

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2025-0087883
(43) 공개일자 2025년06월17일

(51) 국제특허분류(Int. Cl.)
G10L 21/0272 (2013.01) G10L 25/03 (2013.01)
G10L 25/51 (2013.01)
(52) CPC특허분류
G10L 21/0272 (2013.01)
G10L 25/03 (2013.01)
(21) 출원번호 10-2023-0177305
(22) 출원일자 2023년12월08일
심사청구일자 2024년02월06일

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
박형민
서울특별시 강남구 삼성로 212, 3동 1002호 (대치동, 은마아파트)
신의협
서울특별시 강서구 강서로56길 64, 807호 (등촌동, 마곡루나플라체)
(74) 대리인
이준영

전체 청구항 수 : 총 17 항

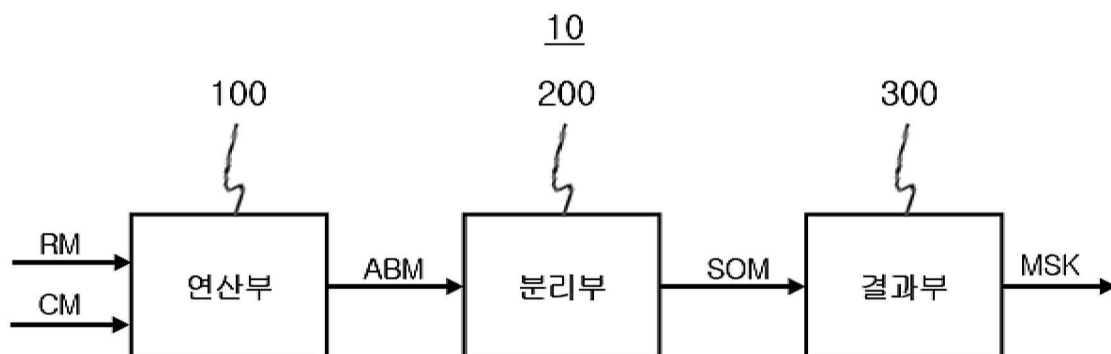
(54) 발명의 명칭 이중 경로 다채널 음원 분리장치

(57) 요약

본 발명의 실시예에 따른 음원 분리장치는 연산부, 분리부 및 결과부를 포함할 수 있다. 연산부는 복수의 채널들 중 기준 마이크로 입력되는 기준 입력신호의 크기를 시간 및 주파수에 따른 2차원 매트릭스로 구현한 기준 매트릭스 및 기준 입력신호의 페이즈 및 기준 마이크를 제외한 복수의 마이크들에 해당하는 비교 마이크로 입력되

(뒷면에 계속)

대표도 - 도1



는 비교 입력신호들의 페이즈에 기초하여 압축 매트릭스를 제공할 수 있다. 분리부는 압축 매트릭스에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다. 결과부는 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 결과 매트릭스를 제공할 수 있다.

본 발명에 따른 음원 분리장치는 압축 매트릭스에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하여 공간정보는 주파수 별로 처리하고, 음원정보는 채널별로 독립적으로 처리함으로써 음원분리 성능을 높일 수 있다.

(52) CPC특허분류

G10L 25/51 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193122
과제번호	2022-0-00989-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	사람중심인공지능핵심원천기술개발
연구과제명	(3세부)대화자 동시 처리를 위한 인공지능 기반 대화 모델링 기술 개발
기 여 율	1/1
과제수행기관명	서강대학교 산학협력단
연구기간	2022.04.01 ~ 2026.12.31

명세서

청구범위

청구항 1

복수의 채널들 중 기준 마이크로 입력되는 기준 입력신호를 시간 및 주파수에 따른 2차원 매트릭스로 구현한 기준 매트릭스 및 상기 복수의 마이크로들에 해당하는 비교 마이크로들로 입력되는 비교 입력신호들의 시간 및 주파수에 따른 2차원 매트릭스로 구현한 비교 매트릭스를 채널 축으로 누적하여 압축 매트릭스를 제공하는 연산부;

상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트랄 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하는 분리부;

상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 마스크 매트릭스를 제공하는 결과부를 포함하는 음원 분리장치.

청구항 2

제1항에 있어서,

상기 연산부는,

상기 기준 매트릭스 및 상기 비교 매트릭스들의 채널축으로의 누적을 통해 입력 매트릭스를 생성하는 입력 생성부;

상기 입력 매트릭스에 대한 코릴레이션(Correlation) 연산에 따라 공간 공분산 매트릭스를 생성하는 공분산 생성부; 및

상기 공간 공분산 매트릭스를 압축하여 압축 매트릭스를 제공하는 압축부를 포함하는 것을 특징으로 하는 음원 분리장치.

청구항 3

제2항에 있어서,

상기 공간 공분산 매트릭스에 포함되는 성분들 각각의 값은 상기 성분들 각각의 크기에 기초하여 노말라이즈(Normalized)되는 것을 특징으로 하는 음원 분리장치.

청구항 4

제3항에 있어서,

상기 분리부는,

상기 압축 매트릭스를 상기 채널을 기준으로 미리 정해진 기준채널 개수로 나누어 복수의 주파수-시간 매트릭스들을 생성하고, 상기 복수의 주파수-시간 매트릭스들에 기초하여 상기 시간 및 상기 주파수 특징을 나타내는 복수의 주파수-시간 특징 매트릭스들을 산출하고, 상기 복수의 주파수-시간 특징 매트릭스들 및 상기 압축 매트릭스를 합성하여 상기 스펙트랄 출력 매트릭스를 제공하는 스펙트랄 분리부; 및

상기 스펙트랄 출력 매트릭스를 상기 주파수를 기준으로 미리 정해진 기준 주파수 개수로 나누어 복수의 채널-시간 매트릭스들을 생성하고, 상기 복수의 채널-시간 매트릭스들에 기초하여 상기 채널 및 상기 시간 특징을 나타내는 복수의 채널-시간 특징 매트릭스들을 산출하고, 상기 복수의 채널-시간 특징 매트릭스들 및 상기 스펙트랄 출력 매트릭스를 합성하여 상기 스페셜 출력 매트릭스를 제공하는 스페셜 분리부를 포함하는 것을 특징으로 하는 음원 분리장치.

청구항 5

제4항에 있어서,

상기 스펙트랄 분리부는,

상기 압축 매트릭스를 상기 채널을 기준으로 나누어 상기 복수의 주파수-시간 매트릭스들을 생성하는 제1 분리부;

상기 복수의 주파수-시간 매트릭스들에 기초하여 상기 시간 및 상기 주파수 특징을 나타내는 상기 복수의 주파수-시간 특징 매트릭스들을 산출하는 제1 산출부; 및

상기 복수의 주파수-시간 특징 매트릭스들 및 상기 압축 매트릭스를 합성하여 상기 스펙트랄 출력 매트릭스를 제공하는 제1 합성부를 포함하는 것을 특징으로 하는 음원 분리장치.

청구항 6

제5항에 있어서,

상기 제1 산출부는,

상기 복수의 주파수-시간 매트릭스들 각각에 포함되는 성분들의 크기에 따라 상기 복수의 주파수-시간 매트릭스들 중 일부를 선택하여 제1 선택 매트릭스들을 선택하는 제1 선택부;

상기 제1 선택 매트릭스들에 기초하여 상기 시간 및 상기 주파수 특징을 나타내는 주파수-시간 기초 매트릭스들을 산출하는 제1 기초부; 및

상기 주파수-시간 기초 매트릭스들을 업샘플링(up-sampling)하여 상기 복수의 주파수-시간 특징 매트릭스들을 제공하는 제1 업샘플부를 포함하는 것을 특징으로 하는 음원 분리장치.

청구항 7

제6항에 있어서,

상기 스페셜 분리부는,

상기 스펙트랄 출력 매트릭스를 상기 주파수를 기준으로 나누어 상기 복수의 채널-시간 매트릭스들을 생성하는 제2 분리부;

상기 복수의 채널-시간 매트릭스들에 기초하여 상기 채널 및 상기 시간 특징을 나타내는 상기 복수의 채널-시간 특징 매트릭스들을 산출하는 제2 산출부; 및

상기 복수의 채널-시간 특징 매트릭스들 및 상기 스펙트랄 출력 매트릭스를 합성하여 상기 스페셜 출력 매트릭스를 제공하는 제2 합성부를 포함하는 것을 특징으로 하는 음원 분리장치.

청구항 8

제7항에 있어서,

상기 제2 산출부는,

상기 복수의 채널-시간 매트릭스들 각각에 포함되는 성분들의 크기에 따라 상기 복수의 채널-시간 매트릭스들 중 일부를 선택하여 제2 선택 매트릭스들을 선택하는 제2 선택부;

상기 제2 선택 매트릭스들에 기초하여 상기 채널 및 상기 주파수 특징을 나타내는 채널-시간 기초 매트릭스들을 산출하는 제2 기초부; 및

상기 채널-시간 기초 매트릭스들을 업샘플링하여 상기 복수의 채널-시간 특징 매트릭스들을 제공하는 제2 업샘플부를 포함하는 것을 특징으로 하는 음원 분리장치.

청구항 9

제8항에 있어서,

상기 결과부는,

상기 스페셜 출력 매트릭스를 확장하여 확장 매트릭스를 제공하는 확장부;

상기 확장 매트릭스를 상기 시간을 기준으로 풀링(pooling)한 풀링 매트릭스를 제공하는 풀링부; 및

상기 풀링 매트릭스에 기초하여 상기 주파수 및 상기 채널에 대해서 압축하여 마스크를 생성하는 마스크 생성부를 포함하는 것을 특징으로 하는 음원 분리장치.

청구항 10

복수의 채널들 중 기준 마이크로 입력되는 기준 입력신호를 시간 및 주파수에 따른 2차원 매트릭스로 구현한 기준 매트릭스 및 상기 기준 입력신호를 제외한 상기 복수의 마이크로들에 해당하는 비교 마이크로들로 입력되는 비교 입력신호들의 비교 매트릭스들에 기초하여 압축 매트릭스를 제공하는 연산부;

상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트랄 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하는 분리부;

상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 결과 매트릭스를 제공하는 결과부; 및

상기 결과 매트릭스에 기초하여 상기 음원에 상응하는 음성신호를 제공하는 음성부를 포함하는 음원 분리시스템.

청구항 11

제10항에 있어서,

상기 연산부는,

상기 기준 매트릭스 및 상기 기준 입력신호를 제외한 상기 복수의 마이크로들에 해당하는 비교 마이크로들로 입력되는 비교 입력신호들의 비교 매트릭스들의 채널축으로의 누적을 통해 입력 매트릭스를 생성하는 입력 생성부;

상기 입력 매트릭스에 대한 코릴레이션(Correlation) 연산에 따라 공간 공분산 매트릭스를 생성하는 공분산 생성부; 및

상기 공간 공분산 매트릭스를 압축하여 압축 매트릭스를 제공하는 압축부를 포함하는 것을 특징으로 하는 음원 분리 시스템.

청구항 12

연산부가 복수의 채널들 중 기준 마이크로 입력되는 기준 입력신호를 시간 및 주파수에 따른 2차원 매트릭스로 변환한 기준 매트릭스 및 상기 기준 입력신호를 제외한 상기 복수의 마이크로들에 해당하는 비교 마이크로들로 입력되는 비교 입력신호들을 시간 및 주파수에 따른 2차원 매트릭스로 변환한 비교 매트릭스들의 채널축으로의 누적을 통해 압축 매트릭스를 제공하는 단계;

분리부가 상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트랄 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하는 단계;

결과부가 상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 결과 매트릭스를 제공하는 단계를 포함하는 음원 분리장치의 동작방법.

청구항 13

연산부가 복수의 채널들 중 기준 마이크로 입력되는 기준 입력신호의 크기를 시간 및 주파수에 따른 2차원 매트릭스로 구현한 기준 매트릭스 및 상기 기준 입력신호의 페이즈 및 상기 기준 마이크로를 제외한 상기 복수의 마이크로들에 해당하는 비교 마이크로들로 입력되는 비교 입력신호들의 페이즈에 기초하여 압축 매트릭스를 제공하는 단계;

분리부가 상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트랄 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하는 단계;

결과부가 상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 결과 매트릭스를 제공하는 단계; 및

음성부가 상기 결과 매트릭스에 기초하여 상기 음원에 상응하는 음성신호를 제공하는 단계를 포함하는 음원 분리시스템의 동작방법.

청구항 14

제13항에 있어서,

상기 음성부는,

상기 스페셜 출력 매트릭스에 기초하여 생성되는 마스크에 따라 기준 매트릭스를 마스킹하여 상기 상응하는 음성신호(VS)를 제공하는 것을 특징으로 하는 음원 분리시스템의 동작방법.

청구항 15

제13항에 있어서,

상기 음성부는,

상기 스페셜 출력 매트릭스에 기초하여 생성되는 마스크에 따라 입력 매트릭스를 마스크 기반 빔포밍하여 상기 상응하는 음성신호(VS)를 제공하는 것을 특징으로 하는 음원 분리시스템의 동작방법.

청구항 16

제15항에 있어서,

상기 마스크 기반 빔포밍은,

상기 스페셜 출력 매트릭스에 기초하여 생성되는 마스크에 따라 상응하는 음성신호(VS)에 대한 방향 벡터를 추정하는 단계;

상기 스페셜 출력 매트릭스에 기초하여 생성되는 마스크에 따라 상응하는 음성신호(VS)에 대한 잡음 공분산 매트릭스를 추정하는 단계;

상기 상응하는 음성신호(VS)에 대한 방향 벡터와 잡음 공분산 매트릭스를 통해 빔포밍 필터를 계산하는 단계를 포함하는 음원 분리시스템의 동작방법.

청구항 17

제3항에 있어서,

상기 성분들 각각의 크기에 기초하여 생성되는 노말라이즈 팩터(Normalize Factor)는 주파수별로 학습 가능한 0 과 1사이의 보정 지수를 포함하는 것을 특징으로 하는 음원 분리장치.

발명의 설명

기술 분야

[0001] 본 발명은 다채널 입력을 기반으로 한 음원 분리장치에 관한 것이다.

배경 기술

[0002] 기존의 음원 다채널 분리방법들에 있어서 공간과 음원정보를 각각 효과적으로 모델링하는데 어려움이 있었다. 최근 이와 같은 문제를 해결하기 위하여 다양한 연구들이 진행되고 있다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) (한국등록특허) 제10-2590887 호 (등록일자, 2023.10.13)

발명의 내용

해결하려는 과제

[0004] 본 발명이 이루고자 하는 기술적 과제는 압축 매트릭스에 기초하여 시간 및 주파수 특성을 나타내는 스펙트럴(Spectral) 출력 매트릭스를 생성하고, 스펙트럴 출력 매트릭스에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하여 공간정보는 주파수 별로 처리하고, 음원정보는 채널별로 독립적으로 처리함으로써 음원분리 성능을 높일 수 있는 음원 분리장치를 제공하는 것이다.

과제의 해결 수단

[0005] 이러한 과제를 해결하기 위하여 본 발명의 실시예에 따른 음원 분리장치는 연산부, 분리부 및 결과부를 포함할 수 있다. 연산부는 기준 및 복수의 비교 마이크 채널로 입력되는 입력신호를 시간, 주파수 및 채널에 따른 3차원 텐서(tensor) 매트릭스로 구현한 기준 매트릭스에 기초하여 3차원 압축 텐서 매트릭스를 제공할 수 있다. 분리부는 상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트럴(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트럴 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다. 결과부는 상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 결과 매트릭스를 제공할 수 있다.

[0006] 일 실시예에 있어서, 상기 연산부는 입력 생성부, 공분산 생성부 및 압축부를 포함할 수 있다. 입력 생성부는 기준 및 복수의 비교 마이크로 입력되는 상기 기준 입력 신호를 단시간 푸리에 변환(Short-time Fourier transform, STFT)을 통해 시간, 주파수 및 채널에 따른 상기 3차원 텐서 입력 매트릭스를 생성할 수 있다. 공분산 생성부는 상기 입력 매트릭스에 대한 상기 채널 차원의 코릴레이션(Correlation) 연산에 따라 공간 공분산 매트릭스를 생성할 수 있다. 압축부는 상기 공간 공분산 매트릭스를 압축하여 3차원 압축 텐서 매트릭스를 제공할 수 있다.

[0007] 일 실시예에 있어서, 상기 공간 공분산 매트릭스에 포함되는 성분들 각각의 값은 상기 성분들 각각의 크기에 기초하여 노말라이즈(Normalized)될 수 있다.

[0008] 일 실시예에 있어서, 상기 분리부는 스펙트럴 분리부 및 스페셜 분리부를 포함할 수 있다. 스펙트럴 분리부는 상기 3차원 압축 텐서 매트릭스를 상기 채널을 기준으로 미리 정해진 기준채널 개수로 나누어 복수의 주파수-시간 매트릭스들을 생성하고, 상기 복수의 주파수-시간 매트릭스들에 기초하여 상기 시간 및 상기 주파수 특징을 나타내는 복수의 주파수-시간 특징 매트릭스들을 산출하고, 상기 복수의 주파수-시간 특징 매트릭스들 및 상기 압축 매트릭스를 잔차 연결(residual connection)을 통해 합성하여 상기 스펙트럴 출력 매트릭스를 제공할 수 있다. 스페셜 분리부는 상기 스펙트럴 출력 매트릭스를 상기 주파수를 기준으로 미리 정해진 기준 주파수 개수로 나누어 복수의 채널-시간 매트릭스들을 생성하고, 상기 복수의 채널-시간 매트릭스들에 기초하여 상기 채널 및 상기 시간 특징을 나타내는 복수의 채널-시간 특징 매트릭스들을 산출하고, 상기 복수의 채널-시간 특징 매트릭스들 및 상기 스펙트럴 출력 매트릭스를 잔차 연결을 통해 합성하여 상기 스페셜 출력 매트릭스를 제공할 수 있다.

[0009] 일 실시예에 있어서, 상기 스펙트럴 분리부 및 스페셜 분리부는 여러 번 반복적으로 누적될 수 있다. 상기 스페셜 출력 매트릭스는 새로운 스펙트럴 분리부의 입력이 될 수 있다.

[0010] 일 실시예에 있어서, 상기 스펙트럴 분리부는 제1 분리부, 제1 산출부 및 제1 합성부를 포함할 수 있다. 제1 분리부는 상기 압축 매트릭스를 상기 채널을 기준으로 나누어 상기 복수의 주파수-시간 매트릭스들을 생성할 수 있다. 제1 산출부는 상기 복수의 주파수-시간 매트릭스들에 기초하여 상기 시간 및 상기 주파수 특징을 나타내는 상기 복수의 주파수-시간 특징 매트릭스들을 산출할 수 있다. 제1 합성부는 상기 복수의 주파수-시간 특징 매트릭스들 및 상기 압축 매트릭스를 잔차 연결을 통해 합성하여 상기 스펙트럴 출력 매트릭스를 제공할 수 있다.

[0011] 일 실시예에 있어서, 상기 제1 산출부는 제1 선택부, 제1 기초부 및 제1 업샘플부를 포함할 수 있다. 제1 선택부는 상기 복수의 주파수-시간 매트릭스들 각각에 포함되는 성분들의 크기에 따라 상기 복수의 주파수-시간 매트릭스들 중 일부를 선택하여 제1 선택 매트릭스들을 선택할 수 있다. 제1 기초부는 상기 제1 선택 매트릭스들에 기초하여 상기 시간 및 상기 주파수 특징을 나타내는 주파수-시간 기초 매트릭스들을 산출할 수 있다. 제1 업샘플부는 상기 주파수-시간 기초 매트릭스들을 업샘플링(up-sampling)하여 상기 복수의 주파수-시간 특징 매트릭스들을 제공할 수 있다.

[0012] 일 실시예에 있어서, 상기 스페셜 분리부는 제2 분리부, 제2 산출부 및 제2 합성부를 포함할 수 있다. 제2 분리

부는 상기 스펙트럴 출력 매트릭스를 상기 주파수를 기준으로 나누어 상기 복수의 채널-시간 매트릭스들을 생성할 수 있다. 제2 산출부는 상기 복수의 채널-시간 매트릭스들에 기초하여 상기 채널 및 상기 시간 특징을 나타내는 상기 복수의 채널-시간 특징 매트릭스들을 산출할 수 있다. 제2 합성부는 상기 복수의 채널-시간 특징 매트릭스들 및 상기 스펙트럴 출력 매트릭스를 잔차 연결을 통해 합성하여 상기 스페셜 출력 매트릭스를 제공할 수 있다.

[0013] 일 실시예에 있어서, 상기 제2 산출부는 제2 선택부, 제2 기초부 및 제2 업샘플부를 포함할 수 있다. 제2 선택부는 상기 복수의 채널-시간 매트릭스들 각각에 포함되는 성분들의 크기에 따라 상기 복수의 채널-시간 매트릭스들 중 일부를 선택하여 제2 선택 매트릭스들을 선택할 수 있다. 제2 기초부는 상기 제2 선택 매트릭스들에 기초하여 상기 채널 및 상기 주파수 특징을 나타내는 채널-시간 기초 매트릭스들을 산출할 수 있다. 제2 업샘플부는 상기 채널-시간 기초 매트릭스들을 업샘플링하여 상기 복수의 채널-시간 특징 매트릭스들을 제공할 수 있다.

[0014] 일 실시예에 있어서, 상기 결과부가 내는 결과 매트릭스는 마스크 매트릭스일 수 있다.

[0015] 일 실시예에 있어서, 상기 결과부는 확장부, 풀링부 및 마스크 생성부를 포함할 수 있다. 확장부는 상기 스페셜 출력 매트릭스를 확장하여 3차원 확장 매트릭스를 제공할 수 있다. 풀링부는 상기 확장 매트릭스를 상기 시간을 기준으로 풀링(pooling)한 상기 주파수 및 상기 채널의 2차원 풀링 매트릭스를 제공할 수 있다. 마스크 생성부는 상기 풀링 매트릭스에 기초하여 상기 주파수 및 상기 채널에 대해서 압축하여 상기 시간 및 상기 주파수에 대한 2차원 마스크 매트릭스를 생성할 수 있다.

[0016] 이러한 과제를 해결하기 위하여 본 발명의 실시예에 따른 음원 분리시스템은 연산부, 분리부, 결과부 및 음성부를 포함할 수 있다. 연산부는 기준 및 복수의 비교 마이크 채널로 입력되는 입력신호를 시간, 주파수 및 채널에 따른 3차원 텐서 매트릭스로 구현한 기준 매트릭스에 기초하여 3차원 텐서 압축 매트릭스를 제공할 수 있다. 분리부는 상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트럴(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트럴 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다. 결과부는 상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리할 결과 매트릭스를 제공할 수 있다. 음성부는 상기 결과 매트릭스에 기초하여 상기 음원에 상응하는 음성신호를 제공할 수 있다.

[0017] 일 실시예에 있어서, 상기 연산부는 입력 생성부, 공분산 생성부 및 압축부를 포함할 수 있다. 입력 생성부는 기준 및 복수의 비교 마이크로 입력되는 상기 기준 및 비교 입력 신호를 단시간 푸리에 변환을 통해 시간, 주파수 및 채널에 따른 상기 3차원 텐서 입력 매트릭스로 생성할 수 있다. 공분산 생성부는 상기 입력 매트릭스에 대한 상기 채널 차원의 코릴레이션(Correlation) 연산에 따라 공간 공분산 매트릭스를 생성할 수 있다. 압축부는 상기 공간 공분산 매트릭스를 압축하여 압축 매트릭스를 제공할 수 있다.

[0018] 이러한 과제를 해결하기 위하여 본 발명의 실시예에 따른 음원 분리장치의 동작방법에서는, 연산부는 기준 및 복수의 비교 마이크 채널로 입력되는 입력신호를 시간, 주파수 및 채널에 따른 3차원 텐서 매트릭스로 구현한 기준 매트릭스에 기초하여 3차원 텐서 압축 매트릭스를 제공할 수 있다. 분리부는 상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트럴(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트럴 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다. 결과부는 상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 결과 매트릭스를 제공할 수 있다.

[0019] 이러한 과제를 해결하기 위하여 본 발명의 실시예에 따른 음원 분리시스템의 동작방법에서는, 연산부는 기준 및 복수의 비교 마이크 채널로 입력되는 입력신호를 시간, 주파수 및 채널에 따른 3차원 텐서 매트릭스로 구현한 기준 매트릭스에 기초하여 압축 매트릭스를 제공할 수 있다. 분리부는 상기 압축 매트릭스에 기초하여 상기 시간 및 상기 주파수 특성을 나타내는 스펙트럴(Spectral) 출력 매트릭스를 생성하고, 상기 스펙트럴 출력 매트릭스에 기초하여 상기 채널 및 상기 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다. 결과부는 상기 스페셜 출력 매트릭스에 기초하여 음원들을 분리하여 결과 매트릭스를 제공할 수 있다. 음성부는 상기 마스크 매트릭스 및 입력 매트릭스(INM)에 기초하여 상기 음원에 상응하는 음성신호를 제공할 수 있다.

[0020] 일 실시예에 있어서, 상기 결과부가 내는 결과 매트릭스는 마스크 매트릭스일 수 있다.

[0021] 일 실시예에 있어서, 상기 음성부는 상기 스페셜 출력 매트릭스에 기초하여 생성되는 마스크 매트릭스에 따라 상기 기준 매트릭스를 마스크하거나 또는 상기 입력 매트릭스를 마스크 기반 빔포밍(beamforming) 하여 상기 음원에 상응하는 음성 신호를 제공할 수 있다.

[0022] 위에서 언급된 본 발명의 기술적 과제 외에도, 본 발명의 다른 특징 및 이점들이 이하에서 기술되거나, 그러한

기술 및 설명으로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

발명의 효과

[0023] 이상과 같은 본 발명에 따르면 다음과 같은 효과가 있다.

[0024] 본 발명에 따른 음원 분리장치는 압축 매트릭스에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하여 공간정보는 주파수 별로 처리하고, 음원정보는 채널별로 독립적으로 처리함으로써 음원분리 성능을 높일 수 있다.

[0025] 이 밖에도, 본 발명의 실시 예들을 통해 본 발명의 또 다른 특징 및 이점들이 새롭게 파악될 수도 있을 것이다.

도면의 간단한 설명

[0026] 도 1은 본 발명의 실시예들에 따른 음원 분리장치를 나타내는 도면이다.

도 2는 도 1의 음원 분리장치의 동작을 설명하기 위한 음원들 및 마이크들을 나타내는 도면이다.

도 3은 도 1의 음원 분리장치에 포함되는 연산부의 일 예를 나타내는 도면이다.

도 4 및 5는 도 1의 음원 분리장치에 포함되는 연산부의 동작을 설명하기 위한 도면들이다.

도 6은 도 1의 음원 분리장치에 포함되는 분리부의 일 예를 나타내는 도면이다.

도 7은 도 6의 분리부에 포함되는 스펙트랄 분리부의 일 예를 나타내는 도면이다.

도 8은 도 7의 스펙트랄 분리부에 포함되는 제1 산출부의 일 예를 나타내는 도면이다.

도 9는 도 6의 분리부에 포함되는 스펙트랄 분리부의 일 동작을 설명하기 위한 도면이다.

도 10은 도 6의 분리부에 포함되는 스페셜 분리부의 일 예를 나타내는 도면이다.

도 11은 도 7의 스페셜 분리부에 포함되는 제2 산출부의 일 예를 나타내는 도면이다.

도 12는 도 6의 분리부에 포함되는 스페셜 분리부의 일 동작을 설명하기 위한 도면이다.

도 13은 도 1의 음원 분리장치에 포함되는 결과부의 일 예를 나타내는 도면이다.

도 14는 도 1의 음원 분리장치에 포함되는 결과부의 동작을 설명하기 위한 도면이다.

도 15는 본 발명의 실시예들에 따른 음원 분리시스템을 나타내는 도면이다.

도 16은 본 발명의 실시예들에 따른 음원 분리장치의 동작방법을 나타내는 순서도이다.

도 17은 본 발명의 실시예들에 따른 음원 분리시스템의 동작방법을 나타내는 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0027] 본 명세서에서 각 도면의 구성 요소들에 참조번호를 부가함에 있어서 동일한 구성 요소들에 한해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 번호를 가지도록 하고 있음에 유의하여야 한다.

[0028] 한편, 본 명세서에서 서술되는 용어의 의미는 다음과 같이 이해되어야 할 것이다.

[0029] 단수의 표현은 문맥상 명백하게 다르게 정의하지 않는 한, 복수의 표현을 포함하는 것으로 이해되어야 하는 것으로, 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다.

[0030] "포함하다" 또는 "가지다" 등의 용어는 하나 또는 그 이상의 다른 특징이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0031] 이하, 첨부되는 도면을 참고하여 상기 문제점을 해결하기 위해 고안된 본 발명의 바람직한 실시예들에 대해 상세히 설명한다.

[0032] 도 1은 본 발명의 실시예들에 따른 음원 분리장치를 나타내는 도면이고, 도 2는 도 1의 음원 분리장치의 동작을 설명하기 위한 음원들 및 마이크들을 나타내는 도면이고, 도 3은 도 1의 음원 분리장치에 포함되는 연산부의 일

예를 나타내는 도면이고, 도 4 및 5는 도 1의 음원 분리장치에 포함되는 연산부의 동작을 설명하기 위한 도면들이다.

[0033] 도 1 내지 5를 참조하면, 본 발명의 실시예에 따른 음원 분리장치(10)는 연산부(100), 분리부(200) 및 결과부(300)를 포함할 수 있다. 연산부(100)는 기준(M1) 및 복수의 비교 마이크(M2부터 MM)로 입력되는 기준 입력신호를 단시간 푸리에 변환을 통해 기준 매트릭스(RM) 및 비교 매트릭스(CM)를 추출하여 채널 차원으로 누적한 시간, 주파수 및 채널에 따른 3차원 텐서 입력 매트릭스(INM)에 기초하여 압축 매트릭스(ABM)를 제공할 수 있다.

[0034] 일 실시예에 있어서, 연산부(100)는 입력 생성부(110), 공분산 생성부(120) 및 압축부(130)를 포함할 수 있다. 입력 생성부(110)는 복수의 마이크들(M)로 입력되는 기준 및 비교 입력 신호를 단시간 푸리에 변환을 통해 생성된 기준 매트릭스(RM) 및 비교 매트릭스(CM)를 채널 차원으로 누적하여 입력 매트릭스(INM)를 생성할 수 있다.

[0035] 예를 들어, 복수의 음원들(SS)은 제1 음원(SS1) 내지 제K 음원(SSK)을 포함할 수 있고, 복수의 마이크들(M)은 제1 마이크(M1) 내지 제M 마이크(MM)를 포함할 수 있다. 제1 음원(SS1) 내지 제K 음원(SSK)으로부터 전달되는 소리신호는 제1 채널(CH1)을 통해서 제1 마이크(M1)에 전달될 수 있다. 이와 동일한 방식으로 제M 채널(CHM)을 통해서 제M 마이크(MM)에 전달될 수 있다. 여기서, 제1 마이크(M1)로 전달되는 소리신호는 기준 입력신호로 설정하는 경우, 제1 마이크(M1)로 전달되는 소리신호의 크기를 시간 및 주파수 축을 따라서 도시한 매트릭스를 기준 매트릭스(RM)일 수 있다.

[0036] 예를 들어, 제1 마이크(M1) 내지 제M 마이크(MM)로 입력되는 신호는 제1 입력신호 내지 제M 입력신호일 수 있고, 제1 입력신호가 기준 입력신호인 경우, 제2 입력신호 내지 제M 입력신호는 비교 입력신호들일 수 있다. 여기서 비교 입력신호의 크기를 시간 및 주파수 축을 따라서 도시한 매트릭스를 비교 매트릭스(CM)일 수 있다.

[0037] 도 4에 도시되는 바와 같이, 기준 매트릭스(RM)와 복수의 비교 매트릭스(CM)를 누적하여 입력 매트릭스(INM)를 생성할 수 있다.

[0038] 공분산 생성부(120)는 입력 매트릭스(INM)에 대한 코릴레이션(Correlation) 연산에 따라 공간 공분산 매트릭스(GBM)를 생성할 수 있다. 예를 들어, M개의 마이크들을 이용하여 입력 매트릭스(INM)를 구성하는 경우, 채널방향으로 M개의 2차원 매트릭스로 구성될 수 있고, 입력 매트릭스(INM)가 채널방향으로 M개의 2차원 매트릭스로 구성되는 경우, 코릴레이션 연산에 따라 공간 공분산 매트릭스(GBM)는 채널방향으로 M*(M+1)개의 2차원 매트릭스로 구현될 수 있다.

[0039] 일 실시예에 있어서, 공간 공분산 매트릭스(GBM)에 포함되는 성분들 각각의 값은 성분들 각각의 크기에 기초하여 노말라이즈(Normalized)될 수 있다.

[0040] [수학식1]

$$\tilde{C}_{tfmm'} = \frac{C_{tfmm'}}{|C_{tfmm'}|^{\beta}}, 0 \leq \beta \leq 1.$$

[0041]

[0042] 여기서, $\tilde{C}_{tfmm'}$ 는 공간 공분산 매트릭스에 포함되는 성분들 각각의 노말라이즈된 이후의 값,

$C_{tfmm'}$ 는 공간 공분산 매트릭스에 포함되는 성분들 각각의 노말라이즈되기 이전의 값, $|C_{tfmm'}|^{\beta}$ 는 노말라이즈 팩터, β 는 인공지능 학습에 따라 결정되는 값일 수 있다.

[0043] 일 실시예에 있어서, 인공지능 학습에 따라 결정되는 값 β 는 주파수 별로 정의되는 β_f 일 수 있다.

[0044] 압축부(130)는 공간 공분산 매트릭스(GBM)를 압축하여 압축 매트릭스(ABM)를 제공할 수 있다. 예를 들어, 도 5에 도시되는 바와 같이, 압축부(130)의 동작은 선형(Linear, 또는 Fully-Connected) 변환 방식에 따라 공간 공분산 매트릭스(GBM)의 차원을 낮추는 과정일 수 있다.

[0045] 도 6은 도 1의 음원 분리장치에 포함되는 분리부의 일 예를 나타내는 도면이고, 7은 도 6의 분리부에 포함되는 스펙트럴 분리부의 일 예를 나타내는 도면이고, 도 8은 도 7의 스펙트럴 분리부에 포함되는 제1 산출부의 일 예

를 나타내는 도면이고, 도 9는 도 6의 분리부에 포함되는 스펙트랄 분리부의 일 동작을 설명하기 위한 도면이다.

[0046] 도 1 내지 9를 참조하면, 분리부(200)는 압축 매트릭스(ABM)에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스(COM)에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다.

[0047] 일 실시예에 있어서, 분리부(200)는 스펙트랄 분리부(210) 및 스페셜 분리부(220)를 포함할 수 있다. 스펙트랄 분리부(210)는 압축 매트릭스(ABM)를 채널을 기준으로 미리 정해진 기준채널 개수로 나누어 복수의 주파수-시간 매트릭스들(FTM)을 생성하고, 복수의 주파수-시간 매트릭스들(FTM)에 기초하여 시간 및 주파수 특징을 나타내는 복수의 주파수-시간 특징 매트릭스들(FTFM)을 산출하고, 복수의 주파수-시간 특징 매트릭스들(FTFM) 및 압축 매트릭스(ABM)를 합성하여 스펙트랄 출력 매트릭스(COM)를 제공할 수 있다.

[0048] 일 실시예에 있어서, 스펙트랄 분리부(210)는 제1 분리부(211), 제1 산출부(212) 및 제1 합성부(213)를 포함할 수 있다. 제1 분리부(211)는 압축 매트릭스(ABM)를 채널을 기준으로 나누어 복수의 주파수-시간 매트릭스들(FTM)을 생성할 수 있다. 예를 들어, 도 9에 도시되는 바와 같이, 압축 매트릭스(ABM)를 채널별로 나누어 복수의 주파수-시간 매트릭스들(FTM)을 생성할 수 있다.

[0049] 제1 산출부(212)는 복수의 주파수-시간 매트릭스들(FTM)에 기초하여 시간 및 주파수 특징을 나타내는 복수의 주파수-시간 특징 매트릭스들(FTFM)을 산출할 수 있다. 일 실시예에 있어서, 제1 산출부(212)는 제1 선택부(216), 제1 기초부(217) 및 제1 업샘플부(218)를 포함할 수 있다. 제1 선택부(216)는 복수의 주파수-시간 매트릭스들(FTM) 각각에 포함되는 성분들의 크기에 따라 복수의 주파수-시간 매트릭스들(FTM) 중 일부를 선택하여 제1 선택 매트릭스들(SEM1)을 선택할 수 있다. 예를 들어, 복수의 주파수-시간 매트릭스의 각각에 포함되는 성분들 크기의 합이 미리 정해진 제1 기준크기보다 큰 경우, 제1 선택 매트릭스로 선택될 수 있고, 복수의 주파수-시간 매트릭스의 각각에 포함되는 성분들 크기의 합이 미리 정해진 제1 기준크기보다 작은 경우, 제1 선택 매트릭스로 선택될 수 없다. 또는 제1 선택부(216)는 복수의 주파수-시간 매트릭스들(FTM)의 인접한 매트릭스 간의 이동 평균을 통해 일부를 선택할 수 있다. 이동 평균을 통한 선택을 하기 전 성분들의 크기에 따른 가중치를 부여하고자 아래와 같은 전역 반응 정규화(Global Response Normalization, GRN)를 통해 복수의 주파수-시간 매트릭스들(FTM)을 정규화할 수 있다.

[0050] [수학식2]

$$[\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{G_M}] \in \mathbb{R}^{T \times G_F \times G_M}$$

$$\mathcal{N}(r_i) = r_i / \sum_{i=1, \dots, G_M} r_i \mathbf{R}_i = \gamma \cdot \mathcal{N}(r_i) \cdot \mathbf{R}_i + \zeta + \mathbf{R}_i$$

[0052] 여기서 $\mathbf{R}_i$ 는 복수의 주파수-시간 매트릭스들(FTM), r_i 는 복수의 주파수-시간 매트릭스들(FTM)에 대해서 L2-norm 값, γ 와 ζ 는 인공지능 학습에 따라 결정되는 값일 수 있다.

[0053] 제1 기초부(217)는 제1 선택 매트릭스들(SEM1)에 기초하여 시간 및 주파수 특징을 나타내는 주파수-시간 기초 매트릭스들(FTIM)을 산출할 수 있다. 예를 들어, 제1 기초부(217)는 학습 데이터를 이용하여 인공지능 학습된 상태일 수 있고, 제1 기초부(217)는 학습된 인공지능에 기초하여 제1 선택 매트릭스에 상응하는 주파수-시간 기초 매트릭스들(FTIM)을 제공할 수 있다.

[0054] 제1 업샘플부(218)는 주파수-시간 기초 매트릭스들(FTIM)을 업샘플링(up-sampling)하여 복수의 주파수-시간 특징 매트릭스들(FTFM)을 제공할 수 있다. 이후, 제1 합성부(213)는 복수의 주파수-시간 특징 매트릭스들(FTFM) 및 압축 매트릭스(ABM)를 합성하여 스펙트랄 출력 매트릭스(COM)를 제공할 수 있다.

[0055] 도 10은 도 6의 분리부에 포함되는 스페셜 분리부의 일 예를 나타내는 도면이고, 도 11은 도 7의 스페셜 분리부에 포함되는 제2 산출부의 일 예를 나타내는 도면이고, 도 12는 도 6의 분리부에 포함되는 스페셜 분리부의 일 동작을 설명하기 위한 도면이다.

[0056] 도 1 내지 12를 참조하면, 스페셜 분리부(220)는 스펙트랄 출력 매트릭스(COM)를 주파수를 기준으로 미리 정해진 기준 주파수 개수로 나누어 복수의 채널-시간 매트릭스들(CTM)을 생성하고, 복수의 채널-시간 매트릭스들

(CTM)에 기초하여 채널 및 시간 특징을 나타내는 복수의 채널-시간 특징 매트릭스들(CTFM)을 산출하고, 복수의 채널-시간 특징 매트릭스들(CTFM) 및 스펙트럴 출력 매트릭스(COM)를 합성하여 스페셜 출력 매트릭스(SOM)를 제공할 수 있다.

[0057] 일 실시예에 있어서, 스페셜 분리부(220)는 제2 분리부(221), 제2 산출부(222) 및 제2 합성부(223)를 포함할 수 있다. 제2 분리부(221)는 스펙트럴 출력 매트릭스(COM)를 주파수를 기준으로 나누어 복수의 채널-시간 매트릭스들(CTM)을 생성할 수 있다. 예를 들어, 도 12에 도시되는 바와 같이, 스펙트럴 출력 매트릭스(COM)를 주파수별로 나누어 복수의 채널-시간 매트릭스들(CTM)을 생성할 수 있다.

[0058] 제2 산출부(222)는 복수의 채널-시간 매트릭스들(CTM)에 기초하여 채널 및 시간 특징을 나타내는 복수의 채널-시간 특징 매트릭스들(CTFM)을 산출할 수 있다. 일 실시예에 있어서, 제2 산출부(222)는 제2 선택부(226), 제2 기초부(227) 및 제2 업샘플부(228)를 포함할 수 있다. 제2 선택부(226)는 복수의 채널-시간 매트릭스들(CTM) 각각에 포함되는 성분들의 크기에 따라 복수의 채널-시간 매트릭스들(CTM) 중 일부를 선택하여 제2 선택 매트릭스들(SEM2)을 선택할 수 있다. 예를 들어, 복수의 채널-시간 매트릭스의 각각에 포함되는 성분들 크기의 합이 미리 정해진 제2 기준크기보다 큰 경우, 제2 선택 매트릭스로 선택될 수 있고, 복수의 채널-시간 매트릭스의 각각에 포함되는 성분들 크기의 합이 미리 정해진 제2 기준크기보다 작은 경우, 제2 선택 매트릭스로 선택될 수 없다. 또는 제1 선택부(216)는 복수의 채널-시간 매트릭스들(CTM)의 인접한 매트릭스 간의 이동평균을 통해 일부를 선택할 수 있다. 이동 평균을 통한 선택을 하기 전 성분들의 크기에 따른 가중치를 부여하고자 아래와 같은 전역 반응 정규화(Global Response Normalization, GRN)을 통해 복수의 주파수-시간 매트릭스들(FTM)을 정규화할 수 있다.

[0059] [수학식3]

$$[\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{G_F}] \in \mathbb{R}^{T \times G_M \times G_F}$$

[0060]

$$\mathcal{N}(r_i) = r_i / \sum_{i=1, \dots, G_M} r_i \quad \mathbf{R}_i = \gamma \cdot \mathcal{N}(r_i) \cdot \mathbf{R}_i + \zeta + \mathbf{R}_i$$

[0061] 여기서 $\mathbf{R}_i$ 는 복수의 채널-시간 매트릭스들(CTM), r_i 는 복수의 채널-시간 매트릭스들(CTM)에 대해서 L2-norm 값, γ 와 ζ 는 인공지능 학습에 따라 결정되는 값일 수 있다.

[0062] 제2 기초부(227)는 제2 선택 매트릭스들(SEM2)에 기초하여 채널 및 주파수 특징을 나타내는 채널-시간 기초 매트릭스들(CTIM)을 산출할 수 있다. 예를 들어, 제2 기초부(227)는 학습 데이터를 이용하여 인공지능 학습된 상태일 수 있고, 제2 기초부(227)는 학습된 인공지능에 기초하여 제2 선택 매트릭스에 상응하는 채널-시간 기초 매트릭스들(CTIM)을 제공할 수 있다.

[0063] 제2 업샘플부(228)는 채널-시간 기초 매트릭스들(CTIM)을 업샘플링하여 복수의 채널-시간 특징 매트릭스들(CTFM)을 제공할 수 있다. 이후, 제2 합성부(223)는 복수의 채널-시간 특징 매트릭스들(CTFM) 및 스펙트럴 출력 매트릭스(COM)를 합성하여 스페셜 출력 매트릭스(SOM)를 제공할 수 있다.

[0064] 도 13은 도 1의 음원 분리장치에 포함되는 결과부의 일 예를 나타내는 도면이고, 도 14는 도 1의 음원 분리장치에 포함되는 결과부의 동작을 설명하기 위한 도면이다.

[0065] 일 실시예에 있어서, 상기 스펙트럴 분리부(210) 및 스페셜 분리부(220)는 반복적으로 누적될 수 있다. 상기 스페셜 출력 매트릭스(SOM)는 새로운 다음 분리부(200)의 입력이 될 수 있다.

[0066] 도 1 내지 14를 참조하면, 결과부(300)는 스페셜 출력 매트릭스(SOM)에 기초하여 음원들(SS)을 분리하여 마스크 매트릭스(REM)를 제공할 수 있다.

[0067] 일 실시예에 있어서, 결과부(300)는 확장부(310), 풀링부(320) 및 마스크 생성부(330)를 포함할 수 있다. 확장부(310)는 스페셜 출력 매트릭스(SOM)를 확장하여 확장 매트릭스(EXM)를 제공할 수 있다. 예를 들어, 도 5에 도시되는 방식과 반대로, 확장부(310)의 동작은 선형(Linear, 또는 Fully-Connected) 변환 방식에 따라 스페셜 출력 매트릭스(SOM)의 차원을 높이는 과정일 수 있다.

[0068] 풀링부(320)는 확장 매트릭스(EXM)를 시간을 기준으로 풀링(pooling)한 풀링 매트릭스(PLM)를 제공할 수 있다. 마스크 생성부(330)는 풀링 매트릭스(PLM)에 기초하여 주파수 및 채널에 대해서 압축하여 마스크(MSK)를 생성할

수 있다. 예를 들어, 마스크 생성부(330)의 동작은 도 14에 도시된 방식으로 진행되고 있으나, 본 발명은 이 방식에만 한정되지 않는다.

[0069] 일 실시예에 있어서, 결과부(300)는 확장 매트릭스(EXM)에 기초하여 마스크 매트릭스(MSK)를 생성하여 제공할 수 있다.

[0070] 본 발명에 따른 음원 분리장치(10)는 압축 매트릭스(ABM)에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스(COM)에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하여 공간정보는 주파수 별로 처리하고, 음원정보는 채널별로 독립적으로 처리함으로써 음원분리 성능을 높일 수 있다.

[0071] 도 15는 본 발명의 실시예들에 따른 음원 분리시스템을 나타내는 도면이다.

[0072] 도 1 내지 15를 참조하면, 본 발명의 실시예에 따른 음원 분리시스템(20)은 연산부(100), 분리부(200), 결과부(300) 및 음성부(400)를 포함할 수 있다. 연산부(100)는 기준(M1) 및 복수의 비교 마이크(M2부터 MM)로 입력되는 기준 입력신호를 단시간 푸리에 변환을 통해 기준 매트릭스(RM) 및 비교 매트릭스(CM)를 추출하여 누적한 시간, 주파수 및 채널에 따른 3차원 텐서 입력 매트릭스(INM)에 기초하여 압축 매트릭스(ABM)를 제공할 수 있다. 분리부(200)는 압축 매트릭스(ABM)에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스(COM)에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다. 결과부(300)는 스페셜 출력 매트릭스(SOM)에 기초하여 분리될 음원들(SS)과 관련하여 마스크 매트릭스(MSK)를 제공할 수 있다. 음성부(400)는 마스크 매트릭스(MSK)에 기초하여 기준 매트릭스(RM)에 마스크하거나 입력 매트릭스(INM)에 마스크 기반 빔포밍을 하여 음원에 상응하는 음성신호(VS)를 제공할 수 있다.

[0073] 일 실시예에 있어서, 연산부(100)는 입력 생성부(110), 공분산 생성부(120) 및 압축부(130)를 포함할 수 있다. 입력 생성부(110)는 복수의 마이크들(M)로 입력되는 기준 및 비교 입력 신호를 단시간 푸리에 변환을 통해 생성된 기준 매트릭스(RM) 및 비교 매트릭스(CM)를 채널 차원으로 누적하여 입력 매트릭스(INM)를 생성할 수 있다. 공분산 생성부(120)는 입력 매트릭스(INM)에 대한 채널 차원에서의 코릴레이션(Correlation) 연산에 따라 공간 공분산 매트릭스(GBM)를 생성할 수 있다. 압축부(130)는 공간 공분산 매트릭스(GBM)를 압축하여 압축 매트릭스(ABM)를 제공할 수 있다.

[0074] 일 실시예에 있어서, 음성부(400)는 결과부(300)에서 제공된 마스크 매트릭스(MSK)를 기준 매트릭스(RM)에 마스크하여 음원에 상응하는 음성신호(VS)를 제공할 수 있다.

[수학식 4]

$$S_{kft} = \mathcal{M}_{kft} * X_1$$

여기서 X_1 는 기준 매트릭스(RM), S_{kft} 는 제 k번째 분리 음원(SSk)의 f,t번째 시간 주파수 성분, $\mathcal{M}_{kft}$ 는 제 k번째 분리 음원에 대한 마스크 매트릭스(MSK)의 f,t번째 시간 주파수 성분 일 수 있다.

[0075]

[0076] 일 실시예에 있어서, 음성부(400)는 결과부(300)에서 제공된 마스크 매트릭스(MSK)를 입력 매트릭스(INM)에 마스크 기반 빔포밍을 하여 음원에 상응하는 음성신호(VS)를 제공할 수 있다.

[수학식5]

$$S_{kft} = \mathbf{w}_{kf}^H \mathbf{x}_{ft}$$

여기서 $\mathbf{x}_{ft}$ 는 입력 매트릭스(INM), S_{kft} 는 제 k번째 분리 음원(SSK)의 f,t번째 시간 주파수 성분, $\mathbf{w}_{kf}^H$ 는 제 k번째 분리 음원에 대한 의 f번째 주파수 성분 빔포밍 필터일 수 있다

이때, 빔포밍 필터 $\mathbf{w}_{kf}^H$ 는 k번째 분리 음원에 대한 마스크 매트릭스(MSX)로부터 추정할 수 있다.

[수학식6]

$$\begin{aligned} \mathbf{w}_{kf} &= \frac{\mathbf{V}_{kf}^{-1} \mathbf{h}_{kf}}{\mathbf{h}_{kf}^H \mathbf{V}_{kf}^{-1} \mathbf{h}_{kf}} \\ \mathbf{h}_{kf} &= \mathcal{P}(\sum_t \mathcal{M}_{kft} \mathbf{x}_{ft} \mathbf{x}_{ft}^H) \\ \mathbf{V}_{kf} &= \sum_t \tilde{\mathcal{M}}_{kft} \mathbf{x}_{ft} \mathbf{x}_{ft}^H \end{aligned}$$

여기서 $\mathbf{w}_{kf}^H$ 는 k번째 분리 음원에 대한 빔포밍 필터, $\mathbf{h}_{kf}$ 는 k번째 분리 음원에 대한 방향(steering) 벡터, $\mathbf{V}_{kf}$ 는 k번째 분리 음원에 대한 잡음 공분산 매트릭스, $\mathcal{P}(\cdot)$ 는 주요 마이겐 벡터를 찾는 연산, $\mathcal{M}_{kft}$ 는 k번째 분리 음원에 대한 마스크 매트릭스(MSX)의 f,t번째 시간 주파수 성분, $\tilde{\mathcal{M}}_{kft}$ 는 k번째 분리 음원에 대한 잡음 마스크 매트릭스의 f,t번째 시간 주파수 성분 일 수 있다.

일 실시예에 있어서, k번째 분리 음원에 대한 잡음 마스크 매트릭스를 구하기 위해서 k번째 분리 음원에 대한 마스크 매트릭스(MSK)를 사용할 수 있다.

[수학식7]

$$\tilde{\mathcal{M}}_{kft} = 1 - \mathcal{M}_{kft}$$

여기서 $\mathcal{M}_{kft}$ 는 k번째 분리 음원에 대한 마스크 매트릭스(MSX)의 f,t번째 시간 주파수 성분, $\tilde{\mathcal{M}}_{kft}$ 는 k번째 분리 음원에 대한 잡음 마스크 매트릭스의 f,t번째 시간 주파수 성분 일 수 있다.

일 실시예에 있어서, k번째 분리 음원에 대한 잡음 마스크 매트릭스를 구하기 위해서 k번째 분리 음원을 제외한

나머지 복수의 마스크 매트릭스(MSK)를 사용할 수 있다.

[수학식8]

$$\tilde{\mathcal{M}}_{kft} = \sum_{k \neq l} \mathcal{M}_{lft}$$

여기서 $\mathcal{M}_{lft}$ 는 1번째 분리 음원에 대한 마스크 매트릭스(MSK)의

f,t번째 시간 주파수 성분, $\tilde{\mathcal{M}}_{kft}$ 는 k번째 분리 음원에 대한 잡음 마스크

매트릭스의 f,t번째 시간 주파수 성분 일 수 있다.

[0082]

도 16은 본 발명의 실시예들에 따른 음원 분리장치의 동작방법을 나타내는 순서도이다.

[0083]

도 1 내지 16을 참조하면, 본 발명의 실시예에 따른 음원 분리장치(10)의 동작방법에서는, 연산부(100)는 복수의 채널들 중 기준 마이크(M1)로 입력되는 기준 입력신호를 시간 및 주파수에 따른 2차원 매트릭스로 변환한 기준 매트릭스(RM) 및 복수의 마이크들(M2부터 MM)에 해당하는 비교 마이크들로 입력되는 비교 입력신호를 복수의 시간 및 주파수에 따른 2차원 매트릭스로 변환한 비교 매트릭스(CM)를 누적하여 생성된 입력 매트릭스에 기초하여 압축 매트릭스(ABM)를 제공할 수 있다(S100). 분리부(200)는 압축 매트릭스(ABM)에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스(COM)에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다(S200). 결과부(300)는 스페셜 출력 매트릭스(SOM)에 기초하여 음원들(SS)을 분리하여 결과 매트릭스(REM)를 제공할 수 있다(S300).

[0084]

일 실시예에 있어서, 분리부(200)는 사용자가 설정한 반복 횟수 N만큼 반복 배치될 수 있다.

[0085]

일 실시예에 있어서, 결과부(300)가 제공하는 결과 매트릭스(REM)는 마스크 매트릭스(MSK)가 될 수 있다.

[0086]

도 17은 본 발명의 실시예들에 따른 음원 분리시스템의 동작방법을 나타내는 순서도이다.

[0087]

도 1 내지 17을 참조하면, 본 발명의 실시예에 따른 음원 분리시스템(20)의 동작방법에서는, 연산부(100)는 복수의 채널들 중 기준 마이크(M1)로 입력되는 기준 입력신호를 시간 및 주파수에 따른 2차원 매트릭스로 변환한 기준 매트릭스(RM) 및 복수의 마이크들(M2부터 MM)에 해당하는 비교 마이크들로 입력되는 비교 입력신호들의 복수의 시간 및 주파수에 따른 2차원 매트릭스로 변환한 비교 매트릭스(CM)를 누적하여 생성된 입력 매트릭스에 기초하여 압축 매트릭스(ABM)를 제공할 수 있다(S100). 분리부(200)는 압축 매트릭스(ABM)에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스(COM)에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공할 수 있다(S200). 결과부(300)는 스페셜 출력 매트릭스(SOM)에 기초하여 음원들(SS)을 분리하여 결과 매트릭스(REM)를 제공할 수 있다(S300). 음성부(400)는 결과 매트릭스(REM)에 기초하여 음원에 상응하는 음성신호(VS)를 제공할 수 있다(S400).

[0088]

본 발명에 따른 음원 분리장치(10)는 압축 매트릭스(ABM)에 기초하여 시간 및 주파수 특성을 나타내는 스펙트랄(Spectral) 출력 매트릭스를 생성하고, 스펙트랄 출력 매트릭스(COM)에 기초하여 채널 및 시간 특성을 나타내는 스페셜(Spatial) 출력 매트릭스를 제공하여 공간정보는 주파수 별로 처리하고, 음원정보는 채널별로 독립적으로 처리함으로써 음원분리 성능을 높일 수 있다.

[0089]

위에서 언급된 본 발명의 기술적 과제 외에도, 본 발명의 다른 특징 및 이점들이 이하에서 기술되거나, 그러한 기술 및 설명으로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

[0090]

부호의 설명

[0091]

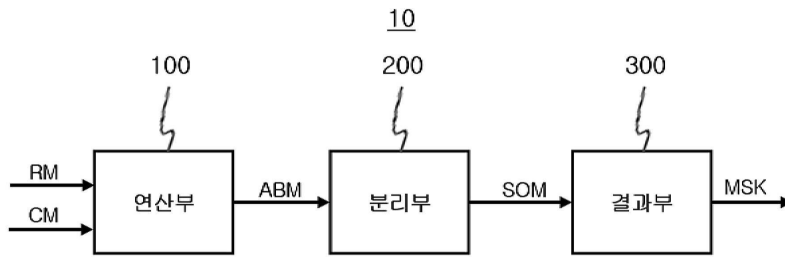
10: 음원 분리장치 100: 연산부

200: 분리부 300: 결과부

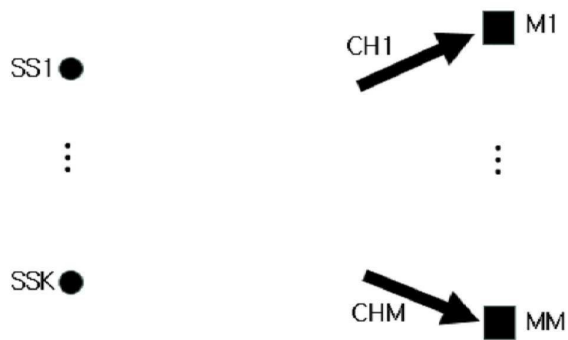
400: 음성부

도면

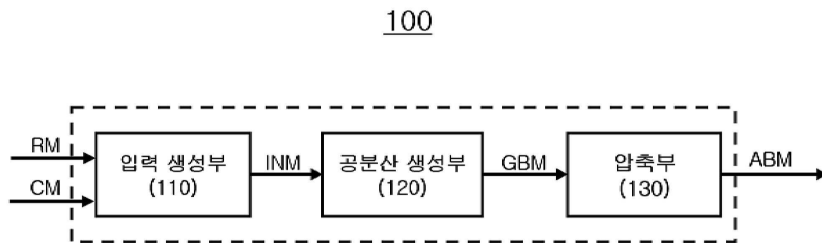
도면1



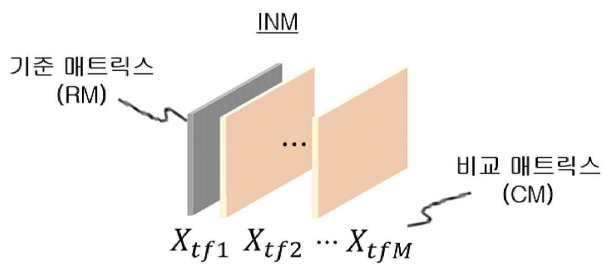
도면2



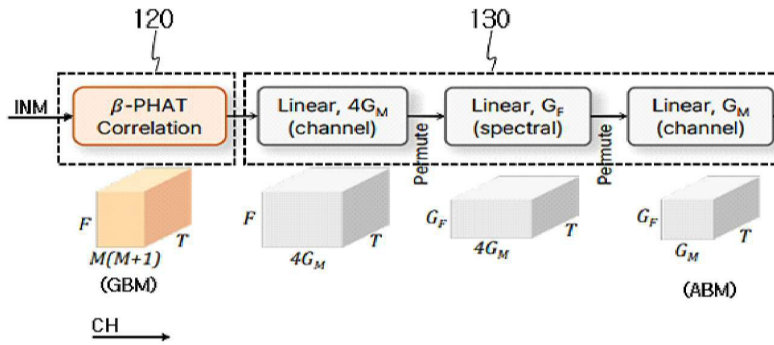
도면3



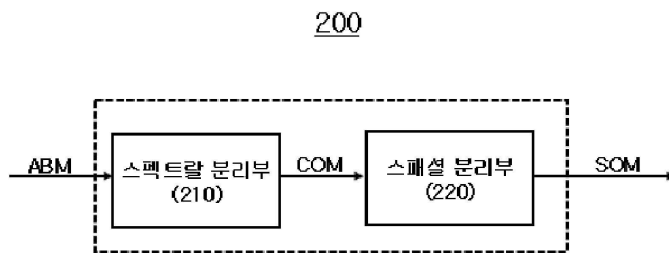
도면4



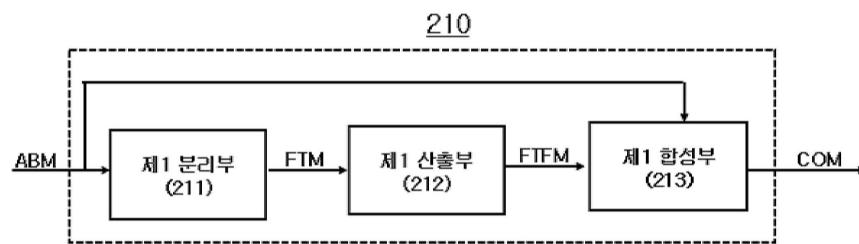
도면5



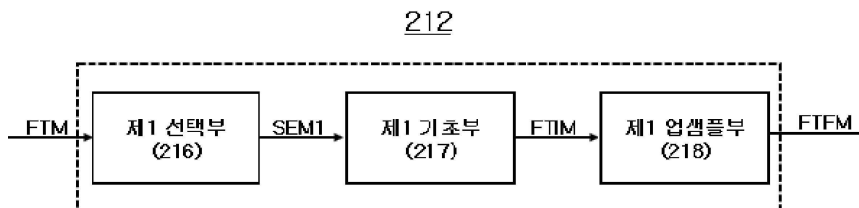
도면6



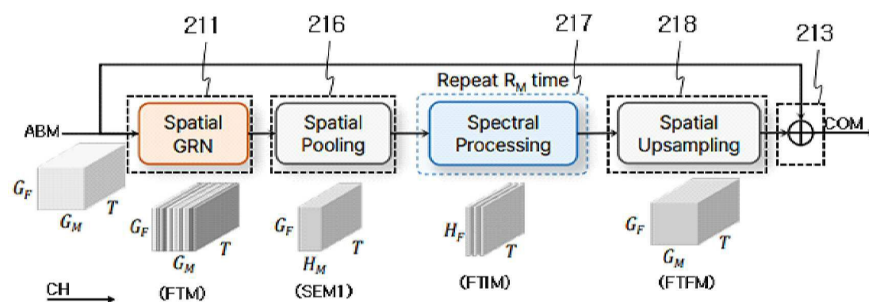
도면7



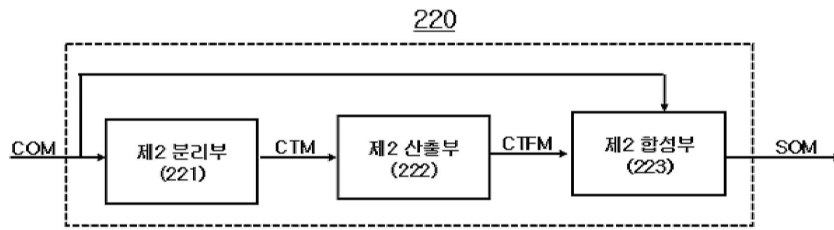
도면8



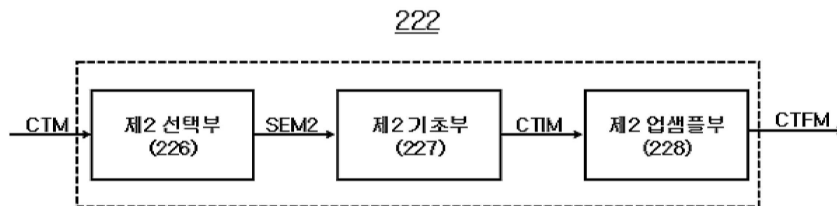
도면9



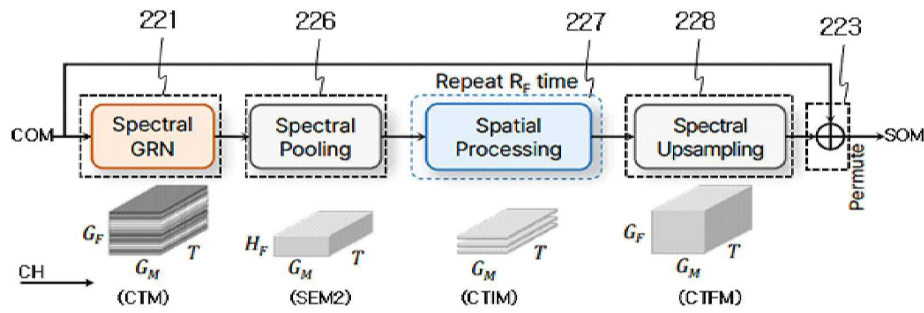
도면10



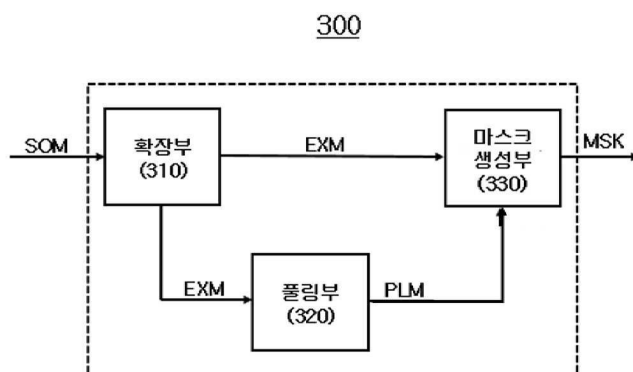
도면11



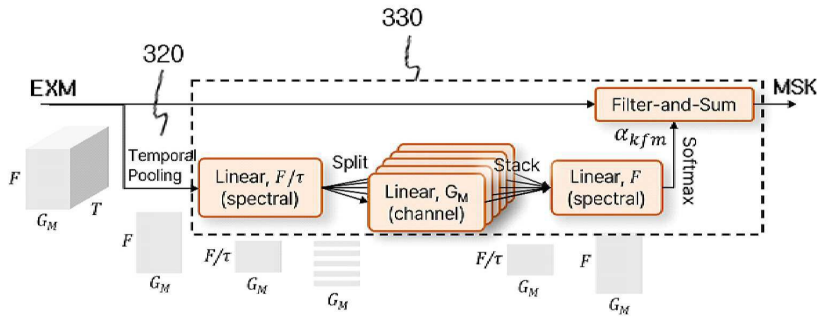
도면12



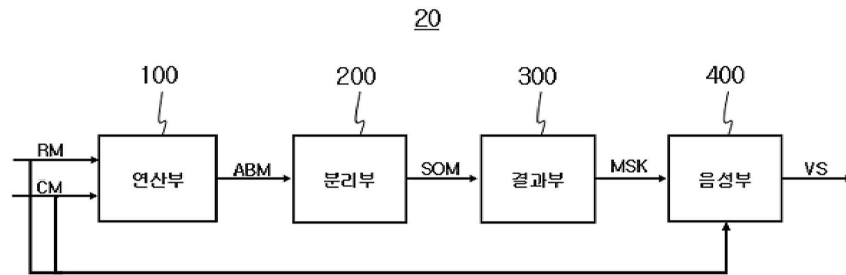
도면13



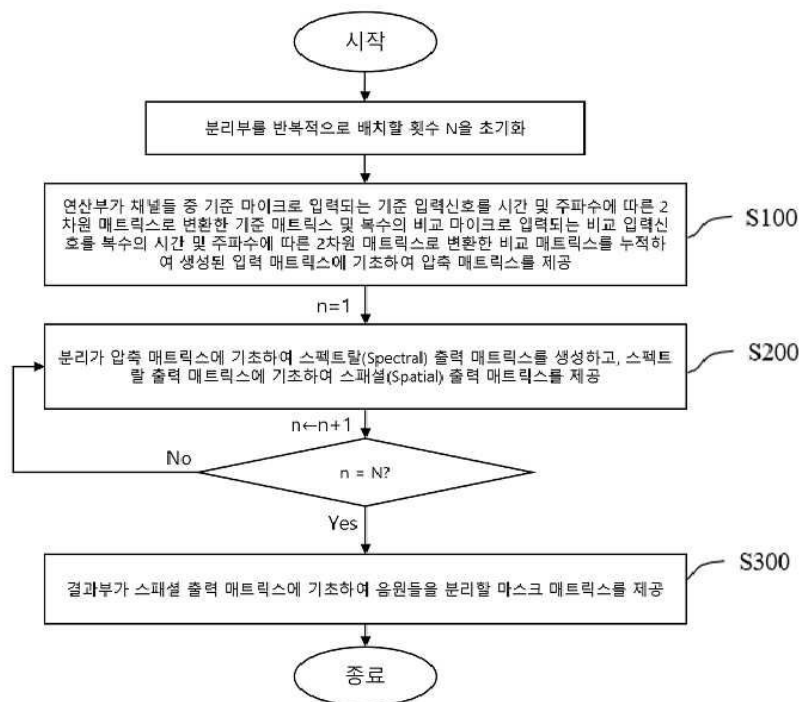
도면14



도면15



도면16



도면17

