

3D 입체음향 변환

5대 분야 Interactive VR/AR • Function Immersive A/V • 기술분야명 실감 오디오

Immersive AV

담당 센터 지능형영상처리 • 연구자 김제우

개념

3D 영상과 함께 실감 몰입형 멀티미디어 구현을 위한 3D 입체 음향 콘텐츠 변환 요소기술 및 인터랙티브 응용 서비스 기술

개발 내용

기술내용

3D 입체 음향 콘텐츠 변환을 위한 요소기술

- 2D 스테레오 오디오 음원(보컬, 타악기, 현악기 등) 객체 분리
- 단일 및 다중 객체(음성, 악기, 움직이는 객체, 사인파 등) 위치 및 방향 추정
- 인터랙티브 3D 오디오 렌더링

모바일 스마트 디바이스 실시간 3D 오디오 재생 S/W

- 안드로이드, iOS 플랫폼에서 3D 오디오 재생 변환 S/W 탑재

차별성

본 기술은 기존 스테레오 콘텐츠(음원)를 콘텐츠 내 각각의 오디오 음원(객체)들로 분리하고, 음원의 위치 및 특성을 반영하여 3D 렌더링 기술을 적용하여 입체 음향으로 변환

* JDX社(佛)서 영화 콘텐츠에 대한 객체오디오 서비스 사업화를 추진 중으로 현재 대부분 2D→3D 오디오 변환은 다채널 Up-mixing 기술을 채택

* 국내는 2008년 오디즌사에서 '객체 오디오 음원 및 서비스 사업'을 시행한바 있음

해외주요기관

INRIA 社(객체오디오 분리 기술), Apple 社(Audition/다채널 업·믹싱 기술)

개발 내용

2D-3D 콘텐츠 변환 솔루션 개념도 PC용 2D-3D 콘텐츠 변환 SW 모습 2D-3D 콘텐츠 변환 3D 렌더링 모습 스마트폰용 2D-3D 콘텐츠 변환 SW 모습

연구원 보유(개발) 핵심기술

오디오 객체 분리 개념도 오디오 객체 분리 알고리즘 블록도

연구원 보유(개발) 핵심기술

KETI 핵심기술

2 채널 음원에 대한 2.1ch, 5.1ch, 7.1ch, 10.2ch로 up-mixing 기술

- PCA(Principal Component Analysis)를 통한 스테레오 오디오의 Primary-Ambient 분리 기법 적용
- Non-negative matrix(NMF) 기반 모델링을 이용한 스테레오 음원 객체 분리 기술
- 선형 믹싱기반 음원 분리 기술 및 객체 음원 파라미터 추출 기술
- 오디오 객체 분리 기술 : 객체 분리도 10.7dB 성능, 음원 객체 품질 MOS 4.0 수준
- 객체 위치 및 방향 추정 정확도(85%), 음원의 자동 장르 분류 정확도(80%)

차별성

INRIA社(佛)의 객체 오디오 분리 기술(대외 논문자료)과 비교 시 동등수준의 객체 오디오 분리 기술 확보

관련기술 보유 IP

글로벌 모델 기반 오디오 객체분리 방법 및 시스템(국내, PCT/출원/2016)

분리된 객체 및 음원에서 목적음과 환경음을 인식하는 방법(국내/출원/2015)

오디오 소스 분리 방법 및 이를 적용한 오디오 시스템(국내/출원/2015)

음원 분리를 포함하는 음원 위치 추정 장치 및 방법(국내/출원/2015)

대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법 및 장치(국내/등록/2014)

Business Model

3D 렌더링 모듈

- 2D 스테레오 음원을 활용한 가상 3D 오디오 콘텐츠 변환 재생 솔루션

- 스피커 및 이어폰의 3D 오디오 재생 효과 기술, 사용자 인터랙티브 기반 3D 오디오 렌더링 변환 기술, 오디오 정보기반 3D 오디오 렌더링 자동 변환

- 디지털 TV 제조사, 스마트 기기 제조사

- 3D TV, 스마트 폰, 스마트 패드, 태블릿 PC 등 실감 미디어 기기

음원 객체 분리 및 위치 추정 S/W

- 음원의 객체를 소스별로 분리를 하고, 각 음원에 대한 정확한 위치를 추정하여 3D로 음원 장면을 재구성하는데 활용

- 음원 스트리밍 서비스 업체

- 3D virtual music player, 네이버 뮤직, 벅스 뮤직 등

사운드 레이더

- 최소한의 마이크로폰을 사용하여 소음원의 방향을 추정함으로써 음원 위치를 파악

- 자동차/건설 제조회사

- 자동차 제조사

관련 연구 분야

Core Technology	Related Technology
<p>4D 객체 복원 기반 실감 서비스 정보통신미디어/지능형영상처리 환경배</p>	<p>실감 오디오 정보통신/VRAR 강훈중</p>
<p>영상 편집 입체영상 제작 가상카메라 시스템 정보통신미디어/VRAR 강훈중</p>	<p>영상 스트리밍 고품질 360 VR 콘텐츠 영상 스트리밍 정보통신미디어/콘텐츠응용 박우철</p>
<p>입체영상 단말 초다시점 3D 디스플레이 단말기 정보통신미디어/콘텐츠응용 이경택</p>	<p>실감음향 단말 고해상도/원음급 오디오 단말 및 서비스 정보통신미디어/스마트미디어 이종실</p>



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년06월05일
(11) 등록번호 10-1864925
(24) 등록일자 2018년05월30일

(51) 국제특허분류(Int. Cl.)
G10L 19/008 (2014.01) G10L 17/04 (2013.01)
G10L 19/20 (2013.01) G10L 25/18 (2013.01)
(52) CPC특허분류
G10L 19/008 (2013.01)
G10L 17/04 (2013.01)
(21) 출원번호 10-2016-0014914
(22) 출원일자 2016년02월05일
심사청구일자 2016년07월04일
(65) 공개번호 10-2017-0093474
(43) 공개일자 2017년08월16일
(56) 선행기술조사문헌
KR101511553 B1
(뒷면에 계속)

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
조충상
경기도 성남시 분당구 장미로 55, 116동 805호
김제우
경기도 성남시 분당구 수내로 181, 309동 905호
(뒷면에 계속)
(74) 대리인
남충우

전체 청구항 수 : 총 6 항

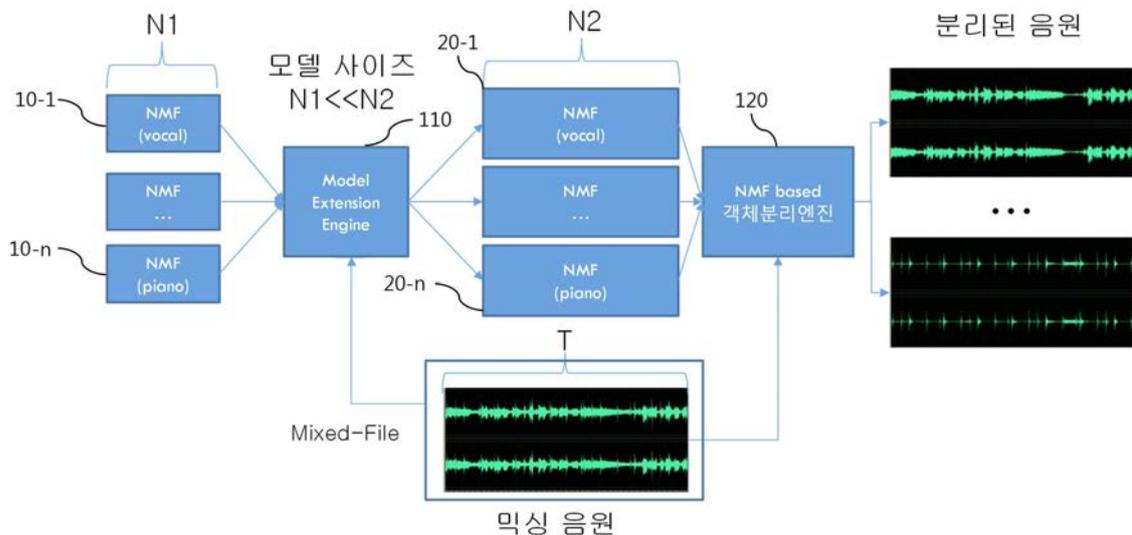
심사관 : 이남숙

(54) 발명의 명칭 글로벌 모델 기반 오디오 객체 분리 방법 및 시스템

(57) 요약

글로벌 모델 기반 오디오 객체 분리 방법 및 시스템이 제공된다. 본 발명의 실시예에 따른 오디오 분리 방법은, 음원의 분리에 사용되는 제1 모델을 자동으로 확장하여 제2 모델을 생성하고, 생성된 제2 모델을 이용하여 음원을 다수의 오디오 객체로 분리한다. 이에 의해, 작은 길이의 글로벌 NMF 모델을 확장하여 긴 길이의 NMF 모델을 자동 생성하여 오디오 객체 분리에 이용하는 것이 가능해져, 보다 간편하고 짧은 시간으로 모든 음원들에 대해 오디오 객체 분리가 가능해진다.

대표도 - 도1



(52) CPC특허분류

G10L 19/20 (2013.01)

G10L 25/18 (2013.01)

(72) 발명자

이영한

경기도 용인시 기흥구 예현로35번길 21, 105동 105호

이혜인

경기도 안양시 동안구 부림로 13, 603동 1402호

(56) 선행기술조사문헌

KR1020110023688 A

KR1020150142777 A

US20150205575 A1*

Brian King, et al. Optimal cost function and magnitude power for NMF-based speech separation and music interpolation. IEEE International Workshop on Machine Learning for Signal Processing. 2012.09.26.*

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호 1711026836

부처명 미래창조과학부

연구관리전문기관 정보통신기술진흥센터

연구사업명 SW컴퓨팅산업원천기술개발

연구과제명 2D 스테레오 콘텐츠를 3D 입체 음향 콘텐츠로 변환하기 위한 음원 객체 분리/위치 추정 및 3D 렌더링 소프트웨어 기술 개발

기여율 1/1

주관기관 (주)이머시스

연구기간 2013.05.01 ~ 2016.02.29

명세서

청구범위

청구항 1

음원의 분리에 공통적으로 사용되는 제1 NMF 모델(Non-Negative Matrix Factorization Model)의 W 행렬은 그대로 사용하고 H 행렬은 반복 나열하여 제2 NMF 모델을 생성하는 단계; 및

생성된 제2 NMF 모델을 이용하여, 상기 음원을 다수의 오디오 객체로 분리하는 단계;를 포함하는 것을 특징으로 하는 오디오 분리 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

청구항 1에 있어서,

상기 생성 단계는,

상기 H 행렬의 전부 또는 일부 단위로 반복 나열하여 확장하는 것을 특징으로 하는 오디오 분리 방법.

청구항 5

청구항 1에 있어서,

상기 생성 단계는,

상기 H 행렬의 일부를 선택하여 나열함으로써 확장하는 것을 특징으로 하는 오디오 분리 방법.

청구항 6

청구항 5에 있어서,

상기 생성 단계는,

상기 H 행렬의 일부를 랜덤하게 선택하는 것을 특징으로 하는 오디오 분리 방법.

청구항 7

청구항 5에 있어서,

상기 생성 단계는,

상기 음원에 대한 STFT(Short Term Fourier Transform) 연산 결과의 절대값을 산출하고, 산출된 결과와 유사도를 분석하여 가장 유사한 상기 H 행렬의 일부를 선택하는 것을 특징으로 하는 오디오 분리 방법.

청구항 8

음원의 분리에 공통적으로 사용되는 제1 NMF 모델(Non-Negative Matrix Factorization Model)의 W 행렬은 그대로 사용하고 H 행렬은 반복 나열하여 제2 NMF 모델을 생성하는 생성부; 및

생성된 제2 NMF 모델을 이용하여, 상기 음원을 다수의 오디오 객체로 분리하는 분리부;를 포함하는 것을 특징으로 하는 오디오 분리 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 오디오 처리 기술에 관한 것으로, 더욱 상세하게는 음원을 다수의 오디오 객체들로 분리하는 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 오디오 음원은 다수의 오디오 객체들, 이를 테면, 보컬, 드럼, 기타, 피아노 등으로 구성된다. 이러한 오디오 음원을 오디오 객체들로 분리하는 것이 가능하다.

[0003] 현재, 오디오 객체 분리에 있어 가장 많이 사용되는 기법들 중 하나는 NMF 모델(Non-Negative Matrix Factorization Model) 기반의 오디오 객체 분리 기법이다.

[0004] NMF 모델 기반으로 오디오 객체를 분리하기 위해서는, 분리하고자 하는 오디오 음원과 동일한 길이의 NMF 모델이 필요한데, 음원 마다 길이가 다르기 때문에, 이용되는 NMF 모델은 음원 마다 다르다.

[0005] 또한, 분리도를 높이기 위해, 분리하고자 하는 오디오 음원의 길이 외에 속성을 더 고려하여 오디오 엔지니어가 NMF 모델을 직접 설계하고 있는데, 매우 어렵고 장시간이 소요되는 작업이다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 음원 분리에 사용되는 모델을 간편하게 자동으로 생성하여 이용하는 글로벌 모델 기반 오디오 객체 분리 방법 및 시스템을 제공함에 있다.

과제의 해결 수단

[0007] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 오디오 분리 방법은, 음원의 분리에 사용되는 제1 모델을 자동으로 확장하여 제2 모델을 생성하는 단계; 및 생성된 제2 모델을 이용하여, 상기 음원을 다수의 오디오 객체로 분리하는 단계;를 포함한다.

[0008] 그리고, 본 발명의 일 실시예에 따른 오디오 분리 방법은, 상기 음원의 길이를 파악하는 단계;를 더 포함하고, 상기 길이를 참조로, 상기 제1 모델을 확장할 수 있다.

[0009] 또한, 상기 제1 모델은, 제1 NMF 모델(Non-Negative Matrix Factorization Model)이고, 상기 제2 모델은, 제2 NMF 모델이며, 상기 생성 단계는, 상기 제1 NMF 모델의 W 행렬을 반복 나열하여 확장할 수 있다.

[0010] 그리고, 상기 생성 단계는, 상기 W 행렬의 전부 또는 일부 단위로 반복 나열하여 확장할 수 있다.

[0011] 또한, 상기 생성 단계는, 상기 W 행렬의 일부를 선택하여 나열함으로써 확장할 수 있다.

[0012] 그리고, 상기 생성 단계는, 상기 W 행렬의 일부를 랜덤하게 선택할 수 있다.

[0013] 또한, 상기 생성 단계는, 상기 음원의 분석 결과를 기초로, 상기 W 행렬의 일부를 선택할 수 있다.

[0014] 한편, 본 발명의 다른 실시예에 따른, 오디오 분리 시스템은, 음원의 분리에 사용되는 제1 모델을 자동으로 확장하여 제2 모델을 생성하는 생성부; 및 생성된 제2 모델을 이용하여, 상기 음원을 다수의 오디오 객체로 분리

하는 분리부;를 포함한다.

발명의 효과

[0015] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 작은 길이의 글로벌 NMF 모델을 확장하여 긴 길이의 NMF 모델을 자동 생성하여 오디오 객체 분리에 이용하는 것이 가능해져, 보다 간편하고 짧은 시간으로 모든 음원들에 대해 오디오 객체 분리가 가능해진다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시예에 따른 오디오 객체 분리 시스템의 설명에 제공되는 도면,
 도 2는 NMF 모델의 설명에 제공되는 도면,
 도 3은 글로벌 NMF 모델의 확장에 대한 설명에 제공되는 도면,
 도 4는, 도 1에 도시된 NMF 모델 확장 엔진의 상세 설명에 제공되는 도면,
 도 5 내지 도 8은, H를 H'으로 확장/변환하는 방법의 설명에 제공되는 도면들, 그리고,
 도 9는 오디오 분석 모듈에 의한 인덱스 결정 방법의 상세 설명에 제공되는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0017] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0018] 도 1은 본 발명의 일 실시예에 따른 오디오 객체 분리 시스템의 설명에 제공되는 도면이다. 본 발명의 실시예에 따른 오디오 객체 분리 시스템은, 입력되는 길이 T의 음원을 오디오 객체들로 분리하기 위해 필요한 NMF 모델들을 그보다 작은 사이즈의 NMF 모델들을 확장하여 생성하는 시스템이다.

[0019] 이와 같은 기능을 수행하는 본 발명의 실시예에 따른 오디오 객체 분리 시스템은, 도 1에 도시된 바와 같이, NMF 모델 확장 엔진(110) 및 NMF 모델 기반 객체 분리 엔진(120)을 포함한다.

[0020] NMF 모델 확장 엔진(110)은 음원 분리에 사용되는 글로벌 NMF 모델들(10-1, ... 10-n)을 음원 길이에 따라 자동으로 확장하여, 음원 분리에 사용할 NMF 모델들(20-1, ... 20-n)을 생성한다.

[0021] NMF 모델 기반 객체 분리 엔진(120)은 NMF 모델 확장 엔진(110)에 의해 생성된 NMF 모델들(20-1, ... 20-n)을 이용하여, 음원을 다수의 오디오 객체들로 분리한다.

[0022] 글로벌 NMF 모델들(10-1, ... 10-n)은 오디오 객체(보컬, 드럼, 기타, 피아노) 별로 마련되어 있는 작은 길이의 NMF 모델들로, 입력되는 모든 음원들에 대해 공통적으로 사용된다.

[0023] NMF 모델은, 도 2에 도시된 바와 같이, '여러 오디오 객체들이 믹싱된 음원'(이하, '믹싱 음원'으로 표기)에 대한 STFT(Short Term Fourier Transform) 연산 결과로부터 $W(F \text{ by } k)$ 와 $H(k \text{ by } N)$ 을 설정함으로써 결정된다.

[0024] N은 STFT 연산에 사용된 윈도우 사이즈, F는 STFT 연산에서 frequency bin의 개수, k는 STFT 연산에 적용된 차원이다. 믹싱 음원의 길이 T가 영향을 미치는 파라미터는 "N"이다.

[0025] 이에, 글로벌 NMF 모델들(10-1, ... 10-n)로부터 확장된 NMF 모델들(20-1, ... 20-n)을 생성함에 있어, 도 3에 도시된 바와 같이 글로벌 NMF 모델들(10-1, ... 10-n)의 $H(k \text{ by } N_1)$ 행렬을 길이 T의 믹싱 음원에 대해 요구되는 $H'(k \text{ by } N_2)$ 행렬로 확장한다.

[0026] 그리고, 글로벌 NMF 모델들(10-1, ... 10-n)의 W는 NMF 모델들(20-1, ... 20-n)을 생성함에 있어 그대로 사용한다. 즉, 글로벌 NMF 모델들(10-1, ... 10-n)의 W와 NMF 모델들(20-1, ... 20-n)의 W는 동일하게 구현한다.

[0027] 도 4는, 도 1에 도시된 NMF 모델 확장 엔진(110)의 상세 설명에 제공되는 도면이다. 도 4에는 NMF 모델 확장 엔진(110)이 글로벌 NMF 모델들(10-1, ... 10-n)을 NMF 모델들(20-1, ... 20-n)로 변환하는 과정이 나타나 있다.

[0028] 구체적으로, 도 4에는, NMF 모델 확장 엔진(110)은 믹싱 음원의 길이 T를 파악하고, 파악된 길이 T를 기초로 글로벌 NMF 모델들(10-1, ... 10-n)을 구성하는 작은 사이즈의 H를 긴 사이즈의 H'로 변환하여, NMF 모델들(20-1, ... 20-n)을 생성하는 과정이 나타나 있다.

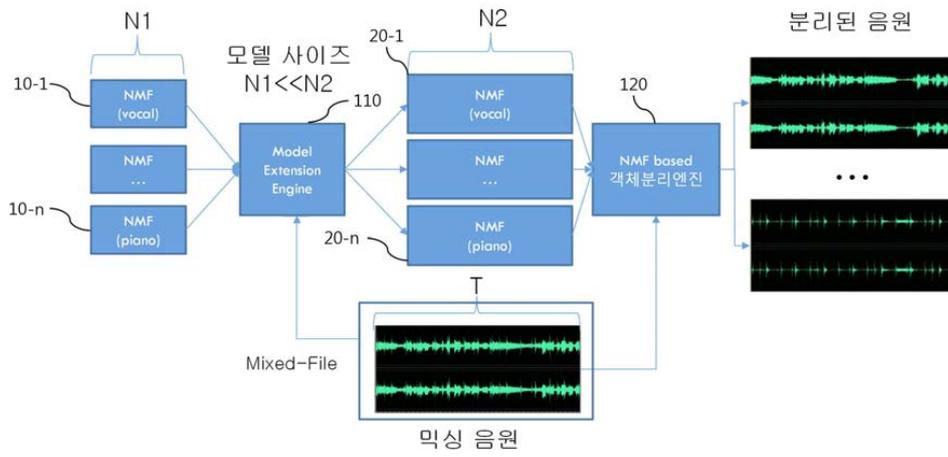
- [0029] H를 H'으로 확장하여 변환하는 방법은 매우 다양하며, 이하에서 상세히 설명한다. H를 H'으로 변환함에 있어서는, H 자체만을 이용할 수도 있지만, 오디오 분석 모듈(115)에 의한 믹싱 음원의 분석 결과를 이용할 수도 있다.
- [0030] 도 5는 H를 H'으로 확장/변환하는 방법을 나타내었다. 도 5에 제시된 확장/변환 방법은, 열 길이가 N1인 H를 반복 나열하되, 요구되는 N2를 초과하는 부분은 삭제하여, H'를 생성하는 방법이다.
- [0031] 도 6은 H를 H'으로 확장/변환하는 다른 방법을 나타내었다. 도 6에 제시된 확장/변환 방법은, 열 길이가 N1인 H의 열들을 열 단위로 반복 나열하여, H'을 생성하는 방법이다. 이때, 요구되는 N2를 맞추기 위해 반복 횟수는 열 마다 다르게 설정할 수 있다.
- [0032] 도 7은 H를 H'으로 확장/변환하는 또 다른 방법을 나타내었다. 도 7에 제시된 확장/변환 방법은, 열 길이가 N1인 H의 열들 중 하나를 랜덤하게 선택하여 나열하는 것을 반복하여 H'를 생성하는 방법이다.
- [0033] 도 8은 H를 H'으로 확장/변환하는 또 다른 방법을 나타내었다. 도 8에 제시된 확장/변환 방법은, H의 열들 중 하나를 선택하여 나열하는 것을 반복하여 H'를 생성한다는 점에서 도 7에 제시된 방법과 동일하다.
- [0034] 하지만, 도 8에 제시된 방법에서는 H'에 나열할 H의 열들 중 하나를 랜덤하게 선택하는 것이 아니라, 오디오 분석 모듈(115)에 의해 믹싱 음원의 분석 결과를 기초로 결정된 인덱스에 따라 선택한다는 점에서, 도 7에 제시된 방법과 차이가 있다.
- [0035] 도 9는 오디오 분석 모듈(115)에 의한 인덱스 결정 방법의 상세 설명에 제공되는 도면이다. 도 9에 도시된 바와 같이, 오디오 분석 모듈(115)은 믹싱 음원에 대한 SIFT 연산 결과의 절대값을 산출하고, 산출된 결과에서 윈도우를 이동시키면서 유사도 분석을 통해 가장 유사한 H의 열을 선택하는 것을 반복하여 인덱스들을 생성한다.
- [0036] 지금까지, 글로벌 모델 기반 오디오 객체 분리 방법 및 시스템에 대해 바람직한 실시예들을 들어 상세히 설명하였다.
- [0037] 위 실시예들에서는 NMF 모델들을 이용한 오디오 객체 분리를 상정하였는데 예시를 위한 것이다. NMF 모델이 아닌 그로부터 변형된 모델 또는 그와 다른 종류의 모델을 적용하는 경우에도, 본 발명의 기술적 사상이 적용될 수 있음은 물론이다.
- [0038] 또한, 위 실시예들에서, 오디오 객체들로 언급한 보컬, 드럼, 기타, 피아노 역시 예시적인 것에 불과하다. 이보다 더 다양한 오디오 객체들로 음원을 분리하는 경우에도 본 발명의 기술적 사상이 적용가능하다.
- [0039] 본 발명의 실시예들에서 제시한 오디오 객체 분리 방법 및 시스템은, 오디오 효과, 콘텐츠 제작, 감시 시스템 등과 같은 분야는 물론, 음성 분리나 그 밖의 다른 종류의 음원 분리가 필요한 분야에 적용될 수 있다.
- [0040] 또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

부호의 설명

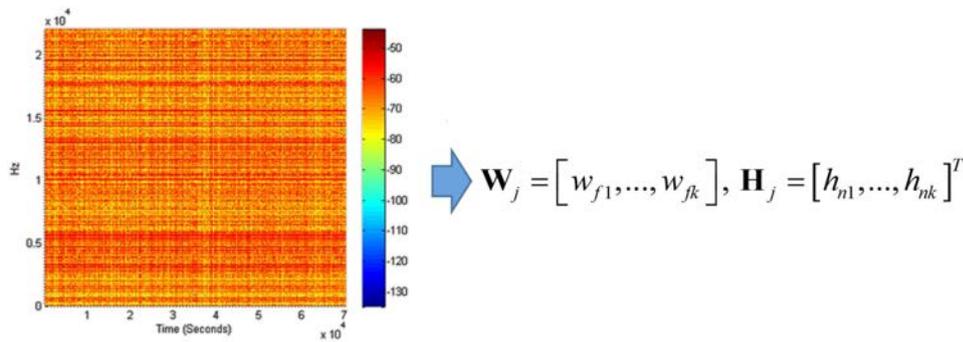
- [0041] 10-1, ... 10-n : 글로벌 NMF 모델
- 20-1, ... 20-n : NMF 모델
- 110 : NMF 모델 확장 엔진
- 120 : NMF 모델 기반 객체 분리 엔진

도면

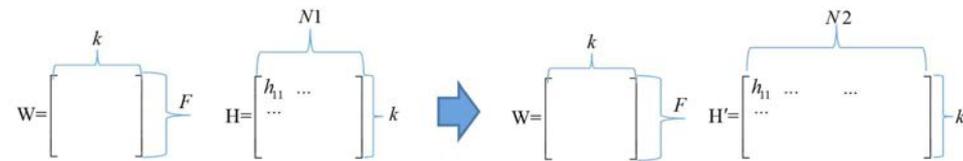
도면1



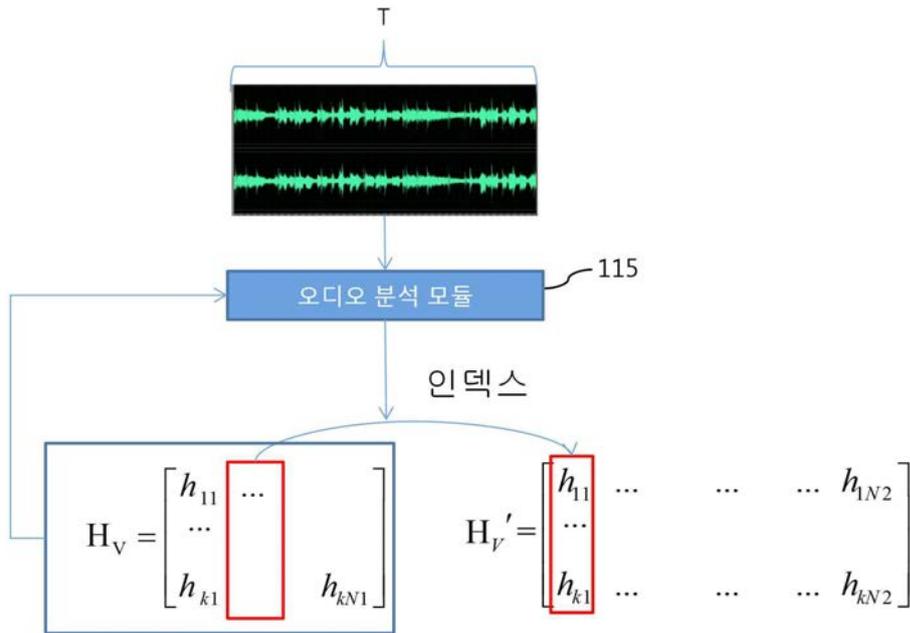
도면2



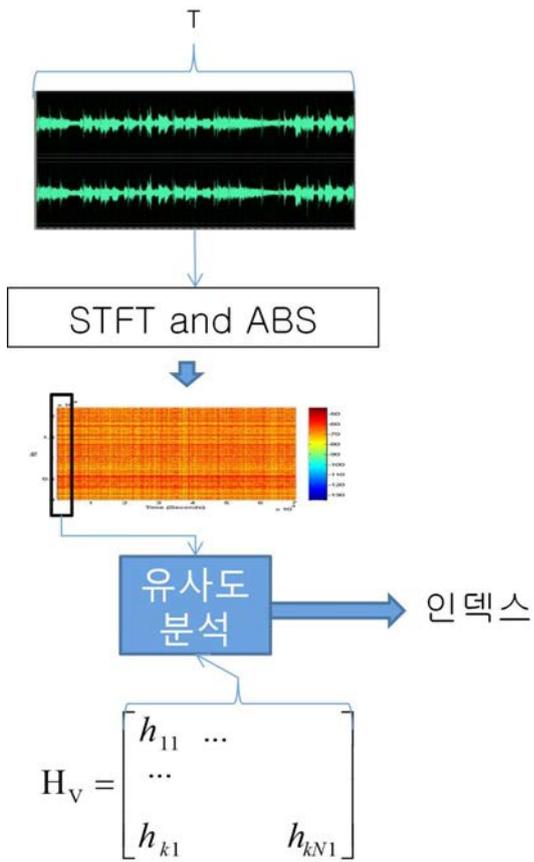
도면3



도면8



도면9





(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2015년04월07일
(11) 등록번호 10-1509649
(24) 등록일자 2015년04월01일

(51) 국제특허분류(Int. Cl.)
H04S 3/00 (2006.01) H04S 7/00 (2006.01)
(21) 출원번호 10-2014-0023339
(22) 출원일자 2014년02월27일
심사청구일자 2014년02월27일
(56) 선행기술조사문헌
JP2013507048 A
KR1020090050380 A

(73) 특허권자
전자부품연구원
경기도 성남시 분당구 새나리로 25 (야탑동)
(72) 발명자
송재종
경기도 수원시 장안구 장안로 199, 302동 701호
(정자동, 동신아파트)
박성주
경기도 용인시 기흥구 죽현로 8-13, 310동 501호
(보정동, 동원로얄듀크아파트)
(뒷면에 계속)
(74) 대리인
특허법인다래

전체 청구항 수 : 총 5 항

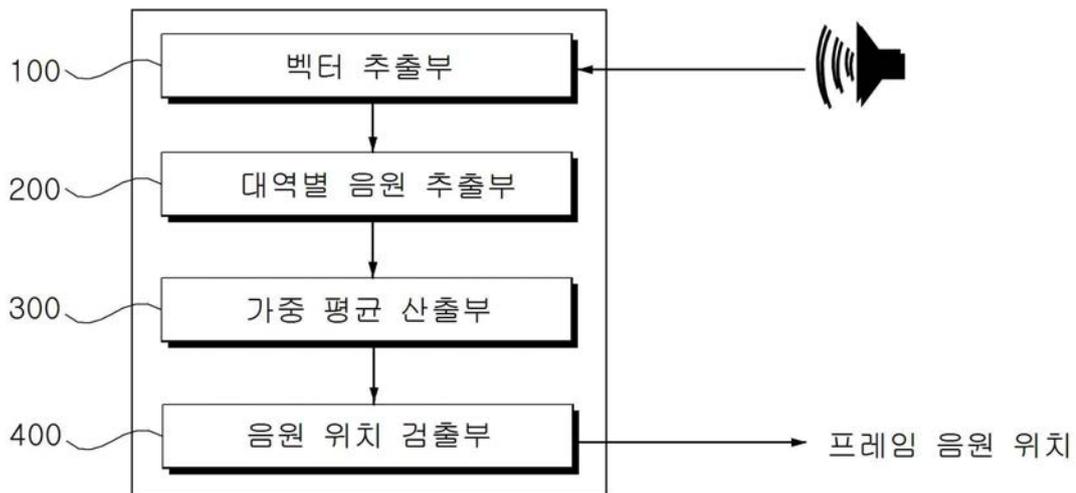
심사관 : 권영학

(54) 발명의 명칭 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법 및 장치

(57) 요약

본 발명은 보컬을 포함하여 여러 악기 음원이 혼합된 음악 신호에서, 패닝 기법을 통해 방향성을 가지고 있는 음원 위치의 검출 방법 및 장치에 관한 것으로, 본 발명에 따른 가중 평균을 이용한 음원 위치 검출 방법은 입력된 스테레오 신호를 PCA(Principle Component Analysis) 기법으로 주성분과 주변성분으로 분리하여 상기 스테레오 (뒷면에 계속)

대표도 - 도1



오 신호의 주성분에 대한 고유벡터를 추출하는 단계; 상기 추출된 고유벡터를 이용하여 상기 스테레오 신호의 대역별 음원 위치를 추출하는 단계; 및 상기 스테레오 신호의 대역별 에너지를 이용하여 상기 추출된 음원 위치의 정확도에 따른 가중 평균을 산출하는 단계를 포함한다. 본 발명에 따르면 음원 객체의 위치 추정 정확도를 향상할 수 있기 때문에 다음원 음악 신호에서 음원 객체별 위치를 보다 정확하게 구현하여 주며, 종래 기술에 사용하는 균일 평균 방식을 사용하지 않고 추정 SNR과 밴드별 에너지를 이용하여 효과적으로 음원 객체 위치 추정 정확도를 향상할 수 있다. 또한, 다양한 종류의 음악에 대하여 음원 객체의 위치를 추정함에 따라 음원 객체별 재배치가 가능하다.

(72) 발명자

양창모

경기도 고양시 일산서구 강선로 137, 1604동 1101호 (일산동, 후곡마을16단지)

박호중

서울특별시 서초구 명달로4길 30, 501-905

구자성

서울특별시 강북구 노해로32길 13-7, 302

한택진

경기도 의정부시 동일로466번길 3, 108-1206

이 발명을 지원한 국가연구개발사업

과제고유번호 10044569

부처명 산업자원부

연구관리전문기관 한국산업기술평가원

연구사업명 산업융합원천기술개발사업

연구과제명 2D 스테레오 콘텐츠를 3D 입체 음향 콘텐츠로 변환하기 위한 음원 객체 분리/위치 추정 및 3D 렌더링 소프트웨어 기술 개발

기여율 1/1

주관기관 ㈜이머시스

연구기간 2013.05.01 ~ 2014.04.30

명세서

청구범위

청구항 1

입력된 스테레오 신호를 PCA(Principle Component Analysis) 기법으로 주성분과 주변성분으로 분리하여 상기 스테레오 신호의 주성분에 대한 고유벡터를 추출하는 단계;

상기 추출된 고유벡터를 이용하여 상기 스테레오 신호의 대역별 음원 위치를 추출하는 단계;

상기 스테레오 신호의 대역별 에너지에 따라 가중 평균의 산출 여부를 결정하는 단계; 및

상기 가중 평균의 산출이 결정된 대역에 대하여 SNR을 이용한 가중 평균을 산출하는 단계를 포함하는 것을 특징으로 하는 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법

청구항 2

제 1 항에 있어서,

상기 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법은,

상기 산출된 가중 평균을 이용하여 상기 스테레오 신호의 프레임 음원 위치를 검출하는 단계를 더 포함하는 것을 특징으로 하는 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법

청구항 3

삭제

청구항 4

제 2 항에 있어서,

상기 SNR을 이용한 가중 평균을 산출하는 단계는 추정 원본 신호와 상기 PCA 기법으로 분리된 주변성분 신호를 이용하여 추정 SNR을 산출하고,

상기 산출된 추정 SNR을 이용하여 상기 프레임 음원 위치의 검출을 위한 가중 평균을 산출하는 것을 특징으로 하는 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법

청구항 5

입력된 스테레오 신호를 PCA(Principle Component Analysis) 기법으로 주성분과 주변성분으로 분리하여 상기 스테레오 신호의 주성분에 대한 고유벡터를 추출하는 벡터 추출부;

상기 추출된 고유벡터를 이용하여 상기 스테레오 신호의 대역별 음원 위치를 추출하는 대역별 음원 위치 추출부;

상기 스테레오 신호의 대역별 에너지를 이용하여 상기 추출된 음원 위치의 정확도에 따른 가중 평균을 산출하는 가중 평균 산출부; 및

상기 산출된 가중 평균을 이용하여 상기 스테레오 신호의 프레임 음원 위치를 검출하는 프레임 음원 위치 검출부를 포함하고,

상기 가중 평균 산출부는 상기 대역별 에너지에 따라 가중 평균의 산출 여부를 결정하는 에너지 가중 평균 산출부; 및

상기 가중 평균의 산출이 결정된 대역에 대하여 SNR을 이용한 가중 평균을 산출하는 SNR 가중 평균 산출부를 포함하는 것을 특징으로 하는 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 장치

청구항 6

삭제

청구항 7

제 5 항에 있어서,
 상기 SNR 가중 평균 산출부는 추정 원본 신호와 상기 PCA 기법으로 분리된 주변성분 신호를 이용하여 추정 SNR 을 산출하고,
 상기 산출된 추정 SNR을 이용하여 상기 프레임 음원 위치의 검출을 위한 가중 평균을 산출하는 것을 특징으로 하는 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 장치

발명의 설명

기술 분야

[0001] 본 발명은 보컬을 포함하여 여러 악기 음원이 혼합된 음악 신호에서, 패닝 기법을 통해 방향성을 가지고 있는 음원 위치의 검출 방법 및 장치에 관한 것이다.

배경 기술

[0002] 일반적으로 스테레오 음원에서 위치 추정은 방향성을 가지고 있는 주성분(primary)의 음원 위치를 추정한다. 하지만 음원에 주성분과 방향성이 없는 주변 성분이 더해지면 주성분의 음원 위치가 흔들린다.

[0003] 따라서, 주성분으로 이루어진 원본 신호에 주변 성분이 더해지면 주성분의 음원 위치를 정확하게 추정할 수 없으므로 PCA를 이용하여 주성분을 분리하고 주성분의 음원 위치를 추정한다. 밴드별로 음원 위치가 추정되면 추정된 음원 위치를 하나의 프레임 위치로 변환하여, 실제 음원 위치를 검출하게 된다.

[0004] 이때 밴드별로 추정된 음원 위치의 정확도는 서로 다르므로 균일 평균을 이용하여 프레임 위치를 추정하는 경우 정확도가 감소하는 문제점이 있다. 즉, 주변 성분은 낮은 주파수에 더 많은 에너지를 가지고 있어, 낮은 밴드의 추정 음원 위치의 정확도를 감소시킨다. 주성분과 주변 성분 모두 높은 주파수로 갈수록 에너지가 감소하기 때문이다. 에너지가 낮은 구간에서 추정된 음원 위치는 정확도가 떨어지게 되고 따라서, 정확도를 고려하지 않고 균일 평균을 이용하여 프레임 위치를 추정하는 경우 검출된 음원 위치의 정확도가 떨어지게 된다.

[0005] 즉, 종래 기술은 여러 음원을 포함하는 음악 신호에서 정확한 음원 객체의 위치의 정확도를 효율적으로 추정하지 못한다.

발명의 내용

해결하려는 과제

[0006] 본 발명의 목적은 효율적이고 정확하게 음원 객체 위치 추정의 정확도의 향상 장치 및 방법을 제공함에 있다.

[0007] 보다 상세하게는 본 발명의 다른 목적은 음원 위치의 추정에 있어 추정된 음원 위치 중 정확도가 높은 구간을 판단하여 정확도가 높은 구간은 PCA를 이용한 추정 SNR과 밴드별 에너지를 이용하여 구할 수 있으며, 정확도에 따른 가중 평균(weighted average)을 통하여 좀 더 정밀하게 음원의 위치를 추정하는 것을 목적으로 한다.

과제의 해결 수단

[0008] 상기 기술적 과제를 해결하기 위한 본 실시예에 따른 가중 평균을 이용한 음원 위치 검출 방법은 입력된 스테레오 신호를 PCA(Principle Component Analysis) 기법으로 주성분과 주변성분으로 분리하여 상기 스테레오 신호의 주성분에 대한 고유벡터를 추출하는 단계; 상기 추출된 고유벡터를 이용하여 상기 스테레오 신호의 대역별 음원 위치를 추출하는 단계; 및 상기 스테레오 신호의 대역별 에너지를 이용하여 상기 추출된 음원 위치의 정확도에 따른 가중 평균을 산출하는 단계를 포함한다.

[0009] 상기 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법은, 상기 산출된 가중 평균을 이용하여 상기 스테레오 신호의 프레임 음원 위치를 검출하는 단계를 더 포함하는 것이 바람직하다.

[0010] 상기 가중 평균을 산출하는 단계는 상기 대역별 에너지에 따라 가중 평균의 산출 여부를 결정하는 단계; 및 상

기 가중 평균의 산출이 결정된 대역에 대하여 SNR을 이용한 가중 평균을 산출하는 단계를 포함한다.

[0011] 상기 SNR을 이용한 가중 평균을 산출하는 단계는 추정 원본 신호와 상기 PCA 기법으로 분리된 주변성분 신호를 이용하여 추정 SNR을 산출하고, 상기 산출된 추정 SNR을 이용하여 상기 프레임 음원 위치의 검출을 위한 가중 평균을 산출하는 것이 바람직하다.

[0012] 상기 기술적 과제를 해결하기 위한 본 실시예에 따른 가중 평균을 이용한 음원 위치 검출 장치는 입력된 스테레오 신호를 PCA(Principle Component Analysis) 기법으로 주성분과 주변성분으로 분리하여 상기 스테레오 신호의 주성분에 대한 고유벡터를 추출하는 벡터 추출부; 상기 추출된 고유벡터를 이용하여 상기 스테레오 신호의 대역별 음원 위치를 추출하는 대역별 음원 위치 추출부; 상기 스테레오 신호의 대역별 에너지를 이용하여 상기 추출된 음원 위치의 정확도에 따른 가중 평균을 산출하는 가중 평균 산출부; 및 상기 산출된 가중 평균을 이용하여 상기 스테레오 신호의 프레임 음원 위치를 검출하는 프레임 음원 위치 검출부를 포함한다.

발명의 효과

[0013] 본 발명에 따르면 음원 객체의 위치 추정 정확도를 향상할 수 있기 때문에 다음원 음악 신호에서 음원 객체별 위치를 보다 정확하게 구현하여 주며, 종래 기술에 사용하는 균일 평균 방식을 사용하지 않고 추정 SNR과 밴드별 에너지를 이용하여 효과적으로 음원 객체 위치 추정 정확도를 향상할 수 있다.

[0014] 또한, 다양한 종류의 음악에 대하여 음원 객체의 위치를 추정함에 따라 음원 객체별 재 배치가 가능하다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 일 실시예에 따른 가중 평균을 이용한 음원 위치 검출 장치를 나타내는 블록도이다.

도 2는 본 발명의 일 실시예에 따른 가중 평균을 이용한 음원 위치 검출 방법을 나타내는 흐름도이다.

도 3은 본 발명의 일 실시예에 따른 가중 평균을 이용한 음원 위치 검출 방법의 SNR에 따른 추정 음원 위치 오차를 나타내는 예시도이다.

도 4는 본 발명의 일 실시예에 따른 가중 평균을 이용한 음원 위치 검출 방법의 SNR에 따른 가중치를 나타내는 예시도이다.

도 5는 본 발명의 일 실시예에 따른 가중 평균을 이용한 음원 위치 검출 방법의 에너지에 따른 추정 음원 위치 오차를 나타내는 예시도이다.

발명을 실시하기 위한 구체적인 내용

[0016] 이하의 내용은 단지 발명의 원리를 예시한다. 그러므로 당업자는 비록 본 명세서에 명확히 설명되거나 도시되지 않았지만 발명의 원리를 구현하고 발명의 개념과 범위에 포함된 다양한 장치를 발명할 수 있는 것이다. 또한, 본 명세서에 열거된 모든 조건부 용어 및 실시예들은 원칙적으로, 발명의 개념이 이해되도록 하기 위한 목적으로만 명백히 의도되고, 이와같이 특별히 열거된 실시예들 및 상태들에 제한적이지 않는 것으로 이해되어야 한다.

[0017] 상술한 목적, 특징 및 장점은 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이며, 그에 따라 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다.

[0018] 또한, 발명을 설명함에 있어서 발명과 관련된 공지 기술에 대한 구체적인 설명이 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에 그 상세한 설명을 생략하기로 한다. 이하에는 첨부한 도면을 참조하여 본 발명의 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법 및 장치에 대하여 설명한다.

[0019] 먼저 도 1을 참조하면, 도 1은 본 발명의 일 실시예에 따른 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 장치를 나타내는 블록도이다.

[0020] 본 실시예에 따른 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 장치는 벡터 추출부(100), 대역별 음원 추출부(200), 가중 평균 산출부(300), 음원 위치 검출부를 포함한다.

[0021] 본 실시예에서 벡터 추출부(100)는 입력된 스테레오 신호를 PCA(Principle Component Analysis) 기법으로 주성분과 주변성분으로 분리하여 스테레오 신호의 주성분에 대한 고유벡터를 추출한다.

- [0022] 고유벡터가 추출되면 대역별 음원 추출부(200)는 추출된 고유벡터를 이용하여 스테레오 신호의 대역별 음원 위치를 추출한다.
- [0023] 다음, 가중 평균 산출부(300)는 스테레오 신호의 대역별 에너지를 이용하여 추출된 음원 위치의 정확도에 따른 가중 평균을 산출한다.
- [0024] 음원 위치 검출부는 산출된 가중 평균을 이용하여 스테레오 신호의 프레임 음원 위치를 검출하게 된다.
- [0025] 또한, 가중 평균 산출부(300)는 대역별 에너지에 따라 가중 평균의 산출 여부를 결정하는 에너지 가중 평균 산출부(300)와, 가중 평균의 산출이 결정된 대역에 대하여 SNR을 이용한 가중 평균을 산출하는 SNR 가중 평균 산출부(300)를 더 포함하고, 본 실시예에서 SNR 가중 평균 산출부(300)는 추정 원본 신호와 PCA 기법으로 분리된 주변성분 신호를 이용하여 추정 SNR을 산출하고, 산출된 추정 SNR을 이용하여 프레임 음원 위치의 검출을 위한 가중 평균을 산출한다.
- [0026] 다음 음원 위치 검출부가 산출된 가중 평균을 이용하여 스테레오 신호의 프레임 음원 위치를 검출하게 된다.
- [0027] 이하 도 2를 참조하여 본 발명에 따른 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 장치의 동작에 대하여 보다 상세히 설명한다.
- [0028] 도 2는 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법을 나타내는 흐름도이다.
- [0029] 도 2를 참조하면, 본 실시예에 따른 대역별 정확도에 따른 가중 평균을 이용한 음원 위치 검출 방법은 벡터 추출 단계(S100), 대역별 음원 추출 단계(S200), 가중 평균 산출 단계(S300), 음원 위치 검출 단계(S400)를 포함한다.
- [0030] 상술한 바와 같이 벡터 추출 단계(S100)는 벡터 추출부가 PCA(Principle Component Analysis) 기법으로 주성분과 주변성분으로 분리하여 스테레오 신호의 주성분에 대한 고유벡터를 추출한다.
- [0031] 구체적으로, 고유벡터의 추출을 위하여 먼저 먼저 입력된 스테레오 신호의 프레임별 N-point DFT(Discrete Fourier Transform)를 수행한 후 주파수 영역에서의 공분산을 구한다. 구해진 공분산을 이용하여 PCA 기법으로 주성분과 주변 성분을 분리한 후 좌, 우 신호의 고유벡터를 추정한다.
- [0032] 다음 대역별 음원 추출 단계(S200)는 대역별 음원 추출부(200)가, 추정된 고유벡터를 이용해 패닝 계인을 대역별로 추정하고 패닝 계인을 음원 위치로 변환하여 스테레오 신호의 대역별 음원 위치를 추출한다.
- [0033] 이때 대역별로 추정된 음원 위치의 정확도는 상술한 바와 같이 서로 다르므로 균일 평균(uniform average)으로 음원 위치를 결정하면 정확한 위치를 찾을 수 없으므로, 추정된 음원 위치 중 정확도가 높은 구간을 판단하고, 이에 따라 가중치를 얻고 프레임에 대한 가중 평균을 산출하여야 한다.
- [0034] 따라서, 가중 평균 산출 단계(S300)는 가중 평균 산출부(300)가 스테레오 신호의 대역별 에너지를 이용하여 추출된 음원 위치의 정확도에 따른 가중 평균을 산출한다. 이하 본 실시예에 따른 가중 평균의 산출방법에 대하여 상세히 설명한다.
- [0035] 먼저 본 실시예에서 가중 평균 산출 단계(S300)는 대역별 에너지를 추출할 수 있다. 즉, 대역별 에너지에 따라 가중 평균의 산출 여부를 결정하고 가중 평균의 산출이 결정된 대역에 대하여 SNR을 이용한 가중 평균을 산출한다.
- [0036] 대역별 에너지가 너무 적은 구간에 대해서는 오차가 커지고 정확도가 떨어지는 문제가 발생하는 바, 임의의 에너지 이하의 구간에 대하여 추출된 음원 위치는 가중 평균을 통한 프레임 위치 검출시 제외하거나, 매우 작은 가중치를 설정하여 영향을 최소화 할 수 있다.
- [0037] 다음, 대역별 에너지가 가중 평균의 산출에 적합한 수준 이상인 경우 본 실시예에서 가중 평균 산출 단계(S300)는 PCA를 이용한 추정 SNR 가중 평균을 산출 한다.
- [0038] 추정 SNR은 PCA로 분리한 주변 성분 신호와 입력신호를 이용하여 추정한다. 추정 SNR이 높을수록 추정치의 정확도가 높아지며 에너지가 클수록 추정치의 정확도가 높아진다. 학습을 통해 미리 얻은 SNR-weight 곡선을 이용하여 가중 평균의 가중치를 산출한다.
- [0039] 즉, 낮은 밴드에서는 주변 성분이 많이 더해져 있어 SNR을 이용하여 가중치를 추정할 수 있다. SNR은 입력 신호와 PCA로 분리한 주변 신호를 이용하여 추정하고 밴드별 SNR을 구하기 위해 주파수 도메인에서 동작한다.

[0040] [수학식 1]

$$SNR_{dB} = 10 \log_{10} \left(\frac{P_{original}}{P_{noise}} \right)$$

[0041]

[0042] 수학식 1을 참조하면 본 실시예에서 SNR은 dB스케일로 추출될 수 있다. 수학식 1에서 $P_{original}$ 은 원본 신호에 대한 파워, P_{noise} 는 노이즈 신호에 대한 파워를 의미한다.

[0043] 즉, SNR은 원본 신호와 잡음에 대한 파워의 비로 나타난다. 주성분으로 이루어진 신호를 원본 신호, 주변 성분을 노이즈라 가정한다. 다만, 음원 위치 추정 기술에서 원본 신호에 대한 정보를 알 수 없으므로 추정 SNR을 사용한다. 이때 원본 신호의 추정은 입력 신호의 주성분과 주변 성분을 PCA를 이용하여 분리하고 수학식 2를 통해 추정한다.

[0044] [수학식 2]

$$\begin{aligned} X_{input} &= X_{primary} + X_{ambient} \\ \hat{X}_{original} &= X_{input} - \hat{X}_{ambient} \end{aligned}$$

[0045]

[0046] 위 식에서 X_{input} 은 입력신호이고 $X_{primary}$ 와 $X_{ambient}$ 는 주성분과 주변 성분이다. 그리고 $\hat{X}_{original}$ 은 입력신호와 주변 성분을 이용해 추정한 원본신호이고 $\hat{X}_{ambient}$ 는 PCA를 이용하여 분리한 주변성분이다. 입력 신호 X_{input} 에 주성분과 주변성분이 있으므로 입력 신호에서 추출한 $\hat{X}_{ambient}$ 을 입력신호와 빼면 추정 원본신호를 구할 수 있다. 추정 원본 신호와 입력신호에서 분리한 주변 성분을 이용해 추정 SNR을 구한다.

[0047] 즉, 도 3을 참조하면 도 3은 SNR에 따른 추정 음원 위치 오차를 나타내는 예시도로서, 오차가 0에 가까울수록 정확도가 높은 음원 위치를 추정한 것이다. 도 3에서 SNR이 감소하면 오차는 커지고 SNR이 상승하면 오차는 감소한다. 따라서 SNR값에 대응하는 가중치는 도 4에 따른 그래프를 이용하여 추정할 수 있다.

[0048] 도 4는 SNR에 따른 가중치를 나타내는 예시도로서, 도 3을 training 데이터로 사용하여 추정하여 산출된 결과이다. 즉, 도 4에 나타난 그래프를 이용하여 추정 SNR에 따른 가중치를 구할 수 있고 가중치를 구하기 위한 식은 다음과 같다.

[0049] [수학식 3]

$$weight = \sqrt{1 - \frac{1}{(1 + (SNR_{estimated}/60)^{10})}}$$

[0050]

[0051] 위 식에서 $weight$ 는 가중치, $SNR_{estimate}$ 은 추정 SNR이다. 위 식에 추정 SNR을 대입하여 SNR에 따른 가중치를 얻고 프레임에 대한 가중 평균을 얻을 수 있다.

[0052] 이하, 또 다른 가중치 판별 방법으로서 높은 대역에 대한 가중 평균 산출 방법에 대하여 설명한다.

[0053] 즉, 높은 밴드에서의 정확도 감소는 밴드별 에너지 감소 때문에 발생한다. 에너지가 작은 구간에서는 방향성을 찾기 어렵기 때문이다. 따라서 에너지가 작은 구간에서는 음원 위치 추정을 할 수 없다.

[0054] 본 실시예에서는 SNR 가중 평균 산출과 병행하여 에너지 가중 평균 산출을 수행할 수 있다. 이때 SNR 가중 평균과 다르게 양자택일로 진행된다. 즉, 특정 밴드의 에너지가 일정 값 이하로 떨어지면 그 밴드의 위치 추정값은 프레임 평균을 추출하는 데 사용하지 않는다.

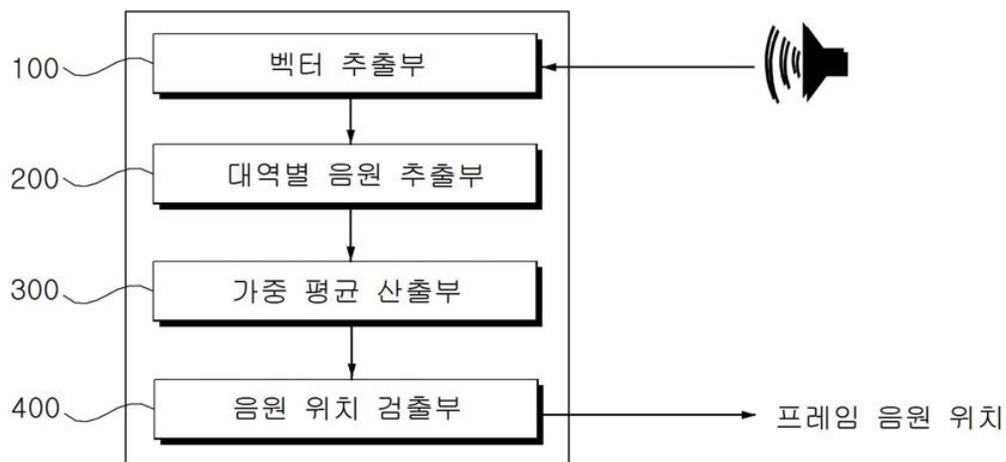
[0055] 도 5를 참조하면, 도 5는 에너지에 따른 추정 음원 위치 오차를 나타내는 예시도로서, 오차 값은 각도이고 0에서 가장 작은 오차를 나타낸다. (a)는 에너지 전대역에 대한 그래프를 그린 것이고 (b)는 (a) 그래프에서 에너지 0 ~ 0.1 사이의 부분을 확대하여 그린 것이다. 에너지와 작을 때 정확도가 함께 떨어지고 특히 밴드 에너지 값이 0.02 이하부터 오차 값이 커지기 시작한다. 따라서 본 실시예에서는 가중 평균시 에너지 0.02 이하의 밴드

는 제외하고 가중 평균을 추정한다.

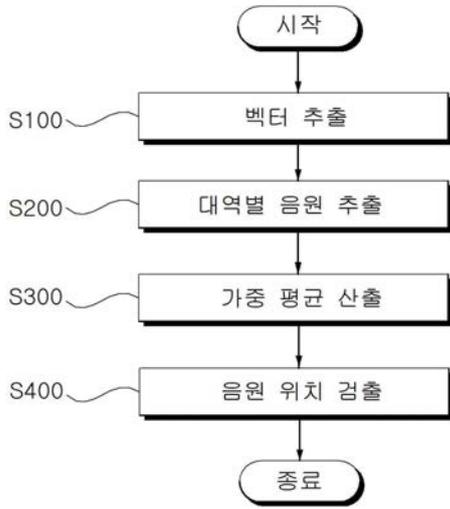
- [0056] 즉, 본 실시예에서 가중 평균 산출 단계(S300)는 가중 평균 산출부(300)가 스테레오 신호의 대역별 에너지를 이용하여 상기 추출된 음원 위치의 정확도에 따른 가중 평균을 산출함에 있어, 먼저 대역별 에너지를 추출하고 가중 평균의 산출 여부를 결정한 다음, 임의의 기준에 따라 낮은 밴드에 대해서는 SNR 가중 평균 산출을 수행할 수 있으며 또한 높은 밴드에 대해서는 양자 택일 적인 에너지 가중 평균 산출을 수행할 수 있다.
- [0057] 이상의 실시예에 따라 가중 평균이 산출되면 음원 위치 검출 단계는 음원 위치 검출부가 산출된 가중 평균을 이용하여 상기 스테레오 신호의 프레임 음원 위치를 검출하게 된다.
- [0058] 본 발명에 따르면 음원 객체의 위치 추정 정확도를 향상할 수 있기 때문에 다음원 음악 신호에서 음원 객체별 위치를 보다 정확하게 구현하여 주며, 종래 기술에 사용하는 균일 평균 방식을 사용하지 않고 추정 SNR과 밴드별 에너지를 이용하여 효과적으로 음원 객체 위치 추정 정확도를 향상할 수 있다.
- [0059] 또한, 다양한 종류의 음악에 대하여 음원 객체의 위치를 추정함에 따라 음원 객체별 재 배치가 가능하다.
- [0060] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다.
- [0061] 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구 범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리 범위에 포함되는 것으로 해석되어야 할 것이다.

도면

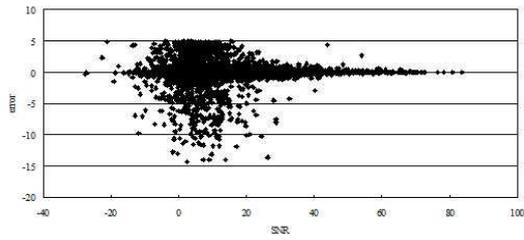
도면1



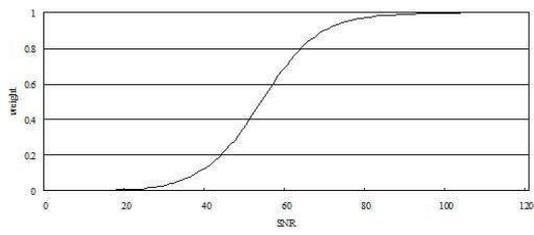
도면2



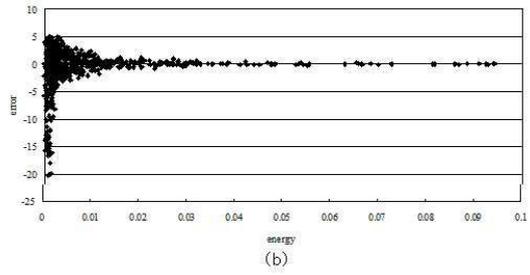
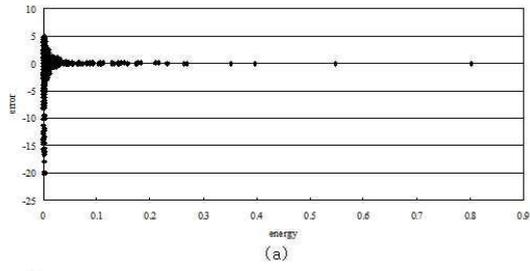
도면3



도면4



도면5





(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년07월22일
 (11) 등록번호 10-1641645
 (24) 등록일자 2016년07월15일

(51) 국제특허분류(Int. Cl.)
 G10L 19/04 (2006.01)
 (21) 출원번호 10-2014-0070876
 (22) 출원일자 2014년06월11일
 심사청구일자 2014년06월11일
 (65) 공개번호 10-2015-0142777
 (43) 공개일자 2015년12월23일
 (56) 선행기술조사문헌
 KR1020070107615 A*
 KR1020130086486 A*
 *는 심사관에 의하여 인용된 문헌

(73) 특허권자
 전자부품연구원
 경기도 성남시 분당구 새나리로 25 (야탑동)
 (72) 발명자
 조충상
 경기도 성남시 분당구 장미로 55 116동 805호 (야탑동, 장미마을코오롱아파트)
 김제우
 경기도 성남시 분당구 수내로 181 309동 905호 (분당동, 샛별마을우방아파트)
 (뒷면에 계속)
 (74) 대리인
 남충우

전체 청구항 수 : 총 6 항

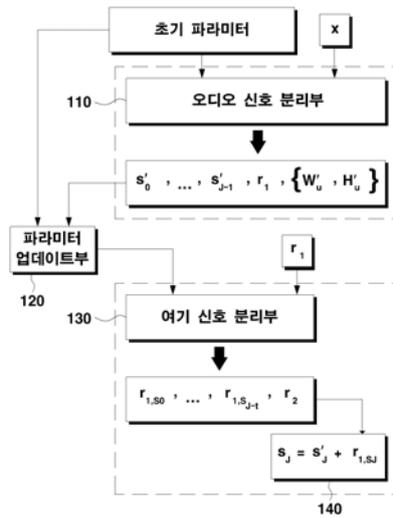
심사관 : 옥윤철

(54) 발명의 명칭 오디오 소스 분리 방법 및 이를 적용한 오디오 시스템

(57) 요약

오디오 소스 분리 방법 및 이를 적용한 오디오 시스템이 제공된다. 본 발명의 실시예들에 따른 오디오 분리 방법은, 믹싱된 오디오 신호를 오디오 소스들로 분리함에 있어, 잔여 신호라는 개념을 도입하여, 오디오 소스들 중 적어도 2개에 해당되는 오디오 신호를 잔여 신호로 별도 분리하여 처리한다. 이에 의해, 오디오 분리 성능 향상을 기대할 수 있다. 또한, 분리된 잔여 신호를 재분리하여 해당 오디오 소스들에 부가할 수 있어, 보다 완전하게 오디오 소스들을 분리할 수 있게 된다.

대표도 - 도3



(72) 발명자

최병호

경기도 용인시 수지구 대지로 27 103동 1306호 (죽전동, 한신아파트)

신화선

경기도 용인시 기흥구 보정로 26 101동 1601호 (보정동, 상록데시앙아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호 10044569

부처명 미래부

연구관리전문기관 한국산업기술평가관리원

연구사업명 지식경제 산업융합원천기술개발사업

연구과제명 2D 스테레오 콘텐츠를 3D 입체 음향 콘텐츠로 변환하기 위한 음원 객체 분리/위치 추정 및 3D 렌더링 소프트웨어 기술 개발

기 여 율 1/1

주관기관 (주)이머시스

연구기간 2013.06.01 ~ 2016.05.31

명세서

청구범위

청구항 1

믹싱된 오디오 신호를 입력받는 단계; 및
 입력된 믹싱된 오디오 신호를 다수의 오디오 소스들과 제1 여기 신호로 분리하는 제1 분리단계;를 포함하고,
 상기 제1 여기 신호는,
 상기 다수의 오디오 소스들 중 적어도 2개에 공통되는 오디오 신호인 것을 특징으로 하는 오디오 분리 방법.

청구항 2

삭제

청구항 3

제 1항에 있어서,
 상기 제1 분리단계에서 분리된 상기 여기 신호를, 상기 오디오 소스들 각각에 해당하는 여기 신호들과 제2 여기 신호로 분리하는 제2 분리단계; 및
 상기 여기 신호들을 상기 오디오 소스들에 각각 추가하는 단계;를 더 포함하는 것을 특징으로 하는 오디오 분리 방법.

청구항 4

제 3항에 있어서,
 상기 제1 분리 단계 및 상기 제2 분리단계는, NMF-EM(Nonnegative Matrix Factorization - Expectation Maximization) 기법을 이용하여, 분리 작업을 수행하고,
 상기 제2 분리단계는,
 상기 제1 분리단계에서 사용한 초기 파라미터들 및 상기 제1 분리단계에 의해 업데이트된 파라미터들을 기초로 결정한 파라미터들을 이용하는 것을 특징으로 하는 오디오 분리 방법.

청구항 5

제 4항에 있어서,
 상기 제2 분리단계는,
 상기 결정한 파라미터들에 가중치를 부가한 파라미터들을 이용하는 것을 특징으로 하는 오디오 분리 방법.

청구항 6

제 5항에 있어서,
 상기 가중치는,
 상기 믹싱된 오디오 신호의 절대 파워 평균과 상기 제1 여기 신호의 절대 파워 평균을 기초로 결정되는 것을 특

징으로 하는 오디오 분리 방법.

청구항 7

믹싱된 오디오 신호를 입력받는 입력부; 및
 입력된 믹싱된 오디오 신호를 다수의 오디오 소스들과 제1 여기 신호로 분리하는 분리부;를 포함하고,
 상기 제1 여기 신호는,
 상기 다수의 오디오 소스들 중 적어도 2개에 공통되는 오디오 신호인 것을 특징으로 하는 오디오 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 오디오 소스 분리 방법에 관한 것으로, 더욱 상세하게는 믹싱된 오디오 신호로부터 오디오 소스들을 분리하는 방법 및 이를 적용한 오디오 시스템에 관한 것이다.

배경 기술

[0002] 도 1은 기존의 오디오 소스 분리 기술을 개념적으로 도시한 도면이다. 도 1에서, s_1, s_2, s_3 은 3개의 서로 다른 오디오 소스들을 의미하고, x 는 믹싱된 오디오 신호를 의미한다. 즉, x 는 s_1, s_2, s_3 가 합쳐진 신호이다.

[0003] 도 1에 도시된 바와 같이, 각각의 오디오 소스들 s_1, s_2, s_3 은 겹쳐진 성분이 존재하지 않는다. 즉, 오디오 소스들 s_1, s_2, s_3 은 서로 독립적인 관계라고 할 수 있다.

[0004] 이와 같은 상황에서, 오디오 신호 x 를 오디오 소스들 s_1, s_2, s_3 로 분리함에 있어서는 아무런 문제가 없다. 오디오 신호 x 를 구성하는 오디오 성분을 오디오 소스들 s_1, s_2, s_3 중 어느 하나로 매칭시킬 수 있기 때문이다.

[0005] 하지만, 도 1에 도시된 상태의 오디오 신호 x 와 오디오 소스들 s_1, s_2, s_3 은 이상적인 경우이거나 매우 특수한 경우에 해당한다. 실제로, 오디오 신호 x 와 오디오 소스들 s_1, s_2, s_3 은, 도 2에 도시된 바와 같다.

[0006] 즉, 오디오 소스들 s_1, s_2, s_3 은 완전하게 독립적이지 않으며, 이에 따라 겹쳐진 영역이 존재하게 된다. 이와 같은 상황은, 오디오 소스들 s_1, s_2 및 s_3 을 하나의 오디오 신호 x 로 믹싱함에 있어서는, 아무런 문제가 발생되지 않는다.

[0007] 하지만, 믹싱된 오디오 신호 x 를 오디오 소스들 s_1, s_2, s_3 로 분리하는 데에는 문제가 발생한다. 오디오 소스들 s_1, s_2, s_3 의 겹쳐진 영역에 해당하였던 오디오 성분을 오디오 소스들 s_1, s_2, s_3 중 어느 하나에 매칭시킬 수 없기 때문이다.

[0008] 이와 같은 문제로 인해, 실제 오디오 신호 x 와 오디오 소스들 s_1, s_2, s_3 은 도 2에 도시된 바와 같음에도 불구하고, 오디오 소스 분리 알고리즘은 오디오 신호 x 와 오디오 소스들 s_1, s_2, s_3 을 도 1에 도시된 바와 같은 상태로 가정하고 처리하고 있는 실정이다.

[0009] 실제 오디오 신호와 오디오 소스들의 상태가 반영되지 않은 채로 오디오 소스 분리가 이루어지기 때문에, 오디오 소스 분리 성능이 좋을 리 없음은 충분히 예측할 수 있으며, 실제로도 그러하다.

발명의 내용

해결하려는 과제

[0010] 본 발명은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 본 발명의 목적은, 믹싱된 오디오 신호를 오디오 소스들로 분리함에 있어, 오디오 소스들 중 적어도 2개에 해당되는 오디오 신호를 잔여 신호로 별도 분리하는 기법에 기반한 오디오 소스 분리 방법 및 이를 적용한 오디오 시스템을 제공함에 있다.

과제의 해결 수단

[0011] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른, 오디오 분리 방법은, 믹싱된 오디오 신호를 입력받는 단계; 및 입력된 믹싱된 오디오 신호를 다수의 오디오 소스들과 제1 여기 신호로 분리하는 제1 분리단계;를 포함한다.

[0012] 그리고, 상기 제1 여기 신호는, 상기 다수의 오디오 소스들 중 적어도 2개에 공통되는 오디오 신호일 수 있다.

[0013] 또한, 본 발명의 일 실시예에 따른 오디오 분리 방법은, 상기 제1 분리단계에서 분리된 상기 여기 신호를, 상기 오디오 소스들 각각에 해당하는 여기 신호들과 제2 여기 신호로 분리하는 제2 분리단계; 및 상기 여기 신호들을 상기 오디오 소스들에 각각 부가하는 단계;를 더 포함할 수 있다.

[0014] 그리고, 상기 제1 분리 단계 및 상기 제2 분리단계는, NMF-EM(Nonnegative Matrix Factorization - Expectation Maximization) 기법을 이용하여, 분리 작업을 수행하고, 상기 제2 분리단계는, 상기 제1 분리단계에서 사용한 초기 파라미터들 및 상기 제1 분리단계에 의해 업데이트된 파라미터들을 기초로 결정한 파라미터들을 이용할 수 있다.

[0015] 또한, 상기 제2 분리단계는, 상기 결정한 파라미터들에 가중치를 부가한 파라미터들을 이용할 수 있다.

[0016] 그리고, 상기 가중치는, 상기 믹싱된 오디오 신호의 절대 파워 평균과 상기 제1 잔여 신호의 절대 파워 평균을 기초로 결정될 수 있다.

[0017] 한편, 본 발명의 다른 실시예에 따른, 오디오 시스템은, 믹싱된 오디오 신호를 입력받는 입력부; 및 입력된 믹싱된 오디오 신호를 다수의 오디오 소스들과 제1 여기 신호로 분리하는 분리부;를 포함한다.

발명의 효과

[0018] 이상 설명한 바와 같이, 본 발명의 실시예들에 따르면, 믹싱된 오디오 신호를 오디오 소스들로 분리함에 있어, 잔여 신호라는 개념을 도입하여, 오디오 소스들 중 적어도 2개에 해당되는 오디오 신호를 잔여 신호로 별도 분리하여 처리하게 되므로, 오디오 분리 성능 향상을 기대할 수 있다.

[0019] 또한, 본 발명의 실시예들에 따르면, 분리된 잔여 신호를 재분리하여 해당 오디오 소스들에 부가할 수 있어, 보다 완전하게 오디오 소스들을 분리할 수 있게 된다.

도면의 간단한 설명

[0020] 도 1은 기존의 오디오 소스 분리 기술을 개념적으로 도시한 도면,
 도 2는 실제 오디오 신호와 오디오 소스들 간의 관계를 나타낸 도면,
 도 3은 본 발명의 일 실시예에 따른 오디오 시스템의 블럭도, 그리고,
 도 4 내지 도 7에는 오디오 분리 성능 평가 결과를 나타낸 그래프들이다.

발명을 실시하기 위한 구체적인 내용

[0021] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

[0022] 도 3은 본 발명의 일 실시예에 따른 오디오 시스템의 블럭도이다. 본 실시예에 따른 오디오 시스템은, 오디오 신호를 오디오 소스들로 분리하기 위한 시스템이다.

- [0023] 이와 같은 기능을 수행하는 본 실시예에 따른 오디오 시스템은, 도 3에 도시된 바와 같이, 오디오 신호 분리부(110), 파라미터 업데이트부(120), 여기 신호 분리부(130) 및 오디오 소스 합성부(140)를 포함한다.
- [0024] 본 발명의 실시예에서, 오디오 신호 x 는 J 개의 오디오 소스(객체)들 s_0, \dots, s_{J-1} 이 믹싱된 신호인 것을 상정한다.
- [0025] 오디오 신호 분리부(110)는 입력되는 오디오 신호 x 를 다수의 오디오 소스들 s'_0, \dots, s'_{J-1} 과 여기 신호 r_1 로 분리한다. 여기 신호 r_1 는 오디오 소스들 s_0, \dots, s_{J-1} 중 적어도 2개에 공통된(겹쳐진) 오디오 신호에 해당한다.
- [0026] 오디오 신호 x 로부터 여기 신호 r_1 가 분리되는 관계로, 오디오 신호 분리부(110)를 통해 오디오 신호 x 로부터 분리되는 오디오 소스들 s'_0, \dots, s'_{J-1} 은 오디오 신호 x 를 믹싱하는데 기초가 된 원래의 오디오 소스들 s_0, \dots, s_{J-1} 과 차이가 있다.
- [0027] 오디오 신호 분리부(110)는 NMF-EM(Nonnegative Matrix Factorization - Expectation Maximization) 기법을 이용하여, 오디오 신호 x 에 대한 분리 작업을 수행한다.
- [0028] NMF-EM 기법은 오디오 분리에 널리 사용되는 기지의 방법으로, 이에 대한 상세한 설명은 생략한다.
- [0029] 기존 방식의 경우, NMF-EM 기법에 의한 오디오 분리는 오디오 소스들에 대한 초기 파라미터들 $\{W'H'\}$ 로부터 업데이트된 파라미터들 $\{W_u'H_u'\}$ 이 생성되며, 업데이트된 파라미터들 $\{W_u'H_u'\}$ 에 의해 오디오 소스들이 결정된다.
- [0030] 하지만, 본 발명의 실시예에서는, 오디오 신호로부터 오디오 소스들 외에 잔여 신호 r_1 을 더 분리하기 때문에, 초기 파라미터들 $\{W'H'\}$ 와 업데이트된 파라미터들 $\{W_u'H_u'\}$ 에는, 오디오 소스들에 대한 파라미터들 외에 잔여 신호 r_1 에 대한 파라미터가 더 포함됨에 유념하여야 한다.
- [0031] 여기 신호 분리부(130)는 오디오 신호 분리부(110)에서 분리된 여기 신호 r_1 를 재분리한다. 구체적으로, 여기 신호 분리부(130)는 여기 신호 r_1 를 오디오 소스들에 대한 여기 신호들 $r_{1,s_0}, \dots, r_{1,s_{J-1}}$ 과 여기 신호 r_2 로 분리한다.
- [0032] 여기 신호 r_2 는 오디오 소스들에 대한 여기 신호들 $r_{1,s_0}, \dots, r_{1,s_{J-1}}$ 에 포함시킬 수 없는 신호이다. 개념적으로, 여기 신호 r_2 는, 오디오 소스들 s_0, \dots, s_{J-1} 중 적어도 2개에 공통된(겹쳐진) 여기 신호 r_1 로 이해할 수 있다.
- [0033] 여기 신호 분리부(130)도 NMF-EM 기법을 이용하여, 여기 신호 r_1 에 대한 분리 작업을 수행한다. 단, NMF-EM 기법을 적용함에 있어 사용되는 초기 파라미터들 $\{W_n'H_n'\}$ 은 파라미터 업데이트부(120)가 아래의 수학적 식 1에 따라 산출한다.
- [0034] [수학적 식 1]
- [0035]
$$\{W_n'H_n'\} = w_2 \times [w_1\{W'H'\} + (1-w_1)\{W_u'H_u'\}]$$
- [0036] 여기서, $\{W'H'\}$ 은 오디오 신호 분리부(110)가 오디오 신호 x 를 분리하는데 이용한 초기 파라미터들이고, $\{W_u'H_u'\}$ 은 오디오 신호 분리부(110)에 의한 오디오 분리 과정에서 업데이트된 파라미터들이다.
- [0037] 이와 같이, 여기 신호 r_1 을 분리하는데 이용하는 파라미터들은, 오디오 신호 x 를 분리하는 과정에서 이용하였던 초기 파라미터들과 분리 결과로 생성된 업데이트된 파라미터들의 가중 합으로부터 획득된다.
- [0038] 가중치 w_1 은 초기 파라미터들 $\{W'H'\}$ 과 업데이트된 파라미터들 $\{W_u'H_u'\}$ 의 비중을 결정하기 위한 가중치로, $0 \leq w_1 \leq 1$ 이다. 가중치 w_2 는 초기 파라미터들 $\{W'H'\}$ 과 업데이트된 파라미터들 $\{W_u'H_u'\}$ 의 비중을 결정하기 위한 가중치로, $0 \leq w_2 \leq 1$ 이다.
- [0039] 가중치 w_2 는 오디오 신호 x 의 절대 파워 평균과 잔여 신호 r_1 의 절대 파워 평균의 비율로 결정되며, 구체적으로

는 아래의 수학적식 2와 같다.

[수학적식 2]

$$w_2 = \frac{\frac{1}{F \times N} \sum_{f,n} |X_{f,n}|}{\frac{1}{F \times N} \sum_{f,n} |R_{1,f,n}|}$$

[0041]

[0042]

오디오 소스 합성부(140)는 오디오 신호 분리부(110)에서 분리된 오디오 소스들 s'_0, \dots, s'_{J-1} 에, 여기 신호 분리부(130)에서 분리된 오디오 소스들에 대한 여기 신호들 $r_{1,s_0}, \dots, r_{1,s_{J-1}}$ 을, 각각 부가하여, 최종 오디오 소스들을 생성한다.

[0043]

한편, 여기 신호 분리부(130)에서 분리된 여기 신호 r_2 에 대해서는 폐기할 수 있지만, 재분리하는 것도 가능하다. 구체적으로, 오디오 소스 합성부(140)가 여기 신호 r_2 를 여기 신호 분리부(130)에 인가하여, 여기 신호 r_1 과 마찬가지로 여기 신호 분리부(130)에 의해 여기 신호 r_2 가 분리되도록 하는 것이다.

[0044]

이 경우, 오디오 소스 합성부(140)는 최종 오디오 소스들에 대해, 여기 신호 r_2 로부터 분리된 오디오 소스들에 대한 여기 신호들 $r_{2,s_0}, \dots, r_{2,s_{J-1}}$ 을 각각 부가할 것이다. 아울러, 여기 신호 분리부(130)에 의해 여기 신호 r_2 로부터 여기 신호 r_3 이 분리된다.

[0045]

이후, 여기 신호 r_3 에 대해서도 재분리 과정을 반복하는 것이 가능하며, 궁극적인 재분리 반복 여부는 여기 신호와 오디오 소스들의 파라미터를 기초로 결정할 수 있다.

[0046]

지금까지, 믹싱된 오디오 신호를 오디오 소스들로 분리함에 있어, 잔여 신호라는 개념을 도입하여 오디오 소스들 중 적어도 2개에 해당되는 오디오 신호를 이 잔여 신호로 별도 분리하는 기법에 따른 오디오 분리에 대해 바람직한 실시예를 들어 상세히 설명하였다.

[0047]

위 기법에 따른 오디오 분리는, 감시 시스템에 적용되어, 오디오 신호로부터 특정 오디오 소스(예를 들면, 음성)만을 추출하거나 특정 오디오 소스(예를 들면, 바람 소리, 자동차 경적 소리)를 제거하는데 활용될 수 있다. 나아가, 오디오 소스별 오디오 효과 부여나, 콘텐츠 제작에도 적용될 수 있음은 물론이다.

[0048]

도 4 내지 도 7에는 오디오 분리 성능 평가 결과를 나타내었다. 도 4 내지 도 7에 도시된 바와 같이, 잔여 신호를 이용한 오디오 소스 분리의 성능이 그렇지 않은 경우보다 우수함을 알 수 있다. 또한, 잔여 신호 분리 기법까지 적용한다면, 그 성능은 더욱 더 우수해짐을 확인할 수 있다.

[0049]

또한, 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

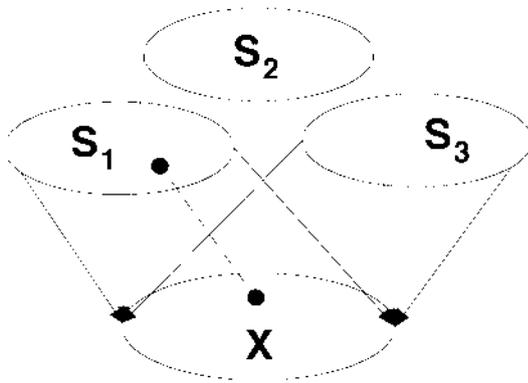
부호의 설명

[0050]

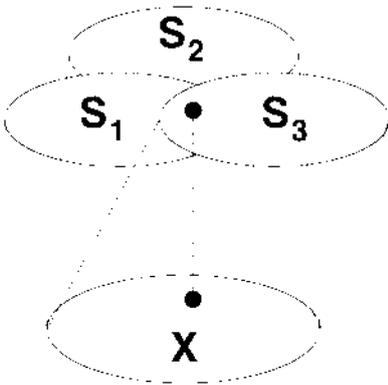
- 110 : 오디오 신호 분리부
- 120 : 파라미터 업데이트부
- 130 : 여기 신호 분리부
- 140 : 오디오 소스 합성부

도면

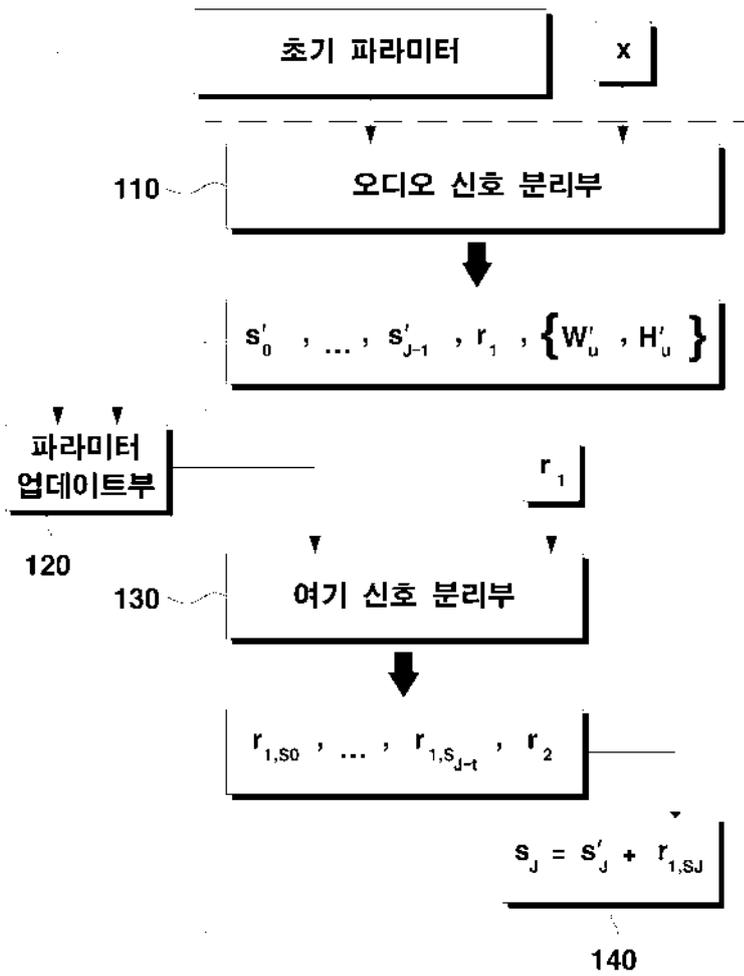
도면1



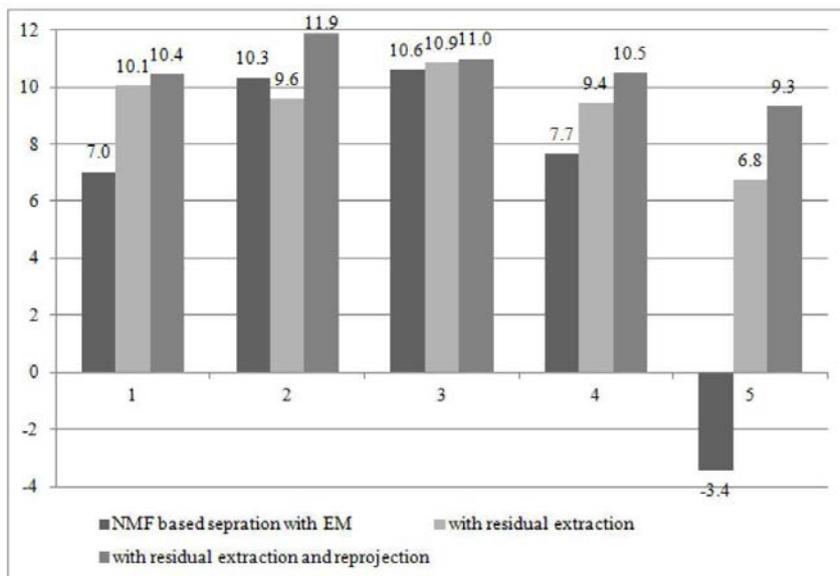
도면2



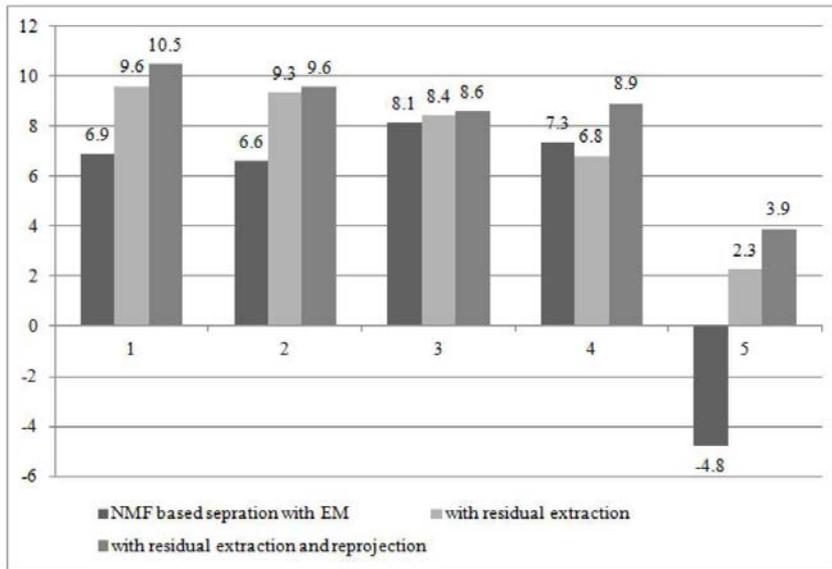
도면3



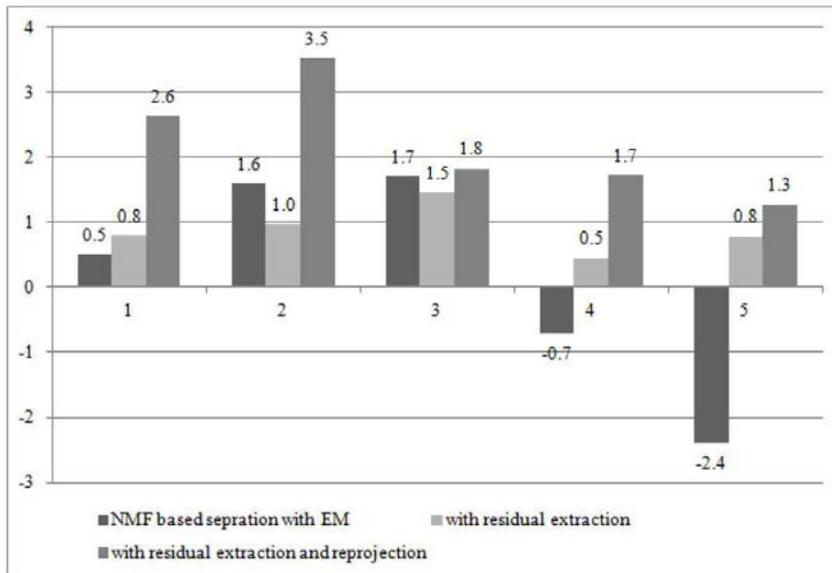
도면4



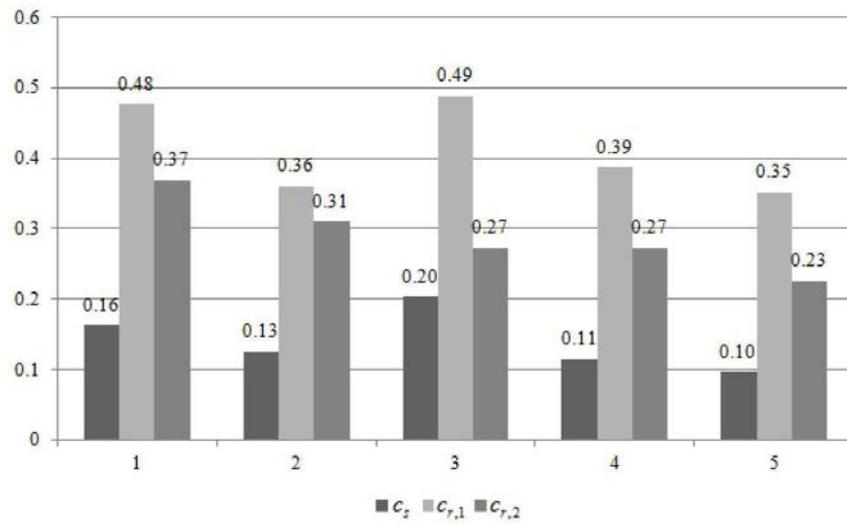
도면5



도면6



도면7





(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년02월09일
 (11) 등록번호 10-1825949
 (24) 등록일자 2018년01월31일

(51) 국제특허분류(Int. Cl.)
G01S 3/808 (2006.01) **G01S 5/22** (2006.01)

(52) CPC특허분류
G01S 3/8083 (2013.01)
G01S 5/22 (2013.01)

(21) 출원번호 10-2015-0140500

(22) 출원일자 2015년10월06일

심사청구일자 2015년10월06일

(65) 공개번호 10-2017-0041328

(43) 공개일자 2017년04월17일

(56) 선행기술조사문헌

JP2010286685 A

한택진. 스테레오 신호에서의 향상된 음원 위치 추정 방법. 광운대학교 대학원 석사학위논문. 2015.08, pp.1-32..

KR101509649 B1*

JP2010541350 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

전자부품연구원

경기도 성남시 분당구 새나리로 25 (야탑동)

(72) 발명자

송재중

경기도 수원시 장안구 장안로 200 동신아파트 302동 701호

양창모

경기도 고양시 일산서구 강선로 141 후곡마을16단지아파트 1605동 302호

(뒷면에 계속)

(74) 대리인

특허법인지명

전체 청구항 수 : 총 8 항

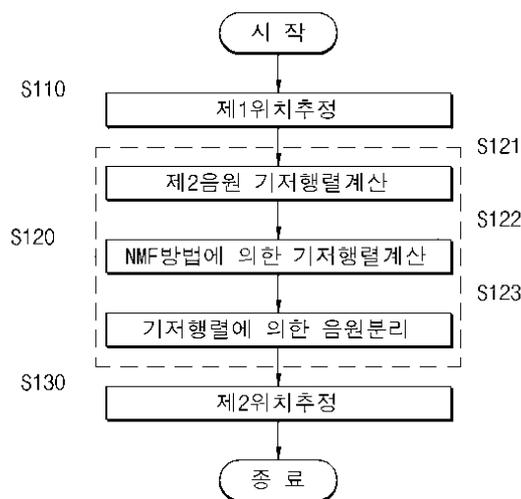
심사관 : 변영석

(54) 발명의 명칭 **음원 분리를 포함하는 음원 위치 추정 장치 및 방법**

(57) 요약

본 발명은 오디오 신호에 포함된 특정 음원의 위치를 추정하는 방법과 그장치에 관한 발명으로, 본 발명의 일면에 따른 음원 위치 추정 방법은 둘 이상의 음원을 포함하는 오디오 신호에서 한 음원의 제1위치를 추정하는 단계; 상기 추정된 제1위치를 이용하여 상기 오디오 신호에서 상기 한 음원만 포함하는 오디오 신호를 분리하는 단계; 상기 분리된 한 음원만 포함하는 오디오 신호에서 상기 한 음원의 제2위치를 추정하는 단계를 포함하는 것을 특징으로 한다.

대표도 - 도1



(52) CPC특허분류
H04S 2400/11 (2013.01)

(72) 발명자
박성주
경기도 용인시 기흥구 죽현로 12 동원로얄듀크 31
0동 501호

김동철
서울특별시 노원구 광운로2가길 17-6 303호

김기준
전라남도 목포시 하당로289번길 6 제일3차아파트
302동 306호

박호중

서울특별시 강남구 삼성로 151 8동 106호 (대치
동, 선경아파트)

한택진

경기도 의정부시 동일로466번길 5 서해아파트 108
동 1206호

이 발명을 지원한 국가연구개발사업

과제고유번호 R0101-15-0061

부처명 미래창조과학부

연구관리전문기관 정보통신기술진흥센터

연구사업명 정보통신·방송 연구개발 사업

연구과제명 2D 스테레오 콘텐츠를 3D 입체 음향 콘텐츠로 변환하기 위한 음원 객체 분리/위치 추정 및
3D 렌더링 소프트웨어 기술 개발

기 여 율 1/1

주관기관 ㈜이머시스

연구기간 2015.03.01 ~ 2016.02.29

공지예외적용 : 있음

명세서

청구범위

청구항 1

둘 이상의 음원을 포함하는 오디오 신호에서 한 음원의 제1위치를 추정하는 단계;

상기 추정된 제1위치를 이용하여 상기 오디오 신호에서 상기 한 음원만 포함하는 오디오 신호를 분리하는 단계;
및

상기 분리된 한 음원만 포함하는 오디오 신호에서 상기 한 음원의 제2위치를 추정하는 단계;를 포함하되,

상기 오디오 신호를 분리하는 단계는,

상기 둘 이상의 음원을 포함하는 오디오 신호에서 상기 한 음원과 상이한 다른 음원의 기저행렬을 산출하는 단계;

상기 다른 음원의 기저행렬로부터 전체 기저행렬을 산출하는 단계; 및

상기 전체 기저행렬을 이용하여 상기 한 음원을 분리하는 단계를 포함하는 것

인 음원 위치 추정 방법.

청구항 2

제1항에 있어서, 상기 오디오 신호를 분리하는 단계는,

상기 추정된 제1위치에 의해 상기 한 음원의 특성을 추출하고,

상기 추출된 음원의 특성을 비음수행렬분해(Non-negative Matrix Factorization, NMF) 계산 과정에 적용하여
상기 음원의 기저행렬을 결정하고,

상기 기저행렬을 비음수행렬분해방법에 적용하여 상기 오디오 신호를 분리하는 것

인 음원 위치 추정 방법.

청구항 3

제2항에 있어서, 상기 비음수행렬분해방법은,

유클리디안 디스턴스(Euclidean Distance)를 이용하는 방법인 것

인 음원 위치 추정 방법.

청구항 4

제1항에 있어서, 상기 제1위치 및 제2위치를 추정하는 단계는,

주성분분석(Principal Component Analysis, PCA)방법으로 이루어지는 것

인 음원 위치 추정 방법.

청구항 5

적어도 하나 이상의 프로세서를 포함하는 음원 위치 추정 장치에 있어서, 상기 프로세서는

둘 이상의 음원을 포함하는 오디오 신호에서 한 음원의 제1위치를 추정하는 제1위치추정부;

상기 제1위치를 이용하여 상기 오디오 신호에서 상기 한 음원만 포함하는 오디오 신호를 분리하는 음원분리부; 및

상기 음원분리부에서 분리된 오디오신호에서 상기 한 음원의 제2위치를 추정하는 제2위치추정부를 포함하되,

상기 음원 분리부는 상기 둘 이상의 음원을 포함하는 오디오 신호에서 상기 한 음원과 상이한 다른 음원의 기저행렬을 산출하고, 상기 다른 음원의 기저행렬로부터 전체 기저행렬을 산출한 다음, 상기 전체 기저행렬을 이용하여 상기 한 음원을 분리하는 것

인 음원 위치 추정 장치.

청구항 6

제5항에 있어서, 상기 음원분리부는,

상기 추정된 제1위치에 의해 상기 한 음원의 특성을 추출하고,

상기 추출된 음원의 특성을 비음수행렬분해(Non-negative Matrix Factorization, NMF) 계산 과정에 적용하여 상기 음원의 기저행렬을 결정하고,

상기 결정된 기저행렬을 비음수행렬분해방법에 적용하여 상기 오디오 신호를 분리하는 것

인 음원 위치 추정 장치.

청구항 7

제6항에 있어서, 상기 비음수행렬분해방법은

유클리디안 디스턴스(Euclidean Distance)를 이용하는 방법인 것

인 음원 위치 추정 장치.

청구항 8

제5항에 있어서, 상기 제1위치추정부 및 제2위치추정부는

주성분분석방법으로 음원의 위치를 추정하는 것

인 음원 위치 추정 장치.

발명의 설명

기술 분야

[0001] 본 발명은 다수의 음원으로 구성된 오디오 신호에서 특정 음원의 위치를 추정하는 방법에 관한 것으로서, 오디오 신호를 구성하는 음원을 각각 분리한 후에 분리된 각 음원의 위치를 추정함으로써 추정된 위치의 정확도를 향상시키는 장치 및 방법에 관한 것이다.

배경 기술

[0003] 오디오 신호에서 공간정보(Spatial Information)를 추출해 내는 것은 오디오 신호의 장면을 분석하거나 공간감을 가지는 오디오(Spatial audio) 구현 등을 위해 필요하다.

[0004] 이러한 공간정보를 추출해 내기 위한 오디오 신호는 한 개의 음원으로 구성되는 경우도 있지만 일반적으로는 여러 개의 음원의 조합에 의해 이루어진다. 이러한 오디오 신호에서 원하는 음원만 분리해 내거나(Sound source separation) 각 음원의 위치를 추정(Location estimation)하는 여러 가지 기술들이 연구되어왔다.

- [0005] 주성분분석(Principal Component Analysis, PCA)방법은 음원의 위치를 추정하기 위한 방법 중의 하나로, 복수의 데이터들을 차원이 낮은 데이터로 변환시키는 것인데, 오디오 신호를 직교변환을 사용하여 첫 번째 주성분과 이와 직교하는 두 번째 주성분으로 선형 변환하는 것이다.
- [0006] 그러나 주성분분석방법을 다수의 음원으로 구성된 오디오 신호에 적용하면, 음원들 사이의 간섭에 의하여 각 음원의 위치를 정확하게 추정할 수 없다. 즉, 특정 음원의 위치를 추정할 때 다른 음원들은 상기 특정 음원의 위치추정을 방해하는 잡음 신호로 동작하고, 그에 따라 상기 음원의 위치추정에 오류가 발생한다.
- [0007] 특히 다른 음원들의 위치가 고정인 경우에 이 음원들이 잡음신호로 동작하면 음원 위치추정에 항상 일정한 방향으로 영향을 미치게 되고, 측정하려는 음원의 위치가 잡음방향으로 편중되어 잘못 측정되는 오류가 발생한다.
- [0008] 따라서 주성분분석방법을 사용하여 음원의 위치를 추정하기 위해서는 추정을 방해하는 잡음역할을 하는 다른 음원들을 제거하고 측정하고자 하는 음원만 분리하여 음원의 위치를 추정하는 방법이 필요하다.
- [0009] 특정 음원만 분리하는 방법은 비음수행렬분해(Non-negative Matrix Factorization, NMF)방법을 사용할 수 있는데, NMF방법은 여러 음원을 포함한 오디오 신호에 대해 NMF 기저행렬(Basis Matrix)을 구하고, 기저행렬에서 각 음원에 해당하는 기저행렬을 분리해 내고, 이를 이용하여 오디오 신호를 합성하는 방법으로 음원을 분리하는 방법이다.
- [0010] 그러나 이 과정에서 각 음원에 대한 기저행렬을 추출하기 위하여는 각 음원의 특성을 미리 알아야 하고, 이를 위해서 각 음원에 대한 훈련과정을 독립적으로 수행해야 한다. 이렇게 사전에 각 음원에 대한 훈련과정을 독립적으로 수행하여 각 음원의 특징을 미리 파악해야 하므로 아무런 정보 없이 음원들이 섞여있는 일반 오디오 신호에는 적용하기 어려운 한계가 있다.
- [0011] 이를 극복하기 위해 이미 알려진 음원들의 고유 특성을 이용하여 기저행렬을 추출하는 방법도 있으나 이 방법은 해당 특성을 가지는 음원의 분리에만 제한적으로 적용할 수 있는 문제점이 있다.

발명의 내용

해결하려는 과제

- [0013] 본 발명은 전술한 바와 같은 기술적 배경에서 안출된 것으로서, 다수의 음원으로 구성된 오디오 신호에서 특정 음원의 위치를 추정할 때 다른 음원에 의한 오류를 제거하여 특정 음원의 위치를 정확하게 추정하는 장치와 방법을 제공하는 것을 목적으로 한다.
- [0014] 이를 위해 본 발명은 특정 음원만 분리하여 이로부터 위치를 추정하는데, 음원의 분리는 사전 훈련 과정 또는 음원 고유 특성을 알지 못하는 상태에서도 각 음원의 특성을 자체적으로 분석하고, 분석결과를 활용하여 특정 음원을 분리할 수 있다.
- [0015] 본 발명의 목적은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0017] 전술한 목적을 달성하기 위한 본 발명의 일면에 따른 음원 위치 추정 방법은 둘 이상의 음원을 포함하는 오디오 신호에서 한 음원의 제1위치를 추정하는 단계; 상기 추정된 제1위치를 이용하여 상기 오디오 신호에서 상기 한 음원만 포함하는 오디오 신호를 분리하는 단계; 상기 분리된 한 음원만 포함하는 오디오 신호에서 상기 한 음원의 제2위치를 추정하는 단계를 포함하는 것을 특징으로 한다.
- [0018] 상기 오디오 신호를 분리하는 단계는, 상기 추정된 제1위치에 의해 상기 한 음원의 특성을 추출하고, 상기 추출된 음원의 특성을 비음수행렬분해(Non-negative Matrix Factorization, NMF) 계산 과정에 적용하여 상기 음원의 기저행렬을 결정하고, 상기 기저행렬을 비음수행렬분해방법에 적용하여 상기 오디오 신호를 분리하는 것을 특징으로 한다.
- [0019] 본 발명의 다른 일면에 따른 음원 위치 추정 장치는 적어도 하나 이상의 프로세서를 포함하고, 상기 프로세서는 둘 이상의 음원을 포함하는 오디오 신호에서 한 음원의 제1위치를 추정하는 제1위치추정부; 상기 제1위치를 이용하여 상기 오디오 신호에서 상기 한 음원만 포함하는 오디오 신호를 분리하는 음원분리부; 및 상기 음원분리부에서 분리된 오디오신호에서 상기 한 음원의 제2위치를 추정하는 제2위치추정부를 포함하여 구현하는 것을 특

징으로 한다.

[0020] 상기 음원분리부는, 상기 추정된 제1위치에 의해 상기 한 음원의 특성을 추출하고, 상기 추출된 음원의 특성을 비음수행렬분해(Non-negative Matrix Factorization, NMF) 계산 과정에 적용하여 상기 음원의 기저행렬을 결정하고, 상기 결정된 기저행렬을 비음수행렬분해방법에 적용하여 상기 오디오 신호를 분리하는 것을 특징으로 한다.

발명의 효과

[0022] 본 발명에 따르면, 위치를 추정하고자 하는 특정한 음원을 분리해내어 이를 이용하여 위치를 추정하므로 여러 음원이 혼재된 상태에서 특정 음원의 위치를 추정할 때 보다 훨씬 정확하게 해당 음원의 위치를 추정할 수 있는 효과가 있다.

도면의 간단한 설명

[0024] 도 1은 본 발명의 일실시예에 따른 음원 위치 추정 방법의 흐름도.

도 2는 본 발명의 다른 실시예에 따른 음원 위치 추정 장치의 구조도.

발명을 실시하기 위한 구체적인 내용

[0025] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 한편, 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소, 단계, 동작 및/또는 소자는 하나 이상의 다른 구성요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

[0027] 이하, 본 발명의 바람직한 실시예에 대하여 첨부한 도면을 참조하여 상세히 설명하기로 한다. 도 1은 본 발명에 의한 위치추정 방법을 위한 흐름도를 나타낸다.

[0028] 특정 음원의 위치를 추정하기 위한 방법은 크게 세 단계로 이루어지는데, 우선 종래의 PCA방법을 이용하여 특정 음원의 초기 위치를 대략적으로 추정하는 제1위치추정단계(S110), 이렇게 추정된 위치를 기반으로 특정음원을 분리해 내어 그로부터 기저행렬을 구하는 음원분리단계(S120), 마지막으로 기저행렬을 이용하여 보다 정확한 특정 음원의 위치를 추정하는 제2위치추정단계(S130)를 거쳐 이루어진다.

[0029] 제1위치추정단계(S110)에서는 우선 입력신호의 좌측채널 신호와 우측채널 신호인 x_L 과 x_R 로부터 종래의 PCA 방법을 이용하여 특정 음원 P_1 의 대략적인 위치를 구하는데, 위치 추정 결과는 P_1 의 패닝이득값인 a_{L1} 과 a_{R1} 으로 나타난다.

[0030] 제1위치추정단계(S110)에서 구한 패닝이득값을 이용하여 위치를 추정하고자 하는 음원 P_1 을 입력 오디오신호로부터 분리하는 단계(S120)를 거치는데 이는 다시 오디오신호에 포함된 다른 음원인 P_2 의 기저행렬을 계산하는 단계(S121), P_2 의 기저행렬로부터 전체 기저행렬을 구하는 단계(S122), 마지막으로 전체 기저행렬을 이용하여 P_1 을 분리해내는 단계(S123)로 이루어진다.

[0031] 이하에서 제2음원인 P_2 의 기저행렬을 구하는 단계(S121)를 설명하기로 한다.

[0032] 입력신호로부터 P_1 을 분리해 내기 위해서는 NMF방법을 이용한다. 오디오 신호 행렬인 V 는 NMF 윈도우를 이용하여 기저행렬 B와 이득행렬 G의 곱으로 나타낼 수 있는데, 근사화된 오디오 신호 행렬 V' 는 $V' = BG$ 로 나타낼 수 있고 이때 모든 행렬의 요소(element)들은 음수가 아니어야 한다. 오디오 신호를 이루는 각 채널의 오디오 신호 행렬은 K개의 프레임으로 이루어질 수 있는데 이를 NMF 윈도우라 한다.

[0033] 이때 기저행렬 B 와 이득행렬 G 는 반복 알고리즘에 의해 구할 수 있는데 V 와 V' 사이의 오차가 최소가 될 때까지 반복해서 B 와 G 를 계산하는 방법이다.

[0034] NMF방법은 오디오 신호 행렬인 V 를 근사화하여 표현하는 것이므로 오차가 있을 수밖에 없는데 오차를 줄이기 위해서 Orthogonal NMF, Discriminative NMF, Convolution NMF, Group sparsity NMF 등의 NMF 방법이 사용될 수 있다.

[0035] 오디오 신호가 좌측 채널과 우측 채널로 이루어진 스테레오 신호라고 하면 입력 신호는 각각 x_L 과 x_R 로 나타내고 K프레임으로 이루어진 NMF원도우에 의한 기저행렬과 이득행렬의 곱으로 나타나는 신호행렬은 V_L 과 V_R 로 나타낸다. 이 때 좌측 채널 오디오 신호행렬인 V_L 과 우측 채널 오디오 신호행렬인 V_R 은 공통적인 음원을 가지고 있고, 또한, 같은 기저행렬을 가진다. 따라서 신호행렬 $V_c, c = L, R,$ 는 다음 수학적 식 1과 같이 나타낼 수 있다.

수학적 식 1

[0036]
$$V_c = [|x_c(t - (K - 1))| |x_c(t - (K - 2))| \dots |x_c(t)|]$$

[0037] 구하려는 특정 음원을 P_1 이라 하고 p_1 을 음원 신호라 하면 다른 음원과 음원 신호는 P_2 와 p_2 로 표현할 수 있다.

[0038] 이들 사이의 관계를 수학적 식 2 및 수학적 식 3과 같이 나타낼 수 있다.

수학적 식 2

[0039]
$$x_L = a_{L1}p_1 + a_{L2}p_2 = y_{L1} + y_{L2}$$

수학적 식 3

[0040]
$$x_R = a_{R1}p_1 + a_{R2}p_2 = y_{R1} + y_{R2}$$

[0041] y_{ci} 는 i 번째 음원 신호의 c 채널에 해당하는 값으로 음원신호 p_i 에 이득값 a_{ci} 를 곱한 값이다.

[0042] 또한 i 번째 음원 신호를 p_i 라 하면 이에 대한 기저행렬은 B_i 로 나타낼 수 있고 따라서 V_L 과 V_R 의 공통적인 기저행렬은 B 로 나타내고 $B = [B_1 B_2]$ 로 나타낼 수 있다.

[0043] 따라서 c 채널에 대한 오디오 행렬신호를 V_c 로 나타내면 B 와 G_c 로 근사화 하여 나타낸 V_c' 는 수학적 식 4로 나타낼 수 있다.

수학적 식 4

[0044]
$$V_c' = BG_c = [B_1 B_2] \begin{bmatrix} G_{c1} \\ G_{c2} \end{bmatrix}$$

[0045] G_c 는 c 채널에 대한 이득행렬이고 G_{c1} 과 G_{c2} 는 각각 c 채널의 p_1 , p_2 소스에 대한 이득행렬을 나타낸다.

[0046] B 와 G_c 를 구하기 위해서는 종래의 NMF방법을 사용하는데 B 를 고정시키고 G_c 에 대해 최적화를 수행하고, 다음

으로 G_c 를 고정하고 B 에 대해 최적화를 수행하는데 이 두 단계를 반복하여 최적화된 B 와 G_c 를 구한다.

[0047] 이렇게 B 와 G_c 를 구하면 결과적으로 음원 신호 p_i 는 B_i 와 G_{ci} 의 곱으로부터 구할 수 있으므로 B 로부터 B_i 를 구하는 것이 대단히 중요한 단계이다.

[0048] 특정 음원인 P_1 의 위치를 추정한다는 것은 결국 스테레오 입력신호인 x_L 과 x_R 로부터 P_1 의 좌측채널 신호인 y_{L1} 과 우측 채널 신호인 y_{R1} 을 구하는 것인데 $y_{L1} = a_{L1}p_1$ 이고 $y_{R1} = a_{R1}p_1$ 으로 나타낸다. a_{L1} 과 a_{R1} 은 각각 좌측채널과 우측채널의 p_1 에 대한 이득 값을 나타낸다.

[0049] 즉, y_{L1} 과 y_{R1} 는 모두 p_1 에 이득값을 곱한 형태이므로 y_{L1} , y_{R1} , p_1 은 모두 같은 기저행렬을 가진다. 따라서 입력신호로부터 구한 s_i 는 수학적 5와 같이 나타낼 수 있고 s_i 역시 p_i 와 같은 기저행렬 B_i 를 포함한다.

수학적 5

[0050]
$$s_i = y_{Li} + y_{Ri} = (a_{Li} + a_{Ri})p_i$$

[0051] 이러한 성질을 이용하여 입력신호 x_L 과 x_R 로부터 s_i 를 구하고 s_i 를 이용해 B_i 를 결정한다. 물론 p_i 를 이용하여 B_i 를 구하는 것도 같은 원리로 가능하지만, p_i 는 s_i 에서 이득값을 제거하는 단계를 더 거쳐야 하기 때문에 오류가 더 커지므로 s_i 를 이용하는 것이 정확도를 높일 수 있다.

[0052] 본 발명에서 P_1 의 위치를 구하기 위해서 B_1 을 구하는 것이 목표이지만, s_2 의 추정값인 \hat{s}_2 를 구하여 이로부터 B_2 를 먼저 구하고, B_2 와 V_L , V_R 을 이용하여 최종적으로 B_1 을 구하는 방법을 사용한다. 이렇게 B_2 와 V_L , V_R 을 먼저 고정해 두고 클로즈드 루프(Closed-loop) 최적화 방식을 사용하여 B_1 을 구하는 것이 훨씬 효율적이기 때문이다.

[0053] 입력신호인 x_L 과 x_R 로부터 초기 P_1 과 P_2 의 위치를 대략적으로 추정하는 것은 기존의 PCA방법을 통해 가능하고 이렇게 추정된 이득값은 \hat{a}_{L1} 과 \hat{a}_{R1} , \hat{a}_{L2} 와 \hat{a}_{R2} 로 나타낼 수 있다.

[0054] 수학적 2, 수학적 3 및 수학적 5를 이용하여 추정된 \hat{s}_2 를 나타내면 다음 수학적 6과 같이 나타낼 수 있다.

수학적 6

[0055]
$$\hat{s}_2 = \hat{a}_{L2}\hat{p}_2 + \hat{a}_{R2}\hat{p}_2 = (x_L - \hat{a}_{L1}\hat{p}_1) + (x_R - \hat{a}_{R1}\hat{p}_1)$$

[0056] p_1 의 추정값인 \hat{p}_1 는 수학적 7과 같이 나타낼 수 있다.

수학적 7

[0057]
$$\hat{p}_1 = \hat{a}_{L1}x_L + \hat{a}_{R1}x_R$$

[0058] 이렇게 구한 \hat{s}_2 로부터 NMF 원도우를 이용하면 기저행렬 B_2 를 구할 수 있고, 이런 식으로 P_2 신호에 대한 아무런

사전 정보 없이 기저행렬 B_2 를 구하는 것이 가능하다.

[0059] 이렇게 구한 B_2 를 이용하면 다음 단계(S122)에서는 NMF방법에 의한 기저행렬 계산 단계(S122)를 거쳐 전체 기저행렬인 B 를 구할 수 있다.

[0060] 구체적으로는, V_L 과 V_R 은 수학식 1과 같이 입력신호인 x_L 과 x_R 로부터 구할 수 있고, V_L 과 V_R 은 공통의 기저행렬 B 를 가지고 있으며 각각의 이득행렬인 G_L 과 G_R 을 가지므로 NMF 방법으로 B 와 G_c 를 구할 수 있다. 이하의 실시예에서는 유클리디안 디스턴스(Euclidean distance)방법을 사용하여 오차를 최소화 하는 방법을 사용하였는데 이에 한정되지 않고 다른 NMF 방법들도 사용될 수 있다.

[0061] NMF 최적화를 위한 반복계산은 수학식 8 및 수학식 9와 같이 나타낼 수 있는데, $V_{ave} = 0.5(V_L + V_R)$, $G_{ave} = 0.5(G_L + G_R)$ 이고 $(\cdot)_{mn}$ 은 NMF 행렬의 각 요소(element)를 나타낸다.

수학식 8

[0062]
$$(B)_{mn} \leftarrow (B)_{mn} \frac{(V_{ave} G^T_{ave})_{mn}}{(B G_{ave} G^T_{ave})_{mn}}$$

수학식 9

[0063]
$$(G_c)_{mn} \leftarrow (G_c)_{mn} \frac{(B^T V_c)_{mn}}{(B^T B G_c)_{mn}}, c = L, R$$

[0064] 수학식 8 및 수학식 9에서 초기 $B = [Q B_2]$ 로 설정되는데 B_2 는 앞에서 구한 행렬을 사용하고, Q 는 임의의 값을 가지는 행렬이다. NMF 최적화 반복이 수행되는 과정에서 B_2 는 고정되어 있고 Q 는 계속 갱신되기 때문에 B 의 계산이 완료되고 나면 Q 가 곧 B_1 이 되는 것이다.

[0065] 수학식 8 및 수학식 9에 의해 G_L 과 G_R 이 결정되고 나면 B_1 의 각 채널별 이득행렬에 해당하는 G_{L2} 과 G_{R1} 도 구할 수 있다.

[0066] 음원 분리의 마지막 단계(S123)에서는 이렇게 구한 각 채널 별 기저행렬과 이득행렬을 곱하여 P_1 의 각 채널에 대한 신호인 y_{L1} 과 y_{R1} 을 구할 수 있다.

[0067] 이렇게 입력신호인 x_L 과 x_R 로부터 P_1 에 해당하는 신호를 분리해 내고 나면 분리해 낸 신호인 y_{L1} 과 y_{R1} 를 이용하여 다시 PCA방법을 이용하여 P_1 의 위치 추정을 하는 단계(S130)를 거쳐 최종적으로 P_1 의 위치를 구하는 것이다.

[0068] 이렇게 P_1 에 대한 신호만 분리하여 위치를 추정하는 경우에는 P_1 과 P_2 가 섞여 있는 x_L 과 x_R 로부터 P_1 을 추정할 때와 달리 P_2 가 노이즈 역할을 하지 않기 때문에 추정오류가 발생하지 않고 따라서 정확한 P_1 의 위치를 구하는 것이 가능하다.

[0069] 본 발명의 다른 실시예에 따른 음원 추정 장치(200)는 제1위치추정부(210), 음원분리부(220) 및 제2위치추정부(230)를 포함하여 이루어진다.

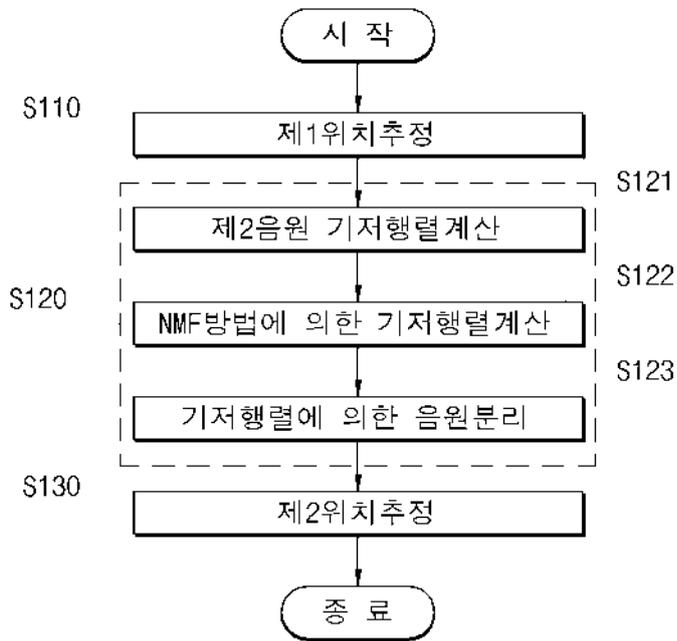
- [0070] 제1위치추정부(210)에서는 입력신호인 x_L 과 x_R 로부터 PCA방법을 이용하여 대략의 P_1 의 위치를 추정하게 된다.
- [0071] 음원분리부(220)에서는 제1위치추정부에서 추정된 P_1 의 특성을 이용하여 NMF방법을 사용하여 입력신호로부터 P_1 만을 분리해낼 수 있는데 상세한 과정은 전술한 바와 같다.
- [0072] 마지막으로 제2위치추정부(230)에서는 음원분리부(220)에서 분리해 낸 P_1 에 PCA방법을 적용하여 P_1 의 위치를 다시 추정하는데, 이때는 P_1 이외의 음원이 포함되지 않으므로 보다 정확한 P_1 의 위치를 추정할 수 있는 효과가 있다.
- [0074] 한편, 본 발명의 일실시예에 따른 음원 위치 추정 방법은 컴퓨터 시스템에서 구현되거나, 또는 기록매체에 기록될 수 있다. 컴퓨터 시스템은 적어도 하나 이상의 프로세서와, 메모리와, 사용자 입력 장치와, 데이터 통신 버스, 사용자 출력 장치와, 저장소를 포함할 수 있다. 전술한 각각의 구성 요소는 데이터 통신 버스를 통해 데이터 통신을 한다.
- [0075] 컴퓨터 시스템은 네트워크에 커플링된 네트워크 인터페이스를 더 포함할 수 있다. 상기 프로세서는 중앙처리 장치(central processing unit)이거나, 혹은 메모리 및/또는 저장소에 저장된 명령어를 처리하는 반도체 장치일 수 있다.
- [0076] 상기 메모리 및 상기 저장소는 다양한 형태의 휘발성 혹은 비휘발성 저장매체를 포함할 수 있다. 예컨대, 상기 메모리는 ROM 및 RAM을 포함할 수 있다.
- [0077] 따라서, 본 발명의 실시예에 따른 음원 위치 추정 방법은 컴퓨터에서 실행 가능한 방법으로 구현될 수 있다. 본 발명의 실시예에 따른 음원 위치 추정 방법이 컴퓨터 장치에서 수행될 때, 컴퓨터로 판독 가능한 명령어들이 본 발명에 따른 인식 방법을 수행할 수 있다.
- [0078] 한편, 상술한 본 발명에 따른 음원 위치 추정 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현되는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체로는 컴퓨터 시스템에 의하여 해독될 수 있는 데이터가 저장된 모든 종류의 기록 매체를 포함한다. 예를 들어, ROM(Read Only Memory), RAM(Random Access Memory), 자기 테이프, 자기 디스크, 플래시 메모리, 광 데이터 저장장치 등이 있을 수 있다. 또한, 컴퓨터로 판독 가능한 기록매체는 컴퓨터 통신망으로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 읽을 수 있는 코드로서 저장되고 실행될 수 있다.
- [0080] 이상, 본 발명의 구성에 대하여 첨부 도면을 참조하여 상세히 설명하였으나, 이는 예시에 불과한 것으로서, 본 발명이 속하는 기술분야에 통상의 지식을 가진자라면 본 발명의 기술적 사상의 범위 내에서 다양한 변형과 변경이 가능함은 물론이다. 따라서 본 발명의 보호 범위는 전술한 실시예에 국한되어서는 아니되며 이하의 특허청구 범위의 기재에 의하여 정해져야 할 것이다.

부호의 설명

- [0082] S110: 제1위치 추정단계 S120: 음원분리단계
- S130: 제2위치 추정단계
- 200: 음원 위치 추정 장치 210: 제1위치추정부
- 220: 음원분리부 230: 제2위치추정부

도면

도면1



도면2

