(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0119151
(43) 공개일자 2015년10월23일

(51) 국제특허분류(Int. Cl.)
*G10L 21/038* (2013.01) *G10L 17/00* (2006.01)
*H04M 3/22* (2006.01) *H04M 9/08* (2006.01)
(52) CPC특허분류
*G10L 21/038* (2013.01)
*G10L 17/00* (2013.01)
(21) 출원번호 10-2015-7024776
(22) 출원일자(국제) 2014년02월11일
심사청구일자 없음
(85) 번역문제출일자 2015년09월10일
(86) 국제출원번호 PCT/US2014/015852
(87) 국제공개번호 WO 2014/126933
국제공개일자 2014년08월21일
(30) 우선권주장
13/768,946 2013년02월15일 미국(US)

(71) 출원인
**퀄컴 인코포레이티드**
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
(72) 발명자
**김 래훈**
미국 92121 캘리포니아주 샌디에고 모어하우스 드라이브 5775
**류 상욱**
미국 92121 캘리포니아주 샌디에고 모어하우스 드라이브 5775
**신 종원**
미국 92121 캘리포니아주 샌디에고 모어하우스 드라이브 5775
(74) 대리인
**특허법인코리아나**

전체 청구항 수 : 총 44 항

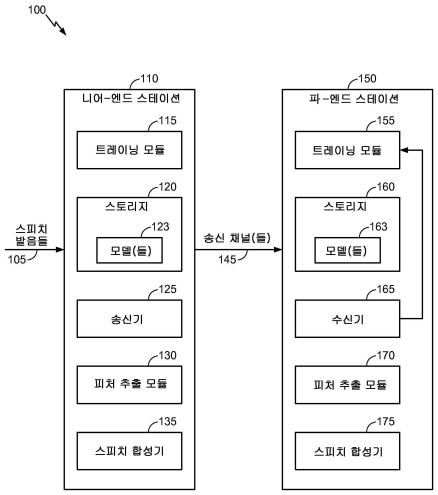
(54) 발명의 명칭 **개인화된 대역폭 확장**

## (57) 요 약

대역폭 확장을 위해 사용된 모델이 각각의 특정 사용자에게 개인화되는 (맞춰지는) 개인화된 (즉, 화자-도출가능한) 대역폭 확장이 제공된다. 트레이닝 페이즈는 사용자에게 개인화되는 대역폭 확장 모델을 생성하도록 수행된다. 이 모델은 사용자를 수반한 전화 호출 동안 대역폭 확장 페이즈에서 후속적으로 사용될 수도 있다.

개인화된 대역폭 확장 모델을 사용하는 대역폭 확장 페이즈는, 더 높은 대역 (예를 들어, 광대역) 이 이용 가능하지 않고 더 낮은 대역 (예를 들어, 협대역) 에서 호출이 발생하는 경우 활성화될 것이다.

*대 표 도* - 도1



(52) CPC특허분류
*H04M 3/22* (2013.01)
*H04M 9/08* (2013.01)

# 명 세 서

## 청구범위

### 청구항 1

대역폭 확장을 위한 방법으로서,

식별된 화자의 데이터로 대역폭 확장 모델을 생성하는 단계; 및

상기 식별된 화자를 포함하는 협대역 호출에 대해 상기 대역폭 확장 모델을 사용하여 대역폭 확장을 수행하는 단계를 포함하는, 대역폭 확장을 위한 방법.

### 청구항 2

제 1 항에 있어서,

상기 대역폭 확장 모델은 상기 식별된 화자의 상기 데이터로 개인화되는, 대역폭 확장을 위한 방법.

### 청구항 3

제 1 항에 있어서,

상기 대역폭 확장 모델은 화자-도출가능 정보에 기초하는, 대역폭 확장을 위한 방법.

### 청구항 4

제 1 항에 있어서,

상기 대역폭 확장 모델을 생성하는 단계는, 상기 식별된 화자의 상기 데이터로 상기 대역폭 확장 모델을 트레이닝하는 단계를 포함하는, 대역폭 확장을 위한 방법.

### 청구항 5

제 1 항에 있어서,

상기 대역폭 확장 모델을 생성하는 단계는, 상기 화자를 식별하는 단계 및 상기 화자의 음성 데이터로부터 광대역 피처들을 추출하는 단계를 포함하는, 대역폭 확장을 위한 방법.

### 청구항 6

제 5 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하는 단계는, 광대역 호출의 파-엔드 스테이션에서 상기 광대역 피처들을 추출하는 단계를 포함하는, 대역폭 확장을 위한 방법.

### 청구항 7

제 5 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하는 단계는, 호출의 니어-엔드 스테이션에서 상기 광대역 피처들을 추출하는 단계를 포함하는, 대역폭 확장을 위한 방법.

### 청구항 8

제 5 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하는 단계는, 오프-라인 이동국에서 상기 광대역 피처들을 추출하는 단계를 포함하는, 대역폭 확장을 위한 방법.

### 청구항 9

제 1 항에 있어서,

상기 협대역 호출에 대해 대역폭 확장을 수행하는 단계는,

상기 화자를 식별하는 단계;

상기 화자와 연관된 상기 대역폭 확장 트레이닝 모델을 스토리지로부터 취출하는 단계;

상기 협대역 호출로부터 협대역 피처들을 추출하는 단계; 및

추출된 상기 협대역 피처들과 결합된 상기 대역폭 확장 트레이닝 모델을 사용하여 광대역 스피치 합성을 수행하는 단계를 포함하는, 대역폭 확장을 위한 방법.

**청구항 10**

제 9 항에 있어서,

상기 광대역 스피치 합성을 수행한 후에 광대역 스피치를 출력하는 단계를 더 포함하는, 대역폭 확장을 위한 방법.

**청구항 11**

제 1 항에 있어서,

상기 협대역 호출에 대해 대역폭 확장을 수행하는 단계는, 광대역이 이용 가능하지 않고 상기 호출이 협대역에서 발생하는 경우 활성화되는, 대역폭 확장을 위한 방법.

**청구항 12**

대역폭 확장을 위한 장치로서,

식별된 화자의 데이터로 대역폭 확장 모델을 생성하기 위한 수단; 및

상기 식별된 화자를 포함하는 협대역 호출에 대해 상기 대역폭 확장 모델을 사용하여 대역폭 확장을 수행하기 위한 수단을 포함하는, 대역폭 확장을 위한 장치.

**청구항 13**

제 12 항에 있어서,

상기 대역폭 확장 모델은 상기 식별된 화자의 상기 데이터로 개인화되는, 대역폭 확장을 위한 장치.

**청구항 14**

제 12 항에 있어서,

상기 대역폭 확장 모델은 화자-도출가능 정보에 기초하는, 대역폭 확장을 위한 장치.

**청구항 15**

제 12 항에 있어서,

상기 대역폭 확장 모델을 생성하기 위한 수단은, 상기 식별된 화자의 상기 데이터로 상기 대역폭 확장 모델을 트레이닝하기 위한 수단을 포함하는, 대역폭 확장을 위한 장치.

**청구항 16**

제 12 항에 있어서,

상기 대역폭 확장 모델을 생성하기 위한 수단은, 상기 화자를 식별하기 위한 수단 및 상기 화자의 음성 데이터로부터 광대역 피처들을 추출하기 위한 수단을 포함하는, 대역폭 확장을 위한 장치.

**청구항 17**

제 16 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하기 위한 수단은, 광대역 호출의 파-엔드 스테이션에서 상기 광대역 피처들을 추출하기 위한 수단을 포함하는, 대역폭 확장을 위한 장치.

**청구항 18**

제 16 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하기 위한 수단은, 호출의 니어-엔드 스테이션에서 상기 광대역 피처들을 추출하기 위한 수단을 포함하는, 대역폭 확장을 위한 장치.

**청구항 19**

제 16 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하기 위한 수단은, 오프-라인 이동국에서 상기 광대역 피처들을 추출하기 위한 수단을 포함하는, 대역폭 확장을 위한 장치.

**청구항 20**

제 12 항에 있어서,

상기 협대역 호출에 대해 대역폭 확장을 수행하기 위한 수단은,

상기 화자를 식별하기 위한 수단;

상기 화자와 연관된 상기 대역폭 확장 트레이닝 모델을 스토리지로부터 취출하기 위한 수단;

상기 협대역 호출로부터 협대역 피처들을 추출하기 위한 수단; 및

추출된 상기 협대역 피처들과 결합된 상기 대역폭 확장 트레이닝 모델을 사용하여 광대역 스피치 합성을 수행하기 위한 수단을 포함하는, 대역폭 확장을 위한 장치.

**청구항 21**

제 20 항에 있어서,

상기 광대역 스피치 합성을 수행한 후에 광대역 스피치를 출력하기 위한 수단을 더 포함하는, 대역폭 확장을 위한 장치.

**청구항 22**

제 12 항에 있어서,

상기 협대역 호출에 대해 대역폭 확장을 수행하는 것은, 광대역이 이용 가능하지 않고 상기 호출이 협대역에서 발생하는 경우 활성화되는, 대역폭 확장을 위한 장치.

**청구항 23**

명령들을 포함하는 컴퓨터 판독가능 매체로서,

상기 명령들은 컴퓨터로 하여금,

식별된 화자의 데이터로 대역폭 확장 모델을 생성하게 하며;

상기 식별된 화자를 포함하는 협대역 호출에 대해 상기 대역폭 확장 모델을 사용하여 대역폭 확장을 수행하게 하는, 컴퓨터 판독가능 매체.

**청구항 24**

제 23 항에 있어서,

상기 대역폭 확장 모델은 상기 식별된 화자의 상기 데이터로 개인화되는, 컴퓨터 판독가능 매체.

**청구항 25**

제 23 항에 있어서,

상기 대역폭 확장 모델은 화자-도출가능 정보에 기초하는, 컴퓨터 판독가능 매체.

**청구항 26**

제 23 항에 있어서,

상기 컴퓨터로 하여금 상기 대역폭 확장 모델을 생성하게 하는 명령들은, 상기 컴퓨터로 하여금 상기 식별된 화자의 상기 데이터로 상기 대역폭 확장 모델을 트레이닝하게 하는 명령들을 포함하는, 컴퓨터 판독가능 매체.

**청구항 27**

제 23 항에 있어서,

상기 컴퓨터로 하여금 상기 대역폭 확장 모델을 생성하게 하는 명령들은, 상기 컴퓨터로 하여금 상기 화자를 식별하게 하고 상기 화자의 음성 데이터로부터 광대역 피처들을 추출하게 하는 명령들을 포함하는, 컴퓨터 판독가능 매체.

**청구항 28**

제 27 항에 있어서,

상기 컴퓨터로 하여금 상기 화자의 음성 데이터로부터 광대역 피처를 추출하게 하는 명령들은, 상기 컴퓨터로 하여금 광대역 호출의 파-엔드 스테이션에서 상기 광대역 피처들을 추출하게 하는 명령들을 포함하는, 컴퓨터 판독가능 매체.

**청구항 29**

제 27 항에 있어서,

상기 컴퓨터로 하여금 상기 화자의 음성 데이터로부터 광대역 피처를 추출하게 하는 명령들은, 상기 컴퓨터로 하여금 호출의 니어-엔드 스테이션에서 상기 광대역 피처들을 추출하게 하는 명령들을 포함하는, 컴퓨터 판독가능 매체.

**청구항 30**

제 27 항에 있어서,

상기 컴퓨터로 하여금 상기 화자의 음성 데이터로부터 광대역 피처를 추출하게 하는 명령들은, 상기 컴퓨터로 하여금 오프-라인 이동국에서 상기 광대역 피처들을 추출하게 하는 명령들을 포함하는, 컴퓨터 판독가능 매체.

**청구항 31**

제 23 항에 있어서,

상기 컴퓨터로 하여금 상기 협대역 호출에 대해 대역폭 확장을 수행하게 하는 명령들은, 상기 컴퓨터로 하여금,

상기 화자를 식별하게 하고;

상기 화자와 연관된 상기 대역폭 확장 트레이닝 모델을 스토리지로부터 취출하게 하고;

상기 협대역 호출로부터 협대역 피처들을 추출하게 하며;

추출된 상기 협대역 피처들과 결합된 상기 대역폭 확장 트레이닝 모델을 사용하여 광대역 스피치 합성을 수행하게 하는 명령들을 포함하는, 컴퓨터 판독가능 매체.

**청구항 32**

제 31 항에 있어서,

상기 컴퓨터로 하여금 상기 광대역 스피치 합성을 수행한 후에 광대역 스피치를 출력하게 하는 명령들을 더 포함하는, 컴퓨터 판독가능 매체.

**청구항 33**

제 23 항에 있어서,

상기 협대역 호출에 대해 대역폭 확장을 수행하는 것은, 광대역이 이용 가능하지 않고 상기 호출이 협대역에서 발생하는 경우 활성화되는, 컴퓨터 판독가능 매체.

**청구항 34**

대역폭 확장을 위한 시스템으로서,

식별된 화자의 데이터로 대역폭 확장 모델을 생성하도록 구성된 트레이닝 모듈; 및

상기 식별된 화자를 포함하는 협대역 호출에 대해 상기 대역폭 확장 모델을 사용하여 대역폭 확장을 수행하기 위한 프로세서를 포함하는, 대역폭 확장을 위한 시스템.

**청구항 35**

제 34 항에 있어서,

상기 대역폭 확장 모델은 상기 식별된 화자의 상기 데이터로 개인화되는, 대역폭 확장을 위한 시스템.

**청구항 36**

제 34 항에 있어서,

상기 대역폭 확장 모델은 화자-도출가능 정보에 기초하는, 대역폭 확장을 위한 시스템.

**청구항 37**

제 34 항에 있어서,

상기 트레이닝 모듈은 상기 식별된 화자의 상기 데이터로 상기 대역폭 확장 모델을 트레이닝하도록 구성되는, 대역폭 확장을 위한 시스템.

**청구항 38**

제 34 항에 있어서,

상기 화자를 식별하고, 상기 화자의 음성 데이터로부터 광대역 피처들을 추출하도록 구성된 피처 추출 모듈을 더 포함하는, 대역폭 확장을 위한 시스템.

**청구항 39**

제 38 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하는 것은, 광대역 호출의 파-엔드 스테이션에서 상기 광대역 피처들을 추출하는 것을 포함하는, 대역폭 확장을 위한 시스템.

**청구항 40**

제 38 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하는 것은, 호출의 니어-엔드 스테이션에서 상기 광대역 피처들을 추출하는 것을 포함하는, 대역폭 확장을 위한 시스템.

**청구항 41**

제 38 항에 있어서,

상기 화자의 음성 데이터로부터 광대역 피처를 추출하는 것은, 오프-라인 이동국에서 상기 광대역 피처들을 추출하는 것을 포함하는, 대역폭 확장을 위한 시스템.

**청구항 42**

제 34 항에 있어서,

상기 프로세서는, 상기 화자를 식별하고, 상기 화자와 연관된 상기 대역폭 확장 트레이닝 모델을 스토리지로부터 취출하도록 구성되고,

상기 시스템은,

상기 협대역 호출로부터 협대역 피처들을 추출하도록 구성된 피처 추출 모듈; 및

추출된 상기 협대역 피처들과 결합된 상기 대역폭 확장 트레이닝 모델을 사용하여 광대역 스피치 합성을 수행하도록 구성된 스피치 합성기를 더 포함하는, 대역폭 확장을 위한 시스템.

**청구항 43**

제 42 항에 있어서,

상기 프로세서는 또한, 상기 광대역 스피치 합성을 수행한 후에 광대역 스피치를 출력하도록 구성되는, 대역폭 확장을 위한 시스템.

**청구항 44**

제 34 항에 있어서,

상기 협대역 호출에 대해 대역폭 확장을 수행하는 것은, 광대역이 이용 가능하지 않고 상기 호출이 협대역에서 발생하는 경우 활성화되는, 대역폭 확장을 위한 시스템.

## *발명의 설명*

### *배 경 기 술*

[0001] 신호들의 대역폭 확장은 잘 알려져 있고, 유용한 콘텐트 또는 효과들을 포함하는 신호의 주파수 범위 (대역폭)를 확장하는 프로세스로서 정의될 수도 있다. 최근에, 대역폭 확장은 전화기 시스템들에서의 스피치와 같은 코딩된 스피치 및 오디오의 고 주파수 향상에서 사용되고 있다. 따라서, 예를 들어 협대역 전화 스피치의 품질은 대역폭 확장에 의해 개선될 수 있다.

[0002] 알려진 대역폭 확장의 하나의 타입은 블라인드 대역폭 확장으로서 지칭된다. 블라인드 대역폭 확장 기법들이 사용되어 협대역 오디오 신호의 고 주파수 대역들을 재생성한다. 이들 방법들은 전화 또는 라디오 통신 시스템들에서와 같이 협대역 주파수 대역을 통해 송신되는 신호들의 감지된 품질을 증가시킨다.

[0003] 그러나, 기존의 블라인드 대역폭 확장 기법들은 대역폭 확장을 수행하는 일반화된 모델들을 사용한다. 이러한 일반화된 모델들은 통상적으로, 임의의 화자들에 대해 대역폭 확장에서 사용되고, 이것은 부정확한 또는 열악한 사용자 경험을 초래할 수도 있다.

## *발명의 내용*

### *과제의 해결 수단*

[0004] 본 시스템들 및 방법들은, 대역폭 확장에 사용된 모델이 각각의 특정 사용자에게 개인화 (예를 들어, 맞춰짐)되는 개인화된 (즉, 화자-도출가능) 대역폭 확장에 관한 것이다.

[0005] 일 구현에서, 트레이닝 페이즈가 수행되어 사용자에게 개인화되는 대역폭 확장 모델을 생성한다. 이 모델은 사용자를 포함하는 전화 통화 동안 대역폭 확장 페이즈에서 후속적으로 사용될 수도 있다. 개인화된 대역폭 확장 모델을 사용하는, 대역폭 확장 페이즈는, 고 대역 (예를 들어, 광대역) 이 이용 가능하지 않고 저 대역 (예를 들어, 협대역) 에서 호출 (call) 이 발생할 때 활성화될 것이다.

[0006] 일 구현에서, 모델은 사용자가 (예를 들어, 호출자 (caller) ID 에 의해) 식별될 수 있는 광대역 호출 동안 파-엔드 스테이션 (예를 들어, 모바일 폰) 에서 트레이닝된다 (화자-도출가능 정보에 기초하도록 사용자에게 개인화된다).

[0007] 다른 구현에서, 모델은 호출 동안 (니어-엔드는 네트워크 컨디션에 관계 없이 광대역 스피치를 캡처할 수 있음) 니어-엔드 스테이션 (예를 들어, 사용자의 모바일 폰) 에서 트레이닝된다.

[0008] 다른 구현에서, 모델은, 이동국의 오너에 의해 선택된 타겟 음성로 또는 사용자의 음성로 (전화 호출 동안이 아닌) 이동국에서 오프 라인으로 트레이닝된다. 오프-라인 트레이닝은, 전화가 전화 호출을 위해 사용되고 있지 않은 경우 사용자의 음성 또는 타겟 음성을 레코딩하는 것을 수반한다.

[0009] 이 요약은 상세한 설명에서 이하에서 추가로 설명되는 단순화된 형태로 개념들의 선택을 도입하기 위해 제공된다. 이 요약은 청구된 주제의 중요한 피처들 또는 필수적인 피처들을 식별하도록 의도되지도 않고, 청구된 주제의 범위를 제한하는데 사용되도록 의도되지도 않는다.

## *도면의 간단한 설명*

[0010] 상기 요약, 뿐만 아니라 예시적인 실시형태들에 대한 다음의 상세한 설명은 첨부된 도면들과 함께 읽을 경우 더 잘 이해된다. 실시형태들을 예시할 목적으로, 도면들에서 실시형태들의 예시적인 구성들이 도시된다; 그러나, 실시형태들은 개시된 특정 방법들 및 수단들로 한정되지 않는다. 도면에서,

도 1 은 개인화된 대역폭 확장이 발생되어 사용될 수도 있는 예시적인 시스템의 블록도이다.

도 2 는 개인화된 대역폭 확장을 위한 트레이닝 방법의 구현의 동작 흐름이다.

도 3 은 개인화된 대역폭 확장을 수행하는 방법의 구현의 동작 흐름이다.

도 4 는 개인화된 대역폭 확장을 위한 트레이닝 방법의 다른 구현의 동작 흐름이다.

도 5 는 개인화된 대역폭 확장을 위한 트레이닝 방법의 다른 구현의 동작 흐름이다.

도 6 은 개인화된 대역폭 확장을 위한 트레이닝 방법의 다른 구현의 동작 흐름이다.

도 7 은 스펙트럼 인벨로프/여기-기반의 개인화된 대역폭 확장을 위해 트레이닝을 수행하는 방법의 구현의 동작 흐름이다.

도 8 은 스펙트럼 인벨로프/여기-기반의 개인화된 대역폭 확장과 사용될 수도 있는 예시적인 맵핑 테이블의 도면이다.

도 9 는 스펙트럼 인벨로프/여기-기반의 개인화된 대역폭 확장과 사용될 수도 있는 다른 예시적인 맵핑 테이블의 도면이다.

도 10 은 스펙트럼 인벨로프/여기-기반의 개인화된 대역폭 확장을 수행하는 방법의 구현의 동작 흐름이다.

도 11 은 예시적인 이동국의 다이아그램이다.

도 12 는 예시적인 컴퓨팅 환경을 나타낸다.

## *발명을 실시하기 위한 구체적인 내용*

[0011] 도면들을 참조하고 포함하는 다음의 상세한 설명은 하나 이상의 구체적인 실시형태들을 설명하고 예시한다. 제한하는 것이 아니라 단지 예시하고 교시하기 위해 제공되는 이들 실시형태들은 충분히 상세하게 도시되고 설명되어 당업자들이 청구된 것을 실시하는 것을 가능하게 한다. 따라서, 간단함을 위해, 설명은 당업자들에게 알려진 소정의 정보를 생략할 수도 있다.

[0012] 도 1 은 개인화된 대역폭 확장이 생성되고 (화자-도출가능 정보에 기초함) 사용되어 수신된 협대역 (NB) 스피치 신호들로부터 광대역 (WB) 스피치 신호들을 개발하는 예시적인 시스템 (100) 의 블록도이다. 시스템 (100) 은 대응하는 수신된 아날로그 신호들의 주파수 범위 보다 더 넓은 주파수 범위를 갖는 아날로그 신호들을 생성하는데 사용될 수도 있다. 이와 같이, 신호가 광대역 신호인지 또는 협대역 신호인지 여부는 다른 것들에 대한 그 관계에 의존한다. 본원에 설명된 예들 및 구현들은 NB 스피치 신호들로부터 WB 스피치 신호들을 전개하는 것을 지칭하지만, 본 시스템들 및 방법들은 이에 제한되지 않으며, 본원에 개시된 시스템들 및 방법들은 슈퍼 광대역 (SWB) 으로부터 WB 또는 SWB 로부터 NB 로 가는 것과 같은 스피치 신호들의 임의의 다른 세트로부터 스피치 신호들의 임의의 세트를 전개하는데 사용될 수 있다.

[0013] 도시된 바와 같이, 시스템 (100) 은 수신 파티에 의해 사용될 수도 있는 파-엔드 (far-end) 스테이션 (150) 및

송신 파티에 의해 사용될 수도 있는 니어-엔드 (near-end) 스테이션 (110) 을 포함한다. 니어-엔드 스테이션 (110) 및 파-엔드 스테이션 (150) 은 각각 이동국 또는 컴퓨팅 디바이스와 같은 통신하기 위한 디바이스일 수도 있다. 예시의 이동국 (800) 은 도 11 에 대하여 설명되고, 예시의 컴퓨팅 디바이스 (900) 는 도 12 에 대하여 설명된다.

[0014] 니어-엔드 스테이션 (110) 및 파-엔드 스테이션 (150) 은 하나 이상의 송신 채널들 (145) 을 통해 통신할 수도 있다. 통신은, 화자의 스피치 발음들 (105)(음성 데이터) 의 협대역 버전에 대응하는 협대역 스피치 신호의, 니어-엔드 스테이션 (110) 으로부터 파-엔드 스테이션 (150) 으로의 송신을 포함할 수도 있다. 파-엔드 스테이션 (150) 은 화자의 스피치 발음들 (105) 의 대응하는 광대역 버전들을 생성 및 출력하도록 화자에게 트레이닝된 (즉, 화자-도출가능한 정보에 기초한) 모델 (163) 을 사용할 수도 있다. 본원에 추가로 설명된 바와 같이, 모델 (123, 163) 은 구현에 따라, 니어-엔드 스테이션 (110) 또는 파-엔드 스테이션 (150) 에서 트레이닝, 생성, 및 저장될 수도 있다.

[0015] 따라서, 스피치 발음들 (105) 및 트레이닝된 모델 (123, 163) 또는 데이터 (구현에 의존) 는 하나 이상의 송신 채널들 (145) 을 통해 송신될 수도 있다. 보다 구체적으로, 송신기 (125) 는 화자의 스피치 발음들 (105)(음성 데이터) 을, 예를 들어 하나 이상의 텔레통신 송신 표준들에 따라 송신을 위한 원래의 스피치 발음들 (105) 의 협대역 버전으로 변환한다. 트레이닝된 모델 또는 데이터 및 원래의 스피치 발음들의 협대역 버전의 송신들은 일반적으로, 병행 방식으로 발생할 수도 있고, 또는 예를 들어 송신 프로세스 동안 별개의 시간들에서 발생할 수도 있다. 본 예 뿐만 아니라 이하에서 설명되는 예들에서 사용하기에 적합한 송신 채널들은 전화 네트워크 채널들, 무선 셀룰러 네트워크 채널들, 무선 워키-토키 시스템들, 유선 네트워크들 등을 포함한다. 이러한 송신 시스템들에서 사용된 협대역 스피치 신호들은 300 Hz-3.4 kHz 의 대역폭에 제한될 수도 있고, 이것은 모바일 통신 (GSM) 네트워크용 글로벌 시스템을 사용하여 스피치 신호들을 송신하는데 사용된 대역폭에 대응한다.

[0016] 구현에 따라, 니어-엔드 스테이션 (110) 은 트레이닝 모듈 (115), 하나 이상의 트레이닝된 모델들 (123) 을 저장할 수도 있는 스토리지 (120)(예를 들어, 메모리), 송신기 (125), 피처 추출 모듈 (130), 및 스피치 합성기 (135) 를 포함할 수도 있다. 유사하지만 동일하지는 않게, 구현에 따라, 파-엔드 스테이션 (150) 은 트레이닝 모듈 (155), 하나 이상의 트레이닝된 모델들 (163) 을 저장할 수도 있는 스토리지 (160)(예를 들어, 메모리), 수신기 (165), 피처 추출 모듈 (170), 및 스피치 합성기 (175) 를 포함할 수도 있다.

[0017] 일부 구현들에서, 니어-엔드 스테이션 (110) 에서 스피치 합성기 (135) 는 필요하거나 사용되지 않을 수도 있다. 예를 들어, 니어-엔드 스테이션 (110) 에서 단지 WB 입력으로부터의 피처 도메인 비교 메커니즘을 사용하는 구현에서, 실제 스피치가 생성되지 않을 수도 있고 따라서 스피치 합성기 (135) 는 필요하거나 사용되지 않을 수도 있다. 이러한 경우 (각각의 기존의 엔트리와 입력 간의 유사성을 컴퓨팅하기 위해 피처 도메인 유사성 비교를 사용함) 에서, 모델 및 피처 (예컨대, NB 피처) 는 파-엔드 스테이션 (150) 으로 전송되어 WB 스피치는 파-엔드 스테이션 (150) 에서 합성될 수 있다. 그러나, 실제 WB 입력 및 합성된 WB 입력을 비교함으로써 모델을 학습하기 위한 "합성에 의한 분석 (analysis-by-synthesis)" 메커니즘을 사용하는 구현에서, 스피치 합성기 (135) 가 사용될 수도 있다. 이 타입의 구현은 합성된 스피치 신호 도메인 유사성 비교를 사용하여, 각각의 기존의 엔트리와 입력 간의 유사성을 연산한다.

[0018] 스피치 발음들 (105) 은 니어-엔드 스테이션 (110) 의 사용자로부터 니어-엔드 스테이션 (110) 에서 수신된다. 스피치 발음들 (105) 은, 대역폭 확장을 위해 사용된 모델이 각각의 특정 사용자에게 개인화되는 (예를 들어, 맞춰지는) 개인화된 대역폭 확장을 제공하는데 사용된다.

[0019] 일 구현에서, 트레이닝 페이즈가 수행되어 사용자에게 개인화되는 대역폭 확장 모델 (또한, 본원에서 "모델" 로서 지칭됨) 을 생성한다. 트레이닝 페이즈는 트레이닝 모듈 (115) 을 사용하여 니어-엔드 스테이션 (110) 에서, 또는 트레이닝 모듈 (155) 을 사용하여 파-엔드 스테이션 (150) 에서 수행될 수도 있다. 트레이닝 모듈 (115) 에 의해 생성된 개인화된 대역폭 확장 모델 (123) 은 스토리지 (120) 에 저장될 수도 있고, 트레이닝 모듈 (155) 에 의해 생성된 모델 (163) 은 스토리지 (160) 에 저장될 수도 있다. 스토리지들 (120, 160) 은 (예를 들어, 상이한 사용자들에게 개인화된) 하나 보다 많은 모델을 저장할 수도 있고, 예를 들어 데이터베이스의 형태를 취할 수도 있다.

[0020] 구현에 따라, 니어-엔드 스테이션 (110) 및/또는 파-엔드 스테이션 (150) 은 피처 추출 모듈 (예를 들어, 피처 추출 모듈들 (130, 170) 각각) 을 포함할 수도 있다. 피처 추출 모듈은 사용자의 스피치로부터 광대역 피처들을 추출하고, 이들 광대역 피처들은 트레이닝 모델을 생성하는데 사용된다.

[0021] 트레이닝된 모델 (123, 163) 은 사용자의 전화기 (예를 들어, 니어-엔드 스테이션 (110)) 에 저장되거나, 다른 사용자의 전화기들 (예를 들어, 파-엔드 스테이션 (150)) 으로 전송될 수 있다. 광대역 트레이닝 데이터에 기초한, 모델 (123, 163) 은 그 후, 사용자를 수반하는 협대역 호출 동안 사용될 수도 있다. 모델은 전화 호출의 재생산 측 (즉, 파-엔드 스테이션 (150)) 에서 이용 가능해진다.

[0022] 따라서, 모델 (123) 또는 모델 (163) 과 같은 모델은 후속적으로, 모델과 연관된 사용자를 수반하는 전화 호출 동안 대역폭 확장 페이즈에서 사용될 수도 있다. 전화 호출은 하나 이상의 송신 채널들 (145) 을 통해 니어-엔드 스테이션 (110) 과 파-엔드 스테이션 (150) 사이에서 발생할 수도 있다. 특히, 니어-엔드 스테이션 (110) 의 송신기 (125) 는, 그것이 송신 채널(들)(145) 을 통해 파-엔드 스테이션의 수신기 (165) 로 수신했던 스피치를 송신할 수도 있다. 개인화된 대역폭 확장 모델을 사용하는, 대역폭 확장 페이즈는, 고 대역 (예를 들어, 광대역) 이 이용 가능하지 않고 저 대역 (예를 들어, 협대역) 에서 통화가 발생할 때 활성화될 것이다. 협대역 호출 동안, 수신기 (165) 는 (이전에 생성 및 저장된) 트레이닝된 모델 (123, 163) 과 함께 화자로부터의 협대역 스피치 신호들을 사용하여, 화자의 스피치 발음들 (105) 의 광대역 버전에 대응하는 광대역 스피치 신호를 생성한다.

[0023] 피처 추출 모듈 (130, 170) 은 또한, 전화 호출 동안 전화 호출로부터 협대역 피처들을 추출하고, 이들 협대역 피처들을 스피치 합성기 (135, 175) 에 제공할 수도 있다. 구현에 따라, 니어-엔드 스테이션 (110) 및/또는 파-엔드 스테이션 (150) 은 스피치 합성기 (예를 들어, 스피치 합성기들 (135, 175)) 를 포함할 수도 있다. 스피치 합성기 (135, 175) 는 추출된 협대역 피처들과 결합된 개인화된 대역폭 확장 트레이닝 모델을 사용하여 광대역 스피치 합성을 수행한다. 광대역 스피치는 그 후, 출력될 수도 있다.

[0024] 도 2 는 개인화된 대역폭 확장을 위한 트레이닝 방법 (200) 의 구현의 동작 흐름이다. 트레이닝 페이즈는 협대역 호출에서 후속적으로 사용될 수도 있는 모델을 생성하는데 사용된다. 이 구현에 따라, 본원에서 추가로 설명되는 바와 같이 트레이닝 페이즈는 광대역 호출, 또는 오프-라인 동안 니어-엔드 스테이션에서 인입 (incoming) 광대역 호출과 함께 파-엔드 스테이션에서 수행된다. 210 에서, 화자 (또한, 사용자로서 본원에 지칭됨) 는, 예를 들어 사용자의 이름 또는 다른 식별자를 스테이션에 엔터 또는 입력하는 사용자에 의해, 또는 예를 들어 호출자 ID 에 의해 식별된다.

[0025] 220 에서, 화자는, 스테이션에 의해 캡처되거나 다르게는 수신되는 발음들 (또한, 스피치 발음들로서 지칭됨) 을 말한다. 230 에서, 광대역 피처들은 발음들로부터 (예를 들어, 발음들에 대응하는 음성 데이터로부터) 추출된다. 피처 트레이닝은 240 에서 트레이닝 모델을 생성하도록 (충분한 데이터가 적절한 트레이닝 동안 수집될 때까지) 수행된다. 데이터는, 하나 이상의 미리결정된 컨디션들이 충족될 때까지, 예컨대 데이터가 충분히 다양하고 (다이버시티 임계 또는 기준을 충족) 및/또는 데이터가 긴 충분한 시간 (예를 들어, 구현에 따라 분, 시, 날들) 동안 수집될 때까지 수집될 수도 있다. 이 포인트에서, 250 에서 트레이닝 모델은 저장될 수도 있고, 협대역 호출에서 사용될 수도 있다.

[0026] 충분한 데이터가 개인화된 대역폭 확장에 대한 트레이닝 모델을 생성하도록 수집되지 않으면, 일반 대중에게 트레이닝되어 있는 대역폭 확장 모델 (즉, "일반화된 모델") 은 따라서 멀리서 수집되었던 개인화된 데이터로 적응될 수도 있다. 예를 들어, 테이블 (예컨대, 도 7 내지 도 10 에 대하여 이하에서 설명된 테이블) 가 일반화된 모델에 대한 데이터를 포함하고, 그 테이블이 개인화된 트레이닝 모델로서 (또는 이와 함께) 사용하기 위해 개인화된 데이터를 포함하는 테이블로 수정되는 구현에서, 그 테이블은 이에 따라 멀리서 수신되었던 개인화된 데이터를 사용하여 적응될 수도 있다. 이것은, 개인화된 데이터의 완전한 세트가 수집되고 (예를 들어, 전술된 하나 이상의 미리결정된 컨디션들이 충족되는 경우) 테이블로 통합되는 (즉, 수집되었던 개인화된 데이터의 완전한 세트를 사용하여 이 테이블이 수정되는 경우) 경우로의 더 매끄러운 트랜지션을 허용한다. 다른 구현에서, 개인화된 대역폭 확장을 위한 트레이닝 모델을 생성하기 위해 충분한 데이터가 수집되지 않으면, 충분한 개인화된 데이터가 수집될 때까지 일반화된 모델이 사용될 수도 있다.

[0027] 도 3 은 개인화된 대역폭 확장을 수행하는 방법 (300) 의 구현의 동작 흐름이다. 310 에서, 협대역 호출은 니어-엔드 스테이션과 파-엔드 스테이션 사이에서 발생하고, 대역폭 확장 페이즈가 시작된다. 320 에서 (예를 들어, 호출자 ID 를 사용하여) 화자가 식별되고, 330 에서 (그 화자와 연관된) 적합한 대역폭 확장 트레이닝 모델이 (예를 들어, 재생 스테이션 측에서, 예컨대 파-엔드 스테이션에서) 스토리지로부터 취출되거나 송신 스테이션 측 (예를 들어, 협대역 스피치 소스, 예컨대 니어-엔드 스테이션) 으로부터 수신된다. 화자 식별의 일부 방식이 이용되어, 현재 말하는 사람이 모델의 생성 동안 말했던 사람과 동일한 개인이도록 확실히 할 수도 있다. 화자가 식별될 수 없으면, 개인화된 대역폭 확장은 거절 또는 종료될 수도 있다.

[0028] 340 에서, 협대역 피처들은 전화 호출에서의 스피치 발음으로부터 (예를 들어, 음성 데이터로부터) 추출되고, 이 모델을 사용하는 화자의 개인화된 광대역 피처들로 맵핑된다. 광대역 스피치 합성은, 추출된 협대역 피처들과 결합된 개인화된 대역폭 확장 트레이닝 모델을 사용하여 350 에서 수행된다. 360 에서 광대역 스피치가 출력된다.

[0029] 도 4 는 개인화된 대역폭 확장을 위한 트레이닝 방법의 다른 구현의 동작 흐름이다. 이 구현에서, 모델은, 사용자가 (예를 들어, 호출자 ID 에 의해) 식별될 수 있는 광대역 호출 동안 파-엔드 스테이션 (150) 에서 트레이닝된다.

[0030] 410 에서, 화자의 음성 (예를 들어, 스피치 발음들) 는 니어-엔드 스테이션으로부터 광대역 호출 동안 파-엔드 스테이션에서 수신된다. 415 에서, 화자의 음성 또는 스피치 발음들의 음향 컨디션이 트레이닝에 사용되기에 매우 충분한 품질인지가 결정된다. 이 품질은 예를 들어, 신호대 잡음비 (SNR) 또는 다른 잡음-기반 통계치들 및/또는 측정치들에 기초할 수도 있다. 따라서, 예를 들어 SNR 이 미리결정된 임계 위이면, 발음들의 품질은 개인화된 대역폭 확장 트레이닝 모델의 생성에서 사용되기에 매우 충분하다. 그렇지 않으면, 프로세싱은 410 으로 리턴하여 음향 컨디션이 허용 가능할 때까지 계속한다.

[0031] 415 에서 음향 컨디션이 허용 가능한 것으로 결정되면, 420 에서 스피치 발음들로부터 협대역 및 광대역 피처들이 추출되고, 430 에서 화자, 파-엔드 스테이션의 사용자, 또는 호출자 ID (425) 에 의해 화자가 식별된다. 화자에 속하는 대역폭 확장 (BWE) 피처들은, 이것은 반복적인 프로세스이기 때문에, 임의의 대역폭 확장 피처들이 이전에 생성 및 저장되었으면, 435 에서 스토리지 (예컨대, 로컬 데이터베이스) 로부터 취출될 수도 있다.

[0032] 440 에서, 피처 트레이닝은 420 으로부터 추출된 피처들 및 435 로부터 취출된 대역폭 확장 피처들을 사용하여 수행된다. 이 피처 트레이닝은 피처 공간을 업데이트하는 것 및 화자에 대한 협대역-광대역 맵핑을 적응시키는 것을 포함할 수도 있다. 전술된 피처 도메인 유사성 비교, 전술된 합성된 스피치 신호 도메인 유사성 비교, 뿐만 아니라 (시간 도메인 또는 주파수 도메인에서) 허용 가능한 시간 조정을 갖는 최소 제곱법, (예컨대, LPC/LPCC, MFCC, 또는 오디오 핑거프린트들을 사용하는) 피처 기반 접근법, 또는 상위 순서 기반 접근법 (예컨대, 크로스-누적률 (cross-cumulant), 실증적인 쿨백-라이블러 발산, 또는 이타쿠라-사이토 거리 (Itakura-Saito distance)) 를 포함하지만 이에 한정되지는 않는, 임의의 거리 메트릭과 같은 유사성 측정 및 매칭을 결정하는 임의의 기법이 사용될 수도 있다.

[0033] 일 구현에서, 초기화 피처 공간 (예를 들어, 일반 대중에 대한 유니버셜 코드북) 으로부터 업데이트하는 것은 개인화된 피처 공간 (예를 들어, 화자에 대해 개인화된 코드북) 을 향해 수행될 수도 있다. 코드북 엔트리는 현재 피처 벡터와 최근접 기존의 코드북 엔트리 간의 가중된 합으로 대체될 수도 있다. 일 구현에서, 너무 많은 메모리 소비가 이슈가 되면, 메모리 소비의 우선순위는 유니버셜 피처 공간으로부터 업데이트 컨디션을 용이하게 함으로써 소정 화자들에게 제공될 수 있다. 대안으로 또는 부가적으로, 단지 상이한 피처 공간 또는 단지 유니버셜 피처 공간으로부터의 델타 또는 변화가 저장될 수도 있다. 일반화된 피처 공간은 스테이션에서 미리-저장될 수 있거나 또는 후속의 사용을 위해 스테이션으로 전송되거나 다르게는 제공될 수 있다.

[0034] 피처 공간의 업데이트가 없더라도, 개인화된 대역폭 확장은 개인화된 협대역-광대역 맵핑 모델을 사용하여 여전히 획득될 수도 있다는 것이 고려된다.

[0035] 450 에서, 대역폭 확장 피처 공간 및 화자에 대한 맵핑은 (예를 들어, 트레이닝 모델의 완료 전에 발생하게 될 맵핑 및/또는 피처들의 양의 미리결정된 임계에 대해) 트레이닝의 완료 퍼센티지로 (예를 들어, 로컬 데이터베이스에) 저장될 수도 있다. 일 구현에서, 모델 트레이닝 완료 퍼센티지는, 넌-스피치 부분을 배제하도록 적절한 VAD (voice activity detection) 로 트레이닝하기 위해 사용된 시간을 측정하는 것, 및 얼마나 많은 새로운 협대역-광대역 맵핑들이 구성되는지의 카운트를 측정하는 것에 의해, 또는 이에 기초하여 획득될 수 있다.

[0036] 일 구현에서, 인디케이터는 사용자에게 액세스 가능한, 주소록에 제공되어, 광대역이 트레이닝 완료의 일부 퍼센티지로 이용 가능하다는 것을 나타낼 수 있다. 주소록은, 예를 들어 광대역이 각각의 사용자 및/또는 트레이닝 완료의 퍼센티지에 대해 이용 가능한지 여부를 나타내기 위한 표시자 (indicator) 및 각종 사용자들을 저장 및 나열하는 데이터베이스를 포함할 수도 있다. 사용자는, 트레이닝 완료 퍼센티지와 사용자 인터페이스에 기초하여 광대역으로의 대역폭 확장을 사용할지 여부를 결정할 수 있다. 대안으로 또는 부가적으로, 트레이닝 완료 퍼센티지가 소정 레벨 (예를 들어, 99%) 에 도달할 때는 언제든지, 광대역이 이용 가능하지 않은

경우, 또는 예를 들어 일부 코딩 리소스가 절감 또는 보존되는 경우 (예를 들어, 호출 동안 나쁜 네트워크 컨디션 또는 헤비 데이터 전송) 광대역으로의 대역폭 확장이 자동으로 스위치되도록 디폴트가 설정될 수도 있다.

[0037] 460 에서, 광대역 접속이 이용 가능하지 않은 호출에서 사용되도록 (예를 들어, 미리결정된 임계에 대해) 저장된 충분한 대역폭 확장 데이터가 존재하는지가 결정된다. 그렇지 않으면, 프로세싱은 410 을 계속하여 더 많은 발음들을 모으고 더 많은 대역폭 확장 데이터를 생성한다. 그러나, 충분한 대역폭 확장 데이터가 생성 및 저장되었으면, 프로세싱은 정지하고 결과의 모델은 470 에서 협대역 음성 통신에서 사용되도록 준비된다.

[0038] 일 구현에 따라, 개인화된 대역폭 확장은 여러 경우들, 예컨대 (1) 광대역이 호출 전체에서 이용 가능하지 않은 경우; (2) 광대역이 호출의 일부에 대해 이용 가능하지만, 호출의 다른 부분에 대해 이용 가능하지 않고, 개인화된 대역폭 확장이 광대역이 이용 가능하지 않을 때 활성화되는 경우; 및 (3) 광대역이 호출의 일부 부분에 대해 이용 가능하거나 항상 이용 가능하고, 개인화된 대역폭 확장 모델이 아주 충분하고 캐리어 회사가 개인화된 대역폭 확장 모드로 스위치하기를 결정하더라도 개인화된 대역폭 확장이 이용 가능한 경우 중 어느 하나에서 사용될 수도 있다.

[0039] 도 5 는 개인화된 대역폭 확장을 위한 트레이닝 방법 (500) 의 다른 구현의 동작 흐름이다. 이 구현에서, 모델은 호출 동안 니어-엔드 스테이션 (110) 에서 (예를 들어, 사용자의 모바일 폰에서) 트레이닝된다. 니어-엔드 스테이션 (110) 은 네트워크 컨디션에 관계없이 광대역 스피치를 캡처할 수 있다.

[0040] 510 에서, 전화 호출 (즉, 온-라인 호출) 이 발생하고, 여기서 화자는 호출 중이고 화자의 스테이션에 말한다 (스피치 발음들을 제공). 515 에서, 화자의 음성 또는 스피치 발음들의 음향 컨디션이 트레이닝에 사용되기에 매우 충분한 품질인지가 결정된다. 전술된 바와 같이, SNR 기법은 음향 컨디션이 허용 가능한지 아닌지 여부를 결정하는데 사용될 수도 있다. 그렇지 않으면, 프로세싱은 510 으로 리턴하여 음향 컨디션이 허용 가능할 때까지 계속한다.

[0041] 515 에서 음향 컨디션이 허용 가능한 것으로 결정되면, 520 에서 예를 들어, 사용자 입력 (525) 을 통해 화자에 의해 화자가 식별된다. 530 에서, 식별된 화자가 (예를 들어, 화자의 스테이션과 연관된 스토리지에서 등록된 사용자들의 리스트에 화자의 식별을 비교함으로써) 등록된 사용자인지가 결정된다. 식별된 화자가 등록된 사용자인지를 결정하는 것은, 스테이션의 오너들 (예를 들어, 모바일 전화의 오너들) 로서 등록된 사용자들과는 다른 사람들에 대한 원하지 않는 트레이닝을 방지하는데 사용된다. 화자가 등록된 사용자가 아니면, 프로세싱은 510 으로 리턴한다. 그러나, 식별된 화자가 등록된 사용자라는 것이 결정되면, 협대역 및 광대역 피처들은 540 에서 스피치 발음들로부터 추출되고, 임의의 대역폭 확장 피처들이 이전에 생성 및 저장되었으면, 이것은 반복적 프로세스이기 때문에 화자에 속하는 대역폭 확장 피처들은 545 에서 스토리지 (예컨대, 로컬 데이터베이스) 로부터 취출될 수도 있다.

[0042] 550 에서, 피처 트레이닝은 540 으로부터 추출된 피처들 및 545 로부터 취출된 대역폭 확장 피처들을 사용하여 수행된다. 이 피처 트레이닝은 피처 공간을 업데이트하는 것 및 방법 (400) 에 대하여 전술된 것과 유사한 화자에 대한 협대역-광대역 맵핑을 적응시키는 것을 포함할 수도 있다. 555 에서, 대역폭 확장 피처 공간 및 화자에 대한 맵핑은 (예를 들어, 트레이닝 모델의 완료 전에 발생하게 될 맵핑 및/또는 피처들의 양의 미리결정된 임계에 대해) 트레이닝의 완료 퍼센티지로 (예를 들어, 로컬 데이터베이스에) 저장될 수도 있다.

[0043] 560 에서, 광대역 접속이 이용 가능하지 않은 호출에서 사용되도록 (예를 들어, 미리결정된 임계에 대해) 저장된 충분한 대역폭 확장 데이터가 존재하는지가 결정된다. 그렇지 않으면, 프로세싱은 510 을 계속하여 더 많은 발음들을 모으고 더 많은 대역폭 확장 데이터를 생성한다. 그러나, 충분한 대역폭 확장 데이터가 생성 및 저장되었으면, 프로세싱은 정지하고 결과의 모델은 협대역 음성 통신에서 사용되도록 준비된다.

[0044] 570 에서, 협대역 음성 통신에서의 사용을 위해 모델은 파-엔드 스테이션으로 전송될 수도 있다. 사용자-대-사용자 데이터 접속이 이용 가능하고 이전의 데이터로부터 델타가 존재할때는 언제든지, 이러한 데이터 송신은 백그라운드에서 이루어질 수 있다. 대안으로, 메일 또는 단문 메시지 푸시에서와 같이, 이용 가능한 사용자의 모델은, 모델의 오너가 그녀/그의 모델을 업데이트하기를 원할 때는 언제든지 그 전화 주소록에서 콘택트를 갖는 사용자에게 통지될 수 있다. 대안으로, 호출의 시작에서, 트레이닝된 모델이 이용 가능하고 최종 싱크업 (last sync-up) 이래로 새롭거나 변경되면, 이것은 송신될 수 있다.

[0045] 도 6 은 개인화된 대역폭 확장을 위한 트레이닝 방법 (600) 의 다른 구현의 동작 흐름이다. 이 구현에서, 모델은 이동국에서 (예를 들어, 니어-엔드 스테이션 (110) 이나 파-엔드 스테이션 (150) 중 어느 하나 또는 적합한 능력을 갖는 임의의 이동국에서) 사용자의 음성로 또는 이동국의 오너에 의해 선택된 타겟 음성로, (예를

들어, 전화 호출 동안이 아닌) 오프-라인으로 트레이닝된다 . 오프-라인 트레이닝은, 전화가 전화 호출 동안 사용되고 있지 않은 경우 사용자의 음성 또는 타겟 음성을 레코딩하는 것을 수반한다.

[0046] 610 에서, 스테이션이 오프-라인 (즉, 전화 호출이 발생하지 않음) 인 동안, 화자는 스테이션에 말하여 스피치 발음들을 화자의 스테이션에 제공한다. 615 에서, 화자의 음성 또는 스피치 발음들의 음향 컨디션이 트레이닝에 사용되기에 매우 충분한 품질인지가 결정된다. 그렇지 않으면, 프로세싱은 610 으로 리턴하여 음향 컨디션이 허용 가능할 때까지 계속한다.

[0047] 음향 컨디션이 615 에서 허용 가능하지 않은 것으로 결정되면, 사용자 입력 (630)(예를 들어, 스테이션과 연관된 미리결정된 전화번호부로부터 식별의 사용자 선택) 을 통해 화자에 의해 625 에서 화자가 식별된다.

[0048] 620 에서 협대역 및 광대역 피처들이 스피치 발음들로부터 추출되고, 임의의 대역폭 확장 피처들이 이전에 생성 및 저장되었다면, 이것은 반복적 프로세스이기 때문에 635 에서 식별된 화자에 속하는 대역폭 확장 피처들은 스토리지 (예컨대, 로컬 데이터베이스) 로부터 취출될 수도 있다.

[0049] 640 에서, 피처 트레이닝은 620 으로부터 추출된 피처들 및 635 로부터 취출된 대역폭 확장 피처들을 사용하여 수행된다. 이 피처 트레이닝은 피처 공간을 업데이트하는 것 및 화자에 대한 협대역-광대역 맵핑을 적응시키는 것을 포함할 수도 있다. 650 에서, 방법 (400) 에 대하여 전술된 것과 유사하게, 대역폭 확장 피처 공간 및 화자에 대한 맵핑은 (예를 들어, 트레이닝 모델의 완료 전에 발생하게 될 맵핑 및/또는 피처들의 양의 미리결정된 임계에 대해) 트레이닝의 완료 퍼센티지로 (예를 들어, 로컬 데이터베이스에) 저장될 수도 있다.

[0050] 660 에서, 광대역 접속이 이용 가능하지 않은 호출에서 사용되도록 (예를 들어, 미리결정된 임계에 대해) 저장된 충분한 대역폭 확장 데이터가 존재하는지가 결정된다. 그렇지 않으면, 프로세싱은 610 을 계속하여 더 많은 발음들을 모으고 더 많은 대역폭 확장 데이터를 생성한다. 그러나, 충분한 대역폭 확장 데이터가 생성 및 저장되었으면, 프로세싱은 정지하고 결과의 모델은 670 에서 협대역 음성 통신에서 사용되도록 준비된다.

[0051] 도 7 은 스펙트럼 인벨로프/여기-기반의 개인화된 대역폭 확장을 위해 트레이닝을 수행하는 방법 (700) 의 구현의 동작 흐름이다. 705 에서, 스피치 소스 (예를 들어, 스피치 발음들) 로부터의 스피치가 생성되어 스테이션에 제공된다. 710 에서 광대역 (WB) 피처 추출이 수행되고, 720 에서 협대역 (NB) 피처 추출이 수행된다. 715 에서 WB 라인 스펙트럼 주파수 (LSF) 및 NB LSF 가 WB LSF 코드북 업데이트에 제공되고, NB LSF 코드들로부터 WB LSF 코드들로 맵핑하는 맵핑 테이블을 구성한다. 예시의 맵핑 테이블 (740), 예컨대 715 에서 생성된 테이블은 도 8 에 도시된다. 도 8 에는 대응하는 WB LSF (747) 와 함께 NB LSF 인덱스 (745) 가 제공된다. WB 코드북은 (예를 들어, 일반 대중의 사람에 속하는) 일반화된 LSF들로 초기화된다. NB LSF 코드 i 에 대해, 현재 프레임의 WB LSF 에 기초하여 연관된 WB LSF 코드북 엔트리를 업데이트한다.

[0052] 도 7 로 돌아가면, 725 에서 NB LSF 및 NB 여기와 함께 WB 여기가 WB 여기 코드북 업데이트에 제공되고, NB 로부터 WB 여기들로의 코드북 맵핑 테이블을 구성한다. 730 에서, (예를 들어, 충분한 트레이닝 데이터가 획득되었는지 여부에 기초하여) 트레이닝을 정지할지 안할지 여부가 결정된다. 그렇다면, 735 에서 트레이닝은 정지되고; 그렇지 않으면 705 에서 프로세싱이 계속되어 추가의 스피치 발음들이 트레이닝에 사용된다. 예시의 맵핑 테이블 (750), 예컨대 725 에서 생성된 테이블은 도 9 에 도시된다. 대응하는 여기 인덱스 (757) 와 함께, LSF 인덱스 (755) 가 제공된다. 코드북은 일반화된 여기들 (즉, 일반 대중으로부터 도출된 여기들) 로 초기화된다. NB LSF 코드 i 및 NB 여기 코드 j 에 대해, 연관된 WB 여기 코드북 엔트리를 업데이트한다.

[0053] 도 10 은 단지 협대역 소스만이 이용 가능한 경우 스펙트럼 인벨로프/여기-기반의 개인화된 대역폭 확장을 수행하는 방법 (760) 의 구현의 동작 흐름이다. 765 에서, 협대역 스피치가 스테이션, 예컨대 파-엔드 스테이션 (150) 에 제공된다. 770 에서, 협대역 피처들이 추출되어, NB 여기 코드들 및 NB LSF 코드들을 초래한다.

[0054] 775 에서, NB 여기 코드들 및 NB LSF 코드들을 사용하여 WB 여기들로의 맵핑이 수행된다. 780 에서, NB LSF 로부터 WB LSF 로 맵핑이 또한 수행된다. 맵핑들을 사용하여, WB 스피치는 785 에서 합성되고, 790 에서 출력된다.

[0055] NB 와 WB 간에 고유하지 않은 맵핑 문제가 존재할 수도 있다. 이 이슈를 해결하기 위해, N-그램으로 지칭된 스피치 인식에 대한 언어 모델과 유사한 기법이 사용될 수도 있다. 일 구현에서, 맵핑을 구성하기 위해 과거 (또는 지연이 허용된다면 가능한 미래) 의 프레임들을 보자. 예를 들어, 2-그램이 구성되면, 맵핑은

NB_{i}-WB{i} 대신에 NB_{i-1},NB{i}-WB{i} 가 된다.

[0056] 스펙트럼 인벨로프 (예를 들어, LPC (linear prediction coefficients), LSF, MFCC (Mel-frequency cepstrum coefficients)) 에 대한 트레이닝 페이즈의 구현에서, 광대역 신호가 이용 가능한 경우, 예를 들어 코드북 맵핑 또는 HMM/GMM (Hidden Markov model/Gaussian mixture model) 을 구성하기 위해 대응하는 광대역 신호의 SE 및 협대역 스피치에 대한 스펙트럼 인벨로프 (SE) 코드북을 찾는다. 더 낮은 대역 SE들을 유지하고/하거나 너무 가까운 SE들을 방지하기 위한 임의의 방법이 또한 적용될 수도 있다.

[0057] 일 구현에서, 코드북 맵핑에 있어서, WE SE 테이블은 일반화된 대중을 위해 BWE 에 대한 테이블로 초기화된다. 테이블에서의 각각의 엔트리는, 대응하는 NB SE 가 관측되는 경우 대체/적응된다. 업데이트된 엔트리들의 수 및/또는 엔트리 당 업데이트들의 수가 유지 및 분석되어, 충분한 트레이닝 데이터가 획득되었는지를 결정할 수도 있다.

[0058] 스펙트럼 인벨로프에 대한 개인화된 대역폭 확장 페이즈의 구현에서, 소정의 NB SE 는 코드북 맵핑 (테이블 검색 또는 필터링) 에 의해 또는 확률 모델 (GMM/HMM) 로부터 대응하는 WB SE 를 결정한다.

[0059] 여기 구현에 대한 트레이닝 페이즈에 관하여, NB-여기 코드북 (FCB; fixed codebook) 또는 FCB/ACB (적응적 코드북) 및 NB SE 가 주어지면, 본 시스템들 및 방법들은 NB SE 여기으로부터 WB 여기으로의 맵핑/확률 모델을 구성할 수도 있다. 예를 들어, N*M 엔트리들이 있는 WB 여기 테이블을 구성할 수도 있고, 여기서 N 은 NB 여기 코드북에서 엔트리들의 수이고 M 은 NB SE 코드북에서 엔트리들의 수이다. 이 테이블은 일반화된 대중에 대한 대역폭 확장에 대한 엔트리들로 초기화된다. 테이블에서의 각각의 엔트리는, 대응하는 NB SE/여기 조합이 관측되는 경우 대체/적응된다. 업데이트된 엔트리들의 수 및/또는 엔트리 당 업데이트들의 수가 유지 및 분석되어, 충분한 트레이닝 데이터가 획득되었는지를 결정할 수도 있다.

[0060] 대역폭 확장 페이즈 동안, 신뢰성 측정 (기준) 은 개인화된 대역폭 확장을 사용할지 여부를 결정하는데 사용될 수도 있고, 시간에서 트레이닝 데이터 길이 및/또는 테이블에서의 업데이트된 엔트리들의 수에 기초할 수도 있다. 미리결정된 양 또는 임계 위이면, 예를 들어 개인화된 대역폭 확장이 사용될 수도 있다.

[0061] 더욱이, 일부 구현들에서 신뢰성 측정의 다수의 레벨들이 사용될 수 있다. 예를 들어, (1) 신뢰성 측정이 매우 높으면, 개인화된 대역폭 확장은 네트워크 리소스가 (네트워크 용량을 감소시키기 위해) WB 에 대해 충분한 경우에도 사용될 수도 있고; (2) 신뢰성 측정이 중간 범위에 있으면, 개인화된 대역폭 확장은 네트워크가 WB 를 지원할 수 없는 경우에도 사용될 수도 있으며; (3) 신뢰성 측정이 낮으면, 협대역 통신 또는 일반화된 대역폭 확장이 사용될 수도 있다.

[0062] 다르게 명시되지 않으면, 특정 피처를 갖는 장치의 동작의 임의의 개시는 또한, 유사한 피처 (및 그 반대) 를 갖는 방법을 개시하도록 명확히 의도되고, 특정 구성에 따른 장치의 동작의 임의의 개시는 유사한 구성 (및 그 반대) 에 따른 방법을 개시하도록 명확히 의도된다.

[0063] 본원에서 사용된 바와 같이, 용어 "결정하기" (및 그 문법상 변형들) 은 예시적인 넓은 의미에서 사용된다. 용어 "결정하기" 는 매우 다양한 액션들을 망라하므로, "결정하기" 는 산출하기, 계산하기, 처리하기, 도출하기, 조사하기, 검색하기 (예를 들어, 테이블, 데이터베이스, 또는 다른 데이터 구조 내 검색하기), 확인하기 등을 포함할 수 있다. 또한, "결정하기" 는 수신하기 (예를 들어, 정보 수신하기), 액세스하기 (예를 들어, 메모리 내의 데이터에 액세스하기) 등을 포함할 수도 있다. 또한, "결정하기" 는 해결하기, 선택하기, 고르기, 확립하기 등을 포함할 수 있다.

[0064] 용어 "신호 프로세싱" (및 그 문법상 변형들) 은 신호들의 프로세싱 및 해석을 지칭할 수도 있다. 관심있는 신호들은 사운드, 이미지들, 및 많은 다른 것들을 포함할 수도 있다. 이러한 신호들의 프로세싱은 저장 및 재구성, 잡음으로부터 정보의 분리, 압축, 및 피처 추출을 포함할 수도 있다. 용어 "디지털 신호 프로세싱" 은 디지털 표현으로 신호들의 연구 및 이들 신호들의 프로세싱 방법들을 지칭할 수도 있다. 디지털 신호 프로세싱은 이동국들, 비-이동국들, 및 인터넷과 같은 많은 통신 기술들의 엘리먼트이다. 디지털 신호 프로세싱에 대해 이용되는 알고리즘들은 특수 컴퓨터들을 사용하여 수행될 수도 있고, 이 컴퓨터들은 디지털 신호 프로세서들 (때때로, DSP들로 약칭됨) 로 지칭된 특수 마이크로프로세서들을 사용할 수도 있다.

[0065] 본원에서 개시된 실시형태들과 연계하여 설명된 방법, 프로세스 또는 알고리즘의 단계들은 하드웨어에서, 프로세서에 의해 실행되는 소프트웨어 모듈에서, 또는 이들 둘의 조합에서 직접적으로 구현될 수도 있다. 방법 또는 프로세스에서 각종 단계들 또는 액트들은 도시된 순서로 수행될 수도 있고, 또는 다른 순서로 수행될 수도 있다. 부가적으로, 하나 이상의 프로세스 또는 방법 단계들이 생략될 수도 있고, 또는 하나 이상의 프로세

스 또는 방법 단계들은 방법들 및 프로세스들에 추가될 수도 있다. 추가의 단계, 블록, 또는 액션이 방법들 및 프로세스들의 시작, 끝, 또는 중간의 기존 엘리먼트들에 추가될 수도 있다.

[0066] 도 11 은 무선 통신 시스템에서의 예시적인 이동국 (800) 의 설계의 블록도를 도시한다. 이동국 (800) 은 스마트 폰, 셀룰러 폰, 단말기, 핸드셋, PDA, 무선 모뎀, 무선 전화기 등일 수도 있다. 무선 통신 시스템은 코드 분할 다중 액세스 (CDMA) 시스템, GSM 시스템 등일 수도 있다.

[0067] 이동국 (800) 은 수신 경로 및 송신 경로를 통해 양방향 통신을 제공할 수 있다. 수신 경로에서, 기지국들에 의해 송신된 신호들이 안테나 (812) 에 의해 수신되어, 수신기 (RCVR) (814) 로 제공된다. 수신기 (814) 는 수신된 신호를 컨디셔닝 및 디지털화하여, 추가적인 프로세싱을 위해 디지털 섹션 (820) 으로 샘플들을 제공한다. 송신 경로에서, 송신기 (TMTR) (816) 가 디지털 섹션 (820) 으로부터 송신될 데이터를 수신하여, 그 데이터를 프로세싱하고 컨디셔닝하며, 변조된 신호를 생성하고, 이 변조된 신호는 안테나 (812) 를 통해 기지국들로 송신된다. 수신기 (814) 및 송신기 (816) 는 CDMA, GSM 등을 지원할 수도 있는 트랜시버의 일부분일 수도 있다.

[0068] 디지털 섹션 (820) 은, 예를 들어, 모뎀 프로세서 (822), RISC/DSP (reduced instruction set computer/digital signal processor)(824), 제어기/프로세서 (826), 내부 메모리 (828), 일반화된 오디오 인코더 (832), 일반화된 오디오 디코더 (834), 그래픽/디스플레이 프로세서 (836), 및 외부 버스 인터페이스 (external bus interface; EBI) (838) 와 같은 다양한 프로세싱, 인터페이스, 및 메모리 유닛들을 포함한다. 모뎀 프로세서 (822) 는 데이터 송신 및 수신을 위한 프로세싱, 예를 들어 인코딩, 변조, 복조, 및 디코딩을 수행할 수도 있다. RISC/DSP (824) 는 이동국 (800) 에 대한 일반적인 프로세싱 및 특수한 프로세싱을 수행할 수도 있다. 제어기/프로세서 (826) 는 디지털 섹션 (820) 내의 다양한 프로세싱 및 인터페이스 유닛들의 동작을 총괄할 수도 있다. 내부 메모리 (828) 는 디지털 섹션 (820) 내의 다양한 유닛들에 대한 데이터 및/또는 명령들을 저장할 수도 있다.

[0069] 일반화된 오디오 인코더 (832) 는 오디오 소스 (842), 마이크로폰 (843) 등으로부터의 입력 신호들에 대한 인코딩을 수행할 수도 있다. 일반화된 오디오 디코더 (834) 는 코딩된 오디오 데이터에 대한 디코딩을 수행할 수도 있고, 출력된 신호들을 스피커/헤드셋 (844) 에 제공할 수도 있다. 그래픽스/디스플레이 프로세서 (836) 는 디스플레이 유닛 (846) 에 제시될 수도 있는 그래픽들, 비디오들, 이미지들, 및 텍스트들에 대한 프로세싱을 수행할 수도 있다. EBI (838) 는 디지털 섹션 (820) 과 메인 메모리 (848) 간의 데이터의 전송을 용이하게 할 수도 있다.

[0070] 디지털 섹션 (820) 은 하나 이상의 프로세서들, DSP 들, 마이크로프로세서들, RISC 들 등으로 구현될 수도 있다. 디지털 섹션 (820) 은 또한, 하나 이상의 주문형 집적 회로 (application specific integrated circuit; ASIC) 들 및/또는 일부 다른 타입의 집적 회로 (IC) 들 상에 제작될 수도 있다.

[0071] 도 12 는 예시의 구현들 및 양태들이 구현될 수도 있는 예시적인 컴퓨팅 환경을 도시한다. 컴퓨팅 시스템 환경은 단지 적합한 컴퓨팅 환경의 일 예이며, 사용 또는 기능성의 범위에 관한 임의의 제한을 제안하도록 의도되지 않는다.

[0072] 컴퓨터에 의해 실행되는, 컴퓨터 실행가능 명령들, 예컨대 프로그램 모듈들이 사용될 수도 있다. 일반적으로, 프로그램 모듈들은 특정 태스크들을 수행하거나 특정의 추상적 데이터 타입들을 구현하는 루틴들, 프로그램들, 오브젝트들, 컴포넌트들, 데이터 구조들 등을 포함한다. 분산형 컴퓨팅 환경들이 사용될 수도 있고, 여기서 태스크들은 통신 네트워크 또는 다른 데이터 송신 매체를 통해 링크되는 원격 프로세싱 디바이스들에 의해 수행된다. 분산형 컴퓨팅 환경에서, 프로그램 모듈들 및 다른 데이터는 메모리 저장 디바이스들을 포함하는 로컬 및 원격 컴퓨터 저장 매체 양자 모두에 위치될 수도 있다.

[0073] 도 12 를 참조하면, 본원에 설명된 양태들을 구현하기 위한 예시적인 시스템은 컴퓨팅 디바이스, 예컨대 컴퓨팅 디바이스 (900) 를 포함한다. 그 가장 기본적인 구성에서, 컴퓨팅 디바이스 (900) 는 통상적으로, 적어도 하나의 프로세싱 유닛 (902) 및 메모리 (904) 를 포함한다. 컴퓨팅 디바이스의 정확한 구성 및 타입에 따라, 메모리 (904) 는 휘발성 (예컨대, 랜덤 액세스 메모리 (RAM)), 비휘발성 (예컨대, 판독 전용 메모리 (ROM), 플래시 메모리 등) 또는 이 둘의 일부 조합일 수도 있다. 이 가장 기본적인 구성은 점선 (906) 에 의해 도 12 에 예시된다.

[0074] 컴퓨팅 디바이스 (900) 는 추가적인 피처들 및/또는 기능을 가질 수도 있다. 예를 들어, 컴퓨팅 디바이스 (900) 는 자기 또는 광학 디스크들 또는 테이프를 포함하지만 이에 한정되지는 않는 추가의 스토리지 (착탈형

및/또는 비-착탈형) 를 포함할 수도 있다. 이러한 추가의 스토리지는 착탈형 스토리지 (908) 및 비-착탈형 스토리지 (910) 에 의해 도 12 에 예시된다.

[0075] 컴퓨팅 디바이스 (900) 는 통상적으로, 다양한 컴퓨터 판독가능 매체를 포함한다. 컴퓨터 판독가능 매체는 디바이스 (900) 에 의해 액세스될 수 있고 휘발성 및 비-휘발성 매체, 및 착탈형 및 비-착탈형 매체 양자 모두를 포함하는 임의의 이용 가능한 매체일 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령들, 데이터 구조들, 프로그램 모듈들, 또는 다른 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술에서 구현된 휘발성 및 비휘발성, 착탈형 및 비착탈형 매체를 포함할 수도 있다. 메모리 (904), 착탈형 스토리지 (908), 및 비착탈형 스토리지 (910) 는 컴퓨터 저장 매체의 모든 예들이다. 컴퓨터 저장 매체는, RAM, ROM, EEPROM, 플래시 메모리 또는 다른 메모리 기술, CD-ROM, 디지털 다기능 디스크들 (DVD) 또는 다른 광학 스토리지, 자기 카세트들, 자기 테이프, 자기 디스크 스토리 또는 다른 자기 스토리지 디바이스들, 또는 원하는 정보를 저장하기 위해 사용될 수도 있고 컴퓨팅 디바이스 (900) 에 의해 액세스될 수도 있는 임의의 다른 매체를 포함하지만, 이에 한정되지는 않는다. 임의의 이러한 컴퓨터 저장 매체는 컴퓨팅 디바이스 (900) 의 일부일 수도 있다.

[0076] 컴퓨팅 디바이스 (900) 는, 디바이스가 다른 디바이스들과 통신하는 것을 허용하는 통신 접속(들)(912) 을 포함할 수도 있다. 컴퓨팅 디바이스 (900) 는 또한, 입력 디바이스(들)(914), 예컨대 키보드, 마우스, 펜, 음성 입력 디바이스, 터치 입력 디바이스 등을 가질 수도 있다. 출력 디바이스(들)(916), 예컨대 디스플레이, 스피커들, 프린터 등이 또한, 포함될 수도 있다. 이들 디바이스들 모두는 이 기술 분야에 잘 알려져 있고 여기서 길게 논의될 필요가 없다.

[0077] 일반적으로, 본원에 설명된 임의의 디바이스는, 무선 또는 유선 전화기, 셀룰러 전화기, 랩톱 컴퓨터, 무선 멀티미디어 디바이스, 무선 통신 PC 카드, PDA, 외부 모뎀 또는 내부 모뎀, 무선 또는 유선 채널을 통해 통신하는 디바이스 등과 같은 다양한 타입의 디바이스들을 나타낼 수도 있다. 디바이스는, 액세스 단말기 (AT), 액세스 유닛, 가입자 유닛, 이동국, 모바일 디바이스, 모바일 유닛, 모바일 전화기, 모바일, 원격국, 원격 단말기, 원격 유닛, 사용자 디바이스, 사용자 장비, 핸드헬드 디바이스, 비-이동국, 비-이동 디바이스, 엔드포인트 등과 같은 다양한 이름들을 가질 수도 있다. 본원에 설명된 임의의 디바이스는 명령들 및 데이터를 저장하기 위한 메모리, 뿐만 아니라 하드웨어, 소프트웨어, 펌웨어, 또는 그 조합들을 가질 수도 있다.

[0078] 본원에 설명된 기법들은 다양한 수단으로 구현될 수도 있다. 예를 들어, 이러한 기법들은 하드웨어, 펌웨어, 소프트웨어, 또는 그 조합으로 구현될 수도 있다. 당업자는 또한, 본원에 개시된 예시적인 실시형태들과 연계하여 설명된 다양한 예시적인 논리 블록들, 모듈들, 회로들, 및 알고리즘 단계들이 전자 하드웨어, 컴퓨터 소프트웨어, 또는 양자 모두의 조합으로서 구현될 수도 있음을 이해할 것이다. 하드웨어와 소프트웨어의 이러한 상호교환성을 명확하게 예시하기 위해, 다양한 예시적인 콤포넌트들, 블록들, 모듈들, 회로들, 및 단계들이 그들의 기능성에 대해 전반적으로 전술되었다. 그러한 기능이 하드웨어 또는 소프트웨어로 구현되는지 여부는 특정 애플리케이션 및 전체 시스템에 부과되는 설계 제약들에 따라 달라진다. 당업자는, 상기 상술한 기능성을 각각의 특정 애플리케이션에 대해 다양한 방식으로 구현할 수도 있지만, 이러한 구현 결정은 본 발명의 범위를 벗어나게 하는 것으로 해석되어서는 안 된다.

[0079] 하드웨어 구현에서, 기법들을 수행하는데 사용된 프로세싱 유닛들은, 하나 이상의 ASIC 들, DSP 들, 디지털 신호 프로세싱 디바이스 (digital signal processing device; DSPD) 들, 프로그램가능 논리 디바이스 (rogrammable logic device; PLD) 들, 필드 프로그램가능 게이트 어레이 (field programmable gate array; FPGA), 프로세서들, 제어기들, 마이크로 제어기들, 마이크로프로세서들, 전자 디바이스들, 본원에 설명된 기능들을 수행하도록 설계된 다른 전자 유닛들, 컴퓨터, 또는 그들의 조합 내에서 구현될 수도 있다.

[0080] 따라서, 본원의 개시물과 관련하여 설명된 다양한 예시적인 논리 블록들, 모듈들, 및 회로들은 범용 프로세서, DSP, ASIC, FPGA 나 다른 프로그램 가능 논리 디바이스, 이산 게이트나 트랜지스터 로직, 이산 하드웨어 컴포넌트들, 또는 본원에 설명된 기능들을 수행하도록 설계된 것들의 임의의 조합으로 구현되거나 수행될 수도 있다. 범용 프로세서는 마이크로프로세서일 수도 있지만, 대안으로 프로세서는 임의의 프로세서, 제어기, 마이크로제어기, 또는 상태 머신일 수도 있다. 프로세서는 또한, 컴퓨팅 디바이스들의 조합, 예를 들어 DSP와 마이크로프로세서의 조합, 복수의 마이크로프로세서들, DSP 코어와 연계한 하나 이상의 마이크로프로세서들, 또는 임의의 다른 그러한 구성으로 구현될 수도 있다.

[0081] 펌웨어 및/또는 소프트웨어 구현에 있어서, 랜덤 액세스 RAM, ROM, 비-휘발성 RAM, 프로그래머블 ROM, EEPROM, 플래시 메모리, 콤팩트 디스크 (CD), 자기 또는 광학 데이터 저장 디바이스 등과 같은 컴퓨터 판독가능 매체 상에서 명령들로서 구현될 수도 있다. 명령들은 하나 이상의 프로세서들에 의해 실행가능할 수도 있고, 프로

세서(들)로 하여금 본원에 설명된 기능의 특정 양태들을 수행하도록 할 수도 있다.

[0082] 소프트웨어로 구현되는 경우, 기능들은 하나 이상의 명령들 또는 코드로서 타입의 컴퓨터 판독 가능한 매체 상에 저장되거나 또는 전송될 수도 있다. 컴퓨터 판독가능 매체들은 한 장소에서 다른 장소로 컴퓨터 프로그램의 전송을 가능하게 하는 임의의 매체를 포함하여 컴퓨터 저장 매체들 및 통신 매체들 양자를 포함한다. 저장 매체는 범용 컴퓨터 또는 특수 목적의 컴퓨터에 의해 액세스될 수 있는 임의의 이용가능한 매체들일 수도 있다. 비제한적인 예로서, 이러한 컴퓨터 판독가능 매체는 RAM, ROM, EEPROM, CD-ROM 또는 다른 광학 디스크 저장 디바이스, 자기 디스크 저장 디바이스 또는 다른 자기 저장 디바이스, 플래시 메모리, 또는 원하는 프로그램 코드를 명령들 또는 데이터 구조들의 형태로 저장하는데 사용될 수 있으며 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수 있다. 또한, 임의의 접속이 컴퓨터 판독가능 매체라고 적절히 칭해진다. 예를 들어, 소프트웨어가 동축 케이블, 광섬유 케이블, 연선, 디지털 가입자 회선 (DSL), 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들을 사용하여 웹사이트, 서버, 또는 다른 원격 소스로부터 명령들이 송신되면, 동축 케이블, 광섬유 케이블, 연선, DSL, 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들은 매체의 정의 내에 포함된다. 본원에서 사용된 디스크 (disk) 와 디스크 (disc) 는, CD, 레이저 디스크, 광학 디스크, 디지털 다기능 디스크 (DVD), 플로피 디스크, 및 블루레이 디스크를 포함하며, 여기서 디스크 (disk) 들은 통상 자기적으로 데이터를 재생하는 반면, 디스크 (disc) 들은 레이저들을 이용하여 광학적으로 데이터를 재생한다. 상기의 조합들이 또한, 컴퓨터 판독가능 매체들의 범위 내에 포함되어야 한다.

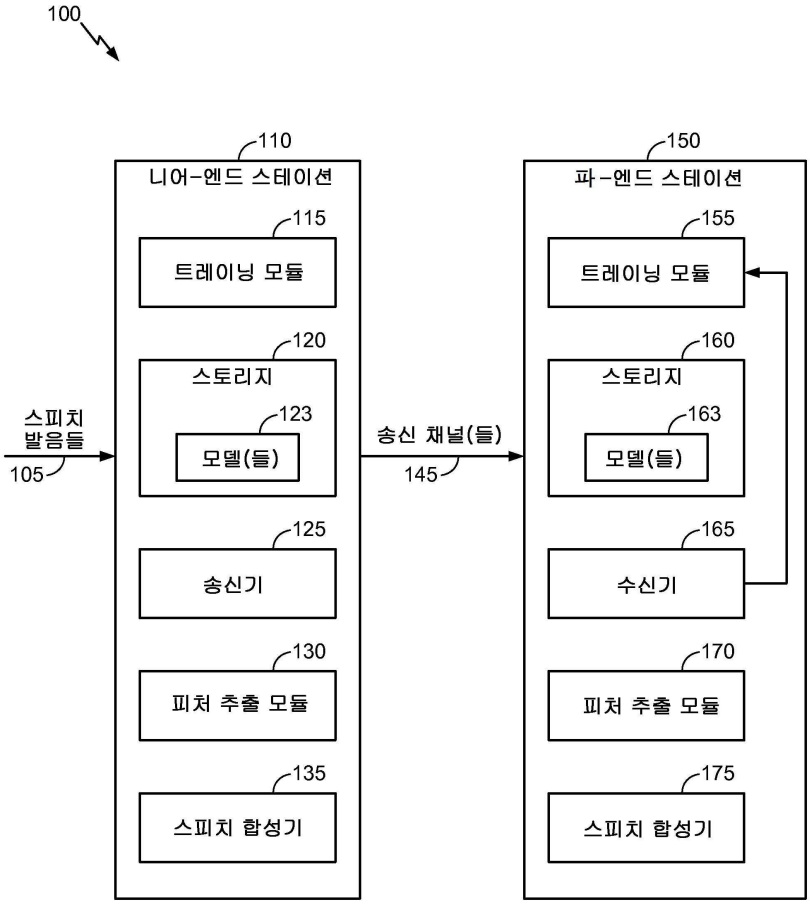
[0083] 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터들, 하드 디스크, 착탈형 디스크, CD-ROM, 또는 공지된 임의의 다른 형태의 저장 매체 내에 상주할 수도 있다. 예시적인 저장 매체는 프로세서에 커플링되어, 프로세가 저장 매체로부터 정보를 판독하거나 저장 매체에 정보를 기록할 수 있다. 대안으로 저장 매체는 프로세서에 통합될 수도 있다. 프로세서와 저장 매체는 ASIC 내에 있을 수도 있다. ASIC 는 사용자 단말기 내에 있을 수도 있다. 대안으로 프로세서와 저장 매체는 사용자 단말기에서 개별 컴포넌트들로 있을 수도 있다.

[0084] 비록 예시적인 실시형태들이 하나 이상의 독립형 컴퓨터 시스템의 맥락에서 현재 개시된 주제의 양태들을 이용하는 것을 언급할 수도 있으나, 본 주제는 그렇게 제한되지 않고, 오히려 네트워크나 분산된 컴퓨팅 환경과 같은 임의의 컴퓨팅 환경과 연계하여 구현될 수도 잇다. 또한, 현재 개시된 주제의 양태들은 복수의 프로세싱 칩들이나 디바이스들에서 또는 그 전체에 걸쳐 구현될 수도 있고, 스토리지는 복수의 디바이스들에 걸쳐 유사하게 영향을 받을 수도 있다. 그러한 디바이스는, 예를 들어, PC 들, 네트워크 서버들, 및 핸드헬드 디바이스들을 포함할 수도 있다.

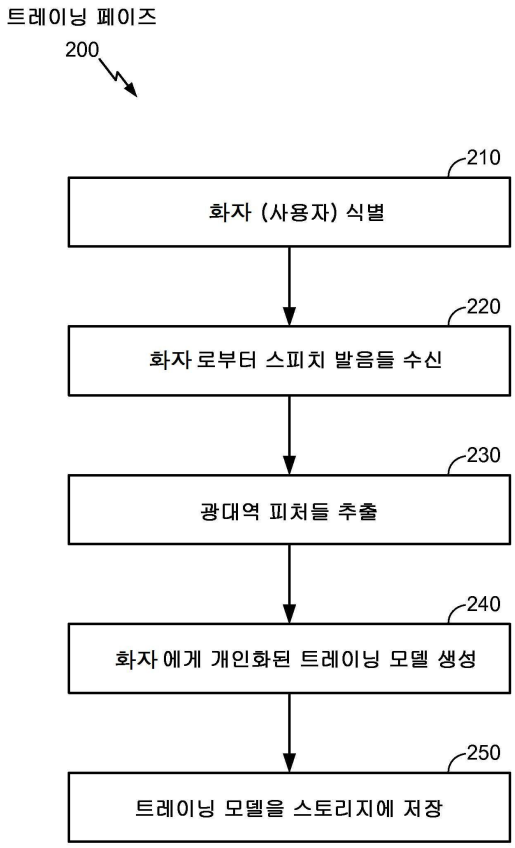
[0085] 본 주제가 구조적 특징들 및/또는 방법론적 작용들에 대해 언어 특정적으로 설명되었으나, 첨부된 청구항들에서 정의된 주제가 위에서 설명된 특정 특징들 또는 작용들로 반드시 제한되는 것은 아님이 이해될 것이다. 차라리, 위에서 설명된 특정 특징들 및 작용들은 청구항들을 구현하는 예시적인 형태로서 개시된다.

## 도면

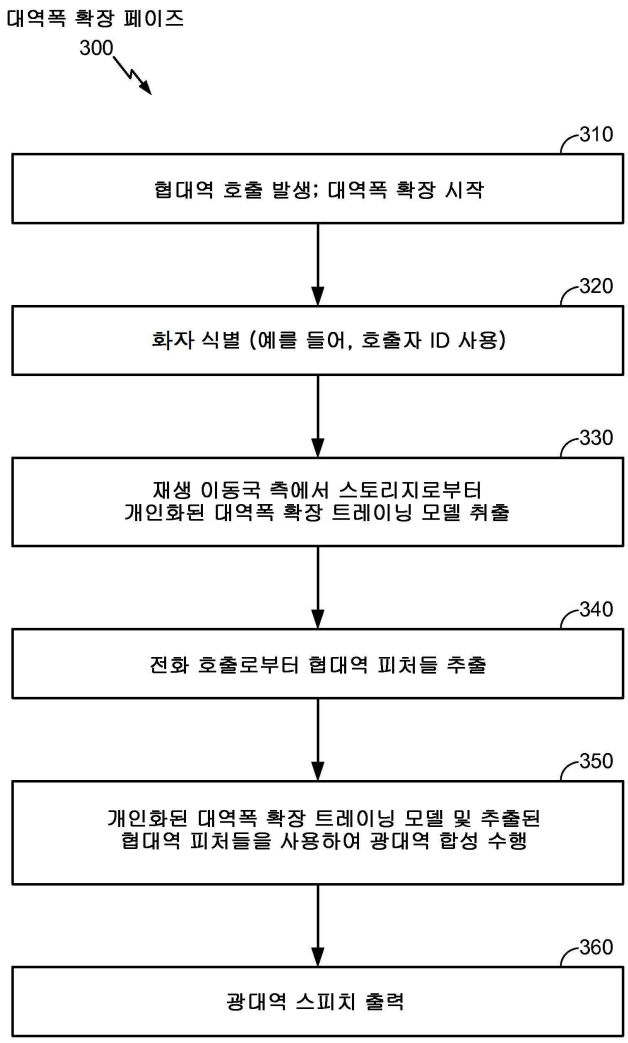
### 도면1


100
110
니어-엔드 스테이션
115
트레이닝 모듈
120
스토리지
123
모델(들)
125
송신기
130
피처 추출 모듈
135
스피치 합성기
스피치
발음들
105
송신 채널(들)
145
150
파-엔드 스테이션
155
트레이닝 모듈
160
스토리지
163
모델(들)
165
수신기
170
피처 추출 모듈
175
스피치 합성기

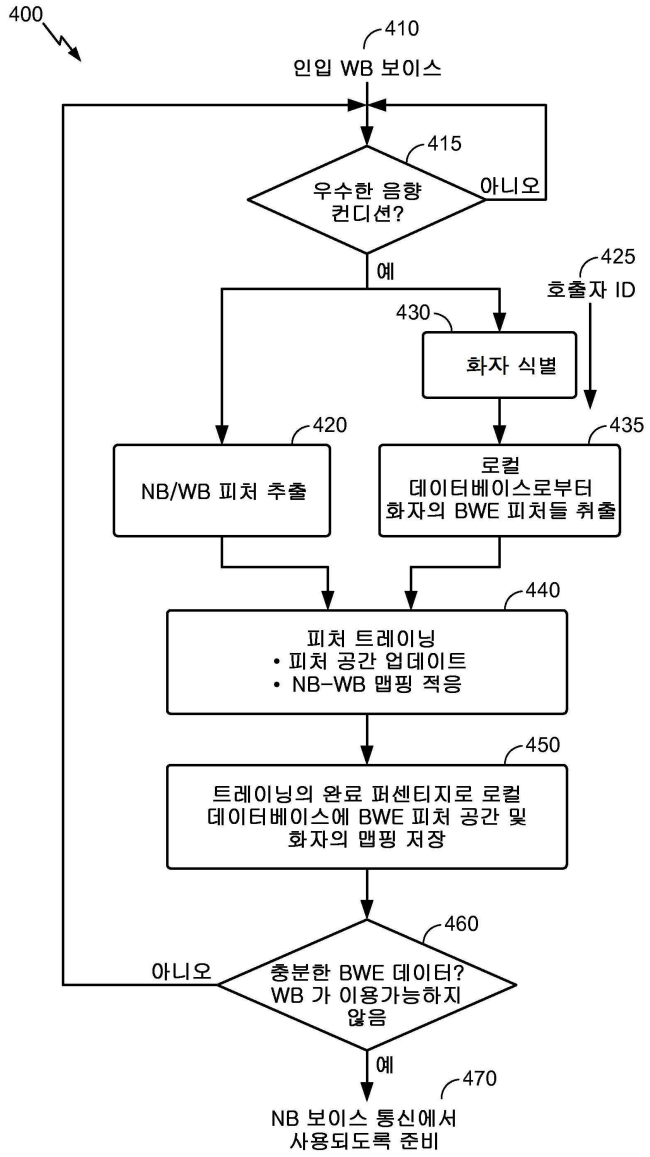
도면2


트레이닝 페이즈
200
210
화자 (사용자) 식별
220
화자 로부터 스피치 발음들 수신
230
광대역 피처들 추출
240
화자 에게 개인화된 트레이닝 모델 생성
250
트레이닝 모델을 스토리지에 저장

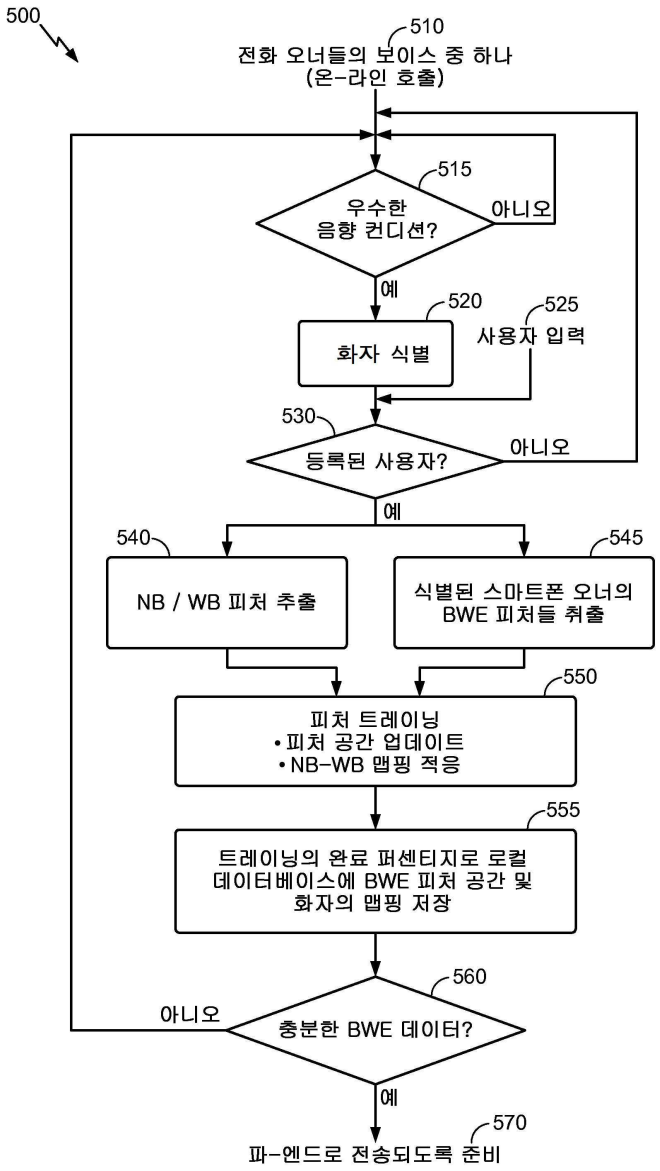
도면3


대역폭 확장 페이즈
300
310
협대역 호출 발생; 대역폭 확장 시작
320
화자 식별 (예를 들어, 호출자 ID 사용)
330
재생 이동국 측에서 스토리지로부터
개인화된 대역폭 확장 트레이닝 모델 취출
340
전화 호출로부터 협대역 피처들 추출
350
개인화된 대역폭 확장 트레이닝 모델 및 추출된
협대역 피처들을 사용하여 광대역 합성 수행
360
광대역 스피치 출력

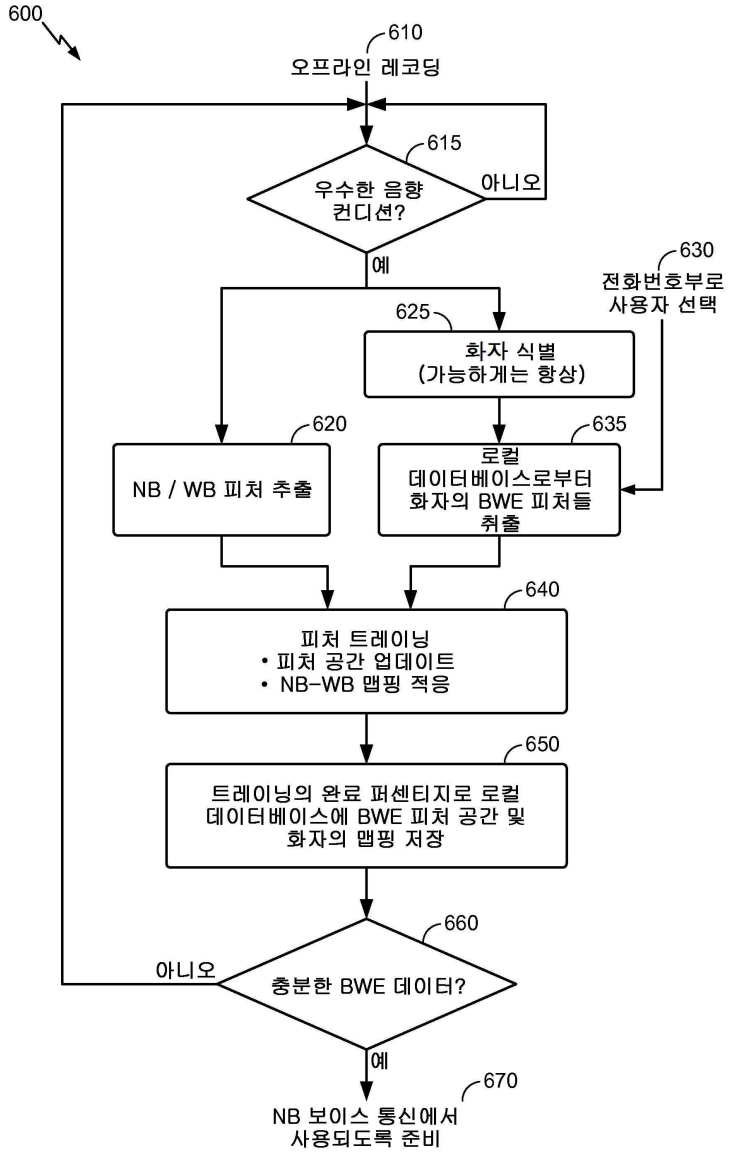
도면4


400
410
인입 WB 보이스
415
우수한 음향 컨디션?
아니오
예
425
호출자 ID
430
화자 식별
420
NB/WB 피처 추출
435
로컬 데이터베이스로부터 화자의 BWE 피처들 취출
440
피처 트레이닝
• 피처 공간 업데이트
• NB-WB 맵핑 적응
450
트레이닝의 완료 퍼센티지로 로컬 데이터베이스에 BWE 피처 공간 및 화자의 맵핑 저장
460
충분한 BWE 데이터? WB 가 이용가능하지 않음
아니오
예
470
NB 보이스 통신에서 사용되도록 준비

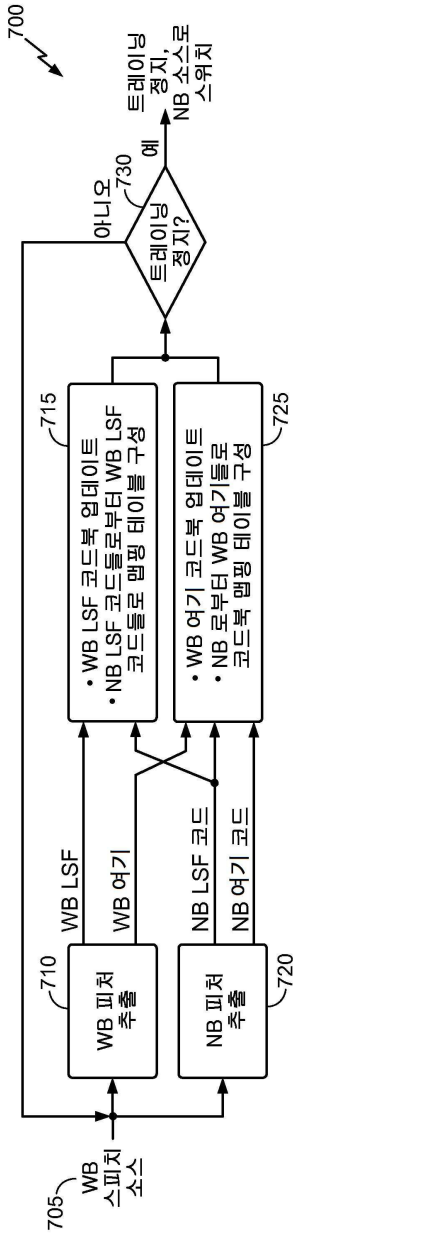
도면5


500
510
전화 오너들의 보이스 중 하나
(온-라인 호출)
515
우수한
음향 컨디션?
아니오
예
520
화자 식별
525
사용자 입력
530
등록된 사용자?
아니오
예
540
NB / WB 피처 추출
545
식별된 스마트폰 오너의
BWE 피처들 취출
550
피처 트레이닝
• 피처 공간 업데이트
• NB-WB 맵핑 적응
555
트레이닝의 완료 퍼센티지로 로컬
데이터베이스에 BWE 피처 공간 및
화자의 맵핑 저장
560
아니오
충분한 BWE 데이터?
예
570
파-엔드로 전송되도록 준비

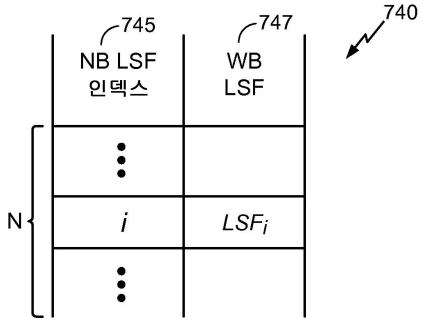
도면6


600
610
오프라인 레코딩
615
우수한 음향 컨디션?
아니오
예
630
전화번호부로 사용자 선택
625
화자 식별 (가능하게는 항상)
620
NB / WB 피처 추출
635
로컬 데이터베이스로부터 화자의 BWE 피처들 취출
640
피처 트레이닝
• 피처 공간 업데이트
• NB-WB 맵핑 적응
650
트레이닝의 완료 퍼센티지로 로컬 데이터베이스에 BWE 피처 공간 및 화자의 맵핑 저장
660
아니오
충분한 BWE 데이터?
예
670
NB 보이스 통신에서 사용되도록 준비

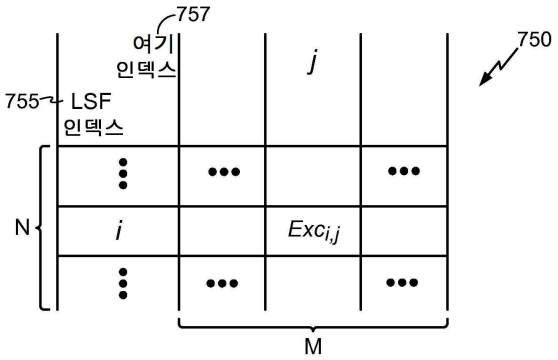
도면7


700
트레이닝 정지, NB 소스로 스위치
예
730
아니오
트레이닝 정지?
715
• WB LSF 코드북 업데이트
• NB LSF 코드들로부터 WB LSF 코드들로 맵핑 테이블 구성
725
• WB 여기 코드북 업데이트
• NB 로부터 WB 여기들로 코드북 맵핑 테이블 구성
WB LSF
WB 여기
NB LSF 코드
NB 여기 코드
710
WB 피처 추출
720
NB 피처 추출
705
WB 스피치 소스

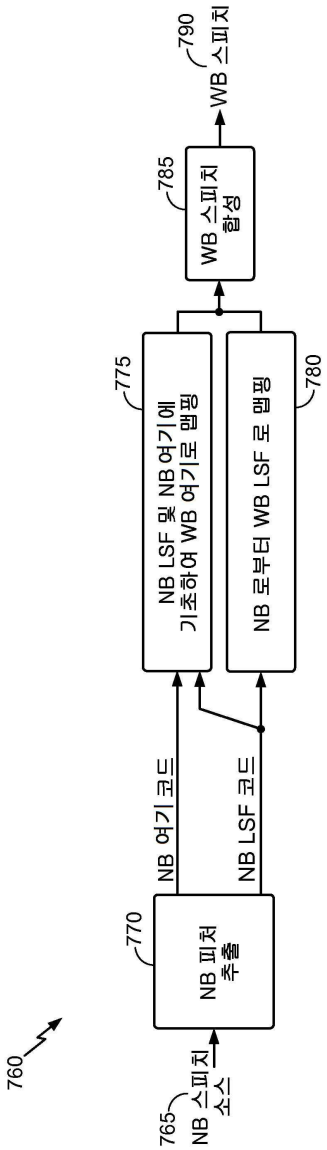
도면8


740
745
NB LSF 인덱스
747
WB LSF
N
i
$LSF_i$

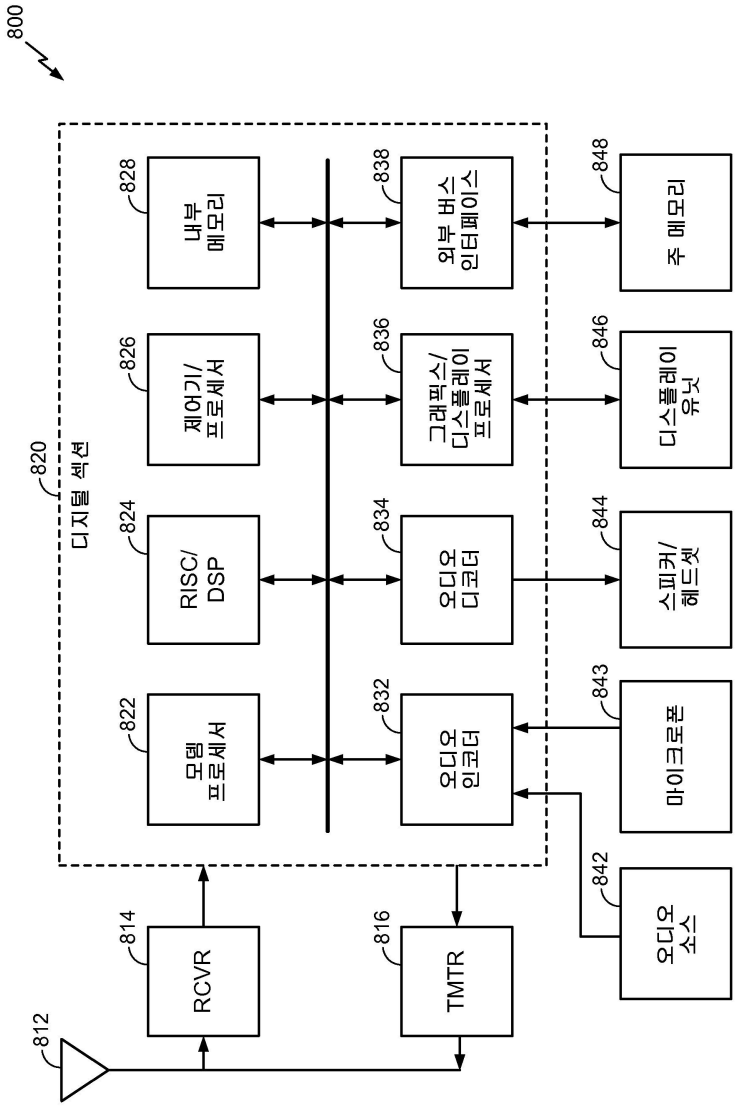
도면9


757
여기
인덱스
750
j
755
LSF
인덱스
N
i
$Exc_{i,j}$
M

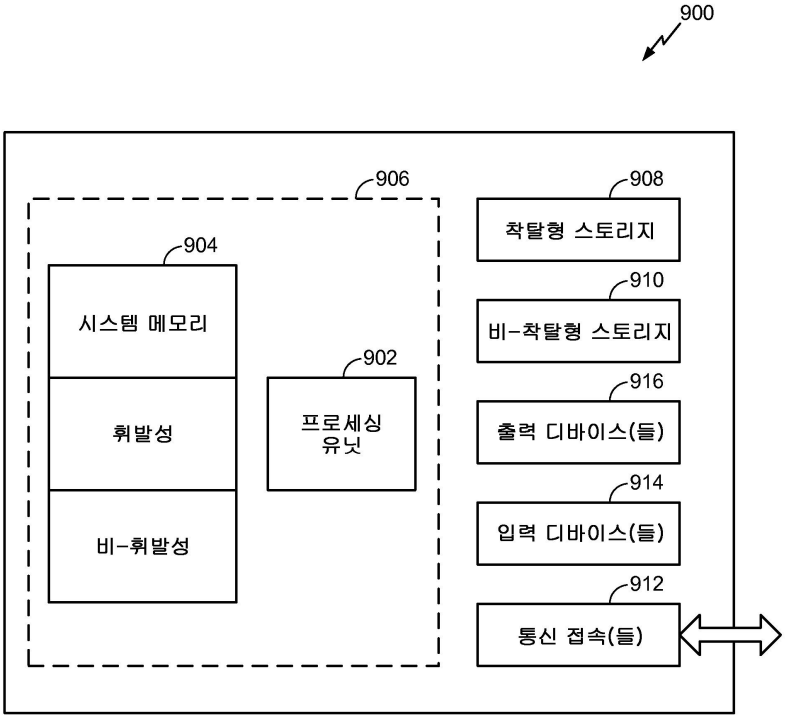
도면10


790
WB 스피치
785
WB 스피치
합성
775
NB LSF 및 NB 여기에
기초하여 WB 여기로 맵핑
780
NB 로부터 WB LSF 로 맵핑
NB 여기 코드
NB LSF 코드
770
NB 피치
추출
760
765
NB 스피치
소스

도면11


800
디지털 섹션
820
내부 메모리
828
제어기/ 프로세서
826
RISC/ DSP
824
모뎀 프로세서
822
외부 버스 인터페이스
838
그래픽스/ 디스플레이 프로세서
836
오디오 디코더
834
오디오 인코더
832
주 메모리
848
디스플레이 유닛
846
스피커/ 헤드셋
844
마이크로폰
843
오디오 소스
842
RCVR
814
TMTR
816
812

도면12


900
906
908
904
착탈형 스토리지
910
시스템 메모리
비-착탈형 스토리지
902
916
휘발성
프로세싱 유닛
출력 디바이스(들)
914
입력 디바이스(들)
비-휘발성
912
통신 접속(들)