



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2017년10월19일  
 (11) 등록번호 10-1788484  
 (24) 등록일자 2017년10월13일

(51) 국제특허분류(Int. Cl.)  
 G10L 19/005 (2013.01) G10L 19/008 (2014.01)

(52) CPC특허분류  
 G10L 19/005 (2013.01)  
 G10L 19/008 (2013.01)

(21) 출원번호 10-2016-7001580

(22) 출원일자(국제) 2014년06월23일  
 심사청구일자 2016년01월19일

(85) 번역문제출일자 2016년01월19일

(65) 공개번호 10-2016-0021295

(43) 공개일자 2016년02월24일

(86) 국제출원번호 PCT/EP2014/063176

(87) 국제공개번호 WO 2014/202789  
 국제공개일자 2014년12월24일

(30) 우선권주장  
 13173154.9 2013년06월21일  
 유럽특허청(EPO)(EP)  
 14166998.6 2014년05월05일  
 유럽특허청(EPO)(EP)

(56) 선행기술조사문헌  
 EP01775717 A1  
 EP02026330 A1\*  
 ISO/IEC FDIS 23003-3:2011(E), Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding. ISO/IEC JTC 1/SC 29/WG 11. 2011.09.20.\*  
 Audio codec processing functions, Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec. 3GPP TS 26.290 version 9.0.0 Release 9, 2009.09.  
 \*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
 프라운호퍼 게젤샤프트 쥘르 프뢰데롱 데어 안겐  
 반텐 포르슈 에. 베.  
 독일 80686 뮌헨 한자슈트라세 27 체

(72) 발명자  
 슈나벨, 미하엘  
 독일 95179 게롤드스그룬, 슈나이트베르크베그 4  
 에이  
 마르코비치, 고란  
 독일 90425 뉘른베르크, 아헨 스트라세 19  
 (뒷면에 계속)

(74) 대리인  
 특허법인 정안

전체 청구항 수 : 총 13 항

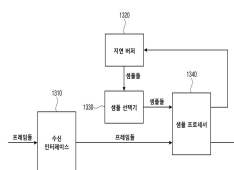
심사관 : 이남숙

(54) 발명의 명칭 T C X L T P를 이용하여 붕괴되거나 붕괴되지 않은 수신된 프레임들의 재구성성을 갖는 오디오 디코딩

**(57) 요약**

재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치가 제공된다. 장치는 복수의 프레임들을 수신하기 위한 수신 인터페이스(1310), 디코딩된 오디오 신호의 오디오 신호 샘플들을 저장하기 위한 지연 버퍼(1020; 1320), 지연 버퍼(1020; 1320)에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 (뒷면에 계속)

**대표도**



호 샘플들을 선택하기 위한 샘플 선택기(1030; 1330), 및 재구성된 오디오 신호의 재구성된 오디오 신호 샘플들을 얻기 위해 선택된 오디오 신호 샘플들을 처리하기 위한 샘플 프로세서(1040; 1340)를 포함한다. 샘플 선택기(1030; 1330)는, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되면, 그리고 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 피치 래그 정보에 따라 지연 버퍼(1020; 1320)에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성된다. 더욱이, 샘플 선택기(1030; 1330)는, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되지 않으면, 또는 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되면, 수신 인터페이스(1310)에 의해 이전에 수신되는 다른 프레임에 의해 포함되는 피치 래그 정보에 따라 지연 버퍼(1020; 1320)에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성된다.

(52) CPC특허분류

*G10L 2019/00* (2013.01)

(72) 발명자

**스피슈나يدر, 랄프**

독일 91320 에버만슈타트, 테베르트 75

**르콩트, 제레미**

독일 90763 쾰르트, 투른스트라쎄 7

**험리히, 크리스티안**

독일 91054 에를랑겐, 하우프트스트라쎄 68

## 명세서

### 청구범위

#### 청구항 1

재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치로서, 상기 장치는 복수의 프레임들을 수신하도록 구성되고, 상기 장치는:

상기 디코딩된 오디오 신호의 오디오 신호 샘플들을 얻기 위해 역 변형된 이산 코사인 변환(inverse modified discrete cosine transform)을 수행함으로써 상기 복수의 프레임들을 디코딩하기 위한 역 변형된 이산 코사인 변환 모듈(683), 및

장기간 예측을 수행하기 위한 장기간 예측 유닛(684)을 포함하고,

상기 장기간 예측 유닛(684)은

상기 디코딩된 오디오 신호의 상기 오디오 신호 샘플들을 저장하기 위한 지연 버퍼(1020),

상기 지연 버퍼(1020)에 저장되는 상기 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하기 위한 샘플 선택기(1030), 및

상기 재구성된 오디오 신호의 재구성된 오디오 신호 샘플들을 얻기 위해 상기 선택된 오디오 신호 샘플들을 처리하기 위한 샘플 프로세서(1040)를 포함하고,

상기 샘플 선택기(1030)는, 현재 프레임이 상기 장치에 의해 수신되면, 그리고 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되지 않으면, 상기 현재 프레임에 의해 포함되는 피치 래그(pitch lag) 정보에 따라 상기 지연 버퍼(1020)에 저장되는 상기 오디오 신호 샘플들로부터 상기 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성되고,

상기 샘플 선택기(1030)는, 상기 현재 프레임이 상기 장치에 의해 수신되지 않으면, 또는 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되면, 상기 장치에 의해 이전에 수신되는 다른 프레임에 의해 포함되는 피치 래그 정보에 따라 상기 지연 버퍼(1020)에 저장되는 상기 오디오 신호 샘플들로부터 상기 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성되고,

상기 샘플 선택기(1030)는, 변형된 이득에 따라 상기 선택된 오디오 신호 샘플들을 제스케일링함으로써 상기 재구성된 오디오 신호 샘플들을 얻도록 구성되고, 상기 변형된 이득은 수학적 식  $gain = gain\_past * damping$ 에 따라 정의되고,

여기서 gain은 상기 변형된 이득이고,

상기 샘플 선택기(1030)는, gain이 계산된 후에 gain\_past를 gain으로 설정하도록 구성되고,

damping은  $0 \leq damping \leq 1$ 을 갖는 실수인, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

#### 청구항 2

제 1항에 있어서,

상기 샘플 프로세서(1040)는, 상기 현재 프레임이 상기 장치에 의해 수신되면, 그리고 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되지 않으면, 상기 현재 프레임에 의해 포함되는 이득 정보에 따라 상기 선택된 오디오 신호 샘플들을 제스케일링(rescaling)함으로써 상기 재구성된 오디오 신호 샘플들을 얻도록 구성되고,

상기 샘플 선택기(1030)는, 상기 현재 프레임이 상기 장치에 의해 수신되지 않으면, 또는 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되면, 상기 장치에 의해 이전에 수신되는 상기 다른 프레임에 의해 포함되는 상기 이득 정보에 따라 상기 선택된 오디오 신호 샘플들을 제스케일링함으로써 상기 재구성된 오디오 신호 샘플들을 얻도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 3**

제 2항에 있어서,

상기 샘플 프로세서(1040)는, 상기 현재 프레임이 상기 장치에 의해 수신되면, 그리고 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되지 않으면, 상기 현재 프레임에 의해 포함되는 상기 이득 정보에 따른 값과 상기 선택된 오디오 신호 샘플들을 곱함으로써, 상기 재구성된 오디오 신호 샘플들을 얻도록 구성되고,

상기 샘플 선택기(1030)는, 상기 현재 프레임이 상기 장치에 의해 수신되지 않으면, 또는 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되면, 상기 장치에 의해 이전에 수신되는 상기 다른 프레임에 의해 포함되는 상기 이득 정보에 따른 값과 상기 선택된 오디오 신호 샘플들을 곱함으로써, 상기 재구성된 오디오 신호 샘플들을 얻도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 4**

제 1항에 있어서, 상기 샘플 프로세서(1040)는 상기 재구성된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 5**

제 4항에 있어서, 상기 샘플 프로세서(1040)는, 추가 프레임이 상기 장치에 의해 수신되기 전에, 상기 재구성된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 6**

제 4항에 있어서, 상기 샘플 프로세서(1040)는, 추가 프레임이 상기 장치에 의해 수신된 후에 상기 재구성된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 7**

제 1항에 있어서, 상기 샘플 프로세서(1040)는, 재스케일링된 오디오 신호 샘플들을 얻기 위해 상기 이득 정보에 따라, 그리고 상기 처리된 오디오 신호 샘플들을 얻기 위해 상기 재스케일링된 오디오 신호 샘플들과 입력 오디오 신호 샘플들을 조합함으로써, 상기 선택된 오디오 신호 샘플들을 재스케일링하도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 8**

제 7항에 있어서,

상기 샘플 프로세서(1040)는, 상기 현재 프레임이 상기 장치에 의해 수신되면, 그리고 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되지 않으면, 상기 재스케일링된 오디오 신호 샘플들과 상기 입력 오디오 신호 샘플들의 조합을 나타내는 상기 처리된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하고, 상기 재스케일링된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하지 않도록 구성되고,

상기 샘플 프로세서(1040)는, 상기 현재 프레임이 상기 장치에 의해 수신되지 않으면, 또는 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되면, 상기 재스케일링된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하고, 상기 처리된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하지 않도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 9**

제 7항에 있어서, 상기 샘플 프로세서(1040)는, 상기 현재 프레임이 상기 장치에 의해 수신되지 않으면, 또는 상기 장치에 의해 수신되는 상기 현재 프레임이 붕괴되면, 상기 처리된 오디오 신호 샘플들을 상기 지연 버퍼(1020)에 저장하도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 10**

제 1항에 있어서, 상기 샘플 선택기(1030)는 상기 변형된 이득을 계산하도록 구성되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 11**

제 1항에 있어서, 마지막 프레임이 상기 장치에 의해 수신되었기 때문에, 프레임들의 적어도 미리 한정된 개수가 상기 장치에 의해 수신되지 않으면, 상기 변형된 이득 gain은 0으로 설정되는, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치.

**청구항 12**

재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 방법으로서,

복수의 프레임들을 수신하는 단계,

상기 디코딩된 오디오 신호의 오디오 신호 샘플들을 얻기 위해 역 변형된 이산 코사인 변환을 수행함으로써 상기 복수의 프레임들을 디코딩하는 단계,

상기 디코딩된 오디오 신호의 상기 오디오 신호 샘플들을 저장함으로써,

지연 버퍼(1020)에 저장되는 상기 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택함으로써, 및

상기 재구성된 오디오 신호의 재구성된 오디오 신호 샘플들을 얻기 위해 상기 선택된 오디오 신호 샘플들을 처리함으로써

장기간 예측을 수행하는 단계를 포함하고,

현재 프레임이 수신되면, 그리고 수신되는 상기 현재 프레임이 붕괴되지 않으면, 상기 지연 버퍼(1020)에 저장되는 상기 오디오 신호 샘플들로부터 상기 복수의 선택된 오디오 신호 샘플들을 선택하는 단계는 상기 현재 프레임에 의해 포함되는 피치 래그 정보에 따라 수행되고,

상기 현재 프레임이 수신되지 않으면, 또는 수신되는 상기 현재 프레임이 붕괴되면, 상기 지연 버퍼(1020)에 저장되는 상기 오디오 신호 샘플들로부터 상기 복수의 선택된 오디오 신호 샘플들을 선택하는 단계는 이전에 수신되는 다른 프레임에 의해 포함되는 피치 래그 정보에 따라 수행되고,

상기 방법은 변형된 이득에 따라 상기 선택된 오디오 신호 샘플들을 재스케일링하는 단계를 더 포함하고, 상기 변형된 이득은 수학적 식  $gain = gain\_past * damping$ 에 따라 정의되고,

여기서 gain은 상기 변형된 이득이고,

gain이 계산된 후에 gain\_past는 gain으로 설정되고,

damping은  $0 \leq damping \leq 1$ 을 갖는 실수인, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 방법.

**청구항 13**

컴퓨터 또는 신호 프로세서 상에서 실행될 때 제12항의 방법을 구현하기 위한 컴퓨터 프로그램을 포함하는 컴퓨터 판독가능 기록매체.

**청구항 14**

삭제

**청구항 15**

삭제

**발명의 설명**

**기술 분야**

[0001] 본 발명은 오디오 신호 인코딩, 처리 및 디코딩에 관한 것으로, 특히 에러 은닉(error concealment) 동안 스위칭된 오디오 코딩 시스템들에 대해 개선된 신호 페이드 아웃(fade out)을 위한 장치 및 방법에 관한 것이다.

**배경 기술**

[0002] 다음에서, 종래 기술은 패킷 손실 은닉(PLC) 동안 음성 및 오디오 코덱 페이드 아웃에 관해 기재된다. 종래 기술에 관한 설명들은 G-시리즈(G.718, G.719, G.722, G.722.1, G.729, G.729.1)의 ITU-T에서 시작하여, 뒤이어 3GPP 코덱들(AMR, AMR-WB, AMR-WB+) 및 하나의 IETF 코덱(OPUS)이 후속하고, 2개의 MPEG 코덱들(HE-AAC, HIL N)에서 종료한다(ITU=International Telecommunication Union; 3GPP=3rd Generation Partnership Project; AMR=Adaptive Multi-Rate; WB=Wideband; IETF=Internet Engineering Task Force). 후속하여, 배경 잡음 레벨의 추적에 관한 종래 기술이 분석되고, 뒤이어 개요를 제공하는 요약이 후속한다.

[0003] 먼저, G.718이 고려된다. G.718은 협대역이고 광역 음성 코덱이고, 이것은 DTX/CNG(DTX = Digital Theater Systems; CNG=Comfort Noise Generation)를 지원한다. 실시예들이 특히 저 지연 코드에 관한 것일 때, 저 지연 버전 모드가 본 명세서에서 더 구체적으로 기재될 것이다.

[0004] ACELP(층 1)(ACELP=Algebraic code Excited Linear Prediction)을 고려하여, ITU-T는 G.718[ITU08a, 섹션 7.11]에 대해 페이딩 속도를 제어하기 위해 선형 예측 도메인에서 적응형 페이드 아웃을 권고한다. 일반적으로, 은닉은 이러한 원리에 따른다:

[0005] G.718에 따라, 프레임 소거들(erasures)의 경우에, 은닉 전략은 배경 잡음의 추정된 파라미터들로의 신호 에너지 및 스펙트럼 엔벨로프(spectral envelope)의 수렴으로서 요약될 수 있다. 신호의 주기성은 0으로 수렴된다. 수렴의 속도는 마지막으로 정확히 수신된 프레임의 파라미터들과, 연속 소거된 프레임들의 개수에 의존하고, 감쇄 인자( $\alpha$ )에 의해 제어된다. 감쇄 인자( $\alpha$ )는 무성음(UNVOICED) 프레임들에 대해 LP 필터(LP=Linear Prediction)의 안정도( $\theta$ )에 추가로 의존한다. 일반적으로, 수렴은, 마지막으로 양호하게 수신된 프레임이 안정된 세그먼트에 있는 경우 느려지고, 프레임이 전이 세그먼트에 있는 경우 빨라진다.

[0006] 감쇄 인자( $\alpha$ )는 음성 신호 클래스에 의존하고, 이러한 음성 신호 클래스는 [ITU08a, 섹션 6.8.1.3.1 및 7.11.1.1]에 기재된 신호 클래스에 의해 도출된다. 안정도 인자( $\theta$ )는 인접한 ISF(Immittance Spectral Frequency) 필터들 사이의 거리 측정에 기초하여 계산된다[ITU08a, 섹션 7.1.2.4.2].

[0007] 표 1은  $\alpha$ 의 계산 계획을 보여준다.

**표 1**

감쇄 인자  $\alpha$ 의 값들, 값  $\theta$ 은 인접한 LP 필터들 사이의 거리 측정으로부터 계산된 안정도 인자이다.  
[ITU08a, 섹션 7.1.2.4.2].

마지막으로 양호하게 수신된 프레임	연속적인 소거된 프레임들의 개수	$\alpha$
인공 은셋		0.6
은셋, 유성음	$\leq 3$	1.0
	$> 3$	0.4
유성음 전이		0.4
무성음 전이		0.8
무성음	= 1	$0.2 \cdot \theta + 0.8$
	= 2	0.6
	$> 2$	0.4

[0008]

[0009] 더욱이, G.718은 스펙트럼 엔벨로프를 변형하기 위해 페이딩 방법을 제공한다. 일반적인 생각은 마지막 ISF 파라미터들을 적응형 ISF 평균 벡터쪽으로 수렴하는 것이다. 처음에, 평균 ISF 벡터는 마지막의 3개의 알려진 ISF 벡터들로부터 계산된다. 그런 후에 평균 ISF 벡터는 오프라인으로 트레이닝된(trained) 장기간 ISF 벡터(상수 벡터임)로 다시 평균화된다[ITU08a, 섹션 7.11.1.2].

[0010] 더욱이, G.718은 장기간 행위를 제어하여, 배경 잡음과의 상호 작용을 제어하는 페이딩 방법을 제공하며, 여기서 피치(pitch) 여기 에너지(이에 따라 여기 주기성)는 0으로 수렴하는 한편, 랜덤 여기 에너지는 CNG 여기 에너지로 수렴한다[ITU08a, 섹션 7.11.1.6]. 혁신(innovation) 이득 감쇄는 수학적 1로서 계산된다:

수학식 1

$$g_s^{[1]} = \alpha g_s^{[0]} + (1 - \alpha)g_n$$

[0011]

[0012] 여기서  $g_s^{[1]}$ 는 다음 프레임의 시작에서의 혁신 이득이고,  $g_s^{[0]}$ 은 현재 프레임의 시작에서의 혁신 이득이고,  $g_n$ 은 안락 잡음 생성 동안 사용된 여기의 이득이고,  $\alpha$ 는 감쇄 인자이다.

[0013] 주기 여기 감쇄와 유사하게, 이득은  $g_s^{[0]}$ 에서 시작하는 샘플 간에 기초하여 프레임 전체에 선형으로 감쇄되고, 다음 프레임의 시작에서  $g_s^{[1]}$ 에 도달한다.

[0014] 도 2는 G.718의 디코더 구조를 개설했다. 특히, 도 2는 PLC에 대한 높은 레벨의 G.718 디코더 구조를 도시하며, 이것은 고역 통과 필터를 특징으로 한다.

[0015] G.718의 전술한 접근법에 의해, 혁신 이득( $g_s$ )은 패킷 손실들의 긴 버스트들(bursts)에 대한 안락 잡음 생성( $g_n$ ) 동안 사용된 이득으로 수렴한다. [ITU08a, 섹션 6.12.3]에 기재된 바와 같이, 안락 잡음 이득( $g_n$ )은 에너지 ( $\tilde{E}$ )의 제공근으로서 주어진다.  $\tilde{E}$ 의 업데이트의 조건들은 구체적으로 기재되지 않는다. 기준 구현(부동 소수점 C-코드, stat\_noise\_uv\_mod.c)에 후속하여,  $\tilde{E}$ 는 다음과 같이 도출된다:

```
if(unvoiced_vad == 0){
    if( unv_cnt > 20 ){
        ftmp = lp_gainc * lp_gainc;
        lp_ener = 0.7f * lp_ener + 0.3f * ftmp;
    }
    else{
        unv_cnt++;
    }
}
else{
    unv_cnt = 0;
}
```

[0016]

[0017] unvoiced\_vad는 보이스 활동 검색을 나타내고, unv\_cnt는 행에서 무성음 프레임들의 개수를 나타내고, lp\_gainc는 고정된 코드북의 저역 통과된 이득들을 나타내고, lp\_ener는 저역 통과된 CNG 에너지 추정치( $\tilde{E}$ )를 나타내고, 이것은 0으로 초기화된다.

[0018] 더욱이, G.718은 무성음 여기의 신호 경로에 도입된 고역 통과 필터를 제공하고, 마지막으로 양호한 프레임의 신호가 무성음과 상이하게 분류되었으면, 도 2를 참조하고, 또한 [ITU08a, 섹션 7.11.1.6]을 참조하자. 이 필터는 나이키스트 주파수보다 약 5 dB 낮은 DC에서 주파수 응답을 갖는 낮은 셸프(shelf) 특징을 갖는다.

[0019] 더욱이, G.718은 결합 해제된(decoupled) LTP 피드백 루프(LTP=Long-Term Prediction)를 제안한다: 정상 동작 동안 적응형 코드북에 대한 피드백 루프는 전체 여기에 기초하여 서브프레임-방식으로([ITU08a, 섹션 7.1.2.1.4]) 업데이트된다. 은닉 동안, 이러한 피드백 루프는 유성음 여기에만 기초하여 프레임-방식으로([ITU08a, 섹션들 7.11.1.4, 7.11.2.4, 7.11.1.6, 7.11.2.6; dec\_GV\_exc@dec\_gen\_voic.c 및 syn\_bfi\_post@syn\_bfi\_pre\_post.c]) 업데이트되고, 적응형 코드북은 무작위로 선택된 혁신 여기에 의해 그 기원

을 갖는 잡음으로 "오염(polluted)"되지 않는다.

- [0020] G.718의 변환 코딩된 개선 층들(3-5)에 관해, 은닉 동안, 디코더는 MDCT 스펙트럼이 0으로 설정되는 것처럼, 정상 동작과 유사한 하이(high) 층에 관해 작용한다. 은닉 동안 특수한 페이드-아웃 행위는 적용되지 않는다.
- [0021] G.718에서 CNG에 대해, CNG 합성은 다음의 순서로 이루어진다. 먼저, 안락 잡음 프레임의 파라미터들이 디코딩된다. 그런 후에, 안락 잡음 프레임은 합성된다. 그 후에 피치 버퍼가 재설정된다. 그런 후에, FER(Frame Error Recovery) 분류에 대한 합성이 저장된다. 그 후에, 스펙트럼 디엠퍼시스가 수행된다. 그런 후에, 저주파수 후치-필터링이 수행된다. 그런 후에, CNG 변수들이 업데이트된다.
- [0022] 은닉의 경우에, CNG 파라미터들이 비트스트림으로부터 디코딩되지 않는다는 점을 제외하고 정확히 동일한 것이 수행된다. 이것은, 파라미터들이 프레임 손실 동안 업데이트되지 않고, 마지막으로 양호한 SID(Silence Insertion Descriptor) 프레임으로부터의 디코딩된 파라미터들이 사용된다는 것을 의미한다.
- [0023] 이제, G.719가 고려된다. 사이렌(Siren) 22에 기초한 G.719는 전체-대역 오디오 코덱에 기초한 변환이다. ITU-T는 G.719에 대해 스펙트럼 도메인에서 프레임 반복을 갖는 페이드-아웃을 권고한다[ITU08b, 섹션 8.6]. G.719에 따라, 프레임 소거 은닉 메커니즘은 디코더에 병합된다. 프레임이 정확히 수신될 때, 재구성된 변환 계수들은 버퍼에 저장된다. 프레임이 손실되었다는 것 또는 프레임이 붕괴되었다는 것이 디코더에게 통보되면, 가장 최근에 수신된 프레임에서 재구성된 변환 계수들은 인자 0.5로 점점 줄어들게 스케일링(scaled)되고, 그런 후에 현재 프레임에 대한 재구성된 변환 계수들로서 사용된다. 디코더는 재구성된 변환 계수들을 시간 도메인으로 변환하고 윈도우-중첩-추가(windowing-overlap-add) 동작을 수행함으로써 진행된다.
- [0024] 다음에서, G.722가 기재된다. G.722는 최대 64 kbit/s의 비트율 내에서 서브밴드 적응형 차동 펄스 코드 변조(SB-ADPCM)를 이용하는 70 내지 7000 Hz 코딩 시스템이다. 신호는 QMF 분석(QMF = Quadrature Mirror Filter)을 이용하여 높은 및 낮은 서브밴드로 분할된다. 결과적인 2개의 밴드들은 ADPCM-코딩된다(ADPCM = Adaptive Differential Pulse Code Modulation).
- [0025] G.722에 대해, 패킷 손실 은닉을 위한 높은-복잡도의 알고리즘은 첨부 III[ITU06a]에서 규정되고, 패킷 손실 은닉을 위한 낮은-복잡도의 알고리즘은 첨부 IV[ITU07]에서 규정된다. G.722-첨부 III([ITU06a, 섹션 III.5])는 점진적으로 수행된 뮤팅(muting)을 제안하며, 이것은 프레임-손실의 20ms 이후에 시작하고, 프레임-손실의 60ms 이후에 완료된다. 더욱이, G.722-첨부 IV는 "샘플마다 계산되고 적용된 이득 인자를 각 샘플에" 적용하는 페이드-아웃 기술을 제안한다[ITU07, 섹션 IV.6.1.2.7].
- [0026] G.722에서, 뮤팅 프로세스는 QMF 합성 바로 직전에 그리고 PLC 모듈의 마지막 단계로서 서브밴드 도메인에서 발생한다. 뮤팅 인자의 계산은 또한 PLC 모듈의 부분인 신호 분류기로부터의 클래스 정보를 이용하여 수행된다. 구별은 클래스들(TRANSIENT, UV\_TRANSITION)과 다른 것들 사이에서 이루어진다. 더욱이, 구별은 10-ms 프레임들의 단일 손실들과 다른 경우들(10-ms 프레임들의 다중 손실들 및 20-ms 프레임들의 단일/다중 손실들) 사이에서 이루어진다.
- [0027] 이것은 도 3에 의해 예시된다. 특히, 도 3은, G.722의 페이드-아웃 인자가 클래스 정보에 의존하고 80개의 샘플들이 10ms와 동등한 시나리오를 도시한다.
- [0028] G.722에 따라, PLC 모듈은 손실 프레임에 대한 신호, 및 다음 양호한 프레임으로 교차-페이딩되도록 가정되는 몇몇 추가 신호(10ms)를 생성한다. 추가 신호에 대한 뮤팅은 동일한 규칙들에 따른다. G.722의 높은 밴드의 은닉에서, 교차-페이딩이 발생하지 않는다.
- [0029] 다음에서, G.722.1이 고려된다. 사이렌 7에 기초한 G.722.1은 초광역 대역 연장 모드를 갖는 광역 오디오 코덱에 기초한 변환이고, 이것은 G.722.1C로 언급된다. G.722.1C 자체는 사이렌 14에 기초한다. ITU-T는 G.722.1에 대해 후속 뮤팅을 갖는 프레임-반복을 권고한다[ITU05, 섹션 4.7]. 이러한 권고에 정의되지 않은 외부 신호 발신(signaling) 메커니즘에 의해, 프레임이 손실되었거나 붕괴되었다는 것을 디코더가 통보받으면, 이전 프레임의 디코딩된 MLT(Modulated Lapped Transform) 계수들을 반복한다. 이들 이전 프레임의 디코딩된 MLT 계수들을 시간 도메인으로 변환하고 이전 및 다음 프레임의 디코딩된 정보로 중첩 및 추가 동작을 수행함으로써 진행된다. 이전 프레임이 또한 손실되었거나 붕괴되었으면, 디코더는 모든 현재 프레임들의 MLT 계수들을 0으로 설정한다.
- [0030] 이제, G.729가 고려된다. G.729는 10 ms 지속기간의 패킷들에서 디지털 보이스를 압축하는 보이스를 위한 오디오 데이터 압축 알고리즘이다. 이것은 코드-여기된 선형 예측 음성 코딩(CS-ACELP)[ITU12]을 이용하여 8 kbit/s



에서 음성의 코딩으로서 공식적으로 기재된다.

[0031] [CPK08]에서 개설된 바와 같이, G.729는 LP 도메인에서 페이드-아웃을 권고한다. G.729 표준에 이용된 PLC 알고리즘은 이전에-수신된 음성 정보에 기초하여 현재 프레임에 대한 음성 신호를 재구성한다. 즉, PLC 알고리즘은 잃어버린 여기를 이전에 수신된 프레임의 등가 특성으로 대체하고, 여기 에너지가 마지막으로 점차 감소하더라도, 적응형 및 고정된 코드북들의 이득들은 상수 인자에 의해 감소된다.

[0032] 감소된 고정된-코드북 이득은  $g_c^{(m)} = 0.98 \cdot g_c^{(m-1)}$ 에 의해 주어지고,

[0033] m은 서브프레임 인덱스이다.

[0034] 적응성-코드북 이득은 이전의 적응성-코드북 이득의 감소된 버전에 기초한다:

[0035]  $g_p^{(m)} = 0.9 \cdot g_p^{(m-1)}$ ,  $g_p^{(m)} < 0.9$ 에 의해 한계가 정해짐

[0036] 남 인 박(Nam in Park) 등은 G.729에 대해 선형 회귀에 의한 예측을 이용하는 신호 진폭 제어를 제안한다 [CPK08, PJK+11]. 이것은 버스트 패킷 손실로 다루어지고, 핵심 기술로서 선형 회귀를 이용한다. 선형 회귀는 수학적 2로서 선형 모델에 기초한다

### 수학식 2

[0037] 
$$g_i' = a + bi$$

[0038] 여기서  $g_i'$ 는 새롭게 예측된 현재 진폭이고, a 및 b는 1차 선형 함수에 대한 계수들이고, i는 프레임의 인덱스이다. 최적화된 계수들( $a^*$  및  $b^*$ )을 찾기 위해, 제공 예측 에러의 합산은 최소화된다:

### 수학식 3

[0039] 
$$\varepsilon = \sum_{j=i-4}^{i-1} (g_j - g_j')^2$$

[0040]  $\varepsilon$ 는 제곱근 에러이고,  $g_j$ 는 원점을 지나는 j번째 진폭이다. 이러한 에러를 최소화하기 위해, 간단히 a 및 b에 관한 미분이 0으로 설정된다. 최적화된 파라미터들( $a^*$  및  $b^*$ )을 이용함으로써, 각  $g_i^*$ 의 추정치는 수학식 4로 표시된다

### 수학식 4

[0041] 
$$g_i^* = a^* + b^*i$$

[0042] 도 4는 진폭 예측, 특히 선형 회귀를 이용함으로써 진폭( $g_i^*$ )의 예측을 도시한다.

[0043] 손실된 패킷(i)의 진폭( $A_i'$ )을 얻기 위해, 비율( $\sigma_i$ )

수학식 5

$$\sigma_i = \frac{g_i^*}{g_i - 1}$$

[0044]

[0045] 는 스케일 인자(scale factor)( $S_i$ )와 곱해진다:

수학식 6

$$A_i' = S_i * \sigma_i$$

[0046]

[0047] 여기서 스케일 인자( $S_i$ )는 연속 은닉된 프레임들{ $l(i)$ }의 개수에 의존한다:

수학식 7

$$S_i = \begin{cases} 1.0, & \text{if } l(i) = 1,2, \\ 0.9, & \text{if } l(i) = 3,4, \\ 0.8, & \text{if } l(i) = 5,6, \\ 0, & \text{그 외의 경우} \end{cases}$$

[0048]

[0049] [PKJ+11]에서, 약간 상이한 스케일링이 제안된다.

[0050] G.729에 따라, 그 후에,  $A_i'$ 는 프레임 경계들에서 이산 감쇄를 방지하도록 평활화될 것이다. 마지막의 평활화된 진폭{ $A_i(n)$ }은 여기와 곱해지고, 이것은 이전의 PLC 성분들로부터 얻어진다.

[0051] 다음에서, G.729.1이 고려된다. G.729.1은 G.729-기반의 내장된 변수 비트율 코더이다: G.729와 상호 동작가능한 8-32 kbit/s 스케일가능 광역 코더 비트스트림[ITU06b].

[0052] G.729.1에 따라, G.718(위를 참조)에서와 같이, 적응형 페이드 아웃이 제안되며, 이것은 신호 특징들의 안정도에 의존한다([ITU06b, 섹션 7.6.1]). 은닉 동안, 신호는 일반적으로 감쇄 인자( $a$ )에 기초하여 감쇄되고, 이러한 감쇄 인자( $a$ )는 마지막으로 양호하게 수신된 프레임 클래스의 파라미터들 및 연속 소거된 프레임들의 개수에 의존한다. 감쇄 인자( $a$ )는 무성음 프레임들에 대해 LP 필터의 안정도에 추가로 의존한다. 일반적으로, 감쇄는 마지막으로 양호하게 수신된 프레임이 안정된 세그먼트인 경우 느려지고, 프레임이 전이 세그먼트인 경우 빨라진다.

[0053] 더욱이, 감쇄 인자( $a$ )는 서브프레임당 평균 피치 이득( $\overline{g_p}$ )에 의존한다([ITU06b, eq. 163. 164]):

수학식 8

$$\overline{g_p} = 0.1g_p^{(0)} + 0.2g_p^{(1)} + 0.3g_p^{(2)} + 0.4g_p^{(3)}$$

[0054]

[0055] 여기서  $g_p^{(i)}$ 는 서브프레임(i)에서의 피치 이득이다.

[0056] 표 2는  $a$ 의 계산 계획을 도시하며, 여기서

수학식 9

$0.85 \geq \beta \geq 0.98$ 인 상태에서  $\beta = \sqrt{g_p}$

은닉 프로세스 동안,  $\alpha$ 는 다음의 은닉 툴들(tools)에 사용된다:

표 2

감쇄 인자  $\alpha$ 의 값들, 값  $\theta$ 은 인접한 LP 필터들 사이의 거리 측정으로부터 계산된 안정도 인자이다. [ITU06b, 섹션 7.6.1].

마지막으로 양호하게 수신된 프레임	연속 소거된 프레임들의 개수	$\alpha$
유성음	1	$\beta$
	2,3	$\bar{g}_p$
	> 3	0.4
은셋	1	$0.8\beta$
	2,3	$\bar{g}_p$
	> 3	0.4
인공 은셋	1	$0.6\beta$
	2,3	$\bar{g}_p$
	> 3	0.4
유성음 전이	$\leq 2$	0.8
	> 2	0.2
무성음 전이		0.88
무성음	1	0.95
	2,3	$0.6\theta + 0.4$
	> 3	0.4

G.729.1에 따라, 성문 펄스 재동기화에 관해, 이전 프레임의 여기의 마지막 펄스가 주기 부분의 구성에 사용될 때, 그 이득은 은닉된 프레임의 시작에서 거의 정확하고, 1로 설정될 수 있다. 이득은 이 후 프레임의 마지막에서  $\alpha$ 의 값을 달성하기 위해 샘플간에 기초하여 프레임 전체에 선형으로 감쇄된다. 유성음 세그먼트들의 에너지 전개(energy evolution)는 마지막으로 양호한 프레임의 각 서브프레임의 피치 여기 이득 값들을 이용함으로써 외삽된다. 일반적으로, 이들 이득들이 1보다 크면, 신호 에너지는 증가하고, 이들 이득들이 1보다 작으면, 에너지는 감소한다.  $\alpha$ 는 이에 따라 전술한 바와 같이  $\beta = \sqrt{g_p}$ 로 설정되고, [ITU06b, eq. 163, 164]를 참조하자.  $\beta$ 의 값은 강한 에너지 증가 및 감소를 피하기 위해 0.98과 0.85 사이에서 클리핑(clipped)되고, [ITU06b, 섹션 7.6.4]를 참조하자.

G.729.1에 따라, 여기의 랜덤 부분의 구성에 관해, 소거된 블록의 시작에서, 혁신 이득( $g_s$ )은 마지막으로 양호한 프레임의 각 서브프레임의 혁신 여기 이득들을 이용함으로써 초기화된다:

$\bar{g}_s = 0.1g^{(0)} + 0.2g^{(1)} + 0.3g^{(2)} + 0.4g^{(3)}$

여기서  $g^{(0)}$ ,  $g^{(1)}$ ,  $g^{(2)}$  및  $g^{(3)}$ 은 마지막으로 정확히 수신된 프레임의 4개의 서브프레임들의 고정된 코드북, 또는 혁신, 이득들이다. 혁신 이득 감쇄는 다음과 같이 이루어진다:

$g_s^{(1)} = \alpha \cdot g_s^{(0)}$

여기서,  $g_s^{(1)}$ 은 다음 프레임의 시작에서의 혁신 이득이고,  $g_s^{(0)}$ 은 현재 프레임의 시작에서의 혁신 이득이고,  $\alpha$ 는 위의 표 2에 정의된 바와 같다. 주기 여기 감쇄와 유사하게, 이득은 이에 따라  $g_s^{(0)}$ 에서 시작하고  $g_s^{(1)}$ 의 값

으로 가는 샘플간에 기초하여 프레임 전체에서 선형으로 감쇄되고, 이것은 다음 프레임의 시작에서 달성된다.

[0066] G.279.1에 따라, 마지막으로 양호한 프레임이 무성음이면, 혁신 여기만이 사용되고, 0.8의 인자만큼 추가로 감쇄된다. 이 경우에, 과거 여기 버퍼는, 여기의 주기 부분이 이용가능하지 않기 때문에 혁신 여기로 업데이트되고, [ITU06b, 섹션 7.6.5]를 참조하자.

[0067] 다음에서, AMR이 고려된다. 3GPP AMR[3GP12b]는 ACELP 알고리즘을 이용하는 음성 코덱이다. AMR은 8000 샘플들/s의 샘플링을 및 4.75와 12.2 kbit/s 사이의 비트율로 음성을 코딩할 수 있고, 신호 발신 무음(silence) 설명자 프레임들(DTX/CNG)을 지원한다.

[0068] AMR에서, 에러 은닉([3GP12a]를 참조) 동안, 에러 유발(error prone)(비트 에러들)이 있는 프레임들과 완전히 손실되는(데이터가 전혀 없음) 프레임들 사이가 구별된다.

[0069] ACELP 은닉에 대해, AMR은 채널의 품질을 추정하는 상태 기계를 도입한다: 상태 카운터의 값이 더 커질수록, 채널 품질은 더 악화된다. 시스템은 상태 0에서 시작한다. 불량 프레임이 검출될 때마다, 상태 카운터는 1만큼 증분되고, 6에 도달할 때 포화된다. 양호한 음성 프레임이 검출될 때마다, 상태 카운터는, 상태가 6이어서 상태 카운터가 5로 설정된 경우를 제외하고, 0으로 재설정된다. 상태 기계의 제어 흐름은 다음의 C 코드(BFI는 불량 프레임 표시자이고, 상태는 상태 변수이다)에 의해 기재될 수 있다:

```

if(BFI != 0 ) {
    State = State + 1;
}
else if(State == 6) {
    State = 5;
}
else {
    State = 0;
}
if(State > 6 ) {
    State = 6;
}

```

[0070] 이러한 상태 기계 외에도, AMR에서, 현재 및 이전 프레임들로부터의 불량 프레임 플래그들이 체크된다 (prevBFI).

[0072] 3가지 상이한 조합들이 가능하다:

[0073] 3가지 조합들 중 첫 번째 조합은 BFI=0, prevBFI=0, State=0이다: 수신된 또는 이전에 수신된 음성 프레임에서는 에러가 검출되지 않는다. 수신된 음성 파라미터들은 음성 합성에서 정상적인 방식으로 사용된다. 음성 파라미터들의 현재 프레임은 저장된다.

[0074] 3가지 조합들 중 두 번째 조합은 BFI=0, prevBFI=1, State=0 또는 5이다: 수신된 음성 프레임에서는 에러가 검출되지 않고, 이전에 수신된 음성 프레임은 불량이다. LTP 이득 및 고정된 코드북 이득은 마지막으로 수신된 양호한 서브프레임에 사용된 값들 아래에 제한된다:

수학식 10

$$g_p = \begin{cases} g_p, & g_p \leq g_p(-1) \\ g_p(-1), & g_p > g_p(-1) \end{cases}$$

[0075]

[0076] 여기서,  $g_p$  = 현재 디코딩된 LTP 이득이고,  $g_p(-1)$  = 마지막으로 양호한 서브프레임(BFI=0)에 사용된 LTP 이득이고,

수학식 11

$$g_c = \begin{cases} g_c, & g_c \leq g_c(-1) \\ g_c(-1), & g_c > g_c(-1) \end{cases}$$

[0077]

[0078] 여기서,  $g_c$  = 현재 디코딩된 고정된 코드북 이득이고,  $g_c(-1)$  = 마지막으로 양호한 서브프레임(BFI=0)에 사용된 고정된 코드북 이득이다.

[0079] 수신된 음성 파라미터들의 나머지는 음성 합성에서 통상적으로 사용된다. 음성 파라미터들의 현재 프레임은 저장된다.

[0080] 3가지 조합들 중 세 번째 조합은 BFI=1, prevBFI= 0 또는 1, State = 1...6:이다. 수신된 음성 프레임에서는 에러가 검출되고, 교체 및 뮤팅 절차가 시작된다. LTP 이득 및 고정된 코드북 이득은 이전 서브프레임들로부터 감쇄된 값들로 교체된다:

수학식 12

$$g_p = \begin{cases} P(state) \cdot g_p(-1), & g_p(-1) \leq \text{median5}(g_p(-1), \dots, g_p(-5)) \\ P(state) \cdot \text{median5}(g_p(-1), \dots, g_p(-5)) & g_p(-1) > \text{median5}(g_p(-1), \dots, g_p(-5)) \end{cases}$$

[0081]

[0082] 여기서,  $g_p$ 는 현재 디코딩된 LTP 이득을 나타내고,  $g_p(-1), \dots, g_p(-n)$ 은 마지막  $n$ 개의 서브프레임들에 사용된 LTP 이득들을 나타내고,  $\text{median5}()$ 는 5-포인트 중간값 연산(5-point median operation)을 나타내고,

[0083]  $P(state)$  = 감쇄 인자,

[0084] 여기서 ( $P(1)=0.98, P(2)=0.98, P(3)=0.8, P(4)=0.3, P(5)=0.2, P(6)=0.2$ )이고,  $state$ =상태 번호이고,

수학식 13

$$g_c = \begin{cases} C(state) \cdot g_c(-1), & g_c(-1) \leq \text{median5}(g_c(-1), \dots, g_c(-5)) \\ C(state) \cdot \text{median5}(g_c(-1), \dots, g_c(-5)) & g_c(-1) > \text{median5}(g_c(-1), \dots, g_c(-5)) \end{cases}$$

[0085]

[0086] 여기서  $g_c$ 는 현재 디코딩된 고정된 코드북 이득을 나타내고,  $g_c(-1), \dots, g_c(-n)$ 은 마지막  $n$ 개의 서브프레임들에 사용된 고정된 코드북 이득들을 나타내고,  $\text{median5}()$ 는 5-포인트 중간값 연산을 나타내고,  $C(state)$ =감쇄 인자이고, 여기서 ( $C(1)=0.98, C(2)=0.98, C(3)=0.98, C(4)=0.98, C(5)=0.98, C(6)=0.7$ )이고,  $state$ =상태 번호이다.

[0087] AMR에서, LTP-래그 값들(LTP=Long-Term Prediction)은 이전 프레임의 제 4 서브프레임으로부터 과거 값으로 교체되거나(12.2 모드), 마지막으로 정확히 수신된 값에 기초하여 약간 변형된 값들로 교체된다(모든 다른 모드들).

- [0088] AMR에 따라, 잘못된 프레임으로부터의 수신된 고정된 코드북 혁신 펄스들은, 붕괴된 데이터가 수신될 때 수신된 상태에서 사용된다. 어떠한 데이터도 수신되지 않은 경우에, 랜덤 고정된 코드북 인덱스들이 사용되어야 한다.
- [0089] AMR에서 CNG에 관해, [3GP12a, 섹션 6.4]에 따라, 각각의 제 1 손실된 SID 프레임은 더 일찍 수신된 유효 SID 프레임들로부터의 SID 정보를 이용함으로써 교체되고, 유효 SID 프레임들에 대한 절차가 적용된다. 후속 손실된 SID 프레임들에 대해, 감쇄 기술은 출력 레벨을 점차 감소시킬 안락 잡음에 적용된다. 그러므로, 마지막 SID 업데이트가 이전에 50을 초과하는 프레임들(=1s)이었던 지가 체크되고, 만약 그렇다면, 출력은 뮤팅될 것이다(프레임당 -6/8 dB만큼의 레벨 감쇄 [3GP12d,  $dtx\_dec\{\ }@sp\_dec.c$ ], 이것은 초당 37.5 dB를 산출한다). CNG에 적용된 페이드-아웃이 LP 도메인에서 수행되는 것이 주지된다.
- [0090] 다음에서, AMR-WB가 고려된다. 적응형 멀티레이트 - WB [ITU03, 3GP09c]는 AMR에 기초한 음성 코덱, ACELP이다 (섹션 1.8을 참조). 이것은 파라미터성 대역폭 확장을 이용하고, 또한 DTX/CNG를 지원한다. 표준 [3GP12g]의 설명에서, 사소한 편차들을 갖는 AMR [3GP12a]와 동일한 은닉 예의 해법들이 주어진다. 그러므로, AMR에 대한 차이들만이 본 명세서에 기재된다. 표준 설명에 대해, 위의 설명을 참조하자.
- [0091] ACELP에 관해, AMR-WB에서, ACELP 페이드-아웃은 피치 이득( $g_p$ )(위의 AMR에 대해 LTP 이득으로 언급됨)을 변형함으로써 그리고 코드 이득( $g_c$ )을 변형함으로써 기준 소스 코드 [3GP12c]에 기초하여 수행된다.
- [0092] 손실된 프레임의 경우에, 제 1 서브프레임에 대한 피치 이득( $g_p$ )은 0.95와 0.5 사이에 제한된다는 점을 제외하고 마지막으로 양호한 프레임에서와 동일하다. 제 2, 제 3 및 후속하는 서브프레임들에 대해, 피치 이득( $g_p$ )은 0.95의 인자만큼 감소되고, 다시 제한된다.
- [0093] AMR-WB는, 은닉된 프레임( $g_c$ )에서, 마지막  $g_c$ 에 기초한다는 것을 제안한다:

**수학식 14**

[0094] 
$$g_{c,current} = g_{c,past} * (1.4 - g_{p,past})$$

**수학식 15**

[0095] 
$$g_c = g_{c,current} * g_{c,mov}$$

**수학식 16**

[0096] 
$$g_{c,mov} = \frac{1.0}{\sqrt{\frac{ener_{inov}}{subframe\_size}}}$$

**수학식 17**

[0097] 
$$ener_{inov} = \sum_{i=0}^{subframe\_size-1} code[i]$$

[0098] LTP-래그들을 은닉하기 위해, AMR-WB에서, 5개의 마지막으로 양호한 LTP-래그들 및 LTP-이득들의 이력은 프레임 손실의 경우에 업데이트하기 위한 최상의 방법을 찾는데 사용된다. 프레임이 비트 에러들을 가지고 수신되는 경우에, 수신된 LTP 래그가 사용가능하거나 사용가능하지 않는지의 여부에 상관없이 수행된다[3GP12g].

[0099] CNG에 관해, AMR-WB에서, 마지막으로 정확히 수신된 프레임이 SID 프레임이었고 프레임이 손실된 것으로 분류되

면, 마지막 유효 SID 프레임 정보로 교체될 수 있고, 유효 SID 프레임들에 대한 절차가 적용되어야 한다.

- [0100] 후속 손실된 SID 프레임들에 대해, AMR-WB는 출력 레벨을 점차 감소시킬 안락 잡음에 감쇄 기술을 적용하는 것을 제안한다. 그러므로, 마지막 SID 업데이트가 이전에 50을 초과하는 프레임들(=1s)이었던 지가 체크되고, 만약 그렇다면, 출력은 뮤팅될 것이다(프레임당 -3/8 dB만큼의 레벨 감쇄 [3GP12f, dtx\_dec{ }@sp\_dtx.c], 이것은 초당 18.75 dB를 산출한다}. CNG에 적용된 페이드-아웃이 LP 도메인에서 수행되는 것이 주지된다.
- [0101] 이제, AMR-WB+가 고려된다. 적응형 멀티레이트 - WB+ [3GP09a]는 핵심 코덱들로서 ACELP 및 TCX(TCX=Transform Coded Excitation)을 이용하여 스위칭된 코덱이다. 이것은 파라미터성 대역폭 확장을 이용하고, 또한 DTX/CNG를 지원한다.
- [0102] AMR-WB+에서, 모드 외삽 논리는 왜곡된 서브프레임 내에서 손실된 프레임들의 모드들을 외삽하도록 적용된다. 이러한 모드 외삽은 모드 표시자들의 정의에서 여분이 존재한다는 점에 기초한다. AMR-WB+에 의해 제안된 결정 논리([3GP09a, 도 18]에 주어짐)는 다음과 같다:
- [0103] - 벡터 모드,  $(m_{-1}, m_0, m_1, m_2, m_3)$ 가 정의되고, 여기서  $m_{-1}$ 은 이전 슈퍼프레임의 마지막 프레임의 모드를 나타내고,  $m_0, m_1, m_2, m_3$ 은 현재 슈퍼프레임(비트스트림으로부터 디코딩된)에서의 프레임들의 모드들을 나타내고, 여기서  $m_k = -1, 0, 1, 2$  또는  $3(-1$ : 손실,  $0$ : ACELP,  $1$ : TCX20,  $2$ : TXC40,  $3$ : TCX80)이고, 손실된 프레임들의 개수(nloss)는 0 내지 4이다.
- [0104] -  $m_{-1} = 3$ 이고, 프레임들 0 내지 3의 모드 표시자들 중 2개가 3이면, 모든 표시자들은 3으로 설정될 것인데, 이는 하나의 TCX80 프레임이 슈퍼프레임 내에서 표시되었다는 것이 확실하기 때문이다.
- [0105] - 프레임들 0 내지 3의 하나의 표시자만이 3(및 손실된 프레임들의 개수(nloss)가 3이면)이면, 모드는 (1,1,1,1)로 설정될 것인데, 이는 TCX80 목표 스펙트럼의 3/4가 손실되고 범용 TCX 이득이 손실될 가능성이 매우 있기 때문이다.
- [0106] - 모드가  $(x, 2, -1, x, x)$  또는  $(x, -1, 2, x, x)$ 를 나타내면,  $(x, 2, 2, x, x)$ 로 외삽될 것이고, 이것은 TCX40 프레임을 나타낸다. 모드가  $(x, x, x, 2, -1)$  또는  $(x, x, -1, 2)$ 이면,  $(x, x, x, 2, 2)$ 로 외삽될 것이고, 이것은 또한 TCX40 프레임을 나타낸다.  $(x, [0, 1], 2, 2, [0, 1])$ 이 무효 구성들이라는 것이 주지되어야 한다.
- [0107] - 그 후에, 손실된(mode = -1) 각 프레임에 대해, 모드는 이전 프레임이 ACELP인 경우 ACELP(mode=0)으로 설정되고, 모드는 모든 다른 경우들에 대해 TCX20(mode=1)로 설정된다.
- [0108] ACELP에 관해, AMR-WB+에 따라, 손실된 프레임들 모드가 모드 외삽 이후에  $m_k=0$ 을 초래하면, [3GP12g]에서와 동일한 접근법은 이러한 프레임에 대해 적용된다(위를 참조).
- [0109] AMR-WB+에서, 손실된 프레임들의 개수 및 외삽된 모드에 따라, 다음의 TCX 관련 은닉 접근법들이 구별된다(TCX=Transform Coded Excitation):
- [0110] - 전체 프레임이 손실되면, ACELP형 은닉이 적용된다: 마지막 여기가 반복되고, 은닉된 ISF 계수들(적응성 평균 쪽으로 약간 시프트함)은 시간 도메인 신호를 합성하는데 사용된다. 추가적으로, 프레임당 0.7의 페이드-아웃 인자(20ms)[3GP09b, dec\_tcx.c]는 LPC(Linear Predictive Coding) 합성 바로 직전에 선형 예측 도메인에서 곱해진다.
- [0111] - 마지막 모드가 TCX80이고 (부분적으로 손실된) 슈퍼프레임의 외삽된 모드가 TCX80(nloss=[1,2], mode=(3,3,3,3,3))이면, 은닉은 위상 및 진폭 외삽을 이용하여, 마지막으로 정확히 수신된 프레임을 고려하여, FFT 도메인에서 수행된다. 위상 정보의 외삽 접근법은 본 명세서에서 어떠한 관심도 없고(페이딩 전략과 관련 없음), 그러므로 기재되지 않는다. 추가 세부사항들에 대해, [3GP09a, 섹션 6.5.1.2.4]를 참조하자. AMR-WB+의 진폭 변형에 대해, TCX 은닉을 위해 수행된 접근법은 다음의 단계들로 구성된다[3GP09a, 섹션 6.5.1.2.3]:
- [0112] - 이전 프레임 크기 스펙트럼이 계산된다:

[0113] 
$$oldA[k] = |old\hat{X}[k]|$$

[0114] - 현재 프레임 크기 스펙트럼이 계산된다:

$$A[k] = |\hat{X}[k]|$$

[0115]

[0116] - 이전 및 현재 프레임 사이의 비-손실 스펙트럼 계수들의 에너지의 이득 차이가 계산된다:

$$\text{이득} = \sqrt{\frac{\sum A[k]^2}{\sum \text{old}A[k]^2}}$$

[0117]

[0118] - 잃어버린 스펙트럼 계수들의 진폭은 다음을 사용하여 외삽된다:

$$\text{if}(\text{lost}[k]) \quad A[k] = \text{gain} \cdot \text{old}A[k]$$

[0119]

[0120] -  $m_k = [2,3]$ 을 갖는 손실된 프레임의 모든 다른 경우에서, TCX 목표{디코딩된 스펙트럼 플러스 잡음 충전(fill-in)의 역 FFT(비트스트림으로부터 디코딩된 잡음 레벨을 이용하여)}는 모든 이용가능한 정보(범용 TCX 이득을 포함)를 이용하여 합성된다. 이 경우에 페이드-아웃은 적용되지 않는다.

[0121]

AMR-WB+에서 CNG에 관해, AMR-WB와 동일한 접근법이 사용된다(위를 참조).

[0122]

다음에서, OPUS가 고려된다. OPUS [IET12]는 2가지 코덱들로부터 기술을 병합한다: 음성-지향 SILK(스카이프 코덱으로서 알려짐) 및 저-레이턴시(low-latency) CELT(CELT= Constrained-Energy Lapped Transform). Opus는 높은 및 낮은 비트율들 사이에서 끊김 없이 조정될 수 있고, 내부적으로, 낮은 비트율(SILK)에서의 선형 예측 코덱과 높은 비트율(CELT)에서의 변환 코덱뿐 아니라 짧은 중첩을 위한 하이브리드 사이를 스위칭한다.

[0123]

SILK 오디오 데이터 압축 및 압축 해제에 관해, OPUS에서, SILK 디코더 루틴에서의 은닉 동안 감쇄되는 여러 개의 파라미터들이 존재한다. 연속 손실된 프레임들의 개수에 따라, LTP 이득 파라미터는 프레임당 0.99, 0.95 또는 0.90과 모든 LPC 계수들을 곱함으로써 감쇄되고, 여기서 여기는 이전 프레임의 여기로부터 마지막 피치 사이클을 이용하여 구축된다. 피치 래그 파라미터는 연속 손실들 동안 매우 느리게 증가한다. 단일 손실들에 대해,

마지막 프레임에 비해 일정하게 유지된다. 더욱이, 여기 이득 파라미터는 프레임당  $0.99^{\text{lost}_{cnt}}$  으로 지수적으로 감쇄되어, 여기 이득 파라미터는 제 1 여기 이득 파라미터에 대해 0.99이고, 여기 이득 파라미터는 제 2 여기 이득 파라미터에 대해 0.992이고, 이후에도 이와 같이 이루어진다(and so on). 여기는 가변 오버플로우(overflow)에 의해 백색 잡음을 생성하는 난수(random number) 생성기를 이용하여 생성된다. 더욱이, LPC 계수들은 계수들의 마지막으로 정확히 수신된 세트에 기초하여 외삽/평균화된다. 감쇄된 여기 벡터를 생성한 후에, 은닉된 LPC 계수들은 시간 도메인 출력 신호를 합성하기 위해 OPUS에 사용된다.

[0124]

이제, OPUS의 정황에서, CELT가 고려된다. CELT는 변환 기반의 코덱이다. CELT의 은닉은 피치 기반의 PLC 접근법을 특징으로 하는데, 이러한 피치 기반의 PLC 접근법은 최대 5개의 연속적으로 손실된 프레임들에 대해 적용된다. 프레임 6에서 시작하여, 잡음형 은닉 접근법이 적용되고, 이것은 배경 잡음을 생성하고, 이러한 특징은 이전 배경 잡음과 같은 사운드를 받기로 되어 있다.

[0125]

도 5는 CELT의 버스트 손실 행위를 도시한다. 특히, 도 5는 CEL 은닉된 음성 세그먼트의 스펙트로그램(x-축: 시간; y-축: 주파수)을 도시한다. 열은 회색 박스는 처음 5개의 연속적으로 손실된 프레임들을 나타내고, 여기서 피치 기반의 PLC 접근법이 적용된다. 이 외에, 잡음형 은닉이 도시된다. 스위칭이 즉시 수행되는 것이 주지되어야 하고, 매끄럽게 전이하지 않는다.

[0126]

피치 기반의 은닉에 관해, OPUS에서, 피치 기반의 은닉은 자동 상관(autocorrelation)에 의해 디코딩된 신호에서 주기성을 발견하는 것과, 피치 오프셋(피치 래그)을 이용하여 윈도우링된 파형(LPC 분석 및 합성을 이용하여 여기 도메인에서)을 반복하는 것으로 구성된다. 윈도우링된 파형은 이전 프레임 및 다음 프레임과의 시간-도메인 얼라이싱(aliasing) 최소를 확보하는 방식으로 중첩된다[IET12]. 추가적으로, 페이드-아웃 인자는 다음의 코드에 의해 도출되고 적용된다:

[0127]

opus\_val32 E1=1, E2=1;

[0128]

int period;



```

[0129]   if (pitch_index <= MAX_PERIOD/2) {
[0130]       period = pitch_index;
[0131]   }
[0132]   else {
[0133]       period = MAX_PERIOD/2;
[0134]   }
[0135]   for (i=0;i<period;i++)
[0136]   {
[0137]       E1 += exc[MAX_PERIOD- period+i] * exc[MAX_PERIOD- period+i];
[0138]       E2 += exc[MAX_PERIOD-2*period+i] * exc[MAX_PERIOD-2*period+i];
[0139]   }
[0140]   if (E1 > E2) {
[0141]       E1 = E2;
[0142]   }
[0143]   decay = sqrt(E1/E2));
[0144]   attenuation = decay;

```

[0145] 이 코드에서, exc는 손실 이전에 최대 MAX\_PERIOD 샘플들까지 여기 신호를 포함한다.

[0146] 여기 신호는 나중에 감쇄와 곱해지고, 그 후에 LPC 합성을 통해 합성되고 출력된다.

[0147] 시간 도메인 접근법에 대한 페이딩 알고리즘은 이와 같이 요약될 수 있다:

- [0148] - 손실 이전에 마지막 피치 사이클의 피치 동기화 에너지를 발견.
- [0149] - 손실 이전에 제 2 마지막 피치 사이클의 피치 동기화 에너지를 발견.
- [0150] - 에너지가 증가하면, 일정하게 유지하기 위해 에너지를 제한: 감쇄 = 1
- [0151] - 에너지가 감소하면, 은닉 동안 동일한 감쇄를 계속

[0152] 잡음형 은닉에 관해, OPUS에 따라, 제 6 및 후속하는 연속적인 손실된 프레임들에 대해, MDCT 도메인에서의 잡음 교체 접근법은 안락 배경 잡음을 시플레이팅하기 위해 수행된다.

[0153] 배경 잡음 레벨 및 형태의 추적에 관해, OPUS에서, 배경 잡음 추정은 다음과 같이 수행된다: MDCT 분석 이후에, MDCT 대역 에너지들의 제공근은 주파수 대역마다 계산되고, 여기서 MDCT 빈들(bins)의 그룹화는 [IET12, 표 55]에 따라 바크 스케일(bark scale)에 뒤따른다. 그런 후에 에너지들의 제공근은 수학적 18에 의해  $\log_2$  도메인으로 변환된다:

### 수학적 18

[0154]  $i = 0 \dots 21$ 에 대해  $bandLogE[i] = \log_2(e) \cdot \log_e(bandE[i] - eMeans[i])$

[0155] 여기서  $e$ 는 오일러의 수이고,  $bandE$ 는 MDCT 대역의 제공근이고,  $eMeans$ 는 상수들의 벡터(결과적인 0의 평균을 얻는데 필요함, 이것은 개선된 코딩 이득을 초래한다)이다.

[0156] OPUS에서, 배경 잡음은 이와 같은 디코더 측 상에서 로깅(logged)된다[IET12, amp2Log2 및 log2Amp@quant\_bands.c]:

수학식 19

$$i = 0 \dots 21 \text{에 대해 } backgroundLogE[i] = \min(backgroundLogE[i] + 8 \cdot 0.001, bandLogE[i])$$

[0157]

[0158] 추적된 최소 에너지는 기본적으로 현재 프레임의 대역의 에너지의 제공근에 의해 결정되지만, 하나의 프레임으로부터 다음 프레임까지의 증가는 0.05 dB만큼 제한된다.

[0159] 배경 잡음 레벨 및 형태의 적용에 관해, OPUS에 따라, 잡음형 PLC가 적용되면, 마지막의 양호한 프레임에서 도출된 *backgroundLogE*가 사용되고, 선형 도메인으로 다시 변환된다:

수학식 20

$$i = 0 \dots 21 \text{에 대해 } bandE[i] = e^{(\log_e(2) \cdot (backgroundLogE[i] + eMeans[i]))}$$

[0160]

[0161] 여기서 *e*는 오일러의 수이고, *eMeans*는 "선형-로그" 변환에 대한 것과 동일한 상수들의 벡터이다.

[0162] 현재 은닉 절차는 난수 생성기에 의해 발생된 백색 잡음으로 MDCT 프레임을 채우고, 대역 방식(wise)을 *bandE*의 에너지에 매칭하는 방식으로 이러한 백색 잡음을 스케일링하는 것이다. 순차적으로, 역 MDCT가 적용되고, 이것은 시간 도메인 신호를 초래한다. 중첩 추가 및 디엠퍼시스(정상적인 디코딩에서와 같이) 이후에, 풋 아웃(put out)된다.

[0163] 다음에서, MPEG-4 HE-AAC가 고려된다(MPEG=Moving Picture Experts Group; HE-AAC=High Efficiency Advanced Audio Coding). 고효율 진보된 오디오 코딩은 변환 기반의 오디오 코덱(AAC)으로 구성되고, 이것은 파라미터성 대역폭 확장(SBR)에 의해 보충된다.

[0164] AAC(AAC=Advanced Audio Coding)에 관해, DAB 협회는 DAB+에서의 AAC에 대해 주파수 도메인에서의 0으로의 페이드-아웃을 규정한다[EBU10, 섹션 A1.2](DAB=Digital Audio Broadcasting). 페이드-아웃 행위, 예를 들어, 감쇄 램프는 사용자에 의해 고정되거나 조정가능할 수 있다. 마지막 AU(AU=Access Unit)로부터의 스펙트럼 계수들은 페이드-아웃 특징들에 대응하는 인자에 의해 감쇄되고, 그런 후에 주파수-시간 매핑으로 전달된다. 감쇄 램프에 따라, 은닉은 연속적인 무효 AU들의 개수 이후에 뮤팅으로 스위칭하고, 이것은 완전한 스펙트럼이 0으로 설정될 것이라는 것을 의미한다.

[0165] DRM(DRM=Digital Rights Management) 협회는 DRM에서의 AAC에 대해 주파수 도메인에서의 페이드-아웃을 규정한다[EBU12, 섹션 5.3.3]. 은닉은 최종 주파수-시간 변환 바로 직전에 스펙트럼 데이터 상에서 작용한다. 다중 프레임들이 붕괴되면, 은닉은 먼저 마지막 유효 프레임으로부터 약간 변형된 스펙트럼 값들에 기초하여 페이드 아웃을 구현한다. 더욱이, DAB+와 유사하게, 페이드-아웃 행위, 예를 들어 감쇄 램프는 사용자에 의해 고정되거나 조정가능할 수 있다. 마지막 프레임으로부터의 스펙트럼 계수들은 페이드-아웃 특징들에 대응하는 인자에 의해 감쇄되고, 그런 후에 주파수-시간 매핑에 전달된다. 감쇄 램프에 따라, 은닉은 연속적인 무효 프레임들의 개수 이후에 뮤팅으로 스위칭하고, 이것은 완전한 스펙트럼이 0으로 설정될 것이라는 것을 의미한다.

[0166] 3GPP는 개선된 aacPlus에서의 AAC에 대해 DRM과 유사한 주파수 도메인에서의 페이드-아웃을 소개한다[3GP12e, 섹션 5.1]. 은닉은 최종 주파수-시간 변환 바로 직전에 스펙트럼 데이터 상에서 작용한다. 다중 프레임들이 붕괴되면, 은닉은 마지막 양호한 프레임으로부터 약간 변형된 스펙트럼 값들에 기초하여 페이드 아웃을 구현한다. 완전한 페이딩 아웃은 5개의 프레임들을 취한다. 마지막 양호한 프레임으로부터의 스펙트럼 계수들은 복제되고, 마지막 양호한 프레임 이래로 프레임 카운터로서 *nFadeOutFrame*을 갖는  $fadeOutFac = 2^{-(nFadeOutFrame/2)}$ 의 인자에 의해 감쇄된다. 페이딩 아웃의 4개의 프레임들 이후에, 은닉은 뮤팅으로 스위칭하고, 이것은 완전한 스펙트럼이 0으로 설정될 것이라는 것을 의미한다.

[0167] 라우버(Lauber) 및 스페르슈나이더(Sperschneider)는 에너지 외삽에 기초하여 MDCT 스펙트럼의 프레임-방식의 페이드-아웃을 소개한다[LS01, 섹션 4.4]. 이전 스펙트럼의 에너지 형태들은 추정된 스펙트럼의 형태를 외삽하는데 사용될 수 있다. 에너지 외삽은 일종의 후치 은닉으로서 은닉 기술들과 독립적으로 수행될 수 있다.

[0168] AAC에 관해, 에너지 계산은 인간 청각 시스템의 중요한 대역들에 가깝게 하기 위해 스케일 인자 대역 기반으로

수행된다. 개별적인 에너지 값들은 부피를 매끄럽게 감소하기 위해, 예를 들어 신호를 페이드 아웃하기 위해 프레임간에 기초하여 감소된다. 이것은, 추정된 값들이 현재 신호를 나타낼 확률이 시간이 지남에 따라 급격히 감소하기 때문에 필요하다.

- [0169] 공급될 스펙트럼의 생성을 위해, 프레임 반복 또는 잡음 교체를 제안한다[LS01, 섹션들 3.2 및 3.3].
- [0170] 퀴켄부쉬(Quackenbusch) 및 드리센(Driesen)은 ACC에 대해 0으로의 지수적 프레임-방식의 페이드-아웃을 제안한다[QD03]. 시간/주파수 계수들의 인접 세트의 반복이 제안되며, 여기서 각 반복은 지수적으로 증가하는 감쇄를 가져, 연장된 정지(outages)의 경우에 뮤팅하기 위해 점차 페이딩한다.
- [0171] MPEG-4 HE-AAC에서의 SBR(SBR=Spectral Band Replication)에 관해, 3GPP는 개선된 aacPlus에서의 SBR에 대해 디코딩된 엔벨로프 데이터를 버퍼링하는 것과, 프레임 손실의 경우에, 전송된 엔벨로프 데이터의 버퍼링된 에너지들을 재사용하는 것과, 은닉된 프레임마다 3 dB의 상수 비만큼 버퍼링된 에너지들을 감소시키는 것을 제안한다. 그 결과는, 엔벨로프 조정기가 HF 생성기에 의해 생성된 패칭된(patchd) 고대역들을 조정하는데 사용된 이득들을 계산하기 위해 이를 이용하는 정상 디코딩 프로세스에 공급된다. SBR 디코딩은 이 후 정상처럼 발생한다. 더욱이, 텔타 코딩된 잡음 플로어 및 사인 레벨 값들이 삭제된다. 이전 정보에 대한 차이가 이용가능하게 남아있지 않기 때문에, 디코딩된 잡음 플로어 및 사인 레벨들은 HF 생성된 신호의 에너지에 비례하여 남아있다 [3GP12e, 섹션 5.2].
- [0172] DRM 협회는 AAC와 연계하여 SBR에 대해 3GPP와 동일한 기술을 규정하였다[EBU12, 섹션 5.6.3.1]. 더욱이, DAB 협회는 DAB+에서의 SBR에 대해 3GPP와 동일한 기술을 규정한다[EBU10, 섹션 A2].
- [0173] 다음에서, MPEG-4 CELP 및 MPEG-4 HVXC(HVXC=Harmonic Vector Excitation Coding)가 고려된다. DRM 협회는 CELP 및 HVXC와 연계하여 SBR에 대해[EBU12, 섹션 5.6.3.2] 음성 코덱들을 위한 SBR에 대한 최소 요건 은닉이, 붕괴된 SBR 프레임이 삭제될 때마다 데이터 값들의 미리 결정된 세트를 적용하기 위한 것임을 규정한다. 이들 값들은 낮은 상대적인 재생 레벨에서 정적의 고대역 스펙트럼 엔벨로프를 산출하고, 이것은 더 높은 주파수쪽으로의 롤-오프(roll-off)를 나타낸다. 그 목적은 간단히 "안락 잡음"(엄격한 뮤팅에 대조적으로)을 삽입함으로써, 잘못된-행위의, 잠재적으로 시끄러운, 오디오 버스트들이 청취자의 귀에 도달하지 않는 것을 보장하는 것이다. 이것은 사실상 실제 페이드-아웃이 이루어지지 않고, 오히려 몇몇 종류의 안락 잡음을 삽입하기 위해 특정 에너지 레벨로의 도약이 이루어진다.
- [0174] 후속하여, 대안이 언급되고[EBU12, 섹션 5.6.3.2], 이 대안은 마지막으로 정확히 디코딩된 데이터를 재사용하고, AAC + SBR 경우와 유사하게 레벨들(L)을 0쪽으로 느리게 페이딩한다.
- [0175] 이제, MPEG-4 HILN이 고려된다(HILN = Harmonic and Individual Lines plus Noise). 마인(Meine) 등은 파라미터성 도메인[MEP01]에서 파라미터성 MPEG-4 HILN 코덱[ISO09]을 위한 페이드-아웃을 소개한다. 지속적인 고조파 성분들에 대해, 붕괴된 차동적으로 인코딩된 파라미터들을 교체하기 위한 양호한 디폴트 행위는 주파수를 일정하게 유지하는 것이고, 감쇄 인자(예를 들어, -6 dB)만큼 진폭을 감소시키는 것이고, 스펙트럼 엔벨로프가 평균화된 저역 통과 특징의 것쪽으로 수렴하도록 하는 것이다. 스펙트럼 엔벨로프에 대한 대안은 스펙트럼 엔벨로프를 변하지 않은 상태로 유지하는 것이다. 진폭들 및 스펙트럼 엔벨로프들에 대해, 잡음 성분들은 고조파 성분들과 동일한 방식으로 처리될 수 있다.
- [0176] 다음에서, 종래 기술에서의 배경 잡음 레벨의 추적이 고려된다. Rangachari 및 Loizou [RL06]는 여러 방법들의 양호한 개요를 제공하고, 그 제한들 중 몇몇을 논의한다. 배경 잡음 레벨을 추적하기 위한 방법들은 예를 들어, 최소 추적 절차[RL06] [Coh03] [SFB00] [Dob95], VAD 기반(VAD=voice activity detection); 칼맨 필터링 [Gan05] [BJH06], 서브스페이스 분해들[BP06] [HJH08]; 소프트 결정[SS98] [MPC89] [HE95], 및 최소 통계이다.
- [0177] 최소 통계 접근법은 USAC-2(USAC=Unified Speech and Audio Coding)을 위한 범주 내에서 사용되도록 선택되었고, 후속하여 더 구체적으로 개술된다.
- [0178] 최적의 평활화(smoothing) 및 최소 통계[Mar01]에 기초한 잡음 전력 스펙트럼 밀도 추정은 잡음 추정기를 도입하고, 이러한 잡음 추정기는 능동 음성 또는 배경 잡음인 신호와 독립적으로 작용할 수 있다. 다른 방법들과 대조적으로, 최소 통계 알고리즘은 음성 활동과 음성 정지 사이를 구별하기 위해 임의의 명백한 임계치를 이용하지 않으므로, 전형적인 보이스 활동 검출 방법들보다 소프트-결정 방법들에 더 밀접하게 관련된다. 소프트-결정 방법들과 유사하게, 음성 활동 동안 추정된 잡음 PSD(Power Spectral Density)를 또한 업데이트할 수 있다.
- [0179] 최소 통계 방법은 즉, 음성과 잡음이 일반적으로 통계적으로 독립적이고, 잡음 음성 신호의 전력이 종종 잡음의

전력 레벨로 지연되는 2가지 관찰들에 있다. 그러므로, 잡음 신호 PSD의 최소치를 트래킹(tracking)함으로써 정밀한 잡음 PSD(PSD=power spectral density) 추정을 도출하는 것이 가능하다. 최소치가 평균값보다 작기(또는 다른 경우들에서 동일하기) 때문에, 최소치 트래킹 방법은 편향 보상을 요구한다.

- [0180] 편향은 평활화된 신호 PSD의 변동의 함수이고, 이와 같이, PSD 추정기의 평활화 파라미터에 의존한다. 일정한 평활화 파라미터 및 일정한 최소치 편향 정정을 이용하는 최소치 트래킹 상에서 더 일찍 작용하는 것에 비해, 시간 및 주파수에 의존하는 PSD 평활화가 사용되고, 이것은 또한 시간 및 주파수 의존 편향 보상을 요구한다.
- [0181] 최소치 트래킹을 이용하는 것은 잡음 전력의 어림 셈(rough estimate)을 제공한다. 하지만, 몇몇 결점들이 존재한다. 고정된 평활화 파라미터를 이용한 평활화는 평활화된 PSD 추정치의 음성 활동의 피크들을 확장한다. 이것은, 최소치 검색을 위한 슬라이딩 윈도우가 넓은 피크들로 미끄러질 수 있기 때문에 부적절한 잡음 추정들을 초래할 것이다. 따라서, 1에 가까운 평활화 파라미터들이 사용될 수 없고, 그 결과, 잡음 추정치는 비교적 큰 변동을 가질 것이다. 더욱이, 잡음 추정치는 낮은 값들쪽으로 편향된다. 더욱이, 증가하는 잡음 전력의 경우에, 최소치 트래킹은 뒤쳐진다.
- [0182] 낮은 복잡도를 갖는 MMSE 기반의 잡음 PSD 트래킹[HHJ10]은 DFT(Discrete Fourier Transform) 상에서 사용된 MMSE 검색을 이용하여 배경 잡음 PSD 접근법을 도입한다. 알고리즘은 이들 처리 단계들로 구성된다:
- [0183] - 최대 우도(maximum likelihood) 추정기는 이전 프레임의 잡음 PSD에 기초하여 계산된다.
- [0184] - 최소 평균 제곱 추정기가 계산된다.
- [0185] - 최대 우도 추정기는 결정-지향 접근법[EM84]을 이용하여 추정된다.
- [0186] - 역방향 편향 인자는, 음성 및 잡음 DFT 계수들이 가우시안 분배된다는 것을 가정하여 계산된다.
- [0187] - 추정된 잡음 전력 스펙트럼 밀도는 평활화된다.
- [0188] 알고리즘의 완전한 교착(dead lock)을 피하기 위해 적용된 안전망의(safety-net) 접근법이 또한 존재한다.
- [0189] 데이터-구동 회귀적 잡음 전력 추정에 기초한 비-고정 잡음의 트래킹[Eh08]은 크게 비-고정 잡음 소스들에 의해 오염된 음성 신호들로부터의 잡음 스펙트럼 변동의 추정을 위한 방법을 소개한다. 이 방법은 또한 시간/주파수 방향에서의 평활화를 이용한다.
- [0190] 잡음 전력 추정 및 추정 편향 정정의 평활화에 기초한 저-복잡도 잡음 추정 알고리즘[Yu09]은 [EH08]에 소개된 접근법을 개선한다. 주요 차이점은, 잡음 전력 추정을 위한 스펙트럼 이득 함수가 반복적인 데이터-구동 방법에 의해 발견된다는 것이다.
- [0191] 잡음 음성의 개선을 위한 통계 방법들[Mar03]은 소프트-결정 이득 변형[MCA99]에 의해, 선형적(a-priori) SNR의 추정[MCA99]에 의해, 적응형 이득 제한[MC99]에 의해, 그리고 MMSE 로그 스펙트럼 진폭 추정기[EM85]에 의해 [Mar01]에 주어진 최소치 통계치 접근법을 조합한다.
- [0192] 페이드 아웃은 특히 복수의 음성 및 오디오 코덱들, 특히 AMR([3GP12b]를 참조)(ACELP 및 CNG를 포함), AMR-WB([3GP09c]를 참조)(ACELP 및 CNG를 포함), AMR-WB+([3GP09a]를 참조)(ACELP, TCX 및 CNG를 포함), G.718([ITU08a]를 참조), G.719([ITU08b]를 참조), G.722([ITU07]을 참조), G.722.1([ITU05]를 참조), G.729([ITU12, CPK08, PKJ+11]을 참조), MPEG-4 HE-AAC/Enhanced AACPlus([EBU10, EBU12, 3GP12e, LS01, QD03]을 참조)(AAC 및 SBR을 포함), MPEG-4 HILN([ISO09, MEP01]을 참조) 및 OPUS([IET12]를 참조)(SILK 및 CELT를 포함)에 관심이 있다.
- [0193] 코덱에 따라, 페이드-아웃은 상이한 도메인들에서 수행된다:
- [0194] LPC를 이용하는 코덱들에 대해, 페이드-아웃은 선형 예측 도메인(또한 여기 도메인으로서 알려짐)에서 수행된다. 이것은 ACELP, 예를 들어, AMR, AMR-WB, AMR-WB+의 ACELP 코어, G.718, G.729, G.729.1, OPUS에서의 SILK 코어에 기초한 코덱들; 시간-주파수 변환, 예를 들어, AMR-WB+의 TCX 코어, OPUS에서의 CELT 코어를 이용하여 여기 신호를 추가로 처리하는 코덱들; 및 선형 예측 도메인, 예를 들어 AMR에서의 CNG, AMR-WB에서의 CNG, AMR-WB+에서의 CNG에서 동작하는 안락 잡음 생성(CNG) 계획들에 대해 그러하다.
- [0195] 시간 신호를 주파수 도메인으로 직접 변환하는 코덱들에 대해, 페이드-아웃 스펙트럼/서브대역 도메인에서 수행된다. 이것은 MDCT, 또는 MPEG-4 HE-AAC에서의 AAC, G.719, G.722(서브대역 도메인) 및 G.722.1과 같은 유사한 변환에 기초한 코덱들에 대해 그러하다.

[0196] 파라미터성 코덱들에 대해, 페이드-아웃은 파라미터성 도메인에 적용된다. 이것은 MPEG-4 HILN에 대해 그러하다.

[0197] 페이드-아웃 속도 및 페이드-아웃 곡선에 관해, 페이드-아웃은 감쇄 인자의 적용에 의해 공통적으로 실현되고, 이것은 적절한 도메인에서의 신호 표현에 적용된다. 감쇄 인자의 크기는 페이드-아웃 속도 및 페이드-아웃 곡선을 제어한다. 대부분의 경우에, 감쇄 인자는 프레임 방식으로 적용되지만, 또한 샘플 방식의 적용이 이용되고, 예를 들어, G.718 및 G.722를 참조하자.

[0198] 특정 신호 세그먼트에 대한 감쇄 인자는 2가지 방식들, 절대적 및 상대적으로 제공될 수 있다.

[0199] 감쇄 인자가 절대적으로 제공되는 경우에, 기준 레벨은 항상 마지막으로 수신된 프레임 중 하나이다. 절대 감쇄 인자들은 일반적으로 마지막 양호한 프레임 직후에 신호 세그먼트에 대해 1에 가까운 값에서 시작하고, 그런 후에 0쪽으로 더 빠르거나 더 느리게 저하된다. 페이드-아웃 곡선은 이들 인자들에 직접적으로 의존한다. 이것은, 예를 들어, G.722의 첨부 IV에 기재된 은닉에 대한 경우이고(특히, [ITU07, 도 IV.7]을 참조), 여기서 가능한 페이드-아웃 곡선들은 선형이거나 점차 선형이 된다. 이득 인자( $g(n)$ ), 반면  $g(0)$ 는 마지막으로 양호한 프레임의 이득 인자를 나타내고, 절대 감쇄 인자( $a_{abs}(n)$ )를 고려하면, 임의의 후속 손실된 프레임의 이득 인자는 수학식 21로서 도출될 수 있다

**수학식 21**

[0200] 
$$g(n) = \alpha_{abs}(n) \cdot g(0)$$

[0201] 감쇄 인자가 상대적으로 제공되는 경우에, 기준 레벨은 이전 프레임으로부터의 레벨이다. 이것은 회귀 은닉 절차의 경우에, 예를 들어 미리 감쇄된 신호가 추가로 처리되고 다시 감쇄되는 경우에 장점들을 갖는다.

[0202] 감쇄 인자가 회귀적으로 적용되면, 이것은 연속적으로 손실된 프레임들의 개수와 독립적인 고정된 값, 예를 들어, G.719에 대해 0.5(위를 참조); 예를 들어, [CPK08]에서 G.729에 대해 제안된 것과 같이, 연속적으로 손실된 프레임들의 개수에 상대적인 고정된 값; 처음 2개의 프레임들에 대해 1.0, 다음 2개의 프레임들에 대해 0.9, 프레임들(5 및 6)에 대해 0.8, 및 모든 후속 프레임들에 대해 0(위를 참조); 또는 연속적으로 손실된 프레임들의 개수에 상대적이고 신호 특징들, 예를 들어 불안정한 신호에 대해 더 빠른 페이드-아웃 및 안정한 신호에 대해 더 느린 페이드-아웃, 예를 들어 G718(위의 섹션 및 [ITU08a, 표 44]를 참조)에 의존하는 값일 수 있다.

[0203] 상대적인 페이드-아웃 인자( $0 \leq a_{rel}(n) \leq 1$ ){ $n$ 은 손실된 프레임( $n \geq 1$ )의 개수임}를 고려하면; 임의의 후속 프레임의 이득 인자는 다음과 같이 도출될 수 있다:

**수학식 22**

[0204] 
$$g(n) = \alpha_{rel}(n) \cdot g(n-1)$$

**수학식 23**

[0205] 
$$g(n) = \left( \prod_{m=1}^n \alpha(m) \right) \cdot g(0)$$

**수학식 24**

[0206] 
$$g(n) = \alpha_{rel}^n \cdot g(0)$$

- [0207] 이것은 지수적 페이딩을 초래한다.
- [0208] 페이드-아웃 절차에 관해, 일반적으로, 감쇄 인자가 규정되지만, 몇몇 응용 표준들(DRM, DAB+)에서, 감쇄 인자는 제조업자에게 위임된다.
- [0209] 상이한 신호 부분들이 개별적으로 페이딩되면, 상이한 감쇄 인자들은 예를 들어, 특정 속도를 갖는 음조(tonal) 성분들 및 다른 속도를 갖는 잡음-형 성분들(예를 들어, AMR, SILK)을 페이딩하는데 적용될 수 있다.
- [0210] 일반적으로, 특정 이득은 전체 프레임에 적용된다. 페이딩이 스펙트럼 도메인에서 수행될 때, 이것은 가능한 유일한 방식이다. 하지만, 페이딩이 시간 도메인 또는 선형 예측 도메인에서 이루어지면, 더욱 과립형의 (granular) 페이딩이 가능하다. 그러한 더욱 과립형의 페이딩은 G.718에서 적용되고, 여기서 개별적인 이득 인자들은 마지막 프레임의 이득 인자와 현재 프레임의 이득 인자 사이에서의 선형 보간에 의해 각 샘플에 대해 도출된다.
- [0211] 가변 프레임 지속기간을 갖는 코덱들에 대해, 상대적인 감쇄 인자는 프레임 지속기간에 따라 상이한 페이드-아웃 속도를 초래한다. 이것은 예를 들어, AAC에 대한 경우인데, 여기서 프레임 지속기간은 샘플링율에 의존한다.
- [0212] 적용된 페이딩 곡선을 마지막으로 수신된 신호의 시간 형태에 채택하기 위해, (정적) 페이드-아웃 인자들이 추가로 조정될 수 있다. 그러한 추가 동적 조정은 예를 들어, 이전 5개의 이득 인자들의 평균이 고려되는 AMR에 대해 적용된다([3GP12b] 및 섹션 1.8.1을 참조). 임의의 감쇄가 수행되기 전에, 현재 이득은 평균이 마지막 이득보다 작으면 평균으로 설정되고, 그렇지 않으면 마지막 이득이 사용된다. 더욱이, 그러한 추가 동적 조정은 예를 들어, 진폭이 이전의 이득 인자들의 선형 회귀를 이용하여 예측되는 G729에 대해 적용된다([CPK08, PKJ+11] 및 섹션 1.6을 참조). 이 경우에, 제 1 은닉된 프레임들에 대한 결과적인 이득 인자는 마지막으로 수신된 프레임의 이득 인자를 초과할 수 있다.
- [0213] 페이드-아웃의 목표 레벨에 관해, G.718 및 CELT를 제외하고, 목표 레벨은 이들 코덱들의 안락 잡음 생성(CNG)을 포함하는 모든 분석된 코덱들에 대해 0이다.
- [0214] G.718에서, 피치 여기(음조 성분들을 나타냄)의 페이딩과 랜덤 여기(잡음-형 성분들을 나타냄)의 페이딩은 개별적으로 수행된다. 피치 이득 인자가 0으로 페이딩되지만, 혁신 이득 인자는 CNG 여기 에너지로 페이딩된다.
- [0215] 상대적인 감쇄 인자들이 주어진다고 가정하면, 이것은 - 수학적 식 23에 기초하여 - 다음의 절대 감쇄 인자를 초래한다:

**수학적 식 25**

[0216] 
$$g(n) = \alpha_{rel} \cdot g(0-1) + (1 - \alpha_{rel}(n)) \cdot g_n$$

- [0217] 여기서,  $g_n$ 은 안락 잡음 생성 동안 사용된 여기의 이득이다. 이러한 수학적 식은  $g_n=0$ 일 때 수학적 식 23에 대응한다.
- [0218] G.718은 DTX/CNG의 경우에서 페이드-아웃을 수행하지 않는다.
- [0219] 목표 레벨쪽에서의 페이딩이 없고 음조 은닉(페이드-아웃을 포함)의 5개의 프레임들 이후에 페이딩되는 CELT에서, 레벨은 제 6 연속적으로 손실된 프레임에서 목표 레벨로 순간적으로 스위칭된다. 레벨은 수학적 식 19를 이용하여 밴드 방식으로 도출된다.
- [0220] 페이드-아웃의 목표 스펙트럼 형태에 관해, 모든 분석된 순 변환 기반의 코덱들(AAC, G.719, G.722, G.722.1)뿐 아니라 SBR은 간단히 페이드-아웃 동안 마지막으로 양호한 프레임의 스펙트럼 형태를 연장한다.
- [0221] 다양한 음성 코덱들은 LPC 합성을 이용하여 스펙트럼 형태를 평균에 페이딩한다. 평균은 정적(AMR) 또는 적응성(AMR-WB, AMR-WB+, G.718)일 수 있는 반면, 적응성은 정적 평균 및 단기간 평균(마지막  $n$  LP 계수 세트들을 평균화함으로써 도출됨)(LP=Linear Prediction)으로부터 도출된다.
- [0222] 논의된 코덱들(AMR, AMR-WB, AMR-WB+, G.718)에서의 모든 CNG 모듈들은 페이드-아웃 동안 마지막으로 양호한 프레임의 스펙트럼 형태를 연장한다.
- [0223] 배경 잡음 레벨 추적에 관해, 논문으로부터 알려진 5개의 상이한 접근법들이 존재한다:

- [0224] - SNR/VAD에 기초하지만, 낮은 SNR 음성에 사용하기에 어렵고 튜닝하기에 매우 어려운 보이스 활동도 검출기.
- [0225] - 소프트-결정 계획: 소프트-결정 접근법은 음성 존재의 확률을 고려한다[SS98][MPC89][HE95].
- [0226] - 최소 통계: PSD의 최소치가 트래킹되어 버퍼에서 시간이 지남에 따라 값들의 특정 양을 유지하고, 이에 따라 과거 샘플들로부터 최소 잡음을 발견하도록 한다[Mar01][HHJ10][EH08][Yu09].
- [0227] - 칼만 필터링: 알고리즘은 잡음(랜덤 변동들)을 포함하는, 시간이 지남에 따라 관찰된 일련의 측정치들을 이용하고, 단일 측정치에만 기초한 것들보다 더 정밀한 경향을 갖는 잡음 PSD의 추정치들을 발생시킨다. 칼만 필터는 시스템 상태의 통계적으로 최적의 추정치를 발생시키기 위해 잡음 입력 데이터의 스트림들 상에서 회귀적으로 동작한다[Gan05][BJH06].
- [0228] - 서브스페이스 분해: 이러한 접근법은 예를 들어 KLT(Karhunen-Loeve transform, 또한 주요 성분 분석으로서 알려짐) 및/또는 DFT(Discrete Time Fourier Transform)을 이용하여 잡음형 신호를 깨끗한 음성 신호 및 잡음 부분으로 분해하려고 시도한다. 그런 후에, 고유 벡터들/고유 값들은 임의의 평활화 알고리즘을 이용하여 추적될 수 있다[BP06][HJH08].

**발명의 내용**

**해결하려는 과제**

- [0229] 본 발명의 목적은 오디오 코딩 시스템들을 위한 개선된 개념들을 제공하는 것이다. 본 발명의 목적은 제1항에 따른 장치에 의해, 제14항에 따른 방법에 의해, 그리고 제15항에 따른 컴퓨터 프로그램에 의해 해결된다.

**과제의 해결 수단**

- [0230] 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치가 제공된다.
- [0231] 장치는 복수의 프레임들을 수신하기 위한 수신 인터페이스, 디코딩된 오디오 신호의 오디오 신호 샘플들을 저장하기 위한 지연 버퍼, 지연 버퍼에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하기 위한 샘플 선택기, 및 재구성된 오디오 신호의 재구성된 오디오 신호 샘플들을 얻기 위해 선택된 오디오 신호 샘플들을 처리하기 위한 샘플 프로세서를 포함한다. 샘플 선택기는, 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 피치 래그 정보에 따라 지연 버퍼에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성된다. 더욱이, 샘플 선택기는, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 수신 인터페이스에 의해 이전에 수신되는 다른 프레임에 의해 포함되는 피치 래그 정보에 따라 지연 버퍼에 저장된 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성된다.
- [0232] 일실시예에 따라, 샘플 프로세서는 예를 들어, 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 이득 정보에 따라 선택된 오디오 신호 샘플들을 재스케일링함으로써 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있다. 더욱이, 샘플 선택기는 예를 들어, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 수신 인터페이스에 의해 이전에 수신되는 상기 다른 프레임에 의해 포함되는 이득 정보에 따라 선택된 오디오 신호 샘플들을 재스케일링함으로써 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있다.
- [0233] 일실시예에서, 샘플 프로세서는, 예를 들어 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 이득 정보에 따른 값과 선택된 오디오 신호 샘플들을 곱함으로써 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있다. 더욱이, 샘플 선택기는, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 수신 인터페이스에 의해 이전에 수신되는 상기 다른 프레임에 의해 포함되는 이득 정보에 따른 값과 선택된 오디오 신호 샘플들을 곱함으로써 재구성된 오디오 신호 샘플들을 얻도록 구성된다.
- [0234] 일실시예에 따라, 샘플 프로세서는 예를 들어, 재구성된 오디오 신호 샘플들을 지연 버퍼에 저장하도록 구성될 수 있다.
- [0235] 일실시예에서, 샘플 프로세서는 예를 들어, 추가 프레임이 수신 인터페이스에 의해 수신되기 전에 재구성된 오

디오 신호 샘플들을 지연 버퍼에 저장하도록 구성될 수 있다.

- [0236] 일실시예에 따라, 샘플 프로세서는 예를 들어, 추가 프레임이 수신 인터페이스에 의해 수신된 후에 재구성된 오디오 신호 샘플들을 지연 버퍼에 저장하도록 구성될 수 있다.
- [0237] 일실시예에서, 샘플 프로세서는 예를 들어, 재스케일링된 오디오 신호 샘플들을 얻기 위해 이득 정보에 따라, 그리고 처리된 오디오 신호 샘플들을 얻기 위해 재스케일링된 오디오 신호 샘플들과 입력 오디오 신호 샘플들을 조합함으로써 선택된 오디오 신호 샘플들을 재스케일링하도록 구성될 수 있다.
- [0238] 일실시예에 따라, 샘플 프로세서는 예를 들어, 재스케일링된 오디오 신호 샘플들과 입력 오디오 신호 샘플들의 조합을 나타내는 처리된 오디오 신호 샘플들을 지연 버퍼에 저장하고, 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 재스케일링된 오디오 신호 샘플들을 지연 버퍼에 저장하지 않도록 구성될 수 있다. 더욱이, 샘플 프로세서는 재스케일링된 오디오 신호 샘플들을 지연 버퍼에 저장하고, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 처리된 오디오 신호 샘플들을 지연 버퍼에 저장하지 않도록 구성된다.
- [0239] 다른 실시예에 따라, 샘플 프로세서는 예를 들어, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 처리된 오디오 신호 샘플들을 지연 버퍼에 저장하도록 구성될 수 있다.
- [0240] 일실시예에서, 샘플 선택기는 예를 들어, 변형된 이득에 따라 선택된 오디오 신호 샘플들을 재스케일링함으로써 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있고, 변형된 이득은 수학적  $gain = gain\_past * damping$ 에 따라 정의되고,
- [0241] 여기서  $gain$ 은 변형된 이득이고, 샘플 선택기는 예를 들어,  $gain$ 이 계산된 후에  $gain\_past$ 를  $gain$ 으로 설정하도록 구성될 수 있고,  $damping$ 은 실값(real value)이다.
- [0242] 일실시예에 따라, 샘플 선택기는 예를 들어, 변형된 이득을 계산하도록 구성될 수 있다.
- [0243] 일실시예에서,  $damping$ 은 예를 들어,  $0 \leq damping \leq 1$ 에 따라 정의될 수 있다.
- [0244] 일실시예에 따라, 변형된 이득  $gain$ 은 예를 들어, 마지막 프레임이 수신 인터페이스에 의해 수신되었기 때문에 적어도 미리 한정된 수의 프레임들이 수신 인터페이스에 의해 수신되지 않았다면, 0으로 설정될 수 있다.
- [0245] 더욱이, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 방법이 제공된다. 방법은
- [0246] - 복수의 프레임들을 수신하는 단계,
- [0247] - 디코딩된 오디오 신호의 오디오 신호 샘플들을 저장하는 단계,
- [0248] - 지연 버퍼에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하는 단계, 및
- [0249] - 재구성된 오디오 신호의 재구성된 오디오 신호 샘플들을 얻기 위해 선택된 오디오 신호 샘플들을 처리하는 단계를
- [0250] 포함한다.
- [0251] 현재 프레임이 수신되면, 그리고 수신되는 현재 프레임이 붕괴되지 않으면, 지연 버퍼에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하는 단계는 현재 프레임에 의해 포함되는 피치 래그 정보에 따라 수행된다. 더욱이, 현재 프레임이 수신되지 않으면, 또는 수신되는 현재 프레임이 붕괴되면, 지연 버퍼에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하는 단계는 수신 인터페이스에 의해 이전에 수신되는 다른 프레임에 의해 포함되는 피치 래그 정보에 따라 수행된다.
- [0252] 더욱이, 컴퓨터 또는 신호 프로세서 상에서 실행될 때 전송할 방법을 구현하기 위한 컴퓨터 프로그램이 제공된다.
- [0253] 실시예들은 TCX LTP(TCX LTP = Transform Coded Excitation Long-Term Prediction)을 이용한다. 정상 동작 동안, TCX LTP 메모리는 잡음 및 재구성된 음조 성분들을 포함하는 합성된 신호로 업데이트된다.
- [0254] 은닉 동안 TCX LTP를 디스에이블링(disabling)하는 것 대신에, 정상 동작은 마지막으로 양호한 프레임에서 수신된 파라미터들을 통해 은닉 동안 계속될 수 있다. 이것은 신호의 스펙트럼 형태, 특히 LTP 필터에 의해 변형되



는 그러한 음조 성분들을 확보한다.

- [0255] 더욱이, 실시예들은 TCX LTP 피드백 루프를 결합 해제(decouple)한다. 정상 TCX LTP 동작의 간단한 계속은 추가 잡음을 도입하는데, 이는 각 업데이트 단계를 통해, LTP 여기로부터 추가의 무작위적으로 생성된 잡음이 도입되기 때문이다. 음조 성분들은 이를 통해 추가된 잡음에 의해 시간이 지남에 따라 더욱 더 왜곡된다.
- [0256] 이를 극복하기 위해, 업데이트된 TCX LTP 버퍼만이 원하지 않은 랜덤 잡음으로 음조 정보를 오염시키지 않기 위해 피드백될 수 있다(추가 잡음 없이).
- [0257] 더욱이, 실시예들에 따라, TCX LTP 이득은 0으로 페이딩된다.
- [0258] 이들 실시예들은, TCX LTP를 계속하는 것이 단기간에 신호 특징들을 확보하는데 도움을 주지만, 장기간에 대한 결점들을 갖는다는 발견에 기초한다: 은닉 동안 재생된 신호는 손실에 선행하여 존재한 보이싱/음조 정보를 포함할 것이다. 특히 깨끗한 음성 또는 배경 잡음에 걸친 음성에 대해, 음조 또는 고조파가 매우 긴 시간에 걸쳐 매우 느리게 감쇠할 가능성이 거의 없다. 은닉 동안 TCX LTP 동작을 계속함으로써, 특히 LTP 메모리 업데이트가 결합 해제되면(후호 스크램블링된 부분이 아니라 단지 음조 성분들이 피드백된다), 보이싱/음조 정보는 전체 페이드-아웃에 의해서만 안락 잡음 레벨로 감쇄되는 전체 손실에 대한 은닉된 신호에 존재하도록 유지할 것이다. 더욱이, 신호가 LTP의 보이싱 정보를 항상 병합하기 때문에, TCX LTP가 시간이 지남에 따라 감쇄되지 않고도 버스트 손실 동안 적용되면, 버스트 패킷 손실들 동안 안락 잡음 엔벨로프에 도달하는 것이 불가능하다.
- [0259] 그러므로, TCX LTP 이득은 0쪽으로 페이딩되어, LTP에 의해 표현된 음조 성분들은 0으로 페이딩될 것이고, 동시에 신호는 배경 신호 레벨 및 형태로 페이딩되고, 페이드-아웃은 바람직하지 않은 음조 성분들 없이 원하는 스펙트럼 배경 엔벨로프(안락 잡음)에 도달하게 된다.
- [0260] 실시예들에서, 동일한 페이딩 속도는 백색 잡음 페이딩에 대해 LTP 이득 페이딩에 사용된다.
- [0261] 이와 대조적으로, 종래 기술에서, 은닉 동안 LTP를 이용하는 알려진 변환 코덱은 존재하지 않는다. MPEG-4 LTP[ISO9]에 대해, 종래 기술에는 은닉 접근법들이 존재하지 않는다. LTP를 이용하는 종래 기술의 다른 MDCT 기반의 코덱은 CELT이지만, 이러한 코덱은 처음 5개의 프레임들에 대한 ACELP-형 은닉을 이용하고, 후속 프레임들에 대해, 배경 잡음이 생성되고, 이것은 LTP를 이용하지 않는다. TCX LTP를 이용하지 않는 종래 기술의 결함은, LTP로 모델링되는 모든 음조 성분들이 갑자기 사라진다는 것이다. 더욱이, 종래 기술의 ACELP 기반의 코덱들에서, LTP 동작은 은닉 동안 연장되고, 적응형 코덱의 이득은 0쪽으로 페이딩된다. 피드백 루프 동작에 관해, 종래 기술은 2가지 접근법들을 이용하는데, 전체 여기, 예를 들어 혁신 및 적응형 여기의 합이 피드백된다(AMR-WB); 또는 업데이트된 적응형 여기만이, 예를 들어 음조 신호 부분들만이 피드백된다(G.718). 전술한 실시예들은 종래 기술의 단점들을 극복한다.
- [0262] 더욱이, 오디오 신호를 디코딩하기 위한 장치가 제공된다.
- [0263] 장치는 수신 인터페이스를 포함한다. 수신 인터페이스는 복수의 프레임들을 수신하도록 구성되고, 수신 인터페이스는 복수의 프레임들 중 제 1 프레임을 수신하도록 구성되고, 상기 제 1 프레임은 오디오 신호의 제 1 오디오 신호 부분을 포함하고, 상기 제 1 오디오 신호 부분은 제 1 도메인에서 표현되고, 수신 인터페이스는 복수의 프레임들 중 제 2 프레임을 수신하도록 구성되고, 상기 제 2 프레임은 오디오 신호의 제 2 오디오 신호 부분을 포함한다.
- [0264] 더욱이, 장치는 제 2 신호 부분 정보를 얻기 위해 제 2 오디오 신호 부분으로부터 도출된 제 2 오디오 신호 부분 또는 값 또는 신호를 제 2 도메인으로부터 추적 도메인으로 변환하기 위한 변환 유닛을 포함하고, 제 2 도메인은 제 1 도메인과 상이하고, 추적 도메인은 제 2 도메인과 상이하고, 추적 도메인은 제 1 도메인과 동일하거나 상이하다.
- [0265] 더욱이, 장치는 잡음 레벨 추적 유닛을 포함하고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 1 신호 부분 정보를 수신하도록 구성되고, 제 1 신호 부분 정보는 제 1 오디오 신호 부분에 의존한다. 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성되고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 1 신호 부분 정보에 따라 그리고, 추적 도메인에서 표현되는 제 2 신호 부분 정보에 따라 잡음 레벨 정보를 결정하도록 구성된다.
- [0266] 더욱이, 장치는, 복수의 프레임들의 제 3 프레임이 수신 인터페이스에 의해 수신되지 않고 붕괴되면, 잡음 레벨 정보에 따라 오디오 신호의 제 3 오디오 신호 부분을 재구성하기 위한 재구성 유닛을 포함한다.

- [0267] 오디오 신호는 예를 들어, 음성 신호, 또는 음악 신호, 또는 음성 및 음악 등을 포함하는 신호일 수 있다.
- [0268] 제 1 신호 부분 정보가 제 1 오디오 신호 부분에 의존한다는 언급은, 제 1 신호 부분 정보가 제 1 오디오 신호 부분이라는 것, 또는 제 1 신호 부분 정보가 제 1 오디오 신호 부분에 따라 얻어진다는/생성된다는 것, 또는 몇몇 다른 방식으로, 제 1 오디오 신호 부분에 의존한다는 것을 의미한다. 예를 들어, 제 1 오디오 신호 부분은 제 1 신호 부분 정보를 얻기 위해 하나의 도메인으로부터 다른 도메인으로 변환되었을 수 있다.
- [0269] 마찬가지로, 제 2 신호 부분 정보가 제 2 오디오 신호 부분에 의존한다는 언급은, 제 2 신호 부분 정보가 제 2 오디오 신호 부분이라는 것, 또는 제 2 신호 부분 정보가 제 2 오디오 신호 부분에 따라 얻어진다는/생성된다는 것, 또는 몇몇 다른 방식으로 제 2 오디오 신호 부분에 의존한다는 것을 의미한다. 예를 들어, 제 2 오디오 신호 부분은 제 2 신호 부분 정보를 얻기 위해 하나의 도메인으로부터 다른 도메인으로 변환되었을 수 있다.
- [0270] 일실시예에서, 제 1 오디오 신호 부분은 예를 들어, 제 1 도메인으로서 시간 도메인에서 표현될 수 있다. 더욱이, 변환 유닛은 예를 들어, 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 도메인인 여기 도메인으로부터 추적 도메인인 시간 도메인으로 변환하도록 구성될 수 있다. 더욱이, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인으로서 시간 도메인에서 표현되는 제 1 신호 부분 정보를 수신하도록 구성될 수 있다. 더욱이, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인으로서 시간 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성될 수 있다.
- [0271] 일실시예에 따라, 제 1 오디오 신호 부분은 예를 들어, 제 1 도메인으로서 여기 도메인에서 표현될 수 있다. 더욱이, 변환 유닛은 예를 들어, 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 도메인인 시간 도메인으로부터 추적 도메인인 여기 도메인으로 변환하도록 구성될 수 있다. 더욱이, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인으로서 여기 도메인에서 표현되는 제 1 신호 부분 정보를 수신하도록 구성될 수 있다. 더욱이, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인으로서 여기 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성될 수 있다.
- [0272] 일실시예에서, 제 1 오디오 신호 부분은 예를 들어, 제 1 도메인으로서 여기 도메인에서 표현될 수 있고, 잡음 레벨 추적 유닛은 예를 들어, 제 1 신호 부분 정보를 수신하도록 구성될 수 있고, 상기 제 1 신호 부분 정보는 추적 도메인인 FFT 도메인에서 표현되고, 상기 제 1 신호 부분 정보는 여기 도메인에서 표현되는 상기 제 1 오디오 신호 부분에 의존하고, 변환 유닛은 예를 들어, 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값을 추적 도메인인 FFT 도메인으로 변환하도록 구성되고, 잡음 레벨 추적 유닛은 예를 들어, FFT 도메인에서 표현되는 제 2 오디오 신호 부분을 수신하도록 구성될 수 있다.
- [0273] 일실시예에서, 장치는 예를 들어, 제 1 오디오 신호 부분에 따라 제 1 집합된 값을 결정하기 위한 제 1 집합 유닛을 더 포함할 수 있다. 더욱이, 장치는 예를 들어, 제 2 오디오 신호 부분에 따라 제 2 오디오 신호 부분으로부터 도출된 값으로서 제 2 집합된 값을 결정하기 위한 제 2 집합 유닛을 더 포함할 수 있다. 더욱이, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인에서 표현되는 제 1 신호 부분 정보로서 제 1 집합된 값을 수신하도록 구성될 수 있고, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인에서 표현되는 제 2 신호 부분 정보로서 제 2 집합된 값을 수신하도록 구성될 수 있고, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인에서 표현되는 제 1 집합된 값에 따라, 그리고 추적 도메인에서 표현되는 제 2 집합된 값에 따라 잡음 레벨 정보를 결정하도록 구성될 수 있다.
- [0274] 일실시예에 따라, 제 1 집합 유닛은 예를 들어, 제 1 집합된 값이 제 1 오디오 신호 부분 또는 제 1 오디오 신호 부분으로부터 도출된 신호의 제곱 평균 제곱근(root mean square)을 나타내도록 제 1 집합된 값을 결정하도록 구성될 수 있다. 더욱이, 제 2 집합 유닛은 예를 들어, 제 2 집합된 값이 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 신호의 제곱 평균 제곱근을 나타내도록 제 2 집합된 값을 결정하도록 구성될 수 있다.
- [0275] 일실시예에서, 변환 유닛은 예를 들어, 제 2 오디오 신호 부분으로부터 도출된 값 상에 이득 값을 적용함으로써 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 도메인으로부터 추적 도메인으로 변환하도록 구성될 수 있다.
- [0276] 실시예들에 따라, 이득 값은 예를 들어, 선형 예측 코딩 합성에 의해 도입된 이득을 나타낼 수 있거나, 이득 값은 예를 들어, 선형 예측 코딩 합성 및 디엠퍼시스에 의해 도입된 이득을 나타낼 수 있다.
- [0277] 일실시예에서, 잡음 레벨 추적 유닛은 예를 들어, 최소 통계 접근법을 적용함으로써 잡음 레벨 정보를 결정하도록 구성될 수 있다.

- [0278] 일실시예에 따라, 잡음 레벨 추적 유닛은 예를 들어, 잡음 레벨 정보로서 안락 잡음을 결정하도록 구성될 수 있다. 재구성 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0279] 일실시예에서, 잡음 레벨 추적 유닛은 예를 들어, 잡음 레벨 스펙트럼으로부터 도출된 잡음 레벨 정보로서 안락 잡음 레벨을 결정하도록 구성될 수 있고, 상기 잡음 레벨 스펙트럼은 최소 통계 접근법을 적용함으로써 얻어진다. 재구성 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 복수의 선형 예측 계수들에 따라 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0280] 다른 실시예에 따라, 잡음 레벨 추적 유닛은 예를 들어, 잡음 레벨 정보로서 안락 잡음 레벨을 나타내는 복수의 선형 예측 계수들을 결정하도록 구성될 수 있고, 재구성 유닛은 예를 들어, 복수의 선형 예측 계수들에 따라 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0281] 일실시예에서, 잡음 레벨 추적 유닛은 잡음 레벨 정보로서 안락 잡음 레벨을 나타내는 복수의 FFT 계수들을 결정하도록 구성되고, 제 1 재구성 유닛은, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 상기 FFT 계수들로부터 도출된 안락 잡음 레벨에 따라 제 3 오디오 신호 부분을 재구성하도록 구성된다.
- [0282] 일실시예에서, 재구성 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 그리고 제 1 오디오 신호 부분에 따라 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0283] 일실시예에 따라, 재구성 유닛은 예를 들어, 제 1 또는 제 2 오디오 신호 부분으로부터 도출된 신호를 감쇄하거나 증폭함으로써 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0284] 일실시예에서, 장치는 예를 들어, 지연 버퍼를 포함하는 장기간 예측 유닛을 더 포함할 수 있다. 더욱이, 장기간 예측 유닛은 예를 들어, 제 1 또는 제 2 오디오 신호 부분에 따라, 지연 버퍼에 저장된 지연 버퍼 입력에 따라, 그리고 장기간 예측 이득에 따라, 처리된 신호를 생성하도록 구성될 수 있다. 더욱이, 장기간 예측 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 장기간 예측 이득을 0쪽으로 페이딩하도록 구성될 수 있다.
- [0285] 일실시예에 따라, 장기간 예측 유닛은 예를 들어, 장기간 예측 이득을 0쪽으로 페이딩하도록 구성될 수 있고, 장기간 예측 이득이 0으로 페이딩되는 속도는 페이드-아웃 인자에 의존한다.
- [0286] 일실시예에서, 장기간 예측 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 생성된 처리된 신호를 지연 버퍼에 저장함으로써 지연 버퍼 입력을 업데이트하도록 구성될 수 있다.
- [0287] 일실시예에 따라, 변환 유닛은 예를 들어, 제 1 변환 유닛일 수 있고, 재구성 유닛은 제 1 재구성 유닛이다. 장치는 제 2 변환 유닛 및 제 2 재구성 유닛을 더 포함한다. 제 2 변환 유닛은 예를 들어, 복수의 프레임들 중 제 4 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보를 추적 도메인으로부터 제 2 도메인으로 변환하도록 구성될 수 있다. 더욱이, 제 2 재구성 유닛은 예를 들어, 복수의 프레임 중 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 제 2 도메인에서 표현되는 잡음 레벨 정보에 따라 오디오 신호의 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0288] 일실시예에서, 제 2 재구성 유닛은 예를 들어, 잡음 레벨 정보에 따라, 그리고 제 2 오디오 신호 부분에 따라 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0289] 일실시예에 따라, 제 2 재구성 유닛은 예를 들어, 제 1 또는 제 2 오디오 신호 부분으로부터 도출된 신호를 감쇄하거나 증폭함으로써 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0290] 더욱이, 오디오 신호를 디코딩하기 위한 방법이 제공된다.
- [0291] 방법은

- [0292] - 복수의 프레임들의 제 1 프레임을 수신하는 단계로서, 상기 제 1 프레임은 오디오 신호의 제 1 오디오 신호 부분을 포함하고, 상기 제 1 오디오 신호 부분은 제 1 도메인에서 표현되는, 복수의 프레임들의 제 1 프레임을 수신하는 단계;
- [0293] - 복수의 프레임들의 제 2 프레임을 수신하는 단계로서, 상기 제 2 프레임은 오디오 신호의 제 2 오디오 신호 부분을 포함하는, 복수의 프레임들의 제 2 프레임을 수신하는 단계;
- [0294] - 제 2 신호 부분 정보를 얻기 위해 제 2 오디오 신호 부분, 또는 제 2 오디오 신호 부분으로부터 도출된 값 또는 신호를 제 2 도메인으로부터 추적 도메인으로 변환하는 단계로서, 제 2 도메인은 제 1 도메인과 상이하고, 추적 도메인은 제 2 도메인과 상이하고, 추적 도메인은 제 1 도메인과 동일하거나 상이한, 변환하는 단계;
- [0295] - 추적 도메인에서 표현되는 제 1 신호 부분 정보에 따라, 그리고 추적 도메인에서 표현되는 제 2 신호 부분 정보에 따라, 잡음 레벨 정보를 결정하는 단계로서, 제 1 신호 부분 정보는 제 1 오디오 신호 부분에 의존하는, 결정하는 단계, 및
- [0296] - 복수의 프레임들의 제 3 프레임이 수신되지 않으면, 또는 상기 제 3 프레임이 수신되지만, 붕괴되면, 추적 도메인에서 표현되는 잡음 레벨 정보에 따라 오디오 신호의 제 3 오디오 신호 부분을 재구성하는 단계를
- [0297] 포함한다.
- [0298] 더욱이, 컴퓨터 또는 신호 프로세서 상에서 실행될 때 진술한 방법을 구현하기 위한 컴퓨터 프로그램이 제공된다.
- [0299] 본 발명의 몇몇 실시예들은, 편향 보상을 위한 알고리즘을 제공하고 일반적으로 잡음 트래킹을 가속시키기 위해, 평활화된 주기도(periodogram)의 트래킹 성능과 그 변동이 더 양호하게 균형을 이루도록 시간 변화 평활화 파라미터를 제공한다.
- [0300] 본 발명의 실시예들은 페이드-아웃에 관해, 다음의 파라미터들이 관심 있다는 발견에 기초한다: 페이드-아웃 도메인; 페이드-아웃 속도, 또는 더 일반적으로, 페이드-아웃 곡선; 페이드-아웃의 목표 레벨; 페이드-아웃의 목표 스펙트럼 형태; 및/또는 배경 잡음 레벨 추적. 이러한 정황에서, 실시예들은, 종래 기술이 중요한 결함들을 갖는다는 발견에 기초한다.
- [0301] 오디오 은닉 동안 스위칭된 오디오 코딩 시스템들에 대해 개선된 신호 페이드 아웃을 위한 장치 및 방법이 제공된다.
- [0302] 더욱이, 컴퓨터 또는 신호 프로세서 상에서 실행될 때 진술한 방법을 구현하기 위한 컴퓨터 프로그램이 제공된다.
- [0303] 실시예들은 안락 잡음 레벨로의 페이드-아웃을 실현한다. 실시예들에 따라, 여기 도메인에서의 공통적인 안락 잡음 레벨이 실현된다. 버스트 패킷 손실 동안 목표로 정해진 안락 잡음 레벨은 사용시 코어 코더(ACELP/TCX)에도 불구하고 동일할 것이고, 항상 최신형일 것이다. 알려진 종래 기술은 없으며, 여기서 공통적인 잡음 레벨 추적이 필요하다. 실시예들은 버스트 패킷 손실들 동안 안락 잡음형 신호로의 스위칭된 코덱의 페이딩을 제공한다.
- [0304] 더욱이, 실시예들은, 기능들(PROM) 및 메모리가 공유될 수 있기 때문에 전체 복잡도가 2개의 독립적인 잡음 레벨 추적 모듈들을 갖는 것에 비해 더 낮아지는 것을 실현한다.
- [0305] 실시예들에서, 여기 도메인에서의 레벨 도출(시간 도메인에서의 레벨 도출에 비해)은 능동 음성 동안 더 최소치를 제공하는데, 이는 음성 정보의 부분이 LP 계수들에 의해 커버되기 때문이다.
- [0306] ACELP의 경우에, 실시예들에 따라, 레벨 도출은 여기 도메인에서 발생한다. TCX의 경우에, 실시예들에서, 레벨은 시간 도메인에서 도출되고, LPC 합성 및 디엠퍼시스의 이득은 여기 도메인에서 에너지 레벨을 모델링하기 위해 정정 인자로서 적용된다. 예를 들어, FDNS 이전에, 여기 도메인에서 레벨을 추적하는 것은 또한 이론적으로 가능하지만, TCX 여기 도메인과 ACELP 여기 도메인 사이의 레벨 보상은 약간 더 복잡한 것으로 간주된다.
- [0307] 종래 기술은 상이한 도메인들에서 그러한 공통적인 배경 레벨 추적을 병합하지 않는다. 종래 기술은 스위칭된 코덱 시스템에서, 예를 들어 여기 도메인에서, 그러한 공통적인 안락 잡음 레벨 추적을 갖지 않는다. 따라서, 실시예들은 종래 기술에 비해 유리한데, 이는 종래 기술에 대해, 버스트 패킷 손실들 동안 목표로 정해진 안락 잡음 레벨이 레벨이 추적된 이전 코딩 모드(ACELP/TCX)에 따라 상이할 수 있기 때문이고; 종래 기술에서, 각 코

딩 모드에 대해 개별적인 추적이 불필요한 오버헤드(overhead) 및 추가적인 계산 복잡도를 야기하기 때문이고; 종래 기술에서, 이러한 코어로의 최근의 스위칭으로 인해 어느 코어에서도 최신 안락 잡음 레벨이 이용가능하지 않기 때문이다.

[0308] 몇몇 실시예들에 따라, 레벨 추적은 여기 도메인에서 수행되지만, TCX 페이드-아웃은 시간 도메인에서 수행된다. 시간 도메인에서의 페이딩에 의해, TDAC의 장애들이 회피되고, 이것은 얼라이싱을 야기한다. 이것은 특히 음조 신호 성분들이 은닉될 때 관심 있다. 더욱이, ACELP 여기 도메인과 MDCT 스펙트럼 도메인 사이의 레벨 변환이 회피되어, 예를 들어, 계산 리소스들이 절감된다. 여기 도메인과 시간 도메인 사이의 스위칭으로 인해, 레벨 조정은 여기 도메인과 시간 도메인 사이에서 요구된다. 이것은 LPC 합성 및 프리엠퍼시스(preemphasis)에 의해 도입된 이득의 도출에 의해, 그리고 2개의 도메인들 사이에서 레벨을 변환하기 위해 정정 인자로서 이러한 이득을 이용하여 해결된다.

[0309] 이와 대조적으로, 종래 기술은 여기 도메인에서의 레벨 추적과, 시간 도메인에서의 TCX 페이드-아웃을 수행하지 않는다. 종래 기술의 변환 기반의 코덱들에 관해, 감쇄 인자는 여기 도메인(시간-도메인/ACELP형 은닉 접근법들에 대해, [3GP09a]를 참조)에서 또는 주파수 도메인(프레임 반복 또는 잡음 교체와 같은 주파수 도메인 접근법들에 대해, [LS01]을 참조)에 적용된다. 주파수 도메인에서 감쇄 인자를 적용하기 위해 종래 기술의 접근법의 결합은, 얼라이싱이 시간 도메인에서의 중첩-추가 영역에서 야기된다는 것이다. 이것은, 상이한 감쇄 인자들이 적용되는 인접 프레임들에 대해 그러할 것인데, 이는 페이딩 절차가 TDAC(time domain alias cancellation)의 장애를 야기하기 때문이다. 이것은 특히 음조 신호 성분들이 은닉될 때 관련된다. 이에 따라, 전술한 실시예들은 종래 기술에 비해 유리하다.

[0310] 실시예들은 LPC 합성 이득에 대한 고역 통과 필터의 영향을 보상한다. 실시예들에 따라, 고역 통과 필터링된 무성음 여기에 의해 야기된 LPC 분석 및 엠퍼시스의 불필요한 이득 변화를 보상하기 위해, 정정 인자가 도출된다. 이러한 정정 인자는 이러한 불필요한 이득 변화를 고려하고, 여기 도메인에서의 목표 안락 잡음 레벨을 변형하여, 정확한 목표 레벨이 시간 도메인에 도달된다.

[0311] 이와 대조적으로, 종래 기술, 예를 들어 G.718[ITU08a]는, 마지막으로 양호한 프레임의 신호가 무성음으로서 분류되지 않으면, 도 2에 도시된 바와 같이, 무성음 여기의 신호 경로에 고역 통과 필터를 도입한다. 이에 의해, 종래 기술은 불필요한 부작용들을 야기하는데, 이는 후속적인 LPC 합성의 이득이 이러한 고역 통과 필터에 의해 변경되는 신호 특징들에 의존하기 때문이다. 배경 레벨이 여기 도메인에서 추적되고 적용되기 때문에, 알고리즘은 LPC 합성 이득에 의존하고, 이것은 다시 여기 신호의 특징들에 의존한다. 즉, 종래 기술에 의해 수행된 바와 같이, 고역 통과 필터링으로 인한 여기의 신호 특징들의 변형은 LPC 합성의 변형된(일반적으로 감소된) 이득을 초래할 수 있다. 이것은, 여기 레벨이 정확하더라도 잘못된 출력 레벨을 초래한다.

[0312] 실시예들은 종래 기술의 이들 단점들을 극복한다.

[0313] 특히, 실시예들은 안락 잡음의 적응형 스펙트럼 형태를 실현한다. G.718에 대조적으로, 배경 잡음의 스펙트럼 형태를 추적함으로써, 그리고 버스트 패킷 손실들 동안 이러한 형태를 적용(페이딩)함으로써, 선행하는 배경 잡음의 잡음 특징이 매칭되어, 안락 잡음의 쾌적한 잡음 특징을 초래한다. 이것은 오프라인 트레이닝(offline training)에 의해 도출된 스펙트럼 엔벨로프를 이용함으로써 도입될 수 있는 스펙트럼 형태 및/또는 마지막으로 수신된 프레임들의 스펙트럼 형태의 눈에 거슬리는 미스매치들(obtrusive mismatches)을 회피한다.

[0314] 더욱이, 오디오 신호를 디코딩하기 위한 장치가 제공된다. 장치는 수신 인터페이스를 포함하고, 수신 인터페이스는 오디오 신호의 제 1 오디오 신호 부분을 포함하는 제 1 프레임을 수신하도록 구성되고, 수신 인터페이스는 오디오 신호의 제 2 오디오 신호 부분을 포함하는 제 2 프레임을 수신하도록 구성된다.

[0315] 더욱이, 장치는 잡음 레벨 추적 유닛을 포함하고, 잡음 레벨 추적 유닛은 제 1 오디오 신호 부분 및 제 2 오디오 신호 부분 중 적어도 하나에 따라 잡음 레벨 정보를 결정하도록 구성되고(이것은 제 1 오디오 신호 부분 및/또는 제 2 오디오 신호 부분에 따르는 것을 의미한다), 잡음 레벨 정보는 추적 도메인에서 표현된다. 더욱이, 장치는 제 1 재구성 도메인에서, 복수의 프레임들의 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 오디오 신호의 제 3 오디오 신호 부분을 재구성하기 위한 제 1 재구성 유닛을 포함하고, 제 1 재구성 도메인은 추적 도메인과 상이하거나, 동일하다.

[0316] 더욱이, 장치는, 복수의 프레임들의 제 4 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보를 추적 도메인으로부터 제 2 재구성

도메인으로 변환하기 위한 변환 유닛을 포함하고, 제 2 재구성 도메인은 추적 도메인과 상이하고, 제 2 재구성 도메인은 제 1 재구성 도메인과 상이하고,

- [0317] 더욱이, 장치는 제 2 재구성 도메인에서, 복수의 프레임들의 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 제 2 재구성 도메인에서 표현되는 잡음 레벨 정보에 따라 오디오 신호의 제 4 오디오 신호 부분을 재구성하기 위한 제 2 재구성 유닛을 포함한다.
- [0318] 몇몇 실시예들에 따라, 추적 도메인은 예를 들어, 추적 도메인이 시간 도메인, 스펙트럼 도메인, FFT 도메인, MDCT 도메인, 또는 여기 도메인이 되도록 이루어질 수 있다. 제 1 재구성 도메인은 예를 들어, 시간 도메인, 스펙트럼 도메인, FFT 도메인, MDCT 도메인, 또는 여기 도메인일 수 있다. 제 2 재구성 도메인은 예를 들어, 시간 도메인, 스펙트럼 도메인, FFT 도메인, MDCT 도메인, 또는 여기 도메인일 수 있다.
- [0319] 일실시예에서, 추적 도메인은 예를 들어, FFT 도메인일 수 있고, 제 1 재구성 도메인은 예를 들어, 시간 도메인일 수 있고, 제 2 재구성 도메인은 예를 들어, 여기 도메인일 수 있다.
- [0320] 다른 실시예에서, 추적 도메인은 예를 들어, 시간 도메인일 수 있고, 제 1 재구성 도메인은 예를 들어, 시간 도메인일 수 있고, 제 2 재구성 도메인은 예를 들어, 여기 도메인일 수 있다.
- [0321] 일실시예에 따라, 상기 제 1 오디오 신호 부분은 예를 들어, 제 1 입력 도메인에서 표현될 수 있고, 상기 제 2 오디오 신호 부분은 예를 들어, 제 2 입력 도메인에서 표현될 수 있다. 변환 유닛은 예를 들어, 제 2 변환 유닛일 수 있다. 장치는 예를 들어, 제 2 신호 부분 정보를 얻기 위해 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값 또는 신호를 제 2 입력 도메인으로부터 추적 도메인으로 변환하기 위한 제 1 변환 유닛을 더 포함할 수 있다. 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인에서 표현되는 제 1 신호 부분 정보를 수신하도록 구성될 수 있고, 제 1 신호 부분 정보는 제 1 오디오 신호 부분에 의존하고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성되고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 1 신호 부분 정보에 따라, 그리고 추적 도메인에서 표현되는 제 2 신호 부분 정보에 따라 잡음 레벨 정보를 결정하도록 구성된다.
- [0322] 일실시예에 따라, 제 1 입력 도메인은 예를 들어, 여기 도메인일 수 있고, 제 2 입력 도메인은 예를 들어, MDCT 도메인일 수 있다.
- [0323] 다른 실시예에서, 제 1 입력 도메인은 예를 들어, MDCT 도메인일 수 있고, 제 2 입력 도메인은 예를 들어 MDCT 도메인일 수 있다.
- [0324] 일실시예에 따라, 제 1 재구성 유닛은 예를 들어, 잡음형 스펙트럼으로의 제 1 페이딩을 수행함으로써 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다. 제 2 재구성 유닛은 예를 들어, 잡음형 스펙트럼으로의 제 2 페이딩 및/또는 LTP 이득의 제 2 페이딩을 수행함으로써 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다. 더욱이, 제 1 재구성 유닛 및 제 2 재구성 유닛은 예를 들어, 잡음형 스펙트럼으로의 제 1 페이딩 및 제 2 페이딩 및/또는 동일한 페이딩 속도를 갖는 LTP 이득의 제 2 페이딩을 수행하도록 구성될 수 있다.
- [0325] 일실시예에서, 장치는 예를 들어, 제 1 오디오 신호 부분에 따라 제 1 집합된 값을 결정하기 위한 제 1 집합 유닛을 더 포함할 수 있다. 더욱이, 장치는 예를 들어, 제 2 오디오 신호 부분에 따라, 제 2 오디오 신호 부분으로부터 도출된 값으로서 제 2 집합된 값을 결정하기 위한 제 2 집합 유닛을 더 포함할 수 있다. 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인에서 표현되는 제 1 신호 부분 정보로서 제 1 집합된 값을 수신하도록 구성될 수 있고, 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인에서 표현되는 제 2 신호 부분 정보로서 제 2 집합된 값을 수신하도록 구성될 수 있고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 1 집합된 값에 따라, 그리고 추적 도메인에서 표현되는 제 2 집합된 값에 따라 잡음 레벨 정보를 결정하도록 구성된다.
- [0326] 일실시예에 따라, 제 1 집합 유닛은 예를 들어, 제 1 집합된 값이 제 1 오디오 신호 부분 또는 제 1 오디오 신호 부분으로부터 도출된 신호의 제곱 평균 제곱근을 나타내도록 제 1 집합된 값을 결정하도록 구성될 수 있다. 제 2 집합 유닛은, 제 2 집합된 값이 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 신호의 제곱 평균 제곱근을 나타내도록 제 2 집합된 값을 결정하도록 구성된다.
- [0327] 일실시예에서, 제 1 변환 유닛은 예를 들어, 제 2 오디오 신호 부분으로부터 도출된 값 상에 이득 값을 적용함으로써 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 입력 도메인으로부터 추적 도메인으로 변환하도록 구성될 수 있다.

- [0328] 일실시예에 따라, 이득 값은 예를 들어, 선형 예측 코딩 합성에 의해 도입된 이득을 나타낼 수 있거나, 이득 값은 선형 예측 코딩 합성 및 디엠퍼시스에 의해 도입된 이득을 나타낸다.
- [0329] 일실시예에서, 잡음 레벨 추적 유닛은 예를 들어, 최소 통계 접근법을 적용함으로써 잡음 레벨 정보를 결정하도록 구성될 수 있다.
- [0330] 일실시예에 따라, 잡음 레벨 추적 유닛은 예를 들어, 잡음 레벨 정보로서 안락 잡음 레벨을 결정하도록 구성될 수 있다. 재구성 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0331] 일실시예에서, 잡음 레벨 추적 유닛은 예를 들어, 잡음 레벨 스펙트럼으로부터 도출된 잡음 레벨 정보로서 안락 잡음 레벨을 결정하도록 구성될 수 있고, 상기 잡음 레벨 스펙트럼은 최소 통계 접근법을 적용함으로써 얻어진다. 재구성 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 복수의 선형 예측 계수들에 따라 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0332] 일실시예에 따라, 제 1 재구성 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 그리고 제 1 오디오 신호 부분에 따라, 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0333] 일실시예에서, 제 1 재구성 유닛은 예를 들어, 제 1 오디오 신호 부분을 감쇄하거나 증폭함으로써 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0334] 일실시예에 따라, 제 2 재구성 유닛은 예를 들어, 잡음 레벨 정보에 따라 그리고 제 2 오디오 신호 부분에 따라 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0335] 일실시예에서, 제 2 재구성 유닛은 예를 들어, 제 2 오디오 신호 부분을 감쇄하거나 증폭함으로써 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0336] 일실시예에 따라, 장치는 예를 들어, 지연 버퍼를 포함하는 장기간 예측 유닛을 더 포함할 수 있고, 장기간 예측 유닛은 예를 들어, 제 1 또는 제 2 오디오 신호 부분에 따라, 지연 버퍼에 저장된 지연 버퍼 입력에 따라, 그리고 장기간 예측 이득에 따라 처리된 신호를 생성하도록 구성될 수 있고, 장기간 예측 유닛은, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 장기간 예측 이득을 0쪽으로 페이딩하도록 구성된다.
- [0337] 일실시예에서, 장기간 예측 유닛은 예를 들어, 장기간 예측 이득을 0쪽으로 페이딩하도록 구성될 수 있고, 장기간 예측 이득이 0으로 페이딩되는 속도는 페이드-아웃 인자에 의존한다.
- [0338] 일실시예에서, 장기간 예측 유닛은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 생성된 처리된 신호를 지연 버퍼에 저장함으로써 지연 버퍼 입력을 업데이트하도록 구성될 수 있다.
- [0339] 더욱이, 오디오 신호를 디코딩하기 위한 방법이 제공된다. 방법은
- [0340] - 오디오 신호의 제 1 오디오 신호 부분을 포함하는 제 1 프레임을 수신하고, 오디오 신호의 제 2 오디오 신호 부분을 포함하는 제 2 프레임을 수신하는 단계,
- [0341] - 제 1 오디오 신호 부분 및 제 2 오디오 신호 부분 중 적어도 하나에 따라 잡음 레벨 정보를 결정하는 단계로서, 잡음 레벨 정보는 추적 도메인에서 표현되는, 잡음 레벨 정보를 결정하는 단계,
- [0342] - 제 1 재구성 도메인에서, 복수의 프레임들의 제 3 프레임이 수신되지 않으면, 또는 상기 제 3 프레임이 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 오디오 신호의 제 3 오디오 신호 부분을 재구성하는 단계로서, 제 1 재구성 도메인은 추적 도메인과 상이하거나 동일한, 제 3 오디오 신호 부분을 재구성하는 단계,
- [0343] - 복수의 프레임들의 제 4 프레임이 수신되지 않으면, 또는 상기 제 4 프레임이 수신되지만, 붕괴되면, 잡음 레벨 정보를 추적 도메인으로부터 제 2 재구성 도메인으로 변환하는 단계로서, 제 2 재구성 도메인은 추적 도메인과 상이하고, 제 2 재구성 도메인은 제 1 재구성 도메인과 상이한, 변환 단계, 및
- [0344] - 제 2 재구성 도메인에서, 복수의 프레임들의 상기 제 4 프레임이 수신되지 않으면, 또는 상기 제 4 프레임이

수신되지만, 붕괴되면, 제 2 재구성 도메인에서 표현되는 잡음 레벨 정보에 따라 오디오 신호의 제 4 오디오 신호 부분을 재구성하는 단계를

- [0345] 포함한다.
- [0346] 더욱이, 컴퓨터 또는 신호 프로세서 상에서 실행될 때 전술한 방법을 구현하기 위한 컴퓨터 프로그램이 제공된다.
- [0347] 더욱이, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치가 제공된다. 장치는 하나 이상의 프레임들을 수신하기 위한 수신 인터페이스, 계수 생성기, 및 신호 재구성기를 포함한다. 계수 생성기는, 하나 이상의 프레임들의 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 하나 이상의 제 1 오디오 신호 계수들을 결정하도록 구성되고, 상기 하나 이상의 제 1 오디오 신호 계수들은 인코딩된 오디오 신호의 특징을 나타내고, 하나 이상의 잡음 계수들은 인코딩된 오디오 신호의 배경 잡음을 나타낸다. 더욱이, 계수 생성기는, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 하나 이상의 제 1 오디오 신호 계수들에 따라, 그리고 하나 이상의 잡음 계수들에 따라, 하나 이상의 제 2 오디오 신호 계수들을 생성하도록 구성된다. 오디오 신호 재구성기는, 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 하나 이상의 제 1 오디오 신호 계수들에 따라 재구성된 오디오 신호의 제 1 부분을 재구성하도록 구성된다. 더욱이, 오디오 신호 재구성기는, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 하나 이상의 제 2 오디오 신호 계수들에 따라 재구성된 오디오 신호의 제 2 부분을 재구성하도록 구성된다.
- [0348] 몇몇 실시예들에서, 하나 이상의 제 1 오디오 신호 계수들은 예를 들어, 인코딩된 오디오 신호의 하나 이상의 선형 예측 필터 계수들일 수 있다. 몇몇 실시예들에서, 하나 이상의 제 1 오디오 신호 계수들은 예를 들어, 인코딩된 오디오 신호의 하나 이상의 선형 예측 필터 계수들일 수 있다.
- [0349] 일실시예에 따라, 하나 이상의 잡음 계수들은 예를 들어, 인코딩된 오디오 신호의 배경 잡음을 나타내는 하나 이상의 선형 예측 필터 계수들일 수 있다. 일실시예에서, 하나 이상의 선형 예측 필터 계수들은 예를 들어, 배경 잡음의 스펙트럼 형태를 표현할 수 있다.
- [0350] 일실시예에서, 계수 생성기는 예를 들어, 하나 이상의 제 2 오디오 신호 부분들이 재구성된 오디오 신호의 하나 이상의 선형 예측 필터 계수들이 되거나, 하나 이상의 제 1 오디오 신호 계수들이 재구성된 오디오 신호의 하나 이상의 이미턴스 스펙트럼 쌍들(immittance spectral pairs)이 되도록 하나 이상의 제 2 오디오 신호 부분들을 결정하도록 구성될 수 있다.
- [0351] 일실시예에 따라, 계수 생성기는 예를 들어, 수학적  $f_{current}[i]=\alpha \cdot f_{last}[i]+(1-\alpha) \cdot pt_{mean}[i]$  을 적용함으로써 하나 이상의 제 2 오디오 신호 계수들을 생성하도록 구성될 수 있고,
- [0352] 여기서  $f_{current}[i]$ 는 하나 이상의 제 2 오디오 신호 계수들 중 하나를 나타내고,  $f_{last}[i]$ 는 하나 이상의 제 1 오디오 신호 계수들 중 하나를 나타내고,  $pt_{mean}[i]$ 는 하나 이상의 잡음 계수들 중 하나이고,  $\alpha$ 는  $0 \leq \alpha \leq 1$ 을 갖는 실수이고,  $i$ 는 인덱스이다. 일실시예에서,  $0 < \alpha < 1$ 이다.
- [0353] 일실시예에 따라,  $f_{last}[i]$ 는 인코딩된 오디오 신호의 선형 예측 필터 계수를 나타내고,  $f_{current}[i]$ 는 재구성된 오디오 신호의 선형 예측 필터 계수를 나타낸다.
- [0354] 일실시예에서,  $pt_{mean}[i]$ 는 예를 들어, 인코딩된 오디오 신호의 배경 잡음을 나타낼 수 있다.
- [0355] 일실시예에서, 계수 생성기는 예를 들어, 하나 이상의 프레임들의 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 인코딩된 오디오 신호의 잡음 스펙트럼을 결정함으로써 하나 이상의 잡음 계수들을 결정하도록 구성될 수 있다.
- [0356] 일실시예에 따라, 계수 생성기는 예를 들어, 배경 잡음 스펙트럼을 결정하기 위해 신호 스펙트럼 상의 최소 통계 접근법을 이용함으로써, 그리고 배경 잡음 스펙트럼으로부터 배경 잡음 형태를 나타내는 LPC 계수들을 계산함으로써 배경 잡음을 나타내는 LPC 계수들을 결정하도록 구성될 수 있다.



- [0357] 더욱이, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 방법이 제공된다. 방법은
- [0358] - 하나 이상의 프레임들을 수신하는 단계,
- [0359] - 하나 이상의 프레임들의 현재 프레임이 수신되면, 그리고 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 하나 이상의 제 1 오디오 신호 계수들을 결정하는 단계로서, 상기 하나 이상의 제 1 오디오 신호 계수들은 인코딩된 오디오 신호의 특징을 나타내고, 하나 이상의 잡음 계수들은 인코딩된 오디오 신호의 배경 잡음을 나타내는, 하나 이상의 제 1 오디오 신호 계수들을 결정하는 단계,
- [0360] - 현재 프레임이 수신되지 않으면, 또는 수신되는 현재 프레임이 붕괴되면, 하나 이상의 제 1 오디오 신호 계수들에 따라, 그리고 하나 이상의 잡음 계수들에 따라 하나 이상의 제 2 오디오 신호 계수들을 생성하는 단계,
- [0361] - 현재 프레임이 수신되면, 그리고 수신되는 현재 프레임이 붕괴되지 않으면, 하나 이상의 제 1 오디오 신호 계수들에 따라 재구성된 오디오 신호의 제 1 부분을 재구성하는 단계, 및
- [0362] - 현재 프레임이 수신되지 않으면, 또는 수신되는 현재 프레임이 붕괴되면, 하나 이상의 제 2 오디오 신호 계수들에 따라 재구성된 오디오 신호의 제 2 부분을 재구성하는 단계를
- [0363] 포함한다.
- [0364] 더욱이, 컴퓨터 또는 신호 프로세서 상에서 실행될 때 전술한 방법을 구현하기 위한 컴퓨터 프로그램이 제공된다.
- [0365] 페이드 아웃 동안 안락 잡음의 스펙트럼 형태를 추적하고 적용하기 위한 공통 수단을 갖는 것은 여러 장점들을 갖는다. 양쪽의 코어 코덱들에 대해 유사하게 이루어질 수 있도록 스펙트럼 형태를 추적하고 적용함으로써 간단한 공통 접근법을 허용한다. CELT는 스펙트럼 도메인에서의 에너지들의 대역 방식의 추적과, 스펙트럼 도메인에서의 스펙트럼 형태의 대역 방식의 형성만을 가르치고, 이것은 CELP 코어에 대해서는 가능하지 않다.
- [0366] 이와 대조적으로, 종래 기술에서, 버스트 손실들 동안 도입된 안락 잡음의 스펙트럼 형태는 완전히 정적이거나, 부분적으로 정적이고, 스펙트럼 형태의 단기간 평균에 부분적으로 적응성이고(G.718[ITU08a]에서 실현됨), 일반적으로 패킷 손실 이전에 신호에서의 배경 잡음에 매칭하지 않을 것이다. 안락 잡음 특징들의 이러한 미스매치는 교란될 수 있다. 종래 기술에 따라, 오프라인 트레이닝된(정적) 배경 잡음형태가 이용될 수 있고, 특정 신호들에 대해 쾌적한 사운드일 수 있지만, 다른 것들, 예를 들어 사무실 잡음과 완전히 상이한 자동차 잡음 사운드들에 대해 덜 쾌적할 수 있다.
- [0367] 더욱이, 종래 기술에서, 이전에 수신된 프레임들의 스펙트럼 형태의 단기간 평균에 대한 적응이 이용될 수 있고, 신호 특징들을 이전에 수신된 신호, 반드시 그럴 필요는 없지만, 배경 잡음 특징들에 더 가까이 다가갈 수 있다. 종래 기술에서, 스펙트럼 도메인에서 대역 방식으로 스펙트럼 형태를 추적하는 것은(CELT[IET12]에서 실현됨) MDCT 도메인 기반의 코어(TCX)뿐 아니라 ACELP 기반의 코어를 이용하여 스위칭된 코덱에 대해 적용가능하지 않다. 전술한 실시예들은 이에 따라 종래 기술에 비해 유리하다.
- [0368] 더욱이, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치가 제공된다. 장치는 인코딩된 오디오 신호의 오디오 신호 스펙트럼의 복수의 오디오 신호 샘플들에 대한 정보를 포함하는 하나 이상의 프레임들을 수신하기 위한 수신 인터페이스, 및 재구성된 오디오 신호를 생성하기 위한 프로세서를 포함한다. 프로세서는, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 현재 프레임이 수신 인터페이스에 의해 수신되지 않고 붕괴되면, 변형된 스펙트럼을 목표 스펙트럼에 페이딩함으로써 재구성된 오디오 신호를 생성하도록 구성되고, 여기서 변형된 스펙트럼은 복수의 변형된 신호 샘플들을 포함하고, 변형된 스펙트럼의 각 변형된 신호 샘플들에 대해, 상기 변형된 신호 샘플의 절대값은 오디오 신호 스펙트럼의 오디오 신호 샘플들 중 하나의 오디오 신호 샘플의 절대값과 동일하다. 더욱이, 프로세서는, 하나 이상의 프레임들의 현재 프레임이 수신 인터페이스에 의해 수신되면, 그리고 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 변형된 스펙트럼을 목표 스펙트럼에 페이딩하지 않도록 구성된다.
- [0369] 일실시예에 따라, 목표 스펙트럼은 예를 들어, 잡음형 스펙트럼일 수 있다.
- [0370] 일실시예에서, 잡음형 스펙트럼은 예를 들어, 백색 잡음을 나타낼 수 있다.
- [0371] 일실시예에 따라, 잡음형 스펙트럼은 예를 들어, 성형(shaped)될 수 있다.
- [0372] 일실시예에서, 잡음형 스펙트럼의 형태는 예를 들어, 이전에 수신된 신호의 오디오 신호 스펙트럼에 의존할 수

있다.

- [0373] 일실시예에 따라, 잡음형 스펙트럼은 예를 들어, 오디오 신호 스펙트럼의 형태에 따라 성형될 수 있다.
- [0374] 일실시예에서, 프로세서는 예를 들어, 잡음형 스펙트럼을 성형하기 위해 경사 인자를 이용할 수 있다.
- [0375] 일실시예에 따라, 프로세서는 예를 들어, 수학적 식  $\text{shaped\_noise}[i] = \text{noise} * \text{power}(\text{tilt\_factor}, i/N)$ 을 이용할 수 있고,
- [0376] 여기서 N은 샘플들의 개수를 표시하고, i는 인덱스이고,  $0 \leq i < N$ 이고,  $\text{tilt\_factor} > 0$ 이고, power는 멱함수이다.
- [0377]  $\text{power}(x, y)$ 는  $x^y$ 를 표시하고,
- [0378]  $\text{power}(\text{tilt\_factor}, i/N)$ 은  $\text{tilt\_factor}^{\frac{i}{N}}$ 을 표시한다.
- [0379] tilt\_factor가 1보다 작으면, 이것은 증가하는 i를 갖는 감쇄를 의미한다. tilt\_factor가 1보다 크면, 이것은 증가하는 i를 갖는 증폭을 의미한다.
- [0380] 다른 실시예에 따라, 프로세서는 예를 들어, 수학적 식  $\text{shaped\_noise}[i] = \text{noise} * (1 + i/(N-1) * (\text{tilt\_factor} - 1))$ 을 이용할 수 있고,
- [0381] 여기서 N은 샘플들의 개수를 표시하고, i는 인덱스이고,  $0 \leq i < N$ 이고,  $\text{tilt\_factor} > 0$ 이다.
- [0382] tilt\_factor가 1보다 작으면, 이것은 증가하는 i를 갖는 감쇄를 의미한다. tilt\_factor가 1보다 크면, 이것은 증가하는 i를 갖는 증폭을 의미한다.
- [0383] 일실시예에 따라, 프로세서는 예를 들어, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 오디오 신호 스펙트럼의 오디오 신호 샘플들 중 하나 이상의 오디오 신호 샘플의 부호를 변경함으로써 변형된 스펙트럼을 생성하도록 구성될 수 있다.
- [0384] 일실시예에서, 오디오 신호 스펙트럼의 오디오 신호 샘플들 각각은 예를 들어, 실수에 의해 표현되지만, 허수에 의해 표현되지 않을 수 있다.
- [0385] 일실시예에 따라, 오디오 신호 스펙트럼의 오디오 신호 샘플들은 예를 들어, 변형된 이산 코사인 변환 도메인으로 표현될 수 있다.
- [0386] 다른 실시예에서, 오디오 신호 스펙트럼의 오디오 신호 샘플들은 예를 들어, 변형된 이산 사인 변환 도메인으로 표현될 수 있다.
- [0387] 일실시예에 따라, 프로세서는 예를 들어, 제 1 또는 제 2 값을 무작위로 또는 의사-무작위로(pseudo-randomly) 출력하는 랜덤 부호 함수를 이용함으로써 변형된 스펙트럼을 생성하도록 구성될 수 있다.
- [0388] 일실시예에서, 프로세서는 예를 들어, 감쇄 인자를 후속적으로 감소시킴으로써 변형된 스펙트럼을 목표 스펙트럼에 페이딩하도록 구성될 수 있다.
- [0389] 일실시예에 따라, 프로세서는 예를 들어, 감쇄 인자를 후속적으로 증가시킴으로써 변형된 스펙트럼을 목표 스펙트럼에 페이딩하도록 구성될 수 있다.
- [0390] 일실시예에서, 현재 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 수신 인터페이스에 의해 수신되는 현재 프레임이 붕괴되면, 프로세서는 예를 들어, 수학적 식  $x[i] = (1 - \text{cum\_damping}) * \text{noise}[i] + \text{cum\_damping} * \text{random\_sign}() * x\_old[i]$ 을 이용함으로써 재구성된 오디오 신호를 생성하도록 구성될 수 있고,
- [0391] 여기서 i는 인덱스이고, x[i]는 재구성된 오디오 신호의 샘플을 나타내고, cum\_damping은 감쇄 인자이고, x\_old[i]는 인코딩된 오디오 신호의 오디오 신호 스펙트럼의 오디오 신호 샘플들 중 하나를 나타내고, random\_sign()은 1 또는 -1로 리턴(return)하고, noise는 목표 스펙트럼을 나타내는 확률 벡터(random vector)이다.
- [0392] 일실시예에서, 상기 확률 벡터 noise는 예를 들어, 확률 벡터의 이차 평균(quadratic mean)이 수신 인터페이스에 의해 마지막으로 수신되는 프레임들 중 하나의 프레임에 의해 포함되는 인코딩된 오디오 신호의 스펙트럼의 이차 평균과 유사하도록 스케일링될 수 있다.

- [0393] 일반적인 실시예에 따라, 프로세서는, 예를 들어, 확률 벡터의 이차 평균이 수신 인터페이스에 의해 마지막으로 수신되는 프레임들 중 하나의 프레임에 의해 포함되는 인코딩된 오디오 신호의 스펙트럼의 이차 평균과 유사하도록 스케일링되는 확률 벡터를 이용함으로써 재구성된 오디오 신호를 생성하도록 구성될 수 있다.
- [0394] 더욱이, 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 방법이 제공된다. 방법은
- [0395] - 인코딩된 오디오 신호의 오디오 신호 스펙트럼의 복수의 오디오 신호 샘플들에 대한 정보를 포함하는 하나 이상의 프레임들을 수신하는 것, 및
- [0396] - 재구성된 오디오 신호를 생성하는 것을
- [0397] 포함한다.
- [0398] 재구성된 오디오 신호를 생성하는 것은, 현재 프레임이 수신되지 않으면 또는 현재 프레임이 수신되지만 붕괴되면, 변형된 스펙트럼을 목표 스펙트럼에 페이딩함으로써 수행되고, 변형된 스펙트럼은 복수의 변형된 신호 샘플들을 포함하고, 변형된 스펙트럼의 변형된 신호 샘플들 각각에 대해, 상기 변형된 신호 샘플의 절대값은 오디오 신호 스펙트럼의 오디오 신호 샘플들 중 하나의 오디오 신호 샘플의 절대값과 동일하다. 변형된 스펙트럼은, 하나 이상의 프레임들의 현재 프레임이 수신되면, 그리고 수신되는 현재 프레임이 붕괴되지 않으면, 백색 잡음 스펙트럼에 페이딩되지 않는다.
- [0399] 더욱이, 컴퓨터 또는 신호 프로세서 상에서 실행될 때 진술한 방법을 구현하기 위한 컴퓨터 프로그램이 제공된다.
- [0400] 실시예들은 FDNS 적용(FDNS= Frequency Domain Noise Substitution) 이전에 백색 잡음으로의 페이드 MDCT 스펙트럼을 실현한다.
- [0401] 종래 기술에 따라, ACELP 기반의 코덱들에서, 혁신 코드북은 확률 벡터(예를 들어, 잡음)로 교체된다. 실시예들에서, 혁신 코드북을 확률 벡터(예를 들어, 잡음)로 교체하는 것으로 구성된 ACELP 접근법은 TCX 디코더 구조로 채택된다. 여기서, 혁신 코드북의 등가물은, 일반적으로 비트스트림 내에서 수신되고 FDNS에 공급된 MDCT 스펙트럼이다.
- [0402] 고전적인 MDCT 은닉 접근법은 간단히 이러한 스펙트럼을 그 자체로 반복하거나 특정 무작위화(randomization) 프로세스를 적용하기 위한 것이고, 이것은 기본적으로 마지막으로 수신된 프레임의 스펙트럼 형태를 연장한다 [LS01]. 이것은, 단기간 스펙트럼 형태가 연장되어, 이것이 배경 잡음형이 아닌 반복적인 금속 사운드를 종종 초래하여, 안락 잡음으로서 사용될 수 없다는 결함을 갖는다.
- [0403] 제안된 방법을 이용하여, 단기간 스펙트럼 성형은 FDNS 및 TCX, LTP에 의해 수행되고, 오랜 실행시 스펙트럼 성형은 FDNS에 의해서만 수행된다. FDNS에 의한 성형은 단기간 스펙트럼 형태로부터 배경 잡음의 추적된 장기 스펙트럼 형태로 페이딩되고, TCX LTP는 0으로 페이딩된다.
- [0404] FDNS 계수들을 추적된 배경 잡음 계수들로 페이딩하는 것은 오랜 버스트 프레임 손실들의 경우에 쾌적한 배경 잡음을 달성하기 위해 오랜 실행시 목표로 정해져야 하는(targeted) 스펙트럼 배경 엔벨로프와 마지막으로 양호한 스펙트럼 엔벨로프 사이에서 매끄러운 전이를 갖는 것을 초래한다.
- [0405] 이와 대조적으로, 종래 기술에 따라, 변환 기반의 코덱들에 대해, 잡음형 은닉은 주파수 도메인에서 프레임 반복 또는 잡음 교체에 의해 수행된다[LS01]. 종래 기술에서, 잡음 교체는 일반적으로 스펙트럼 빈들의 부호 스크램블링에 의해 수행된다. 종래 기술에서 TCX(주파수 도메인) 부호 스크램블링이 은닉 동안 사용되면, 마지막으로 수신된 MDCT 계수들은 재사용되고, 각 부호는, 스펙트럼이 시간 도메인으로 역 변환되기 전에 무작위화된다. 종래 기술의 이러한 절차의 결함은, 연속적으로 손실된 프레임들에 대해, 동일한 스펙트럼이 상이한 부호 무작위화들 및 범용 감쇄만을 가지고 되풀이하여 사용된다. 거친 시간 그리드(coarse time grid) 상에서 시간이 지남에 따라 스펙트럼 엔벨로프를 탐색할 때, 엔벨로프가 연속적인 프레임 손실 동안 대략 동일한데, 이는 대역 에너지들이 프레임 내에서 서로에 대해 상대적으로 일정하게 유지되고 단지 범용적으로 감쇄되기 때문이라는 것을 알 수 있다. 사용된 코딩 시스템에서, 종래 기술에 따라, 스펙트럼 값들은 원래 스펙트럼을 복구하기 위해 FDNS를 이용하여 처리된다. 이것은, MDCT 스펙트럼을 특정 스펙트럼 엔벨로프(FDNS 계수들을 이용하여, 예를 들어 현재 배경 잡음을 기재하여)로 페이딩하기를 원하면, 그 결과는 FDNS 계수들에 의존할 뿐 아니라, 부호 스크램블링된 이전에 디코딩된 스펙트럼에 의존한다. 진술한 실시예들은 종래 기술의 이들 단점들을 극복한다.
- [0406] 실시예들은, 이를 FDNS 처리에 공급하기 전에 부호 스크램블링에 사용된 스펙트럼을 백색 잡음에 페이딩할 필요

가 있다는 발견에 기초한다. 그렇지 않으면, 출력된 스펙트럼은 FDNS 처리에 사용된 목표로 정해진 엔벨로프에 매칭하지 않을 것이다.

[0407] 실시예들에서, 동일한 페이딩 속도는 백색 잡음 페이딩에 대해서 LTP 이득 페이딩에 사용된다.

[0408] 다음에서, 본 발명의 실시예들은 도면들을 참조하여 더 구체적으로 기재된다.

**도면의 간단한 설명**

[0409] 도 1a는 실시예에 따라 오디오 신호를 디코딩하기 위한 장치를 도시한 도면.

도 1b는 다른 실시예에 따라 오디오 신호를 디코딩하기 위한 장치를 도시한 도면.

도 1c는 다른 실시예에 따라 오디오 신호를 디코딩하기 위한 장치를 도시한 도면으로서, 장치는 제 1 및 제 2 집합 유닛을 더 포함하는, 도면.

도 1d는 추가 실시예에 따라 오디오 신호를 디코딩하기 위한 장치를 도시한 도면으로서, 장치는 더욱이 지연 버퍼를 포함하는 장기간 예측 유닛을 포함하는, 도면.

도 2는 G.718의 디코더 구조를 도시한 도면.

도 3은, G.722의 페이드-아웃 인자가 클래스 정보에 의존하는 시나리오를 도시한 도면.

도 4는 선형 회귀를 이용하여 진폭 예측을 위한 접근법을 도시한 도면.

도 5는 억제된-에너지 래핑된 변환(CELT)의 버스트 손실 행위를 도시한 도면.

도 6은 에러 없는 동작 모드 동안 디코더에서의 실시예에 따른 배경 잡음 레벨 추적을 도시한 도면.

도 7은 실시예에 따라 LPC 합성 및 디엠퍼시스의 이득 도출을 도시한 도면.

도 8은 실시예에 따라 패킷 손실 동안 안락 잡음 레벨 적용을 도시한 도면.

도 9는 실시예에 따라 ACELP 은닉 동안 진보된 고역 통과 이득 보상을 도시한 도면.

도 10은 실시예에 따라 은닉 동안 LTP 피드백 루프의 결합 해제를 도시한 도면.

도 11은 실시예에 따라 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한 도면.

도 12는 다른 실시예에 따라 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한 도면.

도 13은 추가 실시예에 따라 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한 도면.

도 14는 다른 실시예에 따라 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한 도면.

**발명을 실시하기 위한 구체적인 내용**

[0410] 도 1a는 실시예에 따라 오디오 신호를 디코딩하기 위한 장치를 도시한다.

[0411] 장치는 수신 인터페이스(110)를 포함한다. 수신 인터페이스는 복수의 프레임들을 수신하도록 구성되고, 수신 인터페이스(110)는 복수의 프레임들의 제 1 프레임을 수신하도록 구성되고, 상기 제 1 프레임은 오디오 신호의 제 1 오디오 신호 부분을 포함하고, 상기 제 1 오디오 신호 부분은 제 1 도메인에서 표현된다. 더욱이, 수신 인터페이스(110)는 복수의 프레임들의 제 2 프레임을 수신하도록 구성되고, 상기 제 2 프레임은 오디오 신호의 제 2 오디오 신호 부분을 포함한다.

[0412] 더욱이, 장치는 제 2 신호 부분 정보를 얻기 위해 제 2 오디오 신호 부분, 또는 제 2 오디오 신호 부분으로부터 도출된 값 또는 신호를 제 2 도메인으로부터 추적 도메인으로 변환하기 위한 변환 유닛(120)을 포함하고, 제 2 도메인은 제 1 도메인과 상이하고, 추적 도메인은 제 2 도메인과 상이하고, 추적 도메인은 제 1 도메인과 동일하거나 상이하다.

[0413] 더욱이, 장치는 잡음 레벨 추적 유닛(130)을 포함하고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 1

신호 부분 정보를 수신하도록 구성되고, 제 1 신호 부분 정보는 제 1 오디오 신호 부분에 의존하고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성되고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 1 신호 부분 정보에 따라, 그리고 추적 도메인에서 표현되는 제 2 신호 부분 정보에 따라 잡음 레벨 정보를 결정하도록 구성된다.

- [0414] 더욱이, 장치는, 복수의 프레임들의 제 3 프레임이 수신 인터페이스에 의해 수신되지 않고, 붕괴되면, 잡음 레벨 정보에 따라 오디오 신호의 제 3 오디오 신호 부분을 재구성하기 위한 재구성 유닛을 포함한다.
- [0415] 예를 들어, 제 1 및/또는 제 2 오디오 신호 부분에 관해, 제 1 및/또는 제 2 오디오 신호 부분은 예를 들어, 하나 이상의 스피커들(loudspeakers)을 위한 하나 이상의 스피커 신호들을 생성하기 위해 하나 이상의 처리 유닛들(미도시)에 공급될 수 있어서, 제 1 및/또는 제 2 오디오 신호 부분에 의해 포함되는 수신된 사운드 정보는 재생될 수 있다.
- [0416] 하지만, 더욱이, 제 1 및 제 2 오디오 신호 부분은 또한 예를 들어, 후속 프레임들이 수신기에 도달하지 않는 경우에, 또는 후속 프레임들이 에러가 있는 경우에, 은닉에 사용된다.
- [0417] 특히, 본 발명은, 잡음 레벨 추적이 본 명세서에서 "추적 도메인"으로 언급되는 공통 도메인에서 수행되어야 한다는 발견에 기초한다. 추적 도메인은 예를 들어, 여기 도메인, 예를 들어, 신호가 AMR-WB 및 AMR-WB+([3GP12a], [3GP12b], [3GP09a], [3GP09b], [3GP09c]를 참조)에 기재된 바와 같이 LPC들(LPC= Linear Predictive Coefficient) 또는 ISP들(ISP= Immittance Spectral Pair)에 의해 표현되는 도메인일 수 있다. 단일 도메인에서 잡음 레벨을 추적하는 것은 특히, 신호가 제 1 도메인에서의 제 1 표현과 제 2 도메인에서의 제 2 표현 사이를 스위칭할 때(예를 들어, 신호 표현이 ACELP로부터 TCX로, 또는 그 반대로 스위칭할 때) 얼라이싱 효과가 회피되는 장점을 갖는다.
- [0418] 변환 유닛(120)에 관해, 제 2 오디오 신호 부분 자체, 또는 제 2 오디오 신호 부분으로부터 도출된 신호(예를 들어, 제 2 오디오 신호는 도출된 신호를 얻도록 처리되었다), 또는 제 2 오디오 신호 부분으로부터 도출된 값(예를 들어, 제 2 오디오 신호 부분은 도출된 값을 얻도록 처리되었다) 중 어느 하나가 변환된다.
- [0419] 제 1 오디오 신호 부분에 관해, 몇몇 실시예들에서, 제 1 오디오 신호 부분은 추적 도메인으로 처리 및/또는 변환될 수 있다.
- [0420] 하지만, 다른 실시예들에서, 제 1 오디오 신호 부분은 추적 도메인에서 미리 표현될 수 있다.
- [0421] 몇몇 실시예들에서, 제 1 신호 부분 정보는 제 1 오디오 신호 부분과 동일하다. 다른 실시예들에서, 제 1 신호 부분 정보는 예를 들어, 제 1 오디오 신호 부분에 따른 집합된 값이다.
- [0422] 이제, 처음으로, 안락 잡음 레벨로의 페이드-아웃이 더 구체적으로 고려된다.
- [0423] 기재된 페이드-아웃 접근법은 예를 들어, xHE-AAC[NMR+12](xHE-AAC=Extended High Efficiency AAC)의 저지연 버전으로 구현될 수 있고, 이것은 프레임 마다 ACELP(음성)과 MDCT(음악/잡음) 코딩 사이에서 끊임없이 스위칭할 수 있다.
- [0424] 매끄러운 페이드-아웃을 패킷 손실 동안 적절한 안락 잡음 레벨에 적용하도록 추적 도메인, 예를 들어, 여기 도메인에서 공통 레벨 추적에 관해, 그러한 안락 잡음 레벨은 정상 디코딩 프로세스 동안 식별될 필요가 있다. 예를 들어, 배경 잡음과 유사한 잡음 레벨이 가장 안락하다고 간주될 수 있다. 따라서, 배경 잡음 레벨이 도출될 수 있고, 정상 디코딩 동안 일정하게 업데이트될 수 있다.
- [0425] 본 발명은, 스위칭된 코어 코덱(예를 들어, ACELP 및 TCX)을 가질 때, 선택된 코어 코더와 독립적인 공통 배경 잡음 레벨을 고려하는 것이 특히 적합하다는 발견에 기초한다.
- [0426] 도 6은 에러 없는 동작 모드 동안, 예를 들어 정상 디코딩 동안, 디코더에서 바람직한 실시예에 따른 배경 잡음 레벨 추적을 도시한다.
- [0427] 추적 자체는 예를 들어, 최소 통계 접근법을 이용하여 수행될 수 있다([Mar01]을 참조).
- [0428] 이러한 추적된 배경 잡음 레벨은 예를 들어, 전술한 잡음 레벨 정보로서 고려될 수 있다.
- [0429] 예를 들어, 최소 통계 잡음 추정은 문헌에 제공된다: 라이너 마르틴(Rainer Martin)의, 최적의 평활화 및 최소 통계에 기초한 잡음 전력 스펙트럼 밀도 추정(Noise power spectral density estimation based on optimal smoothing and minimum statistics), IEEE Transactions on Audio Processing 9 (2001), no.5, 504-

512"[Mar01]은 배경 잡음 레벨 추적을 위해 이용될 수 있다.

- [0430] 이에 대응하여, 몇몇 실시예들에서, 잡음 레벨 추적 유닛(130)은 최소 통계 접근법을 적용함으로써, 예를 들어, [Mar01]의 최소 통계 잡음 추정을 이용함으로써, 잡음 레벨 정보를 결정하도록 구성된다.
- [0431] 후속하여, 이러한 추적 접근법의 몇몇 고려사항들 및 세부사항들이 기재된다.
- [0432] 레벨 추정에 관해, 배경은 잡음-형이 되는 것으로 가정된다. 따라서, LPC에 의해 취득된 추적 전경(foreground) 음조 성분들을 피하기 위해 여기 도메인에서의 레벨 추적을 수행하는 것이 바람직하다. 예를 들어, ACELP 잡음 충전은 또한 여기 도메인에서의 배경 잡음 레벨을 이용할 수 있다. 여기 도메인에서의 추적을 통해, 배경 잡음 레벨의 하나의 단일 추적만이 2가지 목적들을 다룰 수 있고, 이것은 계산 복잡도를 절감한다. 바람직한 실시예에서, 추적은 ACELP 여기 도메인에서 수행된다.
- [0433] 도 7은 실시예에 따라 LPC 합성 및 디엠퍼시스의 이득 도출을 도시한다.
- [0434] 레벨 도출에 관해, 레벨 도출은 예를 들어, 시간 도메인 또는 여기 도메인, 또는 임의의 다른 적합한 도메인에서 수행될 수 있다. 레벨 도출 및 레벨 추적에 대한 도메인들이 상이하면, 예를 들어, 이득 보상이 필요할 수 있다.
- [0435] 바람직한 실시예에서, ACELP에 대한 레벨 도출은 여기 도메인에서 수행된다. 따라서, 이득 보상은 요구되지 않는다.
- [0436] TCX에 대해, 이득 보상은 예를 들어, 도출된 레벨을 ACELP 여기 도메인으로 조정하는데 필요할 수 있다.
- [0437] 바람직한 실시예에서, TCX에 대한 레벨 도출은 시간 도메인에서 발생한다. 관리가능한 이득 보상은 이러한 접근법을 위해 발견되었다: LPC 합성 및 디엠퍼시스에 의해 도입된 이득은 도 7에 도시된 바와 같이 도출되고, 도출된 레벨은 이러한 이득으로 나누어진다.
- [0438] 대안적으로, TCX에 대한 레벨 도출은 TCX 여기 도메인에서 수행될 수 있다. 하지만, TCX 여기 도메인과 ACELP 여기 도메인 사이의 이득 보상은 너무 복잡한 것으로 간주되었다.
- [0439] 따라서, 도 1a로 다시 돌아가면, 몇몇 실시예들에서, 제 1 오디오 신호 부분은 제 1 시간 도메인으로서 시간 도메인에서 표현된다. 변환 유닛(120)은 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 도메인인 여기 도메인으로부터 추적 도메인인 시간 도메인으로 변환하도록 구성된다. 그러한 실시예들에서, 잡음 레벨 추적 유닛(130)은 추적 도메인으로서 시간 도메인에서 표현되는 제 1 신호 부분 정보를 수신하도록 구성된다. 더욱이, 잡음 레벨 추적 유닛(130)은 추적 도메인으로서 시간 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성된다.
- [0440] 다른 실시예들에서, 제 1 오디오 신호 부분은 제 1 도메인으로서 여기 도메인에서 표현된다. 변환 유닛(120)은 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 도메인인 시간 도메인으로부터 추적 도메인인 여기 도메인으로 변환하도록 구성된다. 그러한 실시예들에서, 잡음 레벨 추적 유닛(130)은 추적 도메인으로서 여기 도메인에서 표현되는 제 1 신호 부분 정보를 수신하도록 구성된다. 더욱이, 잡음 레벨 추적 유닛(130)은 추적 도메인으로서 여기 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성된다.
- [0441] 일실시예에서, 제 1 오디오 신호 부분은 예를 들어, 제 1 도메인으로서 여기 도메인에서 표현될 수 있고, 잡음 레벨 추적 유닛(130)은 예를 들어, 제 1 신호 부분 정보를 수신하도록 구성될 수 있고, 상기 제 1 신호 부분 정보는 추적 도메인인 FFT 도메인에서 표현되고, 상기 제 1 신호 부분 정보는 여기 도메인에서 표현되는 상기 제 1 오디오 신호 부분에 의존하고, 변환 유닛(120)은 예를 들어, 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 도메인인 시간 도메인으로부터 추적 도메인인 FFT 도메인으로 변환하도록 구성될 수 있고, 잡음 레벨 추적 유닛(130)은 예를 들어, FFT 도메인에서 표현되는 제 2 오디오 신호 부분을 수신하도록 구성될 수 있다.
- [0442] 도 1b는 다른 실시예에 따른 장치를 도시한다. 도 1b에서, 도 1a의 변환 유닛(120)은 제 1 변환 유닛(120)이고, 도 1a의 재구성 유닛(140)은 제 1 재구성 유닛(140)이다. 장치는 제 2 변환 유닛(121) 및 제 2 재구성 유닛(141)을 더 포함한다.
- [0443] 제 2 변환 유닛(121)은, 복수의 프레임들의 제 4 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보를 추적 도메인으로부터 제 2 도메인으로 변환하도록 구성된다.

- [0444] 더욱이, 제 2 재구성 유닛(141)은, 복수의 프레임들의 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스에 의해 수신되지만, 붕괴되면, 제 2 도메인에서 표현되는 잡음 레벨 정보에 따라 오디오 신호의 제 4 오디오 신호 부분을 재구성하도록 구성된다.
- [0445] 도 1c는 다른 실시예에 따라 오디오 신호를 디코딩하기 위한 장치를 도시한다. 장치는 제 1 오디오 신호 부분에 따라 제 1 집합된 값을 결정하기 위한 제 1 집합 유닛(150)을 더 포함한다. 더욱이, 도 1c의 장치는 제 2 오디오 신호 부분에 따라 제 2 오디오 신호 부분으로부터 도출된 값으로서 제 2 집합된 값을 결정하기 위한 제 2 집합 유닛(160)을 더 포함한다. 도 1c의 실시예에서, 잡음 레벨 추적 유닛(130)은 추적 도메인에서 표현되는 제 1 신호 부분 정보로서 제 1 집합된 값을 수신하도록 구성되고, 잡음 레벨 추적 유닛(130)은 추적 도메인에서 표현되는 제 2 신호 부분 정보로서 제 2 집합된 값을 수신하도록 구성된다. 잡음 레벨 추적 유닛(130)은 추적 도메인에서 표현되는 제 1 집합된 값에 따라, 그리고 추적 도메인에서 표현되는 제 2 집합된 값에 따라 잡음 레벨 정보를 결정하도록 구성된다.
- [0446] 실시예에서, 제 1 집합 유닛(150)은, 제 1 집합된 값이 제 1 오디오 신호 부분, 또는 제 1 오디오 신호 부분으로부터 도출된 신호의 제공 평균 제공근을 나타내도록 제 1 집합된 값을 결정하도록 구성된다. 더욱이, 제 2 집합 유닛(160)은, 제 2 집합된 값이 제 2 오디오 신호 부분, 또는 제 2 오디오 신호 부분으로부터 도출된 신호의 제공 평균 제공근을 나타내도록 제 2 집합된 값을 결정하도록 구성된다.
- [0447] 도 6은 추가 실시예에 따라 오디오 신호를 디코딩하기 위한 장치를 도시한다.
- [0448] 도 6에서, 배경 레벨 추적 유닛(630)은 도 1a에 따른 잡음 레벨 추적 유닛(130)을 구현한다.
- [0449] 더욱이, 도 6에서, RMS 유닛(650)(RMS = root mean square)은 제 1 집합 유닛이고, RMS 유닛(660)은 제 2 집합 유닛이다.
- [0450] 몇몇 실시예들에 따라, 도 1a, 도 1b 및 도 1c의 (제 1) 변환 유닛(120)은 제 2 오디오 신호 부분으로부터 도출된 값 상에 이득 값(x)을 적용함으로써, 예를 들어, 제 2 오디오 신호 부분으로부터 도출된 값을 이득 값(x)으로 나눔으로써, 제 2 오디오 신호 부분으로부터 도출된 값을 제 2 도메인으로부터 추적 도메인으로 변환하도록 구성된다. 다른 실시예들에서, 예를 들어, 이득 값이 곱해질 수 있다.
- [0451] 몇몇 실시예들에서, 이득 값(x)은 예를 들어, 선형 예측 코딩 합성에 의해 도입된 이득을 나타내거나, 이득 값(x)은 예를 들어, 선형 예측 코딩 합성 및 디엠퍼시스에 의해 도입된 이득을 나타낼 수 있다.
- [0452] 도 6에서, 유닛(622)은 선형 예측 코딩 합성 및 디엠퍼시스에 의해 도입된 이득을 나타내는 값(x)을 제공한다. 유닛(622)은 이후 제 2 오디오 신호 부분으로부터 도출된 값인 제 2 집합 유닛(660)에 의해 제공된 값을, 제공된 이득 값(x)으로 나눈다(예를 들어, x로 나눔으로써, 또는 값(1/x)과 곱함으로써). 따라서, 유닛들(621 및 622)을 포함하는 도 6의 유닛(620)은 도 1a, 도 1b 또는 도 1c의 제 1 변환 유닛을 구현한다.
- [0453] 도 6의 장치는 제 1 오디오 신호 부분을 갖는 제 1 프레임을 수신하고, 제 1 오디오 신호 부분은 유성음 여기 및/또는 무성음 여기이고, 추적 도메인, 도 6에서는 (ACELP) LPC 도메인에서 표현된다. 제 1 오디오 신호 부분은 시간-도메인 제 1 오디오 신호 부분 출력을 얻기 위해 처리하기 위한 LPC 합성 및 디엠퍼시스 유닛(671)에 공급된다. 더욱이, 제 1 오디오 신호 부분은 제 1 오디오 신호 부분의 제공 평균 제공근을 나타내는 제 1 값을 얻기 위해 RMS 모듈(650)에 공급된다. 이러한 제 1 값(제 1 RMS 값)은 추적 도메인에서 표현된다. 그 후에, 추적 도메인에서 표현되는 제 1 RMS 값은 잡음 레벨 추적 유닛(630)에 공급된다.
- [0454] 더욱이, 도 6의 장치는 제 2 오디오 신호 부분을 갖는 제 2 프레임을 수신하고, 제 2 오디오 신호 부분은 MDCT 스펙트럼을 포함하고, MDCT 도메인에서 표현된다. 잡음 충전(noise filling)은 잡음 충전 모듈(681)에 의해 수행되고, 주파수-도메인 잡음 성형은 주파수-도메인 잡음 성형 모듈(682)에 의해 수행되고, 시간 도메인에서의 변환은 iMDCT/OLA 모듈(683)(OLA = overlap-add)에 의해 수행되고, 장기간 예측은 장기간 예측 유닛(684)에 의해 수행된다. 장기간 예측 유닛은 예를 들어, 지연 버퍼(도 6에서는 미도시)를 포함할 수 있다.
- [0455] 제 2 오디오 신호 부분으로부터 도출된 신호는 이 후 제 2 오디오 신호 부분으로부터 도출된 그러한 신호의 제공 평균 제공근을 나타내는 제 2 값을 얻기 위해 RMS 모듈(660)에 공급된다. 이러한 제 2 값(제 2 RMS 값)은 여전히 시간 도메인에서 표현된다. 유닛(620)은 이 후 제 2 RMS 값을 시간 도메인으로부터 추적 도메인, 여기서 (ACELP) LPC 도메인으로 변환한다. 추적 도메인에서 표현되는 제 2 RMS 값은 이 후 잡음 레벨 추적 유닛(630)에 공급된다.

- [0456] 실시예들에서, 레벨 추적은 여기 도메인에서 수행되지만, TCX 페이드-아웃은 시간 도메인에서 수행된다.
- [0457] 정상 디코딩 동안, 배경 잡음 레벨이 추적되는 반면, 예를 들어, 마지막으로 수신된 신호는 레벨 방식으로 매끄럽게 페이딩되는 적절한 안락 잡음 레벨의 표시자로서 패킷 손실 동안 사용될 수 있다.
- [0458] 레벨 페이드-아웃을 추적하고 적용하기 위해 레벨을 도출하는 것은 일반적으로 서로 독립적이고, 상이한 도메인들에서 수행될 수 있다. 바람직한 실시예에서, 레벨 적용은 레벨 도출과 동일한 도메인들에서 수행되어, 이것은 ACELP에 대해 이득 보상이 필요하지 않고, TCX에 대해, 레벨 도출(도 6을 참조)에 대해 역 이득 보상이 필요하여, 동일한 이득 도출이 도 7에 도시된 바와 같이 이용될 수 있다는 동일한 이점들을 초래한다.
- [0459] 다음에서, 실시예들에 따라 LPC 합성 이득에 대한 고역 통과 필터의 영향의 보상이 기재된다.
- [0460] 도 8은 이러한 접근법을 개술한다. 특히, 도 8은 패킷 손실 동안 안락 잡음 레벨 적용을 도시한다.
- [0461] 도 8에서, 고역 통과 이득 필터 유닛(643), 곱셈 유닛(644), 페이딩 유닛(645), 고역 통과 필터 유닛(646), 페이딩 유닛(647), 및 조합 유닛(648)은 함께 제 1 재구성 유닛을 형성한다.
- [0462] 더욱이, 도 8에서, 배경 레벨 제공 유닛(631)은 잡음 레벨 정보를 제공한다. 예를 들어, 배경 레벨 제공 유닛(631)은 도 6의 배경 레벨 추적 유닛(630)으로서 동일하게 구현될 수 있다.
- [0463] 더욱이, 도 8에서, LPC 합성 및 디엠퍼시스 이득 유닛(649) 및 곱셈 유닛(641)은 함께 제 2 변환 유닛(630)을 형성한다.
- [0464] 더욱이, 도 8에서, 페이딩 유닛(642)은 제 2 재구성 유닛을 나타낸다.
- [0465] 도 8의 실시예에서, 유성음 및 무성음 여기는 개별적으로 페이딩된다: 유성음 여기는 0으로 페이딩되지만, 무성음 여기는 안락 잡음 레벨쪽으로 페이딩된다. 도 8은 고역 통과 필터를 추가로 도시하고, 이러한 고역 통과 필터는, 신호가 무성음으로서 분류될 때를 제외하고 모든 경우들에 대해 저주파수 성분들을 억제하기 위해 무성음 여기의 신호 체인(chain)에 도입된다.
- [0466] 고역 통과 필터의 영향을 모델링하는 것에 관해, LPC 합성 및 디엠퍼시스 이후의 레벨은 한번은 고역 통과 필터를 가지고 계산되고, 한번은 고역 통과 필터 없이 계산된다. 후속하여, 이들 2가지 레벨들의 비율은 적용된 배경 레벨을 변경하도록 도출되고 이용된다.
- [0467] 이것은 도 9에 의해 도시된다. 특히, 도 9는 실시예에 따라 ACELP 은닉 동안 진보된 고역 통과 이득 보상을 도시한다.
- [0468] 현재 여기 신호 대신에, 단지 간단한 임펄스는 이러한 계산을 위한 입력으로서 사용된다. 이것은 감소된 복잡도를 허용하는데, 이는 임펄스 응답이 빠르게 감쇠되고, 짧은 시간 프레임 상에서 RMS 도출이 수행될 수 있기 때문이다. 사실상, 전체 프레임 대신에 단지 하나의 서브프레임이 사용된다.
- [0469] 실시예에 따라, 잡음 레벨 추적 유닛(130)은 잡음 레벨 정보로서 안락 잡음 레벨을 결정하도록 구성된다. 재구성 유닛(140)은, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 제 3 오디오 신호 부분을 재구성하도록 구성된다.
- [0470] 실시예에 따라, 잡음 레벨 추적 유닛(130)은 잡음 레벨 정보로서 안락 잡음 레벨을 결정하도록 구성된다. 재구성 유닛(140)은, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라 제 3 오디오 신호 부분을 재구성하도록 구성된다.
- [0471] 실시예에서, 잡음 레벨 추적 유닛(130)은 잡음 레벨 스펙트럼으로부터 도출된 잡음 레벨 정보로서 안락 잡음 레벨을 결정하도록 구성되고, 상기 잡음 레벨 스펙트럼은 최소 통계 접근법을 적용함으로써 얻어진다. 재구성 유닛(140)은, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 복수의 선형 예측 계수들에 따라 제 3 오디오 신호 부분을 재구성하도록 구성된다.
- [0472] 실시예에서, (제 1 및/또는 제 2) 재구성 유닛(140, 141)은 예를 들어, 복수의 프레임들의 상기 제 3(제 4) 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 3(제 4) 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보에 따라, 그리고 제 1 오디오 신호 부분에 따라, 제 3 오디오 신호



부분을 재구성하도록 구성될 수 있다.

- [0473] 실시예에 따라, (제 1 및/또는 제 2) 재구성 유닛(140, 141)은 예를 들어, 제 1 오디오 신호 부분을 감쇄 및/또는 증폭함으로써 제 3(또는 제 4) 오디오 신호 부분을 재구성하도록 구성될 수 있다.
- [0474] 도 14는 오디오 신호를 디코딩하기 위한 장치를 도시한다. 장치는 수신 인터페이스(110)를 포함하고, 수신 인터페이스(110)는 오디오 신호의 제 1 오디오 신호 부분을 포함하는 제 1 프레임을 수신하도록 구성되고, 수신 인터페이스(110)는 오디오 신호의 제 2 오디오 신호 부분을 포함하는 제 2 프레임을 수신하도록 구성된다.
- [0475] 더욱이, 장치는 잡음 레벨 추적 유닛(130)을 포함하고, 잡음 레벨 추적 유닛(130)은 제 1 오디오 신호 부분 및 제 2 오디오 신호 부분 중 적어도 하나에 따라(이것은 제 1 오디오 신호 부분 및/또는 제 2 오디오 신호 부분에 따른다는 것을 의미한다) 잡음 레벨 정보를 결정하도록 구성되고, 잡음 레벨 정보는 추적 도메인에서 표현된다.
- [0476] 더욱이, 장치는 제 1 재구성 도메인에서, 복수의 프레임들의 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 잡음 정보 레벨에 따라 오디오 신호의 제 3 오디오 신호 부분을 재구성하기 위한 제 1 재구성 유닛(140)을 포함하고, 제 1 재구성 도메인은 추적 도메인과 상이하거나 동일하다.
- [0477] 더욱이, 장치는, 복수의 프레임들의 제 4 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 잡음 레벨 정보를 추적 도메인으로부터 제 2 재구성 도메인으로 변환하기 위한 변환 유닛(121)을 포함하고, 제 2 재구성 도메인은 추적 도메인과 상이하고, 제 2 재구성 도메인은 제 1 재구성 도메인과 상이하고,
- [0478] 더욱이, 장치는 제 2 재구성 도메인에서, 복수의 프레임들의 상기 제 4 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 4 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 제 2 재구성 도메인에서 표현되는 잡음 레벨 정보에 따라 오디오 신호의 제 4 오디오 신호 부분을 재구성하기 위한 제 2 재구성 유닛(141)을 포함한다.
- [0479] 몇몇 실시예들에 따라, 추적 도메인은 예를 들어, 추적 도메인이 시간 도메인, 스펙트럼 도메인, FFT 도메인, MDCT 도메인, 또는 여기 도메인이 되도록 이루어질 수 있다. 제 1 재구성 도메인은 예를 들어, 시간 도메인, 스펙트럼 도메인, FFT 도메인, MDCT 도메인, 또는 여기 도메인일 수 있다. 제 2 재구성 도메인은 예를 들어, 시간 도메인, 스펙트럼 도메인, FFT 도메인, MDCT 도메인, 또는 여기 도메인일 수 있다.
- [0480] 실시예에서, 추적 도메인은 예를 들어, FFT 도메인일 수 있고, 제 1 재구성 도메인은 예를 들어, 시간 도메인일 수 있고, 제 2 재구성 도메인은 예를 들어, 여기 도메인일 수 있다.
- [0481] 다른 실시예에서, 추적 도메인은 예를 들어, 시간 도메인일 수 있고, 제 1 재구성 도메인은 예를 들어, 시간 도메인일 수 있고, 제 2 재구성 도메인은 예를 들어, 여기 도메인일 수 있다.
- [0482] 실시예에 따라, 상기 제 1 오디오 신호 부분은 예를 들어, 제 1 입력 도메인에서 표현될 수 있고, 상기 제 2 오디오 신호 부분은 예를 들어, 제 2 입력 도메인에서 표현될 수 있다. 변환 유닛은 예를 들어, 제 2 변환 유닛일 수 있다. 장치는 예를 들어, 제 2 신호 부분 정보를 얻기 위해 제 2 오디오 신호 부분 또는 제 2 오디오 신호 부분으로부터 도출된 값 또는 신호를 제 2 입력 도메인으로부터 추적 도메인으로 변환하기 위한 제 1 변환 유닛을 더 포함할 수 있다. 잡음 레벨 추적 유닛은 예를 들어, 추적 도메인에서 표현되는 제 1 신호 부분 정보를 수신하도록 구성될 수 있고, 제 1 신호 부분 정보는 제 1 오디오 신호 부분에 의존하고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 2 신호 부분을 수신하도록 구성되고, 잡음 레벨 추적 유닛은 추적 도메인에서 표현되는 제 1 신호 부분 정보에 따라, 그리고 추적 도메인에서 표현되는 제 2 신호 부분 정보에 따라 잡음 레벨 정보를 결정하도록 구성된다.
- [0483] 실시예에 따라, 제 1 입력 도메인은 예를 들어, 여기 도메인일 수 있고, 제 2 입력 도메인은 예를 들어, MDCT 도메인일 수 있다.
- [0484] 다른 실시예에서, 제 1 입력 도메인은 예를 들어, MDCT 도메인일 수 있고, 제 2 입력 도메인은 예를 들어, MDCT 도메인일 수 있다.
- [0485] 예를 들어, 신호가 시간 도메인에서 표현되면, 이것은 예를 들어, 신호의 시간 도메인 샘플들에 의해 표현될 수 있다. 또는, 예를 들어, 신호가 스펙트럼 도메인에서 표현되면, 이것은 예를 들어, 신호의 스펙트럼의 스펙트럼 샘플들에 의해 표현될 수 있다.

- [0486] 실시예에서, 추적 도메인은 예를 들어, FFT 도메인일 수 있고, 제 1 재구성 도메인은 예를 들어, 시간 도메인일 수 있고, 제 2 재구성 도메인은 예를 들어, 여기 도메인일 수 있다.
- [0487] 다른 실시예에서, 추적 도메인은 예를 들어, 시간 도메인일 수 있고, 제 1 재구성 도메인은 예를 들어, 시간 도메인일 수 있고, 제 2 재구성 도메인은 예를 들어, 여기 도메인일 수 있다.
- [0488] 몇몇 실시예들에서, 도 14에 도시된 유닛들은 예를 들어, 도 1a, 도 1b, 도 1c 및 도 1d에 대해 기재된 바와 같이 구성될 수 있다.
- [0489] 특정 실시예들에 관해, 예를 들어 저 비율 모드에서, 실시예에 따른 장치는 예를 들어, 입력으로서 ACELP 프레임들을 수신할 수 있고, 이러한 ACELP 프레임들은 여기 도메인에서 표현되고, 그 후에 LPC 합성을 통해 시간 도메인으로 변환된다. 더욱이, 저 비율 모드에서, 실시예에 따른 장치는 예를 들어, 입력으로서 TCX 프레임들을 수신할 수 있고, 이러한 TCX 프레임들은 MDCT 도메인에서 표현되고, 그 후에 역 MDCT를 통해 시간 도메인으로 변환된다.
- [0490] 추적은 이 후 FFT-도메인에서 수행되고, 여기서 FFT 신호는 FFT(Fast Fourier Transform)를 수행함으로써 시간 도메인 신호로부터 도출된다. 추적은 예를 들어, 안락 잡음 스펙트럼을 얻기 위해 모든 스펙트럼 라인들에 대해 개별적인 최소 통계 접근법을 수행함으로써 수행될 수 있다.
- [0491] 은닉은 이 후 안락 잡음 스펙트럼에 기초하여 레벨 도출을 수행함으로써 수행된다. 레벨 도출은 안락 잡음 스펙트럼에 기초하여 수행된다. 시간 도메인에서의 레벨 변환은 FD TCX PLC에 대해 수행된다. 시간 도메인에서의 페이딩이 수행된다. 여기 도메인에서의 레벨 도출은 ACELP PLC 및 TD TCX PLC(ACELP 형)에 대해 수행된다. 여기 도메인에서의 페이딩은 이 후 수행된다.
- [0492] 다음의 목록은 이것을 요약한다:
- [0493] 저 비율:
- [0494] - 입력:
- [0495] - acelp(lpc 합성을 통해, 여기 도메인 -> 시간 도메인)
- [0496] - tcx(역 MDCT를 통해, mdct 도메인 -> 시간 도메인)
- [0497] - 추적:
- [0498] - FFT를 통해 시간 도메인으로부터 도출된 fft-도메인
- [0499] - 모든 스펙트럼 라인들에 대해 개별적인 최소 통계 -> 안락 잡음 스펙트럼
- [0500] - 은닉:
- [0501] - 안락 잡음 스펙트럼에 기초한 레벨 도출
- [0502] - 시간 변환으로의 레벨 변환으로서, 이것은
- [0503] - FD TCX PLC
- [0504] -> 시간 도메인에서의 페이딩을 위한 것이고,
- [0505] - 여기 도메인에서의 레벨 변환으로서, 이것은
- [0506] - ACELP PLC
- [0507] - TD TCX PLC (ACELP 형)
- [0508] -> 여기 도메인에서의 페이딩을 위한 것이다.
- [0509] 예를 들어, 고 비율 모드에서, 예를 들어, 입력으로서 TCX 프레임들을 수신할 수 있고, 이러한 TCX 프레임들은 MDCT 도메인에서 표현되고, 이 후 역 MDCT를 통해 시간 도메인으로 변환된다.
- [0510] 추적은 이 후 시간 도메인에서 수행될 수 있다. 추적은 예를 들어, 안락 잡음 레벨을 획득하기 위해 에너지 레벨에 기초하여 최소 통계 접근법을 수행함으로써 수행될 수 있다.
- [0511] 은닉에 대해, FD TCX PLC에 대해, 레벨은 그 상태로 사용될 수 있고, 시간 도메인에서의 페이딩만이 수행될 수

있다. TD TCX PLC(ACELP 형)에 대해, 여기 도메인으로의 레벨 변환 및 여기 도메인에서의 페이딩이 수행된다.

[0512] 다음 목록은 이것을 요약한다:

[0513] 고 비율:

[0514] - 입력:

[0515] - tcx (역 MDCT를 통해, mdct 도메인 -> 시간 도메인)

[0516] - 추적:

[0517] - 시간-도메인

[0518] - 에너지 레벨에 대한 최소 통계 -> 안락 잡음 레벨

[0519] - 은닉:

[0520] - 레벨 사용 "그 자체로(as is)"

[0521] - FD TCX PLC

[0522] -> 시간 도메인에서의 페이딩

[0523] - 여기 도메인으로의 레벨 변환으로서, 이것은

[0524] - TD TCX PLC(ACELP 형)

[0525] -> 여기 도메인에서의 페이딩을 위한 것이다.

[0526] FFT 도메인 및 MDCT 도메인 모두는 특수 도메인들인 반면, 여기 도메인은 몇몇 종류의 시간 도메인이다.

[0527] 실시예에 따라, 제 1 재구성 유닛(140)은 예를 들어, 잡음형 스펙트럼으로의 제 1 페이딩을 수행함으로써 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다. 제 2 재구성 유닛(141)은 예를 들어, 잡음형 스펙트럼으로의 제 2 페이딩 및/또는 LTP 이득의 제 2 페이딩을 수행함으로써 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다. 더욱이, 제 1 재구성 유닛(140) 및 제 2 재구성 유닛(141)은 예를 들어, 잡음형 스펙트럼으로의 제 1 페이딩 및 제 2 페이딩 및/또는 LTP 이득의 제 2 페이딩을 동일한 페이딩 속도로 수행하도록 구성될 수 있다.

[0528] 이제 안락 잡음의 적응형 스펙트럼 성형이 고려된다.

[0529] 제 1 단계로서, 버스트 패킷 손실 동안 안락 잡음에 대한 적응형 성형을 달성하기 위해, 배경 잡음을 나타내는 적절한 LPC 계수들을 찾는 것이 수행될 수 있다. 이들 LPC 계수들은, 배경 잡음 스펙트럼을 발견하고, 그런 후에 문헌으로부터 알려진 LPC 도출을 위한 임의의 알고리즘을 이용함으로써 그로부터 LPC 계수들을 계산하기 위해 최소 통계 접근법을 이용하여 활성 음성 동안 도출될 수 있다. 예를 들어, 몇몇 실시예들은 배경 잡음 스펙트럼을, MDCT 도메인에서 FDNS에 대해 직접 사용될 수 있는 표현으로 직접 변환할 수 있다.

[0530] 안락 잡음으로의 페이딩은 ISF 도메인에서 이루어질 수 있다(또한, LSF 도메인에서 적용가능함; LSF 라인 스펙트럼 주파수):

**수학식 26**

[0531] 
$$f_{current}[i] = \alpha \cdot f_{last}[i] + (1 - \alpha) \cdot pt_{mean}[i] \quad i = 0 \dots 16$$

[0532] 이것은  $pt_{mean}$ 을, 안락 잡음을 기재하는 적절한 LP 계수들로 설정함으로써 이루어진다.

[0533] 안락 잡음의 전술한 적응형 스펙트럼 성형에 관해, 더 일반적인 실시예가 도 11에 도시된다.

[0534] 도 11은 실시예에 따라 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한다.

[0535] 장치는 하나 이상의 프레임들을 위한 수신 인터페이스(110), 계수 생성기(1120), 및 신호 재구성기(1130)를 포함한다.

- [0536] 계수 생성기(1120)는, 하나 이상의 프레임들의 현재 프레임이 수신 인터페이스(1110)에 의해 수신되면, 그리고 수신 인터페이스(1110)에 의해 수신되는 현재 프레임이 붕괴되지 않으면/에러가 없으면, 현재 프레임에 의해 포함되는 하나 이상의 제 1 오디오 신호 계수들을 결정하도록 구성되고, 상기 하나 이상의 제 1 오디오 신호 계수들은 인코딩된 오디오 신호의 특징을 나타내고, 하나 이상의 오디오 계수들은 인코딩된 오디오 신호의 배경 잡음을 나타낸다. 더욱이, 계수 생성기(1120)는, 현재 프레임이 수신 인터페이스(1110)에 의해 수신되지 않으면, 또는 수신 인터페이스(1110)에 의해 수신되는 현재 프레임이 붕괴되면/에러가 있으면, 하나 이상의 제 1 오디오 신호 계수들에 따라, 그리고 하나 이상의 잡음 계수들에 따라, 하나 이상의 제 2 오디오 신호 계수들을 생성하도록 구성된다.
- [0537] 오디오 신호 재구성기(1130)는, 현재 프레임이 수신 인터페이스(1110)에 의해 수신되면, 그리고 수신 인터페이스(1110)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 하나 이상의 제 1 오디오 신호 계수들에 따라 재구성된 오디오 신호의 제 1 부분을 재구성하도록 구성된다. 더욱이, 오디오 신호 재구성기(1130)는, 현재 프레임이 수신 인터페이스(1110)에 의해 수신되지 않으면, 또는 수신 인터페이스(1110)에 의해 수신되는 현재 프레임이 붕괴되면, 하나 이상의 오디오 신호 계수들에 따라 재구성된 오디오 신호의 제 2 부분을 재구성하도록 구성된다.
- [0538] 배경 잡음을 결정하는 것은 종래 기술에 잘 알려져 있다(예를 들어, [Mar01]을 참조): 라이너 마르틴(Rainer Martin)의, 최적의 평활화 및 최소 통계에 기초한 잡음 전력 스펙트럼 밀도 추정(*Noise power spectral density estimation based on optimal smoothing and minimum statistics*), IEEE Transactions on Audio Processing 9 (2001), no.5, 504-512, 실시예에서, 장치는 이에 따라 진행된다.
- [0539] 몇몇 실시예들에서, 하나 이상의 제 1 오디오 신호 계수들은 예를 들어, 인코딩된 오디오 신호의 하나 이상의 선형 예측 필터 계수들일 수 있다. 몇몇 실시예들에서, 하나 이상의 제 1 오디오 신호 계수들은 예를 들어, 인코딩된 오디오 신호의 하나 이상의 선형 예측 필터 계수들일 수 있다.
- [0540] 선형 예측 필터 계수들로부터 또는 이미턴스 스펙트럼 쌍들로부터 오디오 신호, 예를 들어, 음성 신호를 재구성하는 방법이 종래 기술에 잘 알려져 있고(예를 들어, [3GP09c]: 음성 코덱 음성 처리 기능들; 적응형 멀티-레이트 - 광역 (AMRWB) 음성 코덱; 트랜스코딩 기능들, 3GPP TS 26.190, 3세대 파트너십 프로젝트, 2009를 참조), 실시예에서, 신호 재구성기는 이에 따라 진행된다.
- [0541] 실시예에 따라, 하나 이상의 잡음 계수들은 예를 들어, 인코딩된 오디오 신호의 배경 잡음을 나타내는 하나 이상의 선형 예측 필터 계수들일 수 있다. 실시예에서, 하나 이상의 선형 예측 필터 계수들은 예를 들어, 배경 잡음의 스펙트럼 형태를 나타낼 수 있다.
- [0542] 실시예에서, 계수 생성기(1120)는 예를 들어, 하나 이상의 제 2 오디오 신호 부분들이 재구성된 오디오 신호의 하나 이상의 선형 예측 필터 계수들이 되거나, 하나 이상의 제 1 오디오 신호 계수들이 재구성된 오디오 신호의 하나 이상의 이미턴스 스펙트럼 쌍들이 되도록 하나 이상의 오디오 신호 부분들을 결정하도록 구성될 수 있다.
- [0543] 실시예에 따라, 계수 생성기(1120)는 예를 들어, 수학적  $f_{current}[i]=\alpha \cdot f_{last}[i]+(1-\alpha) \cdot pt_{mean}[i]$  을 적용함으로써 하나 이상의 제 2 오디오 신호 계수들을 생성하도록 구성될 수 있고,
- [0544] 여기서  $f_{current}[i]$ 는 하나 이상의 제 2 오디오 신호 계수들 중 하나를 나타내고,  $f_{last}[i]$ 는 하나 이상의 제 1 오디오 신호 계수들 중 하나를 나타내고,  $pt_{mean}[i]$ 는 하나 이상의 잡음 계수들 중 하나이고,  $\alpha$ 는  $0 \leq \alpha \leq 1$ 을 갖는 실수이고,  $i$ 는 인덱스이다.
- [0545] 실시예에 따라,  $f_{last}[i]$ 는 인코딩된 오디오 신호의 선형 예측 필터 계수를 나타내고,  $f_{current}[i]$ 는 재구성된 오디오 신호의 선형 예측 필터 계수를 나타낸다.
- [0546] 실시예에서,  $pt_{mean}[i]$ 는 예를 들어, 인코딩된 오디오 신호의 배경 잡음을 나타내는 선형 예측 필터 계수일 수 있다.
- [0547] 실시예에 따라, 계수 생성기(1120)는 예를 들어, 하나 이상의 제 2 오디오 신호 계수들로서 적어도 10개의 제 2 오디오 신호 계수들을 생성하도록 구성될 수 있다.
- [0548] 실시예에서, 계수 생성기(1120)는 예를 들어, 하나 이상의 프레임들의 현재 프레임이 수신 인터페이스(1110)에 의해 수신되면, 그리고 수신 인터페이스(1110)에 의해 수신된 현재 프레임이 붕괴되지 않으면, 인코딩된 오디오

신호의 잡음 스펙트럼을 결정함으로써 하나 이상의 잡음 계수들을 결정하도록 구성될 수 있다.

[0549] 다음에서, FDNS 적용 이전에 MDCT 스펙트럼을 백색 잡음으로 페이딩하는 것이 고려된다.

[0550] MDCT 빈의 부호를 무작위로 변형(부호 스크램블링)하는 것 대신에, 완전한 스펙트럼은 FDNS를 이용하여 생성되는 백색 잡음으로 충전된다. 스펙트럼 특징들에서의 순간적인 변화를 회피하기 위해, 부호 스크램블링과 잡음 충전 사이의 교차-페이드(cross-fade)가 적용된다. 교차 페이드는 다음과 같이 실현될 수 있다:

```
[0551] for (i=0; i<L_frame; i++) {
[0552]     if (old_x[i] !=0) {
[0553]         x[i] = (i - cum_damping)*noise[i] + cum_damping *
[0554] random_sign() * x_old[i];
[0555]     }
[0556] }
```

[0557] 여기서:

[0558] cum\_damping은 (절대) 감쇄 인자이고 - 1에서 시작하여 0쪽으로 감소하여, 프레임마다 감소하고

[0559] x\_old는 마지막으로 수신된 프레임의 스펙트럼이고

[0560] random\_sign은 1 또는 -1로 리턴하고

[0561] noise는, 이차 평균(RMS)이 마지막으로 양호한 스펙트럼과 유사하도록 스케일링되는 확률 벡터(백색 잡음)를 포함한다.

[0562] 용어 random\_sing()\*old\_x[i]는 위상들을 무작위화하고, 고조파 반복들을 피하기 위해 부호-스크램블링 프로세스를 특징으로 한다.

[0563] 후속하여, 에너지 레벨의 다른 정규화는, 합산 에너지가 2개의 벡터들의 상관으로 인해 편향하지 않는 것을 확인하기 위해 교차-페이드 이후에 수행될 수 있다.

[0564] 실시예들에 따라, 제 1 재구성 유닛(140)은 예를 들어, 잡음 레벨 정보에 따라 그리고 제 1 오디오 신호 부분에 따라 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다. 특정 실시예에서, 제 1 재구성 유닛(140)은 예를 들어, 제 1 오디오 신호 부분을 감쇄 또는 증폭함으로써 제 3 오디오 신호 부분을 재구성하도록 구성될 수 있다.

[0565] 몇몇 실시예들에서, 제 2 재구성 유닛(141)은 예를 들어, 잡음 레벨 정보에 따라 그리고 제 2 오디오 신호 부분에 따라 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다. 특정 실시예에서, 제 2 재구성 유닛(141)은 예를 들어, 제 2 오디오 신호 부분을 감쇄 또는 증폭함으로써 제 4 오디오 신호 부분을 재구성하도록 구성될 수 있다.

[0566] FDNS 적용 이전에 백색 잡음으로의 MDCT 스펙트럼의 전술한 페이딩에 관해, 더 일반적인 실시예는 도 12에 의해 도시된다.

[0567] 도 12는 실시예에 따라 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한다.

[0568] 장치는 인코딩된 오디오 신호의 오디오 신호 스펙트럼의 복수의 오디오 신호 샘플들에 대한 정보를 포함하는 하나 이상의 프레임들을 수신하기 위한 수신 인터페이스(1210)와, 재구성된 오디오 신호를 생성하기 위한 프로세서(1220)를 포함한다.

[0569] 프로세서(1220)는, 현재 프레임이 수신 인터페이스(1210)에 의해 수신되지 않으면, 또는 현재 프레임이 수신 인터페이스(1210)에 의해 수신되지만, 붕괴되면, 변형된 스펙트럼을 목표 스펙트럼으로 페이딩함으로써 재구성된 오디오 신호를 생성하도록 구성되고, 변형된 스펙트럼은 복수의 변형된 신호 샘플들을 포함하고, 변형된 스펙트럼의 변형된 신호 샘플들 각각에 대해, 상기 변형된 신호 샘플의 절대값은 오디오 신호 스펙트럼의 오디오 신호 샘플들 중 하나의 오디오 신호 샘플의 절대값과 동일하다.

- [0570] 더욱이, 프로세서(1220)는, 하나 이상의 프레임들의 현재 프레임이 수신 인터페이스(1210)에 의해 수신되면, 그리고 수신 인터페이스(1210)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 변형된 스펙트럼을 목표 스펙트럼으로 페이딩하지 않도록 구성된다.
- [0571] 실시예에 따라, 목표 스펙트럼은 잡음형 스펙트럼이다.
- [0572] 실시예에 따라, 잡음형 스펙트럼은 백색 잡음을 나타낸다.
- [0573] 실시예에 따라, 잡음형 스펙트럼은 성형된다.
- [0574] 실시예에서, 잡음형 스펙트럼의 형태는 이전에 수신된 신호의 오디오 신호 스펙트럼에 의존한다.
- [0575] 실시예에 따라, 잡음형 스펙트럼은 오디오 신호 스펙트럼의 형태에 따라 성형된다.
- [0576] 실시예에서, 프로세서(1220)는 잡음형 스펙트럼을 성형하기 위해 경사 인자를 이용한다.
- [0577] 실시예에 따라, 프로세서(1220)는 수학적 식  $\text{shaped\_noise}[i] = \text{noise} * \text{power}(\text{tilt\_factor}, i/N)$ 을 이용하고,
- [0578] 여기서 N은 샘플들의 개수를 나타내고,
- [0579] i는 인덱스이고,
- [0580]  $\text{tilt\_factor} > 0$ 인 상태에서,  $0 \leq i < N$ 이고,
- [0581] power는 멱함수이다.
- [0582]  $\text{tilt\_factor}$ 가 1보다 작으면, 이것은 증가하는 i를 갖는 감쇄를 의미한다.  $\text{tilt\_factor}$ 가 1보다 크면, 이것은 증가하는 i를 갖는 증폭을 의미한다.
- [0583] 다른 실시예에 따라, 프로세서(1220)는 수학적 식  $\text{shaped\_noise}[i] = \text{noise} * (1 + i / (N-1) * (\text{tilt\_factor} - 1))$ 를 이용할 수 있고,
- [0584] 여기서 N은 샘플들의 개수를 나타내고,
- [0585] i는 인덱스이고,  $\text{tilt\_factor} > 0$ 인 상태에서,  $0 \leq i < N$ 이다.
- [0586] 실시예에 따라, 프로세서(1220)는, 현재 프레임이 수신 인터페이스(1210)에 의해 수신되지 않으면, 또는 수신 인터페이스(1210)에 의해 수신되는 현재 프레임이 붕괴되면, 오디오 신호 스펙트럼의 하나 이상의 오디오 신호 샘플들의 부호를 변화시킴으로써 변형된 스펙트럼을 생성하도록 구성된다.
- [0587] 실시예에서, 오디오 신호 스펙트럼의 오디오 신호 샘플들 각각은 실수에 의해 표현되지만, 허수에 의해 표현되지 않는다.
- [0588] 실시예에 따라, 오디오 신호 스펙트럼의 오디오 신호 샘플들은 변형된 이산 코사인 변환 도메인에서 표현된다.
- [0589] 다른 실시예에 따라, 오디오 신호 스펙트럼의 오디오 신호 샘플들은 변형된 이산 사인 변환 도메인에서 표현된다.
- [0590] 실시예에 따라, 프로세서(1220)는 제 1 또는 제 2 값을 무작위로 또는 의사-무작위로 출력하는 랜덤 부호 함수를 이용함으로써 변형된 스펙트럼을 생성하도록 구성된다.
- [0591] 실시예에서, 프로세서(1220)는 후속하여 감쇄 인자를 감소시킴으로써 변형된 스펙트럼을 목표 스펙트럼에 페이딩하도록 구성된다.
- [0592] 실시예에 따라, 프로세서(1220)는 후속하여 감쇄 인자를 증가시킴으로써 변형된 스펙트럼을 목표 스펙트럼에 페이딩하도록 구성된다.
- [0593] 실시예에서, 현재 프레임이 수신 인터페이스(1210)에 의해 수신되지 않으면, 또는 수신 인터페이스(1210)에 의해 수신되는 현재 프레임이 붕괴되면, 프로세서(1220)는 수학적 식  $x[i] = (1 - \text{cum\_damping}) * \text{noise}[i] + \text{cum\_damping} * \text{random\_sign}() * x\_old[i]$ 을 이용함으로써 재구성된 오디오 신호를 생성하도록 구성되고,
- [0594] 여기서 i는 인덱스이고, x[i]는 재구성된 오디오 신호의 샘플을 나타내고, cum\_damping은 감쇄 인자이고, x\_old[i]는 인코딩된 오디오 신호의 오디오 신호 스펙트럼의 오디오 신호 샘플들 중 하나를 나타내고, random\_sign()은 1 또는 -1로 리턴하고, noise는 목표 스펙트럼을 나타내는 확률 벡터이다.

- [0595] 몇몇 실시예들은 TCX LTP 동작을 계속한다. 이들 실시예들에서, TCX LTP 동작은 마지막으로 양호한 프레임으로부터 도출된 LTP 파라미터들(LTP 래그 및 LTP 이득)을 이용한 은닉 동안 계속된다.
- [0596] LTP 동작들은 다음과 같이 요약될 수 있다:
- [0597] - 이전에 도출된 출력에 기초하여 LTP 지연 버퍼를 공급.
- [0598] - LTP 래그에 기초: 현재 신호를 성형하기 위해 LTP 기여로서 사용되는 LTP 지연 버퍼로부터 적절한 신호 부분을 선택.
- [0599] - LTP 이득을 이용하여 이러한 LTP 기여를 채스케일링.
- [0600] - LTP 출력 신호를 생성하기 위해 LTP 입력 신호에 이러한 채스케일링된 LTP 기여를 추가.
- [0601] 상이한 접근법들은 LTP 지연 버퍼 업데이트가 수행될 때의 시간에 관해 고려될 수 있다:
- [0602] 마지막 프레임(n-1)으로부터의 출력을 이용하여 프레임(n)에서의 제 1 LTP 동작으로서, 이것은 프레임(n)에서의 LTP 처리 동안 사용될 프레임(n)에서의 LTP 지연 버퍼를 업데이트한다.
- [0603] 현재 프레임(n)으로부터의 출력을 이용하여 프레임(n)에서의 마지막 LTP 동작으로서, 이것은 프레임(n+1)에서의 LTP 처리 동안 사용될 프레임(n)에서의 LTP 지연 버퍼를 업데이트한다.
- [0604] 다음에서, TCX LTP 피드백 루프의 결합 해제가 고려된다.
- [0605] TCX LTP 피드백 루프의 결합 해제는 은닉 모드에 있을 때 LTP 디코더의 각 피드백 루프 동안 추가 잡음(LTP 입력 신호에 적용된 잡음 교체로부터 초래됨)의 도입을 회피한다.
- [0606] 도 10은 이러한 결합 해제를 도시한다. 특히, 도 10은 은닉(bfi=1) 동안 LTP 피드백 루프의 결합 해제를 도시한다.
- [0607] 도 10은 지연 버퍼(1020), 샘플 선택기(1030), 및 샘플 프로세서(1040){샘플 프로세서(1040)는 점선으로 표시됨}를 도시한다.
- [0608] LTP 지연 버퍼(1020) 업데이트가 수행되는 시간 가까이, 몇몇 실시예들은 다음과 같이 진행된다:
- [0609] - 정상 동작에 대해: 제 1 LTP 동작이 바람직하기 때문에 LTP 지연 버퍼(1020)를 업데이트하기 위해, 이는 합산된 출력 신호가 일반적으로 끊임없이 저장되기 때문이다. 이러한 접근법을 통해, 전용 버퍼가 생략될 수 있다.
- [0610] - 결합 해제 동작에 대해: 마지막 LTP 동작이 바람직하기 때문에 LTP 지연 버퍼(1020)를 업데이트하기 위해, 이는 신호에 대한 LTP 기여가 일반적으로 단지 임시로 저장되기 때문이다. 이러한 접근법을 통해, 일시적으로 LTP 기여 신호가 확보된다. 구현-방식의 이러한 LTP 기여 버퍼는 단지 끊임없이 이루어질 수 있다.
- [0611] 후자의 접근법이 임의의 경우(정상 동작 및 은닉)에 사용된다고 가정하면, 실시예들은 예를 들어, 다음을 구현할 수 있다:
- [0612] - 정상 동작 동안: LTP 입력 신호로의 추가 이후에 LTP 디코더의 시간 도메인 신호 출력은 LTP 지연 버퍼에 공급하는데 사용된다.
- [0613] - 은닉 동안: LTP 입력 신호로의 추가 이전에 LTP 디코더의 시간 도메인 신호 출력은 LTP 지연 버퍼에 공급하는데 사용된다.
- [0614] 몇몇 실시예들은 TCX LTP 이득을 0쪽으로 페이딩한다. 그러한 실시예에서, TCX LTP 이득은 예를 들어, 특정한 신호 적응형 페이드-아웃 인자를 가지고 0쪽으로 페이딩될 수 있다. 예를 들어, 이것은 예를 들어, 다음의 의사-코드에 따라 반복적으로 이루어질 수 있다:
- [0615]  $gain = gain\_past * damping;$
- [0616] [...]
- [0617]  $gain\_past = gain;$
- [0618] 여기서 gain은 현재 프레임에 적용된 TCX LTP 디코더 이득이고;
- [0619] gain\_past는 이전 프레임에 적용된 TCX LTP 디코더 이득이고;

- [0620] damping은 (상대적인) 페이드-아웃 인자이다.
- [0621] 도 1d는 추가 실시예에 따른 장치를 도시하고, 장치는 지연 버퍼(180)를 포함하는 장기간 예측 유닛(170)을 더 포함한다. 장기간 예측 유닛(170)은 제 2 오디오 신호 부분에 따라, 지연 버퍼(180)에 저장되는 지연 버퍼 입력에 따라, 그리고 장기간 예측 이득에 따라, 처리된 신호를 생성하도록 구성된다. 더욱이, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 장기간 예측 유닛은 장기간 예측 이득을 0쪽으로 페이딩하도록 구성된다.
- [0622] 다른 실시예들(미도시)에서, 장기간 예측 유닛은 예를 들어, 제 1 오디오 신호 부분에 따라, 지연 버퍼에 저장되는 지연 버퍼 입력에 따라, 그리고 장기간 예측 이득에 따라, 처리된 신호를 생성하도록 구성될 수 있다.
- [0623] 도 1d에서, 제 1 재구성 유닛(140)은 예를 들어, 추가로 처리된 신호에 따라 제 3 오디오 신호 부분을 생성할 수 있다.
- [0624] 실시예에서, 장기간 예측 유닛(170)은 예를 들어, 장기간 예측 이득을 0쪽으로 페이딩하도록 구성될 수 있고, 장기간 예측 이득이 0으로 페이딩되는 속도는 페이드-아웃 인자에 의존한다.
- [0625] 대안적으로 또는 추가로, 장기간 예측 유닛(170)은 예를 들어, 복수의 프레임들의 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지 않으면, 또는 상기 제 3 프레임이 수신 인터페이스(110)에 의해 수신되지만, 붕괴되면, 생성된 처리된 신호를 지연 버퍼(180)에 저장함으로써 지연 버퍼(180)를 업데이트하도록 구성될 수 있다.
- [0626] TCX LTP의 전술한 이용에 관해, 더 일반적인 실시예는 도 13에 도시된다.
- [0627] 도 13은 재구성된 오디오 신호를 얻기 위해 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한다.
- [0628] 장치는 복수의 프레임들을 수신하기 위한 수신 인터페이스(1310)와, 디코딩된 오디오 신호의 오디오 신호 샘플들을 저장하기 위한 지연 버퍼(1320)와, 지연 버퍼(1320)에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하기 위한 샘플 선택기(1330)와, 재구성된 오디오 신호의 재구성된 오디오 신호 샘플들을 얻기 위해 선택된 오디오 신호 샘플들을 처리하기 위한 샘플 프로세서(1340)를 포함한다.
- [0629] 샘플 선택기(1330)는, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되면, 그리고 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 피치 래그 정보에 따라 지연 버퍼(1320)에 저장되는 오디오 신호 샘플들로부터 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성된다. 더욱이, 샘플 선택기(1330)는, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되지 않으면, 또는 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되면, 수신 인터페이스(1310)에 의해 이전에 수신되는 다른 프레임에 의해 포함되는 피치 래그 정보에 따라 지연 버퍼(1320)에 저장되는 오디오 신호로부터 복수의 선택된 오디오 신호 샘플들을 선택하도록 구성된다.
- [0630] 실시예에 따라, 샘플 프로세서(1340)는, 예를 들어 현재 프레임이 수신 인터페이스(1310)에 의해 수신되면, 그리고 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 이득 정보에 따라 선택된 오디오 신호 샘플들을 재스케일링함으로써, 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있다. 더욱이, 샘플 선택기(1330)는 예를 들어, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되지 않으면, 또는 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되면, 수신 인터페이스(1310)에 의해 이전에 수신된 상기 다른 프레임에 의해 포함되는 이득 정보에 따라 선택된 오디오 신호 샘플들을 재스케일링함으로써, 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있다.
- [0631] 실시예에서, 샘플 프로세서(1340)는 예를 들어, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되면, 그리고 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 현재 프레임에 의해 포함되는 이득 정보에 따른 값과 선택된 오디오 신호 샘플들을 곱함으로써, 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있다. 더욱이, 샘플 선택기(1330)는, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되지 않으면, 또는 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되면, 수신 인터페이스(1310)에 의해 이전에 수신되는 상기 다른 프레임에 의해 포함되는 이득 정보에 따른 값과 선택된 오디오 신호 샘플들을 곱함으로써, 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있다.
- [0632] 실시예에 따라, 샘플 프로세서(1340)는 예를 들어, 재구성된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하도록 구성될 수 있다.



- [0633] 실시예에서, 샘플 프로세서(1340)는 예를 들어, 추가 프레임이 수신 인터페이스(1310)에 의해 수신되기 전에 재구성된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하도록 구성될 수 있다.
- [0634] 실시예에 따라, 샘플 프로세서(1340)는 예를 들어, 추가 프레임이 수신 인터페이스(1310)에 의해 수신된 후에 재구성된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하도록 구성될 수 있다.
- [0635] 실시예에서, 샘플 프로세서(1340)는, 예를 들어, 재스케일링된 오디오 신호 샘플들을 얻기 위해 이득 정보에 따라, 그리고 처리된 오디오 신호 샘플들을 얻기 위해 재스케일링된 오디오 신호 샘플들과 입력 오디오 신호 샘플들을 조합함으로써, 선택된 오디오 신호 샘플들을 재스케일링하도록 구성될 수 있다.
- [0636] 실시예에 따라, 샘플 프로세서(1340)는 예를 들어, 재스케일링된 오디오 신호 샘플들과 입력 오디오 신호 샘플들의 조합을 나타내는 처리된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하고, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되면, 그리고 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 재스케일링된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하지 않도록 구성될 수 있다. 더욱이, 샘플 프로세서(1340)는, 재스케일링된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하고, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되지 않으면, 또는 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되지 않으면, 처리된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하지 않도록 구성된다.
- [0637] 다른 실시예에 따라, 샘플 프로세서(1340)는 예를 들어, 현재 프레임이 수신 인터페이스(1310)에 의해 수신되지 않으면, 또는 수신 인터페이스(1310)에 의해 수신되는 현재 프레임이 붕괴되면, 처리된 오디오 신호 샘플들을 지연 버퍼(1320)에 저장하도록 구성될 수 있다.
- [0638] 실시예에서, 샘플 선택기(1330)는 예를 들어, 변형된 이득에 따라 선택된 오디오 신호 샘플들을 재스케일링함으로써 재구성된 오디오 신호 샘플들을 얻도록 구성될 수 있고, 변형된 이득은 수학적 식  $gain = gain\_past * damping$ 에 따라 정의되고;
- [0639] 여기서 gain은 변형된 이득이고, 샘플 선택기(1330)는 예를 들어, gain이 계산된 후에 gain\_past를 gain으로 설정하도록 구성될 수 있고, damping은 실수이다.
- [0640] 실시예에 따라, 샘플 선택기(1330)는 예를 들어, 변형된 이득을 계산하도록 구성될 수 있다.
- [0641] 실시예에서, damping은 예를 들어,  $0 < damping < 1$ 에 따라 정의될 수 있다.
- [0642] 실시예에 따라, 변형된 이득 gain은 예를 들어, 프레임이 수신 인터페이스(1310)에 의해 마지막으로 수신되었기 때문에 적어도 미리 한정된 수의 프레임들이 수신 인터페이스(1310)에 의해 수신되지 않은 경우 0으로 설정될 수 있다.
- [0643] 다음에서, 페이드-아웃 속도가 고려된다. 특정 종류의 페이드-아웃을 적용하는 여러 은닉 모듈들이 존재한다. 이러한 페이드-아웃의 속도가 이들 모듈들 양단에서 상이하게 선택될 수 있지만, 하나의 코어에 대한 모든 은닉 모듈들(ACELP 또는 TCX)에 대해 동일한 페이드-아웃 속도를 이용하는 것이 유리하다. 예를 들어:
- [0644] ACELP에 대해, 동일한 페이드 아웃 속도가, 특히 적응형 코드북(이득을 변경함으로써), 및/또는 혁신 코드북 신호(이득을 변경함으로써)에 대해 사용되어야 한다.
- [0645] 또한, TCX에 대해, 동일한 페이드 아웃 속도가, 특히 시간 도메인 신호, 및/또는 LTP 이득(0으로의 페이딩), 및/또는 LPC 가중치(1로 페이딩), 및/또는 LP 계수들(배경 스펙트럼 형태로 페이딩), 및/또는 백색 잡음으로의 교차 페이드에 사용되어야 한다.
- [0646] 또한 ACELP 및 TCX에 대해 동일한 페이드-아웃 속도를 이용하는 것이 추가로 바람직할 수 있지만, 코어들의 상이한 특성으로 인해, 또한 상이한 페이드-아웃 속도들을 이용하도록 선택될 수 있다.
- [0647] 이러한 페이드-아웃 속도는 정적일 수 있지만, 바람직하게 신호 특징들에 적응성이다. 예를 들어, 페이드-아웃 속도는 예를 들어, LPC 안정도 인자(TCX) 및/또는 분류, 및/또는 연속적으로 손실된 프레임들의 개수에 의존할 수 있다.
- [0648] 페이드-아웃 속도는, 예를 들어, 감쇄 인자에 따라 결정될 수 있는데, 이러한 감쇄 인자는 절대적으로 또는 상대적으로 주어질 수 있고, 또한 특정 페이드-아웃 동안 시간이 지남에 따라 변할 수 있다.
- [0649] 실시예들에서, 동일한 페이딩 속도는 백색 잡음 페이딩에 대한 LTP 이득 페이딩에 사용된다.

- [0650]        기술한 바와 같이, 안락 잠음 신호를 생성하기 위한 장치, 방법 및 컴퓨터 프로그램이 제공되었다.
- [0651]        몇몇 양상들이 장치의 정황에서 기재되었지만, 이들 양상들이 또한 대응하는 방법의 설명을 나타내고, 여기서 블록 또는 디바이스가 방법 단계 또는 방법 단계의 특징에 대응한다는 것이 명확하다. 유사하게, 방법 단계의 정황에서 기재된 양상들은 또한 대응하는 블록 또는 대응하는 장치의 항목 또는 특징의 설명을 나타낸다.
- [0652]        본 발명의 분해된 신호는 디지털 저장 매체 상에 저장될 수 있거나, 인터넷과 같이 무선 송신 매체 또는 유선 송신 매체와 같은 송신 매체 상에서 송신될 수 있다.
- [0653]        특정 구현 요건들에 따라, 본 발명의 실시예들은 하드웨어 또는 소프트웨어로 구현될 수 있다. 구현은 디지털 저장 매체, 예를 들어, 플로피 디스크, DVD, CD, ROM, PROM, EPROM, EEPROM, 또는 FLASH 메모리를 이용하여 수행될 수 있는데, 이러한 디지털 저장 매체는 그 위에 저장된 전자적으로 판독가능한 제어 신호들을 갖고, 각 방법이 수행되도록 프로그래밍가능 컴퓨터 시스템과 협력한다(또는 협력할 수 있다).
- [0654]        본 발명에 따른 몇몇 실시예들은, 본 명세서에 기재된 방법들 중 하나가 수행되도록, 프로그래밍가능 컴퓨터 시스템과 협력할 수 있는, 전자적으로 판독가능한 제어 신호들을 갖는 비-임시 데이터 캐리어를 포함한다.
- [0655]        일반적으로, 본 발명의 실시예들은 프로그램 코드를 갖는 컴퓨터 프로그램 제품으로서 구현될 수 있고, 프로그램 코드는, 컴퓨터 프로그램이 컴퓨터 상에서 실행될 때 방법들 중 하나를 수행하기 위해 동작가능하다. 프로그램 코드는 예를 들어, 기계 판독가능한 캐리어 상에 저장될 수 있다.
- [0656]        다른 실시예들은 기계 판독가능한 캐리어 상에 저장된, 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램을 포함한다.
- [0657]        즉, 그러므로, 본 발명의 방법의 실시예는, 컴퓨터 프로그램이 컴퓨터 상에서 실행될 때, 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 프로그램 코드를 갖는 컴퓨터 프로그램이다.
- [0658]        그러므로, 본 발명의 방법들의 추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램을 그 위에 리코딩되게 포함하는 데이터 캐리어(또는 디지털 저장 매체, 또는 컴퓨터-판독가능 매체)이다.
- [0659]        그러므로, 본 발명의 방법의 추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램을 나타내는 신호들의 시퀀스 또는 데이터 스트림이다. 예를 들어, 신호들의 시퀀스들 또는 데이터 스트림은 데이터 통신 연결부를 통해, 예를 들어, 인터넷을 통해, 전송되도록 구성될 수 있다.
- [0660]        추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하도록 구성되거나 적용된 처리 수단, 예를 들어, 컴퓨터, 또는 프로그래밍가능 논리 디바이스를 포함한다.
- [0661]        추가 실시예는 본 명세서에 기재된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램이 그 위에 설치된 컴퓨터를 포함한다.
- [0662]        몇몇 실시예들에서, 프로그래밍가능 논리 디바이스(예를 들어, 전계 프로그래밍가능 게이트 어레이)는 본 명세서에 기재된 방법들의 기능들 중 몇몇 또는 전부를 수행하는데 사용될 수 있다. 몇몇 실시예들에서, 전계 프로그래밍가능 게이트 어레이는 본 명세서에 기재된 방법들 중 하나를 수행하기 위해 마이크로프로세서와 협력할 수 있다. 일반적으로, 방법들은 임의의 하드웨어 장치에 의해 바람직하게 수행된다.
- [0663]        전술한 실시예들은 본 발명의 원리들을 위해 단지 예시적이다. 본 명세서에 기재된 세부사항들 및 배치들의 변형들 및 변경들이 당업자에게 명백하다는 것이 이해된다. 그러므로, 본 명세서에서 실시예들의 기재 및 설명에 의해 제공된 특정 세부사항들에 의해서가 아니라 다음의 특허 청구항들의 범주에 의해서만 제한되도록 의도된다.
- [0664]        인용 문헌들
- [0665]        [3GP09a]        3GPP; Technical Specification Group Services and System Aspects, *Extended adaptive multi-rate - wideband (AMR-WB+) codec*, 3GPP TS 26.290, 3rd Generation Partnership Project, 2009.
- [0666]        [3GP09b]        *Extended adaptive multi-rate - wideband (AMR-WB+) codec; floating-point ANSI-C code*, 3GPP TS 26.304, 3rd Generation Partnership Project, 2009.
- [0667]        [3GP09c]        *Speech codec speech processing functions; adaptive multi-rate - wideband (AMRWB) speech codec; transcoding functions*, 3GPP TS 26.190, 3rd Generation Partnership Project, 2009.

- [0668] [3GP12a] *Adaptive multi-rate (AMR) speech codec; error concealment of lost frames (release 11)*, 3GPP TS 26.091, 3rd Generation Partnership Project, Sep 2012.
- [0669] [3GP12b] *Adaptive multi-rate (AMR) speech codec; transcoding functions (release 11)*, 3GPP TS 26.090, 3rd Generation Partnership Project, Sep 2012. [3GP12c] , ANSI-C code for the adaptive multi-rate - wideband (AMR-WB) speech codec, 3GPP TS 26.173, 3rd Generation Partnership Project, Sep 2012.
- [0670] [3GP12d] *ANSI-C code for the floating-point adaptive multi-rate (AMR) speech codec (release11)*, 3GPP TS 26.104, 3rd Generation Partnership Project, Sep 2012.
- [0671] [3GP12e] *General audio codec audio processing functions; Enhanced aacPlus general audio codec; additional decoder tools (release 11)*, 3GPP TS 26.402, 3rd Generation Partnership Project, Sep 2012.
- [0672] [3GP12f] *Speech codec speech processing functions; adaptive multi-rate - wideband (amr-wb) speech codec; ansi-c code*, 3GPP TS 26.204, 3rd Generation Partnership Project, 2012.
- [0673] [3GP12g] *Speech codec speech processing functions; adaptive multi-rate - wideband (AMR-WB) speech codec; error concealment of erroneous or lost frames*, 3GPP TS 26.191, 3rd Generation Partnership Project, Sep 2012.
- [0674] [BJH06] I. Batina, J. Jensen, and R. Heusdens, *Noise power spectrum estimation for speech enhancement using an autoregressive model for speech power spectrum dynamics*, in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. 3 (2006), 1064-1067.
- [0675] [BP06] A. Borowicz and A. Petrovsky, *Minima controlled noise estimation for klt-based speech enhancement*, CD-ROM, 2006, Italy, Florence.
- [0676] [Coh03] I. Cohen, *Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging*, IEEE Trans. Speech Audio Process. 11 (2003), no. 5, 466-475.
- [0677] [CPK08] Choong Sang Cho, Nam In Park, and Hong Kook Kim, *A packet loss concealment algorithm robust to burst packet loss for celp- type speech coders*, Tech. report, Korea Electronics Technology Institute, Gwang Institute of Science and Technology, 2008, The 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2008).
- [0678] [Dob95] G. Doblinger, *Computationally efficient speech enhancement by spectral minima tracking in subbands*, in Proc. Eurospeech (1995), 1513-1516.
- [0679] [EBU10] EBU/ETSI JTC Broadcast, *Digital audio broadcasting (DAB); transport of advanced audio coding (AAC) audio*, ETSI TS 102 563, European Broadcasting Union, May 2010.
- [0680] [EBU12] *Digital radio mondiale (DRM); system specification*, ETSI ES 201 980, ETSI, Jun 2012.
- [0681] [EH08] Jan S. Erkelens and Richards Heusdens, *Tracking of Nonstationary Noise Based on Data-Driven Recursive Noise Power Estimation*, Audio, Speech, and Language Processing, IEEE Transactions on 16 (2008), no. 6, 1112 -1123.
- [0682] [EM84] Y. Ephraim and D. Malah, *Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator*, IEEE Trans. Acoustics, Speech and Signal Processing 32 (1984), no. 6, 1109-1121.
- [0683] [EM85] *Speech enhancement using a minimum mean-square error log-spectral amplitude estimator*, IEEE Trans. Acoustics, Speech and Signal Processing 33 (1985), 443-445.
- [0684] [Gan05] S. Gannot, *Speech enhancement: Application of the kalman filter in the estimate-maximize (em framework)*, Springer, 2005.
- [0685] [HE95] H. G. Hirsch and C. Ehrlicher, *Noise estimation techniques for robust speech recognition*, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, no. pp. 153-156, IEEE, 1995.
- [0686] [HHJ10] Richard C. Hendriks, Richard Heusdens, and Jesper Jensen, *MMSE based noise PSD*

*tracking with low complexity*, Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on, Mar 2010, pp. 4266 -4269.

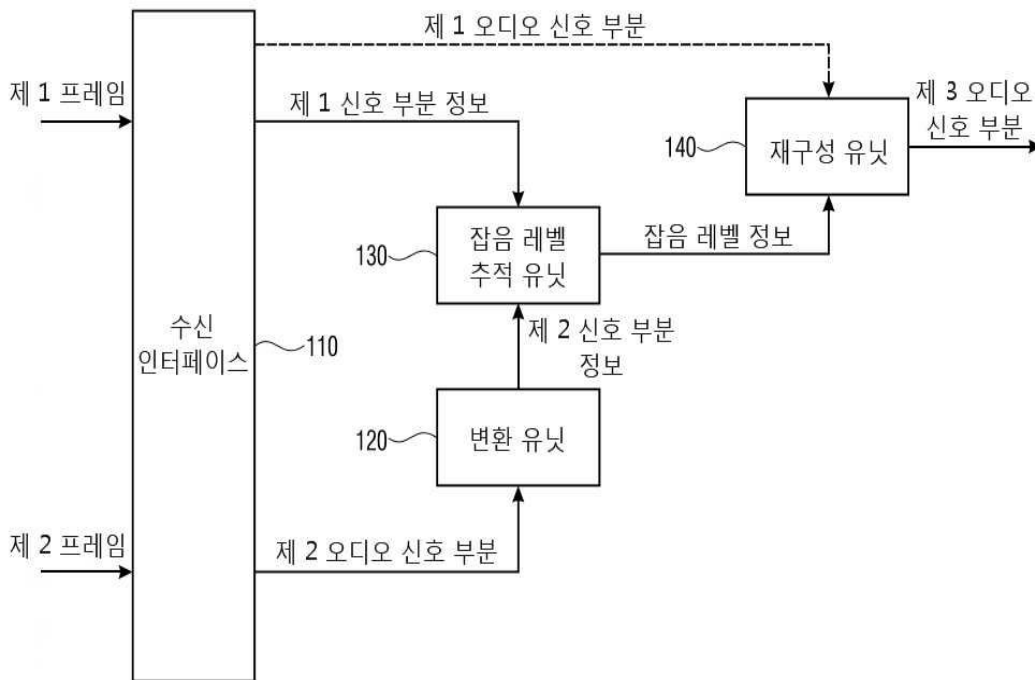
- [0687] [HJH08] Richard C. Hendriks, Jesper Jensen, and Richard Heusdens, *Noise tracking using dft domain subspace decompositions*, IEEE Trans. Audio, Speech, Lang. Process. 16 (2008), no. 3, 541-553.
- [0688] [IET12] IETF, *Definition of the Opus Audio Codec*, Tech. Report RFC 6716, Internet Engineering Task Force, Sep 2012.
- [0689] [ISO09] ISO/IEC JTC1/SC29/WG11, *Information technology - coding of audio-visual objects - part 3: Audio*, ISO/IEC IS 14496-3, International Organization for Standardization, 2009.
- [0690] [ITU03] ITU-T, *Wideband coding of speech at around 16 kbit/s using adaptive multi-rate wideband (amr-wb)*, Recommendation ITU-T G.722.2, Telecommunication Standardization Sector of ITU, Jul 2003.
- [0691] [ITU05] *Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss*, Recommendation ITU-T G.722.1, Telecommunication Standardization Sector of ITU, May 2005.
- [0692] [ITU06a] *G.722 Appendix III: A high-complexity algorithm for packet loss concealment for G.722*, ITU-T Recommendation, ITU-T, Nov 2006.
- [0693] [ITU06b] *G.729.1: G.729-based embedded variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with g.729*, Recommendation ITU-T G.729.1, Telecommunication Standardization Sector of ITU, May 2006.
- [0694] [ITU07] *G.722 Appendix IV: A low-complexity algorithm for packet loss concealment with G.722*, ITU-T Recommendation, ITU-T, Aug 2007.
- [0695] [ITU08a] *G.718: Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s*, Recommendation ITU-T G.718, Telecommunication Standardization Sector of ITU, Jun 2008.
- [0696] [ITU08b] *G.719: Low-complexity, full-band audio coding for high-quality, conversational applications*, Recommendation ITU-T G.719, Telecommunication Standardization Sector of ITU, Jun 2008.
- [0697] [ITU12] *G.729: Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (cs-acelp)*, Recommendation ITU-T G.729, Telecommunication Standardization Sector of ITU, June 2012.
- [0698] [LS01] Pierre Lauber and Ralph Sperschneider, *Error concealment for compressed digital audio*, Audio Engineering Society Convention 111, no. 5460, Sep 2001.
- [0699] [Mar01] Rainer Martin, *Noise power spectral density estimation based on optimal smoothing and minimum statistics*, IEEE Transactions on Speech and Audio Processing 9 (2001), no. 5, 504 -512.
- [0700] [Mar03] *Statistical methods for the enhancement of noisy speech*, International Workshop on Acoustic Echo and Noise Control (IWAENC2003), Technical University of Braunschweig, Sep 2003.
- [0701] [MC99] R. Martin and R. Cox, *New speech enhancement techniques for low bit rate speech coding*, in Proc. IEEE Workshop on Speech Coding (1999), 165-167.
- [0702] [MCA99] D. Malah, R. V. Cox, and A. J. Accardi, *Tracking speech-presence uncertainty to improve speech enhancement in nonstationary noise environments*, Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing (1999), 789-792.
- [0703] [MEP01] Nikolaus Meine, Bernd Edler, and Heiko Purnhagen, *Error protection and concealment for HILN MPEG-4 parametric audio coding*, Audio Engineering Society Convention 110, no. 5300, May 2001.
- [0704] [MPC89] Y. Mahieux, J.-P. Petit, and A. Charbonnier, *Transform coding of audio signals using correlation between successive transform blocks*, Acoustics, Speech, and Signal Processing, 1989.

ICASSP-89., 1989 International Conference on, 1989, pp. 2021-2024 vol.3.

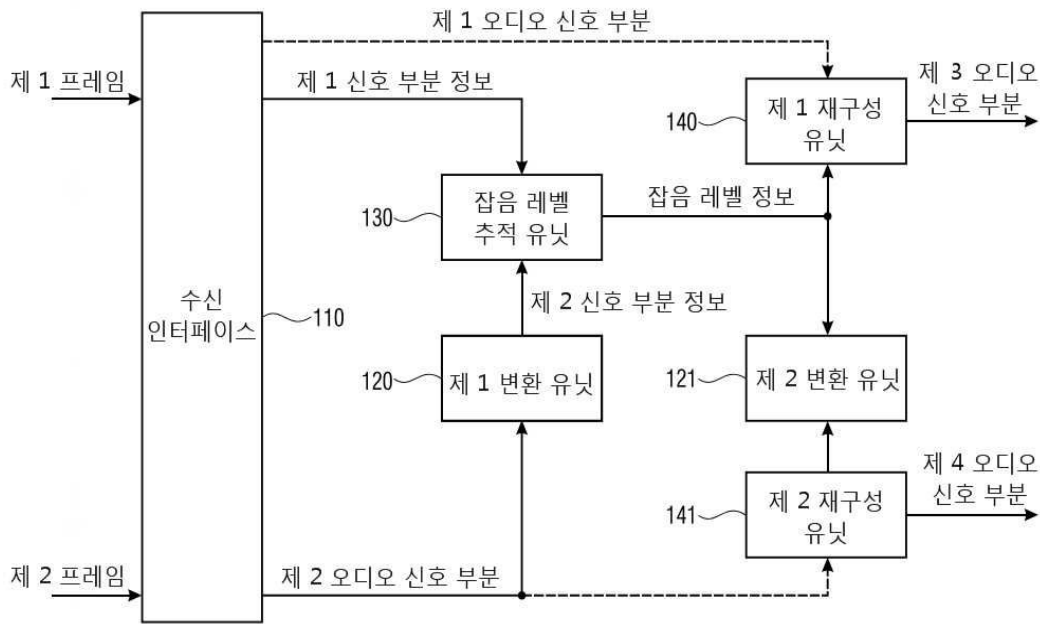
- [0705] [NMR+12] Max Neuendorf, Markus Multrus, Nikolaus Rettelbach, Guillaume Fuchs, Julien Robilliard, J꺁꺁꺁꺁꺁꺁 Lecomte, Stephan Wilde, Stefan Bayer, Sascha Disch, Christian Helmrich, Roch Lefebvre, Philippe Gournay, Bruno Bessette, Jimmy Lapierre, Kristopfer Kjø꺁꺁꺁꺁꺁꺁, Heiko Purnhagen, Lars Villemoes, Werner Oomen, Erik Schuijers, Kei Kikuiiri, Toru Chinen, Takeshi Norimatsu, Chong Kok Seng, Eunmi Oh, Miyoung Kim, Schuyler Quackenbush, and Berndhard Grill, *MPEG Unified Speech and Audio Coding - The ISO / MPEG Standard for High-Efficiency Audio Coding of all Content Types*, Convention Paper 8654, AES, April 2012, Presented at the 132nd Convention Budapest, Hungary.
- [0706] [PKJ+11] Nam In Park, Hong Kook Kim, Min A Jung, Seong Ro Lee, and Seung Ho Choi, *Burst packet loss concealment using multiple codebooks and comfort noise for celp-type speech coders in wireless sensor networks*, *Sensors* 11 (2011), 5323-5336.
- [0707] [QD03] Schuyler Quackenbush and Peter F. Driessen, *Error mitigation in MPEG-4 audio packet communication systems*, *Audio Engineering Society Convention* 115, no. 5981, Oct 2003.
- [0708] [RL06] S. Rangachari and P. C. Loizou, *A noise-estimation algorithm for highly non-stationary environments*, *Speech Commun.* 48 (2006), 220-231.
- [0709] [SFB00] V. Stahl, A. Fischer, and R. Bippus, *Quantile based noise estimation for spectral subtraction and wiener filtering*, in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process.* (2000), 1875-1878.
- [0710] [SS98] J. Sohn and W. Sung, *A voice activity detector employing soft decision based noise spectrum adaptation*, *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, no. pp. 365-368, IEEE, 1998.
- [0711] [Yu09] Rongshan Yu, *A low-complexity noise estimation algorithm based on smoothing of noise power estimation and estimation bias correction*, *Acoustics, Speech and Signal Processing*, 2009. ICASSP 2009. IEEE International Conference on, Apr 2009, pp. 4421-4424.

도면

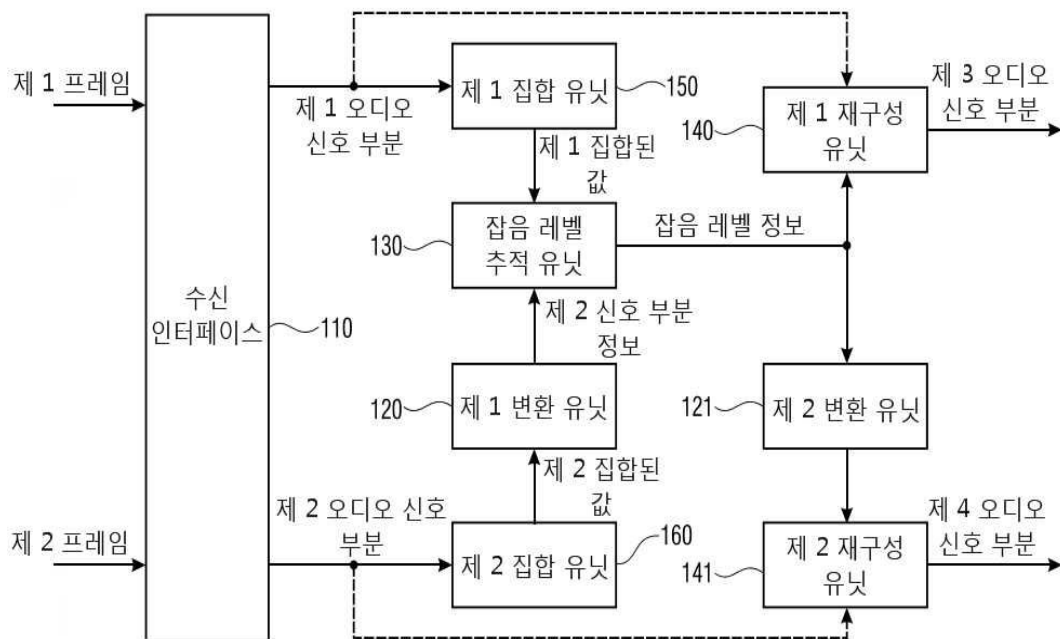
도면1a



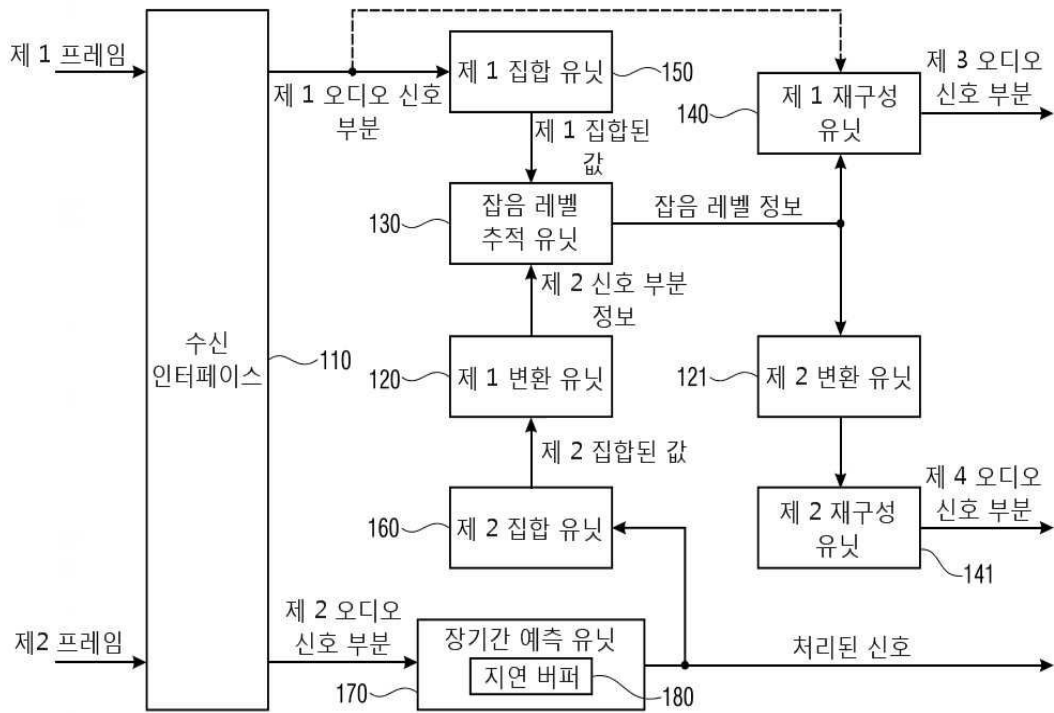
도면1b



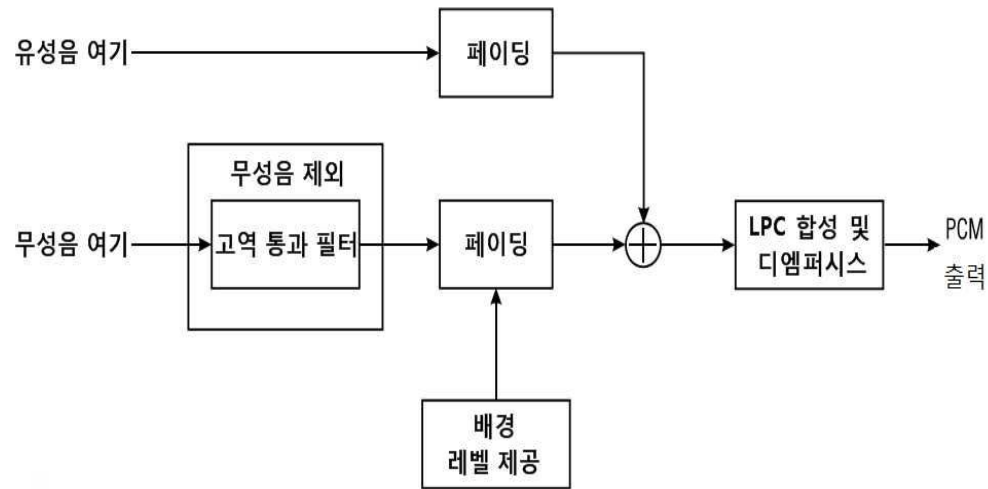
도면1c



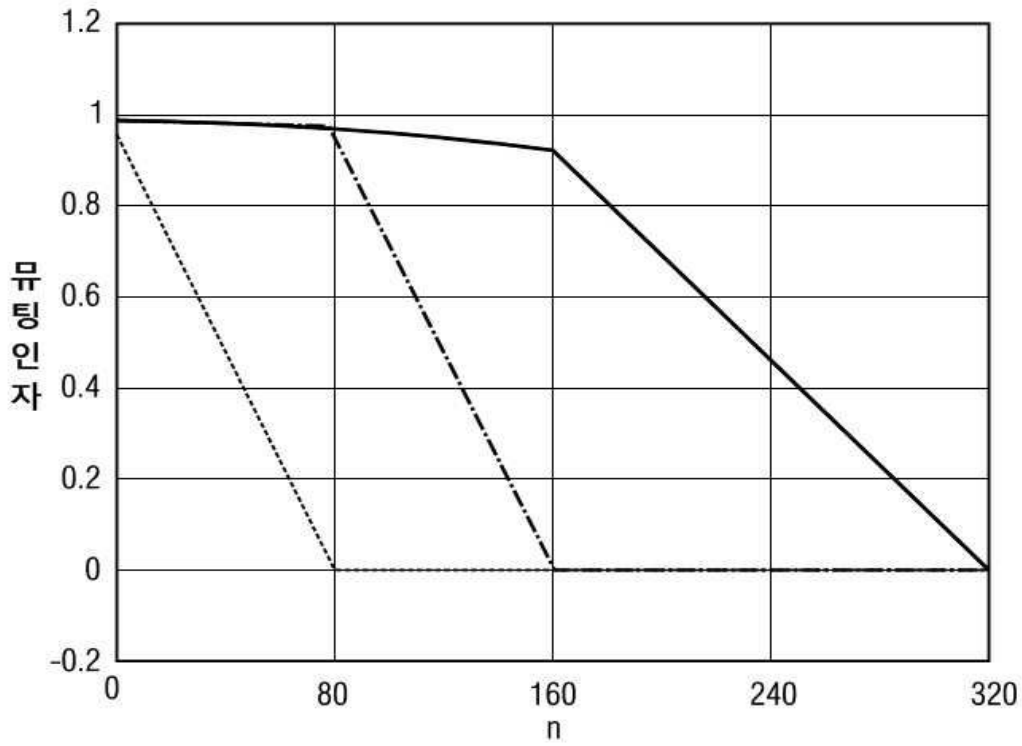
도면1d



도면2



도면3

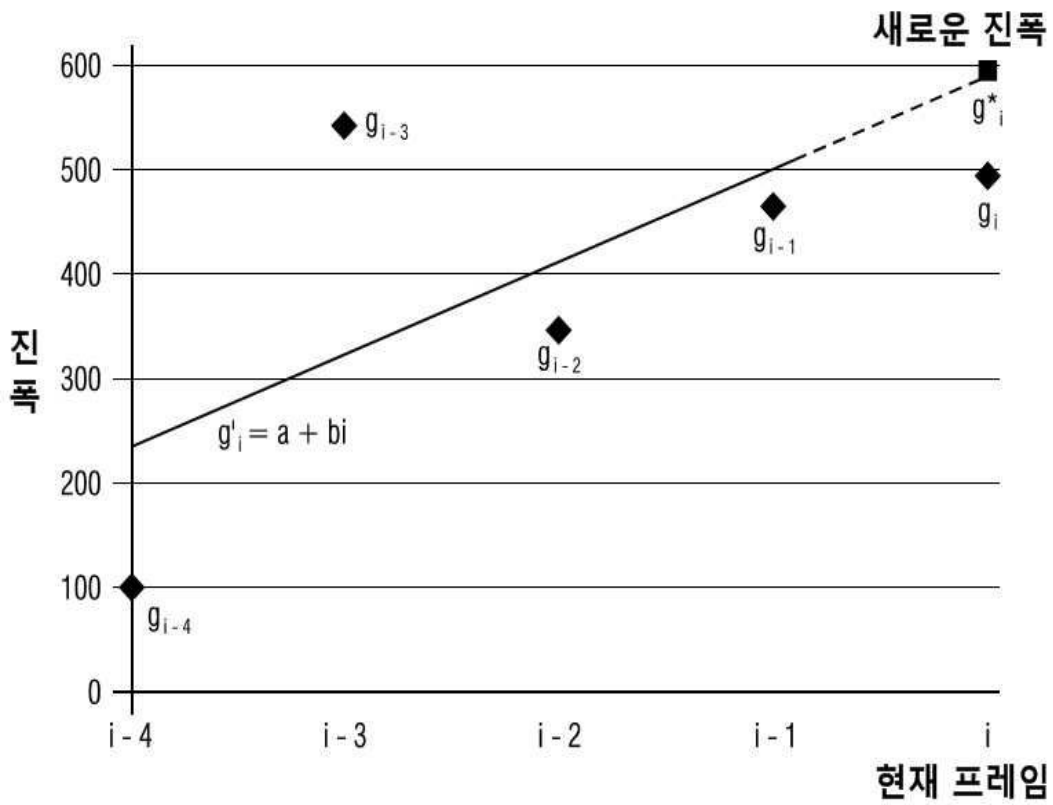


G.722APRIV\_F07

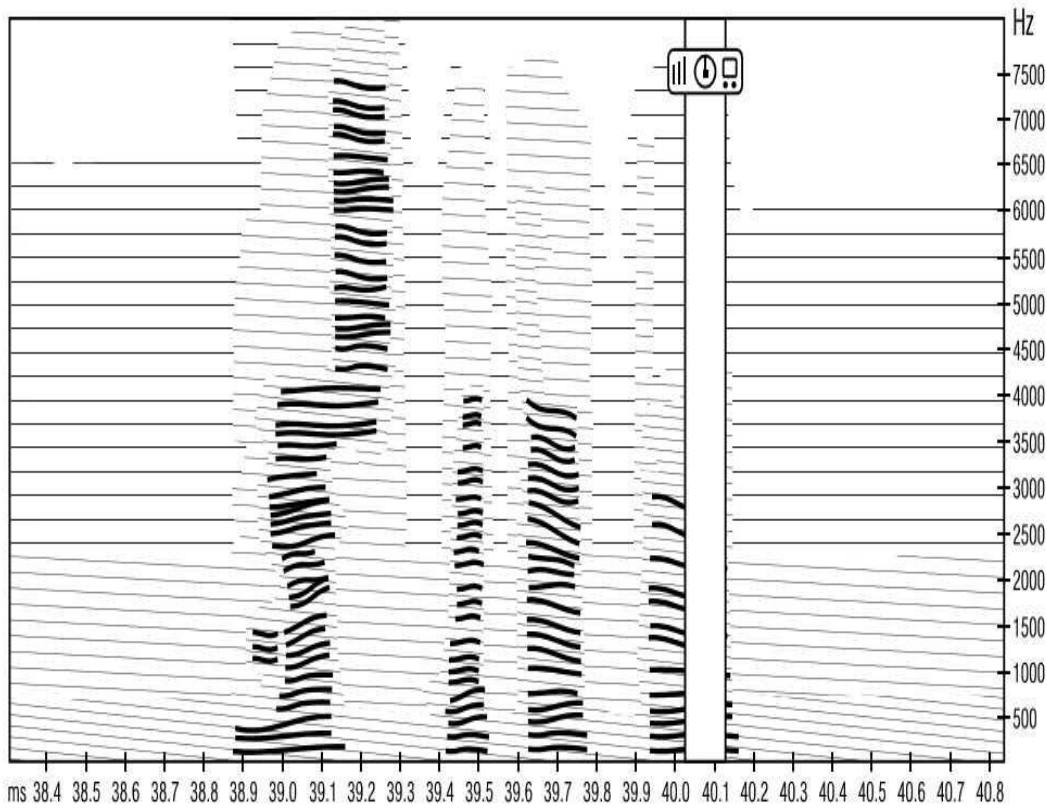
- ..... 과도
- · - · - VUV\_전이
- 다른 클래스들



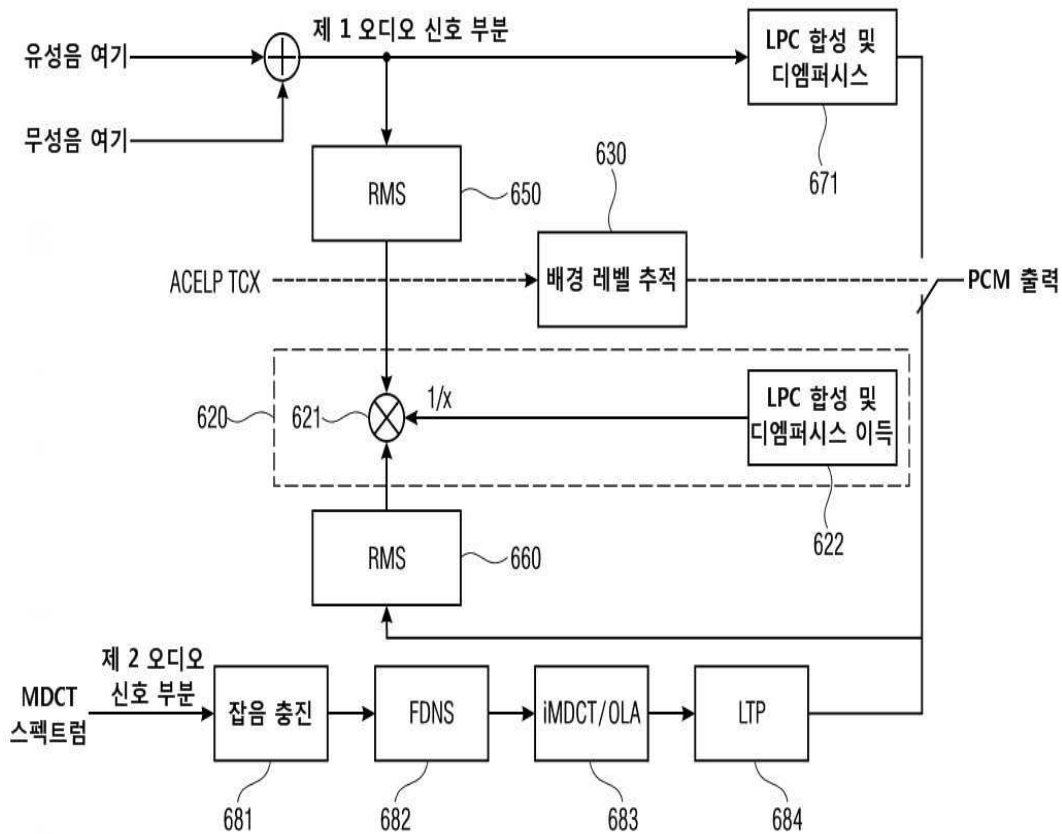
도면4



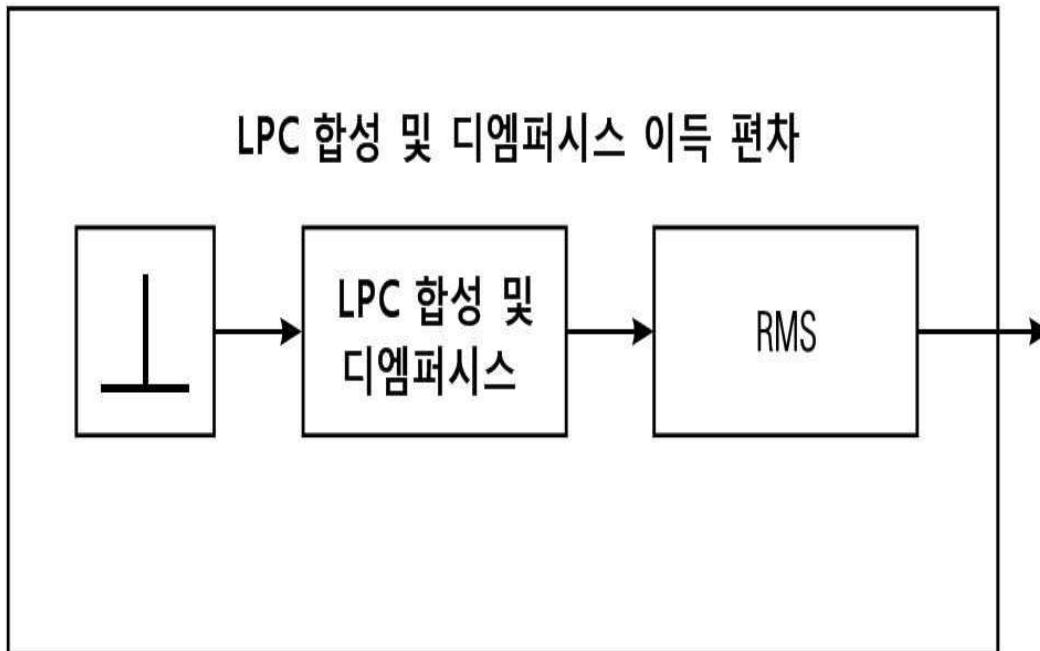
도면5



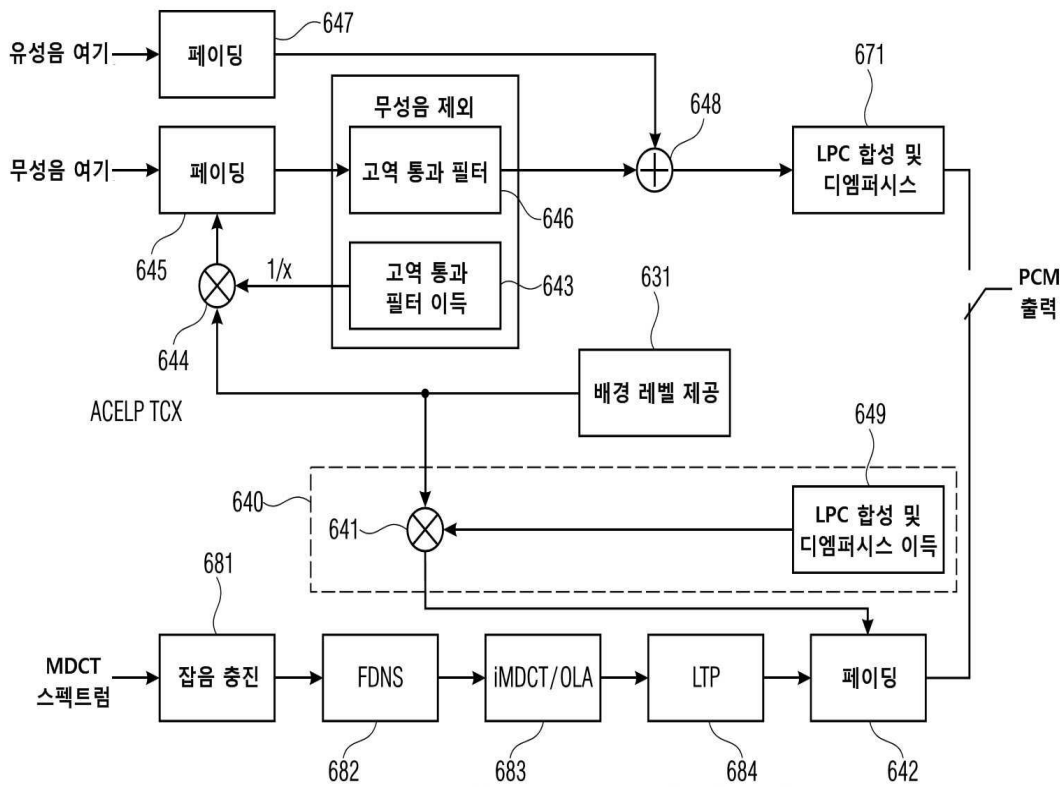
도면6



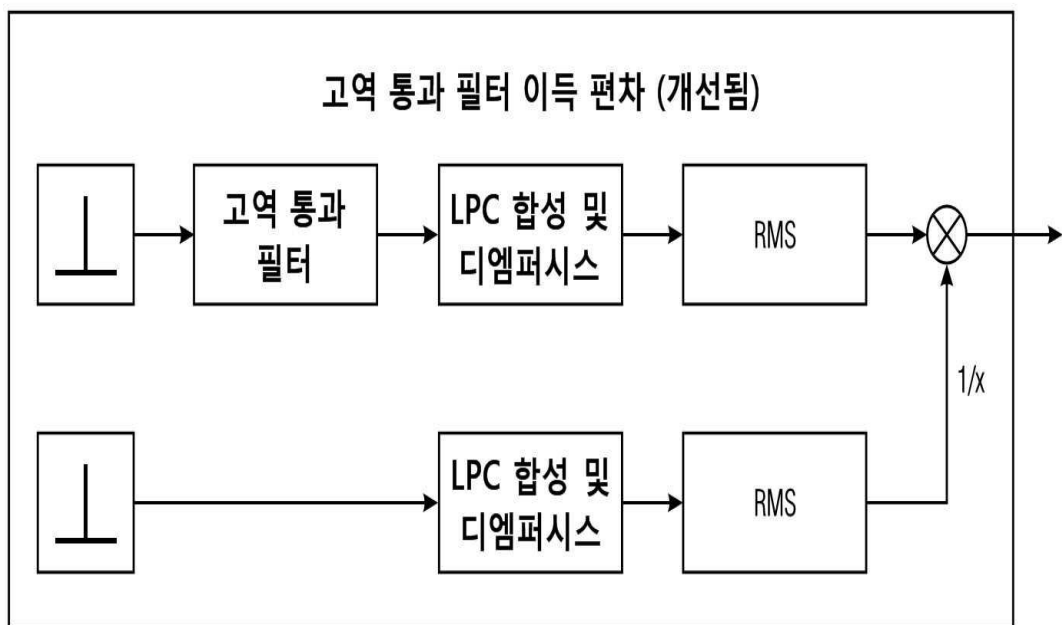
도면7



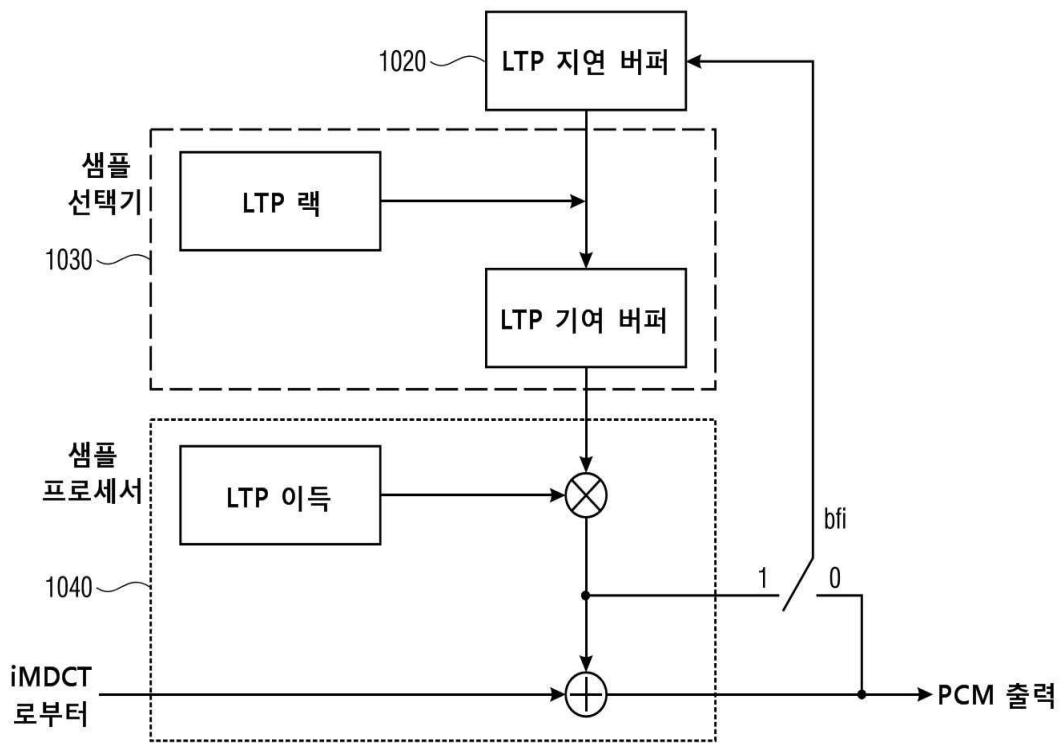
도면8



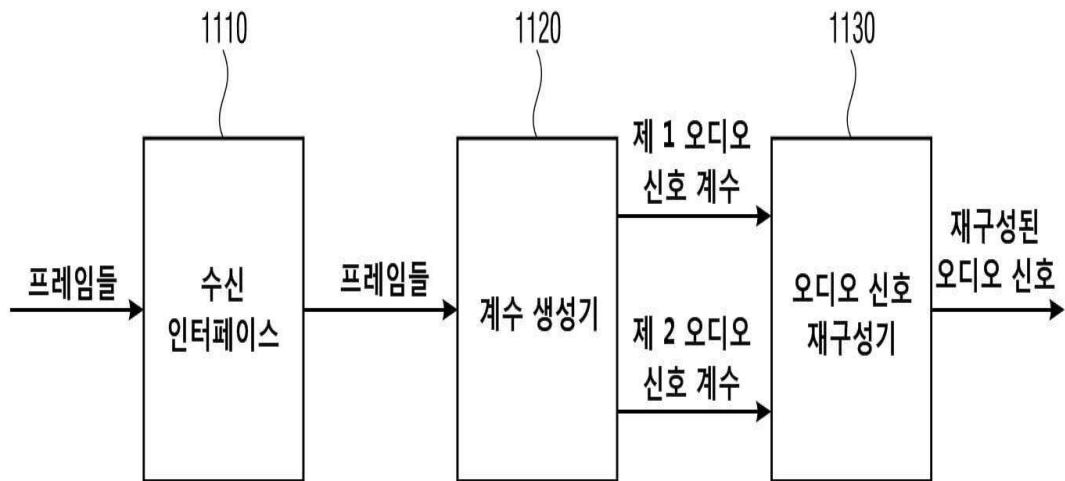
도면9



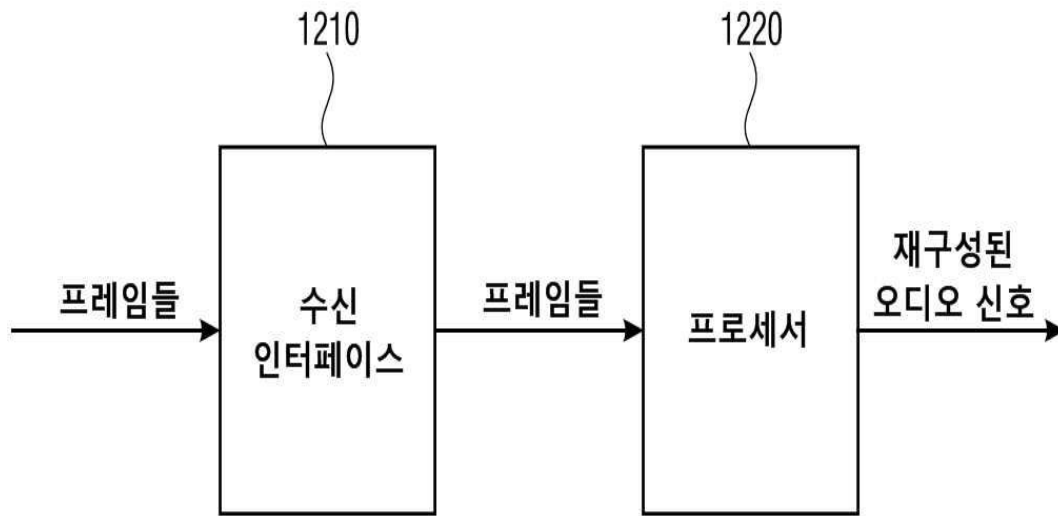
도면10



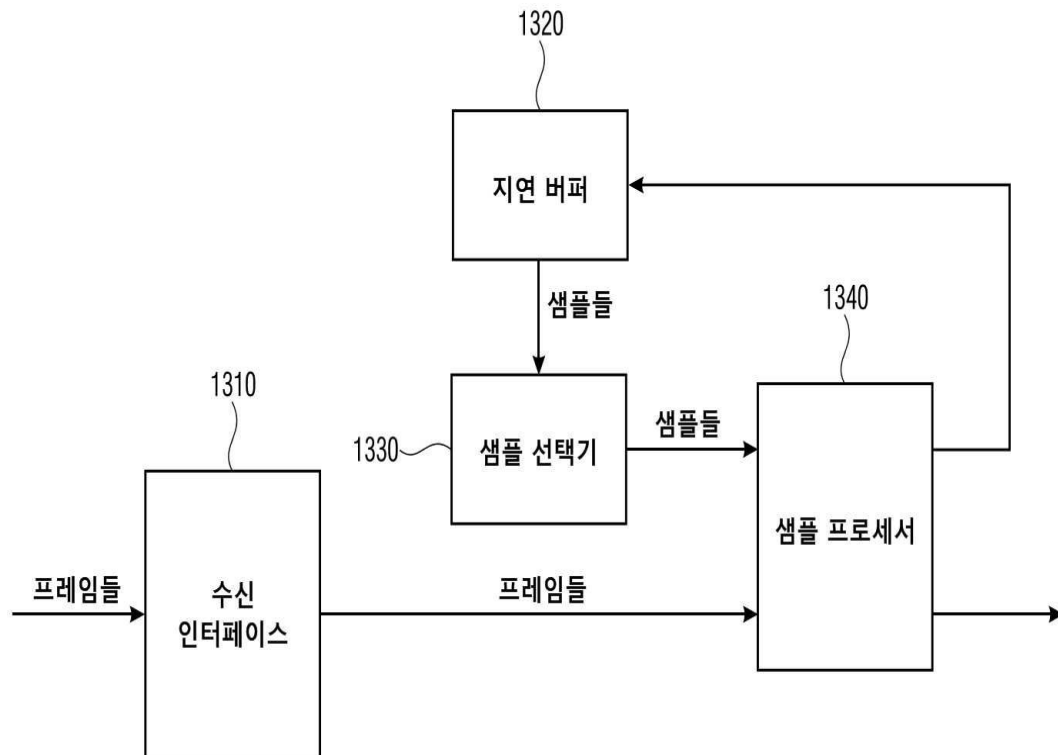
도면11



도면12



도면13



도면14

