



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2021년12월01일

(11) 등록번호 10-2334390

(24) 등록일자 2021년11월29일

(51) 국제특허분류(Int. Cl.)

G06T 9/00 (2019.01) G06N 3/08 (2006.01)

H04N 19/85 (2014.01) H04N 21/2343 (2011.01)

(52) CPC특허분류

G06T 9/002 (2013.01)

G06N 3/08 (2013.01)

(21) 출원번호 10-2020-0077164

(22) 출원일자 2020년06월24일

심사청구일자 2020년06월24일

(56) 선행기술조사문헌

KR1020170122398 A*

KR1020190001280 A*

KR1020200062183 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

강홍구

서울특별시 서대문구 연세로 50, 제3공학관 C523호(신촌동, 연세대학교)

한혜원

서울특별시 서대문구 연세로 50, 제3공학관 C505호(신촌동, 연세대학교)

(뒷면에 계속)

(74) 대리인

특허법인우인

전체 청구항 수 : 총 13 항

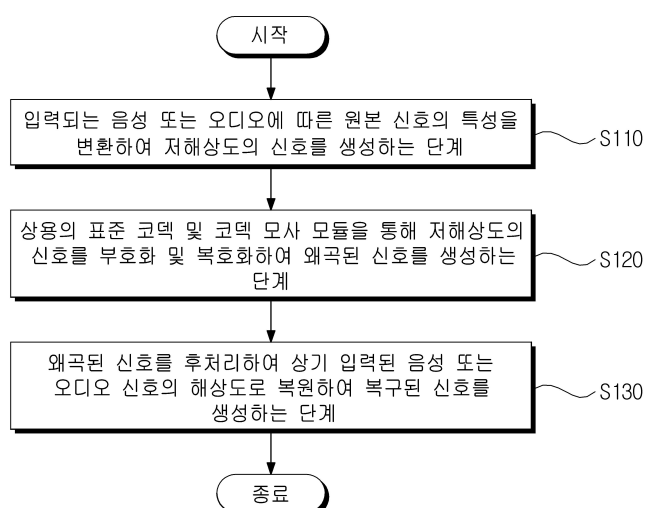
심사관 : 권영학

(54) 발명의 명칭 인공지능 기술을 활용한 코덱 압축 효율 향상 장치 및 방법

(57) 요약

본 실시예들은 하나 이상의 프로세서 및 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 코덱 압축 효율 향상 방법에 있어서, 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 단계, 표준 코덱 및 코덱 모사 모듈을 통해 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계 및 왜곡된 신호를 후처리하여 입력된 원본 신호의 해상도로 복원하여 복구된 신호를 생성하는 단계를 포함하고, 코덱 모사 모듈이 표준 코덱과 유사한 역할을 하도록 딥러닝 네트워크 구조를 기반으로 종단간(end-to-end) 방법으로 학습하도록 구현되는 것을 특징으로 하는 코덱 압축 효율 향상 방법을 제안한다.

대표도 - 도1



(52) CPC특허분류

H04N 19/85 (2015.01)

H04N 21/234381 (2013.01)

G06T 2207/20084 (2013.01)

(72) 발명자

이찬우

서울특별시 서대문구 연세로 50, 제3공학관 C505
호(신촌동, 연세대학교)

이지현

서울특별시 서대문구 연세로 50, 제3공학관 C505
호(신촌동, 연세대학교)

명세서

청구범위

청구항 1

하나 이상의 프로세서 및 상기 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 코덱 압축 효율 향상 방법에 있어서,

입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 단계;

표준 코덱 및 코덱 모사 모듈을 통해 상기 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계; 및

상기 왜곡된 신호를 후처리하여 상기 입력된 원본 신호의 해상도로 복원하여 복구된 신호를 생성하는 단계를 포함하고,

상기 코덱 모사 모듈이 상기 표준 코덱과 유사한 역할을 하도록 딥러닝 네트워크 구조를 기반으로 종단간(end-to-end) 방법으로 학습하도록 구현되며,

상기 복구된 신호를 생성하는 단계는, 상기 왜곡된 신호를 코딩하여 노이즈를 제거하는 단계; 및 상기 노이즈가 제거된 왜곡된 신호를 복구 및 보강하여 상기 복구된 신호를 생성하는 단계를 포함하고,

상기 복구된 신호를 생성하는 단계는 상기 왜곡된 신호가 생성적 대립 신경망에 입력되어 상기 원본 신호와 유사한 신호를 생성하고, 업샘플링 레이어를 통해 상기 원본 신호와 유사한 신호의 원 샘플링 비를 증가시켜 상기 복구된 신호를 생성하는 것을 특징으로 하는 코덱 압축 효율 향상 방법.

청구항 2

제1항에 있어서,

상기 저해상도의 신호를 생성하는 단계는,

네트워크 대역폭에 따라 전송률을 변화하는 네트워크 기반의 조절 또는 상기 입력되는 원본 신호의 특성에 따라 전송률을 변화하는 입력 신호 기반의 조절을 통해 상기 저해상도의 신호를 생성하는 것을 특징으로 하는 코덱 압축 효율 향상 방법.

청구항 3

제2항에 있어서,

상기 저해상도의 신호를 생성하는 단계는,

상기 원본 신호를 분석하여 상기 원본 신호의 프레임에 따른 해상도를 추출하는 단계;

상기 네트워크 대역폭이 입력되고, 상기 추출된 해상도를 상기 네트워크의 상황에 따라 상기 원본 신호의 해상도를 조절하여 상기 저해상도 신호를 생성하는 단계; 및

상기 네트워크를 학습하여 상기 추출된 해상도로 워터마킹 신호를 생성하는 단계를 포함하는 코덱 압축 효율 향상 방법.

청구항 4

제3항에 있어서,

상기 저해상도 신호를 생성하는 단계는,

딥러닝 기법을 사용하여 신호의 시간 축 길이와 주파수 대역폭을 분석 프레임마다 고정 또는 동적으로 변환하거나, 사용자의 청각 특성이 반영되도록 딥러닝 네트워크 구조 및 손실 함수를 사용하는 것을 특징으로 하는 코덱 압축 효율 향상 방법.

청구항 5

제3항에 있어서,

상기 워터마킹 신호를 생성하는 단계는,

오디오 인코더를 통해 상기 원본 신호의 제1 특징 벡터를 추출하는 단계;

조건 파라미터가 완전 연결 레이어(fully-connected layer)에 입력되어 제2 특징 벡터를 추출하는 단계;

상기 제1 특징 벡터 및 상기 제2 특징 벡터를 입력 받아 연결하는 단계; 및

어텐션 레이어를 통해 상기 제1 특징 벡터와 제2 특징 벡터를 이용하여 예측하거나 추정하여 가중 합산된 값을 통해 출력 벡터를 생성하는 단계; 및

디코더를 통해 상기 출력 벡터를 복원하여 열화된 상기 워터마킹 신호를 생성하는 단계를 포함하고,

상기 왜곡된 신호를 생성하는 단계는 상기 저해상도의 신호 대신 상기 워터마킹 신호를 전달 받아 부호화 및 복호화하여 왜곡된 신호를 생성하는 것을 특징으로 하는 코덱 압축 효율 향상 방법.

청구항 6

제1항에 있어서,

상기 왜곡된 신호를 생성하는 단계는,

상기 저해상도의 신호가 제1 신호 경로를 통해 상기 표준 코덱에 입력되어 제1 왜곡된 신호를 출력하는 단계;

상기 저해상도의 신호가 제2 신호 경로를 통해 상기 코덱 모사 모듈에 입력되고, 코덱 설정 변수를 입력 받아 제2 왜곡된 신호를 출력하는 단계를 포함하고,

상기 왜곡된 신호를 생성하는 단계는 상기 제1 왜곡된 신호를 참조값으로 설정하고, 상기 제2 왜곡된 신호가 상기 참조값과 유사한 값이 되도록 오차를 최소화하는 손실 함수를 계산하고,

상기 계산된 손실 함수를 이용하여 역전파 학습을 수행하여 산출한 가중치를 상기 코덱 모사 모듈에 적용하는 것을 특징으로 하는 코덱 압축 효율 향상 방법.

청구항 7

제6항에 있어서,

상기 제1 왜곡된 신호를 출력하는 단계는,

상기 저해상도의 신호가 입력되어 제1 컨볼루션 레이어를 이용하여 컨볼루션 연산을 수행함에 따라 제1 컨볼루션 맵을 생성하는 단계;

다운 샘플링 레이어를 통해 상기 제1 컨볼루션 맵의 원 샘플링 비를 감소시키는 단계;

제2 컨볼루션 레이어를 통해 상기 원 샘플링 비가 감소된 제1 컨볼루션 맵을 컨볼루션 연산을 수행함에 따라 제2 컨볼루션 맵을 생성하는 단계; 및

제1 복원 레이어를 통해 상기 제2 컨볼루션 맵을 복원하여 제1 왜곡된 신호를 출력하는 단계를 포함하는 코덱 압축 효율 향상 방법.

청구항 8

제6항에 있어서,

상기 제2 왜곡된 신호를 출력하는 단계는,

상기 저해상도의 신호가 입력되어 제3 컨볼루션 레이어를 이용하여 컨볼루션 연산을 수행함에 따라 제3 컨볼루션 맵을 생성하는 단계;

레지듀얼 레이어를 통해 상기 제3 컨볼루션 맵에 따른 특징 맵을 생성하는 단계; 및

제2 복원 레이어를 통해 상기 특징 맵을 복원하여 상기 제2 왜곡된 신호를 출력하는 단계를 포함하는 코덱 압축

효율 향상 방법.

청구항 9

제6항에 있어서,

상기 왜곡된 신호를 생성하는 단계는,

코딩 방법, 전송률, 샘플링 주파수, 주파수 대역과 같이 코덱 실행에 필요한 다양한 환경 변수를 상태 벡터(condition vector)를 통해 구현하고,

상기 코덱 모사 모듈은 상기 구현된 환경 변수를 조정함으로써 다양한 표준 코덱에 대한 처리 과정을 모델링 하도록 설계하는 것을 특징으로 하는 코덱 압축 효율 향상 방법.

청구항 10

삭제

청구항 11

제1항에 있어서,

상기 복구된 신호를 생성하는 단계는,

상기 복구된 신호와 상기 원본 신호가 유사한 값이 되도록 손실 함수를 계산하고 상기 계산된 손실 함수를 이용하여 역전파 학습을 수행하여 산출한 가중치를 코덱 압축 효율 향상 장치에 적용하는 것을 특징으로 하는 코덱 압축 효율 향상 방법.

청구항 12

입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 신호 변환 모듈;

표준 코덱 및 코덱 모사 모듈을 통해 상기 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 코덱 모듈; 및

상기 왜곡된 신호를 후처리하여 상기 입력된 원본 신호의 해상도로 복원하여 복구된 신호를 생성하는 신호 복구 모듈을 포함하고,

딥러닝 네트워크 구조를 기반으로 상기 코덱 모사 모듈이 상기 표준 코덱과 유사한 역할을 하도록 종단간(end-to-end) 방법으로 학습하도록 구현되며,

상기 신호 복구 모듈은, 상기 왜곡된 신호를 코딩하여 노이즈를 제거하는 노이즈 제거부; 및 상기 노이즈가 제거된 왜곡된 신호를 복구 및 보강하여 상기 복구된 신호를 생성하는 신호 복구부를 포함하고,

상기 신호 복구 모듈은 상기 왜곡된 신호가 생성적 대립 신경망에 입력되어 상기 원본 신호와 유사한 신호를 생성하고, 업샘플링 레이어를 통해 상기 원본 신호와 유사한 신호의 원 샘플링 비를 증가시켜 상기 복구된 신호를 생성하는 것을 특징으로 하는 코덱 압축 효율 향상 장치.

청구항 13

제12항에 있어서,

상기 신호 변환 모듈은,

상기 원본 신호를 분석하여 상기 원본 신호의 프레임에 따른 해상도를 추출하는 신호 특성 분석부;

네트워크 대역폭이 입력되고, 상기 추출된 해상도를 상기 네트워크의 상황에 따라 상기 원본 신호의 해상도를 조절하여 저해상도 신호를 생성하는 신호 조절부; 및

상기 네트워크를 학습하여 상기 추출된 해상도로 워터마킹 신호를 생성하는 워터마킹부를 포함하는 코덱 압축 효율 향상 장치.

청구항 14

제12항에 있어서,

상기 코덱 모듈은,

상기 저해상도의 신호가 제1 신호 경로를 통해 입력되고, 제1 왜곡된 신호를 출력하는 표준 코덱; 및

상기 저해상도의 신호가 제2 신호 경로를 통해 입력되고, 코덱 설정 변수를 입력 받아 제2 왜곡된 신호를 출력하는 코덱 모사 모듈을 포함하고,

상기 코덱 모듈은 상기 제1 왜곡된 신호를 참조값으로 설정하고, 상기 제2 왜곡된 신호가 상기 참조값과 유사한 값이 되도록 오차를 최소화하는 손실 함수를 계산하고, 상기 계산된 손실 함수를 이용하여 역전과 학습을 수행하여 산출한 가중치를 상기 코덱 모사 모듈에 적용하는 것을 특징으로 하는 코덱 압축 효율 향상 장치.

청구항 15

삭제

발명의 설명

기술 분야

[0001] 본 발명은 코덱 압축 효율 향상 장치 및 방법에 관한 것으로서, 특히 인공지능 기술을 활용한 코덱 압축 효율 향상 장치 및 방법에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 종래에는 높은 압축률, 저 전송률로 전송하기 위해서는 부호화 대상이 되는 입력 신호를 분석하여 인간의 청각으로 쉽게 구분할 수 있는 주파수 성분에 대해서는 많은 정보량을, 그렇지 않은 부분에 대해서는 적은 양의 정보를 할당한다. 즉, 신호 분석을 통해 각 주파수 대역의 중요도를 구하고, 이에 따라 다른 양의 정보를 할당함으로써 부호화에 필요한 총 정보량을 줄일 수 있다. 또한, 다운 샘플링(Down sampling) 신호처리 기법을 통해 부호화에 사용될 입력 신호의 기본 해상도 자체를 줄임으로써 코덱의 부호화 과정에서 필요한 정보량을 자연스럽게 줄이는 방법을 사용할 수 있다.

[0004] 최근, 딥러닝 기술을 음성, 오디오 코덱에 적용하여 성능을 향상시킬 수 있으며, 부호화, 복호화 형태로 이루어진 딥러닝 네트워크 구조를 통해 새로운 형태의 코덱 알고리즘을 개발하는 연구가 이루어지고 있다. 그러나, 딥러닝만을 활용하는 기술은 이미 상용화되어 활발히 사용 중인 표준 코덱과는 호환성이 없으므로 이를 서비스에 활용하기 위해서는 많은 시간이 소요되는 문제가 있다. 또한, 상용화 중인 음성/오디오 코덱을 통해 얻어진 복호화된 신호를 후처리하는 과정을 통해 음질을 향상시키는 방법을 제안한 몇 가지 연구가 있지만, 기존 코덱과의 호환성을 유지하면서 성능을 더욱 높일 수 있는 방법에 대해 개발된 코덱 시스템은 거의 없는 상황이다.

발명의 내용

해결하려는 과제

[0005] 본 발명의 실시예들은 인공지능 기술을 활용하여 음성/오디오 신호의 압축 효율을 높이기 위한 것으로서, 특히 현재 상용 중인 음성/오디오 코덱과의 호환성을 유지할 수 있으므로 특별한 표준화 작업을 하지 않더라도 다양한 응용 분야에 직접 적용하는데 발명의 주된 목적이 있다.

[0006] 또한, 본 발명의 실시예들은 저작권 보호가 필요한 음성/오디오 콘텐츠에 대해서도 복잡한 워터마킹 기술 없이 사용하는데 발명의 주된 목적이 있다.

[0007] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

[0008] 본 실시예의 일 측면에 의하면, 본 발명은 하나 이상의 프로세서 및 상기 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행되는 코덱 압축 효율 향상 방법에 있어

서, 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 단계, 표준 코덱 및 코덱 모사 모듈을 통해 상기 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계 및 상기 왜곡된 신호를 후처리하여 상기 입력된 원본 신호의 해상도로 복원하여 복구된 신호를 생성하는 단계를 포함하고, 상기 코덱 모사 모듈이 상기 표준 코덱과 유사한 역할을 하도록 딥러닝 네트워크 구조를 기반으로 종단간(end-to-end) 방법으로 학습하도록 구현되는 것을 특징으로 하는 코덱 압축 효율 향상 방법을 제안한다.

[0009] 본 발명의 또 다른 실시예에 따르면, 본 발명은 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 신호 변환 모듈, 표준 코덱 및 코덱 모사 모듈을 통해 상기 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 코덱 모듈 및 상기 왜곡된 신호를 후처리하여 상기 입력된 원본 신호의 해상도로 복원하여 복구된 신호를 생성하는 신호 복구 모듈을 포함하고, 딥러닝 네트워크 구조를 기반으로 상기 코덱 모사 모듈이 상기 표준 코덱과 유사한 역할을 하도록 종단간(end-to-end) 방법으로 학습하도록 구현되는 것을 특징으로 하는 코덱 압축 효율 향상 장치를 제공한다.

발명의 효과

[0010] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 본 발명은 인공지능 기술을 활용하여 현재 상용 중인 음성/오디오 코덱과의 호환성을 유지하면서도 코덱의 성능을 최대한 향상시킬 수 있는 효과가 있다.

[0011] 또한, 본 발명은 ITU-T, MPEG과 같이 다수 기관의 동의를 얻기 위해 오랜 시간이 필요한 국제 협의체를 통한 특별한 표준화 작업이 없이도 기존 표준 코덱 시스템과 연계하여 다양한 서비스 분야에 직접 적용할 수 있는 효과가 있다.

[0012] 또한, 본 발명은 신호 특성 변환 모듈의 출력을 효과적으로 설정함으로써 복잡한 워터마킹 기술을 대신하여 저작권 보호가 필요한 음성/오디오 콘텐츠에 대해서 효과적으로 사용할 수 있다.

[0013] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

[0014] 도 1은 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 방법을 나타내는 흐름도이다.
 도 2 내지 도 5는 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 장치를 나타내는 도면이다.
 도 6은 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 장치의 사용 및 학습을 나타내는 도면이다.
 도 7은 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 장치의 코덱 모듈의 학습을 나타내는 도면이다.
 도 8 및 도 9는 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 장치의 네트워크를 나타내는 도면이다.
 도 10은 실시예들에서 사용되기에 적합한 컴퓨팅 기기를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0015] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[0016] 및/또는이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목 중의 어느 항목을 포함한다.

[0017] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다.

[0018] 이하의 설명에서 사용되는 구성요소에 대한 접미사 "모듈" 및 "부"는 명세서 작성의 용이함만이 고려되어 부여되거나 혼용되는 것으로서, 그 자체로 서로 구별되는 의미 또는 역할을 갖는 것은 아니다.

- [0019] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들은 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.
- [0020] 본 발명은 인공지능 기술을 활용한 코덱 압축 효율 향상 장치 및 방법에 관한 것이다.
- [0021] 도 1은 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 방법을 나타내는 흐름도이다. 코덱 압축 효율 향상 방법은 코덱 압축 효율 향상 장치에 의하여 수행될 수 있다.
- [0022] 코덱 압축 효율 향상 방법은 하나 이상의 프로세서 및 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 컴퓨팅 디바이스에 의해 수행될 수 있으며, 반드시 이에 한정되는 것은 아니다.
- [0023] 코덱 압축 효율 향상 방법은 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 단계(S110), 상용의 표준 코덱 및 코덱 모사 모듈을 통해 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계(S120) 및 왜곡된 신호를 후처리하여 입력된 음성 또는 오디오 신호의 해상도로 복원하여 복구된 신호를 생성하는 단계(S130)를 포함한다.
- [0024] 이때, 코덱 압축 효율 향상 방법은 딥러닝 네트워크 구조를 기반으로 코덱 모사 모듈이 표준 코덱과 유사한 역할을 하도록 종단간(end-to-end) 방법으로 학습하도록 구현될 수 있다.
- [0025] 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 단계(S110)는 네트워크 대역폭에 따라 전송률을 변화하는 네트워크 기반의 조절 또는 입력되는 원본 신호의 특성에 따라 전송률을 변화하는 입력 신호 기반의 조절을 통해 저해상도의 신호를 생성할 수 있다.
- [0026] 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 단계(S110)는 원본 신호를 분석하여 상기 원본 신호의 프레임에 따른 해상도를 추출하는 단계, 네트워크 대역폭이 입력되고, 추출된 해상도를 네트워크 상황에 따라 원본 신호의 해상도를 조절하여 저해상도 신호를 생성하는 단계 및 네트워크를 학습하여 추출된 해상도로 워터마킹 신호를 생성하는 단계를 포함할 수 있다.
- [0027] 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성하는 단계(S110)는 신호의 시간 축 길이와 주파수 대역폭을 분석 프레임마다 고정 또는 동적으로 변환하거나, 사용자의 청각 특성이 반영되도록 딥러닝 네트워크 구조 및 비용 함수를 사용할 수 있다.
- [0028] 네트워크를 학습하여 추출된 해상도로 워터마킹 신호를 생성하는 단계는 원본 신호의 특징 벡터를 추출하는 단계, 특징 벡터를 입력으로 모델링하여 출력 벡터를 출력하는 단계 및 출력 벡터를 복원하여 열화된 상기 워터마킹 신호를 생성하는 단계를 포함할 수 있다.
- [0029] 상용의 표준 코덱 및 코덱 모사 모듈을 통해 상기 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계(S120)는 저해상도의 신호 대신 워터마킹 신호를 전달 받아 부호화 및 복호화하여 왜곡된 신호를 생성할 수 있다.
- [0030] 상용의 표준 코덱 및 코덱 모사 모듈을 통해 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계(S120)는 저해상도의 신호가 표준 코덱에 입력되어 제1 왜곡된 신호를 출력하는 단계 및 저해상도의 신호가 상기 코덱 모사 모듈에 입력되어 상기 제2 왜곡된 신호를 출력하는 단계를 포함할 수 있다.
- [0031] 상용의 표준 코덱 및 코덱 모사 모듈을 통해 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계(S120)는 표준 코덱을 이용하여 저해상도의 신호를 압축하고, 코덱 모사 모듈을 통해 압축된 제1 왜곡된 신호 및 표준 코덱을 통해 압축된 제2 왜곡된 신호를 이용하여 왜곡된 신호를 생성할 수 있다.
- [0032] 상용의 표준 코덱 및 코덱 모사 모듈을 통해 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계(S120)는 제1 왜곡된 신호를 참조값으로 설정하고, 제2 왜곡된 신호가 참조값과 유사한 값이 되도록 오차를 최소화하는 손실 함수를 계산하고, 계산된 손실 함수를 역전과 학습을 통해 코덱 모사 모듈에 적용할 수 있다. 여기서, 오차는 가중 평균자승오차일 수 있으며, 반드시 이에 한정되는 것은 아니다.
- [0033] 상용의 표준 코덱 및 코덱 모사 모듈을 통해 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성하는 단계(S120)는 코딩 방법, 전송률, 샘플링 주파수, 주파수 대역과 같이 코덱 실행에 필요한 다양한 환경 변수를 상태 벡터(condition vector)를 통해 구현할 수 있다. 여기서, 코덱 모사 모듈은 구현된 환경 변수를 조정함으로써 다양한 표준 코덱에 대한 처리 과정을 모델링할 수 있도록 설계할 수 있다.
- [0034] 왜곡된 신호를 후처리하여 입력된 음성 또는 오디오 신호의 해상도로 복원하여 복구된 신호를 생성하는 단계

(S130)는 표준 코덱을 통해 신호를 부호화 및 복호화된 저해상도의 왜곡된 신호를 상기 원본 신호의 해상도로 복원할 수 있다. 왜곡된 신호를 후처리하여 입력된 음성 또는 오디오 신호의 해상도로 복원하여 복구된 신호를 생성하는 단계(S130)는 왜곡된 신호를 코딩하여 노이즈를 제거하는 단계 및 노이즈가 제거된 왜곡된 신호를 복구 및 보강하여 복구된 신호를 생성하는 단계를 포함할 수 있다.

[0035] 도 1에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 1에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능하다.

[0036] 도 2 내지 도 5는 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 장치를 나타내는 도면이다.

[0037] 코덱 압축 효율 향상 장치(10)는 신호 변환 모듈(100), 코덱 모듈(200) 및 신호 복구 모듈(300)을 포함한다. 코덱 압축 효율 향상 장치(10)는 도 2에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.

[0038] 코덱 압축 효율 향상 장치(10)는 딥러닝 네트워크 구조를 기반으로 상기 코덱 모사 모듈이 상기 표준 코덱과 유사한 역할을 하도록 종단간(end-to-end) 방법으로 학습하도록 구현될 수 있다.

[0039] 코덱 압축 효율 향상 장치(10)는 인공지능 기술을 활용하여 음성/오디오 신호의 압축 효율을 높이기 위한 것이다.

[0040] 코덱 압축 효율 향상 장치(10)는 특히 현재 상용 중인 음성/오디오 코덱과의 호환성을 유지할 수 있으므로 특별한 표준화 작업을 하지 않더라도 다양한 응용 분야에 직접 적용할 수 있다. 또한, 코덱 압축 효율 향상 장치(10)는 저작권 보호가 필요한 음성/오디오 콘텐츠에 대해서도 복잡한 워터마킹 기술 없이 사용할 수 있는 부가 기능도 포함할 수 있다.

[0041] 코덱 압축 효율 향상 장치(10)는 인공지능 기술을 활용하여 현재 상용 중인 음성/오디오 코덱과의 호환성을 유지하면서도 코덱의 성능을 최대한 향상시킬 수 있다. 특히, 코덱 압축 효율 향상 장치(10)는 ITU-T, MPEG과 같이 다수 기관의 동의를 얻기 위해 오랜 시간이 필요한 국제 협의체를 통한 특별한 표준화 작업이 없이도 기존 표준 코덱 시스템과 연계하여 다양한 서비스 분야에 직접 적용할 수 있다.

[0042] 코덱은 코더(Coder)와 디코더(Decoder)의 합성어로, 용량 대비 품질이 좋아 용량을 감소시키는 경우 손실되는 데이터가 적어 원본과 비슷한 데이터 품질을 나타낼 때 효율이 좋다고 볼 수 있다. 여기서, 데이터 품질은 원본과 청각적으로 차이가 거의 없는 품질을 의미하며, 왜곡이 일어나더라도 청각적으로 큰 차이가 없는 품질을 나타낼 수 있다.

[0043] 신호 특성 변환 모듈의 출력을 효과적으로 설정함으로써 복잡한 워터마킹 기술을 대신할 수 있으므로 저작권 보호가 필요한 음성/오디오 콘텐츠에 대해서도 효과적으로 사용할 수 있는 부가 기능을 포함한다. 따라서, 발명된 코덱 시스템의 산업적 파급 효과는 매우 클 것으로 판단된다.

[0044] 신호 변환 모듈(100)은 입력되는 음성 또는 오디오에 따른 원본 신호의 특성을 변환하여 저해상도의 신호를 생성할 수 있다.

[0045] 신호 변환 모듈(100)은 도 3을 참조하여 자세히 설명한다. 신호 변환 모듈(100)은 신호 특성 분석부(110), 신호 조절부(120) 및 워터 마킹부(130)를 포함한다. 신호 변환 모듈(100)은 도 3에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.

[0046] 신호 변환 모듈(100)은 네트워크 대역폭에 따라 전송률을 변화하는 네트워크 기반의 조절 또는 입력되는 원본 신호의 특성에 따라 전송률을 변화하는 입력 신호 기반의 조절을 통해 저해상도의 신호를 생성할 수 있다.

[0047] 본 발명의 일 실시예에 따르면, 신호 변환 모듈(100)은 입력된 원본 신호를 다음 단계인 코덱 모듈(200)에서 적은 양의 비트를 할당하더라도 효과적으로 압축할 수 있도록 신호의 특성을 변환하는 역할을 담당한다. 따라서, 샘플링 주파수를 낮추어 주파수 해상도를 낮추거나 중요하지 않은 정보를 삭제하는 등 코덱 모듈(200)에서 처리해야 하는 정보량을 최소로 만들 수 있다.

[0048] 대부분의 코덱은 입력의 샘플링 주파수가 고정되고, 프레임별로 일정한 전송률로 압축하는 방법을 사용하지만, 부호화 효율을 높이기 위해서는 입력 신호 혹은 네트워크 상황에 따라 샘플링 주파수 및 전송률이 변화할 수 있도록 조정하는 것이 필수적이다. 본 발명의 일 실시예에 따르면, 신호 변환 모듈(100)은 네트워크 대역폭에 따

라 전송물을 변화하는 네트워크 기반 조절(network-controlled) 방법과 입력 신호의 특성에 따라 전송물을 변화하는 입력 신호 기반 조절(source-controlled) 방법 모두에 적합하도록 딥러닝 네트워크를 설계할 수 있다.

- [0049] 신호 특성 분석부(110)는 원본 신호를 분석하여 상기 원본 신호의 프레임에 따른 해상도를 추출할 수 있다.
- [0050] 본 발명의 일 실시예에 따르면, 신호 특성 분석부(110)는 입력 신호의 특성을 분석하여 해당 프레임 신호에 적합한 해상도를 찾을 수 있다.
- [0051] 신호 조절부(120)는 네트워크 대역폭이 입력되고, 추출된 해상도를 네트워크 상황에 따라 원본 신호의 해상도를 조절하여 저해상도 신호를 생성할 수 있다.
- [0052] 본 발명의 일 실시예에 따르면, 신호 조절부(120)는 네트워크 상황에 따라 이를 조절할 수 있다. 또한, 신호 조절부(120)는 신호 변환 과정에서 사람의 심리음향모델의 특성이 반영될 수 있도록 내부 변수를 조절함으로써 다음 단계 주어진 표준 코덱(210)의 부호화기의 심리음향모델 성능이 높지 않은 경우에도 효과적으로 신호를 압축할 수 있도록 할 수 있다. 심리음향은 물리적 음향의 청각적 자극에 대한 지각 특성을 고려한 것으로, 청각 자극에 대한 지각을 나타낼 수 있다.
- [0053] 워터 마킹부(130)는 네트워크를 학습하여 추출된 해상도로 워터마킹 신호를 생성할 수 있다.
- [0054] 본 발명의 일 실시예에 따르면, 워터 마킹부(130)는 신호 변환 모듈(100)의 출력으로 왜곡된 저해상도 오디오 신호를 얻도록 네트워크를 학습하여 향후 원 신호의 저작권을 보호할 수 있는 워터마킹 기법을 대신하도록 할 수 있다. 즉, 워터 마킹부(130)를 통해 생성된 신호를 직접 청취하면 왜곡된 음질이 얻어지지만, 코덱 모듈(200)에서는 자체 성능의 저하 없이 신호를 생성할 수 있고, 복호화단에서 수행되는 고해상도 신호 변화를 위한 후처리 과정을 통해 원 신호와 유사한 형태로 변환할 수 있다.
- [0055] 본 발명의 일 실시예에 따르면, 신호 변환 모듈(100)은 입력된 음성/오디오 신호를 나타내는 원본 신호가 두 번째 단계인 코덱 모듈(200)에서 효과적으로 처리될 수 있도록 딥러닝 기법을 활용하여 신호의 특성을 변환하는 역할을 담당한다. 즉, 신호 변환 모듈(100)은 코덱의 압축률을 높일 수 있도록 신호의 해상도를 낮추는 과정으로서 신호 복구 모듈(300)을 거친 후에는 원 신호의 고유 특성이 잘 유지될 수 있도록 설계될 수 있다.
- [0056] 코덱 모듈(200)은 표준 코덱 및 코덱 모사 모듈을 통해 저해상도의 신호를 부호화 및 복호화하여 왜곡된 신호를 생성할 수 있다.
- [0057] 코덱 모듈(200)은 도 4를 참조하여 자세히 설명한다. 코덱 모듈(200)은 표준 코덱(210) 및 코덱 모사 모듈(220)을 포함한다. 코덱 모듈(200)은 도 4에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.
- [0058] 코덱 모듈(200)은 제1 왜곡된 신호를 참조값으로 설정하고, 제2 왜곡된 신호가 참조값과 유사한 값이 되도록 가중 평균자승오차를 최소화하는 손실 함수를 계산하고, 계산된 손실 함수를 역전과 학습을 통해 코덱 모사 모듈(220)에 적용할 수 있다.
- [0059] 본 발명의 일 실시예에 따르면, 코덱 모듈(200)은 상용 중인 표준 코덱을 활용하여 이전 모듈을 거쳐 얻어진 변환된 입력 신호를 압축을 담당한다. 그런데, 이 모듈은 신호처리 기반으로 만들어진 알고리즘으로서 양자화 과정 등에서 미분 불가능한 연산을 포함하고 있으므로 역전과 학습이 불가하다. 즉, 본 발명에서처럼 전체 코덱 시스템에 대해 종단간(end-to-end) 딥러닝 학습 방법을 사용하는 구조에는 직접 적용할 수 없다. 상술한 방안을 해결하기 위해 모델 학습에 앞서 도 4와 같이 구현할 수 있다.
- [0060] 본 발명의 일 실시예에 따르면, 코덱 모듈(200)은 상용 코덱(210)의 입출력 관계를 모델링할 수 있는 코덱 모사 모듈(220)을 구현할 수 있다. 여기서, 코덱 모사 모듈(220)은 딥러닝 네트워크(201)로 구현될 수 있으며, 반드시 이에 한정되는 것은 아니다.
- [0061] 본 발명의 일 실시예에 따르면, 코덱 모듈(200)은 표준 코덱(210)의 출력을 참조값으로 하고, 코덱 모사 모듈(220)을 통과한 출력 신호가 참조값과 유사한 값이 되기 위해 가중 평균자승오차(WMSE; Weighted Mean Square Error)를 최소화하도록 비용 함수를 설정함으로써 구현할 수 있다. 여기서 비용 함수는 손실 함수로 구현될 수 있다.
- [0062] 본 발명의 일 실시예에 따르면, 코덱 모사 모듈(220)은 상용 중인 표준 음성/오디오 코덱의 종류는 매우 다양하고, 동일 코덱이라고 하더라도 옵션에 따라 전송물이 다르므로 딥러닝 네트워크의 구조도 이를 반영하도록 설계해야 한다.

- [0063] 따라서, 코텍 모사 모듈(220)은 코딩 방법, 전송률, 샘플링 주파수, 주파수 대역 등과 같이 코텍 실행에 필요한 다양한 환경 변수는 딥러닝 네트워크의 상태 벡터(condition vector)를 통해 구현하고, 그 값을 조정함으로써 한 개의 네트워크로도 다양한 표준 코텍에 대한 처리 과정을 모델링할 수 있도록 설계할 수 있다.
- [0064] 표준 코텍(210)은 저해상도의 신호가 제1 신호 경로를 통해 입력되고, 제1 왜곡된 신호를 출력할 수 있다.
- [0065] 코텍 모사 모듈(220)은 저해상도의 신호가 제2 신호 경로를 통해 코텍 모사 모듈에 입력되고, 코텍 설정 변수를 입력 받아 제2 왜곡된 신호를 출력할 수 있다.
- [0066] 본 발명의 일 실시예에 따르면, 코텍 모사 모듈(220)은 이미 상용화되어 사용 중인 다양한 표준화 코텍 중 어느 것을 사용하여도 무방하며, 이를 통해 표준화 코텍을 사용하는 기존의 모든 코텍 관련 서비스와의 호환성을 유지할 수 있다.
- [0067] 신호 복구 모듈(300)은 왜곡된 신호를 후처리하여 입력된 원본 신호의 해상도로 복원하여 복구된 신호를 생성할 수 있다.
- [0068] 신호 복구 모듈(300)은 도 5를 참조하여 자세히 설명한다. 신호 복구 모듈(300)은 노이즈 제거부(310) 및 신호 복구부(320)를 포함한다. 신호 복구 모듈(300)은 도 5에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.
- [0069] 신호 복구 모듈(300)은 왜곡된 신호가 생성적 대립 신경망에 입력되어 원본 신호와 유사한 신호를 생성하고, 업 샘플링 레이어를 통해 원본 신호와 유사한 신호의 원 샘플링 비를 증가시켜 복구된 신호를 생성할 수 있다.
- [0070] 노이즈 제거부(310)는 왜곡된 신호를 코딩하여 노이즈를 제거할 수 있다.
- [0071] 신호 복구부(320)는 노이즈가 제거된 왜곡된 신호를 복구 및 보강하여 복구된 신호를 생성할 수 있다.
- [0072] 본 발명의 일 실시예에 따르면, 신호 복구 모듈(300)은 표준 코텍(210)의 부호화/복호화 과정을 거쳐 생성된 신호의 해상도를 높인다. 따라서, 신호 복구 모듈(300)의 입력은 표준 코텍(210)을 통해 부호화/복호화된 왜곡된 신호이고, 그 출력은 원본 신호의 해상도로 복원할 수 있도록 설계된 딥러닝 네트워크이다. 도 5는 신호 복구 모듈(300)에 대한 블록 구성도이다.
- [0073] 본 발명의 일 실시예에 따르면 코텍 압축 효율 향상 장치(10)의 본 모듈의 성능을 높이기 위해 가장 좋은 방법은 효과적인 손실 함수를 정하는 것이다. 비록 평균자승오차를 사용하는 것이 가장 간단한 방법이지만, 음성/오디오 신호의 코텍은 사람의 청각 특성이 반영된 방법을 사용하므로 이를 효과적으로 반영하기 위한 비용 함수를 설정하는 것이 매우 중요하다. 같은 맥락에서 출력 신호의 품질을 측정하기 위해 사람들의 주관적 청취 평가를 통해 얻은 결과와 유사하지만, 학습에 도움이 되는 객관적 평가 방법을 도입하는 것이 바람직하다.
- [0074] 평균자승오차는 오차의 제곱에 대해 평균을 취한 것으로, 추정값과 실제 값 간의 평균 제곱 차를 측정할 수 있다. 상술한 바에 따르면, 코텍 압축 효율 향상 장치(10)는 평균자승오차를 이용하는 것으로 도시하였으나 반드시 이에 한정되는 것은 아니다.
- [0075] 본 발명의 일 실시예에 따르면, 신호 복구 모듈(300)은 낮은 해상도로 표현된 신호를 딥러닝 기반의 생성모델 기법을 활용하여 원 신호의 해상도로 복원할 수 있다.
- [0076] 따라서, 본 발명의 일 실시예에 따르면 코텍 압축 효율 향상 장치(10)는 입력부터 출력까지의 일련의 과정을 종단간(end-to-end) 학습 방법을 활용할 수 있도록 구현하였고, 이를 위해 코텍 모듈(200)에 대한 입/출력 관계를 딥러닝 네트워크 구조로 모사할 수 있도록 설계될 수 있다.
- [0077] 종단간(end-to-end) 학습은 모든 매개변수가 하나의 손실 함수에 대해 동시에 훈련되는 경로가 가능한 네트워크로써 역전파 알고리즘(Backpropagation Algorithm)과 함께 최적화 될 수 있다. 종단간(end-to-end) 학습은 딥러닝으로써 훈련 데이터와 시험 데이터로 나눠 학습과 시험을 수행하고, 데이터에서 목표한 결과를 얻을 수 있다. 구체적으로 종단간(end-to-end) 학습은 훈련 데이터만을 사용하여 학습하면서 매개변수(가중치, 편향)를 찾으며, 시험 데이터를 이용하여 훈련된 모델의 성능을 평가할 수 있다. 손실 함수로는 평균 제곱 오차, 교차 엔트로피 오차 등을 사용할 수 있다.
- [0078] 도 6은 본 발명의 일 실시예에 따른 코텍 압축 효율 향상 장치의 사용 및 학습을 나타내는 도면이다.
- [0079] 도 6의 (a)는 본 발명의 일 실시예에 따른 코텍 압축 효율 향상 장치의 실제 사용 시 구성을 나타내는 도면이고, 도 6의 (b)는 본 발명의 일 실시예에 따른 코텍 압축 효율 향상 장치의 훈련 과정의 구성을 나타내는

도면이다.

- [0080] 코덱 압축 효율 향상 장치(10)는 훈련 과정을 거쳐 사용될 수 있다. 이하에서는 훈련 과정 및 사용을 각각 설명한다.
- [0081] 도 6의 (b)를 참조하면, 신호 변환 모듈(100)은 음성 또는 오디오에 따른 원본 신호가 입력될 수 있으며, 원본 신호를 상술한 과정을 거쳐 변환하여 저해상도의 신호를 생성할 수 있다.
- [0082] 저해상도의 신호는 원본 신호의 특성을 분석하여 해당 프레임 신호에 적합한 해상도로 신호 분석을 통해 각 주파수 대역의 중요도를 구하고, 이에 따라 다른 양의 정보가 할당되는 신호일 수 있다. 저해상도의 신호는 낮은 샘플링 주파수로 표현된 신호를 의미할 수 있으며, 상술한 바에 반드시 한정되는 것은 아니다.
- [0083] 훈련 과정에서, 저해상도의 신호가 코덱 모사 모듈(220)에 입력되면 모사 네트워크 사전 훈련이 진행될 수 있다. 모사 네트워크 사전 훈련에 대해서는 도 7을 참조하여 자세히 설명한다.
- [0084] 도 7은 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 장치의 코덱 모듈의 학습을 나타내는 도면이다.
- [0085] 도 7을 참조하면, 코덱 모듈(200)은 표준 코덱(210) 및 코덱 모사 모듈(220)을 포함할 수 있다.
- [0086] 코덱 모듈(200)은 원본 신호가 입력되면, 원본 신호가 표준 코덱(210)과 코덱 모사 모듈(220)로 각각 입력될 수 있다.
- [0087] 표준 코덱(210)은 제1 신호 경로를 통해 원본 신호가 입력되어 제1 왜곡된 신호를 출력할 수 있으며, 코덱 모사 모듈(220)은 제2 신호 경로를 통해 원본 신호가 입력되어 제2 왜곡된 신호를 출력할 수 있다. 이때, 코덱 모듈(200)은 제1 왜곡된 신호 및 제2 왜곡된 신호를 이용하여 역전과 학습을 통해 손실 함수를 계산하고, 손실 함수를 코덱 모사 모듈(220)에 적용할 수 있다.
- [0088] 본 발명의 일 실시예에 따르면, 코덱 모듈(200)은 제1 왜곡된 신호와 제2 왜곡된 신호의 오차를 구해 이 오차에 관여하는 값들의 가중치를 수정하여 오차가 작아지는 방향으로 기 설정된 횟수를 반복하여 수정할 수 있다. 여기서, 가중치는 손실 함수의 결과값을 가장 작게 만드는 매개변수이다.
- [0089] 도 7을 참조하면, 제1 왜곡된 신호와 제2 왜곡된 신호를 이용하여 산출한 가중치는 코덱 모사 모듈(220)에 적용하는 것으로 도시하였으나 반드시 이에 한정되는 것은 아니다.
- [0090] 따라서, 코덱 모듈(200)은 코덱 모사 모듈(220)에 가중치가 적용된 왜곡된 신호를 생성할 수 있으며, 왜곡된 신호는 신호 복구 모듈(300)에 입력될 수 있다. 신호 복구 모듈(300)은 왜곡된 신호를 원본 신호와 유사하게 복구하여 복구된 신호를 생성하며, 복구된 신호와 원본 신호를 이용하여 손실 함수를 계산하고, 역전과 학습을 통해 가중치를 산출하여 코덱 압축 효율 향상 장치(10)에 가중치를 적용시킬 수 있다.
- [0091] 도 6의 (b)와 같이 훈련 과정을 거치면 도 6의 (a)와 같이 실제 사용을 할 수 있다. 도 6의 (a)를 참조하면, 코덱 압축 효율 향상 장치(10)는 원본 신호가 신호 변환 모듈(100)에 입력되면 저해상도의 신호가 출력되며, 저해상도의 신호가 표준 코덱(210)에 입력될 수 있다.
- [0092] 표준 코덱(210)은 부호화기(211) 및 복호화기(213)를 포함한다.
- [0093] 부호화기(211)는 부호화 알고리즘을 구현하는 장치로서, 정보를 비트 또는 부호어로 변환시킬 수 있다. 부호화기(211)는 전화, 음성통신을 위해 음성 정보를 가능한 낮은 속도로 전송할 수 있게 하며, 반드시 이에 한정되는 것은 아니다.
- [0094] 복호화기(213)는 부호화된 정보를 부호화되기 전으로 되돌리는 처리방식으로, 부호화의 절차를 역으로 수행할 수 있으며, 반드시 이에 한정되는 것은 아니다.
- [0095] 표준 코덱(210)은 입력된 저해상도의 신호를 왜곡된 신호로 출력하며, 신호 복구 모듈(300)은 왜곡된 신호를 입력 받아 원본 신호와 유사하도록 복구하여 복구된 신호를 생성할 수 있다.
- [0096] 따라서, 상술한 과정을 통해 코덱 압축 효율 향상 장치(10)는 사용화되어 사용중인 표준 코덱(210)과의 호환성을 유지하면서 성능을 높일 수 있다.
- [0097] 도 8 및 도 9는 본 발명의 일 실시예에 따른 코덱 압축 효율 향상 장치의 네트워크를 나타내는 도면이다.
- [0098] 도 8의 (a)는 오디오 효율적 표현을 위한 네트워크를 나타내는 도면이다.

- [0099] 도 8의 (a)를 참조하면, 오디오 효율적 표현을 위한 네트워크는 제1 컨볼루션 레이어(212), 다운샘플링 레이어(214), 제2 컨볼루션 레이어(216) 및 제1 복원 레이어(218)를 포함한다. 오디오 효율적 표현을 위한 네트워크는 도 8의 (a)에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.
- [0100] 제1 컨볼루션 레이어(212)는 1-D convolutional layer로 구현될 수 있으며, 1차원의 특징을 추출할 수 있다. 제1 컨볼루션 레이어(212)는 원본 신호가 입력되어 컨볼루션 연산을 수행함에 따라 제1 컨볼루션 맵을 생성할 수 있다. 여기서, 제1 컨볼루션 레이어(212)에 입력되는 신호는 원본 신호로 도시하였으나 반드시 이에 한정되는 것은 아니며, 저해상도의 신호가 입력될 수도 있다.
- [0101] 다운샘플링 레이어(Downsampling layer)(214)는 신호 처리 과정 가운데 신호의 원 샘플링 비를 낮추는 일을 수행할 수 있다.
- [0102] 여기서, 제1 컨볼루션 레이어(212) 및 다운샘플링 레이어(214)는 인코딩 과정을 나타내며, 이하의 제2 컨볼루션 레이어(216) 및 제1 복원 레이어(218)는 디코딩 과정을 나타낼 수 있다.
- [0103] 제2 컨볼루션 레이어(216)는 1-D convolutional layer로 구현될 수 있으며, 1차원의 특징을 추출할 수 있다. 제2 컨볼루션 레이어(216)는 다운샘플링 레이어(214)에 의해 원 샘플링 비가 감소된 제1 컨볼루션 맵을 컨볼루션 연산을 수행함에 따라 제2 컨볼루션 맵을 생성할 수 있다.
- [0104] 제1 복원 레이어(Reconstruction layer)(218)는 제2 컨볼루션 맵을 복원하여 제1 왜곡된 신호를 출력할 수 있다. 제1 복원 레이어(218)는 입력되는 신호를 재구성하여 출력하며, 입력되는 원본 신호를 재구성할 수 있다.
- [0105] 도 8의 (b)는 코텍 모사 네트워크를 나타내는 도면이다.
- [0106] 도 8의 (b)를 참조하면, 코텍 모사 네트워크는 제3 컨볼루션 레이어(222), 레지듀얼 레이어(224) 및 제2 복원 레이어(226)를 포함한다. 코텍 모사 네트워크는 도 8의 (b)에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.
- [0107] 제3 컨볼루션 레이어(222)는 1-D convolutional layer로 구현될 수 있으며, 1차원의 특징을 추출할 수 있다. 제3 컨볼루션 레이어(222)는 원본 신호가 입력되어 컨볼루션 연산을 수행함에 따라 제3 컨볼루션 맵을 생성할 수 있다. 여기서, 제3 컨볼루션 레이어(222)에 입력되는 신호는 원본 신호로 도시하였으나 반드시 이에 한정되는 것은 아니며, 저해상도의 신호가 입력될 수도 있다.
- [0108] 레지듀얼 레이어(Residual layer)(224)는 제3 컨볼루션 맵에 따른 특징 맵을 생성할 수 있다. 레지듀얼 레이어(224)는 복잡한 기법 없이 특징 추출과 재구성이 가능하며, 복원이 가능할 수 있다. 레지듀얼 레이어(224)는 복수의 컨볼루션 레이어로 구성되어 있을 수 있으며, 반드시 이에 한정되는 것은 아니다.
- [0109] 제2 복원 레이어(Reconstruction layer)(226)는 특징 맵을 복원하여 상기 제2 왜곡된 신호를 출력할 수 있다. 제2 복원 레이어(226)는 입력되는 신호를 재구성하여 출력하며, 입력되는 원본 신호를 재구성할 수 있다.
- [0110] 도 9의 (a)는 워터마킹 등을 이용한 열화 모델링을 나타내는 도면이다.
- [0111] 도 9의 (a)를 참조하면, 워터마킹 등을 이용한 열화 모델링은 오디오 인코더(132), 인코더 전결합 레이어(135), 연결(134), 어텐션 레이어(136) 및 디코더(138)를 포함한다. 오디오 효율적 표현을 위한 네트워크는 도 9의 (a)에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.
- [0112] 오디오 인코더(Audio encoder)(132)는 원본 신호의 제1 특징 벡터를 추출할 수 있다. 오디오 인코더(132)는 특징을 추출하는 역할을 할 수 있다.
- [0113] 인코더 전결합 레이어(Encoder(fully-connected layer))(135)는 조건 파라미터가 입력되어 제2 특징 벡터를 추출할 수 있다.
- [0114] 연결(Concatenation)(134)은 제1 특징 벡터 및 제2 특징 벡터를 입력 받아 연결할 수 있다.
- [0115] 어텐션 레이어(Attention layer)(136)는 어텐션 함수로 구현될 수 있으며, 제1 특징 벡터와 제2 특징 벡터를 이용하여 예측하거나 추정하여 가중 합산된 값을 통해 출력 벡터를 생성할 수 있다. 어텐션 함수는 주어진 쿼리(Query)에 대해서 키(Key)와의 유사도를 각각 구하고, 유사도를 키와 맵핑되어있는 각각의 값(Value)에 반영해 준다. 유사도가 반영된 값(Value)을 모두 더해서 리턴한다. 이를 어텐션 값(Attention Value)이라고 한다.

- [0116] 연결(134) 및 어텐션 레이어(136)는 열화 모델링을 수행하며, 잡음도 고려하여 열화 모델링을 수행할 수 있다.
- [0117] 디코더(Decoder)(138)는 출력 벡터를 복원하여 열화된 워터마킹 신호를 생성할 수 있다.
- [0118] 워터마킹은 음성 신호의 데이터에 인지되지 않는 정보를 은닉시키는 기술로서, 불법 복제, 사용 등을 막고 저작권을 보호하기 위해 사용될 수 있다.
- [0119] 도 9의 (b)는 오디오 복원 네트워크를 나타내는 도면이다.
- [0120] 도 9의 (b)를 참조하면, 오디오 복원 네트워크는 생성적 대립 신경망(310) 및 업샘플링 레이어(320)를 포함한다. 오디오 복원 네트워크는 도 9의 (b)에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다.
- [0121] 생성적 대립 신경망(Generative model(GAN))(310)은 생성모델과 판별모델이 경쟁하면서 실제와 가까운 음성을 자동으로 만들어 내는 기계학습이다. 생성적 대립 신경망(310)은 왜곡된 신호가 생성적 대립 신경망에 입력되어 상기 원본 신호와 유사한 신호를 생성할 수 있다.
- [0122] 업샘플링 레이어(Upsampling layer)(320)는 원본 신호와 유사한 신호의 원 샘플링 비를 증가시켜 복구된 신호를 생성할 수 있다. 업샘플링 레이어(320)는 다운 샘플링에 의해 원 샘플링 비가 감소된 것을 증가시켜 복구시킬 수 있다.
- [0123] 도 10은 실시예들에서 사용되기에 적합한 컴퓨팅 기기를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도이다.
- [0124] 도 10은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 기기를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술되지 않은 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.
- [0125] 도시된 컴퓨팅 환경은 코덱 압축 효율 향상 장치(10)를 포함한다. 일 실시예에서, 코덱 압축 효율 향상 장치(10)는 타 단말기와 신호를 송수신하는 모든 형태의 컴퓨팅 기기일 수 있다.
- [0126] 코덱 압축 효율 향상 장치(10)는 적어도 하나의 프로세서(1100), 컴퓨터 판독 가능한 저장매체(1200) 및 통신 버스(1600)를 포함한다. 프로세서(1100)는 코덱 압축 효율 향상 장치(10)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(1100)는 컴퓨터 판독 가능한 저장 매체(1200)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(1100)에 의해 실행되는 경우 코덱 압축 효율 향상 장치(10)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0127] 컴퓨터 판독 가능한 저장 매체(1200)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(1200)에 저장된 프로그램(1300)은 프로세서(1100)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능한 저장 매체(1200)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 기기들, 광학 디스크 저장 기기들, 플래시 메모리 기기들, 그 밖에 코덱 압축 효율 향상 장치(10)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0128] 통신 버스(1600)는 프로세서(1100), 컴퓨터 판독 가능한 저장 매체(1200)를 포함하여 코덱 압축 효율 향상 장치(10)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0129] 코덱 압축 효율 향상 장치(10)는 또한 하나 이상의 입출력 장치(미도시)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(1400) 및 하나 이상의 통신 인터페이스(1500)를 포함할 수 있다. 입출력 인터페이스(1400) 및 통신 인터페이스(1500)는 통신 버스(1600)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(1400)를 통해 코덱 압축 효율 향상 장치(10)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(미도시)는 코덱 압축 효율 향상 장치(10)를 구성하는 일 컴포넌트로서 코덱 압축 효율 향상 장치(10)의 내부에 포함될 수도 있고, 코덱 압축 효율 향상 장치(10)와는 구별되는 별개의 장치로 컴퓨팅 기기와 연결될 수도 있다.

[0130] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.

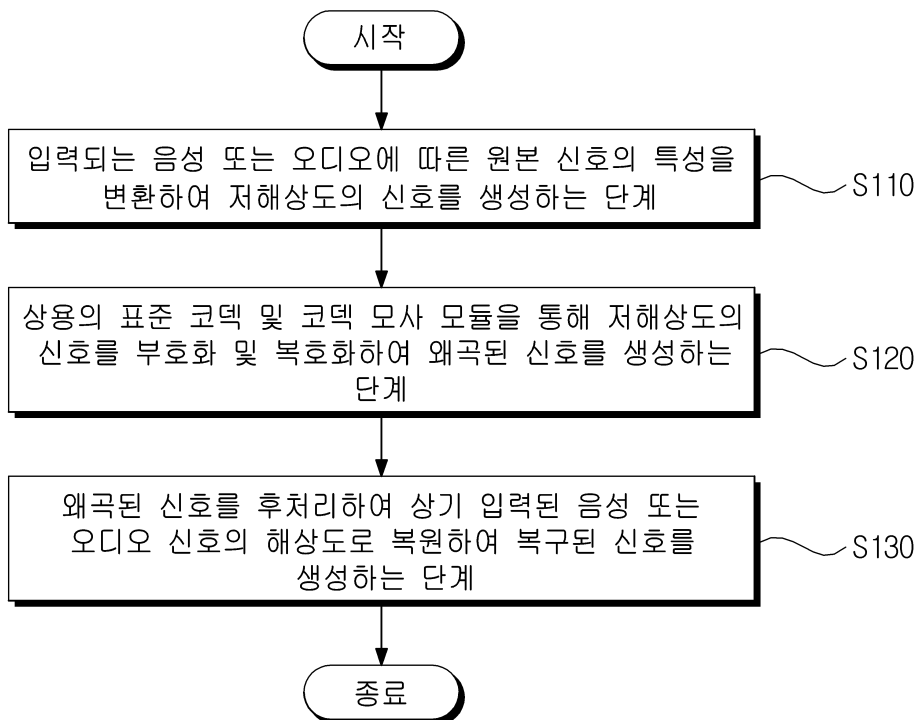
[0131] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0132] 10: 코덱 압축 효율 향상 장치
100: 신호 변환 모듈
200: 코덱 모듈
300: 신호 복구 모듈

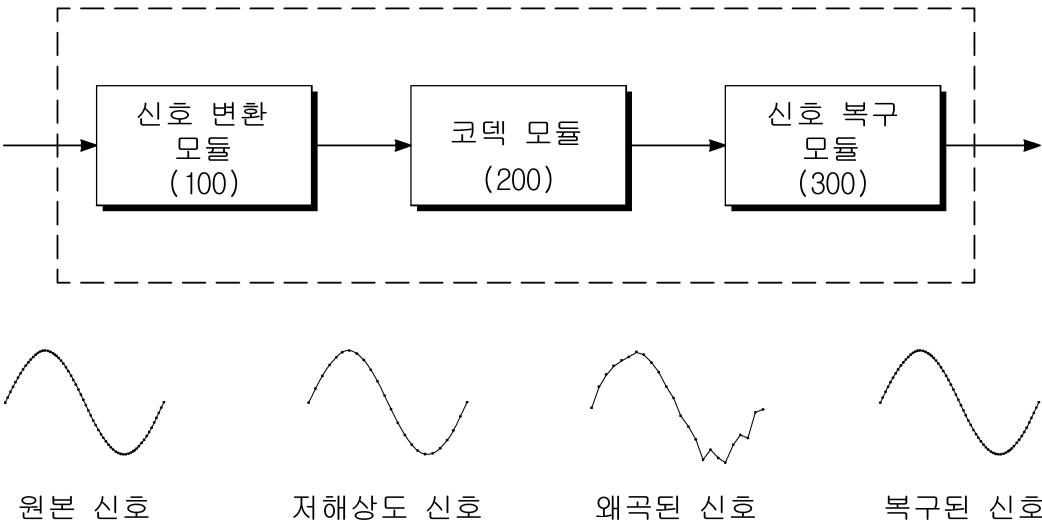
도면

도면1

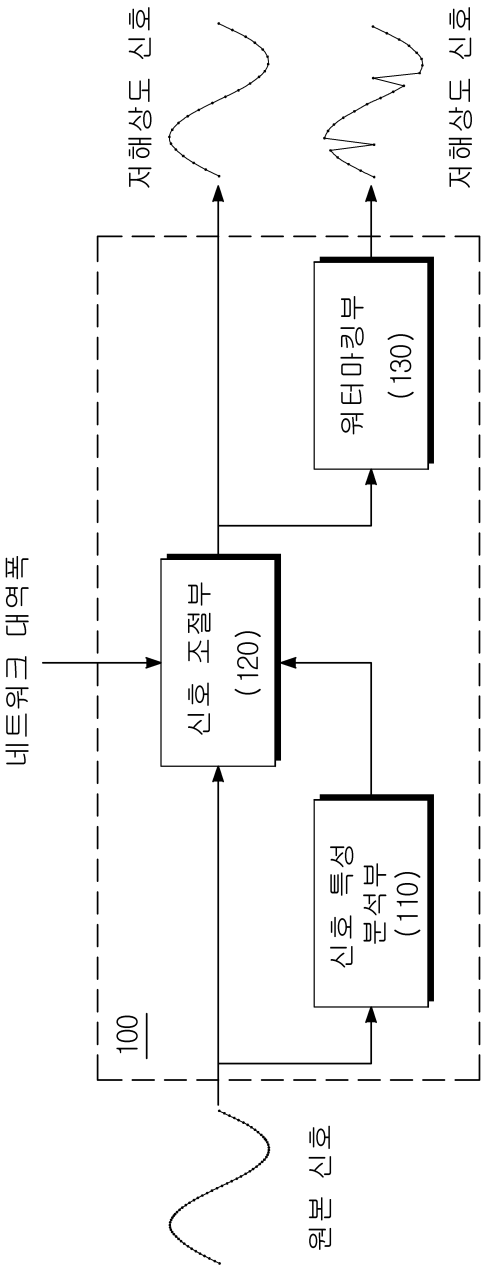


도면2

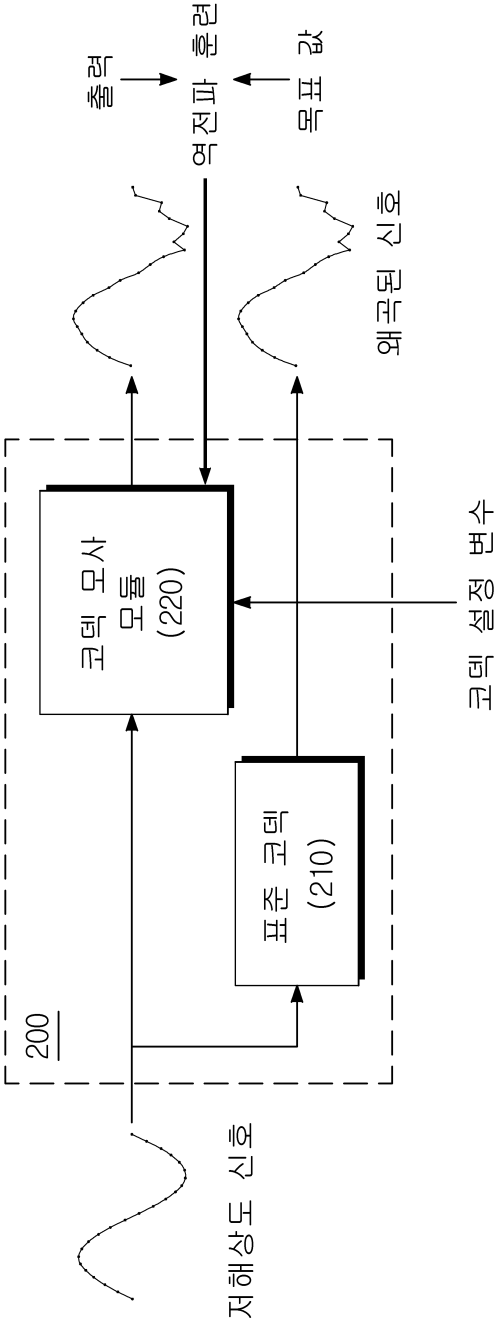
10



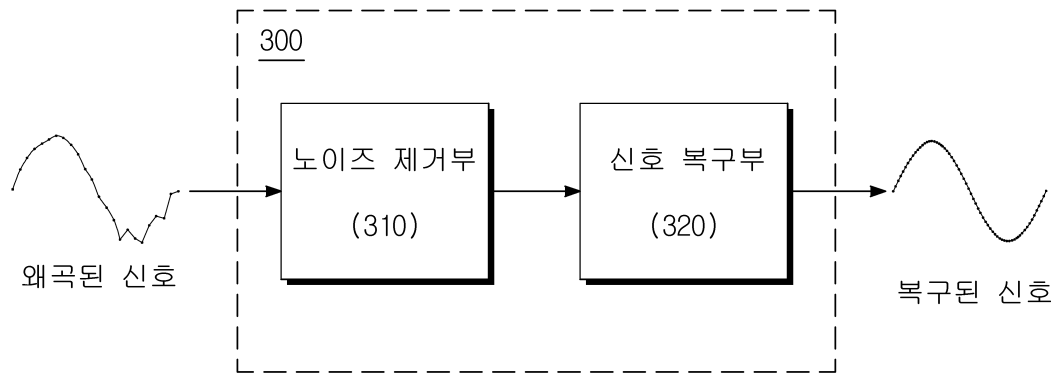
도면3



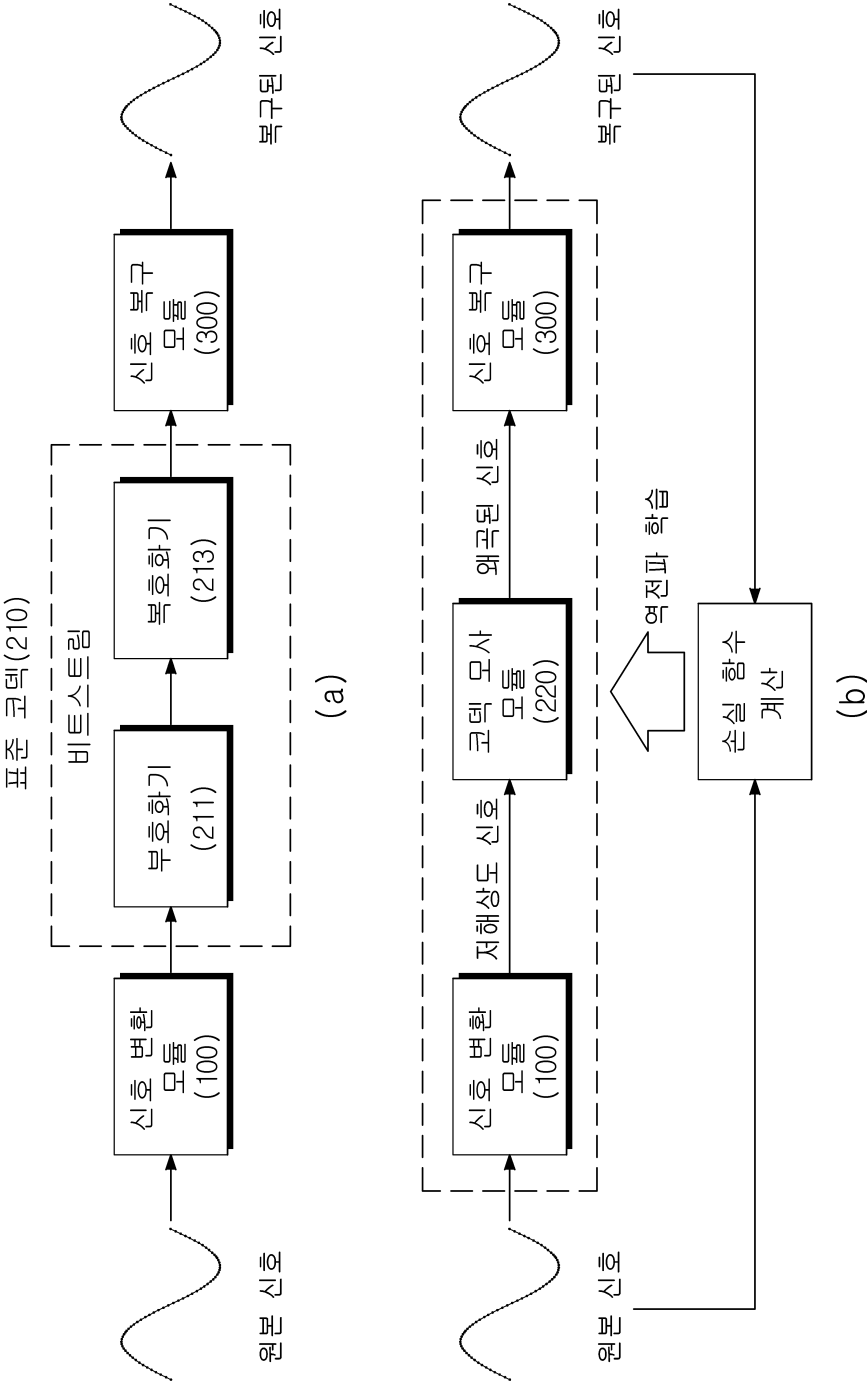
도면4



도면5

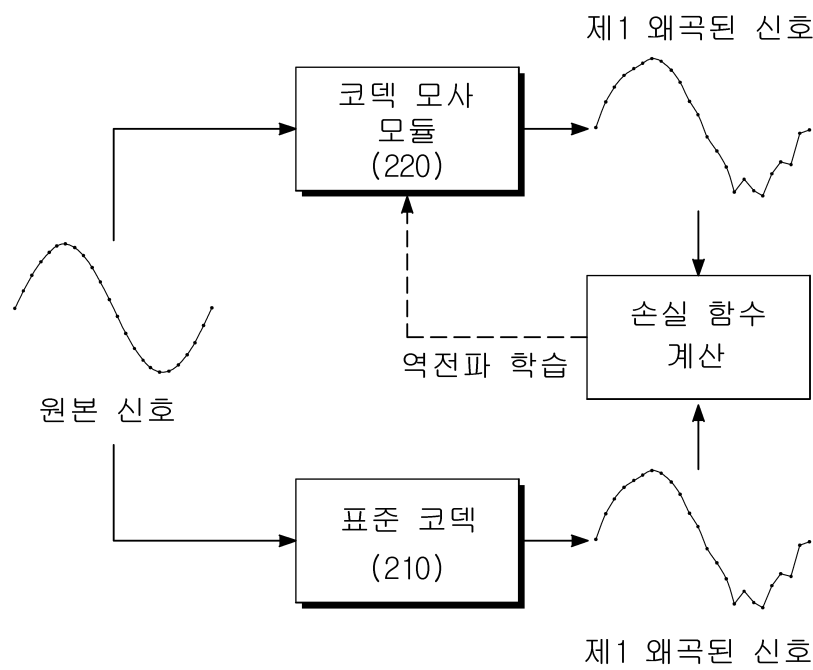


도면6

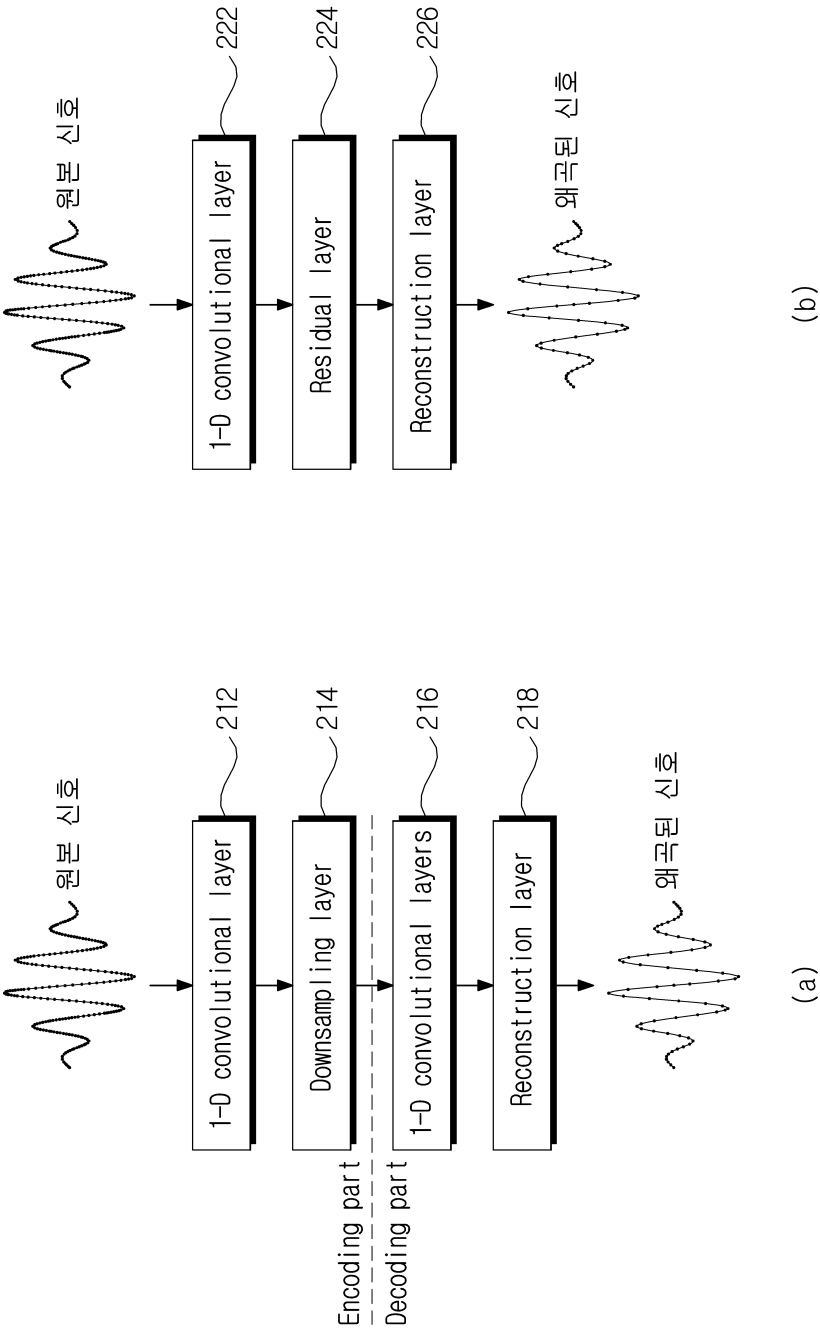


도면7

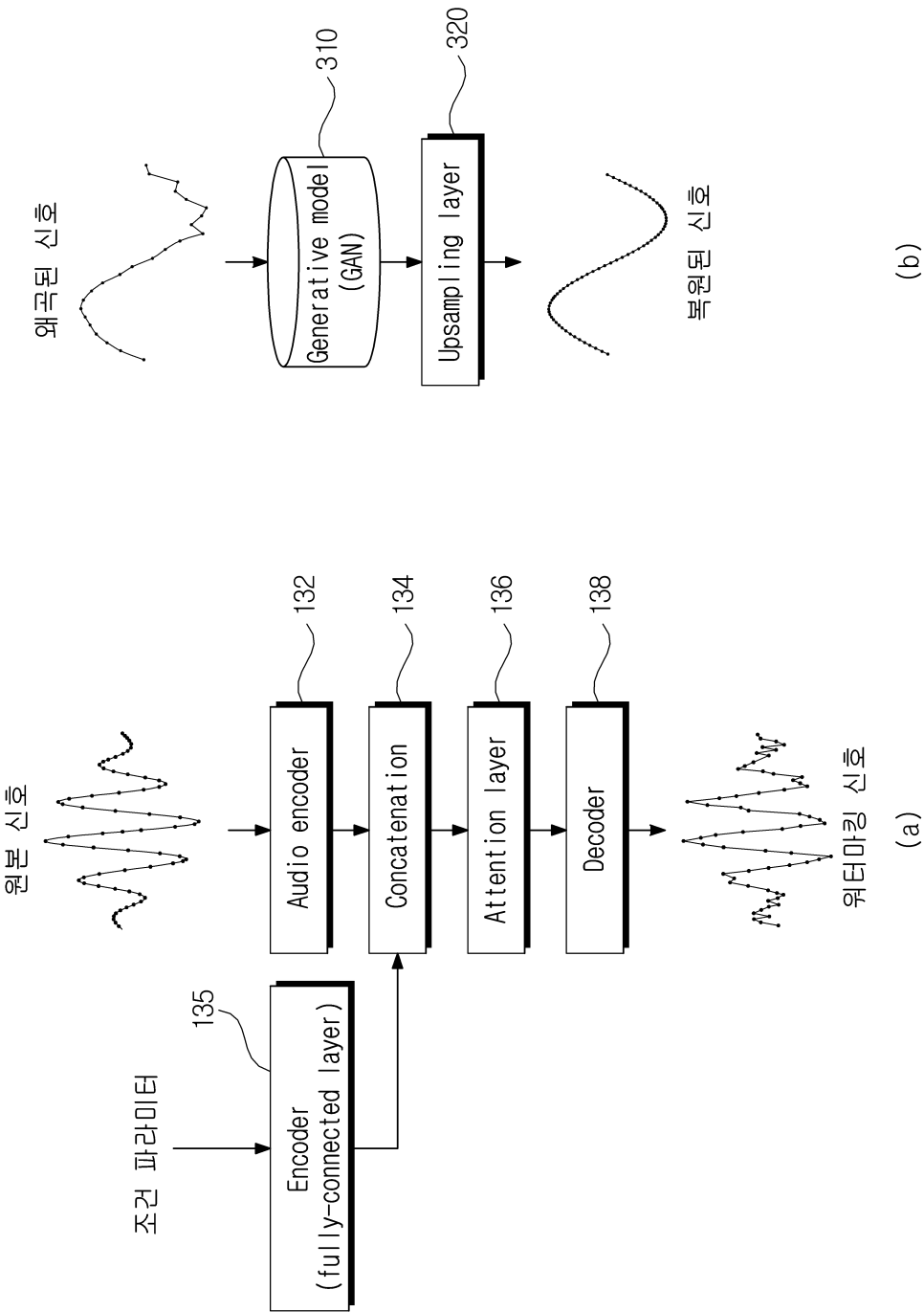
200



도면8



도면9



도면10

