 일일이 사람이 수작업에 의해 음악 데이터를 카테고리 별로 분류하였다. 하지만, 방대한 디지털 음악 데이터를 분류하기 위해서는 작곡가, 가수, 장르 별로 자동 분류하는 알고리즘이 필요하다.
- [0003] 장르는 나라마다 혹은 사람마다 경계가 분명치 않고 문화, 가수, 시장에 따라 정의를 내리기 모호한 점이 있다. 음악 데이터 분석에 의한 자동 장르 분류는 효율적인 데이터관리와 음악 추천 등 다양한 어플리케이션에 적용이 가능하다. 또한, 수작업으로 분류를 하지 않아 경제적으로도 효율성이 있다.
- [0004] 음악 데이터 분석에 의한 장르 분류는 특징 벡터 추출, 분류기 등 다양한 방법으로 연구가 진행되고 발전하고 있다. 음악 장르 또는 무드를 분류하기 위해서는 일단 음악 데이터를 분석하여야 한다. 음악 데이터를 분석하는 방법에는 박(Beat), 리듬, 박자(Meter), 멜로디, 화음(Chord) 등 다양한 기준으로 분석하는 방법이 있다.
- [0005] 음악 데이터 분석 후에는 Gaussian Mixture Model(GMM)을 비롯하여 Hidden Markov Model(HMM), Nearest Neighbor(NN), K-Nearest Neighbor(KNN), Super Vector Machine(SVM)과 같은 다양한 분류기로 장르를 분류한다.
- [0006] Foote는 음악의 12차 MFCC(Mel-Frequency Cepstral Coefficient: 멜주파수 캡스트럴 계수)의 히스토그램을 만들고, 분류기로 NN을 사용하여 장르를 분류하였다. Bagci는 13차의 MFCC와 델타값을 구하고, Inter-Genre Similarity(IGS) 모델과 GMM분류기를 사용하여 장르를 분류하였다.
- [0007] Jiang은 음악의 Octave를 고려하여, MFCC와 다른 특징인 OSC(Octave-based Spectral Contrast: 옥타브기반 스펙트럴 콘트라스트)를 제안하여 장르 분류 성공률을 향상시켰다. 분류기로 GMM을 사용하여 Jazz, Pop, Romantic, Baroque, Rock의 5가지 장르에 대해서 약 82 % 장르 분류 성공률을 얻었다.
- [0008] 기존의 음악 장르 분류 방법 또는 시스템은 MFCC, Chroma, OSC 등 다양한 특징 벡터를 이용하여, SVM으로 장르를 인식한다. 그러나 상기 시스템은 낮은 인식 성공률 때문에 더 높은 성공률이 기대되는 음악 장르 분류 방법 또는 시스템에 대한 기술이 계속 논의되고 있다.
- [0009] 본 발명의 배경이 되는 기술은 국내공개특허 제2011-0013646호(2011.02.10 공개)에 개시되어 있다.

발명의 내용

해결하려는 과제

- [0010] 본 발명은 입력된 음악에 대하여 연산량을 줄이고 장르 또는 무드를 정확하게 구별하기 위한 자기 회귀 모델을 이용한 음악 분류 장치 및 그 방법에 관한 기술을 제공하는데 목적이 있다.

과제의 해결 수단

- [0011] 상기한 바와 같은 목적을 달성하기 위한 본 발명의 하나의 실시예에 따른 음악 분류 장치는, 음원을 입력받는 입력부; 상기 입력된 음원의 음색 특징에 대한 Short-term 특징 벡터를 추출하는 Short-term 특징 추출부; 상기 Short-term 특징 벡터를 이용하여 Long-term 특징 벡터를 추출하는 Long-term 특징 추출부; 상기 추출된 Short-term 특징 벡터를 이용하여 선형예측계수(LPC)를 추출하고, 추출된 선형예측계수(LPC)를 자기 회귀 모델로 모델링하여 차수가 증가된 새로운 특징 벡터를 생성한 뒤, LSP 파라미터로 변환하는 AR 모델링부; 상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터 중에서 인식률이 높은 상위 특징 벡터를 선택하는 특징 선택부; 상기 선택된 상위 특징 벡터를 이용하여 음악의 분류 모델을 생성하는 모델 생성부; 및 상기 분류 모델을 기반으로 입력된 테스트 음악의 장르 또는 무드를 분류하는 음악 분류부를 포함한다.
- [0012] 상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터는 MFCC(Mel-Frequency Cepstral Coefficient) 및 DFB(Decorrelated Filter Bank)를 포함할 수 있다.

[0013] 상기 Short-term 특징 벡터는 Texture Window 기법을 이용하여 추출되는 음색에 대한 통계적 특징을 포함하되, 상기 통계적 특징은 평균값 및 분산값을 포함할 수 있다.

[0014] 상기 Long-term 특징 벡터는, FMS(Feature-based Modulation Spectrum)를 이용하여 추출되는 FMSFM(Feature-based Modulation Spectral Flatness Measures) 및 FMSCM(Feature-based Modulation Spectral Crest Measures)를 포함하며, 상기 FMSFM 및 FMSCM은 아래 수학식으로 정의될 수 있다.

$$FMSFM(k) = \frac{M}{2} \frac{\sqrt{\prod_{m=0}^{M/2-1} FMS(k,m)}}{\sum_{m=0}^{M/2-1} FMS(k,m)}$$

$$FMSCM(k) = \frac{M}{2} \frac{\max_{n=0, \dots, M/2-1} (\overline{FMS}(k,m))}{\sum_{n=0}^{M/2-1} \overline{FMS}(k,m)}$$

[0015] 상기 FMS 및 상기 FMS의 평균은, 아래 수학식으로 정의될 수 있다.

$$FMS_t(k,m) = \sum_{p \in P_t} X_t(k,p) \exp(-j2\pi mp/M)$$

$$\overline{FMS}(k,m) = \frac{1}{T} \sum_{t=0}^{T-1} FMS_t(k,m)$$

[0016] 여기서, T는 총 Texture Window의 수이고, $X_t(k,p)$ 는 t번째 Texture Window내의 p번째 프레임의 Short-term 특징의 k번째 요소이고, P_t 는 t번째 Texture Window내에 속한 프레임의 총수이고, M은 Modulation Fourier-transform의 크기이다.

[0017] 상기 Long-term 특징 벡터는 통계적 특징을 포함하되, 상기 통계적 특징은 평균값 및 분산값을 포함할 수 있다.

[0018] 상기 음악 분류부는, 분류 대상이 되는 상기 테스트 음악이 입력되면, 상기 테스트 음악에 대하여 상기 상위 특징 벡터에 해당하는 특징 벡터를 추출하고, 상기 추출된 특징 벡터를 기초로 상기 분류 모델과 비교하여 상기 입력된 테스트 음악의 장르 또는 무드를 분류할 수 있다.

[0019] 상기 특징 선택부는, SVM(Support Vector Machine) ranker를 이용하여 상기 상위 특징 벡터를 선택하고, 상기 음악 분류부는, One-against-one SVM(Support Vector Machine)를 이용하여 상기 테스트 음악의 장르 또는 무드를 분류할 수 있다.

[0020] 본 발명의 다른 실시예에 따르면, 음악 분류 장치를 이용한 음악 분류 방법에 있어서, 입력된 음원의 음색 특징에 대한 Short-term 특징 벡터를 추출하는 단계; 상기 Short-term 특징 벡터를 이용하여 Long-term 특징 벡터를 추출하는 단계; 상기 추출된 Short-term 특징 벡터를 이용하여 선형예측계수(LPC)를 추출하는 단계; 추출된 선형예측계수(LPC)를 자기 회귀 모델로 모델링하여 차수가 증가된 새로운 특징 벡터를 생성하여 LSP 파라미터로 변환하는 단계; 상기 Short-term 특징 벡터, Long-term 특징 벡터 및 새로운 특징 벡터 중에서 인식률이 높은 상위 특징 벡터를 선택하는 단계; 및 상기 선택된 상위 특징 벡터를 이용하여 음악의 분류 모델을 생성하는 단계; 상기 분류 모델을 기반으로 입력된 테스트 음악의 장르 또는 무드를 분류하는 단계를 포함한다.

발명의 효과

[0025] 본 발명에 따르면, 음악의 특징 벡터 중에서 인식율이 높은 특징 벡터를 선택함으로써 입력된 음악의 장르 분류에 요구되는 계산량을 줄일 수 있으므로 분류 결과를 얻는데 걸리는 시간을 단축할 수 있으며, 높은 인식률을 얻을 수 있다.

도면의 간단한 설명

[0026] 도 1은 본 발명의 실시예에 따른 음악 분류 장치의 구성도이다.
 도 2는 본 발명의 실시예에 따른 음악 분류 방법의 순서도이다.
 도 3은 입력되는 음악 신호의 파형도이다.
 도 4는 Analysis Window를 설명하기 위한 도면이다.
 도 5는 Texture Window를 설명하기 위한 도면이다.
 도 6은 본 발명의 실시예에 따른 MFCC 알고리즘을 이용한 특징 추출 블록도이다.
 도 7은 본 발명의 실시예에 따른 DFB 알고리즘을 이용한 특징 추출 블록도이다.
 도 8은 본 발명의 실시예에 따른 OSC 알고리즘을 이용한 특징 추출 블록도이다.
 도 9는 본 발명의 실시예에 따른 LSP 변환 과정을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0027] 이하에서는 첨부된 도면을 참조하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0028] 먼저 본 발명의 음악 장르 인식 장치에 대하여 설명한다.

[0029] 도 1은 본 발명의 실시예에 따른 음악 분류 장치의 구성도이다.

[0030] 도 1에 도시한 바와 같이, 음악 분류 장치(100)는 입력부(110), Short-term 특징 추출부(120), Long-term 특징 추출부(130), AR 모델링부(140), 특징 선택부(150), 모델 생성부(160) 및 음악 분류부(170))를 포함한다.

[0031] 입력부(110)는 음성 또는 음악 신호를 프레임 단위로 입력 받는다. 본 발명에 따른 하나의 실시예로서 입력부(110)는 마이크를 통한 사람의 음성을 직접 입력 받거나 음악 데이터 베이스로부터 저장된 음악 데이터를 입력 받을 수 있다.

[0032] Short-term 특징 추출부(120)는 입력부(110)로 입력된 음성 또는 음악 신호의 Short-term 특징에 해당하는 음색 특징(Timbre feature) 또는 통계적 특징을 추출한다.

[0033] Long-term 특징 추출부(130)는 Short-term 특징 추출부(120)에서 추출된 음색 특징(Timbre feature)을 이용하여 Long-term 특징을 추출한다. 또한, Long-term 특징 추출부(130)는 통계적 특징도 추출한다.

[0034] AR 모델링부(140)는 Short-term 특징 추출부(120)에서 추출된 음색 특징(Timbre feature)을 이용하여 선형예측계수(LPC)를 추출하고, 추출된 선형예측계수(LPC)를 자기 회귀 모델(AR 모델)로 모델링하여 차수가 증가된 새로운 특징 벡터를 생성한다. 그리고, AR 모델링부(140)는 새로운 특징 벡터를 LSP 파라미터로 변환한다.

[0035] 특징 선택부(150)는 SVM ranker라는 알고리즘을 이용하여 Short-term 특징부(120), Long-term 특징부(130) 및 AR 모델링부(140)에서 추출된 전체 특징 벡터 중에서 인식률이 높은 상위 특징 벡터를 선택한다. 즉, 특징 선택부(150)는 Short-term 특징 벡터, Long-term 특징 벡터 및 LSP 변환된 새로운 특징 벡터에 대하여 인식률을 측정하고, 인식률이 기준값보다 높은 상위 특징 벡터를 선택한다. 전체 특징 벡터 중에서 선택된 상위 특징 벡터만이 음악 장르 분류에 이용되기 때문에 선택부(150)는 특징 추출에 필요한 계산량을 줄일 수 있으며 인식률을 향상시킬 수 있다.

- [0036] 모델 생성부(160)는 선택부(150)가 선택한 상위 특징 벡터를 기초로 각 음악 분류 모델을 생성한다. 구체적인 장르는 Classical, Country, Disco, Hiphop, Jazz, Rock, Blues, Reggae, Pop, Metal 등으로 구분될 수 있다.
- [0037] 음악 분류부(170)는 모델 생성부(160)가 생성한 음악 분류 모델을 기초로 테스트 음악에서 추출한 상위 벡터를 이용하여 테스트 음악의 장르 또는 무드를 인식기를 이용하여 분류한다.
- [0038] 이하에서는 본 발명의 실시예에 따른 음악 분류 장치(100)를 이용한 음악 장르 인식 방법에 관하여 더욱 상세하게 설명한다.
- [0039] 도 2는 본 발명의 실시예에 따른 음악 분류 방법의 순서도이다.
- [0040] 본 발명의 실시예에 따른 음악 장르 인식 방법은, 전체적으로 분류 모델을 생성하기 위한 학습단계와 생성된 모델을 통하여 입력된 테스트 음악의 장르 또는 무드를 분류하기 위한 테스트 단계를 포함한다. 도 2에 나타난 것과 같이 학습단계는 S210 내지 S250 단계를 포함하고, 테스트 단계는 S260 내지 S300 단계를 포함한다.
- [0041] 최종적으로 음악의 장르 또는 무드를 분류하기 위해서는 음악의 장르 또는 무드를 인식하는 것이 선행되어야 하는데 음악의 장르 또는 무드를 인식하기 위하여, 음악의 분류 모델을 정확히 생성해야 한다.
- [0042] 음악 장르의 인식을 위해서 본 발명의 실시예에 따른 음악 분류 장치(100)는 먼저 각 장르의 음악에 대해서 특징 벡터를 추출하는데, 특징 벡터의 종류로는 Short-term 및 Long-term 특징 벡터, LSP 변환된 Short-term에 대한 새로운 특징 벡터를 포함한다.
- [0043] 도 3은 입력되는 음악 신호의 파형도이다. 도 4는 Analysis Window를 설명하기 위한 도면이며, 도 5는 Texture Windows를 설명하기 위한 도면이다.
- [0044] 도 3에 도시한 바와 같이 입력부(110)를 통하여 음악 신호가 입력되면, Short-term 특징 추출부(120)는 도 4와 같은 Window를 통해서 프레임별 Short-term 특징 벡터를 추출한다(S210). Short-term 특징 추출부(120)는 각 프레임(Analysis Window)에서 추출된 Short-term 특징 벡터를 기초로, 도 5의 Texture Window 기법을 이용하여 통계적 특징 벡터를 추출한다. 통계적 특징은 평균(Mean) 또는 분산(Variance) 중에서 적어도 하나를 포함한다.
- [0045] 이하에서는, Short-term 특징 벡터에 해당하는 음색 특징(Timbre Feature)을 추출하는 단계(S210)에 대해 더욱 상세하게 설명한다.
- [0046] 본 발명의 하나의 실시예에 따른 Short-term 특징 추출은, 음성인식과 화자인식에 주로 사용되는 MFCC(Mel-Frequency Cepstral Coefficient: 멜주파수 캡스트럴 계수), mel-scale band pass filter와 high pass filter를 이용하는
- [0047] DFB(Decorrelated Filter Bank: 상관감소 필터 뱅크), 음악 신호의 특성을 octave band 별로 나타내며, 음악 인식에 사용되는 OSC(Octave-based Spectral Contrast: 옥타브기반 스펙트럴 콘트라스트)를 포함한다.
- [0048] 이하 도 6을 참조하여 MFCC 알고리즘에 대해 설명한다.
- [0049] 도 6은 본 발명의 실시예에 따른 MFCC 알고리즘을 이용한 특징 추출 블록도이다.
- [0050] 먼저, Short-term 특징 추출부(120)는 시간 도메인에서 각 프레임에 Hamming window를 적용 후 Fast Fourier Transform(FFT)을 수행한다. 그 후 밴드의 수가 B개인 Mel-scale Band-pass Filter로 스펙트럼을 스케일링 한 후 각 스펙트럼의 가중치 합 $WS(b)_{b=1, \dots, B}$ 를 구한다. 그 후 이 값에 log를 적용한다($1 \leq b \leq B$). 마지막으로, Short-term 특징 추출부(120)는 각 가중치 합에 Discrete Cosine Transform (DCT)을 적용하여 K차원의 MFCC 특징 벡터를 추출한다.
- [0051] Short-term 특징 추출부(120)가 MFCC를 이용해 추출하는 특징은 아래 수학식 1로 정의된다.

수학식 1

$$MFCC(k) = \sum_{b=0}^{B-1} \log(WS(b)) \cos\left(k \frac{\pi}{B}(b+0.5)\right), \quad 0 \leq k \leq K$$

[0052]

[0053] 다음으로, 도 7을 참조하여 DFB(Decorrelated Filter Bank: 상관감소 필터 뱅크)에 대해 설명한다.

[0054] 도 7은 본 발명의 실시예에 따른 DFB 알고리즘을 이용한 특징 추출 블록도이다.

[0055] Short-term 특징 추출부(120)는 DFB 알고리즘을 통해 MFCC와 마찬가지로 각 밴드별 스케일링된 값들을 더한 후, log를 적용하여 $\log(WS(b)) (1 \leq b \leq B)$ 를 생성한다. 그리고, 도 7과 같이 마지막 과정에서 DCT를 대신하여 FIR High-pass-filter $H(z) = 1 - z^{-1}$ 를 통과하게 한다. 최종적으로 Short-term 특징 추출부(120)는 아래의 수학적식 2와 같이 D차원의 DFB 특징 벡터를 추출한다.

수학적식 2

[0056]
$$DFB(d) = \log(WS(d+1)) - \log(WS(d)), \quad 0 \leq d \leq D < B$$

[0057] 다음으로, 도 8를 참조하여 OSC(Octave-based Spectral Contrast: 옥타브 기반 스펙트럴 콘트라스트)에 대해 설명한다.

[0058] 도 8은 본 발명의 실시예에 따른 OSC 알고리즘을 이용한 특징 추출 블록도이다.

[0059] OSC는 MFCC나 DFB와 다르게 청각모델이 아닌 옥타브(Octave)를 기반으로 특징을 추출한다. Short-term 특징 추출부(120)는 OSC 알고리즘을 이용하여 각 밴드별 스펙트럼의 Peak와 Valley값을 고려한 특징 벡터를 추출한다. 대부분의 음악에서 강한 Peak는 Harmonic부분과 연관되며, 강한 Valley는 Non-harmonic부분과 연관된다. 따라서, OSC는 밴드별 스펙트럼의 Peak와 Valley값을 고려함으로써, 음악의 Harmonic과 Non-harmonic 성분을 나타낼 수 있다.

[0060] 도 8에 도시한 바와 같이, OSC는 MFCC나 DFB와 다르게 멜스케일 밴드 패스 필터(Mel-scale Band-pass Filter)를 사용하지 않고, 옥타브스케일 밴드 패스 필터(Octave-scale Band-pass Filter)를 사용한다. 본 발명에서는 아래 표 1과 같이 8개의 Octave-scale Band-pass Filter를 이용한다.

표 1

[0061]

Band	Frequency(Hz)
1	0~100
2	100~200
3	200~400
4	400~800
5	800~1600
6	1600~3200
7	3200~8000
8	8000~22050

[0062] 한 프레임의 길이와 FFT 포인트가 N으로 같을 때, 각 프레임에 대한 FFT 스펙트럼은 $\{x_1, x_2, \dots, x_N\}$ 로 정의할 수 있다. i번째 밴드에 해당하는 FFT포인트 수가 K_i 이면, i번째 밴드의 스펙트럼은 $\{x_{i,1}, x_{i,2}, \dots, x_{i,K_i}\}$ 와 같이 나타낼 수 있다. Peak와 Valley를 구하기 위해선 먼저 각 밴드별 스펙트럼을 내림차순으로 정리한다. 내림차순 정리된 i번째 밴드의 스펙트럼 $\{x_{\text{궁},1}, x_{\text{궁},2}, \dots, x_{\text{궁},K_i}\}$ 을 가지고, 아래 식을 이용하여 Peak와 Valley를 추출할 수 있다.

수학식 3

$$P_i = \log \left(\frac{1}{\alpha K_i} \sum_{j=1}^{\alpha K_i} x'_{i,j} \right)$$

[0063]

수학식 4

$$V_i = \log \left(\frac{1}{\alpha K_i} \sum_{j=1}^{\alpha K_i} x'_{i, K_i - j + 1} \right)$$

[0064]

[0065]

여기서 α 는 주변 값의 범위에 대한 상수로서 0.02부터 0.2까지 실험한 결과 α 는 성능에 중요한 영향을 미치지 않는 바, 0.02로 설정하여 사용할 수 있다. 수학식 3 및 수학식 4에서 구한 Peak값과 Valley값의 차이를 계산하여 다음의 수학식 5와 같이 Spectral Contrast(SC_i)를 구한다.

수학식 5

$$SC_i = P_i - V_i$$

[0066]

[0067]

OSC는 다음과 같이 1개 밴드에 대한 Spectral Contrast와 Valley $\{SC_1, SC_2, \dots, SC_I, V_1, V_2, \dots, V_I\}$ 를 특징 벡터로 사용하게 된다. 본 발명에서는, 8개의 밴드를 사용하기 때문에 16차 OSC 특징 벡터를 사용하게 된다. i 번째 밴드에 대해서 OSC는 Spectral Contrast와 Valley를 $\{SC_i, V_i\}$ 와 같이 특징 벡터로 사용하게 된다.

[0068]

Short-term 특징 추출부(120)는 다음의 수학식 6 및 수학식 7을 통하여 Analysis Window에서 추출된 특징 벡터들의 평균($\mu_t(k)$)과 분산($\sigma_t^2(k)$)을 추출한다.

수학식 6

$$\mu_t(k) = \frac{1}{P} \sum_{p=0}^{P-1} X_t(k, p)$$

[0069]

수학식 7

$$\sigma_t^2(k) = \frac{1}{P} \sum_{p=0}^{P-1} (X_t(k, p) - \mu_t(k))^2$$

[0070]

[0071]

상기 수학식에서, $X_t(k, p)$ 는 t 번째 Texture Window에 p 번째 프레임의 음색 특징(Timbre feature)의 k 번째 요소이고, P 는 Texture Window안에 포함된 총 프레임 개수이다.

[0072] t번째 Texture Window에 구해진 통계적 특징 벡터인 $\mu_t(k)$, $\sigma_t^2(k)$ 는 전체 Texture Window에 대하여 각각 평균을 구하면 아래 수식식 8 및 수식식 9와 같다.

수식식 8

[0073]

$$\bar{\mu}(k) = \frac{1}{T} \sum_{t=0}^{T-1} \mu_t(k)$$

수식식 9

[0074]

$$\bar{\sigma}^2(k) = \frac{1}{T} \sum_{t=0}^{T-1} \sigma_t^2(k)$$

[0075] 상기의 수식식 8 및 수식식 9와 같이, Short-term 특징 추출부(120)가 Short-term 특징을 추출하면, Long-term 특징 추출부(130)는 추출된 Short-term 특징을 이용하여 Long-term 특징을 추출한다(S220). 이하 Long-term 특징 중의 하나인 변조 스펙트럼(Modulation Spectrum)을 사용한 스펙트럼 특징(Spectro-temporal Feature)에 대해 설명한다.

[0076] 본 발명의 실시예에 따르면 Long-term 특징 추출부(130)는 변조 스펙트럼(Modulation Spectrum)을 사용한 특징 벡터로 MSFM(Modulation Spectral Flatness Measures: 변조 스펙트럴 플래트니스 측정), MSCM(Modulation Spectral Crest Measures: 변조 스펙트럴 크레스트 측정)를 추출한다.

[0077] 또한, Long-term 특징 추출(130)는 상기 MSFM/MSCM을 기반으로 FMSFM(Feature-based Modulation Spectral Flatness Measures: 특징기반 변조 스펙트럴 플래트니스 측정) 및 FMSCM(Feature-based Modulation Spectral Crest Measures: 특징기반 변조 스펙트럴 크레스트 측정)을 추출할 수 있다.

[0078] MSFM/MSCM은 옥타브 밴드 합(Octave Band Sum: OBS)의 변조 스펙트럼(Modulation Spectrum)을 이용하는 반면, 본 발명의 실시예에 따른 FMSFM/FMSCM은 FMS(Feature-based Modulation Spectrum: 특징기반 변조 스펙트럼)를 이용하여 추출된다. Long-term 특징 추출부(130)는 FMS를 이용하여 각 Feature Dimension별로 변조 스펙트럼(Modulation Spectrum)을 추출하는데, 이를 통해 각 Feature Dimension이 시간에 따라 어떻게 변화하는지를 알 수 있다.

[0079] S220단계에 대하여 더욱 상세하게 설명하면, 먼저 Long-term 특징 추출부(130)는 다음의 수식식 10과 같이 FMS를 추출한다.

수식식 10

[0080]

$$FMS_t(k, m) = \sum_{p \in P_t} X_t(k, p) \exp(-j2\pi mp/M)$$

[0081] 여기서 $X_t(k, p)$ 는 t번째 texture window내의 p번째 프레임의 음색 특징(Timbre Feature)의 k번째 요소이고, P_t 는 t번째 Texture Window내의 속한 프레임의 총 수이고, M은 변조 푸리에 변환(Modulation Fourier-transform)의 크기이다. 그리고 Long-term 특징 추출부(130)는 전체 Texture Window에 대하여 아래와 같이 평균 FMS를 추출한다.

수학식 11

$$\overline{FMS}(k, m) = \frac{1}{T} \sum_{t=0}^{T-1} FMS_t(k, m)$$

여기서 T는 총 Texture Window의 개수이다.

다음으로, Long-term 특징 추출부(130)는 수학식 10 및 수학식 11의 결과를 이용하여 FMSFM/FMSCM을 다음과 같이 추출한다.

수학식 12

$$FMSFM(k) = \frac{M}{2} \frac{\sqrt{\prod_{m=0}^{M/2-1} \overline{FMS}(k, m)}}{\sum_{m=0}^{M/2-1} \overline{FMS}(k, m)}$$

수학식 13

$$FMSCM(k) = \frac{M}{2} \frac{\max_{n=0, \dots, M/2-1} (\overline{FMS}(k, m))}{\sum_{n=0}^{M/2-1} \overline{FMS}(k, m)}$$

수학식 12에서 연산된 FMSFM의 작은값(예를 들면 0)과 큰값(예를 들면 1 또는 1이상의 값)은 각각 평균 FMS의 Peakiness와 Flatness를 나타낸다. FMSCM은 FMSFM과 반대의 성향을 보인다. 만약 k번째 FMSFM이 매우 작은 값을 가진다면, 입력 음악에서 k번째 Modulation 주파수가 반복되는 패턴을 가진다.

Long-term 특징 추출부(130)는 MSFM/MSCM를 추출하기 위해 Octave Band내의 에너지인 Octave Band Sum (OBS)을 사용한다.

다음으로 AR 모델링부(140)는 Short-term 특징 추출부(120)에서 추출된 음색 특징(Timbre feature)을 이용하여 선형예측계수(LPC)를 추출하고, 추출된 선형예측계수(LPC)를 자기 회귀 모델(AR 모델)로 모델링하여 차수가 증가된 Short-term에 대한 새로운 특징 벡터를 생성한다(S230). 그리고, AR 모델링부(140)는 새로운 특징 벡터를 LSP 파라미터로 변환한다.

이하에서는 도 9를 통하여 S230 단계에서 LSP 변환된 Short-term에 대한 특징 벡터를 생성하는 과정에 대하여 더욱 상세하게 설명한다.

도 9는 본 발명의 실시예에 따른 LSP 변환 과정을 설명하기 위한 도면이다.

AR 모델링부(140)는 modulation spectrum과 유사하지만 푸리에 변환이 아닌 레빈슨(levinson) 알고리즘을 이용하여 시간축의 선형 예측 계수(Linear Prediction Coding Coefficient, LPC)를 추출한다.

[0093] 그리고, 도 9와 같이 선형예측계수(LPC)를 자기 회귀 모델(Autoregressive Model, AR 모델)을 통해 모델링하면 기존의 Short-term특징 벡터에 비하여 차수가 증가한 새로운 특징 벡터가 생성된다.

[0094] 예를 들어, 입력된 Short-term 특징 벡터가 13차인 경우, 자기 회귀 모델(AR 모델)을 이용하면 10배가 증가된 130차수의 새로운 특징 벡터가 생성된다.

[0095] 그러나, 선형예측계수(LPC)는 다이내믹 레인지(dynamic range)가 크기 때문에 인식 에러가 발생할 수 있으므로, AR 모델링부(140)는 자기 회귀 모델(AR 모델)에 의해 생성된 새로운 특징 벡터들을 LSP 파라미터(Line Spectrum Pairs Parameter)로 변환하여 각 장르/무드 별 모델을 생성한다.

[0096] LSP 파라미터는 다이내믹 레인지가 선형예측계수(LPC)보다 상대적으로 적고, 인접한 LSP의 상대적 거리에 의하여 스펙트럼의 폴(pole)과 제로(zero)를 정확하게 유추할 수 있으므로 선형예측계수(LPC)만을 사용하는 경우에 비하여 우수한 성능을 가진다.

[0097] 이와 같이 Short-term 특징 추출부(120), Long-term 특징 추출부(130), AR 모델링부(140)는 각각 특징 벡터를 생성하고, 실제 시스템에서 추출되는 전체 특징에 해당하는 특징 벡터는 아래 표 2와 같이 예시할 수 있다.

표 2

[0098]

Feature vectors			Dimension
Texture window	Mean	MFCC	13
		DFB	13
		OSC	16
	Variance	MFCC	13
		DFB	13
		OSC	16
Feature-based modulation spectrum	FMSFM	MFCC	13
		DFB	13
		OSC	16
	FMSCM	MFCC	13
		DFB	13
		OSC	16
Feature-based autoregressive model	LSP	MFCC	130
		DFB	130
		OSC	160
Total			588

[0099] 이와 같이, 표 2에서 보는 바와 같이 Short-term 특징 추출부(120), Long-term 특징 추출부(130) 및 AR 모델링부(140)는 입력된 음악에 대하여 588개의 특징 벡터를 생성할 수 있다. 즉, Short-term 특징 추출부(120)는 S210단계에서 설명한 것처럼, Texture window 기법을 통하여 입력된 음악 신호 중에서 13개의 차수 각각에 대하여 MFCC의 평균 값(Mean)에 대한 13개의 특징 벡터를 획득하고, 13개의 차수 각각에 대하여 DFB의 평균 값(Mean)에 대한 13개의 특징 벡터를 획득한다.

[0100] 그리고, Long-term 특징 추출부(130)는 S220단계에서 설명한 것처럼, Feature-based modulation spectrum을 이용하여 13개의 차수 각각에 대하여 MFCC의 FMSFM 연산된 13개의 특징 벡터를 획득하고, 13개의 차수 각각에 대하여 DFB의 FMSCM 연산된 13개의 특징 벡터를 획득한다.

[0101] 마찬가지로, AR 모델링부(140)는 S220단계에서 설명한 것처럼, Feature-based autoregressive model을 이용하여 차수가 10배로 증가된 130개의 차수 각각에 대하여 MFCC의 LSP 변환된 130개의 특징 벡터를 획득하고, 130개의 차수 각각에 대하여 DFB의 LSP 변환된 130개의 특징 벡터를 획득한다.

[0102] 다음으로 특징 선택부(150)는 추출된 Short-term 또는 Long-term 특징 벡터 중에서 상위 특징 벡터를 선택(feature selection)한다(S240). 여기서 상위 특징 벡터란 인식률이 기준값보다 높은 특징 벡터를 의미한다.

[0103] 이와 같이 본 발명의 실시예에 따르면 종래 기술과 달리, 전체 특징 벡터들 중에서 우선 순위에 대응하는 특징 벡터를 선택함으로써, 인식률을 향상시키고 계산량을 줄일 수 있다. 표 2를 예로 들면, 특징 선택부(150)는 획득한 588개의 특징 벡터 중에서 인식률이 우수한 160개의 상위 특징 벡터를 선택한다. 따라서, 테스트 단계에서

는 추출된 전체 특징들 중에서 특징 선택부(150)에 의해 선택된 상위 특징 벡터만이 장르 인식에 사용된다.

- [0104] 그리고, 본 발명의 실시예에 따른 특징 선택부(150)는 Best-first search와 Support Vector Machine(SVM) ranker를 이용하는 방법 등을 통하여 우선 순위에 대응하는 상위 특징 벡터를 선택할 수 있다. 즉, 특징 선택부(150)는 SVM-ranker라는 SVM 인식기를 이용하여 각 특징별로 가치를 평가하여 Ranking을 매겨 특징 벡터를 선택할 수 있는데, SVM-ranker를 통하여 Ranking을 매기는 기술은 당업자라면 용이하게 실시할 수 있는 내용이므로 상세한 설명은 생략한다.
- [0105] 다음으로 모델 생성부(160)는 선택된 상위 특징 벡터를 이용하여 음악의 분류 모델을 생성한다(S240). 구체적인 장르는 Classical, Country, Disco, Hiphop, Jazz, Rock, Blues, Reggae, Pop, Metal 등으로 구분될 수 있다. 구체적인 방법에 대해 살펴보면, 모델 생성부(160)는 추출된 상위 특징 벡터를 이용하여 음악의 각 장르별로 모델링을 수행한다. 즉, 모델 생성부(160)는 각각의 음악 장르에 대응하는 상위 특징 벡터의 특징 값들을 그룹핑하여 각 장르별로 모델을 생성한다.
- [0106] 이와 같이 분류 모델이 생성되면, 입력된 테스트 음악의 장르 또는 무드를 분류하기 위한 테스트 단계가 진행된다.
- [0107] 먼저, 도 2와 같이 분류 대상이 되는 테스트 음악 신호가 입력되면(S260), 음악 분류부(170)는 입력된 음악 신호에서 상위 특징 벡터에 포함되는 Short-term 특징 벡터, Long-term 특징 벡터, AR 모델에 의해 생성된 새로운 특징 벡터를 순차적으로 추출한다(S270, S280, S290). 즉, 표 2의 경우에는, 음악 분류부(170)는 588개 전체에 대한 특징 벡터를 추출할 필요가 없이 우선 순위에 해당하는 160개의 특징 벡터만을 추출한다.
- [0108] 예를 들면, Texture window 기법을 통하여 획득한 MFCC의 평균 값(Mean)에 대하여 13개의 차수 중에서 1, 4, 6, 9, 13차에 대응하는 음악 신호의 특징 벡터가 우선 순위에 포함된다고 가정하면, 음악 분류부(170)는 입력된 테스트 음악 신호 중에서 상기 5개에 대응하는 특징 벡터를 추출하도록 한다.
- [0109] 여기서, Short-term 특징 벡터를 추출하는 단계, Long-term 특징 벡터를 추출하는 단계, AR 모델에 의해 생성된 새로운 특징 벡터를 추출하는 단계(S270, S280, S290)는 상위 특징 벡터에 해당하는 Short-term 특징 벡터, Long-term 특징 벡터, 새로운 특징 벡터를 추출한다는 점을 제외하고 상기 S210, S220, S230와 실질적으로 동일한 바, 중복되는 설명은 생략한다.
- [0110] 또한 설명의 편의상 음악 분류부(170)가 Short-term 특징 벡터, Long-term 특징 벡터, 새로운 특징 벡터를 추출하는 것(S270, S280, S290)으로 설명하였으나, Short-term 특징 추출부(120), Long-term 특징 추출부(130) 및 AR 모델링부(140)가 테스트 과정을 직접 수행할 수도 있다.
- [0111] 다음으로, 음악 분류부(170)는 추출된 상위 특징 벡터에 해당하는 특징 벡터를 이용하여 모델 생성부(160)가 생성한 분류 모델을 기반으로 입력된 테스트 음악의 장르 또는 무드를 분류한다(S300). 특히 본 발명의 실시예에 따르면 음악 분류부(170)는 One-against-one SVM을 사용하여 음악 장르 또는 무드를 분류할 수 있다.
- [0112] 이와 같이 본 발명의 실시예에 따르면 테스트 단계에서 상위 특징 벡터만을 추출하여 음악 장르 또는 무드를 분류하는데 이용하기 때문에, 입력되는 테스트 음악의 모든 특징 벡터를 추출하여 음악 장르 또는 무드를 분류하는 종래 기술에 비하여 연산량이 월등히 감소할 수 있다.
- [0113] 본 발명의 실시예에 따른 하나의 실험예로서 음악 장르 분류 성능 평가를 위해 GTZAN 데이터 베이스와 MIREX 무드 클러스터링 방법을 이용하였다.
- [0114] 즉, 장르 인식에는 GTZAN 데이터 베이스를 이용하고, 무드 인식에는 MIREX 무드 클러스터링 방법을 이용하여 성능 평가를 수행하였다.
- [0115] GTZAN은 Classical, Country, Disco, Hiphop, Jazz, Rock, Blues, Reggae, Pop, Metal로 총 10개의 장르를 포함하며, 각 장르당 100곡, 한 곡당 30초로 16bit, 22050Hz, 모노, AU파일 포맷으로 구성되어있다.

표 3

분류기	Accuracy (%)		
	장르	무드	차수
All feature set	82.4	62.2	588
+ Feature selection	85.0	69.7	160

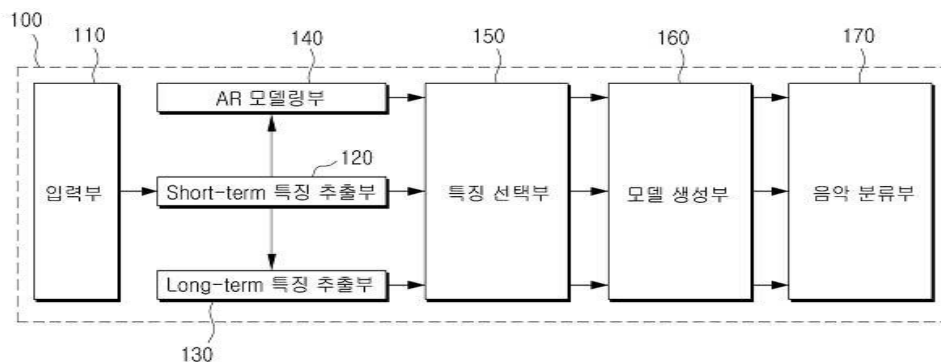
- [0117] 표 3에 나타난 정확성 결과에서 나타났듯이, 588개의 전체 특징 벡터를 모두 추출하여 장르 또는 무드를 인식하는 경우에 비하여, 상위 160개의 특징 벡터를 선별하여, 본 발명의 실시예에 따라서 선별된 특징 벡터에 해당하는 특징 벡터를 추출하여 장르 또는 무드를 인식하는 경우가 장르 또는 무드 인식의 정확성이 높다는 것을 알 수 있다.
- [0118] 이와 같이 본 발명의 실시예에 따른 음악 장르 분류 방법 및 그 장치에 따르면, 음악의 특징 추출 후의 상위 특징 벡터 선택 과정을 통해, 입력된 음악 데이터의 장르 분류에 요구되는 계산량을 줄일 수 있기 때문에 분류 결과를 얻는데 걸리는 시간을 단축 할 수 있다. 또한, 상위 특징 벡터 선택을 통해 입력된 음악 데이터의 장르 또는 무드를 분류하더라도 전체 특징 벡터를 사용하는 종래의 방법과 비교하여 높은 인식률을 얻을 수 있다.
- [0119] 또한, 특징 벡터 추출시, 음색(timbre) 특징 이외에도 자기 회귀 모델(AR 모델)을 이용하여 음색(timbre) 특징을 기반으로 한 LSP 파라미터를 추출하여 특징 벡터로 이용함으로써, 장르 및 무드 인식 시스템의 성능을 향상시킬 수 있다.
- [0120] 이제까지 본 발명에 대하여 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 따라서 본 발명의 범위는 전술한 실시예에 한정되지 않고 특허청구범위에 기재된 내용 및 그와 동등한 범위 내에 있는 다양한 실시 형태가 포함되도록 해석되어야 할 것이다.

부호의 설명

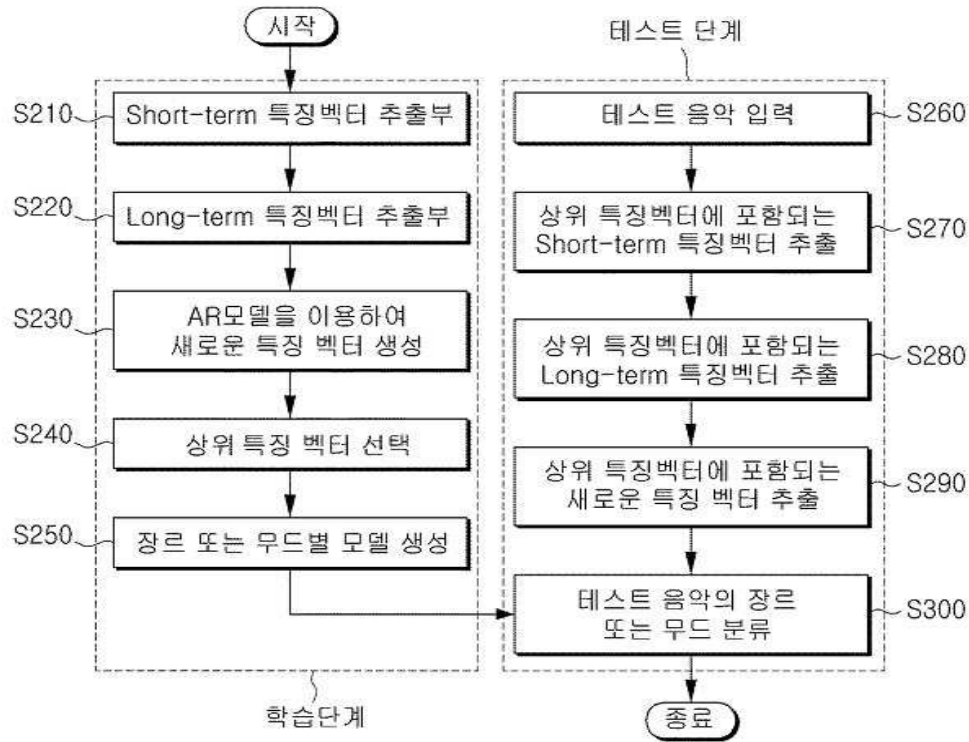
- [0121] 100: 음악 분류 장치, 110: 입력부,
120: Short-term 특징 추출부, 130: Long-term 특징 추출부,
140: AR 모델링부, 150: 특징 선택부,
160: 모델 생성부, 170: 음악 분류부,
180: 모델 생성부, 190: 음악 분류부

도면

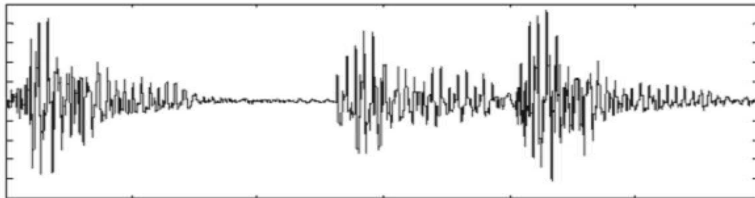
도면1



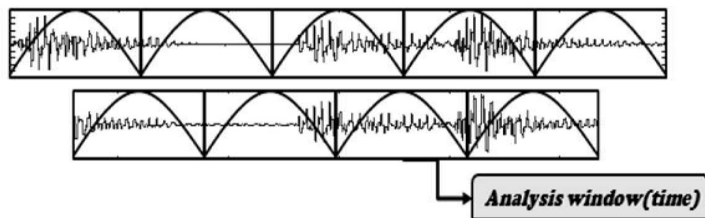
도면2



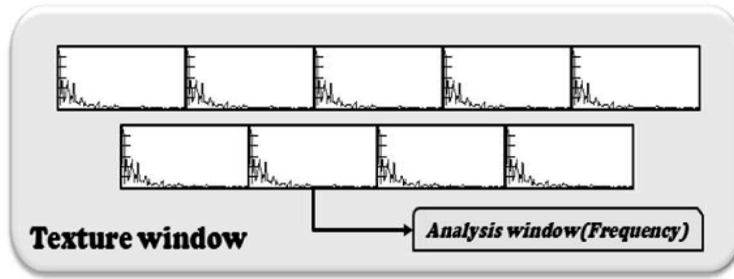
도면3



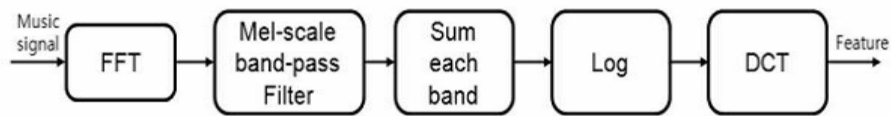
도면4



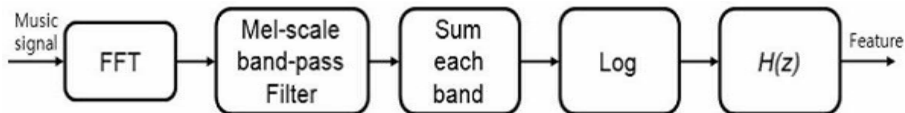
도면5



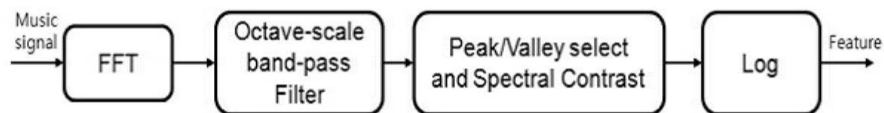
도면6



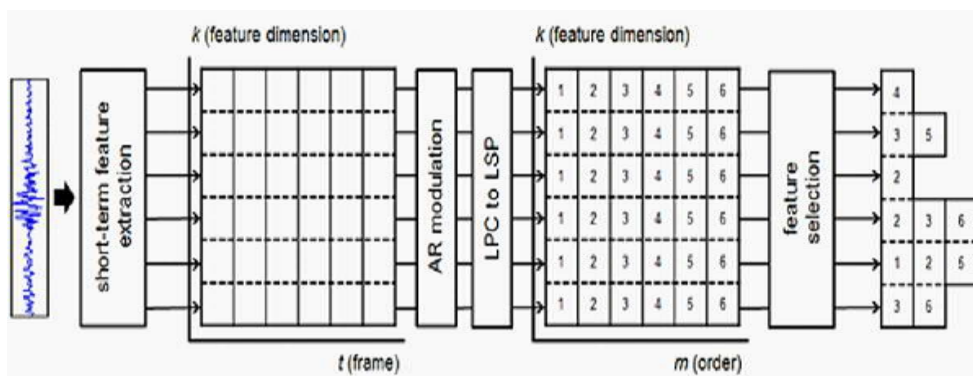
도면7



도면8



도면9



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 8의 7번째 줄

【변경전】

음악의 분류 장치

【변경후】

음악 분류 장치