



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0142787
(43) 공개일자 2020년12월23일

(51) 국제특허분류(Int. Cl.)
G10L 19/018 (2013.01) G10L 21/02 (2006.01)
(52) CPC특허분류
G10L 19/018 (2013.01)
G10L 21/02 (2020.08)
(21) 출원번호 10-2019-0070097
(22) 출원일자 2019년06월13일
심사청구일자 2019년06월13일

(71) 출원인
네이버 주식회사
경기도 성남시 분당구 불정로 6, 그린팩토리 (정자동)
(72) 발명자
전지수
경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)
김동환
경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)
(뒷면에 계속)
(74) 대리인
양성보

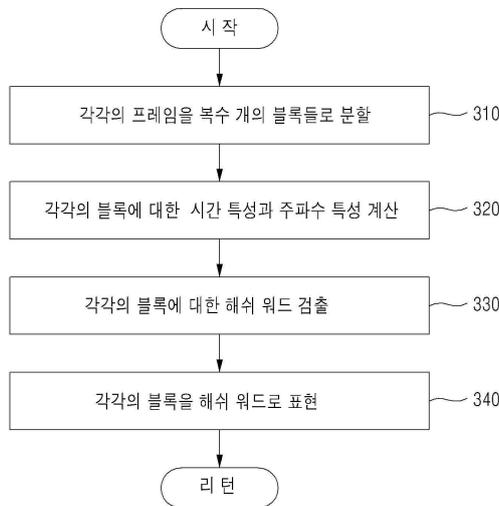
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 멀티미디어 신호 인식을 위한 전자 장치 및 그의 동작 방법

(57) 요약

다양한 실시예들에 따른 멀티미디어 신호 인식을 위한 전자 장치 및 그의 동작 방법은, 검출되는 신호를 복수 개의 프레임들로 분할하고, 프레임들 각각을 복수 개의 블록들로 분할하고, 및 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 블록들 각각을 해쉬 워드로 표현하도록 구성될 수 있다.

대표도 - 도3



(72) 발명자

서동우

경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)

김대황

경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)

박종은

경기도 성남시 분당구 불정로 6(정자동, 그린팩토리)

명세서

청구범위

청구항 1

전자 장치의 동작 방법에 있어서,

검출되는 신호를 복수 개의 프레임들로 분할하는 동작;

상기 프레임들 각각을 복수 개의 블록들로 분할하는 동작; 및

상기 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 상기 블록들 각각을 해쉬 워드로 표현하는 동작을 포함하는 방법.

청구항 2

제 1 항에 있어서,

상기 시간 특성은 상기 블록들 각각에 대한 시간에 따른 에너지 분포를 나타내고,

상기 주파수 특성은 상기 블록들 각각에 대한 주파수에 따른 에너지 분포를 나타내는 방법.

청구항 3

제 2 항에 있어서,

상기 블록들은 교차되는 시간 축과 주파수 축에 의해 정의되는 방법.

청구항 4

제 3 항에 있어서, 상기 해쉬 워드 표현 동작은,

상기 시간 축 상의 양 방향들과 상기 주파수 축 상의 양 방향들의 조합에 따른 미분 값을 기반으로, 상기 블록들 각각에 대한 해쉬 워드를 검출하는 동작; 및

상기 블록들 각각을 상기 검출된 해쉬 워드로 표현하는 동작을 포함하는 방법.

청구항 5

제 4 항에 있어서, 상기 해쉬 워드 검출 동작은,

상기 시간 축의 제 1 시간 방향과 상기 주파수 축의 제 1 주파수 방향에 따른 제 1 미분 값을 계산하는 동작;

상기 시간 축의 제 2 시간 방향과 상기 주파수 축의 제 2 주파수 방향에 따른 제 2 미분 값을 계산하는 동작;

상기 제 1 미분 값과 제 2 미분 값을 합하여, 상기 미분 값을 계산하는 동작; 및

상기 미분 값을 기반으로, 상기 해쉬 워드를 검출하는 동작을 포함하는 방법.

청구항 6

제 5 항에 있어서, 상기 해쉬 워드 검출 동작은,

상기 제 1 시간 방향과 제 2 주파수 방향에 따른 제 3 미분 값을 계산하는 동작; 또는

상기 제 2 시간 방향과 제 1 주파수 방향에 따른 제 4 미분 값을 계산하는 동작 중 적어도 어느 하나를 더 포함하는 방법.

청구항 7

제 6 항에 있어서, 상기 미분 값 계산 동작은,

상기 제 1 미분 값 및 제 2 미분 값과 함께, 상기 제 3 미분 값 또는 제 4 미분 값 중 적어도 어느 하나를 합하여, 상기 미분 값을 계산하는 방법.

청구항 8

제 5 항에 있어서, 상기 해쉬 워드 검출 동작은,

상기 미분 값을 임계 값과 비교하여, 상기 미분 값에 대응하는 상기 해쉬 워드를 검출하는 방법.

청구항 9

제 8 항에 있어서,

상기 임계 값은, 상기 프레임들 중 현재 프레임에서 미분 값과 비교하기 위하여, 상기 프레임들 중 이전 프레임에서 미분 값으로 결정되는 방법.

청구항 10

제 8 항에 있어서, 상기 해쉬 워드 검출 동작은,

상기 미분 값이 상기 임계 값을 초과하면, 상기 해쉬 워드를 1로 검출하고,

상기 미분 값이 상기 임계 값 이하이면, 상기 해쉬 워드를 0으로 검출하는 방법.

청구항 11

전자 장치에 있어서,

입력 모듈; 및

상기 입력 모듈을 통해 검출되는 신호의 핑거프린트를 생성하도록 구성되는 프로세서를 포함하고,

상기 프로세서는,

검출되는 신호를 복수 개의 프레임들로 분할하고,

상기 프레임들 각각을 복수 개의 블록들로 분할하고,

상기 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 상기 블록들 각각을 해쉬 워드로 표현하도록 구성되는 장치.

청구항 12

제 11 항에 있어서,

상기 시간 특성은 상기 블록들 각각에 대한 시간에 따른 에너지 분포를 나타내고,

상기 주파수 특성은 상기 블록들 각각에 대한 주파수에 따른 에너지 분포를 나타내는 장치.

청구항 13

제 12 항에 있어서,

상기 블록들은 교차되는 시간 축과 주파수 축에 의해 정의되는 장치.

청구항 14

제 13 항에 있어서, 상기 프로세서는,

상기 시간 축 상의 양 방향들과 상기 주파수 축 상의 양 방향들의 조합에 따른 미분 값을 기반으로, 상기 블록들 각각에 대한 해쉬 워드를 검출하고,

상기 블록들 각각을 상기 검출된 해쉬 워드로 표현하도록 구성되는 장치.

청구항 15

제 14 항에 있어서, 상기 프로세서는,

상기 시간 축의 제 1 시간 방향과 상기 주파수 축의 제 1 주파수 방향에 따른 제 1 미분 값을 계산하고,

상기 시간 축의 제 2 시간 방향과 상기 주파수 축의 제 2 주파수 방향에 따른 제 2 미분 값을 계산하고,

상기 제 1 미분 값과 제 2 미분 값을 합하여, 상기 미분 값을 계산하고,

상기 미분 값을 기반으로, 상기 해쉬 워드를 검출하도록 구성되는 장치.

청구항 16

제 15 항에 있어서, 상기 프로세서는,

상기 제 1 시간 방향과 제 2 주파수 방향에 따른 제 3 미분 값 또는 상기 제 2 시간 방향과 제 1 주파수 방향에 따른 제 4 미분 값 중 적어도 어느 하나를 계산하고,

상기 제 1 미분 값 및 제 2 미분 값과 함께, 상기 제 3 미분 값 또는 제 4 미분 값 중 적어도 어느 하나를 합하여, 상기 미분 값을 계산하도록 구성되는 장치.

청구항 17

제 15 항에 있어서, 상기 프로세서는,

상기 미분 값을 임계 값과 비교하여, 상기 미분 값에 대응하는 상기 해쉬 워드를 검출하는 장치.

청구항 18

제 17 항에 있어서,

상기 임계 값은, 상기 프레임들 중 현재 프레임에서 미분 값과 비교하기 위하여, 상기 프레임들 중 이전 프레임에서 미분 값으로 결정되는 장치.

청구항 19

제 17 항에 있어서, 상기 프로세서는,
 상기 미분 값이 상기 임계 값을 초과하면, 상기 해쉬 워드를 1로 검출하고,
 상기 미분 값이 상기 임계 값 이하이면, 상기 해쉬 워드를 0으로 검출하는 장치.

청구항 20

비-일시적(non-transitory) 컴퓨터-판독 가능(computer-readable) 저장(storage) 매체(medium)에 있어서,
 검출되는 신호를 복수 개의 프레임들로 분할하는 동작;
 상기 프레임들 각각을 복수 개의 블록들로 분할하는 동작; 및
 상기 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 상기 블록들 각각을 해쉬 워드로 표현하는 동작
 을 실행하기 위한 하나 이상의 프로그램들을 저장하기 위한 비-일시적 컴퓨터-판독 가능 저장 매체.

발명의 설명

기술 분야

[0001] 다양한 실시예들은 멀티미디어 신호 인식을 위한 전자 장치 및 그의 동작 방법에 관한 것이다.

배경 기술

[0002] 최근 통신 환경에서, QbH(query by humming) 서비스가 제공되고 있다. 이러한 서비스를 제공하고 위해, 서버는 많은 양의 오디오 파일들과 관련된 정보를 저장하고 있다. 클라이언트가 검색하고자 하는 멜로디를 갖는 오디오 신호를 서버에 전송하고, 서버는 오디오 신호와 가장 유사한 오디오 파일에 대한 정보를 클라이언트에 피드백한다. 이를 위해, 서버는 오디오 신호를 많은 양의 오디오 파일들과 각각 비교해야 한다. 일반적으로 서버는 오디오 파일들과 비교하기 위하여, 오디오 신호의 특성을 기반으로 오디오 핑거프린트를 생성한다. 이 때 서버는 오디오 신호에서 단지 주파수 방향에 따른 에너지의 분포 만을 기반으로, 핑거프린트를 생성한다. 이로 인하여, 오디오 신호 내 잡음의 주파수 성분에 의해, 핑거프린트에 왜곡이 발생될 수 있다. 이에 따라, 서버에서 핑거프린트를 이용하여 오디오 파일들을 검색하는 데 있어서, 정확도가 저하되는 문제점이 있다.

발명의 내용

해결하려는 과제

[0003] 다양한 실시예들은 멀티미디어 신호로부터 생성되는 핑거프린트의 왜곡을 최소화할 수 있는 전자 장치 및 그의 동작 방법을 제공한다.

[0004] 다양한 실시예들은 핑거프린트를 이용하여 멀티미디어 파일들을 검색하는 데 있어서, 정확도를 향상시킬 수 있는 전자 장치 및 그의 동작 방법을 제공한다.

과제의 해결 수단

[0005] 다양한 실시예들에 따른 전자 장치의 동작 방법은, 검출되는 신호를 복수 개의 프레임들로 분할하는 동작, 상기 프레임들 각각을 복수 개의 블록들로 분할하는 동작, 및 상기 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 상기 블록들 각각을 해쉬 워드로 표현하는 동작을 포함할 수 있다.

[0006] 다양한 실시예들에 따른 전자 장치는, 입력 모듈, 및 상기 입력 모듈을 통해 검출되는 신호의 핑거프린트를 생성하도록 구성되는 프로세서를 포함할 수 있다.

[0007] 다양한 실시예들에 따르면, 상기 프로세서는, 검출되는 신호를 복수 개의 프레임들로 분할하고, 상기 프레임들 각각을 복수 개의 블록들로 분할하고, 상기 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 상기 블

록들 각각을 해쉬 워드로 표현하도록 구성될 수 있다.

[0008] 다양한 실시예들에 따른 비-일시적(non-transitory) 컴퓨터-판독 가능(computer-readable) 저장(storage) 매체(medium)는, 검출되는 신호를 복수 개의 프레임들로 분할하는 동작, 상기 프레임들 각각을 복수 개의 블록들로 분할하는 동작, 및 상기 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 상기 블록들 각각을 해쉬 워드로 표현하는 동작을 실행하기 위한 하나 이상의 프로그램들을 저장할 수 있다.

발명의 효과

[0009] 다양한 실시예들에 따르면, 전자 장치가 검출되는 신호의 주파수 특성 뿐 아니라 시간 특성을 기반으로 해쉬 워드를 표현함으로써, 신호의 핑거프린트를 생성할 수 있다. 이를 통해, 전자 장치는 핑거프린트의 왜곡을 최소화할 수 있다. 즉 신호로부터 주파수 특성 또는 시간 특성 중 어느 하나를 검출하는 데 있어서 예러가 발생하더라도, 주파수 특성 또는 시간 특성 중 다른 하나를 기반으로 핑거프린트의 왜곡을 무효화할 수 있다. 이에 따라, 전자 장치가 핑거프린트를 생성하는 데 있어서, 정확도가 향상될 수 있다. 나아가, 전자 장치가 핑거프린트를 이용하여 멀티미디어 파일들을 검색하는 데 있어서, 정확도가 향상될 수 있다.

도면의 간단한 설명

[0010] 도 1은 다양한 실시예들에 따른 전자 장치를 도시하는 도면이다.
 도 2는 다양한 실시예들에 따른 전자 장치의 동작 방법을 도시하는 도면이다.
 도 3은 도 2의 핑거프린트 생성 동작을 도시하는 도면이다.
 도 4a, 도 4b, 도 5a, 도 5b, 도 6a, 도 6b, 도 6c, 도 6d, 도 7a, 도 7b, 도 7c 및 도 7d는 멀티미디어 신호의 특성을 설명하기 위한 도면들이다.
 도 8은 멀티미디어 신호의 핑거프린트를 설명하기 위한 도면이다.
 도 9a, 도 9b, 도 9c 및 도 9d는 핑거프린트에서 도출되는 특성을 설명하기 위한 도면들이다.

발명을 실시하기 위한 구체적인 내용

[0011] 이하, 본 문서의 다양한 실시예들이 첨부된 도면을 참조하여 설명된다.

[0013] 다양한 실시예들에 따른 전자 장치는 멀티미디어 신호를 인식하도록 구현될 수 있다. 예를 들면, 멀티미디어 신호는 오디오 신호 또는 비디오 신호 중 적어도 어느 하나를 포함할 수 있다. 일 실시예에 따르면, 전자 장치는 복수 개의 멀티미디어 파일들을 저장하고 있으며, 멀티미디어 신호를 기반으로, 멀티미디어 파일들 중 적어도 어느 하나를 검출할 수 있다. 여기서, 전자 장치는 멀티미디어 신호의 특성과 멀티미디어 파일들의 특성들 간의 유사도를 검출하고, 유사도에 따라 멀티미디어 파일들 중 적어도 어느 하나를 검출할 수 있다. 다양한 실시예들에 따르면, 전자 장치는 멀티미디어 신호의 특성을 기반으로, 멀티미디어 신호의 핑거프린트(fingerprint)를 생성할 수 있다.

[0015] 도 1은 다양한 실시예들에 따른 전자 장치(100)를 도시하는 도면이다.

[0016] 도 1을 참조하면, 다양한 실시예들에 따른 전자 장치(100)는, 입력 모듈(110), 출력 모듈(120), 메모리(130) 또는 프로세서(140) 중 적어도 어느 하나를 포함할 수 있다. 어떤 실시예에서는, 전자 장치(100)의 구성 요소들 중 적어도 어느 하나가 생략되거나, 전자 장치(100)에 하나 이상의 다른 구성 요소가 추가될 수 있다.

[0017] 입력 모듈(110)은 전자 장치(100)의 적어도 하나의 구성 요소에 사용될 명령 또는 신호를 입력할 수 있다. 입력 모듈(110)은, 사용자가 전자 장치(100)에 직접적으로 명령 또는 데이터를 입력하도록 구성되는 입력 장치, 주변 환경을 감지하여 신호를 발생하도록 구성되는 센서 장치, 또는 외부 장치(미도시)와 통신을 수행하여 명령 또는 데이터를 수신하도록 구성되는 통신 장치 중 적어도 어느 하나를 포함할 수 있다. 예를 들면, 입력 장치는 마이크로폰(microphone), 카메라(camera), 마우스(mouse) 또는 키보드(keyboard) 중 적어도 어느 하나를 포함할 수 있다. 예를 들면, 통신 장치는 유선 통신 모듈 또는 무선 통신 모듈 중 적어도 어느 하나를 포함하며, 무선 통신 모듈은 원거리 통신망 또는 근거리 통신망 중 적어도 어느 하나를 통하여 외부 장치와 통신을 수행할 수 있

다.

- [0018] 출력 모듈(120)은 전자 장치(100)의 외부로 정보를 제공할 수 있다. 출력 모듈(120)은 정보를 시각적으로 제공하는 표시 장치, 정보를 청각적으로 제공하는 오디오 장치 또는 정보를 송신하도록 구성되는 통신 장치 중 적어도 어느 하나를 포함할 수 있다. 예를 들면, 표시 장치는 디스플레이, 홀로그램 장치 또는 프로젝터 중 적어도 어느 하나를 포함할 수 있다. 어떤 실시예에서, 표시 장치는 터치를 감지하도록 설정된 터치 회로(touch circuitry) 또는 터치에 의해 발생하는 힘의 세기를 측정하도록 설정된 센서 회로 중 적어도 어느 하나를 포함할 수 있다. 예를 들면, 통신 장치는 유선 통신 모듈 또는 무선 통신 모듈 중 적어도 어느 하나를 포함하며, 무선 통신 모듈은 원거리 통신망 또는 근거리 통신망 중 적어도 어느 하나를 통하여 외부 장치와 통신을 수행할 수 있다.
- [0019] 메모리(130)는 전자 장치(100)의 적어도 하나의 구성 요소에 의해 사용되는 다양한 데이터를 저장할 수 있다. 예를 들면, 메모리(130)는 휘발성 메모리 또는 비휘발성 메모리 중 적어도 어느 하나를 포함할 수 있다. 데이터는 프로그램 또는 이와 관련된 입력 데이터 또는 출력 데이터를 포함할 수 있다. 일 실시예에 따르면, 메모리(130)는 복수 개의 멀티미디어 파일들 또는 각각의 멀티미디어 파일에 대한 메타 정보 중 적어도 어느 하나를 저장할 수 있다.
- [0020] 프로세서(140)는 메모리(130)의 프로그램을 실행하여, 전자 장치(100)의 적어도 하나의 구성 요소를 제어할 수 있고, 데이터 처리 또는 연산을 수행할 수 있다. 프로세서(140)는 입력 모듈(110)을 통하여, 멀티미디어 신호를 검출할 수 있다. 이 때 멀티미디어 신호는 오디오 신호 또는 비디오 신호 중 적어도 어느 하나를 포함할 수 있다. 프로세서(140)는 멀티미디어 신호를 복수 개의 프레임들로 분할할 수 있다. 이 때 프레임들은 시간 축과 주파수 축에 의해 정의될 수 있으며, 상호에 중첩될 수 있다. 각각의 프레임은 시간과 주파수에 따른 각각의 에너지 분포를 가질 수 있다. 프로세서(140)는 각각의 프레임에 대한 핑거프린트를 생성할 수 있다. 이 때 핑거프린트는 해쉬 워드(hash word)들의 배열로 표현될 수 있다.
- [0021] 프로세서(140)가 각각의 프레임을 복수 개의 블록들로 분할할 수 있다. 여기서, 블록들은 동일한 프레임 내에 배열되며, 이로 인해 시간 또는 주파수 중 적어도 어느 하나가 상이할 수 있다. 프로세서(140)가 각각의 블록에 대한 시간 특성과 주파수 특성을 계산할 수 있다. 시간 특성은 시간에 따른 에너지 분포를 나타내며, 주파수 특성은 시간에 따른 에너지 분포를 나타낼 수 있다. 이를 위해, 프로세서(140)는 각각의 블록에 대하여, 시간 축상의 양 방향들과 주파수 축상의 양 방향들의 조합(D)에 따른 미분 값을 계산할 수 있다. 프로세서(140)는 각각의 블록에 대한 시간 특성과 주파수 특성을 기반으로, 각각의 블록에 대한 해쉬 워드를 검출할 수 있다. 이 때 프로세서(140)는 각각의 블록에 대한 에너지 미분 값을 기반으로, 각각의 블록에 대한 해쉬 워드를 검출할 수 있다. 예를 들면, 프로세서(140)는 미분 값을 임계 값과 비교하여, 미분 값에 대응하는 해쉬 워드를 검출할 수 있다. 여기서, 임계 값은 현재 프레임에서 미분 값과 비교하기 위하여, 이전 프레임에서 미분 값으로 결정될 수 있다.
- [0023] 다양한 실시예들에 따르면, 전자 장치(100)가 검출되는 신호의 주파수 특성 뿐 아니라 시간 특성을 기반으로 해쉬 워드를 표현함으로써, 신호의 핑거프린트를 생성할 수 있다. 이를 통해, 전자 장치(100)는 핑거프린트의 왜곡을 최소화할 수 있다. 즉 신호로부터 주파수 특성 또는 시간 특성 중 어느 하나를 검출하는 데 있어서 에러가 발생하더라도, 주파수 특성 또는 시간 특성 중 다른 하나를 기반으로 핑거프린트의 왜곡을 무효화할 수 있다. 예를 들면, 신호가 잡음을 포함하더라도, 잡음의 주파수 성분에 의한 핑거프린트의 왜곡이 무효화될 수 있다. 이에 따라, 전자 장치(100)가 핑거프린트를 생성하는 데 있어서, 정확도가 향상될 수 있다. 나아가, 전자 장치(100)가 핑거프린트를 이용하여 멀티미디어 파일들을 검색하는 데 있어서, 정확도가 향상될 수 있다.
- [0025] 다양한 실시예들에 따른 전자 장치(100)는, 입력 모듈(110), 및 입력 모듈(110)을 통해 검출되는 신호의 핑거프린트를 생성하도록 구성되는 프로세서(140)를 포함할 수 있다.
- [0026] 다양한 실시예들에 따르면, 프로세서(140)는, 검출되는 신호를 복수 개의 프레임들로 분할하고, 프레임들 각각을 복수 개의 블록들로 분할하고, 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 블록들 각각을 해쉬 워드로 표현하도록 구성될 수 있다.
- [0027] 다양한 실시예들에 따르면, 시간 특성은 블록들 각각에 대한 시간에 따른 에너지 분포를 나타내고, 주파수 특성

은 블록들 각각에 대한 주파수에 따른 에너지 분포를 나타낼 수 있다.

- [0028] 다양한 실시예들에 따르면, 블록들은 교차되는 시간 축과 주파수 축에 의해 정의될 수 있다.
- [0029] 다양한 실시예들에 따르면, 프로세서(140)는, 시간 축 상의 양 방향들과 주파수 축 상의 양 방향들의 조합에 따른 미분 값을 기반으로, 블록들 각각에 대한 해쉬 워드를 검출하고, 블록들 각각을 검출된 해쉬 워드로 표현하도록 구성될 수 있다.
- [0030] 다양한 실시예들에 따르면, 프로세서(140)는, 시간 축의 제 1 시간 방향과 주파수 축의 제 1 주파수 방향에 따른 제 1 미분 값을 계산하고, 시간 축의 제 2 시간 방향과 주파수 축의 제 2 주파수 방향에 따른 제 2 미분 값을 계산하고, 제 1 미분 값과 제 2 미분 값을 합하여, 미분 값을 계산하고, 미분 값을 기반으로, 해쉬 워드를 검출하도록 구성될 수 있다.
- [0031] 다양한 실시예들에 따르면, 프로세서(140)는, 제 1 시간 방향과 제 2 주파수 방향에 따른 제 3 미분 값 또는 제 2 시간 방향과 제 1 주파수 방향에 따른 제 4 미분 값 중 적어도 어느 하나를 계산하고, 제 1 미분 값 및 제 2 미분 값과 함께, 제 3 미분 값 또는 제 4 미분 값 중 적어도 어느 하나를 합하여, 미분 값을 계산하도록 구성될 수 있다.
- [0032] 다양한 실시예들에 따르면, 프로세서(140)는, 미분 값을 임계 값과 비교하여, 미분 값에 대응하는 해쉬 워드를 검출할 수 있다.
- [0033] 다양한 실시예들에 따르면, 임계 값은, 프레임들 중 현재 프레임에서 미분 값과 비교하기 위하여, 프레임들 중 이전 프레임에서 미분 값으로 결정될 수 있다.
- [0034] 다양한 실시예들에 따르면, 프로세서(140)는, 미분 값이 임계 값을 초과하면, 해쉬 워드를 1로 검출하고, 미분 값이 임계 값 이하이면, 해쉬 워드를 0으로 검출할 수 있다.
- [0036] 도 2는 다양한 실시예들에 따른 전자 장치(100)의 동작 방법을 도시하는 도면이다. 도 4a, 도 4b, 도 5a, 도 5b, 도 6a, 도 6b, 도 6c, 도 6d, 도 7a, 도 7b, 도 7c 및 도 7d는 멀티미디어 신호의 특성을 설명하기 위한 도면들이다. 도 8은 멀티미디어 신호의 핑거프린트를 설명하기 위한 도면이다. 도 9a, 도 9b, 도 9c 및 도 9d는 핑거프린트에서 도출되는 특성을 설명하기 위한 도면들이다.
- [0037] 도 2를 참조하면, 전자 장치(100)는 210 동작에서 멀티미디어 신호를 검출할 수 있다. 프로세서(140)는 입력 모듈(110)을 통하여, 멀티미디어 신호를 검출할 수 있다. 이 때 멀티미디어 신호는 오디오 신호 또는 비디오 신호 중 적어도 어느 하나를 포함할 수 있다.
- [0038] 예를 들면, 오디오 신호는, 도 4a, 도 4b, 도 5a 또는 도 5b에 도시된 바와 같이 시간 영역(time domain)과 주파수 영역(frequency domain)에 따른 에너지의 분포를 나타내는 스펙트럼(spectrum)으로 표현될 수 있다. 여기서, 오디오 신호의 스펙트럼은, 도 4a 또는 도 4b에 도시된 바와 같이 입체적으로 표현될 수 있으며, 도 5a 또는 도 5b에 도시된 바와 같이 평면적으로 표현될 수 있다. 이 때 오디오 신호가 잡음을 포함하지 않는 경우, 오디오 신호의 스펙트럼은, 도 4a 및 도 5a에 도시된 바와 같이 표현될 수 있다. 한편, 오디오 신호가 잡음을 포함하는 경우, 오디오 신호의 스펙트럼은, 도 4b 및 도 5b에 도시된 바와 같이 표현될 수 있다.
- [0039] 전자 장치(100)는 220 동작에서 멀티미디어 신호를 복수 개의 프레임들로 분할할 수 있다. 이 때 프레임들은 시간 축과 주파수 축에 의해 정의될 수 있으며, 상호에 중첩될 수 있다. 각각의 프레임은 시간과 주파수에 따른 각각의 에너지 분포를 가질 수 있다. 일 실시예에 따르면, 프로세서(140)는 프레임링 회로, 푸리에 변환 회로 또는 절댓값 회로 중 적어도 어느 하나를 포함할 수 있다. 프레임링 회로는 프레임들을 구분할 수 있다. 푸리에 변환 회로는 각각의 프레임에서 시간과 주파수에 따른 에너지 분포에 대한 푸리에 계수를 검출할 수 있다. 절댓값 회로는 푸리에 계수의 절대값을 계산할 수 있다.
- [0040] 예를 들면, 오디오 신호로부터 분할된 프레임들 중 어느 하나에 대하여, 각각의 시간에서 주파수에 따른 에너지 분포가 도 6a, 도 6b, 도 6c 및 도 6d에 도시된 바와 같이 표현될 수 있다. 여기서, 각각의 시간에서 주파수에 따른 에너지 분포는, 오디오 신호가 잡음을 포함하는지의 여부에 따라, 다르게 표현될 수 있다. 한편, 오디오 신호로부터 분할된 프레임들 중 어느 하나에 대하여, 각각의 주파수에서 시간에 따른 에너지 분포가 도 7a, 도 7b, 도 7c 및 도 7d에 도시된 바와 같이 표현될 수 있다. 여기서, 각각의 주파수에서 시간에 따른 에너지 분포는, 오디오 신호가 잡음을 포함하는지의 여부에 따라, 다르게 표현될 수 있다. 그러나, 각각의 주파수에서 시

간에 따른 에너지 분포에 있어서의 특성, 즉 스펙트럴 피크(spectral peak)들은 오디오 신호가 잡음을 포함하는지의 여부와 관계 없이, 유지될 수 있다.

[0041] 전자 장치(100)는 230 동작에서 각각의 프레임에 대한 핑거프린트를 생성할 수 있다. 프로세서(140)는 각각의 프레임의 시간과 주파수에 따른 에너지 분포를 핑거프린트로 표현할 수 있다. 이 때 핑거프린트는 해쉬 워드(hash word)들의 배열로 표현될 수 있다. 일 실시예에 따르면, 프로세서(140)는 프레임들 중 인접한 두 개의 프레임들의 에너지 분포들을 비교하여, 각각의 프레임에 대한 핑거프린트를 생성할 수 있다. 이를 위해, 프레임들로부터, 핑거프린트를 생성하고자 하는 현재 프레임과 현재 프레임에 인접한 이전 프레임이 정의될 수 있다.

[0042] 예를 들면, 프로세서(140)는, 도 8에 도시된 바와 같이 핑거프린트를 생성할 수 있다. 핑거프린트에서, 백색(white color)의 화소는, 비트값이 1인 해쉬 워드를 나타내고, 흑색(black color)의 화소는, 비트값이 0인 해쉬 워드를 나타낼 수 있다. 여기서, 오디오 신호가 잡음을 포함하지 않는 경우, 핑거프린트는, 도 8의 (a)에 도시된 바와 같이 표현될 수 있다. 한편, 오디오 신호가 잡음을 포함하지 않는 경우, 핑거프린트는, 도 8의 (b)에 도시된 바와 같이 표현될 수 있다. 한편, 오디오 신호의 잡음이, 도 8의 (c)에 도시된 바와 같이 표현될 수 있다.

[0044] 도 3은 도 2의 핑거프린트 생성 동작을 도시하는 도면이다.

[0045] 도 3을 참조하면, 전자 장치(100)는 310 동작에서 각각의 프레임을 복수 개의 블록들로 분할할 수 있다. 프로세서(140)가 각각의 프레임을 복수 개의 블록들로 분할할 수 있다. 이 때 블록들은 시간 축과 주파수 축에 의해 정의될 수 있다. 여기서, 블록들은 동일한 프레임 내에 배열되며, 이로 인해 시간 또는 주파수 중 적어도 어느 하나가 상이할 수 있다. 일 실시예에 따르면, 프로세서(140)는 대역 분할 회로를 포함하고, 대역 분할 회로가 미리 정해진 시간 대역 및 미리 정해진 주파수 대역을 기반으로, 각각의 프레임을 복수 개의 블록들로 분할할 수 있다. 예를 들면, 대역 분할 회로는 복수 개의 선택기들을 포함하며, 각각의 선택기가 각각의 블록을 검출할 수 있다.

[0046] 전자 장치(100)는 320 동작에서 각각의 블록에 대한 시간 특성과 주파수 특성을 검출할 수 있다. 프로세서(140)가 각각의 블록에 대한 시간 특성과 주파수 특성을 계산할 수 있다. 시간 특성은 시간에 따른 에너지 분포를 나타내며, 주파수 특성은 시간에 따른 에너지 분포를 나타낼 수 있다. 이를 위해, 프로세서(140)는 시간 축 및 주파수 축에 대한 에너지 기울기의 방향을 고려하여, 각각의 블록에 대한 에너지 미분 값을 계산할 수 있다.

[0047] 다양한 실시예들에 따르면, 프로세서(140)는 각각의 블록에 대하여, 시간 축 상의 양 방향들과 주파수 축 상의 양 방향들의 조합(D)에 따른 미분 값을 계산할 수 있다. 이 때 프로세서(140)는, 하기의 [수학식 1]과 같이 미분 값을 계산할 수 있다. 즉 프로세서(140)는 각각의 블록에 대하여, 시간 축 상의 양 방향들 중 적어도 어느 하나와 주파수 축의 양 방향들 중 적어도 어느 하나의 조합(D)에 따른 적어도 하나의 미분 값을 합할 수 있다. 여기서, 시간 축 상의 양 방향들은 제 1 시간 방향과 그에 반대되는 제 2 시간 방향으로 정의되고, 주파수 축 상의 양 방향들은 제 1 주파수 방향과 그에 반대되는 제 2 주파수 방향으로 정의될 수 있다. 예를 들면, 제 1 시간 방향과 제 2 시간 방향은 1과 -1로 표현되고, 제1 주파수 방향과 제 2 주파수 방향은 1과 -1로 표현될 수 있으며, 이러한 경우 시간 축 상의 양 방향들과 주파수 축 상의 양 방향들의 조합(D)은, (1, 1), (1, -1), (-1, 1) 또는 (-1, -1) 중 적어도 어느 하나를 포함할 수 있다.

수학식 1

$$\sum_{(i,j) \in D} E[n+i, m+j] - E[n, m]$$

[0048]

[0049] 여기서, n은 시간 대역을 나타내고, m은 주파수 대역을 나타내며, E는 밴드합(band-summed) 에너지를 나타낼 수 있다.

[0050] 일 실시예에 따르면, 프로세서(140)는 제 1 시간 방향과 제 1 주파수 방향에 따른 제 1 미분 값을 계산하고, 제 2 시간 방향과 제 2 주파수 방향에 따른 제 2 미분 값을 계산할 수 있다. 이를 통해, 프로세서(140)는 제 1 미분 값과 제 2 미분 값을 합하여, 각각의 블록에 대한 에너지 미분 값을 계산할 수 있다. 일 예로, 시간 축 상의

양 방향들과 주파수 축 상의 양 방향들의 조합(D)은, (-1, 1)와 (-1, -1)를 포함할 수 있다.

[0051] 다른 실시예에 따르면, 프로세서(140)는 제 1 시간 방향과 제 1 주파수 방향에 따른 제 1 미분 값을 계산하고, 제 2 시간 방향과 제 2 주파수 방향에 따른 제 2 미분 값을 계산할 수 있다. 아울러, 프로세서(140)는 제 1 시간 방향과 제 2 주파수 방향에 따른 제 3 미분 값 또는 제 2 시간 방향과 제 1 주파수 방향에 따른 제 4 미분 값 중 적어도 어느 하나를 계산할 수 있다. 이를 통해, 프로세서(140)는 제 1 미분 값 및 제 2 미분 값과 함께, 제 3 미분 값 또는 제 4 미분 값 중 적어도 어느 하나를 합하여, 각각의 블록에 대한 에너지 미분 값을 계산할 수 있다. 다른 예로, 시간 축 상의 양 방향들과 주파수 축 상의 양 방향들의 조합(D)은, (1, 1), (-1, 1) 및 (-1, -1)를 포함하거나, (1, -1), (-1, 1) 및 (-1, -1)를 포함할 수 있다.

[0052] 전자 장치(100)는 330 동작에서 각각의 블록에 대한 시간 특성과 주파수 특성을 기반으로, 각각의 블록에 대한 해쉬 워드를 검출할 수 있다. 이 때 프로세서(140)는 각각의 블록에 대한 에너지 미분 값을 기반으로, 각각의 블록에 대한 해쉬 워드를 검출할 수 있다. 예를 들면, 프로세서(140)는, 하기의 [수학식 2]와 같이 미분 값을 임계 값과 비교하여, 미분 값에 대응하는 해쉬 워드를 검출할 수 있다. 여기서, 임계 값은, 하기의 [수학식 2]와 같이 현재 프레임에서 미분 값과 비교하기 위하여, 이전 프레임에서 미분 값으로 결정될 수 있다. 즉 프로세서(140)는, 미분 값이 임계 값을 초과하면, 해쉬 워드를 1로 검출하고, 미분 값이 임계 값 미만이면, 해쉬 워드를 0으로 검출할 수 있다.

수학식 2

$$H[n, m] = \begin{cases} 1 & \text{if } \left(\sum_{(i,j) \in D} E[n+1+i, m+j] - E[n+1, m] \right) > \left(\sum_{(i,j) \in D} E[n+i, m+j] - E[n, m] \right) \\ 0 & \text{if } \left(\sum_{(i,j) \in D} E[n+1+i, m+j] - E[n+1, m] \right) \leq \left(\sum_{(i,j) \in D} E[n+i, m+j] - E[n, m] \right) \end{cases}$$

[0053]

[0054] 전자 장치(100)는 340 동작에서 각각의 블록을 해쉬 워드로 표현할 수 있다. 프로세서(140)가, 도 8에 도시된 바와 같이 각각의 블록을 해쉬 워드로 표현할 수 있다. 예를 들면, 각각의 블록에 대하여, 프로세서(140)는 해쉬 워드가 1이면, 백색의 화소로 표현하고, 해쉬 워드가 0이면, 흑색의 화소로 표현할 수 있다.

[0055] 이를 통해, 각각의 프레임에 대한 핑거프린트가 생성될 수 있다. 예를 들면, 각각의 프레임에 대한 핑거프린트로부터, 각각의 시간에서 주파수에 따른 에너지 분포가 도 9a, 도 9b, 도 9c 및 도 9d에 도시된 바와 같이 도출될 수 있다. 여기서, 각각의 시간에서 주파수에 따른 에너지 분포는, 오디오 신호가 잡음을 포함하는지의 여부에 따라, 다르게 표현될 수 있다. 그러나, 각각의 시간에서 주파수에 따른 에너지 분포에 있어서의 특성, 즉 스펙트럴 피크들은 오디오 신호가 잡음을 포함하는지의 여부와 관계 없이, 유지될 수 있다. 바꿔 말하면, 오디오 신호가 잡음을 포함하더라도, 스펙트럴 피크들이 명확하게 표현될 수 있다.

[0057] 다양한 실시예들에 따르면, 전자 장치(100)가 검출되는 신호의 주파수 특성 뿐 아니라 시간 특성을 기반으로 해쉬 워드를 표현함으로써, 신호의 핑거프린트를 생성할 수 있다. 이를 통해, 전자 장치(100)는 핑거프린트의 왜곡을 최소화할 수 있다. 즉 신호로부터 주파수 특성 또는 시간 특성 중 어느 하나를 검출하는 데 있어서 예러가 발생하더라도, 주파수 특성 또는 시간 특성 중 다른 하나를 기반으로 핑거프린트의 왜곡을 무효화할 수 있다. 예를 들면, 신호가 잡음을 포함하더라도, 잡음의 주파수 성분에 의한 핑거프린트의 왜곡이 무효화될 수 있다. 이에 따라, 전자 장치(100)가 핑거프린트를 생성하는 데 있어서, 정확도가 향상될 수 있다. 나아가, 전자 장치(100)가 핑거프린트를 이용하여 멀티미디어 파일들을 검색하는 데 있어서, 정확도가 향상될 수 있다.

[0059] 다양한 실시예들에 따른 전자 장치(100)의 동작 방법은, 검출되는 신호를 복수 개의 프레임들로 분할하는 동작, 프레임들 각각을 복수 개의 블록들로 분할하는 동작, 및 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 블록들 각각을 해쉬 워드로 표현하는 동작을 포함할 수 있다.

[0060] 다양한 실시예들에 따르면, 시간 특성은 블록들 각각에 대한 시간에 따른 에너지 분포를 나타내고, 주파수 특성은 블록들 각각에 대한 주파수에 따른 에너지 분포를 나타낼 수 있다.

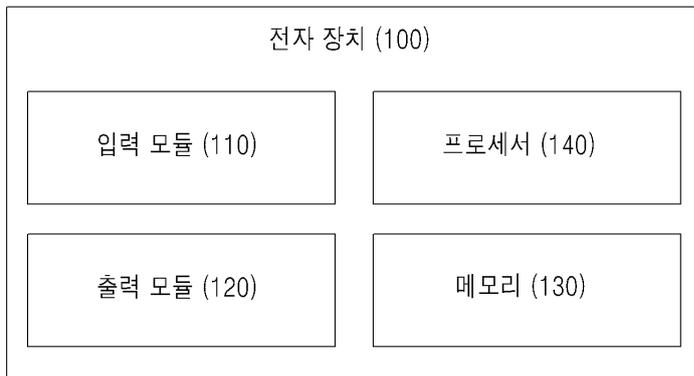
- [0061] 다양한 실시예들에 따르면, 블록들은 교차되는 시간 축과 주파수 축에 의해 정의될 수 있다.
- [0062] 다양한 실시예들에 따르면, 해쉬 워드 표현 동작은, 시간 축 상의 양 방향들과 주파수 축 상의 양 방향들의 조합에 따른 미분 값을 기반으로, 블록들 각각에 대한 해쉬 워드를 검출하는 동작, 및 블록들 각각을 검출된 해쉬 워드로 표현하는 동작을 포함할 수 있다.
- [0063] 다양한 실시예들에 따르면, 해쉬 워드 검출 동작은, 시간 축의 제 1 시간 방향과 주파수 축의 제 1 주파수 방향에 따른 제 1 미분 값을 계산하는 동작, 시간 축의 제 2 시간 방향과 주파수 축의 제 2 주파수 방향에 따른 제 2 미분 값을 계산하는 동작, 제 1 미분 값과 제 2 미분 값을 합하여, 미분 값을 계산하는 동작, 및 미분 값을 기반으로, 해쉬 워드를 검출하는 동작을 포함할 수 있다.
- [0064] 다양한 실시예들에 따르면, 해쉬 워드 검출 동작은, 제 1 시간 방향과 제 2 주파수 방향에 따른 제 3 미분 값을 계산하는 동작, 또는 제 2 시간 방향과 제 1 주파수 방향에 따른 제 4 미분 값을 계산하는 동작 중 적어도 어느 하나를 더 포함할 수 있다.
- [0065] 다양한 실시예들에 따르면, 미분 값 계산 동작은, 제 1 미분 값 및 제 2 미분 값과 함께, 제 3 미분 값 또는 제 4 미분 값 중 적어도 어느 하나를 합하여, 미분 값을 계산할 수 있다.
- [0066] 다양한 실시예들에 따르면, 해쉬 워드 검출 동작은, 미분 값을 임계 값과 비교하여, 미분 값에 대응하는 해쉬 워드를 검출할 수 있다.
- [0067] 다양한 실시예들에 따르면, 임계 값은, 프레임들 중 현재 프레임에서 미분 값과 비교하기 위하여, 프레임들 중 이전 프레임에서 미분 값으로 결정될 수 있다.
- [0068] 다양한 실시예들에 따르면, 해쉬 워드 검출 동작은, 미분 값이 임계 값을 초과하면, 해쉬 워드를 1로 검출하고, 미분 값이 임계 값 이하이면, 해쉬 워드를 0으로 검출할 수 있다.
- [0070] 본 문서의 다양한 실시예들은 기기(machine)(예: 전자 장치(110, 130))에 의해 읽을 수 있는 저장 매체(storage medium)(예: 메모리(219, 239))에 저장된 하나 이상의 명령어들을 포함하는 소프트웨어로서 구현될 수 있다. 예를 들면, 기기의 프로세서(예: 프로세서(221, 241))는, 저장 매체로부터 저장된 하나 이상의 명령어들 중 적어도 하나의 명령어를 호출하고, 그것을 실행할 수 있다. 이것은 기기가 호출된 적어도 하나의 명령어에 따라 적어도 하나의 기능을 수행하도록 운영되는 것을 가능하게 한다. 하나 이상의 명령어들은 컴파일러에 의해 생성된 코드 또는 인터프리터에 의해 실행될 수 있는 코드를 포함할 수 있다. 기기로 읽을 수 있는 저장매체는, 비일시적(non-transitory) 저장 매체의 형태로 제공될 수 있다. 여기서, '비일시적'은 저장매체가 실제(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다.
- [0071] 다양한 실시예들에 따른 비-일시적 컴퓨터-판독 가능(computer-readable) 저장 매체는, 검출되는 신호를 복수 개의 프레임들로 분할하는 동작, 프레임들 각각을 복수 개의 블록들로 분할하는 동작, 및 블록들 각각에 대한 시간 특성과 주파수 특성을 기반으로, 블록들 각각을 해쉬 워드로 표현하는 동작을 실행하기 위한 하나 이상의 프로그램들을 저장할 수 있다.
- [0073] 본 문서의 다양한 실시예들 및 이에 사용된 용어들은 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 해당 실시 예의 다양한 변경, 균등물, 및/또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및/또는 B 중 적어도 하나", "A, B 또는 C" 또는 "A, B 및/또는 C 중 적어도 하나" 등의 표현은 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. "제 1", "제 2", "첫째" 또는 "둘째" 등의 표현들은 해당 구성요소들을, 순서 또는 중요도에 상관없이 수식할 수 있고, 한 구성요소를 다른 구성요소와 구분하기 위해 사용될 뿐 해당 구성요소들을 한정하지 않는다. 어떤(예: 제 1) 구성요소가 다른(예: 제 2) 구성요소에 "(기능적으로 또는 통신적으로) 연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로 연결되거나, 다른 구성요소(예: 제 3 구성요소)를 통하여 연결될 수 있다.
- [0074] 본 문서에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구성된 유닛을 포함하며, 예를 들면,

로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수 있다. 예를 들면, 모듈은 ASIC(application-specific integrated circuit)으로 구성될 수 있다.

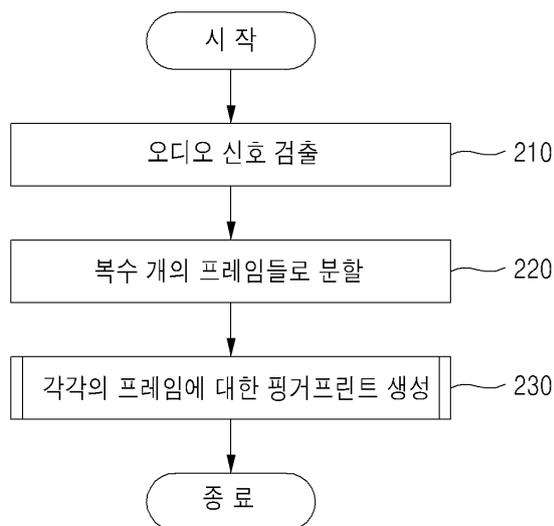
[0075] 다양한 실시예들에 따르면, 기술한 구성요소들의 각각의 구성요소(예: 모듈 또는 프로그램)는 단수 또는 복수의 개체를 포함할 수 있다. 다양한 실시예들에 따르면, 전술한 해당 구성요소들 중 하나 이상의 구성요소들 또는 동작들이 생략되거나, 또는 하나 이상의 다른 구성요소들 또는 동작들이 추가될 수 있다. 대체적으로 또는 추가적으로, 복수의 구성요소들(예: 모듈 또는 프로그램)은 하나의 구성요소로 통합될 수 있다. 이런 경우, 통합된 구성요소는 복수의 구성요소들 각각의 구성요소의 하나 이상의 기능들을 통합 이전에 복수의 구성요소들 중 해당 구성요소에 의해 수행되는 것과 동일 또는 유사하게 수행할 수 있다. 다양한 실시예들에 따르면, 모듈, 프로그램 또는 다른 구성요소에 의해 수행되는 동작들은 순차적으로, 병렬적으로, 반복적으로, 또는 휴리스틱하게 실행되거나, 동작들 중 하나 이상이 다른 순서로 실행되거나, 생략되거나, 또는 하나 이상의 다른 동작들이 추가될 수 있다.

도면

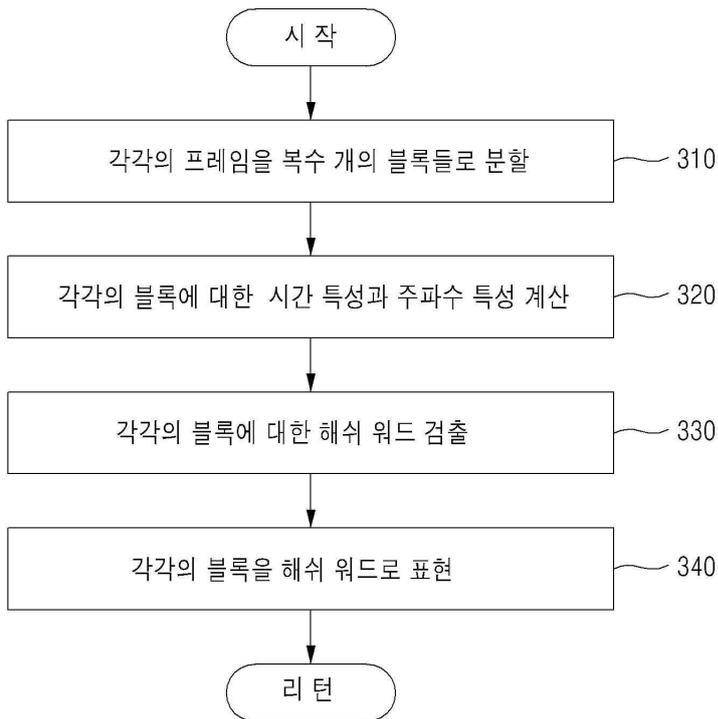
도면1



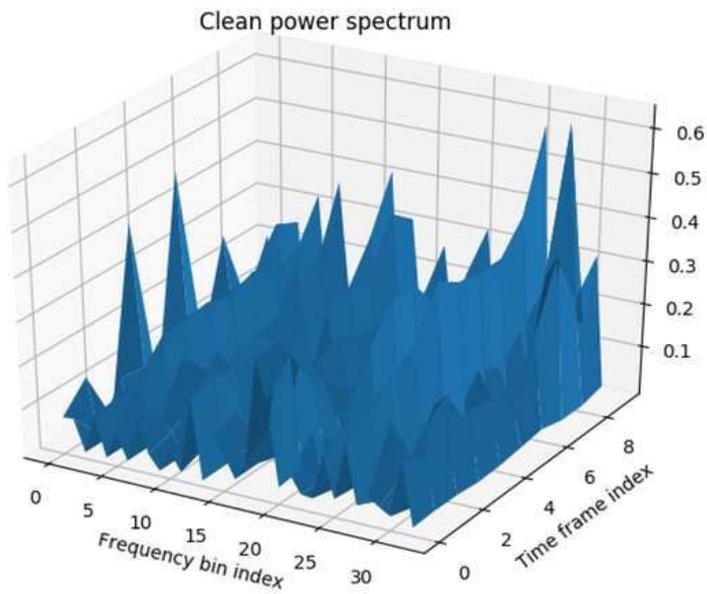
도면2



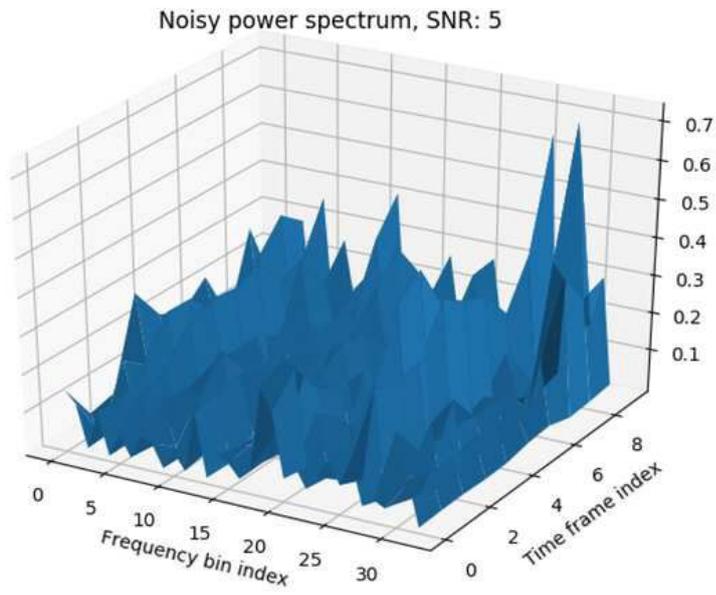
도면3



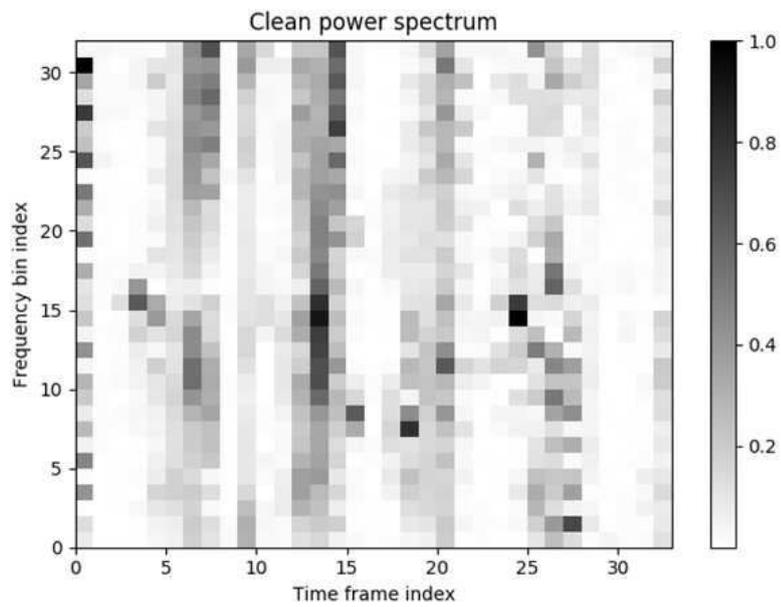
도면4a



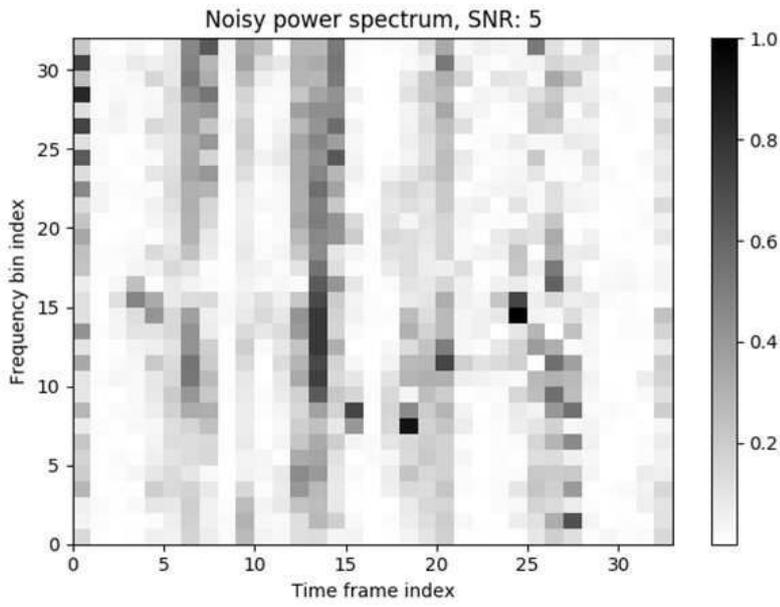
도면4b



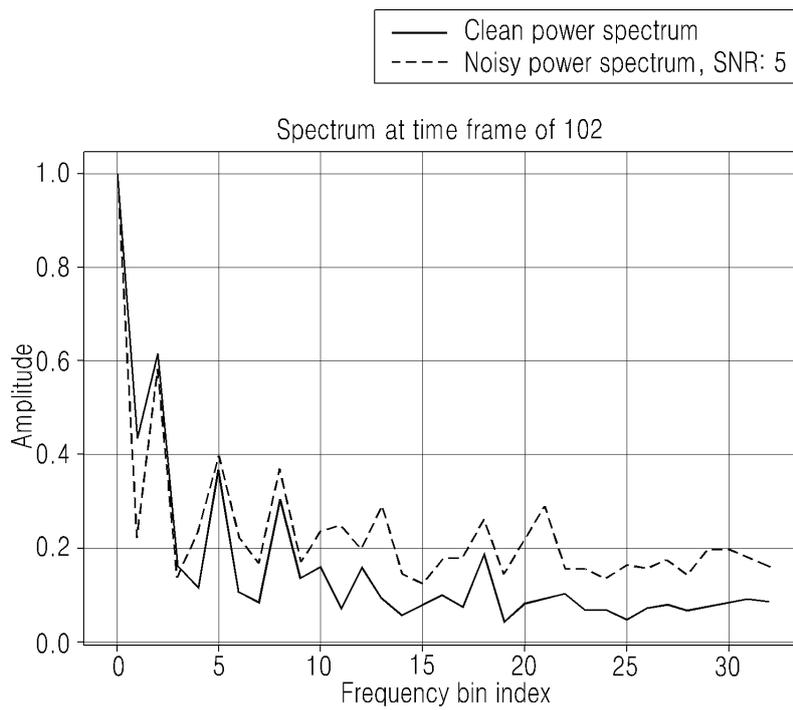
도면5a



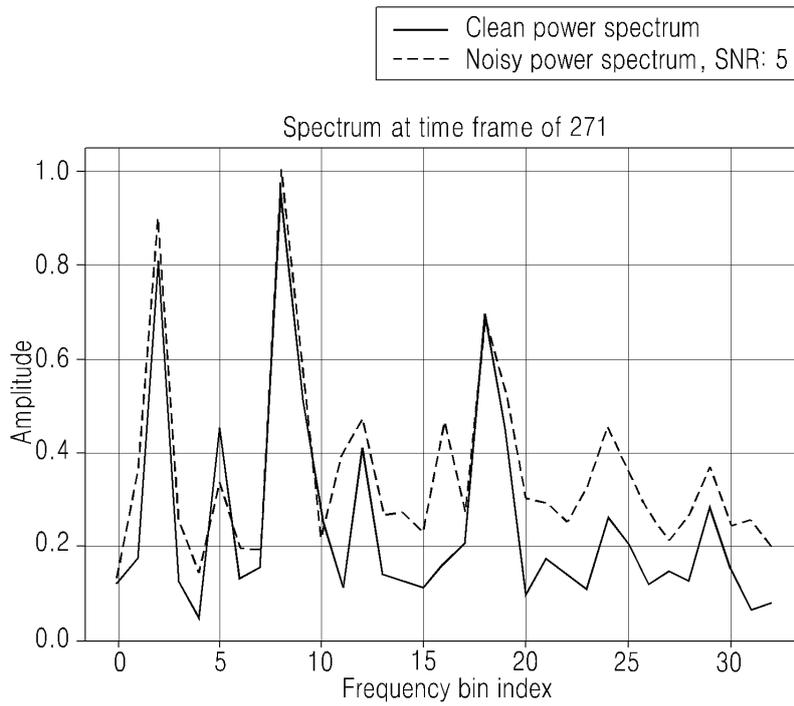
도면5b



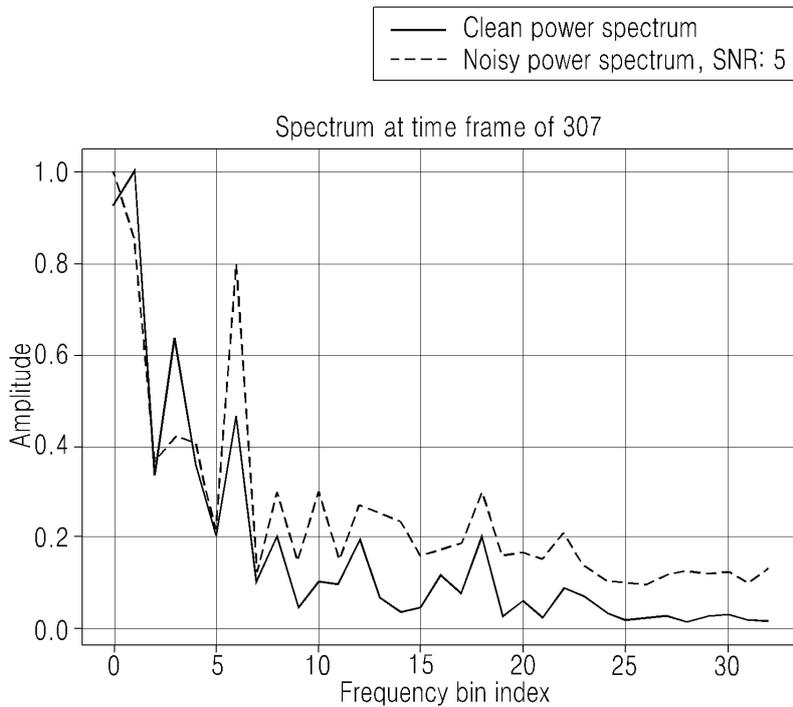
도면6a



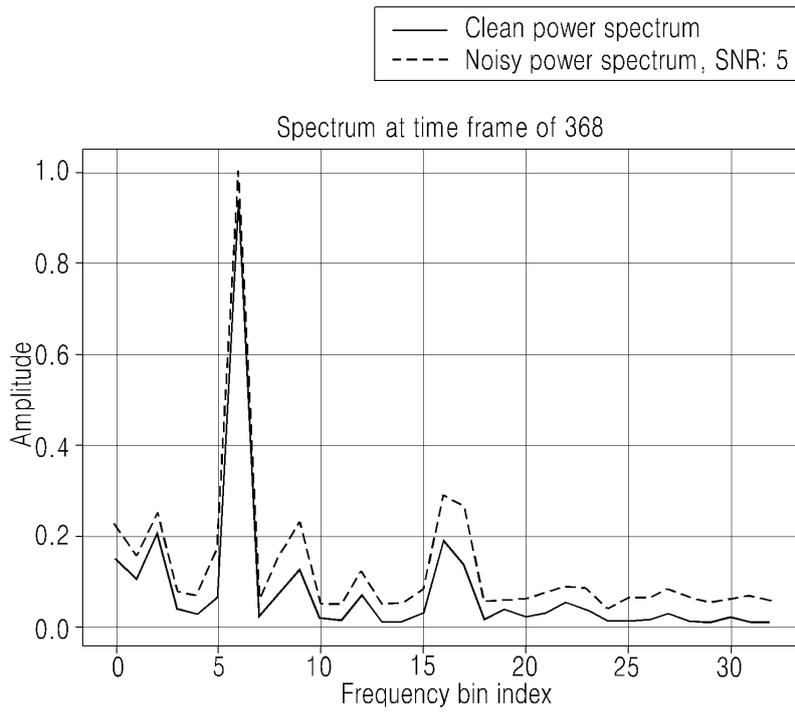
도면6b



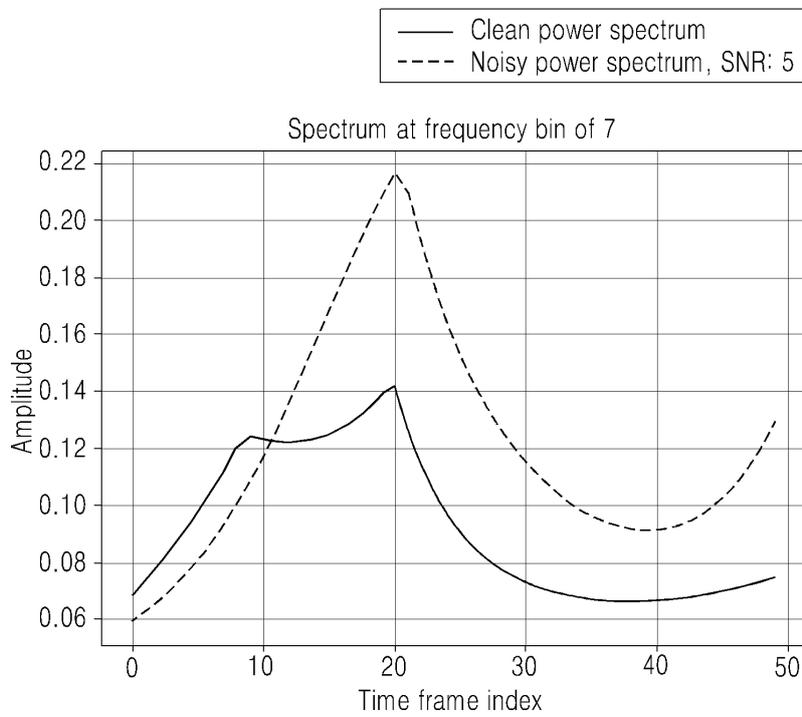
도면6c



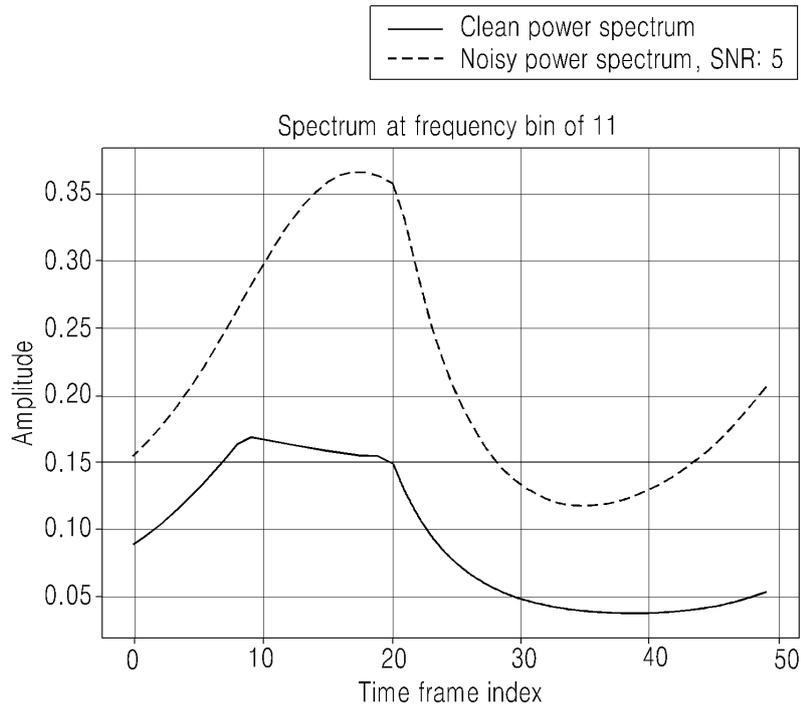
도면6d



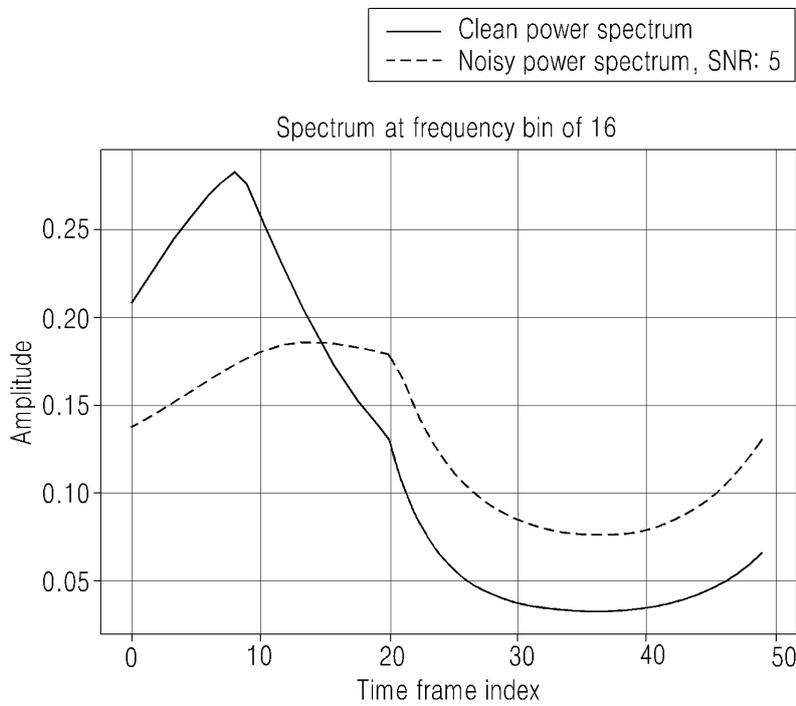
도면7a



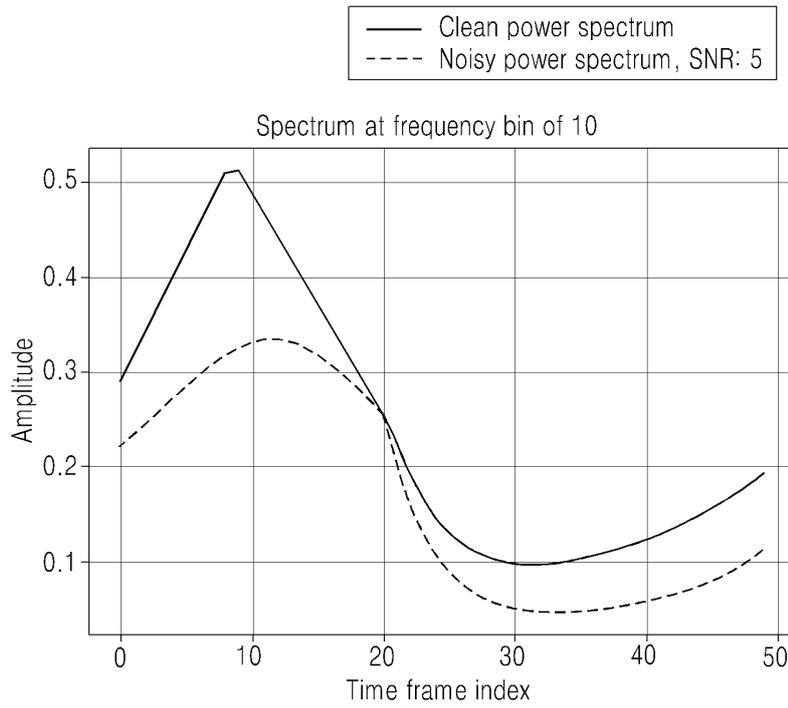
도면7b



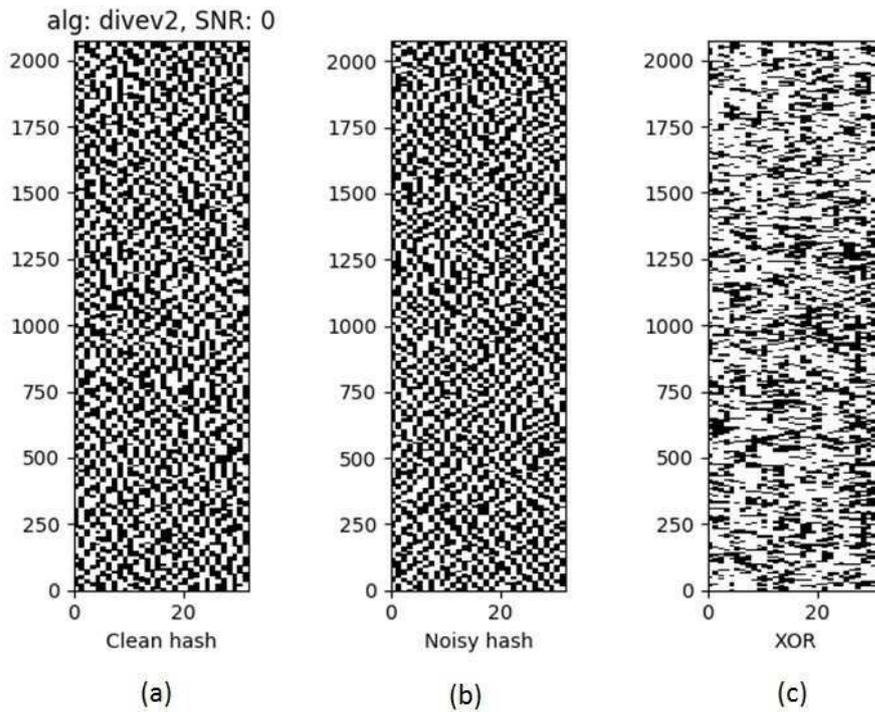
도면7c



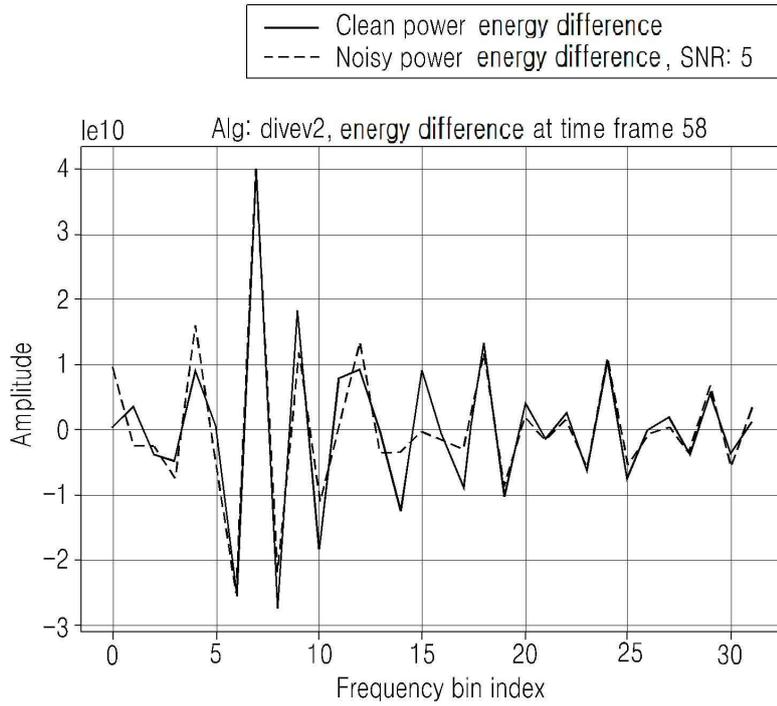
도면7d



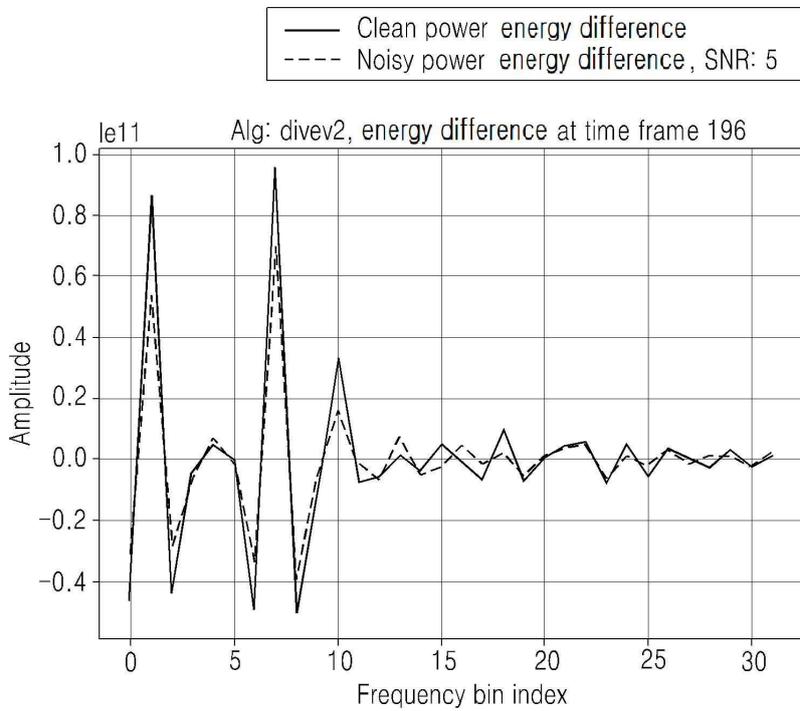
도면8



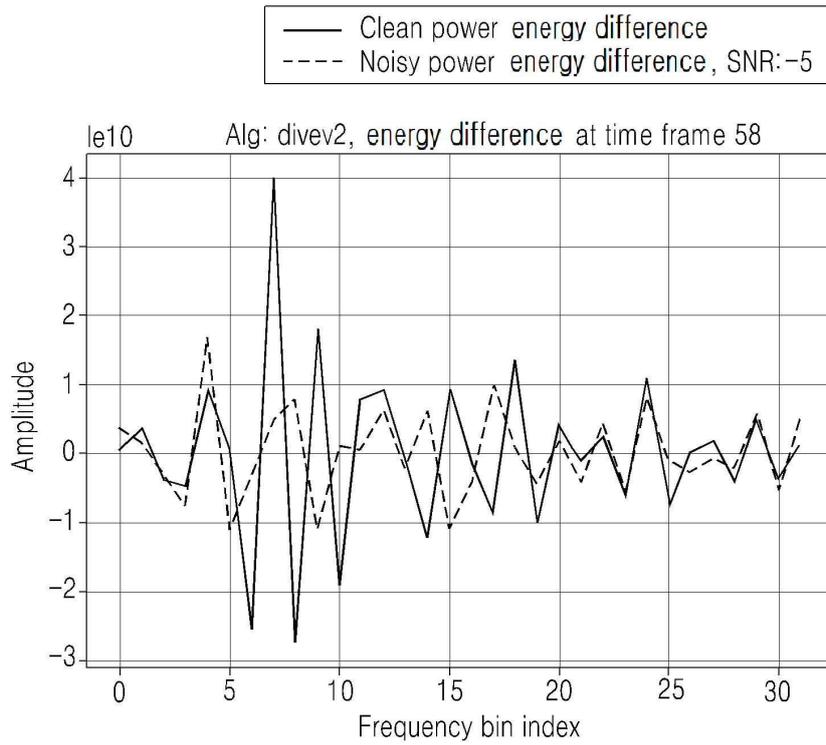
도면9a



도면9b



도면9c



도면9d

