

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2011년 7월 21일 (21.07.2011)



(10) 국제공개번호
WO 2011/087332 A2

- (51) 국제특허분류: 미분류
- (21) 국제출원번호: PCT/KR2011/000324
- (22) 국제출원일: 2011년 1월 17일 (17.01.2011)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
61/295,170 2010년 1월 15일 (15.01.2010) US
61/349,192 2010년 5월 27일 (27.05.2010) US
61/377,448 2010년 8월 26일 (26.08.2010) US
61/426,502 2010년 12월 22일 (22.12.2010) US
- (71) 출원인 (US 을(를) 제외한 모든 지정국에 대하여): 엘지전자 주식회사 (LG ELECTRONICS INC.) [KR/KR]; 서울 영등포구 여의도동 20, 150-721 Seoul (KR). 충북대학교 산학협력단 (CHUNGBUK NATIONAL UNIVERSITY INDUSTRY-ACADEMIC

COOPERATION FOUNDATION) [KR/KR]; 충청북도 청주시 흥덕구 개신동 12, 361-804 ChungBuk (KR).

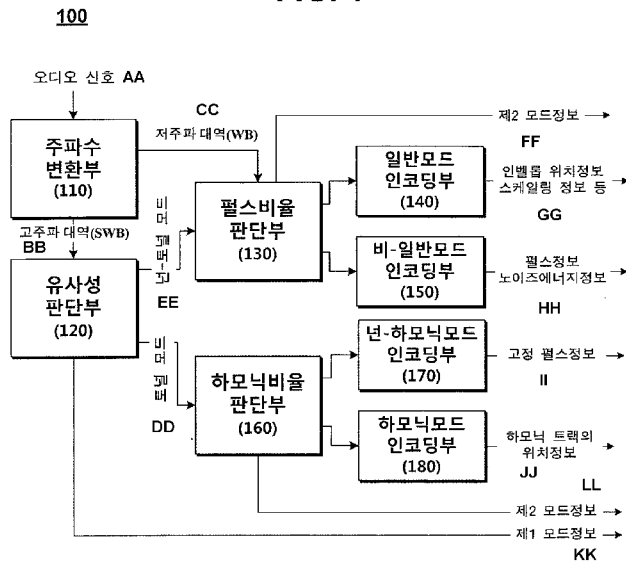
- (72) 발명자: 겸
- (75) 발명자/출원인 (US 에 한하여): 정규혁 (JEONG, Gyuhyeok) [KR/KR]; 서울시 서초구 우면동 16번지 엘지전자 특허센터, 137-724 Seoul (KR). 김대환 (KIM, Dachwan) [KR/KR]; 서울시 서초구 우면동 16번지 엘지전자 특허센터, 137-724 Seoul (KR). 강인규 (KANG, Ingyu) [KR/KR]; 서울시 서초구 우면동 16번지 엘지전자 특허센터, 137-724 Seoul (KR). 김락용 (KIM, Lagyoung) [KR/KR]; 서울시 서초구 우면동 16번지 엘지전자 특허센터, 137-724 Seoul (KR). 홍기봉 (HONG, Kibong) [KR/KR]; 충청북도 청주시 흥덕구 개신동 12, 361-804 ChungBuk (KR). 박지강 (PIAO, Zhigang) [CN/KR]; 충청북도 청주시 흥덕구 개신동 12, 361-804 ChungBuk (KR). 이인성 (LEE, Insung) [KR/KR]; 충청북도 청주시 흥덕구 개신동 12, 361-804 ChungBuk (KR). 임종하 (LIM, Jongha)

[다음 쪽 계속]

(54) Title: METHOD AND APPARATUS FOR PROCESSING AN AUDIO SIGNAL

(54) 발명의 명칭 : 오디오 신호 처리 방법 및 장치

FIG. 1



(57) Abstract: The present invention relates to a method for processing an audio signal, comprising: a step of performing a frequency conversion process on an audio signal to obtain a plurality of frequency transform coefficients; a step of selecting either a general mode or a non-general mode, on the basis of a pulse ratio, for the frequency transform coefficients having a high frequency band from among the plurality of frequency transform coefficients; and a step of performing, if the non-general mode is selected, the following steps: extracting a predetermined number of pulses from the frequency transform coefficients having the high frequency band, and generating pulse information; generating an original noise signal from the frequency transform coefficients having the high frequency band, excluding the pulses; generating a reference noise signal using the frequency transform coefficient having a low frequency band from among the plurality of frequency transform coefficients; and generating noise position information and noise energy information using the original noise signal and the reference noise signal.

(57) 요약서:

[다음 쪽 계속]

- AA ... Audio signal
BB ... High frequency band (SWB)
CC ... Low frequency band (WB)
DD ... Tonal mode
EE ... Non-tonal mode
FF ... 2nd mode information
GG ... Envelope position information
Scaling information, etc.
HH ... Pulse information
Noise energy information
II ... Fixed pulse information
JJ ... Information on the position
of a harmonic track
KK ... 1st mode information
LL ... 2nd mode information
110 ... Frequency converting unit
120 ... Similarity judging unit
130 ... Pulse ratio judging unit
140 ... General mode encoding unit
150 ... Non-general mode encoding unit
160 ... Harmonic ratio judging unit
170 ... Non-harmonic mode encoding unit
180 ... Harmonic mode encoding unit



[KR/KR]; 충청북도 청주시 흥덕구 개신동 12, 361-804 ChungBuk (KR). **문상현 (MOON, Sanghyeon)** [KR/KR]; 충청북도 청주시 흥덕구 개신동 12, 361-804 ChungBuk (KR). **이병석 (LEE, Byungsuk)** [KR/KR]; 서울시 서초구 우면동 16번지 엘지전자 특허센터, 137-724 Seoul (KR). **전혜정 (JEON, Hye-jeong)** [KR/KR]; 서울시 서초구 우면동 16번지 엘지전자 특허센터, 137-724 Seoul (KR).

(74) **대리인: 김용인 (KIM, Yong In)** 등; 서울 송파구 잠실동 175-9 현대빌딩 7층 KBK 특허법률사무소, 138-861 Seoul (KR).

(81) **지정국** (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,

MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **지정국** (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서 없이 공개하며 보고서 접수 후 이를 별도 공개함 (규칙 48.2(g))

본 발명은, 오디오 신호에 대해서 주파수 변환을 수행함으로써, 복수 개의 주파수 변환 계수들을 획득하는 단계; 상기 복수 개의 주파수 변환 계수들 중 고주파 대역의 주파수 변환 계수들에 대해서, 펄스 비율을 근거로 일반 모드 및 비-일반모드 중 하나를 선택하는 단계; 및, 상기 비-일반 모드가 선택된 경우, 다음 단계가 수행되는 단계; 상기 고주파 대역의 주파수 변환 계수들로부터 미리 결정된 개수의 펄스들을 추출하고, 펄스 정보를 생성하는 단계; 상기 고주파 대역의 주파수 변환 계수들로부터 상기 펄스들을 제외한 원본 노이즈 신호를 생성하는 단계; 상기 복수 개의 주파수 변환 계수들 중 저주파 대역의 주파수 변환 계수를 이용하여 참조 노이즈 신호를 생성하는 단계; 및, 상기 원본 노이즈 신호 및 상기 참조 노이즈 신호를 이용하여, 노이즈 위치정보 및 노이즈 에너지정보를 생성하는 단계를 포함하는 오디오 신호 처리 방법이 개시된다.

【DESCRIPTION】**【Invention Title】**

오디오 신호 처리 방법 및 장치

【Technical Field】

- 5 본 발명은 오디오 신호를 인코딩하거나 디코딩할 수 있는 오디오 신호 처리 방법 및 장치에 관한 것이다.

【Background Art】

- 일반적으로, 오디오 신호는 다양한 주파수의 신호가 포함되어 있고, 사람의 가청 주파수는, 20Hz-20kHz 인데 비해, 보통 사람의 음성은 약 200Hz-
10 3kHz 영역에 존재한다.

 사람의 음성이 존재하기 어려운 7kHz 이상의 고주파 영역의 오디오 신호를 인코딩하는 데 있어서, 오디오의 특성에 따라서 복수 개의 코딩 모드 또는 코딩 방식들 중 하나를 적용할 수 있다.

【Disclosure】

- 15 **【Technical Problem】**

 오디오의 특성에 적합하지 않은 코딩 모드 또는 코딩 방식을 적용할 경우, 음질이 왜곡되는 문제점이 있다.

【Technical Solution】

- 본 발명은 상기와 같은 문제점을 해결하기 위해 창안된 것으로서,
20 예컨대 타악기의 소리와 같이 특정 주파수 대역이 강한 신호의 경우, 해당

대역의 펄스를 별도로 인코딩하기 위한 오디오 신호 처리 방법 및 장치를 제공하는 데 있다.

- 본 발명의 또 다른 목적은, 예컨대 현악의 소리와 같이 하모닉(고조파)이 강하게 존재하는 신호의 경우, 해당 하모닉 트랙을 별도로
- 5 인코딩하기 위한 오디오 신호 처리 방법 및 장치를 제공하는 데 있다.

본 발명의 또 다른 목적은, 펄스비율 및/또는 하모닉비율을 기반으로, 오디오 신호의 특성에 적절한 코딩 모드를 적용하기 위한 오디오 신호 처리 방법 및 장치를 제공하는 데 있다.

【Advantageous Effects】

- 10 본 발명은 다음과 같은 효과와 이점을 제공한다.

첫째, 특정 주파수 대역이 강한 신호에 대해서, 해당 대역의 펄스만을 별도로 인코딩으로써, 저주파 대역만을 이용한 인코딩 모드(일반 모드)보다 복원률이 높기 때문에, 음질이 현저히 향상될 수 있다.

- 둘째, 하모닉이 포함된 신호에 대해서는, 하모닉에 해당하는 펄스들을
- 15 각각 인코딩하기 보다는, 하모닉 트랙 전체를 인코딩함으로써, 비트수의 거의 증가시키지 않고 복원율을 높일 수 있다.

셋째, 프레임마다 각 오디오 특성에 맞게 총 4 개의 모드에 해당하는 인코딩 및 디코딩 방식들 중 하나를 적응적으로 적용함으로써, 음질을 향상시킬 수 있다.

넷째, MDCT 를 적용할 경우, MDCT 의 특성을 고려하여, 메인 펄스 그에 인접한 서브 펄스를 추출함으로써, 특정 주파수 대역에 매핑되는 펄스를 정확하게 추출함으로써, 비-일반모드 인코딩 방식의 성능을 보다 높일 수 있다.

다섯째, 하모닉 모드에서 복수의 하모닉 트랙에서 최대 펄스와 그의
5 인접한 펄스만을 추출하여 별도로 양자화함으로써, 비트수를 절감할 수 있다.

여섯째, 하모닉 모드에 있어서, 동일한 피치를 갖는 하나의 그룹에 속하는 하모닉 트랙에 대해서, 정해진 위치 셋 중 하나로 시작 위치를 설정함으로써, 복수의 하모닉 트랙들의 각 시작 위치를 표시하는 데 있어서, 비트수를 절감할 수 있다.

10 【Description of Drawings】

도 1 은 본 발명의 실시예에 따른 오디오 신호 처리 장치 중 인코더의 구성도.

도 2 는 프레임간 유사성(tonality) 판단의 예를 설명하기 위한 도면.

도 3 은 일반모드 또는 비-일반모드로 코딩하기에 적합한 신호의 예들.

15 도 4 는 일반모드 인코딩부(140)의 세부 구성도.

도 5 는 일반모드로 인코딩한 경우의 선택스의 일 예.

도 6 은 비-일반모드 인코딩부(150)의 세부 구성도.

도 7 및 도 8 은 펄스 추출 과정을 설명하기 위한 도면.

20 도 9 는 펄스 추출 이전의 신호(SWB 신호) 및 펄스 추출 이후 신호(원본 노이즈 신호)의 일 예.

도 10 은 참조 노이즈 생성과정을 설명하기 위한 도면.

도 11 은 비-일반모드로 인코딩한 경우의 신택스의 일 예.

도 12 는 특정 오디오 신호에 대해 일반 모드 및 비-일반모드로 인코딩된 결과를 보여주기 위한 도면.

도 13 은 하모닉비율 판단부(160)의 세부 구성도.

5 도 14 는 하모닉비율이 높은 오디오 신호의 예.

도 15 는 년-하모닉모드 인코딩부(170)의 세부 구성도.

도 16 은 년-하모닉모드인 경우, 고정 펄스를 추출하는 규칙을 설명하기 위한 도면.

도 17 은 년-하모닉모드로 인코딩한 경우의 신택스의 일 예.

10 도 18 은 하모닉모드 인코딩부(180)의 세부 구성도.

도 19 는 하모닉트랙의 추출을 설명하기 위한 도면

도 20 은 하모닉트랙 위치정보의 양자화를 설명하기 위한 도면.

도 21 은 하모닉모드로 인코딩한 경우의 신택스의 일 예.

도 22 은 특정 오디오 신호에 대해 년-하모닉 모드 및 하모닉모드로
15 인코딩된 결과를 보여주기 위한 도면.

도 23 은 본 발명의 실시예에 따른 오디오 신호 처리 장치 중 디코더의 구성도.

도 24 는 본 발명의 실시예에 따른 오디오 신호 처리 장치가 구현된 제품의 개략적인 구성도.

20 도 25 은 본 발명의 일 실시예에 따른 오디오 신호 처리 장치가 구현된 제품들의 관계도.

【Best Mode】

상기와 같은 목적을 달성하기 위하여 본 발명에 따른 오디오 신호 처리 방법은, 오디오 신호에 대해서 주파수 변환을 수행함으로써, 복수 개의 주파수 변환 계수들을 획득하는 단계; 상기 복수 개의 주파수 변환 계수들 중 고주파 대역의 주파수 변환 계수들에 대해서, 펄스 비율을 근거로 일반 모드 및 비-일반모드 중 하나를 선택하는 단계; 및, 상기 비-일반 모드가 선택된 경우, 다음 단계가 수행되는 단계; 상기 고주파 대역의 주파수 변환 계수들로부터 미리 결정된 개수의 펄스들을 추출하고, 펄스 정보를 생성하는 단계; 상기 고주파 대역의 주파수 변환 계수들로부터 상기 펄스들을 제외한 원본 노이즈 신호를 생성하는 단계; 상기 복수 개의 주파수 변환 계수들 중 저주파 대역의 주파수 변환 계수를 이용하여 참조 노이즈 신호를 생성하는 단계; 및, 상기 원본 노이즈 신호 및 상기 참조 노이즈 신호를 이용하여, 노이즈 위치정보 및 노이즈 에너지정보를 생성하는 단계를 포함한다.

본 발명에 따르면, 상기 펄스 비율은, 현재 프레임의 전체 에너지에 대한 복수의 펄스들에 대한 에너지의 비율일 수 있다.

본 발명에 따르면, 미리 결정된 개수의 펄스들을 추출하는 단계는 가장 큰 에너지를 갖는 메인 펄스를 추출하는 단계; 상기 메인 펄스에 인접한 서브 펄스를 추출하는 단계; 및 상기 고주파 대역의 주파수 변환 계수에서 상기 메인 펄스 및 상기 서브 펄스를 제외시킴으로써 타깃 노이즈 신호를 생성하는

단계; 상기 타깃 노이즈 신호에 대해서 상기 메인 펄스 및 상기 서브 펄스를 추출하는 것을 미리 정해진 횟수만큼 반복하는 단계를 포함할 수 있다.

본 발명에 따르면, 상기 펄스 정보는, 펄스 위치 정보, 펄스 부호 정보, 펄스 크기 정보, 펄스 서브밴드 정보 중 하나 이상을 포함할 수 있다.

- 5 본 발명에 따르면, 상기 참조 노이즈 신호를 생성하는 단계는, 저주파 대역의 전체 에너지를 근거로 임계치를 설정하는 단계; 및 상기 임계치 이상의 펄스를 제외시킴으로써, 상기 참조 노이즈 신호를 생성하는 단계를 포함할 수 있다.

- 본 발명에 따르면, 상기 노이즈 에너지정보를 생성하는 단계는, 상기
- 10 미리 결정된 개수의 펄스의 에너지를 생성하는 단계; 상기 원본 노이즈 신호의 에너지를 생성하는 단계; 상기 펄스의 에너지 및 상기 원본 노이즈 신호의 에너지를 이용하여 펄스 비율을 획득하는 단계; 및,

 상기 펄스 비율을 노이즈 에너지 정보로서 생성하는 단계를 포함할 수 있다.

- 15 본 발명의 또 다른 측면에 따르면, 오디오 신호에 대해서 주파수 변환을 수행함으로써, 복수 개의 주파수 변환 계수들을 획득하는 주파수 변환부; 상기 복수 개의 주파수 변환 계수들 중 고주파 대역의 주파수 변환 계수들에 대해서, 펄스 비율을 근거로 일반 모드 및 비-일반모드 중 하나를 선택하는 펄스비율 판단부; 상기 비-일반 모드가 선택된 경우에 동작하며, 다음을

- 포함하는 비-일반모드 인코딩부; 상기 고주파 대역의 주파수 변환 계수들로부터 미리 결정된 개수의 펄스들을 추출하고, 펄스 정보를 생성하는 펄스 추출부; 상기 복수 개의 주파수 변환 계수들 중 저주파 대역의 주파수 변환 계수를 이용하여 참조 노이즈 신호를 생성하는 참조노이즈 생성부; 및,
- 5 원본 노이즈 신호 및 상기 참조 노이즈 신호를 이용하여, 노이즈 위치정보 및 노이즈 에너지정보를 생성하는 노이즈 탐색부를 포함하고, 상기 원본 노이즈 신호는, 상기 고주파 대역의 주파수 변환 계수들로부터 상기 펄스들을 제외함으로써 생성되는 오디오 신호 처리 장치가 제공된다.

- 본 발명의 또 다른 측면에 따르면, 현재 프레임이 일반 모드 인지
- 10 아니면 비-일반 모드인지 여부를 지시하는 제 2 모드 정보를 수신하는 단계; 상기 제 2 모드 정보가 현재 프레임이 비-일반 모드임을 지시하는 경우, 펄스 정보, 노이즈 위치정보 및 노이즈 에너지정보를 수신하는 단계; 상기 펄스 정보를 이용하여, 주파수 변환 계수들에 대해서 미리 결정된 개수의 펄스들을 생성하는 단계; 상기 노이즈 위치정보에 대응하는 저주파 대역의 주파수 변환
- 15 계수를 이용하여 참조 노이즈 신호를 생성하는 단계; 상기 노이즈 에너지정보를 이용하여 상기 참조 노이즈 신호의 에너지를 조절하는 단계; 및, 상기 에너지 조절된 참조 노이즈 신호 및 상기 복수의 펄스를 이용하여 고주파 대역에 대응하는 주파수 변환 계수들을 생성하는 단계를 포함하는 오디오 신호 처리 방법이 제공된다.

본 발명에 따르면, 오디오 신호를 수신하는 단계; 오디오 신호에 대해서 주파수 변환을 수행하여, 복수개의 주파수 변환 계수들을 획득하는 단계; 상기 주파수 변환 계수에 대해서, 하모닉 비율을 근거로 년-하모닉 모드 및 하모닉 모드 중 하나를 선택하는 단계; 상기 하모닉 모드가 선택된 경우, 다음 단계가

5 수행되는 단계; 제 1 피치에 해당하는 제 1 그룹의 하모닉 트랙들을 결정하는 단계; 제 2 피치에 해당하는 제 2 그룹의 하모닉 트랙들을 결정하는 단계; 및, 상기 복수의 하모닉 트랙들의 시작 위치 정보를 생성하는 단계를 포함하고, 상기 제 1 그룹의 하모닉 트랙들은 제 1 하모닉 트랙 및 제 2 하모닉 트랙을 포함하고, 상기 제 2 그룹의 하모닉 트랙들은 제 3 하모닉 트랙 및 제 4 하모닉

10 트랙을 포함하고, 상기 제 1 하모닉 트랙 및 제 3 하모닉 트랙의 시작 위치 정보는 제 1 위치 셋 중 하나에 대응되고, 상기 제 2 하모닉 트랙 및 제 4 하모닉 트랙의 시작 위치 정보는 제 2 위치 셋 중 하나에 대응되는 것을 특징으로 하는 오디오 신호 처리 방법이 제공된다.

본 발명에 따르면, 상기 하모닉 비율은, 복수의 하모닉 트랙들의 에너지

15 및 복수 개의 펄스의 에너지를 근거로 생성된 것일 수 있다.

본 발명에 따르면, 상기 제 1 위치 셋은 짝수 위치들에 해당하고, 제 2 위치 셋은 홀수 위치들에 해당할 수 있다.

본 발명에 따르면, 상기 제 1 하모닉 트랙 중에서의 최대 펄스, 및 그의 인접 펄스, 상기 제 2 하모닉 트랙 중에서 최대 펄스, 및 그의 인접 펄스를

포함하는 제 1 타깃 벡터를 생성하는 단계; 상기 제 3 하모닉 트랙 중에서의 최대 펄스, 그의 인접 펄스, 상기 제 4 하모닉 트랙 중에서의 최대 펄스, 그의 인접 펄스를 포함하는 제 2 타깃 벡터를 생성하는 단계; 상기 제 1 타깃 벡터 및 상기 제 2 타깃 벡터를 벡터-양자화하는 단계; 상기 하모닉 트랙들에서 상기

5 제 1 타깃 벡터 및 상기 제 2 타깃 벡터를 제외한 대상에 대해 주파수 변환하는 단계를 더 포함할 수 있다.

본 발명에 따르면, 제 1 하모닉 트랙은 제 1 피치를 갖는 복수의 펄스들의 집합이고, 제 2 하모닉 트랙은 제 1 피치를 갖는 복수의 펄스들의 집합이고, 제 3 하모닉 트랙은 제 2 피치를 갖는 복수의 펄스들의 집합이고,

10 제 4 하모닉 트랙은 제 2 피치를 갖는 복수의 펄스들의 집합일 수 있다.

본 발명에 따르면, 상기 제 1 피치 및 상기 제 2 피치를 지시하는 피치 정보를 생성하는 단계를 더 포함할 수 있다.

본 발명의 또 다른 측면에 따르면, 제 1 피치에 해당하는 제 1 그룹의 하모닉 트랙들 및 제 2 피치에 해당하는 제 2 그룹의 하모닉 트랙을 포함하는

15 복수의 하모닉 트랙들의 시작 위치 정보를 수신하는 단계; 상기 시작 위치 정보에 대응하는 복수의 하모닉 트랙을 생성하는 단계; 및, 상기 복수의 하모닉 트랙을 이용하여 현재 프레임에 해당하는 오디오 신호를 생성하는 단계를 포함하고, 상기 제 1 그룹의 하모닉 트랙들은 제 1 하모닉 트랙 및 제 2 하모닉 트랙을 포함하고, 상기 제 2 그룹의 하모닉 트랙들은 제 3 하모닉 트랙

및 제 4 하모닉 트랙을 포함하고, 상기 제 1 하모닉 트랙 및 제 3 하모닉 트랙의 시작 위치 정보는 제 1 위치 셋 중 하나에 대응되고, 상기 제 2 하모닉 트랙 및 제 4 하모닉 트랙의 시작 위치 정보는 제 2 위치 셋 중 하나에 대응되는 오디오 신호 처리 방법이 제공된다.

- 5 본 발명의 또 다른 측면에 따르면, 오디오 신호에 대해서 주파수 변환을 수행하여, 복수개의 주파수 변환 계수들을 획득하는 단계; 상기 주파수 변환 계수들에 대해서, 프레임간 유사성을 근거로 년-토널 모드 및 토널 모드를 선택하는 단계; 상기 년-토널 모드가 선택된 경우, 펄스 비율을 근거로 일반 모드 및 비-일반모드 중 하나를 선택하는 단계; 상기 토널 모드가 선택된 경우, 하모닉 비율을 근거로 년-하모닉 모드 및 하모닉 모드 중 하나를 선택하는 단계; 및, 상기 선택된 모드에 따라서 상기 오디오 신호를 인코딩함으로써 파라미터를 생성하는 단계를 포함하고, 상기 일반 모드인 경우, 상기 파라미터는 인벨롭 위치정보 및 스케일링 정보를 포함하고, 상기 비-일반 모드인 경우, 상기 파라미터는 펄스정보 및 노이즈 에너지정보를 포함하고
- 10 상기 년-하모닉 모드인 경우, 상기 파라미터는 서브밴드별로 개수가 미리 정해진 고정 펄스들에 대한 정보인 고정 펄스정보를 포함하고, 상기 하모닉 모드인 경우, 상기 파라미터는 제 1 그룹의 하모닉 트랙들의 위치정보들, 제 2 그룹의 하모닉 트랙들의 위치정보들을 포함하는 오디오 신호 처리 방법이 제공된다.

본 발명에 따르면, 상기 선택된 모드에 따라서, 제 1 모드 정보 및 제 2 모드 정보를 생성하는 단계를 더 포함하고, 상기 제 1 모드 정보는, 상기 넌-토널 모드 및 상기 토널 모드 중 하나를 지시하고, 상기 제 2 모드 정보는, 상기 제 1 모드 정보가 넌-토널 모드를 지시하는 경우, 상기 일반모드 또는

5 상기 비-일반 모드 중 하나를 지시하고, 상기 제 1 모드 정보가 토널 모드를 지시하는 경우, 상기 넌-하모닉 모드 및 상기 하모닉 모드 중 하나를 지시할 수 있다.

본 발명의 또 다른 측면에 따르면, 비트스트림을 통해 제 1 모드 정보 및 제 2 모드 정보를 추출하는 단계; 상기 제 1 모드 정보 및 제 2 모드 정보를

10 근거로 하여, 현재 프레임에 해당하는 현재 모드를 결정하는 단계; 상기 현재 모드가 일반 모드인 경우, 인벨롭 위치 정보 및 스케일링 정보를 이용하여 현재 프레임의 오디오 신호를 복원하는 단계; 상기 현재 모드가 비-일반 모드인 경우, 펄스정보 및 노이즈 에너지 정보를 이용하여 현재 프레임의 오디오 신호를 복원하는 단계; 상기 현재 모드가 넌-하모닉 모드인 경우,

15 서브밴드별로 개수가 미리 정해진 고정 펄스들에 대한 정보인 고정펄스정보를 이용하여 현재 프레임의 오디오 신호를 복원하는 단계; 및, 상기 현재 모드가 하모닉 모드인 경우, 제 1 그룹의 하모닉 트랙의 위치정보들 및 제 2 그룹의 하모닉 트랙의 위치정보들을 이용하여 현재 프레임의 오디오 신호를 복원하는 단계를 포함하는 오디오 신호 처리 방법이 제공된다.

【Mode for Invention】

이하 첨부된 도면을 참조로 본 발명의 바람직한 실시예를 상세히 설명하기로 한다. 이에 앞서, 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니되며, 발명자는 그

5 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서, 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 가장 바람직한 일 실시예에 불과할 뿐이고 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서

10 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.

본 발명에서 다음 용어는 다음과 같은 기준으로 해석될 수 있고, 기재되지 않은 용어라도 하기 취지에 따라 해석될 수 있다. 코딩은 경우에 따라 인코딩 또는 디코딩으로 해석될 수 있고, 정보(information)는 값(values),

15 파라미터(parameter), 계수(coefficients), 성분(elements) 등을 모두 아우르는 용어로서, 경우에 따라 의미는 달리 해석될 수 있는 바, 그러나 본 발명은 이에 한정되지 아니한다.

여기서 오디오 신호(audio signal)란, 광의로는, 비디오 신호와 구분되는 개념으로서, 재생 시 청각으로 식별할 수 있는 신호를 지칭하고, 협의로는,

음성(speech) 신호와 구분되는 개념으로서, 음성 특성이 없거나 적은 신호를 의미한다. 본 발명에서의 오디오 신호는 광의로 해석되어야 하며 음성 신호와 구분되어 사용될 때 협의의 오디오 신호로 이해될 수 있다.

또한 코딩이란, 인코딩만을 지칭할 수도 있지만, 인코딩 및 디코딩을

5 모두 포함하는 개념으로 사용될 수도 있다.

도 1 은 본 발명의 실시예에 따른 오디오 신호 처리 장치 중 인코더의 구성을 보여주는 도면이다. 실시예에 따른 인코더(100)는 펄스비율 판단부(130), 하모닉비율 판단부(160), 비-일반모드 인코딩부(150), 및 하모닉모드 인코딩부(180) 중 하나 이상을 포함하고, 주파수 변환부(110), 유사성
10 판단부(120), 일반모드 인코딩부(140), 및 난-하모닉모드 인코딩부(180) 중 하나 이상을 더 포함할 수 있다.

개괄적으로 설명하면, 총 네 가지의 코딩 모드가 존재하는 바, 1) 일반 모드, 2) 비-일반 모드, 3) 난-하모닉모드, 4) 하모닉모드가 존재하는 데, 여기서
1) 일반 모드 및 2) 비-일반 모드는 난-토널 모드에 해당하고, 3) 난-하모닉
15 모드 및 4) 하모닉모드는 토널 모드에 해당한다.

비-토널 모드인지 아니면 토널 모드인지 여부는 프레임간의 유사성의 정도에 따라서 유사성 판단부(120)에 의해 결정된다. 즉 유사성이 높지 않은 경우에는 비-토널 모드이고, 유사성이 높은 경우에는 토널 모드에 해당한다.
한편, 비-토널 모드인 경우, 펄스비율 판단부(130)는 펄스비율(구체적으로, 전체
20 에너지 대비 펄스의 에너지의 비율)이 높은 경우 1) 일반 모드 방식을

적용하고 펄스비율이 낮은 경우에는 2) 비-일반 모드 방식을 적용할 것을 결정한다.

- 또한, 톤널 모드일 때는, 하모닉비율 판단부(160)는 하모닉비율(구체적으로, 펄스의 에너지 대비 하모닉 트랙의 에너지의 비율)이
- 5 높지 않는 경우 3) 난-하모닉 모드를 적용할 것을 결정하고, 4) 하모닉비율이 높은 경우에는 4) 하모닉 모드를 적용할 것을 결정한다.

- 주파수 변환부(110)는 입력되는 오디오 신호에 대해서 주파수 변환을 수행함으로써, 복수 개의 주파수 변환 계수들을 획득한다. 여기서 주파수 변환에는 MDCT(Modified Discrete Cosine Transform), FFT(Fast Fourier Transform)
- 10 방식 등이 적용될 수 있지만 본 발명은 이에 한정되지 아니한다.

- 주파수 변환 계수들은 상대적으로 저주파 대역에 해당하는 주파수 변환 계수들과, 고주파 대역에 해당하는 주파수 변환 계수들이 존재할 수 있다. 저주파 대역의 주파수 변환 계수들을 광대역 밴드 신호, WB(Wide Band) 신호, 또는 WB 계수라 지칭하고, 고주파 대역의 주파수 변환 계수들을 초광대역
- 15 밴드 신호, SWB(Super Wide Band) 신호 또는 WB 계수라고도 지칭하고자 한다. 한편 저주파 대역 및 고주파 대역의 기준은 약 7kHz 가 될 수 있지만, 본 발명은 특정 주파수에 한정되지 아니한다.

- 한편, 주파수 변환 방식이 MDCT 의 경우, 전체 오디오 신호에 대해서 총 640 개의 주파수 변환 계수들이 생성될 수 있다. 이때, 가장 낮은 대역에
- 20 해당하는 약 280 개의 계수를 WB 신호로 지칭하고, 그 다음의 대역에

해당하는 약 280 개의 계수를 SWB 신호라 지칭할 수도 있지만, 역시 본 발명은 이에 한정되지 아니한다.

유사성 판단부(120)는 입력 오디오 신호에 대해서 프레임간의 유사성을 판단하는데, 프레임간 유사성이란, 현재 프레임의 주파수 변환 계수들의
5 스펙트럼이 이전 프레임의 주파수 변환 계수들의 스펙트럼과 얼마나 유사한지에 관한 것이다. 프레임간 유사성은 음조성(tonality)라고 지칭하기도 한다. 프레임간 유사성에 관한 구체적인 수식에 대한 설명은 생략하고자 한다.

도 2 는 프레임간 유사성(tonality) 판단의 예를 설명하기 위한 도면이다. 도 2 의 (A)를 참조하면, 이전 프레임의 스펙트럼 및 현재 프레임의
10 스펙트럼의 예가 도시되어 있다. 직관적으로 판단해볼 때, 주파수 빈이 약 40 에서 60 사이인 구간에서 유사성이 가장 적은 것을 알 수 있다. 도 2 의 (B)를 참조하면, 직관에 의한 결과와 마찬가지로 유사성은 40~60 구간에서 가장 낮은 수치를 보임을 알 수 있다.

유사성 판단부(120)가 프레임간 유사성을 판단한 결과, 유사성이 낮은
15 경우는 노이즈와 유사한 신호로서 넌-토널 모드(non-tonal)로 분류하고, 유사성이 높은 경우는 노이즈와는 다른 특성을 갖는 신호로서 토널 모드(tonal)로 분류한다. 그리고 해당 프레임이 넌-토널 모드에 해당하는지 토널 모드에 해당하는지를 지시하는 제 1 모드 정보를 생성하여 디코더에 전달한다.

넌-토널 모드로 판단한 경우(예를 들어 제 1 모드 정보가 0 인 경우),
20 고주파 대역의 주파수 변환 계수들을 펄스비율 판단부(130)로 전달하고,

반대인 경우(예를 들어, 제 1 모드 정보가 1 인 경우), 상기 계수들을 하모닉비율 판단부(160)에 전달한다.

다시 도 1 을 참고하면, 프레임간 유사성을 판단한 결과 유사성이 낮은 경우, 즉 년-토널 모드인 경우, 펄스비율 판단부(130)가 활성화된다.

- 5 펄스비율 판단부(130)는 펄스 비율 즉, 현재 프레임의 전체 에너지 대비 복수의 펄스들에 대한 에너지 비율을 근거로 하여 일반 모드 또는 비-일반 모드를 결정한다. 여기서 펄스란, 주파수 변환 계수의 도메인(예: MDCT 도메인)에서 상대적으로 에너지가 높은 계수를 의미한다.

- 10 도 3 은 일반모드 또는 비-일반모드로 코딩하기에 적합한 신호의 예들이다. 도 3 의 (A)를 참조하면, 특정 주파수 대역에 치우치지 않고 전체 주파수가 골고루 포함되어 있음을 알 수 있다. 이와 같이 노이즈와 유사한 특성을 갖는 경우, 일반 모드로 코딩하기에 적합한 신호라 할 수 있다. 도 3 의 (B)를 참조하면, 전체 주파수 신호가 고루 포함되기 보다는 특정 대역(라인)이 강한 것을 알 수 있다. 상기 특정 대역은 주파수 변환 계수의 도메인에서는
- 15 펄스로 나타날 수 있는데, 이 펄스의 에너지가 전체 에너지에 비해서 높은 경우, 펄스 비율이 높으므로 비-일반 모드로 인코딩하기 적합한 신호라 할 수 있다. 도 3 의 (A)와 같은 신호는 노이즈에 가까운 신호이고, 도 3 의 (B)는 타악기 소리에 가까운 신호일 수 있다.

- 20 펄스비율 판단부(130)가 주파수 변환 계수의 도메인 상에서 높은 에너지를 갖는 펄스들을 추출하는 과정에 관해서는, 비-일반 모드의 코딩

방식이 적용되었을 때 수행되는 펄스 추출 과정과 동일할 수 있으므로, 이
추후 비-일반모드 인코딩부(150)에서의 세부 구성과 함께 설명하고자 한다.

만약에, 총 8 개의 펄스를 추출한 경우, 다음과 같이 표현될 수 있다.

[수학식 1]

$$5 \quad P(j) = \max(\{M_{32}(k+280)\}^2), \quad j = 0, \dots, 7 \quad k = 280, \dots, 560$$

여기서, $M_{32}(k)$ 는 SWB 계수(고주파 대역의 주파수 변환 계수),

k는 주파수 변환 계수의 인덱스,

$P(j)$ 는 펄스(또는 피크), j는 펄스 인덱스

여기서, 펄스 비율은 다음 수학식과 같이 표현될 수 있다.

10 [수학식 2]

$$R_{peak8} = \frac{E_{peak}}{E_{total}}$$

$$\text{여기서, } E_{peak} = \sum_{k=0}^7 \{P(k)^2\}, \quad E_{total} = \sum_{k=0}^{280} \{P(k+280)^2\}$$

여기서, R_{peak8} 는 펄스 비율, E_{peak} 는 펄스의 총 에너지, E_{total} 는 전체
에너지.

15 이와 같이 펄스 비율(R_{peak8})을 추정 한 후에, 펄스 비율이 특정
기준값(예: 0.6)을 넘지 못할 경우, 일반 모드(generic mode)로 결정하고,
기준값을 초과할 경우에는 비-일반 모드(non-generic mode)로 결정한다.

다시, 도 1 을 참조하면, 펄스비율 판단부(130)는 상기와 같은 과정을
통해 펄스 비율을 기준으로 일반 모드 또는 비-일반 모드를 결정하고, 난-토널

모드일 때, 일반 모드인지 비-일반 모드인지 여부를 지시하는 제 2 모드 정보를 생성하여 디코더에 전송한다. 그러면 일반모드 인코딩부(140)의 세부 구성 및 비일반모드 인코딩부(150)의 세부 구성을 각각 별도의 도면을 참조하면서 설명하고자 한다.

- 5 그런 다음 하모닉비율 판단부(160), 년-하모닉모드 인코딩부(170) 및 하모닉모드 인코딩부(180)의 각 세부구성에 대해서 별도 도면을 참조하면서 설명하고자 한다.

도 4 는 일반모드 인코딩부(140)의 세부 구성을 보여주는 도면이고, 도 5 는 일반모드로 인코딩한 경우의 선택스의 일 예이다.

- 10 우선 도 4 를 참조하면, 일반모드 인코딩부(140)는 정규화부(142), 서브밴드 생성부(144), 및 탐색부(146)를 포함한다. 일반모드인 경우에는, 고주파대역의 신호(SWB 신호)는 부호화된 저주파 대역의 신호(WB 신호)의 인벨롭과의 유사 정도를 이용하여 부호화한다.

- 정규화부(142)는 WB 신호의 인벨롭을 지수 도메인(logarithmic
15 domain)에서 정규화한다. 여기서 WB 신호는 디코더에서도 알 수 있는 신호여야 하므로, 인코딩된 WB 신호를 이용하여 복원된 신호인 것이 바람직하다. WB 신호의 인벨롭(포락선)이 급속도로 변화하기 때문에, 2 개의 스케일링 팩터의 양자화는 정확히 할 수 없을 수 있기 때문에 지수 도메인에서 정규화하는 과정이 필요할 수 있다.

서브밴드 생성부(144)는 SWB 신호를 복수 개(약 4 개)의 서브 밴드로 분할한다. 예를 들어, SWB 신호의 주파수 변환 계수들이 총 280 개인 경우, 각 서브밴드는, 40, 70, 70, 100 개의 계수들을 가질 수 있다.

5 탐색부(146)는 WB 신호의 정규화된 인벨롭을 탐색함으로써, SWB
신호의 각 서브밴드와 유사도를 연산하고, 이 유사도를 기반으로 하여 각 서브밴드와 유사한 인벨롭의 구간을 갖는 최대유사 WB 신호를 결정한다. 이 최대 유사 WB 신호의 시작 위치를 인벨롭 위치정보로서 생성한다.

그런 다음, 탐색부(146)는 원본 SWB 신호와 청각적으로 유사하게 만들기 위해, 두 개의 스케일링 정보를 결정할 수 있다. 이따, 제 1 스케일링
10 정보는 선형 도메인에서 서브밴드별로 결정되며, 제 2 스케일링 정보는 로그 도메인에서 서브밴드별로 결정될 수 있다.

이와 같이 일반모드 인코딩부(140)는 WB 신호의 인벨롭을 이용하여 SWB 신호를 인코딩하며, 인벨롭 위치정보 및 스케일링 정보를 생성한다.

도 5 를 참조하면, 일반 모드인 경우의 선택스의 일 예로서, 년-토널인지
15 토널인지를 지시하는 제 1 모드 정보, 년-토널일 경우 일반 모드인지 비-일반 모드인지를 지시하는 제 2 모드 정보가 각각 1 비트씩 할당된다. 그리고, 인벨롭 위치정보가 각 서브밴드별로 총 30 비트가 할당될 수 있다.

스케일링 정보는 각 서브밴드별 스케일링 부호 정보가 총 4 비트,
서브밴드별 제 1 스케일링 정보(총 4 개)는 총 16 비트가 할당될 수 있고,
20 서브밴드별 제 2 스케일링 정보들 총 4 개는 8 비트 코드북을 기반으로

벡터양자화되어 총 8 비트가 할당될 수 있으나, 본 발명은 이에 한정되지 아니한다.

이하, 도 6 이하를 참조하면서, 비-일반모드에서의 인코딩과정에 대해서 설명하고자 한다. 도 6 은 비-일반모드 인코딩부(150)의 세부 구성을 보여주는
5 도면이다. 도 6 을 참조하면, 비-일반모드 인코딩부(150)는 펄스 추출부(152), 참조 노이즈 생성부(154) 및 노이즈 탐색부(156)를 포함한다.

펄스 추출부(152)는 고주파 대역의 주파수 변환 계수들(SWB
신호)로부터 미리 결정된 개수의 펄스들을 추출하고, 상기 펄스들에 대한
정보인 펄스 정보(예: 펄스 위치 정보, 펄스 부호 정보, 펄스 크기 정보 등)을
10 생성한다. 여기서 펄스는 앞서 펄스비율 판단부(130)에서 정의한 펄스와
마찬가지이다. 이하, 도 7 내지 도 9 를 참조하면서, 펄스 추출에 대한 과정의
일 실시예에 대해서 구체적으로 설명하고자 한다.

우선, 펄스 추출부(152)는 SWB 신호를 아래와 같이 복수 개의 서브밴드
신호로 분할한다. 이때, 각 서브밴드는 총 64 개의 주파수 변환 계수들에
15 대응될 수 있다.

[수학식 3]

$$M_{32}^0(k) = M_{32}(k + 280), \quad k = 0, \dots, 63$$

$$M_{32}^1(k) = M_{32}(k + 344), \quad k = 0, \dots, 63$$

$$M_{32}^2(k) = M_{32}(k + 408), \quad k = 0, \dots, 63$$

$$M_{32}^3(k) = M_{32}(k + 472), \quad k = 0, \dots, 63$$

$M_{32}^0(k)$ 는 SWB 신호 중 제 1 서브밴드

그런 다음, 서브밴드별 에너지를 다음과 같이 산출한다.

[수학식 4]

$$E^0 = \sum_{k=0}^{63} \{M_{32}(k+280)\}^2$$

$$E^1 = \sum_{k=0}^{63} \{M_{32}(k+344)\}^2$$

$$E^2 = \sum_{k=0}^{63} \{M_{32}(k+408)\}^2$$

$$E^3 = \sum_{k=0}^{63} \{M_{32}(k+472)\}^2$$

E^0 은 제 1 서브밴드의 에너지

도 7 및 도 8 은 펄스 추출 과정을 설명하기 위한 도면이다. 우선 도 5 7 의 (A)를 참조하면, SWB 가 총 네 개의 서브밴드가 존재하고, 각 서브밴드별 펄스의 예가 도시되어 있다.

그런 다음, 서브밴드별 에너지(E^0, E^1, E^2, E^3)가 가장 큰 서브밴드($j=0, 1, 2, 3$ 중 하나)를 선택한다. 도 7 의 (B)를 참조하면, 제 1 서브밴드의 에너지(E^0)가 가장 높아서, 제 1 서브밴드($j=0$)가 선택된 예가 도시되어 있다.

10 그런 다음, 해당 서브밴드 내에서 가장 에너지가 높은 펄스를 메인 펄스로 결정한다. 그런 다음 메인 펄스와 인접한 두 개의 펄스들, 즉 메인 펄스의 왼쪽 펄스 및 오른쪽 펄스 중 에너지가 높은 펄스를 서브 펄스로 결정한다. 도 7 의 (C)를 참조하면, 제 1 서브밴드 중에서 메인 펄스와 서브 펄스를 결정한 예가 도시되어 있다.

15 특히, 메인 펄스 및 그에 인접한 서브 펄스를 추출하는 과정은, 주파수 변환 계수들이 MDCT 를 통해 생성된 계수들인 경우에 바람직하다. 왜냐하면, MDCT 의 독특한 특성 때문인데, 이 특성은 시간 이동(time shift)에 민감하고 위상 변화성(phase-variant)을 갖는 특징이다. 따라서, 주파수 분해능(frequency

resolution)이 정확하지 않으므로, 하나의 특정 주파수가 1 개의 MDCT 계수에 대응되지 않고, 2 개 이상의 MDCT 계수에 대응될 수 있다. 따라서, MDCT 도메인에서 보다 정확하게 펄스를 추출하기 위해서, MDCT 의 메인 펄스만을 추출하기 보다는, 그에 인접한 서브 펄스를 부가적으로 추출한다.

- 5 한편, 서브 펄스는 메인 펄스의 좌측 또는 우측에 인접하기 때문에, 서브 펄스의 위치정보는 메인 펄스의 좌측인지 또는 우측인지를 지시하는 1 비트만으로 인코딩이 가능하기 때문에, 상대적으로 적은 비트를 사용하면서도 보다 정확한 펄스를 추정하는 것이 가능하다.

상기 메인 펄스 및 서브 펄스의 추출과정을 논리식으로 정리하면

- 10 아래와 같다. 단 아래 표현에 본 발명은 한정되지 아니한다.

$$M_{32}^{\text{max}}(k) = \text{subband of maximum Energy}$$

$$\text{index} = \text{peak position in subband } M_{32}^{\text{max}}$$

$$\text{if } (\text{index} == 0 \text{ or } |M_{32}^{\text{max}}(\text{index} - 1)| < |M_{32}^{\text{max}}(\text{index} + 1)|)$$

$$P_{\text{pos}}(1) = \text{index} + 1$$

$$P_{\text{amp}}(1) = |M_{32}^{\text{max}}(\text{index} + 1)|$$

$$\text{if } (P_{\text{amp}}(1) < 0)$$

$$P_{\text{sign}}(1) = 1$$

$$\text{else}$$

$$P_{\text{sign}}(1) = 0$$

$$\text{else}$$

$$P_{\text{pos}}(1) = \text{index} - 1$$

$$P_{\text{amp}}(1) = |M_{32}^{\text{max}}(\text{index} - 1)|$$

$$\text{if } (P_{\text{amp}}(1) < 0)$$

$$P_{\text{sign}}(1) = 1$$

$$\text{else}$$

$$P_{\text{sign}}(1) = 0$$

이와 같이 펄스 추출부(152)는 SWB 신호에서 앞서 추출된 제 1 세트의 메인 펄스 및 서브 펄스를 제외시킴으로써 타깃 노이즈 신호를 생성한다.

도 8 의 (A)를 참조하면, 도 7 의 (C)에서 추출된 제 1 세트의 펄스가 제외되어 있음을 알 수 있다. 그 타깃 노이즈 신호에 대해서 상기 메인 펄스
5 및 서브 펄스의 추출과정을 반복한다. 즉, 가장 큰 에너지를 갖는 서브밴드를 결정하고, 해당 서브밴드 내에서 가장 큰 에너지를 갖는 펄스를 메인 펄스로 결정하고, 그 인접한 펄스들 중 하나를 서브 펄스로 결정하는 것이다. 이와 같이 추출된 제 2 세트의 메인 펄스 및 서브 펄스를 제외시킴으로써 타깃
10 노이즈 신호를 다시 정의함으로써, 제 N 세트까지 이 과정을 반복한다. 예를 들어, 상기 과정을 제 3 세트까지 반복하고, 상기 제 3 세트를 제외한 타깃 노이즈 신호에 대해서 2 개의 단독 펄스를 더 추출할 수 있다. 여기서 단독 펄스란 상기와 같이 메인 및 서브의 관계 없이, 타깃 노이즈 신호 중 가장
에너지가 높은 펄스를 의미한다.

펄스 추출부(152)는 상기와 같이 정해진 개수의 펄스들을 추출한 후에,
15 이 펄스에 대한 정보인 펄스 정보를 생성한다. 여기서 정해진 개수의 펄스들은 예를 들어 총 8 개(총 3 세트의 메인 펄스들 및 서브 펄스들, 총 3 개의 단독 펄스들)일 수 있지만, 본 발명은 이에 한정되지 아니한다. 펄스 정보는, 펄스 위치 정보, 펄스 부호 정보, 펄스 크기 정보, 펄스 서브밴드 정보 중 하나 이상을 포함할 수 있다. 여기서 먼저 펄스 서브밴드 정보는, 해당 펄스가 몇
20 번째 서브밴드에 속하는지에 대한 정보이다.

도 11 은 비-일반모드로 인코딩한 경우의 선택스의 일 예인데, 여기서 펄스 관련 정보에 대해서만 참조하고자 한다. 도 11 은 서브밴드가 총 4 개이고, 펄스가 총 8 개일 때(각 3 개씩의 메인 펄스 및 서브 펄스들, 단독 펄스 2 개)의 경우이다. 도 11 의 펄스 서브밴드 정보는 2 비트씩 총 10 비트가 할당되어
5 있음을 알 수 있다. 이는, 서브밴드가 총 4 개인 경우, 하나의 펄스가 서브밴드를 표시하기 위해서는 2 비트가 필요한데, 이는 각 세트의 메인 펄스 및 서브 펄스는 동일한 서브밴드에 속하기 때문에, 한 세트(메인 및 서브)의 펄스를 표현하기 위해 총 2 비트만 소요된다. 그런데 단독 펄스의 경우, 각 하나의 펄스를 표현하기 위해 2 비트씩이 소요된다.

10 따라서, 펄스 서브밴드 정보를 인코딩하기 위해서는 제 1 세트에 2 비트, 제 2 세트에 2 비트, 제 3 세트에 2 비트, 제 1 단독 펄스에 2 비트, 제 2 단독 펄스에 2 비트, 즉 총 10 비트가 필요하다.

또한, 펄스 위치 정보는, 특정 서브밴드 내에서는 펄스가 몇 번째 계수에 존재하는지를 나타내기 위한 정보이므로, 제 1 세트 내지 제 3 세트까지
15 각 6 비트, 제 1 단독 펄스에 6 비트, 제 2 단독 펄스에 6 비트, 총 30 비트가 소요된다.

또한, 펄스 부호 정보는, 총 8 개의 비트 각각 1 비트씩, 총 8 개의 비트가 소요되고, 펄스 크기 정보는, 4 개의 펄스들의 크기 정보를 8 비트의 코드북을 이용하여 벡터양자화하여, 총 16 비트를 할당한다.

20 다시 도 6 을 참조하면, 상기와 같은 과정을 통해 펄스 추출부(152)에 의해 추출된 펄스들이 고주파 대역의 신호(SWB 신호)로부터 제외됨으로써,

원본 노이즈 신호($\tilde{M}_{32}^0(k)$ 등)가 생성된다. 예를 들어, 총 280 개의 계수들
 중에서 총 8 개의 펄스에 해당하는 계수를 제외시킨 경우, 원본 노이즈 신호는
 총 272 개의 계수들에 대응할 수 있다. 도 9 는 펄스 추출 이전의 신호(SWB
 신호) 및 펄스 추출 이후 신호(원본 노이즈 신호)의 일 예가 도시되어 있다. 도
 5 9 의 (A)에서 원본 SWB 신호는 주파수 변환 계수 도메인 상에서 높은 피크
 에너지를 갖는 펄스가 복수 개 포함되어 있는 반면에, 도 9 의 (B)는 그 펄스가
 제외된 노이즈성(noise-like)의 신호만이 남는다.

도 6 의 참조 노이즈 생성부(154)는 저주파 대역의 주파수 변환 계수(WB
 신호)를 근거로, 참조 노이즈 신호를 생성한다. 구체적으로, WB 신호의 전체
 10 에너지를 근거로 임계치를 설정하고, 그 임계치 이상의 펄스를 제외시킴으로써,
 참조 노이즈 신호를 생성한다.

도 10 은 참조 노이즈 생성과정을 설명하기 위한 도면이다. 도 10 의
 (A)를 참조하면, WB 신호의 일 예가 주파수 변환 도메인 상에 표시되어 있다.
 이 중 전체 에너지를 고려하여 임계치가 설정되었을 때, 임계치를 벗어나는
 15 펄스도 존재하고, 임계치 범위 내에 존재하는 펄스도 있다. 임계치를 벗어나는
 펄스를 제외시키면 도 10 의 (B)와 같은 신호가 남게 된다. 이와 같이 참조
 노이즈 신호를 생성한 후, 정규화과정을 거친다. 그러면 도 10 의 (C)와 같이
 표현될 수 있다.

참조 노이즈 생성부(154)는 상기와 같은 과정을 통해 WB 신호를
 20 이용하여 참조 노이즈 신호 $\tilde{M}_{16}$ 를 생성한다.

도 6 의 노이즈 탐색부(156)는 원본 노이즈 신호 및 참조 노이즈 신호 $\tilde{M}_{16}$ 를 비교함으로써, 원본 노이즈 신호($\tilde{M}_{32}^0(k)$ 등)와 가장 유사한 참조 노이즈 신호의 구간을 결정하고, 노이즈 위치 정보 및 노이즈 에너지정보를 생성한다. 이 과정에 대한 일 실시예를 아래와 같이 구체적으로 설명하고자

5 한다.

우선, 원본 노이즈 신호(SWB 신호에서 펄스가 제외된 신호)를 아래와 같이 복수 개의 서브밴드 신호로 분할한다.

[수학식 5]

$$\begin{aligned}\tilde{M}_{32}^0(k) &= \tilde{M}_{32}(k+280), & k &= 0, \dots, 39 \\ \tilde{M}_{32}^1(k) &= \tilde{M}_{32}(k+320), & k &= 0, \dots, 69 \\ \tilde{M}_{32}^2(k) &= \tilde{M}_{32}(k+390), & k &= 0, \dots, 69 \\ \tilde{M}_{32}^3(k) &= \tilde{M}_{32}(k+460), & k &= 0, \dots, 99\end{aligned}$$

10 여기서의 각 서브밴드의 크기는 앞서 일반모드에서의 서브밴드와 동일한 크기를 가질 수 있다. 서브밴드의 길이 $d^j(k)$ $j=0, \dots, 3$ 는 40, 70, 70, 100 개의 주파수 변환 계수들에 대응할 수 있다. 모든 서브밴드는 각 다른 검색시작 위치 k^j 와 각기 다른 검색범위 w^j 를 가지며, 참조 노이즈 신호 $\tilde{M}_{16}$ 와의 유사정도를 검색한다. 검색시작 위치 k^j 는 $j=0, 2$ 일 때 0 으로

15 고정되고, $j=1, 3$ 인 경우 이전 서브밴드의 최대 유사도를 갖는 서브밴드의 시작위치에 종속된다. j 번째 서브밴드에 대한 검색시작 위치 k^j 및 검색범위 w^j 는 다음과 같이 표현될 수 있다.

[수학식 6]

$$k^j = \begin{cases} 0 & j = 0 \\ BestIdx^{j-1} + d^{j-1} - \frac{w^j}{2} & j = 1 \\ 0 & j = 2 \\ BestIdx^{j-1} + d^{j-1} - \frac{w^j}{2} & j = 3 \end{cases}$$

$$w^j = \begin{cases} 240 & j = 0 \\ 128 & j = 1 \\ 210 & j = 2 \\ 128 & j = 3 \end{cases}$$

k^j 는 검색시작 위치, $BestIdx^j$ 는 최대유사 시작위치,

d^j 는 서브밴드의 길이, w^j 는 검색 범위

- 5 만약, k^j 가 음수가 되면 0 으로 수정해주며, $280 - d^j - w^j$ 보다 커지면, $280 - d^j - w^j$ 으로 수정해준다. 최대유사 시작위치는 $BestIdx^j$ 은 다음 과정을 통해 각 서브밴드별로 추정한다.

우선 유사정도지표 k' 에 해당하는 유사도 $corr(k')$ 는 다음 수학적식과 같이 연산된다. 일반 모드와 유사한 방법으로 부호화되나, 한 샘플(하나의

- 10 계수)가 아니라 4 개의 샘플 단위로 검색을 실시한다.

[수학적식 7]

$$corr(k') = \sum_{k=0}^{k < d^j} M_{32}^j(k) \tilde{M}_{16}(k^j + k' + k), \quad k' = 0, 3, 7, \dots, w^j - 1$$

$corr(k')$ 는 유사도, $M_{32}^j(k)$ 는 원본 노이즈(수학적식 5 참조), $\tilde{M}_{16}$ 는 참조 노이즈, k^j 는 검색시작 위치, k' 유사정도지표, w^j 는 검색 범위

- 15 유사정도지표 k' 에 해당하는 에너지는 다음 수학적식과 같이 연산된다.

[수학식 8]

$$Ene(k') = \sum_{k=0}^{k \leq d^j} \tilde{M}_{16}(k^j + k' + k)^2, \quad k' = 0, 3, 7, \dots, w^j - 1$$

실질적인 유사정도 $S(k')$ 는 다음 수학식과 같다.

[수학식 9]

$$S(k') = \left| \frac{corr(k')}{\sqrt{Ene(k')}} \right|$$

실질적인 유사정도 $S(k')$ 가 최대값이 되는 서브밴드의 시작 위치 $BestIdx^j$ 는 다음과 같이 연산되며, $BestIdx^j$ 는 $LagIndex^j$ 파라미터로 변환되어 노이즈 위치 정보로서 비트스트림에 포함된다.

$BestIdx = 0$
 $lagCorr = 0$
 $lagEnergy = 1e30$

for $k' = 0$ to $w^j - 1$

if ($Ene(k') > 0$)

if ($lagCorr^2 Ene(k') < corr(k') lagEnergy$)

$BestIdx^j = k'$

$lagCorr = corr(k')$

$lagEnergy = Ene(k')$

end

end

end

10 앞서 노이즈 탐색부(156)가 노이즈 위치 정보를 생성한 과정에 대해서 설명한 바, 이하 노이즈 에너지 정보를 생성하는 과정에 대해서 설명하고자 한다. 참조 노이즈 신호는 원본 노이즈 신호와 유사한 파형을 유지하나, 그 에너지는 서로 다를 수 있다. 디코더에서 원본 노이즈와 유사한 에너지를 갖는

노이즈 신호를 가질 수 있도록, 원본 노이즈 신호의 에너지에 대한 정보인 노이즈 에너지 정보를 생성하여 디코더에 전송할 필요가 있다.

한편, 상기 노이즈 에너지의 값은 다이내믹 레인지가 크기 때문에, 펄스 비율이라는 값으로 변환하여 전송할 수 있다. 펄스 비율은 백분율 값이므로 0%~100%값을 가지므로 다이내믹 레인지가 작아서, 비트수를 절감할 수 있다. 이 변환 과정에 대해서 설명하고자 한다.

노이즈 신호의 에너지는 다음 수학적식과 같이 SWB 신호의 전체 에너지에서 펄스 에너지를 제외한 값과 같다.

[수학적식 10]

$$Noise_{energy} = \sum_{k=0}^{280} \{M_{32}(280+k)\}^2 - \hat{P}_{energy}$$

$Noise_{energy}$ 는 노이즈 에너지, M_{32} 는 SWB 신호, $\hat{P}_{energy}$ 는 펄스 에너지($\hat{P}_{energy} = \sum_{k=0}^7 \{P_{amp}(k)\}^2$)

상기 수학적식을 백분율인 펄스 비율 $\hat{R}_{percent}$ 로 나타내면 다음과 같다.

[수학적식 11]

$$\hat{R}_{percent} = \frac{\hat{P}_{energy}}{\hat{P}_{energy} + Noise_{energy}} \times 100$$

$\hat{R}_{percent}$ 는 펄스 비율, $\hat{P}_{energy}$ 는 펄스 에너지, $Noise_{energy}$ 는 노이즈 에너지

즉, 인코더에서는 수학적식 10에 나타난 노이즈 에너지 $Noise_{energy}$ 대신에, 수학적식 11에 나타난 펄스 비율 $\hat{R}_{percent}$ 을 전송한다. 이 펄스 비율에

해당하는 노이즈 에너지 정보는 도 11 에 도시된 바와 같이 4 비트로 인코딩될 수 있다.

그러면, 디코더에서는 우선, 펄스 추출부(152)에 의해 생성된 펄스 정보를 근거로 펄스 에너지 ($\hat{P}_{energy} = \sum_{k=0}^7 \{P_{amp}(k)\}^2$)를 생성한다. 그런 다음
 5 펄스 에너지($\hat{P}_{energy}$) 및 전송된 펄스 비율 $\hat{R}_{percent}$ 을 다음 수학식에 대입함으로써, $Noise_{energy}$ 는 노이즈 에너지를 생성한다.

[수학식 12]

$$Noise_{energy} = \frac{(100 - \hat{P}_{energy}) \times \hat{R}_{percent}}{\hat{R}_{percent}}$$

상기 수학식 12는 수학식 11을 다시 정리한 식이다.

10 디코더는 전송된 펄스 비율을 상기와 같이 노이즈 에너지로 변환한 후, 노이즈 에너지는 참조 노이즈 신호의 각 계수에 다음과 같이 곱함으로써, 참조 노이즈 신호를 이용하여, 원본 노이즈 신호와 유사한 에너지 분포를 갖는 노이즈 신호를 획득할 수 있다.

[수학식 13]

15
$$\hat{S}_{amp} = \sqrt{Noise_{energy} \times \frac{1}{272}}$$

$$\tilde{M}_{32}(k+280) = \tilde{M}_{32}(k+280) \times \hat{S}_{amp} \quad k = 0, \dots, 280$$

노이즈 탐색부(156)는 상기와 같은 과정을 통해 노이즈 위치 정보를 생성하고, 노이즈 에너지 값을 펄스 비율로 변환하고, 이를 노이즈 에너지 정보로서 디코더에 전송한다.

도 12 는 특정 오디오 신호에 대해 일반 모드 및 비-일반모드로 인코딩된 결과를 보여주기 위한 도면이다. 도 12 의 (A)를 우선 참조하면, 특정 신호(예: 타악기와 같이 특정 주파수 대역이 강한 신호)에 대해서는 일반
 5 모드로 인코딩한 후 합성한 결과와, 비-일반 모드로 인코딩한 후 다시
 10 복호화한 결과가, 도 12 의 (A)에 나타난 바와 같이, 차이가 있을 수 있다. 도
 12 의 (B)를 살펴보면, 도 12 에 나타난 원본 신호에 대해서는, 비-일반 모드로
 인코딩한 결과가, 일반 모드로 인코딩한 결과보다 우수함을 알 수 있다.

즉, 오디오 신호의 특성에 따라, 미리 결정된 특정 펄스의 에너지가
 높은 경우에는, 본 발명의 실시예에 따른 비-일반 모드로서 인코딩함으로써,
 10 비트수 증가는 거의 없이 보다 음질을 높일 수 있다.

이하에서는, 프레임간의 유사성이 높아서 토널-모드로 분류된 경우, 도
 1 에서 언급한 하모닉비율 판단부(160), 년-하모닉모드 인코딩부(170) 및
 하모닉모드 인코딩부(180)에 대해서 각각 설명하고자 한다.

우선 도 13 은 하모닉비율 판단부(160)의 세부 구성을 보여주는
 15 도면이다. 도 13 을 참조하면, 하모닉비율 판단부(160)는, 하모닉트랙
 추출부(162) 및 고정펄스 추출부(164) 및 하모닉비율 결정부(166)를 포함할 수
 있는 데, 오디오 신호의 하모닉비율을 근거로 년-하모닉 모드 및 하모닉
 모드를 결정한다. 하모닉모드는 예를 들어, 단일 악기에 의한 하모닉 성분이
 강한 신호, 또는 여러 악기에 의해서 발생하는 멀티플 피치(multiple pitch)
 20 신호를 포함하는 신호를 인코딩하기에 적절한 모드이다.

도 14 는 하모닉비율이 높은 오디오 신호의 예이다. 도 14 를 참조하면, 주파수 변환 계수 도메인 상에서 기본 주파수의 배수가 되는 하모닉들이 강하게 나타나고 있음을 알 수 있다. 이러한 하모닉 특성이 강한 신호의 경우, 종래의 방식대로 인코딩을 할 경우, 하모닉에 해당하는 펄스를 모두 인코딩해야 하므로, 비트가 많이 소요되어 부호화기의 성능이 떨어지는 문제점이 있다. 반대로, 정해진 개수의 펄스만을 추출하는 인코딩 방식을 적용하였을 때, 모든 펄스를 추출하기 어렵기 때문에 음질이 저하되는 문제점이 있다. 따라서, 이러한 신호에 적절한 코딩 방식을 본원에서는 제안하고자 한다.

10 하모닉트랙 추출부(162)는 고주파 대역에 해당하는 주파수 변환 계수들에 대해서, 하모닉 트랙을 추출한다. 이 과정은 추후 하모닉모드 인코딩부(180)의 하모닉트랙 추출부(182)와 동일한 과정을 수행하므로, 추후에 구체적으로 설명하고자 한다.

고정펄스 추출부(164)는 정해진 영역에서 정해진 개수의 펄스를 15 추출한다(164). 이 과정은 추후 년-하모닉모드 인코딩부(170)의 고정펄스 추출부(172)와 동일한 과정을 수행하므로, 추후에 구체적으로 설명하고자 한다.

하모닉비율 결정부(166)는, 하모닉 비율 즉, 추출된 하모닉트랙의 에너지의 합에 비한 고정펄스 에너지의 비율인 바, 하모닉 비율이 낮은 경우에는 년-하모닉 모드로 결정하고, 하모닉 비율이 높은 경우에는 하모닉 20 모드로 결정한다. 앞서 설명한 바와 같이 년-하모닉 모드인 경우 년-

하모닉모드 인코딩부(170)가 활성화되고, 하모닉 모드인 경우 하모닉모드 인코딩부(180)가 활성화된다.

도 15 는 년-하모닉모드 인코딩부(170)의 세부 구성을 보여주는 도면이고, 도 16 은 년-하모닉모드인 경우, 고정 펄스를 추출하는 규칙을 설명하기 위한 도면이고, 도 17 은 년-하모닉모드로 인코딩한 경우의 신택스의 일 예이다.

우선, 도 15 를 참조하면, 년-하모닉모드 인코딩부(170)는 고정펄스 추출부(172) 및 펄스위치정보 생성부(174)를 포함한다.

고정펄스 추출부(172)는 도 16 에 나타난 바와 같이 고정된 영역에서 고정된 개수의 고정 펄스를 추출한다.

10 [수학식 14]

$$D(k) = |\ddot{M}_{32}(k) - M_{32}(k)|, \quad k = 280, \dots, 560$$

여기서, $M_{32}(k)$ 는 SWB 신호, $\ddot{M}_{32}(k)$ 는 HF 합성신호.

HF 합성 신호 $\ddot{M}_{32}(k)$ 는 존재하지 않기 때문에 0 으로 결정되고, $M_{32}(k)$ 의 최대값을 찾는 과정이 된다. $D(k)$ 에서 5 개의 서브밴드로
 15 분할하여 D_j 를 만들고 각 서브밴드의 펄스의 개수는 이전에 결정된 개수 N_j 값을 갖는다. 각 서브밴드별 N_j 개의 가장 큰 값을 찾는 과정은 다음과 같다. 다음의 알고리즘은 정렬 알고리즘으로서 최대값 N 을 찾고 이를 input_data 라는 배열에 저장한다.

```

for j = 0 to N
    data_sorted(j) = 0
    data_sorted(j) = 0
    Idx = 0
    for k = 1 to length(input_data)
        if(input_data(j) > data_sorted(j))
            index_sorted(j) = k
            Idx = k
        end
    end
end
end

```

도 16 을 참조하면, 각 서브밴드 별로 다수의 위치 셋 중 하나, 즉, 제 1 위치 셋(예: 짝수 위치들) 또는 제 2 위치 셋(홀수 위치들)에서 정해진 개수(예: 10 개)의 펄스를 추출하는 예가 도시되어 있다. 제 1 서브밴드에서는 짝수의 위치(position: 280 등)에서 두 개의 펄스들(track 0)을, 홀수의 위치(position: 281, 등)에서도 두 개의 펄스들(track 1)을 추출한다. 제 2 서브밴드에서도 마찬가지로, 짝수의 위치(position: 280 등)에서 두 개의 펄스들(track 2)을, 홀수의 위치(position: 281, 등)에서도 두 개의 펄스들(track 3)을 추출한다. 그런 다음 제 3 서브밴드에서는, 위치와 상관없이 1 개의 펄스(track4)를, 제 4 서브밴드에서도, 위치와 상관없이 1 개의 펄스(track5)를 추출한다.

위와 같이, 고정 펄스를 추출하는 이유는, 즉, 정해진 위치에서 정해진 개수만큼의 펄스를 추출하는 이유는, 고정 펄스의 위치정보에 해당하는 비트수를 절약하기 위해서이다.

다시 도 15 를 참조하면, 펄스위치정보 생성부(174)는 추출된 고정펄스에 대해서, 정해진 규칙에 따라 고정펄스 위치정보를 생성한다. 도 17 은 년-

하모닉모드로 인코딩한 경우의 선택스의 일 예이다. 도 17 를 참조하면, 도 16 에 나타난 규칙대로 고정펄스를 추출할 경우, track0 내지 track3 까지의 총 8 개의 펄스는 짝수 또는 홀수로 위치가 정해져 있기 때문에, 고정펄스 위치정보를 인코딩하기 위한 비트수가 64 비트가 아니라 32 비트가 될 수 있다.

- 5 track4 에 해당하는 펄스는 짝수 또는 홀수의 제한이 없으므로 64 비트가 소요되고, track5 의 경우에는 짝수 또는 홀수의 제한이 없지만, 위치가 472 부터 503 으로 제한되므로, 32 비트가 필요하다.

이하, 도 18 내지 도 20 과 함께 하모닉모드 인코딩 과정을 설명하고자 한다.

- 10 도 18 은 하모닉모드 인코딩부(180)의 세부 구성을 보여주는 도면이고, 도 19 는 하모닉트랙의 추출을 설명하기 위한 도면이고, 도 20 은 하모닉트랙 위치정보의 양자화를 설명하기 위한 도면이다.

도 18 을 참조하면, 하모닉모드 인코딩부(180)은 하모닉트랙 추출부(182) 및 하모닉정보 부호화부(184)를 포함한다.

- 15 하모닉트랙 추출부(182)는 고주파 대역에 해당하는 주파수 변환 계수들에 대해서 복수의 하모닉 트랙을 추출한다. 구체적으로, 제 1 피치에 해당하는 제 1 그룹의 하모닉 트랙(제 1 하모닉 트랙 및 제 2 하모닉 트랙)을 추출하고, 제 2 피치에 해당하는 제 2 그룹의 하모닉 트랙(제 3 하모닉 트랙 및 제 4 하모닉 트랙)을 추출한다. 제 1 하모닉 트랙 및 제 3 하모닉 트랙의 시작
20 위치 정보는 제 1 위치 셋(예: 홀수) 중 하나에 대응될 수 있고, 제 2 하모닉

트랙 및 제 4 하모닉 트랙의 시작 위치 정보는 제 2 위치 셋(예: 짝수) 중 하나에 대응될 수 있다.

도 19 의 (A)를 참조하면, 제 1 피치를 갖는 제 1 하모닉 트랙 및 제 1 피치를 갖는 제 2 하모닉 트랙이 표현되어 있다. 예를 들어, 제 1 하모닉 5 트랙의 시작 위치는 짝수, 제 2 하모닉 트랙의 시작 위치는 홀수로 표현될 수 있다. 한편, 도 19 의 (B)를 참조하면, 제 2 피치를 갖는 제 3 하모닉 트랙 및 제 4 하모닉 트랙이 표시되어 있다. 제 3 하모닉 트랙의 시작 위치는 홀수로 설정되고, 제 4 하모닉 트랙의 시작 위치는 짝수로 설정될 수 있다. 한편, 각 그룹의 하모닉 트랙들이 3 개 이상일 경우(즉, 제 1 그룹에 A 하모닉 트랙, B 10 하모닉 트랙, C 하모닉 트랙이 포함, 제 2 그룹에 K 하모닉 트랙, L 하모닉 트랙, M 하모닉 트랙이 포함), A / K 하모닉 트랙에 대응하는 제 1 위치 셋은 $3N$ (N 은 정수), B / L 하모닉 트랙에 대응하는 제 2 위치 셋은 $3N+1$ (N 은 정수), C/M 하모닉 트랙에 대응하는 제 3 위치 셋은 $3N+2$ (N 은 정수)일 수 있다.

상기와 같은 복수의 하모닉 트랙은 다음 수학식을 통해 수행될 수 있다.

15 [수학식 14]

$$D(k) = |\ddot{M}_{32}(k) - M_{32}(k)|, \quad k = 280, \dots, 560$$

여기서, $M_{32}(k)$ 는 SWB 신호, $\ddot{M}_{32}(k)$ 는 HF 합성신호.

HF 합성신호는 존재하지 않기 때문에 0 으로 초기값을 설정하면,

$M_{32}(k)$ 의 최대값을 찾는 과정이 된다.

20 $D(k)$ 는 미리 결정된 개수(예: 총 4 개)의 하모닉 트랙의 합으로 표현된다. 각각의 하모닉 트랙 D_j 는 최대한 2 개 이상의 피치 성분을 포함할

수 있으며, 하나의 피치 성분에 2 개의 하모닉 트랙 D_j 을 추출할 수 있다. 각 피치 성분별 2 개의 가장 큰 값을 갖는 하모닉 트랙 D_j 를 찾는 과정은 다음과 같다.

다음 수식은 오토코릴레이션(autocorrelation) 함수를 이용하여 가장 많은

- 5 에너지를 포함하는 하모닉 트랙 D_j 의 피치 P_i 를 찾는다. 피치의 범위를 주파수 변환 계수들의 20-27 의 계수로 제한하여 추출되는 하모닉의 개수를 제한할 수 있다.

[수학식 15]

$$P_i(m) = \sum_{n=280}^{560-m} (|M_{32}(n)| \times |M_{32}(n+m)|), \quad m = 20, \dots, 27, \quad i = 1, 2$$

- 10 다음 수식은 개별 피치 P_i 마다 가장 많은 에너지를 포함하는 총 2 개의 하모닉 트랙 D_j 의 시작위치 PS_i 를 계산하여 하모닉 트랙 D_j 를 추출하는 과정이다. 하모닉 트랙 D_j 의 시작위치 PS_i 의 범위는 추출되는 하모닉의 개수를 포함하여 계산되며, MDCT 도메인 신호의 특성을 반영하여 피치 P_i 에 의해 2 개의 시작위치 PS_i 에 의해서 총 2 개의 하모닉 트랙 D_j 이
- 15 추출된다.

[수학식 16]

$$PS_i(2m-1) = \sum_{n=1}^{\lfloor 280/P_i \rfloor} |M_{32}[(2m-1) + P_i \times n]|, \quad m = 1, \dots, 16$$

$$PS_i(2m) = \sum_{n=1}^{\lfloor 280/P_i \rfloor} |M_{32}[(2m) + P_i \times n]|, \quad m = 1, \dots, 16$$

추출된 4 개의 하모닉 트랙 D_j 의 피치 P_i 와 시작위치 PS_i 의 범위와 개수는 도 19의 (C)에 도시되어 있다.

하모닉정보 부호화부(184)는 하모닉트랙에 관한 상기 정보들을 인코딩하고 벡터양자화한다.

- 5 위의 과정에서 추출된 각 하모닉 트랙은 피치 P_i 와 시작위치 PS_i 들의 위치정보를 갖고 있다. 추출된 피치 P_i 와 시작위치 PS_i 는 다음과 같이 부호화된다. 피치 P_i 는 HF에서 존재할 수 있는 하모닉의 개수를 제한하여 3 비트를 사용하여 양자화되며, 시작위치 PS_i 는 4 비트를 사용하여 각각 양자화된다. 총 4 개의 하모닉 트랙을 추출하기 위한 위치정보는 2 개의 피치
- 10 P_i 개의 시작위치 PS_i 를 사용하여 총 22 비트가 사용될 수 있으나 본 발명에는 이에 한정되지 아니한다.

- 위 과정에서 추출되는 4 개의 하모닉 트랙은 최대 44 개의 펄스를 포함한다. 44 개의 펄스의 크기값과 부호정보를 모두 양자화하기 위해서는 많은 비트가 필요하다. 따라서 각각의 하모닉 트랙의 펄스 중 펄스 피크 추출
- 15 알고리즘을 이용하여 에너지를 많이 포함하는 펄스를 추출하여 크기값과 부호화 정보를 다음 수식과 같이 별도 부호화한다.

- 다음의 알고리즘은 각각 하모닉 트랙에서 펄스 피크 PP_i 추출 알고리즘으로서 에너지를 많이 포함하는 연속된 펄스를 찾고 크기값을 양자화하고, 부호정보를 다음 수식과 같이 별도로 부호화한다. 각각의 하모닉
- 20 트랙에서 펄스 피크를 추출하기 위해서 3 비트를 사용하며, 2 개의 하모닉

트랙에서 추출된 4 개의 펄스는 8 비트를 사용하여 그 크기값이 양자화되며, 각각의 부호정보를 위해 1 비트를 할당한다. 이를 통해 펄스 피크 추출알고리즘을 통해 추출된 펄스는 총 24 비트로 양자화된다.

[수학식 17]

$$PP_i(n) = (|M_{32}(n)|^2 + |M_{32}(n+1)|^2), \quad n = 1, \dots, 5$$

$$PP_i(n-1) = (|M_{32}(n)|^2 + |M_{32}(n+1)|^2), \quad n = 7$$

$$PP_i(n-2) = (|M_{32}(n)|^2 + |M_{32}(n+1)|^2), \quad n = 9$$

$$5 \quad PP_i(n-3) = (|M_{32}(n)|^2 + |M_{32}(n+1)|^2), \quad n = 11$$

$$Sign_harpulse_j(n) = \begin{cases} 1 & M_{32}(PP_i(n)) \geq 0 \\ -1 & otherwise \end{cases}$$

$$Sign_harpulse_j(n) = \begin{cases} 1 & M_{32}(PP_i(n+1)) \geq 0 \\ -1 & otherwise \end{cases}$$

위 과정에서 추출된 8 개의 펄스를 제외한 하모닉 트랙은 하나의 트랙으로 결합되어 DCT(Discrete Cosine Transform)을 이용하여 크기값과 부호정보를 동시에 양자화한다. DCT 양자화를 위해서는 19 비트가 사용된다.

추출된 4 개의 하모닉 트랙의 펄스 피크 추출 알고리즘을 통해 추출된 펄스와 이를 제외한 하모닉 트랙의 부호화 과정은 도 20 에 나타난 바와 같다. 도 20 을 참조하면, 제 1 그룹의 제 1 하모닉 트랙 중에 최대 펄스와 그의 인접펄스, 제 1 그룹의 제 2 하모닉 트랙 중에서 최대 펄스와 그의 인접 펄스에 대해 제 1 타깃 벡터(targetA)를 생성하고, 제 3 하모닉 트랙 중 최대 펄스와 그의 인접펄스, 제 4 하모닉 트랙 중 최대 펄스와 그의 인접 펄스에 대해서 제 2 타깃 벡터(targetB)를 생성한다. 제 1 타깃 벡터 및 제 2 타깃 벡터에 대해서 벡터-양자화를 수행하고, 각 하모닉 트랙들 중에서 상기 최대 펄스 및

인접 펄스를 제외한 나머지 부분들(residual)은 조합하여 이에 대해 주파수 변환을 수행한다. 이때 주파수 변환에는 앞서 언급한 바와 같이 DCT(Discrete Cosine Transform)가 사용될 수 있다.

그리고 상기에서 설명한 하모닉 트랙에 관한 정보의 일 예는 도 21에
5 나타난 바와 같다.

도 22은 특정 오디오 신호에 대해 년-하모닉 모드 및 하모닉모드로 인코딩된 결과를 보여주기 위한 도면이다. 도 22를 참조하면, 하모닉이 강한 신호에 대해서는 년-하모닉 모드로 인코딩한 결과보다, 하모닉 모드로 인코딩한 결과가 원본 신호에 더욱 가깝기 때문에, 음질이 향상될 수 있음을
10 알 수 있다.

도 23은 본 발명의 실시예에 따른 오디오 신호 처리 장치 중 디코더의 구성을 보여주는 도면이다. 도 23을 참조하면, 본 발명의 실시예에 따른 디코더(200)는 모드결정부(210), 비-일반모드 디코딩부(230), 하모닉모드 디코딩부(250) 중 하나 이상을 포함하고, 일반모드 디코딩부(220), 년-
15 하모닉모드 디코딩부(240)를 더 포함할 수 있다. 한편, 수신된 오디오 신호에 대한 비트스트림을 파싱하는 디-멀티플렉서(미도시)를 더 포함할 수 있다.

모드결정부(210)는, 비트스트림을 통해 수신된 제 1 모드 정보 및 제 2 모드 정보를 근거로 현재 프레임에 대응하는 모드 즉, 현재 모드를 결정한다. 제 1 모드 정보는, 앞서 인코더(100)와 관련하여 언급한 바와 마찬가지로, 제 1
20 모드 정보는, 상기 년-토널 모드 및 상기 토널 모드 중 하나를 지시하고, 제 2 모드 정보는, 상기 제 1 모드 정보가 년-토널 모드를 지시하는 경우, 일반모드

또는 비-일반 모드 중 하나를 지시하고, 제 1 모드 정보가 토널 모드를 지시하는 경우, 년-하모닉 모드 및 하모닉 모드 중 하나를 지시한다.

이 결정된 현재 모드에 따라서, 현재 프레임에 대해 네 가지 디코딩부들(220, 230, 240, 250) 중 하나가 활성화되고, 현재 모드에 따라서 각

5 모드에 대응하는 파라미터가 디멀티플렉서(미도시)에 의해 추출한다.

현재 모드가 일반 모드인 경우에는 인벨롭 위치정보 및 스케일링 정보 등이 추출된다. 그러면 일반모드 디코딩부(220)는 복원된 저주파 대역의 주파수 변환 계수들(WB 신호) 중에서 상기 인벨롭 위치정보에 대응하는 구간, 즉 최대유사 밴드의 인벨롭을 추출한다. 그런 다음, 상기 스케일링 정보를

10 이용하여 인벨롭을 스케일링 함으로써, 현재 프레임의 고주파 대역(SWB 신호)를 복원한다.

현재 모드가 비-일반 모드인 경우에는 펄스 정보, 노이즈 위치정보, 노이즈 에너지 정보 등이 추출된다. 그러면 비-일반모드 디코딩부(230)는, 펄스 정보를 근거로 복수 개의 펄스(예를 들어, 총 3 세트의 메인펄스 및 서브펄스,

15 2 개의 단독 펄스)를 생성한다. 여기서 펄스 정보는 펄스 위치 정보, 펄스 부호 정보, 및 펄스 크기 정보를 포함할 수 있는데, 상기 펄스 부호 정보에 따라 각 펄스들의 부호가 결정된다, 펄스 크기 정보 및 펄스 위치 정보에 따라서 각 펄스들의 크기 및 위치가 결정된다. 그런 다음, 노이즈 위치정보를 이용하여 복원된 WB 신호 중에서 노이즈로 사용할 구간을 결정하고, 상기 노이즈

20 에너지 정보를 이용하여 노이즈 에너지를 조절한 후, 상기 펄스들을 합함으로써, 현재 프레임의 SWB 신호를 복원한다.

현재 모드가 년-하모닉 모드인 경우, 고정펄스들에 대한 정보인 고정펄스정보가 추출된다. 년-하모닉모드 디코딩부(240)는 고정펄스정보를 이용하여 서브밴드별로 정해진 위치 및 정해진 개수만큼의 고정펄스를 획득한다. 고정펄스를 이용하여 현재 프레임의 SWB 신호를 생성한다.

- 5 현재 모드가 하모닉 모드인 경우, 하모닉트랙의 위치정보 등이 추출된다.
- 여기서 하모닉트랙의 위치정보는 제 1 피치를 갖는 제 1 그룹의 하모닉트랙들의 시작위치정보 및, 제 2 피치를 갖는 제 2 그룹의 하모닉트랙들의 시작위치정보를 포함한다. 제 1 그룹의 하모닉 트랙들은 제 1 하모닉 트랙 및 제 2 하모닉 트랙을 포함하고, 상기 제 2 그룹의 하모닉
- 10 트랙들은 제 3 하모닉 트랙 및 제 4 하모닉 트랙을 포함할 수 있는데, 제 1 하모닉 트랙 및 제 3 하모닉 트랙의 시작 위치 정보는 제 1 위치 셋 중 하나에 대응되고, 상기 제 2 하모닉 트랙 및 제 4 하모닉 트랙의 시작 위치 정보는 제 2 위치 셋 중 하나에 대응될 수 있다.

- 여기서, 상기 제 1 피치 및 제 2 피치를 지시하는 피치 정보를 더 수신할
- 15 수 있다. 하모닉모드 디코딩부(250)은 상기 피치 정보 및 상기 시작위치정보를 이용하여 상기 시작 위치 정보에 대응하는 복수의 하모닉 트랙을 생성하고, 복수의 하모닉 트랙을 이용하여 현재 프레임에 해당하는 오디오 신호 즉 SWB 신호를 생성한다.

- 본 발명에 따른 오디오 신호 처리 장치는 다양한 제품에 포함되어
- 20 이용될 수 있다. 이러한 제품은 크게 스탠드 얼론(stand alone) 군과 포터블(portable) 군으로 나뉘 수 있는데, 스탠드 얼론군은 티비, 모니터, 셋탑

박스 등을 포함할 수 있고, 포터블군은 PMP, 휴대폰, 네비게이션 등을 포함할 수 있다.

도 24 는 본 발명의 실시예에 따른 오디오 신호 처리 장치가 구현된 제품의 개략적인 구성을 보여주는 도면이다. 우선 도 24 를 참조하면, 유무선 5 통신부(510)는 유무선 통신 방식을 통해서 비트스트림을 수신한다. 구체적으로 유무선 통신부(510)는 유선통신부(510A), 적외선통신부(510B), 블루투스부(510C), 무선랜통신부(510D) 중 하나 이상을 포함할 수 있다.

사용자 인증부(520)는 사용자 정보를 입력 받아서 사용자 인증을 수행하는 것으로서 지문인식부(520A), 홍채인식부(520B), 얼굴인식부(520C), 및 10 음성인식부(520D) 중 하나 이상을 포함할 수 있는데, 각각 지문, 홍채정보, 얼굴 윤곽 정보, 음성 정보를 입력받아서, 사용자 정보로 변환하고, 사용자 정보 및 기존 등록되어 있는 사용자 데이터와의 일치여부를 판단하여 사용자 인증을 수행할 수 있다.

입력부(530)는 사용자가 여러 종류의 명령을 입력하기 위한 15 입력장치로서, 키패드부(530A), 터치패드부(530B), 리모컨부(530C) 중 하나 이상을 포함할 수 있지만, 본 발명은 이에 한정되지 아니한다.

신호 코딩 유닛(540)는 유무선 통신부(510)를 통해 수신된 오디오 신호 및/또는 비디오 신호에 대해서 인코딩 또는 디코딩을 수행하고, 시간 도메인의 오디오 신호를 출력한다. 오디오 신호 처리 장치(545)를 포함하는데, 이는 앞서 20 설명한 본 발명의 실시예(즉, 제 1 실시예에 따른 인코더(100) 및/또는 디코더(200), 제 2 실시예에 따른 인코더(300) 및/또는 디코더(400))에 해당하는

것으로서, 이와 같이 오디오 처리 장치(545) 및 이를 포함한 신호 코딩 유닛은 하나 이상의 프로세서에 의해 구현될 수 있다.

제어부(550)는 입력장치들로부터 입력 신호를 수신하고, 신호 디코딩부(540)와 출력부(560)의 모든 프로세스를 제어한다. 출력부(560)는 신호
5 디코딩부(540)에 의해 생성된 출력 신호 등이 출력되는 구성요소로서, 스피커부(560A) 및 디스플레이부(560B)를 포함할 수 있다. 출력 신호가 오디오 신호일 때 출력 신호는 스피커로 출력되고, 비디오 신호일 때 출력 신호는 디스플레이를 통해 출력된다.

도 25 는 본 발명의 일 실시예에 따른 오디오 신호 처리 장치가 구현된
10 제품들의 관계도이다. 도 25 는 도 24 에서 도시된 제품에 해당하는 단말 및 서버와의 관계를 도시한 것으로서, 도 25 의 (A)를 참조하면, 제 1 단말(500.1) 및 제 2 단말(500.2)이 각 단말들은 유무선 통신부를 통해서 데이터 내지 비트스트림을 양방향으로 통신할 수 있음을 알 수 있다. 도 16 의 (B)를 참조하면, 서버(600) 및 제 1 단말(500.1) 또한 서로 유무선 통신을 수행할 수
15 있음을 알 수 있다.

본 발명에 따른 오디오 신호 처리 방법은 컴퓨터에서 실행되기 위한 프로그램으로 제작되어 컴퓨터가 읽을 수 있는 기록 매체에 저장될 수 있으며, 본 발명에 따른 데이터 구조를 가지는 멀티미디어 데이터도 컴퓨터가 읽을 수 있는 기록 매체에 저장될 수 있다. 상기 컴퓨터가 읽을 수 있는 기록 매체는
20 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 저장 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-

ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한, 상기 인코딩 방법에 의해 생성된 비트스트림은 컴퓨터가 읽을 수 있는 기록 매체에 저장되거나, 유/무선 통신망을 이용해 전송될 수 있다.

- 5 이상과 같이, 본 발명은 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명은 이것에 의해 한정되지 않으며 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 본 발명의 기술사상과 아래에 기재될 특허청구범위의 균등범위 내에서 다양한 수정 및 변형이 가능함은 물론이다.

10 【Industrial Applicability】

본 발명은 오디오 신호를 인코딩하고 디코딩하는 데 적용될 수 있다.

【청구의 범위】

[청구항 1]

오디오 신호에 대해서 주파수 변환을 수행함으로써, 복수 개의 주파수 변환 계수들을 획득하는 단계;

- 5 상기 복수 개의 주파수 변환 계수들 중 고주파 대역의 주파수 변환 계수들에 대해서, 펄스 비율을 근거로 일반 모드 및 비-일반모드 중 하나를 선택하는 단계; 및,

상기 비-일반 모드가 선택된 경우, 다음 단계가 수행되는 단계;

- 10 상기 고주파 대역의 주파수 변환 계수들로부터 미리 결정된 개수의 펄스들을 추출하고, 펄스 정보를 생성하는 단계;

상기 고주파 대역의 주파수 변환 계수들로부터 상기 펄스들을 제외한 원본 노이즈 신호를 생성하는 단계;

상기 복수 개의 주파수 변환 계수들 중 저주파 대역의 주파수 변환 계수를 이용하여 참조 노이즈 신호를 생성하는 단계; 및,

- 15 상기 원본 노이즈 신호 및 상기 참조 노이즈 신호를 이용하여, 노이즈 위치정보 및 노이즈 에너지정보를 생성하는 단계를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 2]

- 20 제 1 항에 있어서,

상기 펄스 비율은, 현재 프레임의 전체 에너지에 대한 복수의 펄스들에 대한 에너지의 비율인 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 3]

- 25 제 1 항에 있어서,

미리 결정된 개수의 펄스들을 추출하는 단계는

가장 큰 에너지를 갖는 메인 펄스를 추출하는 단계;

상기 메인 펄스에 인접한 서브 펄스를 추출하는 단계; 및

- 30 상기 고주파 대역의 주파수 변환 계수에서 상기 메인 펄스 및 상기 서브 펄스를 제외시킴으로써 타깃 노이즈 신호를 생성하는 단계;

상기 타깃 노이즈 신호에 대해서 상기 메인 펄스 및 상기 서브 펄스를 추출하는 것을 미리 정해진 횟수만큼 반복하는 단계를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

- 35 [청구항 4]

제 1 항에 있어서,

상기 펄스 정보는, 펄스 위치 정보, 펄스 부호 정보, 펄스 크기 정보, 펄스 서브밴드 정보 중 하나 이상을 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

5 [청구항 5]

제 1 항에 있어서,
상기 참조 노이즈 신호를 생성하는 단계는,
저주파 대역의 전체 에너지를 근거로 임계치를 설정하는 단계; 및
상기 임계치 이상의 펄스를 제외시킴으로써, 상기 참조 노이즈 신호를
10 생성하는 단계를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 6]

제 1 항에 있어서,
상기 노이즈 에너지정보를 생성하는 단계는,
15 상기 미리 결정된 개수의 펄스의 에너지를 생성하는 단계;
상기 원본 노이즈 신호의 에너지를 생성하는 단계;
상기 펄스의 에너지 및 상기 원본 노이즈 신호의 에너지를 이용하여
펄스 비율을 획득하는 단계; 및,
상기 펄스 비율을 노이즈 에너지 정보로서 생성하는 단계를 포함하는
20 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 7]

오디오 신호에 대해서 주파수 변환을 수행함으로써, 복수 개의 주파수
변환 계수들을 획득하는 주파수 변환부;
25 상기 복수 개의 주파수 변환 계수들 중 고주파 대역의 주파수 변환
계수들에 대해서, 펄스 비율을 근거로 일반 모드 및 비-일반모드 중 하나를
선택하는 펄스비율 판단부;
상기 비-일반 모드가 선택된 경우에 동작하며, 다음을 포함하는 비-
일반모드 인코딩부;
30 상기 고주파 대역의 주파수 변환 계수들로부터 미리 결정된
개수의 펄스들을 추출하고, 펄스 정보를 생성하는 펄스 추출부;
상기 복수 개의 주파수 변환 계수들 중 저주파 대역의 주파수
변환 계수를 이용하여 참조 노이즈 신호를 생성하는 참조노이즈
생성부; 및,
35 원본 노이즈 신호 및 상기 참조 노이즈 신호를 이용하여,
노이즈 위치정보 및 노이즈 에너지정보를 생성하는 노이즈 탐색부를
포함하고,

상기 원본 노이즈 신호는, 상기 고주파 대역의 주파수 변환 계수들로부터 상기 펄스들을 제외함으로써 생성되는 것을 특징으로 하는 오디오 신호 처리 장치.

5 [청구항 8]

현재 프레임이 일반 모드 인지 아니면 비-일반 모드인지 여부를 지시하는 제 2 모드 정보를 수신하는 단계;

상기 제 2 모드 정보가 현재 프레임이 비-일반 모드임을 지시하는 경우, 펄스 정보, 노이즈 위치정보 및 노이즈 에너지정보를 수신하는 단계;

10 상기 펄스 정보를 이용하여, 주파수 변환 계수들에 대해서 미리 결정된 개수의 펄스들을 생성하는 단계;

상기 노이즈 위치정보에 대응하는 저주파 대역의 주파수 변환 계수를 이용하여 참조 노이즈 신호를 생성하는 단계;

15 상기 노이즈 에너지정보를 이용하여 상기 참조 노이즈 신호의 에너지를 조절하는 단계; 및,

상기 에너지 조절된 참조 노이즈 신호 및 상기 복수의 펄스를 이용하여 고주파 대역에 대응하는 주파수 변환 계수들을 생성하는 단계를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

20 [청구항 9]

오디오 신호를 수신하는 단계;

오디오 신호에 대해서 주파수 변환을 수행하여, 복수개의 주파수 변환 계수들을 획득하는 단계;

25 상기 주파수 변환 계수에 대해서, 하모닉 비율을 근거로 난-하모닉 모드 및 하모닉 모드 중 하나를 선택하는 단계;

상기 하모닉 모드가 선택된 경우, 다음 단계가 수행되는 단계;

제 1 피치에 해당하는 제 1 그룹의 하모닉 트랙들을 결정하는 단계;

제 2 피치에 해당하는 제 2 그룹의 하모닉 트랙들을 결정하는 단계; 및,

30 상기 복수의 하모닉 트랙들의 시작 위치 정보를 생성하는 단계를 포함하고,

상기 제 1 그룹의 하모닉 트랙들은 제 1 하모닉 트랙 및 제 2 하모닉 트랙을 포함하고,

상기 제 2 그룹의 하모닉 트랙들은 제 3 하모닉 트랙 및 제 4 하모닉 트랙을 포함하고,

35 상기 제 1 하모닉 트랙 및 제 3 하모닉 트랙의 시작 위치 정보는 제 1 위치 셋 중 하나에 대응되고,

상기 제 2 하모닉 트랙 및 제 4 하모닉 트랙의 시작 위치 정보는 제 2 위치 셋 중 하나에 대응되는 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 10]

5 제 9 항에 있어서,

상기 하모닉 비율은, 복수의 하모닉 트랙들의 에너지 및 복수 개의 펄스의 에너지를 근거로 생성된 것임을 특징으로 하는 오디오 신호 처리 방법.

[청구항 11]

10 제 9 항에 있어서,

상기 제 1 위치 셋은 짝수 위치들에 해당하고, 제 2 위치 셋은 홀수 위치들에 해당하는 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 12]

15 제 9 항에 있어서,

상기 제 1 하모닉 트랙 중에서의 최대 펄스, 및 그의 인접 펄스, 상기 제 2 하모닉 트랙 중에서 최대 펄스, 및 그의 인접 펄스를 포함하는 제 1 타깃 벡터를 생성하는 단계;

20 상기 제 3 하모닉 트랙 중에서의 최대 펄스, 그의 인접 펄스, 상기 제 4 하모닉 트랙 중에서의 최대 펄스, 그의 인접 펄스를 포함하는 제 2 타깃 벡터를 생성하는 단계;

상기 제 1 타깃 벡터 및 상기 제 2 타깃 벡터를 벡터-양자화하는 단계;

25 상기 하모닉 트랙들에서 상기 제 1 타깃 벡터 및 상기 제 2 타깃 벡터를 제외한 대상에 대해 주파수 변환하는 단계를 더 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 13]

제 9 항에 있어서,

30 제 1 하모닉 트랙은 제 1 피치를 갖는 복수의 펄스들의 집합이고,

제 2 하모닉 트랙은 제 1 피치를 갖는 복수의 펄스들의 집합이고,

제 3 하모닉 트랙은 제 2 피치를 갖는 복수의 펄스들의 집합이고,

제 4 하모닉 트랙은 제 2 피치를 갖는 복수의 펄스들의 집합인 것을 특징으로 하는 오디오 신호 처리 방법.

35 [청구항 14]

제 9 항에 있어서,

상기 제 1 피치 및 상기 제 2 피치를 지시하는 피치 정보를 생성하는 단계를 더 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 15]

- 5 제 1 피치에 해당하는 제 1 그룹의 하모닉 트랙들 및 제 2 피치에 해당하는 제 2 그룹의 하모닉 트랙을 포함하는 복수의 하모닉 트랙들의 시작 위치 정보를 수신하는 단계;
상기 시작 위치 정보에 대응하는 복수의 하모닉 트랙을 생성하는 단계;
및,
10 상기 복수의 하모닉 트랙을 이용하여 현재 프레임에 해당하는 오디오 신호를 생성하는 단계를 포함하고,
상기 제 1 그룹의 하모닉 트랙들은 제 1 하모닉 트랙 및 제 2 하모닉 트랙을 포함하고,
상기 제 2 그룹의 하모닉 트랙들은 제 3 하모닉 트랙 및 제 4 하모닉
15 트랙을 포함하고,
상기 제 1 하모닉 트랙 및 제 3 하모닉 트랙의 시작 위치 정보는 제 1 위치 셋 중 하나에 대응되고,
상기 제 2 하모닉 트랙 및 제 4 하모닉 트랙의 시작 위치 정보는 제 2 위치 셋 중 하나에 대응되는 것을 특징으로 하는 오디오 신호 처리 방법.

20

[청구항 16]

- 오디오 신호에 대해서 주파수 변환을 수행하여, 복수개의 주파수 변환 계수들을 획득하는 단계;
상기 주파수 변환 계수들에 대해서, 프레임간 유사성을 근거로 년-토널
25 모드 및 토널 모드를 선택하는 단계;
상기 년-토널 모드가 선택된 경우, 펄스 비율을 근거로 일반 모드 및 비-일반모드 중 하나를 선택하는 단계;
상기 토널 모드가 선택된 경우, 하모닉 비율을 근거로 년-하모닉 모드 및 하모닉 모드 중 하나를 선택하는 단계; 및,
30 상기 선택된 모드에 따라서 상기 오디오 신호를 인코딩함으로써 파라미터를 생성하는 단계를 포함하고,
상기 일반 모드인 경우, 상기 파라미터는 인벨롭 위치정보 및 스케일링 정보를 포함하고,
상기 비-일반 모드인 경우, 상기 파라미터는 펄스정보 및 노이즈
35 에너지정보를 포함하고
상기 년-하모닉 모드인 경우, 상기 파라미터는 서브밴드별로 개수가 미리 정해진 고정 펄스들에 대한 정보인 고정 펄스정보를 포함하고,

상기 하모닉 모드인 경우, 상기 파라미터는 제 1 그룹의 하모닉 트랙들의 위치정보들, 제 2 그룹의 하모닉 트랙들의 위치정보들을 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

5 [청구항 17]

제 16 항에 있어서,

상기 선택된 모드에 따라서, 제 1 모드 정보 및 제 2 모드 정보를 생성하는 단계를 더 포함하고,

10 상기 제 1 모드 정보는, 상기 년-토널 모드 및 상기 토널 모드 중 하나를 지시하고,

상기 제 2 모드 정보는, 상기 제 1 모드 정보가 년-토널 모드를 지시하는 경우, 상기 일반모드 또는 상기 비-일반 모드 중 하나를 지시하고, 상기 제 1 모드 정보가 토널 모드를 지시하는 경우, 상기 년-하모닉 모드 및 상기 하모닉 모드 중 하나를 지시하는 것을 특징으로 하는 오디오 신호 처리 방법.

[청구항 18]

비트스트림을 통해 제 1 모드 정보 및 제 2 모드 정보를 추출하는 단계;

20 상기 제 1 모드 정보 및 제 2 모드 정보를 근거로 하여, 현재 프레임에 해당하는 현재 모드를 결정하는 단계;

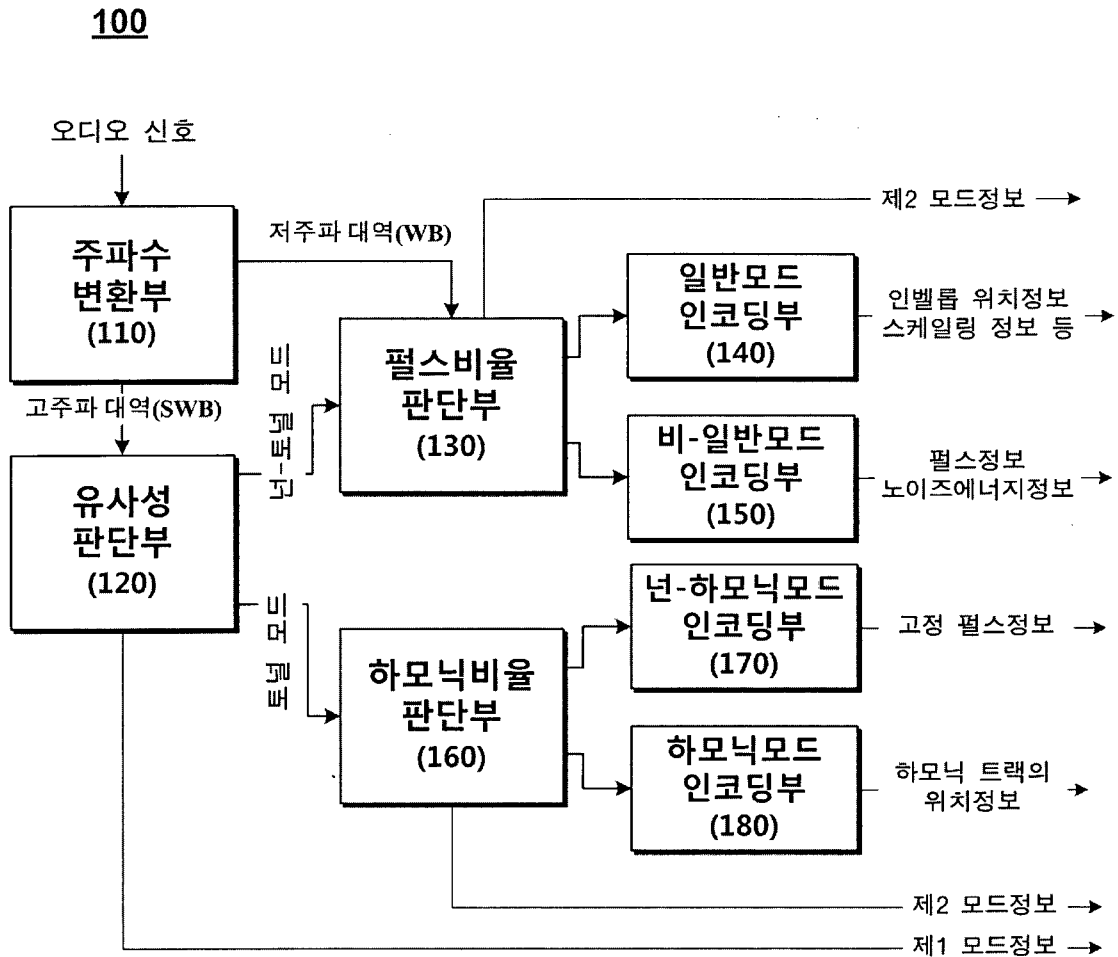
상기 현재 모드가 일반 모드인 경우, 인벨롭 위치 정보 및 스케일링 정보를 이용하여 현재 프레임의 오디오 신호를 복원하는 단계;

25 상기 현재 모드가 비-일반 모드인 경우, 펄스정보 및 노이즈 에너지 정보를 이용하여 현재 프레임의 오디오 신호를 복원하는 단계;

상기 현재 모드가 년-하모닉 모드인 경우, 서브밴드별로 개수가 미리 정해진 고정 펄스들에 대한 정보인 고정펄스정보를 이용하여 현재 프레임의 오디오 신호를 복원하는 단계; 및,

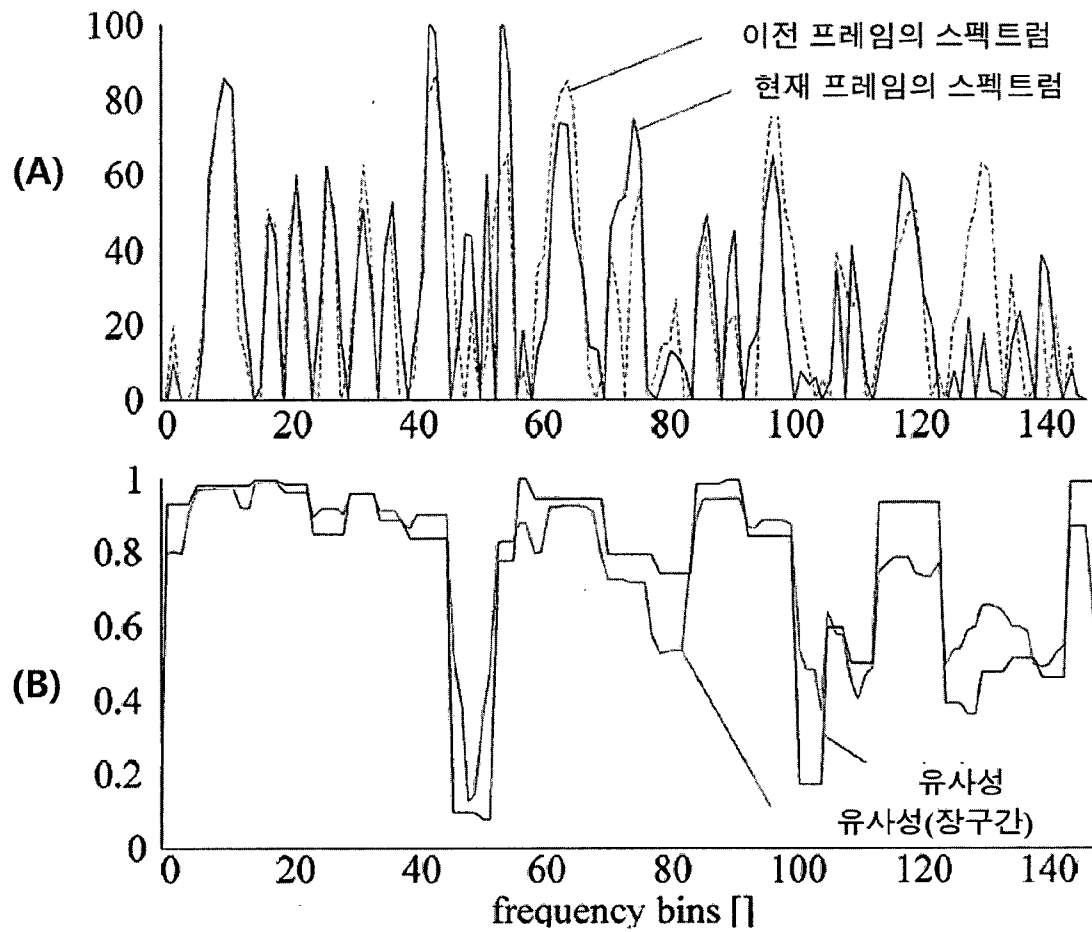
30 상기 현재 모드가 하모닉 모드인 경우, 제 1 그룹의 하모닉 트랙의 위치정보들 및 제 2 그룹의 하모닉 트랙의 위치정보들을 이용하여 현재 프레임의 오디오 신호를 복원하는 단계를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

FIG. 1



2/25

FIG. 2



<프레임간 유사성(tonality) 판단의 예>

3/25

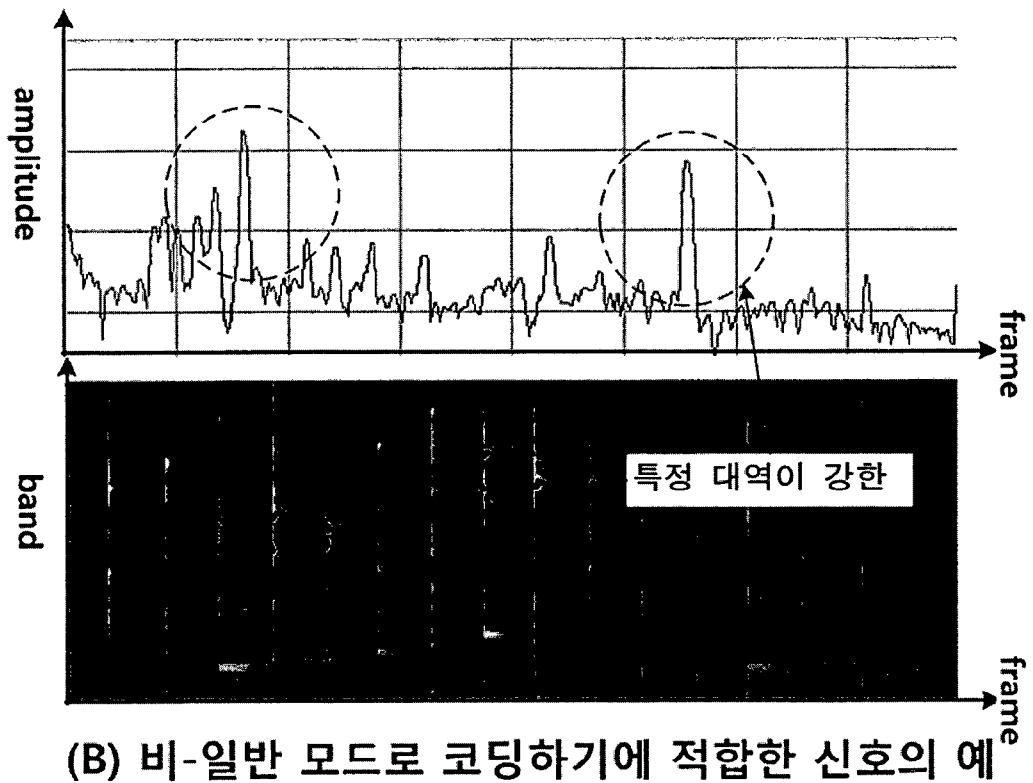
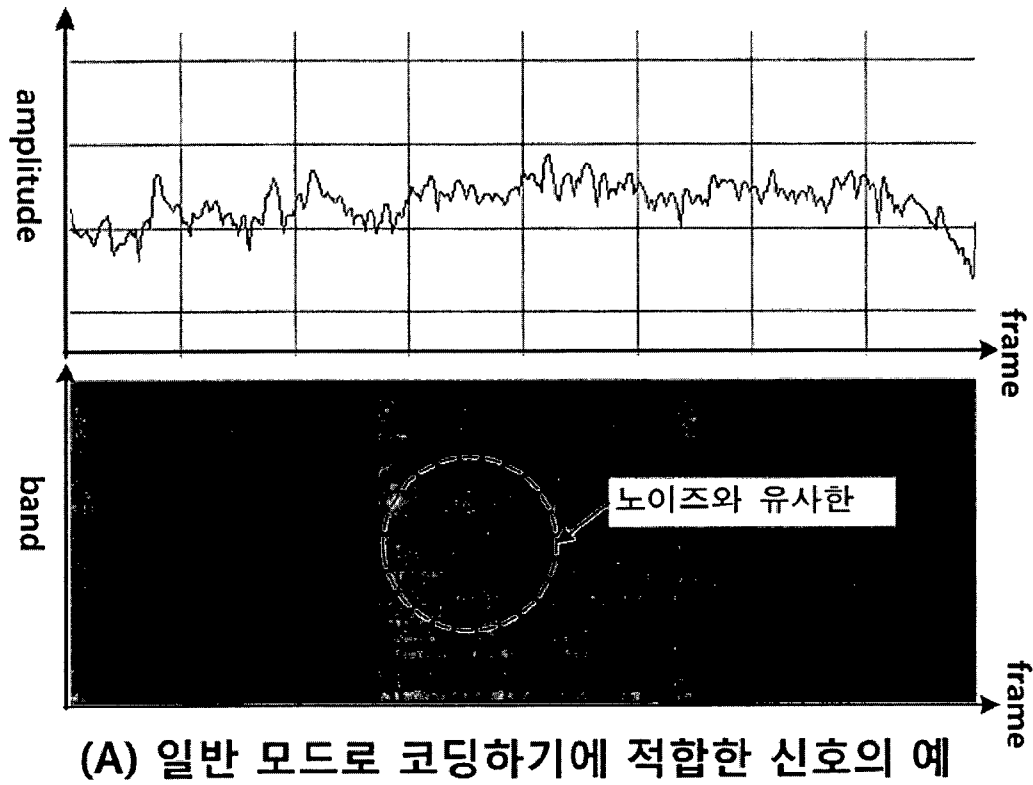
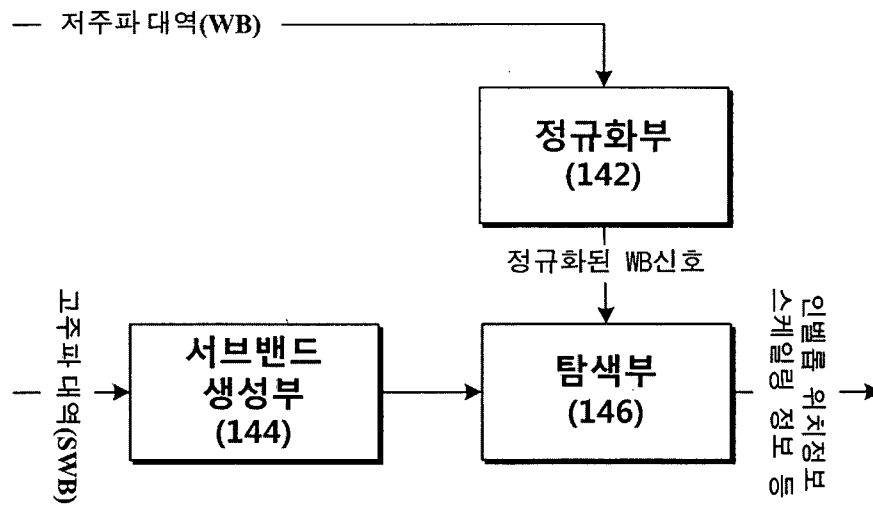
FIG. 3

FIG. 4

140

5/25

FIG. 5

설명	Parameter	Bits, breakdown	Bits, total
	SWB / Stereo	1	1
제1 모드 정보	tonal/non-tonal	1	1
제2 모드 정보	generic/non-generic	1	1
인벨롭 위치정보	Subband lags	8 + 7 + 8 + 7	30
스케일링 정보	Gain signs	1 + 1 + 1 + 1	4
	Gains, 1 st scaling	8 + 8	16
	Gains, 2 nd scaling	8	8
	Sinusoidal positions	5 + 5	10
	Sinusoidal signs	1	1
	Sinusoidal amplitudes	4 + 4	8
	Unused	1	1

일반 모드의 신택스의 예

FIG. 6

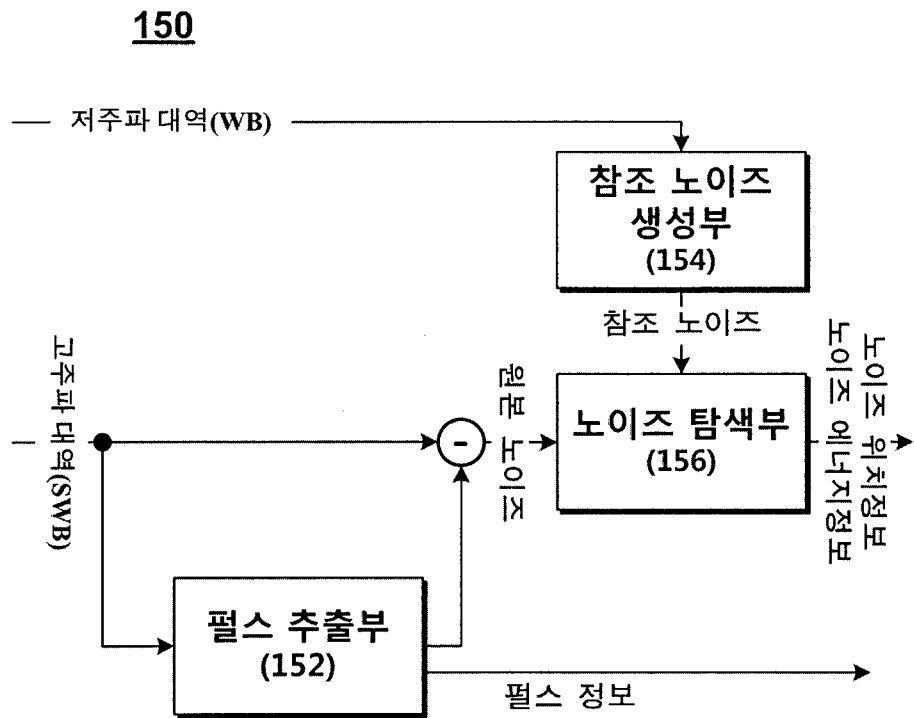


FIG. 7

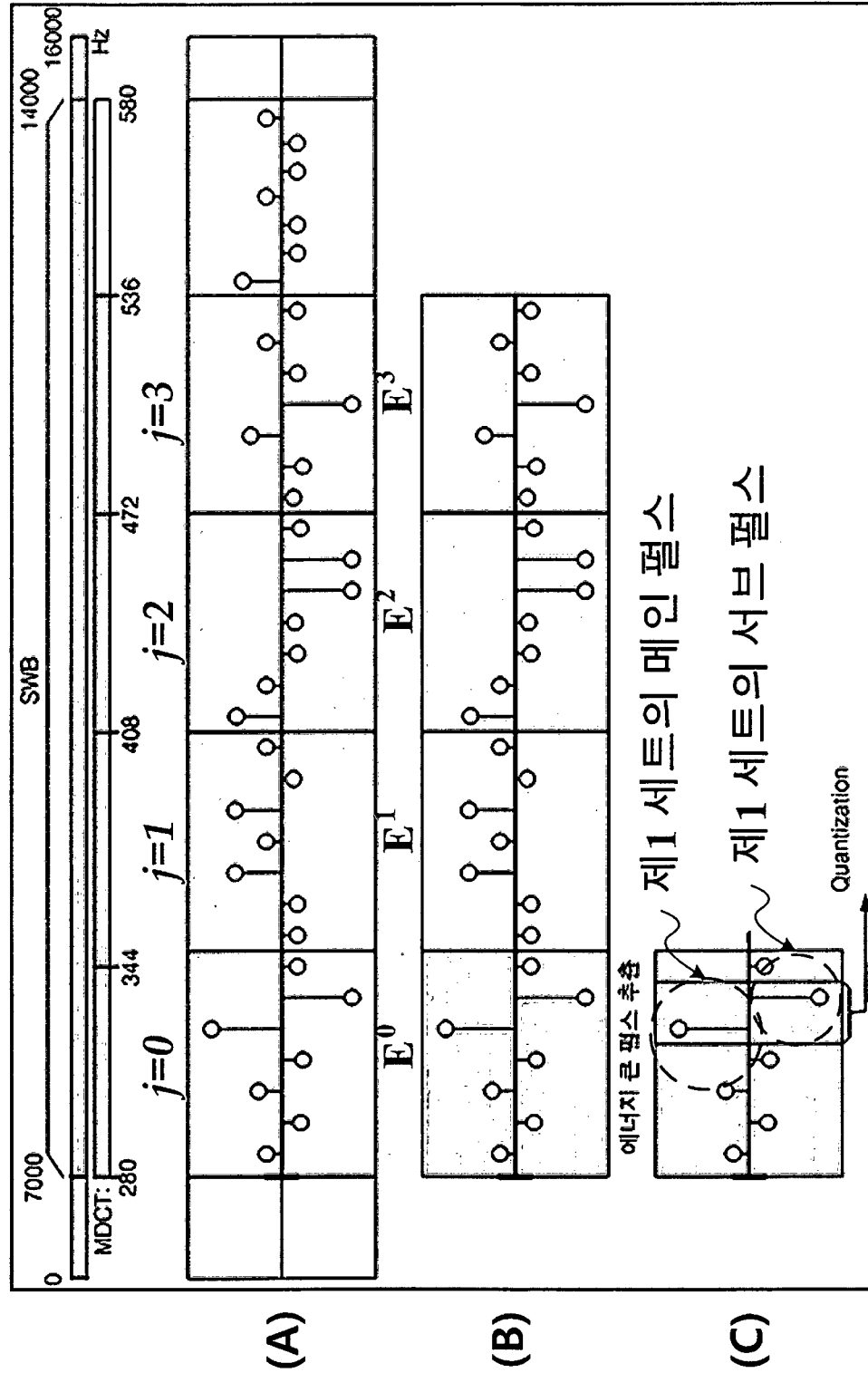


FIG. 8

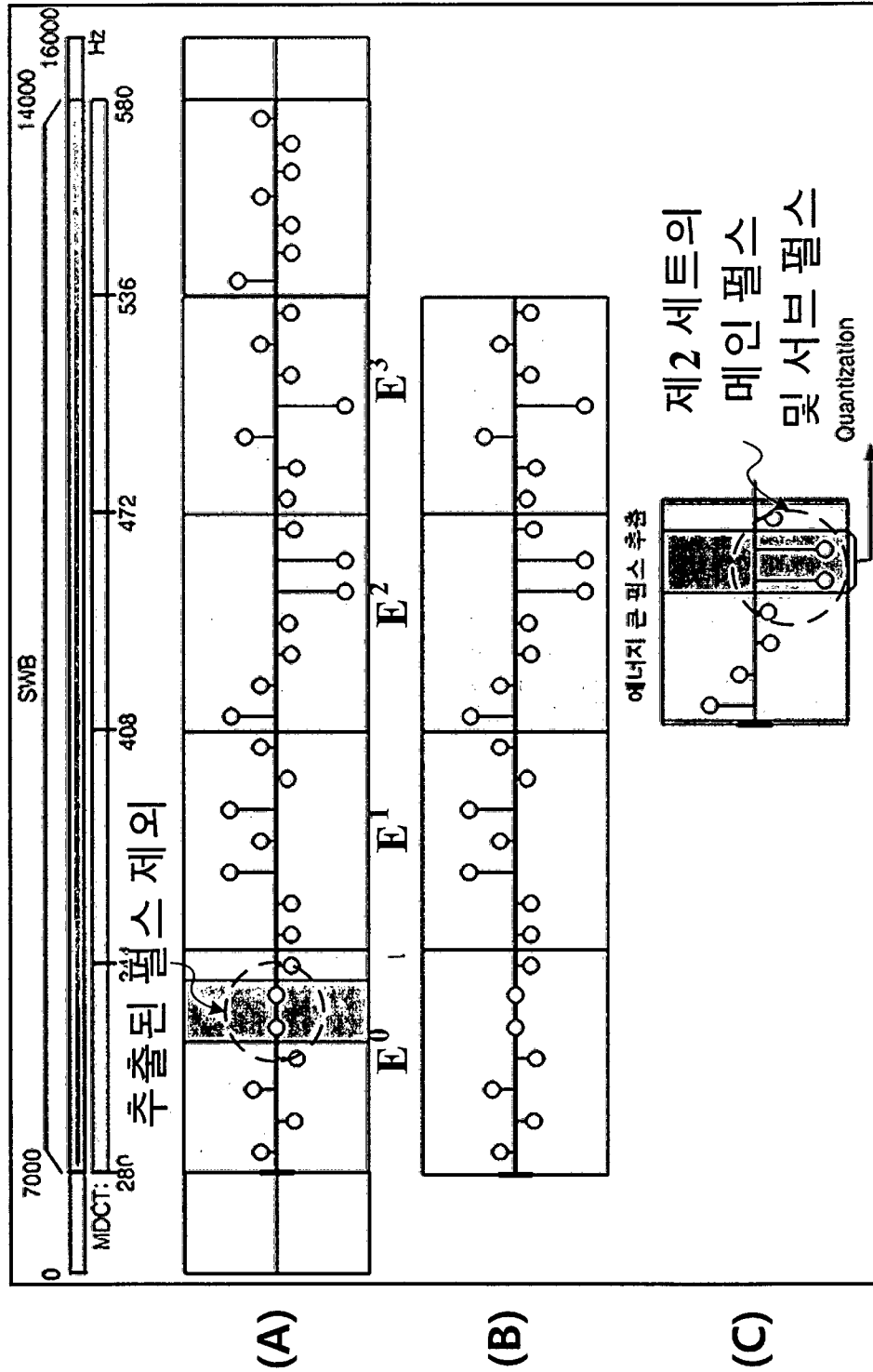
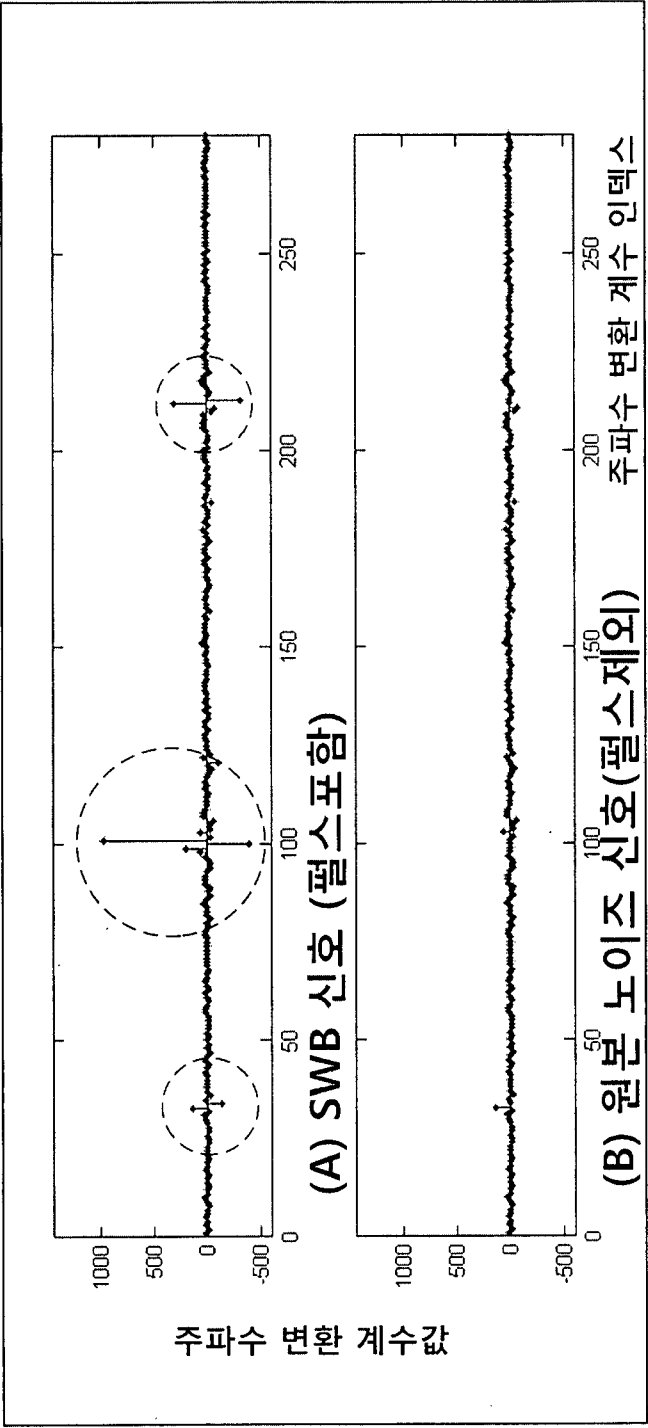
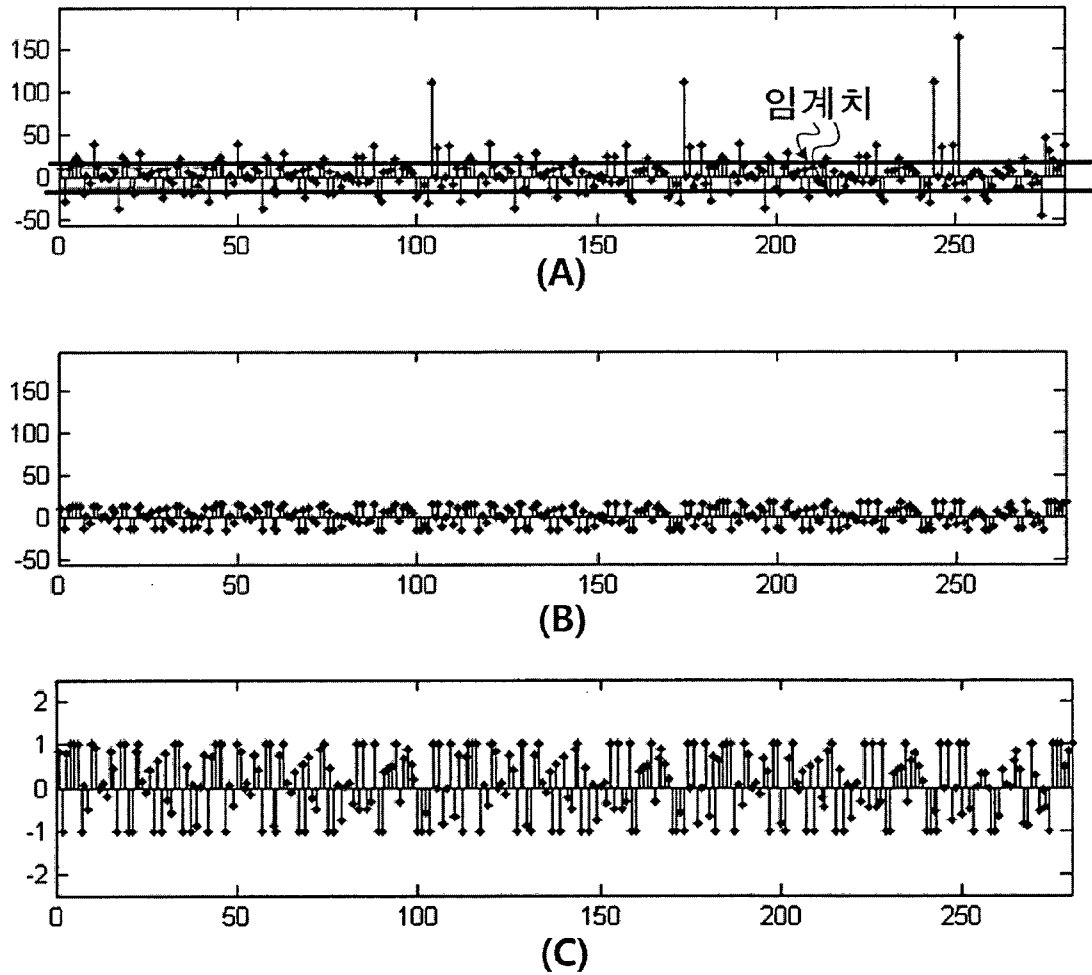


FIG. 9



10/25

FIG. 10



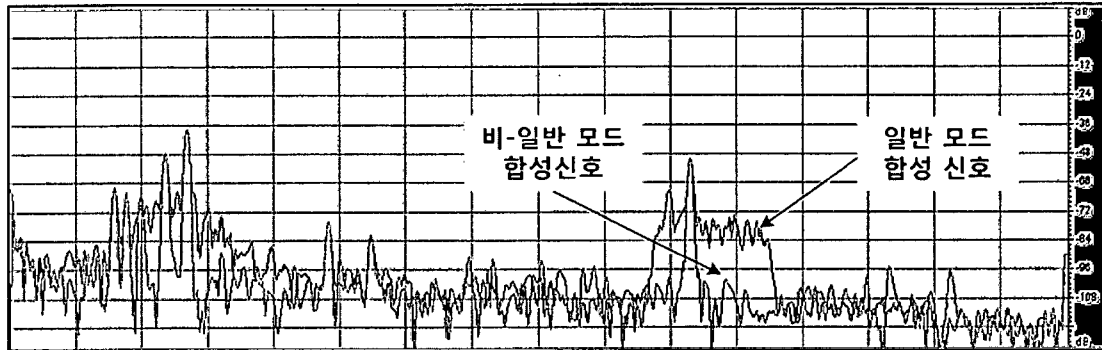
11/25

FIG. 11

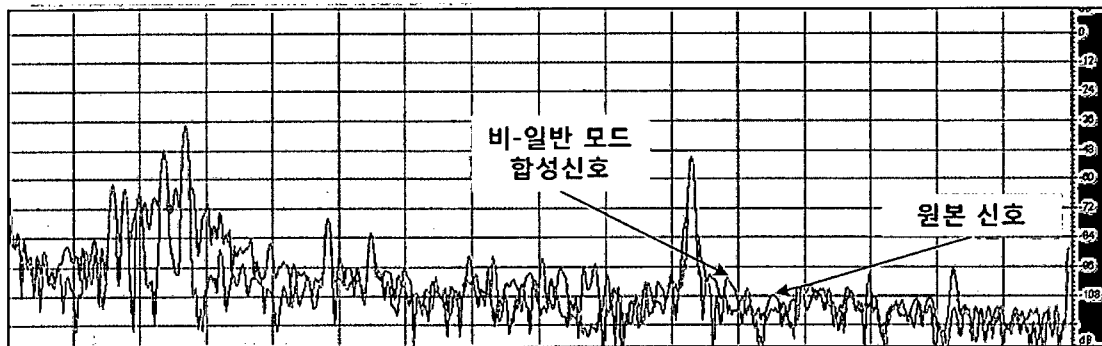
	Parameter	Bits, breakdown	Bits, total
	SWB / Stereo	1	1
제1 모드 정보	tonal/non-tonal	1	1
제2 모드 정보	generic/non-generic	1	1
노이즈 위치정보	Subband lags	3 + 2 + 2 + 2	9
펄스 위치 정보	Sinusoidal positions	6 + 6 + 6 + 6 + 6	30
펄스 부호 정보	Sinusoidal signs	1 + 1 + 1 + 1 + 1 + 1 + 1 + 1	8
펄스 크기 정보	Sinusoidal amplitudes	8 + 8	16
펄스 서브밴드 정보	Subband position	2 + 2 + 2 + 2 + 2	10
노이즈 에너지 정보	Subband average energy	4	4
	TOTAL		80 bits

12/25

FIG. 12

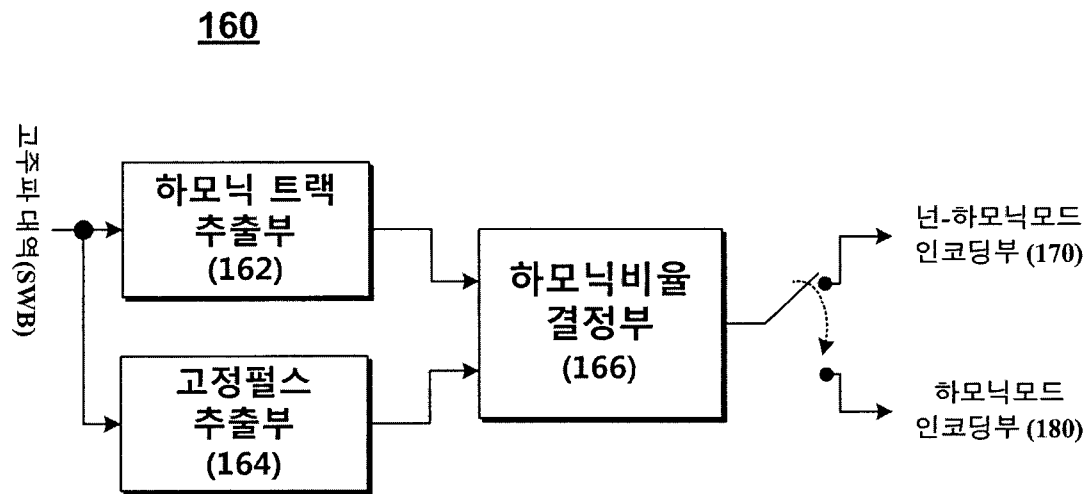


(A)



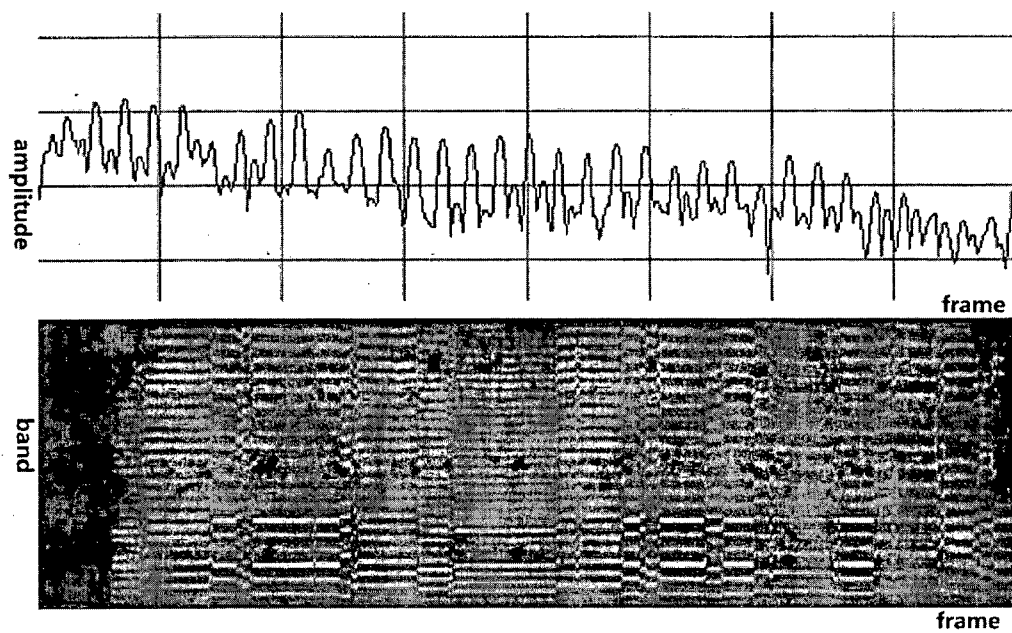
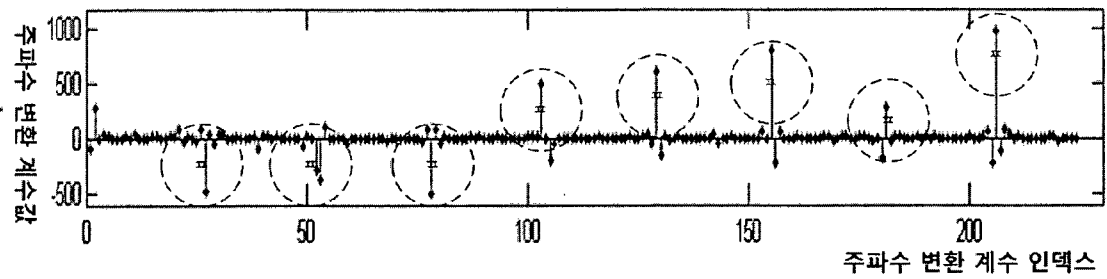
(B)

FIG. 13



14/25

FIG. 14



15/25

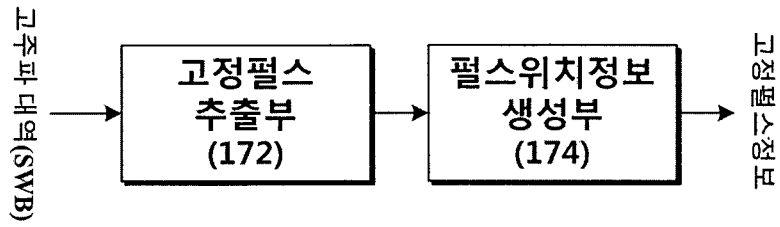
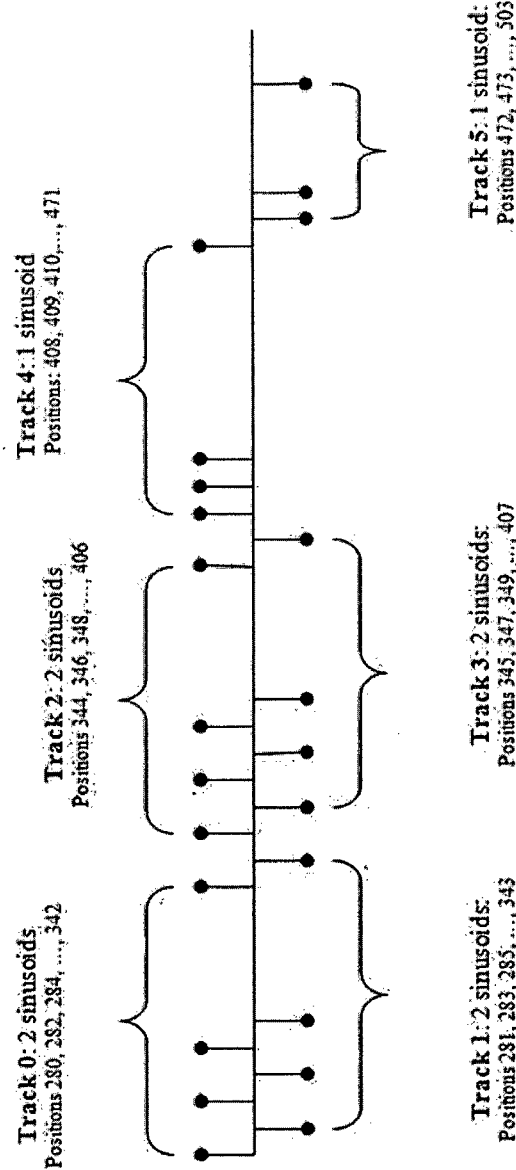
FIG. 15**170**

FIG. 16

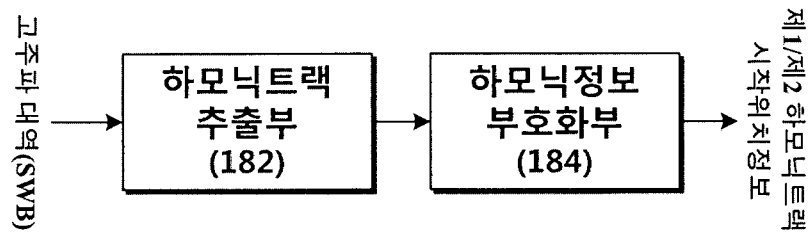


17/25

FIG. 17

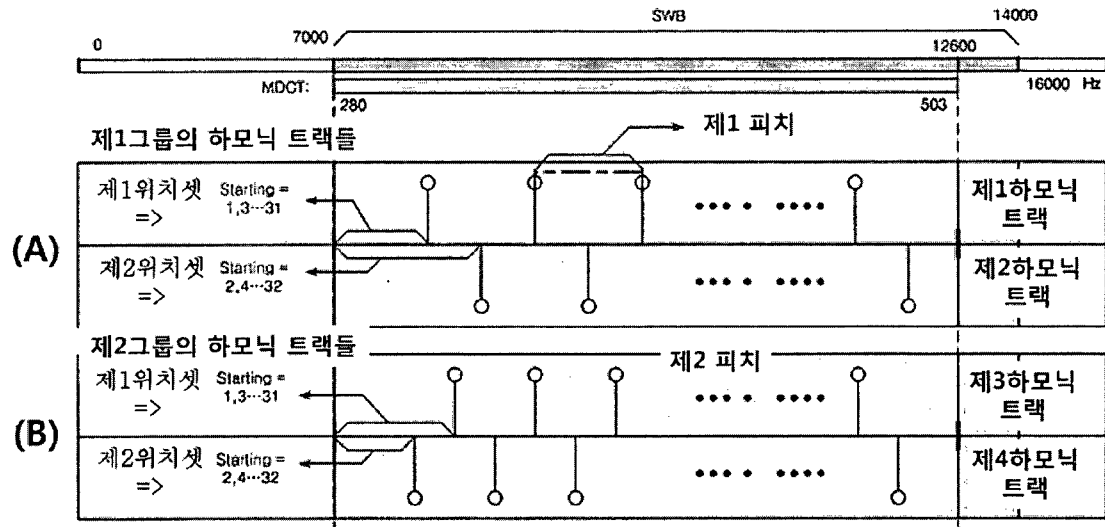
Track	Num of sinusoids	Starting position	Position step size	Length
0	2	280	2	32
1	2	281	2	32
2	2	344	2	32
3	2	345	2	32
4	1	408	1	64
5	1	472	1	32

18/25

FIG. 18**180**

19/25

FIG. 19



Pitch	Harmonic		Bit
제1 피치 20 ~ 27 [3bit]	제1하모닉트랙의 시작위치	1~31(홀수)[4bit]	7 bit
	제2하모닉트랙의 시작위치	2~32(짝수)[4bit]	4 bit
제2 피치 20 ~ 27 [3bit]	제3하모닉트랙의 시작위치	1~31(홀수)[4bit]	7 bit
	제4하모닉트랙의 시작위치	2~32(짝수)[4bit]	4 bit

(C)

Pitch	Track	Starting position	Position step size	Length
20-27	1	280	2	16
	2	281	2	16
20-27	3	280	2	16
	4	281	2	16

FIG. 20

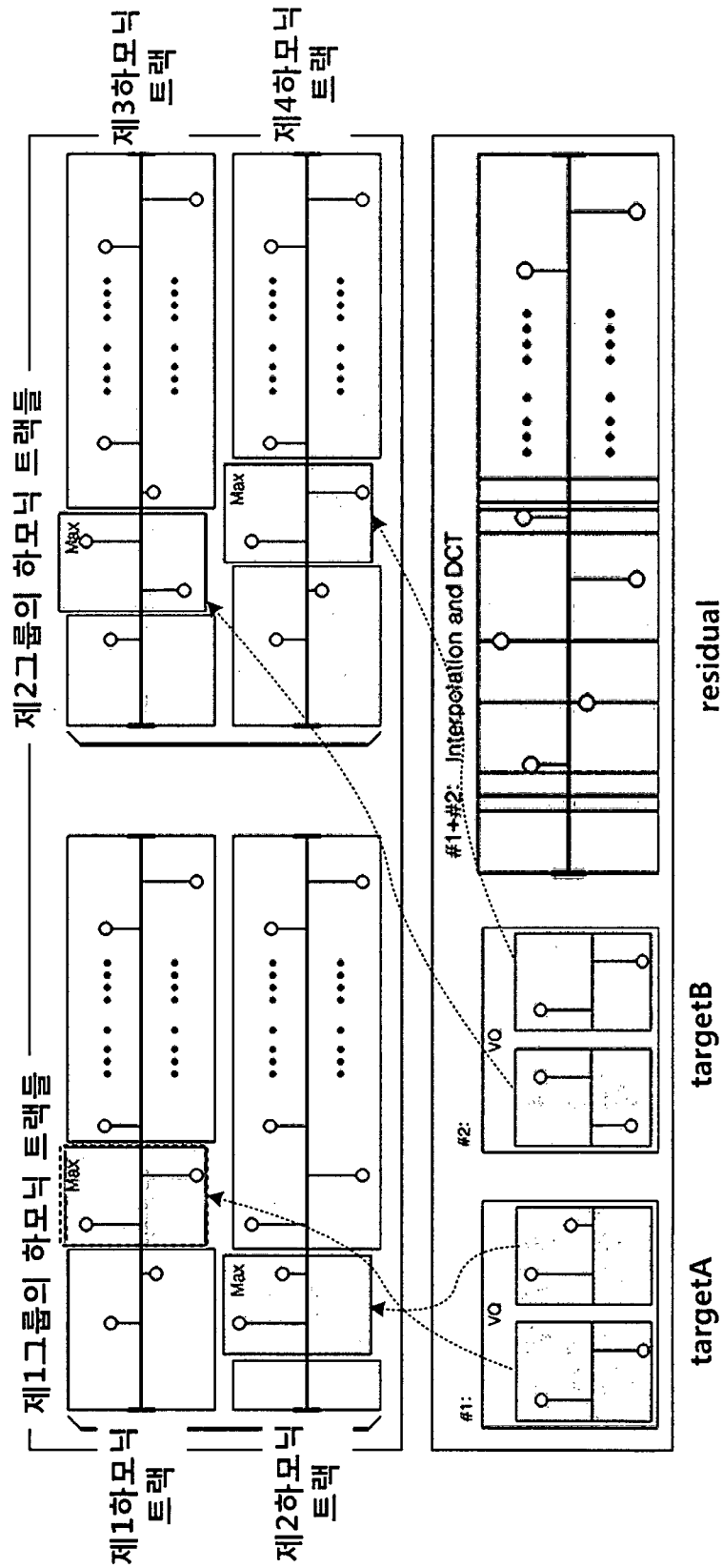
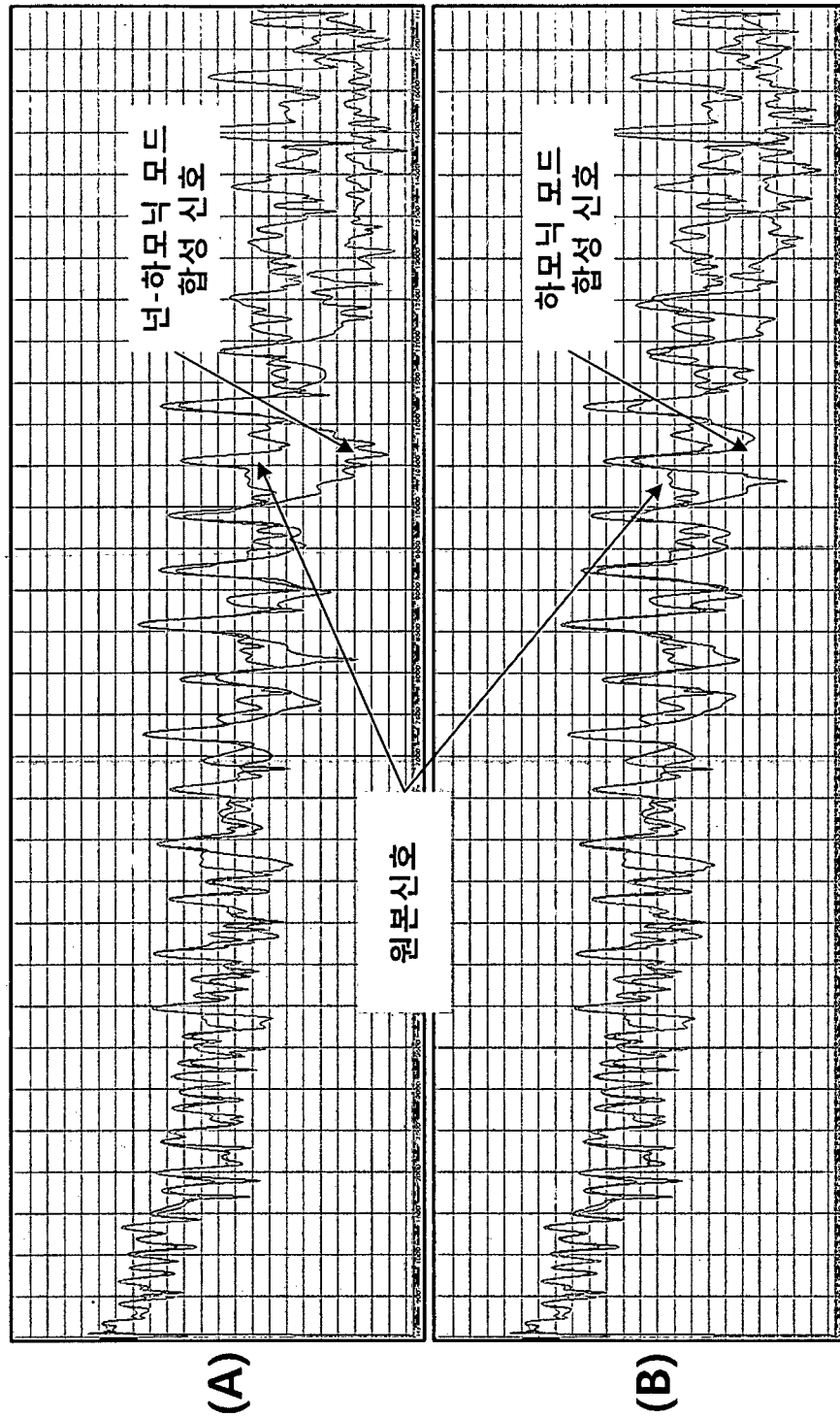


FIG. 21

	Parameter	Bits, breakdown	Bits, total
	SWB / Stereo	1	1
제1 모드 정보	tonal/non-tonal	1	1
제2 모드 정보	generic/non-generic	1	1
피치 정보(제1피치/제2 피치)	Pitch extraction	3 + 3	6
하모닉 트랙 정보	Harmonic extraction	4 + 4 + 4 + 4	16
하모닉 시작 위치 정보	Harmonic positions	2 + 2 + 2 + 2	8
펄스 크기 정보	Pulse amplitudes	12 + 12	24
	DCT coefficients	23	23
	TOTAL		80bits

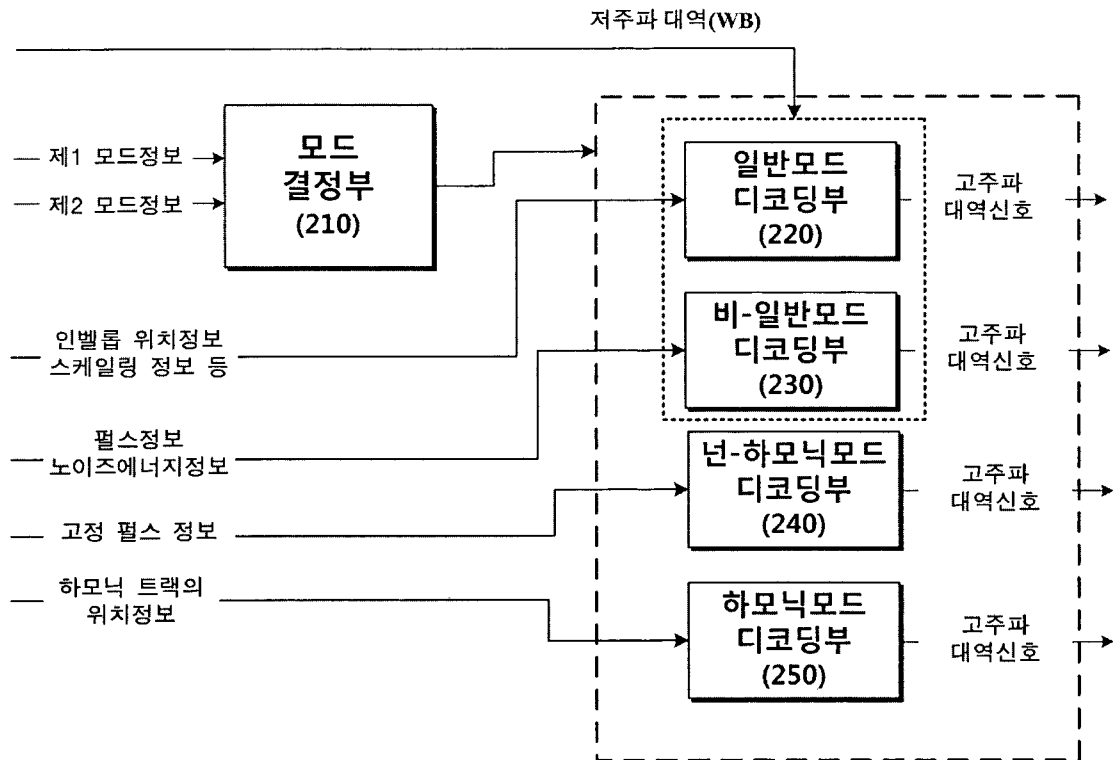
FIG. 22



23/25

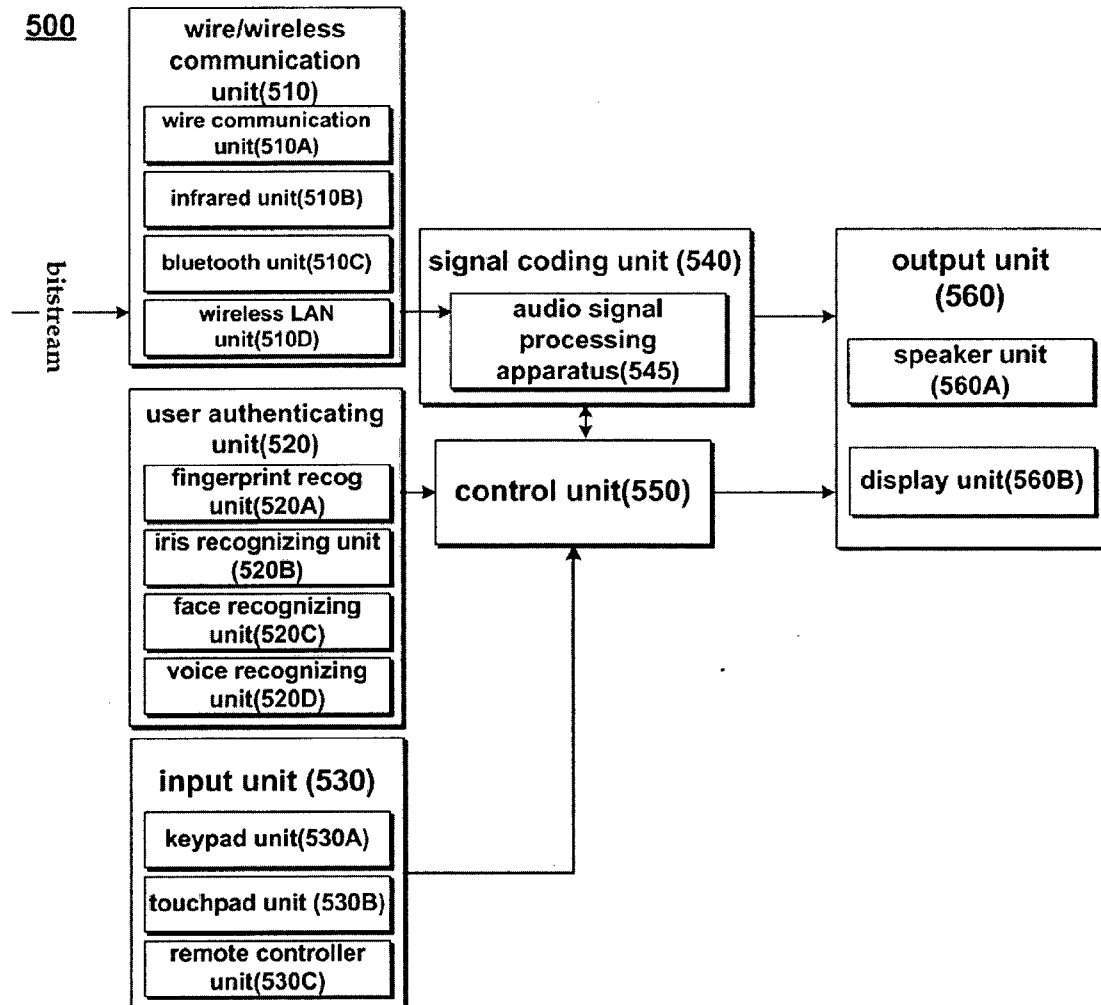
FIG. 23

200



24/25

FIG. 24

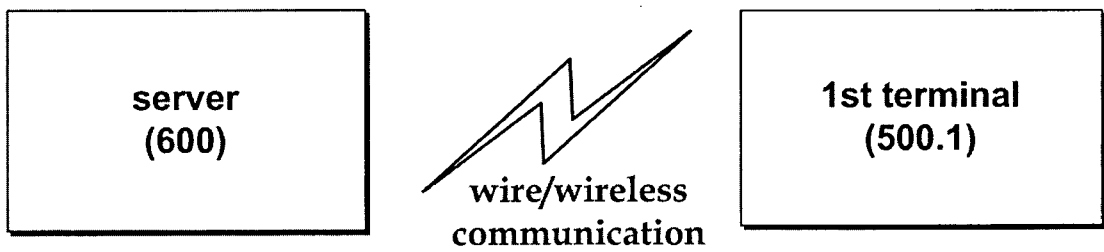


25/25

FIG. 25



(A)



(B)