



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0044849
(43) 공개일자 2017년04월26일

(51) 국제특허분류(Int. Cl.)
G10L 13/08 (2006.01) G10L 13/04 (2006.01)
(52) CPC특허분류
G10L 13/08 (2013.01)
G10L 13/04 (2013.01)
(21) 출원번호 10-2015-0144462
(22) 출원일자 2015년10월16일
심사청구일자 없음

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
(72) 발명자
성준식
서울특별시 관악구 남부순환로239길 47, 4층
(74) 대리인
윤동열

전체 청구항 수 : 총 20 항

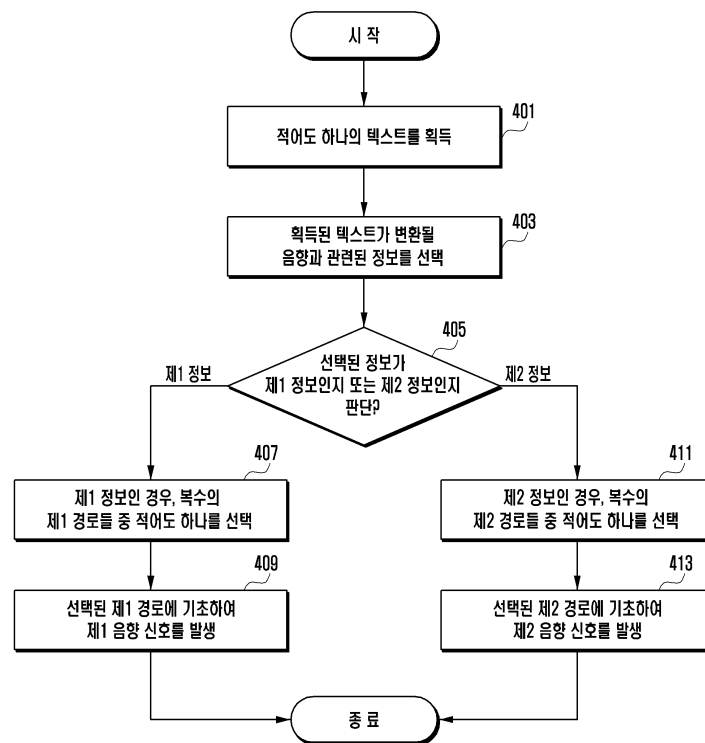
(54) 발명의 명칭 전자 장치 및 다국어/다화자의 공통 음향 데이터 셋을 활용하는 TTS 변환 방법

(57) 요약

본 발명의 다양한 실시예에 따른 전자 장치는 프로세서, 및 상기 프로세서와 전기적으로 연결된 메모리를 포함하고, 상기 메모리는, 공통 음향 데이터 셋을 저장하고, 상기 메모리는, 실행 시에, 상기 프로세서가, 적어도 하나의 텍스트를 획득하고, 상기 획득된 텍스트가 변환될 음향과 관련된 정보를 선택하고, 상기 선택된 정보가 제 1

(뒷면에 계속)

대표도 - 도4



정보인 경우, 복수의 제 1 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 1 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 일부를 로드하고, 상기 로드된 일부에 기초하여 제 1 음향 신호를 발생시키고, 상기 선택된 정보가 제 2 정보인 경우, 복수의 제 2 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 2 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 상기 일부 또는 다른 일부를 로드하고, 상기 로드된 일부 또는 다른 일부에 기초하여 제 2 음향 신호를 발생시키도록 하는 인스트럭션들을 저장하는 것을 특징으로 한다.

그 이외에 다양한 실시예가 가능하다.

명세서

청구범위

청구항 1

전자 장치에 있어서,
프로세서; 및
상기 프로세서와 전기적으로 연결된 메모리를 포함하고,
상기 메모리는, 공통 음향 데이터 셋을 저장하고,
상기 메모리는, 실행 시에, 상기 프로세서가,
적어도 하나의 텍스트를 획득하고,
상기 획득된 텍스트가 변환될 음향과 관련된 정보를 선택하고,
상기 선택된 정보가 제 1 정보인 경우, 복수의 제 1 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 1 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 일부를 로드하고, 상기 로드된 일부에 기초하여 제 1 음향 신호를 발생시키고,
상기 선택된 정보가 제 2 정보인 경우, 복수의 제 2 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 2 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 상기 일부 또는 다른 일부를 로드하고, 상기 로드된 일부 또는 다른 일부에 기초하여 제 2 음향 신호를 발생시키도록 하는 인스트럭션들을 저장하는 것을 특징으로 하는 장치.

청구항 2

제 1 항에 있어서,
상기 음향과 관련된 정보는,
상기 음향의 언어 정보 및/또는 화자 정보를 포함하는 것을 특징으로 하는 장치.

청구항 3

제 1 항에 있어서,
상기 인스트럭션들은, 상기 프로세서가,
사용자로부터 상기 적어도 하나의 텍스트를 획득하거나 외부 장치로부터 상기 적어도 하나의 텍스트를 포함하는 문자 메시지를 수신하도록 하는 것을 특징으로 하는 장치.

청구항 4

제 1 항에 있어서,
상기 인스트럭션들은, 상기 프로세서가,
상기 입력된 텍스트에 기초하여, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택하고,
상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여, 상기 제 1 음향 신호 또는 상기 제 2 음향 신호를 발생시키도록 하는 것을 특징으로 하는 장치.

청구항 5

제 4 항에 있어서,
상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부는,

상기 발생된 음향 신호의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응하는 것을 특징으로 하는 장치.

청구항 6

제 1 항에 있어서,

상기 복수의 제 1 경로들 또는 상기 복수의 제 2 경로들은,

상기 공통 음향 데이터 셋 중의 일부를 지시(indicate) 하는 것을 특징으로 하는 장치.

청구항 7

전자 장치에 있어서,

프로세서; 및

상기 프로세서와 전기적으로 연결된 메모리를 포함하고,

상기 메모리는, 실행 시에, 상기 프로세서가,

음향과 관련된 제 1 정보에 대응하는 제 1 음향 데이터 셋, 및 상기 음향과 관련된 제 2 정보에 대응하는 제 2 음향 데이터 셋을 획득하고,

상기 제 1 음향 데이터 셋의 적어도 일부, 및/또는 상기 제 2 음향 데이터 셋의 적어도 일부의 유사도를 판단하고,

상기 판단에 기초하여, 상기 제 1 음향 데이터 셋의 적어도 일부, 및/또는 상기 제 2 음향 데이터 셋의 적어도 일부와 관련된 공통 음향 데이터 셋을 생성하도록 하는 인스트럭션들을 저장하는 것을 특징으로 하는 장치.

청구항 8

제 7 항에 있어서,

상기 제 1 정보 또는 상기 제 2 정보는,

상기 음향의 언어 정보 및/또는 화자 정보를 포함하는 것을 특징으로 하는 장치.

청구항 9

제 7 항에 있어서,

상기 인스트럭션들은, 상기 프로세서가,

상기 판단에 기초하여, 상기 유사도가 선택된 쓰레쉬 홀드(threshold) 값 이상인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부, 및 상기 제 2 음향 데이터 셋의 적어도 일부 모두에 대응하는 제 1 파라미터를 결정하고,

상기 유사도가 상기 쓰레쉬 홀드 값 미만인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부에 대응하는 제 2 파라미터, 및 상기 제 2 음향 데이터 셋의 적어도 일부에 대응하는 제 3 파라미터를 결정하고,

상기 제 1 파라미터, 상기 제 2 파라미터, 또는 상기 제 3 파라미터에 기초하여 상기 공통 음향 데이터 셋을 생성하도록 하는 인스트럭션들을 저장하는 것을 특징으로 하는 장치.

청구항 10

제 9 항에 있어서,

상기 제 1 파라미터, 상기 제 2 파라미터, 또는 상기 제 3 파라미터는,

상기 음향의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응하는 것을 특징으로 하는 장치.

청구항 11

적어도 하나의 텍스트를 획득하는 동작;

상기 획득된 텍스트가 변환될 음향과 관련된 정보를 선택하는 동작;

상기 선택된 정보가 제 1 정보인 경우, 복수의 제 1 경로들 중 적어도 하나를 선택하는 동작, 상기 선택된 적어도 하나의 제 1 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 일부를 로드하는 동작, 상기 로드된 일부에 기초하여 제 1 음향 신호를 발생시키는 동작; 및

상기 선택된 정보가 제 2 정보인 경우, 복수의 제 2 경로들 중 적어도 하나를 선택하는 동작, 상기 선택된 적어도 하나의 제 2 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 상기 일부 또는 다른 일부를 로드하는 동작, 상기 로드된 일부 또는 다른 일부에 기초하여 제 2 음향 신호를 발생시키는 동작을 포함하는 전자 장치의 TTS 변환 방법.

청구항 12

제 11 항에 있어서,

상기 음향과 관련된 정보는,

상기 음향의 언어 정보 및/또는 화자 정보를 포함하는 전자 장치의 TTS 변환 방법.

청구항 13

제 11 항에 있어서,

상기 텍스트를 획득하는 동작은,

사용자로부터 상기 적어도 하나의 텍스트를 획득하거나 외부 장치로부터 상기 적어도 하나의 텍스트를 포함하는 문자 메시지를 수신하는 동작을 포함하는 전자 장치의 TTS 변환 방법.

청구항 14

제 11 항에 있어서,

상기 제 1 음향 신호 또는 상기 제2 음향 신호를 발생시키는 동작은,

상기 입력된 텍스트에 기초하여, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택하는 동작; 및

상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여, 상기 제 1 음향 신호 또는 상기 제 2 음향 신호를 발생시키는 동작을 포함하는 전자 장치의 TTS 변환 방법.

청구항 15

제 14 항에 있어서,

상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부는,

상기 발생된 음향 신호의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응하는 전자 장치의 TTS 변환 방법..

청구항 16

제 11 항에 있어서,

상기 복수의 제 1 경로들 또는 상기 복수의 제 2 경로들은,

상기 공통 음향 데이터 셋 중의 일부를 지시(indicate) 하는 전자 장치의 TTS 변환 방법.

청구항 17

적어도 하나의 텍스트가 변환될 음향과 관련된 제 1 정보에 대응하는 제 1 음향 데이터 셋, 및/또는 상기 음향과 관련된 제 2 정보에 대응하는 제 2 음향 데이터 셋을 획득하는 동작;

상기 제 1 음향 데이터 셋의 적어도 일부, 및/또는 상기 제 2 음향 데이터 셋의 적어도 일부의 유사도를 판단하는 동작; 및

상기 판단에 기초하여, 상기 제 1 음향 데이터 셋의 적어도 일부, 및/또는 상기 제 2 음향 데이터 셋의 적어도

일부와 관련된 공통 음향 데이터 셋을 생성하는 동작을 포함하는 전자 장치의 TTS 변환 방법.

청구항 18

제 17 항에 있어서,

상기 제 1 정보 또는 상기 제 2 정보는,

상기 음향의 언어 정보 및/또는 화자 정보를 포함하는 전자 장치의 TTS 변환 방법.

청구항 19

제 17 항에 있어서,

상기 공통 음향 데이터 셋을 생성하는 동작은,

상기 판단에 기초하여, 상기 유사도가 선택된 쓰레쉬 홀드(threshold) 값 이상인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부, 및 상기 제 2 음향 데이터 셋의 적어도 일부 모두에 대응하는 제 1 파라미터를 결정하는 동작;

상기 유사도가 상기 쓰레쉬 홀드 값 미만인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부에 대응하는 제 2 파라미터, 및 상기 제 2 음향 데이터 셋의 적어도 일부에 대응하는 제 3 파라미터를 결정하는 동작; 및

상기 제 1 파라미터, 상기 제 2 파라미터, 또는 상기 제 3 파라미터에 기초하여 상기 공통 음향 데이터 셋을 생성하는 동작을 포함하는 전자 장치의 TTS 변환 방법.

청구항 20

제 19 항에 있어서,

상기 제 1 파라미터, 상기 제 2 파라미터, 또는 상기 제 3 파라미터는,

상기 음향의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응하는 전자 장치의 TTS 변환 방법.

발명의 설명

기술 분야

[0001] 본 발명의 다양한 실시예는 파라미터 기반의 문자-음성 변환(TTS, Text To Speech)을 수행하는 전자 장치에 관한 것으로서, 공통 음향 데이터 셋을 활용하여 다언어/다화자를 지원하는 공통 음향 데이터 셋을 활용하여 TTS 변환을 수행하는 전자 장치 및 그의 TTS 변환 방법에 관한 것이다.

배경 기술

[0002] 파라미터 기반의 TTS 변환은 각 언어별로 언어 처리부와 음성 데이터를 가질 수 있고, 입력된 문장에 대한 문장 분석 결과에 기초하여 적절한 음성 데이터를 선택하여 이들의 연결 및 변환을 통해 합성음을 생성할 수 있다. TTS 변환은 CODEC(Coder-Decoder)과 같이 음성이 입력으로 들어가는 것이 아니라, 텍스트를 입력으로 받기 때문에 텍스트에 맞는 음성 데이터를 추정해서 음향 모델의 형태로 저장하는 과정이 선행될 수 있다. 파라미터 기반의 TTS는 각 언어 및 화자 별로 음향모델을 가질 수 있으며, 각각의 음향모델은 5MB 내외의 크기로 구성될 수 있다.

발명의 내용

해결하려는 과제

[0003] TTS를 다국어에 대해 상용 서비스하는 경우, 서비스 언어의 수 및 언어별 지원 화자 수가 증가함에 따라 언어의 종류 또는 화자의 종류에 관한 음향 모델의 음성 데이터가 비례하여 증가하므로, 전자 장치의 용량 부담이 커지는 문제점이 존재할 수 있다. 또한, 결정-트리(decision-tree) 기반의 음향 모델은 음소 단위를 분할한 세분 음소 단위의 음향 데이터를 나타내는 리프 노드(leaf node)를 다량으로 생성할 수 있고, 이러한 세분 음소 단위의 음향 신호는 사람의 귀로 구분하기 쉽지 않다. 이와 같이 유사한 형태의 리프 노드가 다량으로 생성되는 현상은

이중 언어 및 화자 사이에 두드러지게 나타날 수 있고, 이는 언어 및 화자 별로 구분해서 저장하는 음향모델 자체에 높은 리던던시(redundancy)를 포함하는 문제점이 존재할 수 있다.

[0004] 본 발명의 다양한 실시예에는 다국어 및 다화자가 공유하는 공통 음향 데이터(super-clustered common acoustic data, SCCAD)를 구성할 수 있고, 다국어 및 다화자를 지원하는 공통 음향 데이터를 통하여 파라미터 기반의 TTS 변환을 수행함으로써, 용량이 크게 절감된 TTS 변환 방법 및 장치를 제공할 수 있다.

과제의 해결 수단

[0005] 본 발명의 다양한 실시예들에 따른 전자 장치는 프로세서, 및 상기 프로세서와 전기적으로 연결된 메모리를 포함하고, 상기 메모리는, 공통 음향 데이터 셋을 저장하고, 상기 메모리는, 실행 시에, 상기 프로세서가, 적어도 하나의 텍스트를 획득하고, 상기 획득된 텍스트가 변환될 음향과 관련된 정보를 선택하고, 상기 선택된 정보에 제 1 정보인 경우, 복수의 제 1 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 1 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 일부를 로드하고, 상기 로드된 일부에 기초하여 제 1 음향 신호를 발생시키고, 상기 선택된 정보가 제 2 정보인 경우, 복수의 제 2 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 2 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 상기 일부 또는 다른 일부를 로드하고, 상기 로드된 일부 또는 다른 일부에 기초하여 제 2 음향 신호를 발생시키도록 하는 인스트럭션들을 저장할 수 있다.

발명의 효과

[0006] 본 발명의 다양한 실시예들에 따른 전자 장치는 다국어 및 다화자를 지원하는 하나의 공통 음향 데이터 셋을 통하여 TTS 변환을 수행함으로써, 전자 장치가 복수개의 음향 데이터 셋을 저장하는데 필요한 저장 공간을 감소시킬 수 있다.

[0007] 본 발명의 다양한 실시예들에 따른 전자 장치는 새로운 언어 또는 새로운 화자의 음향 모델을 전자 장치에 추가로 설치할 때, 이미 생성된 공통 음향 데이터 셋에 대한 추가 음향 모델의 링커(linker)만 다운받음으로써, 전자 장치가 데이터 전송에 필요한 부담을 감소시킬 수 있다.

도면의 간단한 설명

- [0008] 도 1는 본 발명의 실시예에 따른 전자 장치를 포함하는 네트워크 환경을 도시한다.
- 도 2는 본 발명의 다양한 실시예들에 따른 전자 장치의 블록도이다.
- 도 3은 다양한 실시예에 따른 프로그램 모듈의 블록도이다.
- 도 4는 본 발명의 다양한 실시예들에 따른 전자 장치가 텍스트가 변환될 음향과 관련된 정보를 선택하고, 선택된 정보에 기초하여 음향 신호를 발생시키는 동작을 나타내는 순서도이다.
- 도 5는 본 발명의 다양한 실시예들에 따른 전자 장치가 음향 데이터 셋의 적어도 하나의 경로를 공통 음향 데이터 셋의 적어도 일부에 맵핑하는 동작을 나타내는 도면이다.
- 도 6은 본 발명의 다양한 실시예들에 따른 전자 장치가 공통 음향 데이터를 생성하는 동작을 나타내는 순서도이다.
- 도 7a는 본 발명의 다양한 실시예들에 따른 전자 장치가 제 1 음향 데이터 셋의 적어도 일부, 및 제 2 음향 데이터 셋의 적어도 일부의 유사도를 판단하고, 상기 유사도 판단에 기초하여 공통 음향 데이터 셋을 생성하는 동작을 나타내는 도면이다.
- 도 7b는 본 발명의 다양한 실시예들에 따른 전자 장치가 적어도 하나의 음향 데이터 셋을 수집한 전체 음향 데이터 셋에서 클러스터링(clustering) 알고리즘을 수행하는 동작을 나타내는 도면이다.
- 도 8은 본 발명의 다양한 실시예들에 따른 전자 장치가 공통 음향 데이터 셋을 생성하고 특정 음향 데이터 셋의 복수의 경로들을 상기 공통 음향 데이터 셋에 매칭하는 동작을 나타내는 도면이다.
- 도 9는 본 발명의 다양한 실시예들에 따른 제 1 전자 장치의 블록도 및 제 2 전자 장치의 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0009] 이하, 본 문서의 다양한 실시예가 첨부된 도면을 참조하여 기재된다. 그러나, 이는 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 문서의 실시예의 다양한 변경(modifications), 균등물(equivalents), 및/또는 대체물(alternatives)을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다.
- [0010] 본 문서에서, "가진다," "가질 수 있다," "포함한다," 또는 "포함할 수 있다" 등의 표현은 해당 특징(예: 수치, 기능, 동작, 또는 부품 등의 구성요소)의 존재를 가리키며, 추가적인 특징의 존재를 배제하지 않는다.
- [0011] 본 문서에서, "A 또는 B," "A 또는/및 B 중 적어도 하나," 또는 "A 또는/및 B 중 하나 또는 그 이상" 등의 표현은 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. 예를 들면, "A 또는 B," "A 및 B 중 적어도 하나," 또는 "A 또는 B 중 적어도 하나"는, (1) 적어도 하나의 A를 포함, (2) 적어도 하나의 B를 포함, 또는 (3) 적어도 하나의 A 및 적어도 하나의 B 모두를 포함하는 경우를 모두 지칭할 수 있다.
- [0012] 본 문서에서 사용된 "제 1," "제 2," "첫째," 또는 "둘째," 등의 표현들은 다양한 구성요소들을, 순서 및/또는 중요도에 상관없이 수식할 수 있고, 한 구성요소를 다른 구성요소와 구분하기 위해 사용될 뿐 해당 구성요소들을 한정하지 않는다. 예를 들면, 제 1 사용자 기기와 제 2 사용자 기기는, 순서 또는 중요도와 무관하게, 서로 다른 사용자 기기를 나타낼 수 있다. 예를 들면, 본 문서에 기재된 권리 범위를 벗어나지 않으면서 제 1 구성요소는 제 2 구성요소로 명명될 수 있고, 유사하게 제 2 구성요소도 제 1 구성요소로 바꾸어 명명될 수 있다.
- [0013] 어떤 구성요소(예: 제 1 구성요소)가 다른 구성요소(예: 제 2 구성요소)에 "(기능적으로 또는 통신적으로) 연결되어((operatively or communicatively) coupled with/to)" 있거나 "접속되어(connected to)" 있다고 언급된 때에는, 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로 연결되거나, 다른 구성요소(예: 제 3 구성요소)를 통하여 연결될 수 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소(예: 제 1 구성요소)가 다른 구성요소(예: 제 2 구성요소)에 "직접 연결되어" 있거나 "직접 접속되어" 있다고 언급된 때에는, 상기 어떤 구성요소와 상기 다른 구성요소 사이에 다른 구성요소(예: 제 3 구성요소)가 존재하지 않는 것으로 이해될 수 있다.
- [0014] 본 문서에서 사용된 표현 "~하도록 구성된(또는 설정된)(configured to)"은 상황에 따라, 예를 들면, "~에 적합한(suitable for)," "~하는 능력을 가지는(having the capacity to)," "~하도록 설계된(designed to)," "~하도록 변경된(adapted to)," "~하도록 만들어진(made to)," 또는 "~를 할 수 있는(capable of)"과 바꾸어 사용될 수 있다. 용어 "~하도록 구성된(또는 설정된)"은 하드웨어적으로 "특별히 설계된(specifically designed to)" 것만을 반드시 의미하지 않을 수 있다. 대신, 어떤 상황에서는, "~하도록 구성된 장치"라는 표현은, 그 장치가 다른 장치 또는 부품들과 함께 "~할 수 있는" 것을 의미할 수 있다. 예를 들면, 문구 "A, B, 및 C를 수행하도록 구성된(또는 설정된) 프로세서"는 해당 동작을 수행하기 위한 전용 프로세서(예: 임베디드 프로세서), 또는 메모리 장치에 저장된 하나 이상의 소프트웨어 프로그램들을 실행함으로써, 해당 동작들을 수행할 수 있는 범용 프로세서(generic-purpose processor)(예: CPU 또는 application processor)를 의미할 수 있다.
- [0015] 본 문서에서 사용된 용어들은 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 다른 실시예의 범위를 한정하려는 의도가 아닐 수 있다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 용어들은 본 문서에 기재된 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가질 수 있다. 본 문서에 사용된 용어들 중 일반적인 사전에 정의된 용어들은, 관련 기술의 문맥상 가지는 의미와 동일 또는 유사한 의미로 해석될 수 있으며, 본 문서에서 명백하게 정의되지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다. 경우에 따라서, 본 문서에서 정의된 용어일지라도 본 문서의 실시예들을 배제하도록 해석될 수 없다.
- [0016] 본 문서의 다양한 실시예들에 따른 전자 장치는, 예를 들면, 스마트폰(smartphone), 태블릿 PC(tablet personal computer), 이동 전화기(mobile phone), 영상 전화기, 전자책 리더기(e-book reader), 데스크탑 PC(desktop personal computer), 랩탑 PC(laptop personal computer), 넷북 컴퓨터(netbook computer), 워크스테이션(workstation), 서버, PDA(personal digital assistant), PMP(portable multimedia player), MP3 플레이어, 모바일 의료기기, 카메라(camera), 또는 웨어러블 장치(wearable device) 중 적어도 하나를 포함할 수 있다. 다양한 실시예에 따르면, 웨어러블 장치는 액세서리형(예: 시계, 반지, 팔찌, 발찌, 목걸이, 안경, 콘택트 렌즈, 또는 머리 착용형 장치(head-mounted-device(HMD)), 직물 또는 의류 일체형(예: 전자 의복), 신체 부착형(예: 스킨 패드(skin pad) 또는 문신), 또는 생체 이식형(예: implantable circuit) 중 적어도 하나를 포함할 수 있다.

- [0017] 어떤 실시예들에서, 전자 장치는 가전 제품(home appliance)일 수 있다. 가전 제품은, 예를 들면, 텔레비전, DVD(digital video disk) 플레이어, 오디오, 냉장고, 에어컨, 청소기, 오븐, 전자레인지, 세탁기, 공기 청정기, 셋톱 박스(set-top box), 홈 오토메이션 컨트롤 패널(home automation control panel), 보안 컨트롤 패널(security control panel), TV 박스(예: 삼성 HomeSync™, 애플TV™, 또는 구글 TV™), 게임 콘솔(예: Xbox™, PlayStation™), 전자 사전, 전자 키, 캠코더(camcorder), 또는 전자 액자 중 적어도 하나를 포함할 수 있다.
- [0018] 다른 실시예에서, 전자 장치는, 각종 의료기기(예: 각종 휴대용 의료측정기기(혈당 측정기, 심박 측정기, 혈압 측정기, 또는 체온 측정기 등), MRA(magnetic resonance angiography), MRI(magnetic resonance imaging), CT(computed tomography), 촬영기, 또는 초음파기 등), 네비게이션(navigation) 장치, 위성 항법 시스템(GNSS(global navigation satellite system)), EDR(event data recorder), FDR(flight data recorder), 자동차 인포테인먼트(infotainment) 장치, 선박용 전자 장비(예: 선박용 항법 장치, 자이로 콤팩스 등), 항공 전자 기기(avionics), 보안 기기, 차량용 헤드 유닛(head unit), 산업용 또는 가정용 로봇, 금융 기관의 ATM(automatic teller's machine), 상점의 POS(point of sales), 또는 사물 인터넷 장치(internet of things)(예: 전구, 각종 센서, 전기 또는 가스 미터기, 스프링클러 장치, 화재경보기, 온도조절기(thermostat), 가로등, 토스터(toaster), 운동기구, 온수탱크, 히터, 보일러 등) 중 적어도 하나를 포함할 수 있다.
- [0019] 어떤 실시예에 따르면, 전자 장치는 가구(furniture) 또는 건물/구조물의 일부, 전자 보드(electronic board), 전자 사인 수신 장치(electronic signature receiving device), 프로젝터(projector), 또는 각종 계측 기기(예: 수도, 전기, 가스, 또는 전파 계측 기기 등) 중 적어도 하나를 포함할 수 있다. 다양한 실시예에서, 전자 장치는 전술한 다양한 장치들 중 하나 또는 그 이상의 조합일 수 있다. 어떤 실시예에 따른 전자 장치는 플렉서블 전자 장치일 수 있다. 또한, 본 문서의 실시예에 따른 전자 장치는 전술한 기기들에 한정되지 않으며, 기술 발전에 따른 새로운 전자 장치를 포함할 수 있다.
- [0020] 이하, 첨부 도면을 참조하여, 다양한 실시예에 따른 전자 장치가 설명된다. 본 문서에서, 사용자라는 용어는 전자 장치를 사용하는 사람 또는 전자 장치를 사용하는 장치(예: 인공지능 전자 장치)를 지칭할 수 있다.
- [0021] 도 1을 참조하여, 다양한 실시예에서의, 네트워크 환경(100) 내의 전자 장치(101)가 기재된다. 전자 장치(101)는 버스(110), 프로세서(120), 메모리(130), 입출력 인터페이스(150), 디스플레이(160), 및 통신 인터페이스(170)를 포함할 수 있다. 어떤 실시예에서는, 전자 장치(101)는, 구성요소들 중 적어도 하나를 생략하거나 다른 구성요소를 추가적으로 구비할 수 있다.
- [0022] 버스(110)는, 예를 들면, 구성요소들(110-170)을 서로 연결하고, 구성요소들 간의 통신(예: 제어 메시지 및/또는 데이터)을 전달하는 회로를 포함할 수 있다.
- [0023] 프로세서(120)는, 중앙처리장치(central processing unit(CPU)), 어플리케이션 프로세서(application processor(AP)), 또는 커뮤니케이션 프로세서(communication processor(CP)) 중 하나 또는 그 이상을 포함할 수 있다. 프로세서(120)는, 예를 들면, 전자 장치(101)의 적어도 하나의 다른 구성요소들의 제어 및/또는 통신에 관한 연산이나 데이터 처리를 실행할 수 있다.
- [0024] 메모리(130)는, 휘발성 및/또는 비휘발성 메모리를 포함할 수 있다. 메모리(130)는, 예를 들면, 전자 장치(101)의 적어도 하나의 다른 구성요소에 관계된 명령 또는 데이터를 저장할 수 있다. 한 실시예에 따르면, 메모리(130)는 소프트웨어 및/또는 프로그램(140)을 저장할 수 있다. 프로그램(140)은, 예를 들면, 커널(kernel)(141), 미들웨어(middleware)(143), 어플리케이션 프로그래밍 인터페이스(application programming interface(API))(145), 및/또는 어플리케이션 프로그램(또는 "어플리케이션")(147) 등을 포함할 수 있다. 커널(kernel)(141), 미들웨어(middleware)(143), 또는 API(API)(145)의 적어도 일부는, 운영 시스템(operating system(OS))으로 지칭될 수 있다.
- [0025] 커널(kernel)(141)은, 예를 들면, 다른 프로그램들(예: 미들웨어(middleware)(143), API(API)(145), 또는 어플리케이션 프로그램(application program)(147))에 구현된 동작 또는 기능을 실행하는 데 사용되는 시스템 리소스들(예: 버스(bus)(110), 프로세서(processor)(120), 또는 메모리(memory)(130) 등)를 제어 또는 관리할 수 있다. 또한, 커널(kernel)(141)은 미들웨어(middleware)(143), API(API)(145), 또는 어플리케이션 프로그램(application program)(147)에서 전자 장치(101)의 개별 구성요소에 접근함으로써, 시스템 리소스들을 제어 또는 관리할 수 있는 인터페이스를 제공할 수 있다.
- [0026] 미들웨어(middleware)(143)는, 예를 들면, API(API)(145) 또는 어플리케이션 프로그램(application program)(147)이 커널(kernel)(141)과 통신하여 데이터를 주고받을 수 있도록 중개 역할을 수행할 수 있다.
- [0027] 또한, 미들웨어(middleware)(143)는 어플리케이션 프로그램(application program)(147)으로부터 수신된 하나 이상의 작업 요청들을 우선 순위에 따

라 처리할 수 있다. 예를 들면, 미들웨어(143)는 어플리케이션 프로그램(147) 중 적어도 하나에 전자 장치(101)의 시스템 리소스(예: 버스(110), 프로세서(120), 또는 메모리(130) 등)를 사용할 수 있는 우선 순위를 부여할 수 있다. 예컨대, 미들웨어(143)는 상기 적어도 하나에 부여된 우선 순위에 따라 상기 하나 이상의 작업 요청들을 처리함으로써, 상기 하나 이상의 작업 요청들에 대한 스케줄링 또는 로드 밸런싱 등을 수행할 수 있다.

[0028] API(145)는, 예를 들면, 어플리케이션(147)이 커널(141) 또는 미들웨어(143)에서 제공되는 기능을 제어하기 위한 인터페이스로, 예를 들면, 파일 제어, 창 제어, 영상 처리, 또는 문자 제어 등을 위한 적어도 하나의 인터페이스 또는 함수(예: 명령어)를 포함할 수 있다.

[0029] 입출력 인터페이스(150)는, 예를 들면, 사용자 또는 다른 외부 기기로부터 입력된 명령 또는 데이터를 전자 장치(101)의 다른 구성요소(들)에 전달할 수 있는 인터페이스의 역할을 할 수 있다. 또한, 입출력 인터페이스(150)는 전자 장치(101)의 다른 구성요소(들)로부터 수신된 명령 또는 데이터를 사용자 또는 다른 외부 기기로부터 출력할 수 있다.

[0030] 디스플레이(160)는, 예를 들면, 액정 디스플레이(liquid crystal display(LCD)), 발광 다이오드(light-emitting diode(LED)) 디스플레이, 유기 발광 다이오드(organic light-emitting diode(OLED)) 디스플레이, 또는 마이크로 전자기계 시스템(microelectromechanical systems(MEMS)) 디스플레이, 또는 전자종이(electronic paper) 디스플레이를 포함할 수 있다. 디스플레이(160)는, 예를 들면, 사용자에게 각종 콘텐츠(예: 텍스트, 이미지, 비디오, 아이콘, 또는 심볼 등)를 표시할 수 있다. 디스플레이(160)는, 터치 스크린을 포함할 수 있으며, 예를 들면, 전자 펜 또는 사용자의 신체의 일부를 이용한 터치, 제스처, 근접, 또는 호버링 입력을 수신할 수 있다.

[0031] 통신 인터페이스(170)는, 예를 들면, 전자 장치(101)와 외부 장치(예: 제1 외부 전자 장치(102), 제2 외부 전자 장치(104), 또는 서버(106)) 간의 통신을 설정할 수 있다. 예를 들면, 통신 인터페이스(170)는 무선 통신 또는 유선 통신을 통해서 네트워크(162)에 연결되어 외부 장치(예: 제2 외부 전자 장치(104) 또는 서버(106))와 통신할 수 있다.

[0032] 무선 통신은, 예를 들면, 셀룰러 통신 프로토콜로서, 예를 들면, LTE(long-term evolution), LTE-A(LTE Advance), CDMA(code division multiple access), WCDMA(wideband CDMA), UMTS(universal mobile telecommunications system), WiBro(Wireless Broadband), 또는 GSM(Global System for Mobile Communications) 등 중 적어도 하나를 사용할 수 있다. 또한, 무선 통신은, 예를 들면, 근거리 통신(164)을 포함할 수 있다. 근거리 통신(164)은, 예를 들면, WiFi(wireless fidelity), 블루투스(Bluetooth), NFC(near field communication), 또는 GNSS(global navigation satellite system) 등 중 적어도 하나를 포함할 수 있다. GNSS는 사용 지역 또는 대역폭 등에 따라, 예를 들면, GPS(Global Positioning System), Glonass(Global Navigation Satellite System), Beidou Navigation Satellite System(이하 "Beidou") 또는 Galileo, the European global satellite-based navigation system 중 적어도 하나를 포함할 수 있다. 이하, 본 문서에서는, "GPS"는 "GNSS"와 혼용되어 사용(interchangeably used)될 수 있다. 유선 통신은, 예를 들면, USB(universal serial bus), HDMI(high definition multimedia interface), RS-232(recommended standard232), 또는 POTS(plain old telephone service) 등 중 적어도 하나를 포함할 수 있다. 네트워크(162)는 통신 네트워크(telecommunications network), 예를 들면, 컴퓨터 네트워크(computer network)(예: LAN 또는 WAN), 인터넷, 또는 전화 망(telephone network) 중 적어도 하나를 포함할 수 있다.

[0033] 제1 및 제2 외부 전자 장치(102, 104) 각각은 전자 장치(101)와 동일한 또는 다른 종류의 장치일 수 있다. 한 실시예에 따르면, 서버(106)는 하나 또는 그 이상의 서버들의 그룹을 포함할 수 있다. 다양한 실시예에 따르면, 전자 장치(101)에서 실행되는 동작들의 전부 또는 일부는 다른 하나 또는 복수의 전자 장치(예: 전자 장치(102, 104), 또는 서버(106))에서 실행될 수 있다. 한 실시예에 따르면, 전자 장치(101)가 어떤 기능이나 서비스를 자동으로 또는 요청에 의하여 수행해야 할 경우에, 전자 장치(101)는 기능 또는 서비스를 자체적으로 실행시키는 대신에 또는 추가적으로, 그와 연관된 적어도 일부 기능을 다른 장치(예: 전자 장치(102, 104), 또는 서버(106))에게 요청할 수 있다. 다른 전자 장치(예: 전자 장치(102, 104), 또는 서버(106))는 요청된 기능 또는 추가 기능을 실행하고, 그 결과를 전자 장치(101)로 전달할 수 있다. 전자 장치(101)는 수신된 결과를 그대로 또는 추가적으로 처리하여 요청된 기능이나 서비스를 제공할 수 있다. 이를 위하여, 예를 들면, 클라우드 컴퓨팅, 분산 컴퓨팅, 또는 클라이언트-서버 컴퓨팅 기술이 이용될 수 있다.

[0034] 도 2는 다양한 실시예에 따른 전자 장치(201)의 블록도이다. 전자 장치(201)는, 예를 들면, 도 1에 도시된 전자 장치(101)의 전체 또는 일부를 포함할 수 있다. 전자 장치(201)는 하나 이상의 프로세서(예: AP(application

processor))(210), 통신 모듈(220), (가입자 식별 모듈(224), 메모리(230), 센서 모듈(240), 입력 장치(250), 디스플레이(260), 인터페이스(270), 오디오 모듈(280), 카메라 모듈(291), 전력 관리 모듈(295), 배터리(296), 인디케이터(297), 및 모터(298) 를 포함할 수 있다.

- [0035] 프로세서(210)는, 예를 들면, 운영 체제 또는 응용 프로그램을 구동하여 프로세서(210)에 연결된 다수의 하드웨어 또는 소프트웨어 구성요소들을 제어할 수 있고, 각종 데이터 처리 및 연산을 수행할 수 있다. 프로세서(210)는, 예를 들면, SoC(system on chip) 로 구현될 수 있다. 한 실시예에 따르면, 프로세서(210)는 GPU(graphic processing unit) 및/또는 이미지 신호 프로세서(image signal processor)를 더 포함할 수 있다. 프로세서(210)는 도 2에 도시된 구성요소들 중 적어도 일부(예: 셀룰러 모듈(221))를 포함할 수도 있다. 프로세서(210)는 다른 구성요소들(예: 비휘발성 메모리) 중 적어도 하나로부터 수신된 명령 또는 데이터를 휘발성 메모리에 로드(load)하여 처리하고, 다양한 데이터를 비휘발성 메모리에 저장(store)할 수 있다.
- [0036] 통신 모듈(220)은, 도 1의 통신 인터페이스(170)와 동일 또는 유사한 구성을 가질 수 있다. 통신 모듈(220)은, 예를 들면, 셀룰러 모듈(221), WiFi 모듈(223), 블루투스 모듈(225), GNSS 모듈(227)(예: GPS 모듈, Glonass 모듈, Beidou 모듈, 또는 Galileo 모듈), NFC 모듈(228) 및 RF(radio frequency) 모듈(229)를 포함할 수 있다.
- [0037] 셀룰러 모듈(221)은, 예를 들면, 통신망을 통해서 음성 통화, 영상 통화, 문자 서비스, 또는 인터넷 서비스 등을 제공할 수 있다. 한 실시예에 따르면, 셀룰러 모듈(221)은 가입자 식별 모듈(예: SIM 카드)(224)을 이용하여 통신 네트워크 내에서 전자 장치(201)의 구별 및 인증을 수행할 수 있다. 한 실시예에 따르면, 셀룰러 모듈(221)은 프로세서(210)가 제공할 수 있는 기능 중 적어도 일부 기능을 수행할 수 있다. 한 실시예에 따르면, 셀룰러 모듈(221)은 커뮤니케이션 프로세서(CP: communication processor)를 포함할 수 있다.
- [0038] WiFi 모듈(223), 블루투스 모듈(225), GNSS 모듈(227) 또는 NFC 모듈(228) 각각은, 예를 들면, 해당하는 모듈을 통해서 송수신되는 데이터를 처리하기 위한 프로세서를 포함할 수 있다. 어떤 실시예에 따르면, 셀룰러 모듈(221), WiFi 모듈(223), 블루투스 모듈(225), GNSS 모듈(227) 또는 NFC 모듈(228) 중 적어도 일부(예: 두 개 이상)는 하나의 integrated chip(IC) 또는 IC 패키지 내에 포함될 수 있다.
- [0039] RF 모듈(229)은, 예를 들면, 통신 신호(예: RF 신호)를 송수신할 수 있다. RF 모듈(229)은, 예를 들면, 트랜시버(transceiver), PAM(power amp module), 주파수 필터(frequency filter), LNA(low noise amplifier), 또는 안테나 등을 포함할 수 있다. 다른 실시예에 따르면, 셀룰러 모듈(221), WiFi 모듈(223), 블루투스 모듈(225), GNSS 모듈(227) 또는 NFC 모듈(228) 중 적어도 하나는 별개의 RF 모듈을 통하여 RF 신호를 송수신할 수 있다.
- [0040] 가입자 식별 모듈(224)은, 예를 들면, 가입자 식별 모듈을 포함하는 카드 및/또는 내장 SIM(embedded SIM)을 포함할 수 있으며, 고유한 식별 정보(예: ICCID(integrated circuit card identifier)) 또는 가입자 정보(예: IMSI(international mobile subscriber identity))를 포함할 수 있다.
- [0041] 메모리(230)(예: 메모리(130))는, 예를 들면, 내장 메모리(232) 또는 외장 메모리(234)를 포함할 수 있다. 내장 메모리(232)는, 예를 들면, 휘발성 메모리(예: DRAM(dynamic RAM), SRAM(static RAM), 또는 SDRAM(synchronous dynamic RAM) 등), 비휘발성 메모리(non-volatile Memory)(예: OTPROM(one time programmable ROM), PROM(programmable ROM), EPROM(erasable and programmable ROM), EEPROM(electrically erasable and programmable ROM), mask ROM, flash ROM, 플래시 메모리(예: NAND flash 또는 NOR flash 등), 하드 드라이브, 또는 솔리드 스테이트 드라이브(solid state drive(SSD)) 중 적어도 하나를 포함할 수 있다.
- [0042] 외장 메모리(234)는 플래시 드라이브(flash drive), 예를 들면, CF(compact flash), SD(secure digital), Micro-SD(micro secure digital), Mini-SD(mini secure digital), xD(extreme digital), MMC(multi-media card) 또는 메모리 스틱(memory stick) 등을 더 포함할 수 있다. 외장 메모리(234)는 다양한 인터페이스를 통하여 전자 장치(201)와 기능적으로 및/또는 물리적으로 연결될 수 있다.
- [0043] 본 발명의 다양한 실시예들에 따른 메모리(230)는, 실행 시에, 상기 프로세서(210)가, 적어도 하나의 텍스트를 획득하고, 상기 획득된 텍스트가 변환될 음향과 관련된 정보를 선택하고, 상기 선택된 정보가 제 1 정보인 경우, 복수의 제 1 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 1 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 일부를 로드하고, 상기 로드된 일부에 기초하여 제 1 음향 신호를 발생시키고, 상기 선택된 정보가 제 2 정보인 경우, 복수의 제 2 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 2 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 상기 일부 또는 다른 일부를 로드하고, 상기 로드된 일부 또는 다른 일부에 기초하여 제 2 음향 신호를 발생시키도록 하는 인스트럭션들을 저장할 수 있다.
- [0044] 본 발명의 다양한 실시예들에 따른 메모리(230)는, 실행 시에, 상기 프로세서(210)가, 사용자로부터 상기 적어

도 하나의 텍스트를 획득하거나 외부 장치로부터 상기 적어도 하나의 텍스트를 포함하는 문자 메시지를 수신하도록 하는 인스트럭션들을 저장할 수 있다.

[0045] 본 발명의 다양한 실시예들에 따른 메모리(230)는, 실행 시에, 상기 프로세서(210)가, 상기 입력된 텍스트에 기초하여, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택하고, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여, 상기 제 1 음향 신호 또는 상기 제 2 음향 신호를 발생시키도록 하는 인스트럭션들을 저장할 수 있다.

[0046] 본 발명의 다양한 실시예들에 따른 메모리(230)는, 실행 시에, 상기 프로세서(210)가, 음향과 관련된 제 1 정보에 대응하는 제 1 음향 데이터 셋, 및/또는 상기 음향과 관련된 제 2 정보에 대응하는 제 2 음향 데이터 셋을 획득하고, 상기 제 1 음향 데이터 셋의 적어도 일부, 및/또는 상기 제 2 음향 데이터 셋의 적어도 일부를 유사도를 판단하고, 상기 판단에 기초하여, 상기 제 1 음향 데이터 셋의 적어도 일부, 및/또는 상기 제 2 음향 데이터 셋의 적어도 일부와 관련된 공통 음향 데이터 셋을 생성하도록 하는 인스트럭션들을 저장할 수 있다.

[0047] 본 발명의 다양한 실시예들에 따른 메모리(230)는, 실행 시에, 상기 프로세서(210)가, 상기 판단에 기초하여, 상기 유사도가 선택된 쓰레쉬 홀드(threshold) 값 이상인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부, 및 상기 제 2 음향 데이터 셋의 적어도 일부 모두에 대응하는 제 1 파라미터를 결정하고, 상기 유사도가 상기 쓰레쉬 홀드 값 미만인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부에 대응하는 제 2 파라미터, 및 상기 제 2 음향 데이터 셋의 적어도 일부에 대응하는 제 3 파라미터를 결정하고, 상기 제 1 파라미터, 상기 제 2 파라미터, 또는 상기 제 3 파라미터에 기초하여 상기 공통 음향 데이터 셋을 생성하도록 하는 인스트럭션들을 저장할 수 있다.

[0048] 본 발명의 다양한 실시예들에 따른 메모리(230)는 공통 음향 데이터 셋, 적어도 하나의 결정 트리에 관한 정보, 및 상기 결정 트리의 인덱스가 지시하는 적어도 하나의 음향 데이터 셋을 저장할 수 있다.

[0049] 센서 모듈(240)은, 예를 들면, 물리량을 계측하거나 전자 장치(201)의 작동 상태를 감지하여, 계측 또는 감지된 정보를 전기 신호로 변환할 수 있다. 센서 모듈(240)은, 예를 들면, 제스처 센서(240A), 자이로 센서(240B), 기압 센서(240C), 마그네틱 센서(240D), 가속도 센서(240E), 그립 센서(240F), 근접 센서(240G), 컬러(color) 센서(240H)(예: RGB(red, green, blue) 센서), 생체 센서(240I), 온/습도 센서(240J), 조도 센서(240K), 또는 UV(ultra violet) 센서(240M) 중의 적어도 하나를 포함할 수 있다. 추가적으로 또는 대체적으로(Additionally or alternatively), 센서 모듈(240)은, 예를 들면, 후각 센서(E-nose sensor), EMG 센서(electromyography sensor), EEG 센서(electroencephalogram sensor), ECG 센서(electrocardiogram sensor), IR(infrared) 센서, 홍채 센서 및/또는 지문 센서를 포함할 수 있다. 센서 모듈(240)은 그 안에 속한 적어도 하나 이상의 센서들을 제어하기 위한 제어 회로를 더 포함할 수 있다. 어떤 실시예에서는, 전자 장치(201)는 프로세서(210)의 일부로서 또는 별도로, 센서 모듈(240)을 제어하도록 구성된 프로세서를 더 포함하여, 프로세서(210)가 슬립(sleep) 상태에 있는 동안, 센서 모듈(240)을 제어할 수 있다.

[0050] 입력 장치(250)는, 예를 들면, 터치 패널(touch panel)(252), (디지털) 펜 센서(pen sensor)(254), 키(key)(256), 또는 초음파(ultrasonic) 입력 장치(258)를 포함할 수 있다. 터치 패널(252)은, 예를 들면, 정전식, 감압식, 적외선 방식, 또는 초음파 방식 중 적어도 하나의 방식을 사용할 수 있다. 또한, 터치 패널(252)은 제어 회로를 더 포함할 수도 있다. 터치 패널(252)은 택타일 레이어(tactile layer)를 더 포함하여, 사용자에게 촉각 반응을 제공할 수 있다.

[0051] (디지털) 펜 센서(254)는, 예를 들면, 터치 패널의 일부이거나, 별도의 인식용 시트(sheet)를 포함할 수 있다. 키(256)는, 예를 들면, 물리적인 버튼, 광학식 키, 또는 키패드를 포함할 수 있다. 초음파 입력 장치(258)는 마이크(예: 마이크(288))를 통해, 입력 도구에서 발생된 초음파를 감지하여, 상기 감지된 초음파에 대응하는 데이터를 확인할 수 있다.

[0052] 디스플레이(260)(예: 디스플레이(160))는 패널(262), 홀로그램 장치(264), 또는 프로젝터(266)를 포함할 수 있다. 패널(262)은, 도 1의 디스플레이(160)와 동일 또는 유사한 구성을 포함할 수 있다. 패널(262)은, 예를 들면, 유연하게(flexible), 투명하게(transparent), 또는 착용할 수 있게(wearable) 구현될 수 있다. 패널(262)은 터치 패널(252)과 하나의 모듈로 구성될 수도 있다. 홀로그램 장치(264)는 빛의 간섭을 이용하여 입체 영상을 허공에 보여줄 수 있다. 프로젝터(266)는 스크린에 빛을 투사하여 영상을 표시할 수 있다. 스크린은, 예를 들면, 전자 장치(201)의 내부 또는 외부에 위치할 수 있다. 한 실시예에 따르면, 디스플레이(260)는 패널(262), 홀로그램 장치(264), 또는 프로젝터(266)를 제어하기 위한 제어 회로를 더 포함할 수 있다.

- [0053] 인터페이스(270)는, 예를 들면, HDMI(high-definition multimedia interface)(272), USB(universal serial bus)(274), 광 인터페이스(optical interface)(276), 또는 D-sub(D-subminiature)(278)를 포함할 수 있다. 인터페이스(270)는, 예를 들면, 도 1에 도시된 통신 인터페이스(170)에 포함될 수 있다. 추가적으로 또는 대체적으로(additionally and alternatively), 인터페이스(270)는, 예를 들면, MHL(mobile high-definition link) 인터페이스, SD(secure digital) 카드/MMC(multi-media card) 인터페이스, 또는 IrDA(infrared data association) 규격 인터페이스를 포함할 수 있다.
- [0054] 오디오 모듈(280)은, 예를 들면, 소리(sound)와 전기 신호를 쌍방향으로 변환시킬 수 있다. 오디오 모듈(280)의 적어도 일부 구성요소는, 예를 들면, 도 1에 도시된 입출력 인터페이스(145)에 포함될 수 있다. 오디오 모듈(280)은, 예를 들면, 스피커(282), 리시버(284), 이어폰(286), 또는 마이크(288) 등을 통해 입력 또는 출력되는 소리 정보를 처리할 수 있다.
- [0055] 카메라 모듈(291)은, 예를 들면, 정지 영상 및 동영상을 촬영할 수 있는 장치로서, 한 실시예에 따르면, 하나 이상의 이미지 센서(예: 전면 센서 또는 후면 센서), 렌즈, ISP(image signal processor), 또는 플래시(flash)(예: LED 또는 xenon lamp 등)를 포함할 수 있다.
- [0056] 전력 관리 모듈(295)은, 예를 들면, 전자 장치(201)의 전력을 관리할 수 있다. 한 실시예에 따르면, 전력 관리 모듈(295)은 PMIC(power management integrated circuit), 충전 IC(charger integrated circuit), 또는 배터리 또는 연료 게이지(battery or fuel gauge)를 포함할 수 있다. PMIC는, 유선 및/또는 무선 충전 방식을 가질 수 있다. 무선 충전 방식은, 예를 들면, 자기공명 방식, 자기유도 방식 또는 전자기파 방식 등을 포함하며, 무선 충전을 위한 부가적인 회로, 예를 들면, 코일 루프, 공진 회로, 또는 정류기 등을 더 포함할 수 있다. 배터리 게이지는, 예를 들면, 배터리(296)의 잔량, 충전 중 전압, 전류, 또는 온도를 측정할 수 있다. 배터리(296)는, 예를 들면, 충전식 전지(rechargeable battery) 및/또는 태양 전지(solar battery)를 포함할 수 있다.
- [0057] 인디케이터(297)는 전자 장치(201) 또는 그 일부(예: 프로세서(210))의 특정 상태, 예를 들면, 부팅 상태, 메시지 상태 또는 충전 상태 등을 표시할 수 있다. 모터(298)는 전기적 신호를 기계적 진동으로 변환할 수 있고, 진동(vibration), 또는 햅틱(haptic) 효과 등을 발생시킬 수 있다. 도시되지는 않았으나, 전자 장치(201)는 모바일 TV 지원을 위한 처리 장치(예: GPU)를 포함할 수 있다. 모바일 TV 지원을 위한 처리 장치는, 예를 들면, DMB(digital multimedia broadcasting), DVB(digital video broadcasting), 또는 미디어플로(mediaFlo™) 등의 규격에 따른 미디어 데이터를 처리할 수 있다.
- [0058] 본 문서에서 기술된 구성요소들 각각은 하나 또는 그 이상의 부품(component)으로 구성될 수 있으며, 해당 구성요소의 명칭은 전자 장치의 종류에 따라서 달라질 수 있다. 다양한 실시예에서, 전자 장치는 본 문서에서 기술된 구성요소 중 적어도 하나를 포함하여 구성될 수 있으며, 일부 구성요소가 생략되거나 또는 추가적인 다른 구성요소를 더 포함할 수 있다. 또한, 다양한 실시예에 따른 전자 장치의 구성요소들 중 일부가 결합되어 하나의 개체(entity)로 구성됨으로써, 결합되기 이전의 해당 구성요소들의 기능을 동일하게 수행할 수 있다.
- [0059] 도 3은 다양한 실시예에 따른 프로그램 모듈의 블록도이다. 한 실시예에 따르면, 프로그램 모듈(310)(예: 프로그램(140))은 전자 장치(예: 전자 장치(101))에 관련된 자원을 제어하는 운영 체제(operating system(OS)) 및/또는 운영 체제 상에서 구동되는 다양한 어플리케이션(예: 어플리케이션 프로그램(147))을 포함할 수 있다. 운영 체제는, 예를 들면, 안드로이드(android), iOS, 윈도우즈(windows), 심비안(symbian), 타이젠(tizen), 또는 바다(bada) 등이 될 수 있다.
- [0060] 프로그램 모듈(310)은 커널(320), 미들웨어(330), 어플리케이션 프로그래밍 인터페이스(application programming interface (API))(360), 및/또는 어플리케이션(370)을 포함할 수 있다. 프로그램 모듈(310)의 적어도 일부는 전자 장치 상에 프리로드(preload) 되거나, 외부 전자 장치(예: 전자 장치(102, 104), 서버(106) 등)로부터 다운로드(download) 가능하다.
- [0061] 커널(320)(예: 커널(141))은, 예를 들면, 시스템 리소스 매니저(321) 및/또는 디바이스 드라이버(323)를 포함할 수 있다. 시스템 리소스 매니저(321)는 시스템 리소스의 제어, 할당, 또는 회수 등을 수행할 수 있다. 한 실시예에 따르면, 시스템 리소스 매니저(321)는 프로세스 관리부, 메모리 관리부, 또는 파일 시스템 관리부 등을 포함할 수 있다. 디바이스 드라이버(323)는, 예를 들면, 디스플레이 드라이버, 카메라 드라이버, 블루투스 드라이버, 공유 메모리 드라이버, USB 드라이버, 키패드 드라이버, WiFi 드라이버, 오디오 드라이버, 또는 IPC(inter-process communication) 드라이버를 포함할 수 있다.
- [0062] 미들웨어(330)는, 예를 들면, 어플리케이션(370)이 공통적으로 필요로 하는 기능을 제공하거나, 어플리케이션

(370)이 전자 장치 내부의 제한된 시스템 자원을 효율적으로 사용할 수 있도록 API(360)를 통해 다양한 기능들을 어플리케이션(370)으로 제공할 수 있다. 한 실시예에 따르면, 미들웨어(330)(예: 미들웨어(143))는 런타임 라이브러리(335), 어플리케이션 매니저(application manager)(341), 윈도우 매니저(window manager)(342), 멀티미디어 매니저(multimedia manager)(343), 리소스 매니저(resource manager)(344), 파워 매니저(power manager)(345), 데이터베이스 매니저(database manager)(346), 패키지 매니저(package manager)(347), 연결 매니저(connectivity manager)(348), 통지 매니저(notification manager)(349), 위치 매니저(location manager)(350), 그래픽 매니저(graphic manager)(351), 또는 보안 매니저(security manager)(352) 중 적어도 하나를 포함할 수 있다.

[0063] 런타임 라이브러리(335)는, 예를 들면, 어플리케이션(370)이 실행되는 동안에 프로그래밍 언어를 통해 새로운 기능을 추가하기 위해 컴파일러가 사용하는 라이브러리 모듈을 포함할 수 있다. 런타임 라이브러리(335)는 입출력 관리, 메모리 관리, 또는 산술 함수에 대한 기능 등을 수행할 수 있다.

[0064] 어플리케이션 매니저(341)는, 예를 들면, 어플리케이션(370) 중 적어도 하나의 어플리케이션의 생명 주기(life cycle)를 관리할 수 있다. 윈도우 매니저(342)는 화면에서 사용하는 GUI 자원을 관리할 수 있다. 멀티미디어 매니저(343)는 다양한 미디어 파일들의 재생에 필요한 포맷을 파악하고, 해당 포맷에 맞는 코덱(codec)을 이용하여 미디어 파일의 인코딩(encoding) 또는 디코딩(decoding)을 수행할 수 있다. 리소스 매니저(344)는 어플리케이션(370) 중 적어도 어느 하나의 어플리케이션의 소스 코드, 메모리 또는 저장 공간 등의 자원을 관리할 수 있다.

[0065] 파워 매니저(345)는, 예를 들면, 바이오스(BIOS: basic input/output system) 등과 함께 동작하여 배터리(battery) 또는 전원을 관리하고, 전자 장치의 동작에 필요한 전력 정보 등을 제공할 수 있다. 데이터베이스 매니저(346)는 어플리케이션(370) 중 적어도 하나의 어플리케이션에서 사용할 데이터베이스를 생성, 검색, 또는 변경할 수 있다. 패키지 매니저(347)는 패키지 파일의 형태로 배포되는 어플리케이션의 설치 또는 업데이트를 관리할 수 있다.

[0066] 연결 매니저(348)는, 예를 들면, WiFi 또는 블루투스 등의 무선 연결을 관리할 수 있다. 통지 매니저(349)는 도착 메시지, 약속, 근접성 알림 등의 사건(event)을 사용자에게 방해되지 않는 방식으로 표시 또는 통지할 수 있다. 위치 매니저(350)는 전자 장치의 위치 정보를 관리할 수 있다. 그래픽 매니저(351)는 사용자에게 제공될 그래픽 효과 또는 이와 관련된 사용자 인터페이스를 관리할 수 있다. 보안 매니저(352)는 시스템 보안 또는 사용자 인증 등에 필요한 제반 보안 기능을 제공할 수 있다. 한 실시예에 따르면, 전자 장치(예: 전자 장치(101))가 전화 기능을 포함할 경우, 미들웨어(330)는 전자 장치의 음성 또는 영상 통화 기능을 관리하기 위한 통화 매니저(telephony manager)를 더 포함할 수 있다.

[0067] 미들웨어(330)는 전문화된 구성요소들의 다양한 기능의 조합을 형성하는 미들웨어 모듈을 포함할 수 있다. 미들웨어(330)는 차별화된 기능을 제공하기 위해 운영 체제의 종류 별로 특화된 모듈을 제공할 수 있다. 또한, 미들웨어(330)는 동적으로 기존의 구성요소를 일부 삭제하거나 새로운 구성요소들을 추가할 수 있다.

[0068] API(360)(예: API(145))는, 예를 들면, API 프로그래밍 함수들의 집합으로, 운영 체제에 따라 다른 구성으로 제공될 수 있다. 예를 들면, 안드로이드 또는 iOS의 경우, 플랫폼 별로 하나의 API 셋을 제공할 수 있으며, 타이젠(tizen)의 경우, 플랫폼 별로 두 개 이상의 API 셋을 제공할 수 있다.

[0069] 어플리케이션(370)(예: 어플리케이션 프로그램(147))은, 예를 들면, 홈(371), 다이얼러(372), SMS/MMS(373), IM(instant message)(374), 브라우저(375), 카메라(376), 알람(377), 컨택트(378), 음성 다이얼(379), 이메일(380), 달력(381), 미디어 플레이어(382), 앨범(383), 또는 시계(384), 건강 관리(health care)(예: 운동량 또는 혈당 등을 측정), 또는 환경 정보 제공(예: 기압, 습도, 또는 온도 정보 등을 제공) 등의 기능을 수행할 수 있는 하나 이상의 어플리케이션을 포함할 수 있다.

[0070] 한 실시예에 따르면, 어플리케이션(370)은 전자 장치(예: 전자 장치(101))와 외부 전자 장치(예: 전자 장치(102, 104)) 사이의 정보 교환을 지원하는 어플리케이션(이하, 설명의 편의 상, "정보 교환 어플리케이션")을 포함할 수 있다. 정보 교환 어플리케이션은, 예를 들면, 외부 전자 장치에 특정 정보를 전달하기 위한 알림 전달(notification relay) 어플리케이션, 또는 외부 전자 장치를 관리하기 위한 장치 관리(device management) 어플리케이션을 포함할 수 있다.

[0071] 예를 들면, 알림 전달 어플리케이션은 전자 장치의 다른 어플리케이션(예: SMS/MMS 어플리케이션, 이메일 어플리케이션, 건강 관리 어플리케이션, 또는 환경 정보 어플리케이션 등)에서 발생된 알림 정보를 외부 전자 장치

(예: 전자 장치(102, 104))로 전달하는 기능을 포함할 수 있다. 또한, 알림 전달 어플리케이션은, 예를 들면, 외부 전자 장치로부터 알림 정보를 수신하여 사용자에게 제공할 수 있다.

- [0072] 장치 관리 어플리케이션은, 예를 들면, 전자 장치와 통신하는 외부 전자 장치(예: 전자 장치(102, 104))의 적어도 하나의 기능(예: 외부 전자 장치 자체(또는, 일부 구성 부품)의 턴-온/턴-오프 또는 디스플레이의 밝기(또는, 해상도) 조절), 외부 전자 장치에서 동작하는 어플리케이션 또는 외부 전자 장치에서 제공되는 서비스(예: 통화 서비스 또는 메시지 서비스 등)를 관리(예: 설치, 삭제, 또는 업데이트)할 수 있다.
- [0073] 한 실시예에 따르면, 어플리케이션(370)은 외부 전자 장치(예: 전자 장치(102, 104))의 속성(에 따라 지정된 어플리케이션(예: 모바일 의료 기기의 건강 관리 어플리케이션 등)을 포함할 수 있다. 한 실시예에 따르면, 어플리케이션(370)은 외부 전자 장치(예: 서버(106) 또는 전자 장치(102, 104))로부터 수신된 어플리케이션을 포함할 수 있다. 한 실시예에 따르면, 어플리케이션(370)은 프리로드 어플리케이션(preloaded application) 또는 서버로부터 다운로드 가능한 제3자 어플리케이션(third party application)을 포함할 수 있다. 도시된 실시예에 따른 프로그램 모듈(310)의 구성요소들의 명칭은 운영 체제의 종류에 따라서 달라질 수 있다.
- [0074] 다양한 실시예에 따르면, 프로그램 모듈(310)의 적어도 일부는 소프트웨어, 펌웨어, 하드웨어, 또는 이들 중 적어도 둘 이상의 조합으로 구현될 수 있다. 프로그램 모듈(310)의 적어도 일부는, 예를 들면, 프로세서(예: 프로세서(210))에 의해 구현(implement)(예: 실행)될 수 있다. 프로그램 모듈(310)의 적어도 일부는 하나 이상의 기능을 수행하기 위한, 예를 들면, 모듈, 프로그램, 루틴, 명령어 세트(sets of instructions) 또는 프로세스 등을 포함할 수 있다.
- [0075] 도 4는 본 발명의 다양한 실시예들에 따른 전자 장치(201)가 텍스트가 변환될 음향과 관련된 정보를 선택하고, 선택된 정보에 기초하여 음향 신호를 발생시키는 동작을 나타내는 순서도이다.
- [0076] 전자 장치(201)는 401 동작에서, 적어도 하나의 텍스트를 획득할 수 있다. 전자 장치(201)는 입력 장치(250)를 통하여 사용자로부터 적어도 하나의 텍스트를 획득할 수 있고, 또는 외부 장치로부터 적어도 하나의 텍스트를 포함하는 문자 메시지를 수신할 수 있다.
- [0077] 전자 장치(201)는 403 동작에서, 획득된 텍스트가 변환될 음향과 관련된 정보를 선택할 수 있다. 음향과 관련된 정보는 음향의 언어 정보 또는 음향의 화자 정보를 포함할 수 있다. 예를 들어, 음향의 언어 정보는 한국어, 영어, 프랑스어 등과 같이 음향 데이터 셋이 어떤 국가의 언어로 구성되어 있는지에 관한 정보를 포함할 수 있고, 음향의 화자 정보는 남성 화자, 여성 화자, 각 나이대별 화자, 각 지역별 화자(다양한 사투리를 쓰는 화자) 등과 같이 음향 데이터 셋이 어떤 화자의 말투로 구성되어 있는지에 관한 정보를 포함할 수 있다. 전자 장치(201)는 음향과 관련된 정보를 선택하기 위해 사용자로부터 음향과 관련된 정보를 수신할 수 있고, 또는 상기 전자 장치(201)는 획득된 텍스트를 분석하여 음향과 관련된 정보를 판단할 수 있다. 예를 들어, 전자 장치(201)는 획득된 텍스트가 변환될 음향이 한국어로 재생될 것인지 남성 목소리로 재생될 것인지에 관한 선택을 사용자로부터 수신할 수 있고, 또는 텍스트를 분석하여 상기 텍스트가 어떤 국가의 언어로 구성되어 있는지 판단할 수 있다. 다양한 실시예에 따라, 403 동작은 텍스트를 획득하기 전, 즉, 401 동작 이전에 사용자가 선택할 수 있다. 다양한 실시예에 따라, 선택된 정보는 메모리(230)에 저장될 수 있다.
- [0078] 전자 장치(201)는 405 동작에서, 선택된 정보를 확인할 수 있다. 전자 장치(201)는 상기 선택된 정보가 제 1 정보인지 또는 제 2 정보인지를 판단할 수 있다. 전자 장치(201)는 선택된 정보에 대응하는 결정 트리(decision tree)를 확인할 수 있다. 전자 장치(201)는 외부 장치(예를 들어, 공통 음향 데이터 제공 서버)로부터 상기 결정 트리에 관한 데이터를 수신하여 메모리(230)에 저장할 수 있다. 상기 결정 트리는 복수의 경로들로 구성될 수 있고 각 경로들의 가장 마지막 부분(leaf node)에는 공통 음향 데이터 셋 중 특정 음향 데이터를 지시하는 인덱스 정보가 포함될 수 있다. 예를 들어 도 5를 참조하면, 제1 결정 트리는(510)은 여성 목소리의 영어의 언어 처리 결과를 나타내는 복수의 경로들로 구성될 수 있고, 각 경로들의 가장 마지막 부분에는 음소 단위의 음향 데이터(예를 들어 여성 목소리 "g"에 해당하는 음향 데이터)를 지시하는 인덱스 정보가 포함될 수 있다. 본 발명의 다양한 실시예들에 따른 결정 트리에 포함된 인덱스 정보들은 음소 단위의 음향 데이터를 지시할 수 있고, 또는 상기 음소 단위의 음향 데이터를 소정 시간 간격으로 분할한 세분 음소 단위의 음향 데이터를 지시할 수 있다.
- [0079] 전자 장치(201)는 407 동작에서, 텍스트가 변환될 음향과 관련된 정보가 제 1 정보인 경우, 복수의 제 1 경로들 중 적어도 하나를 선택할 수 있다. 제 1 정보는 음향의 언어 정보 또는 음향의 화자 정보 중 적어도 하나를 포함할 수 있다. 예를 들어 도 5를 참조하면, 선택된 정보가 여성 목소리의 영어이고, 획득한 텍스트가 "go"이고,

선택된 정보에 대응하는 제 1 결정 트리(510)가 여성 목소리의 영어에 대한 음향 데이터들을 지시하는 인덱스 정보들로 구성된 경우, 전자 장치(201)는 상기 획득한 텍스트를 음향 신호로 변환하기 위하여 제 1 결정 트리(510)에 포함된 여성 목소리 "g"에 대한 경로(예를 들어, 인덱스 A4까지의 경로)와, 제1 결정 트리(510)에 포함된 여성 목소리 "o"에 대한 경로(예를 들어, 인덱스 An-1까지의 경로)를 선택할 수 있다. 결정 트리의 적어도 하나의 인덱스는 공통 음향 데이터 셋을 구성하는 적어도 하나의 음향 데이터를 지시(indicate)할 수 있다. 본 발명의 다양한 실시예들에 따른 복수의 제 1 경로들은 공통 음향 데이터 셋 중의 일부를 지시할 수 있다. 예를 들어 도 5를 참조하면, 제 1 결정 트리(510)의 하나의 경로(인덱스 A1까지의 경로)는 공통 음향 데이터 셋(500)의 음향 데이터(S2)를 지시할 수 있고, 다른 하나의 경로(인덱스 A2까지의 경로)는 공통 음향 데이터 셋(500)의 음향 데이터(S3)를 지시할 수 있다. 공통 음향 데이터 셋(SCCAD, Super-Clustered Common Acoustic Data)은 적어도 하나의 음향 데이터 셋에 기반하여 생성될 수 있다. 공통 음향 데이터 셋의 생성에 관한 내용은 하기의 도 6에서 후술하도록 한다.

[0080] 전자 장치(201)는 409 동작에서, 선택된 적어도 하나의 제 1 경로에 기초하여 제 1 음향 신호를 발생시킬 수 있다. 전자 장치(201)는 선택된 적어도 하나의 제 1 경로에 기초하여 공통 음향 데이터 셋 중의 일부를 로드하고, 상기 로드된 일부에 기초하여 제 1 음향 신호를 발생시킬 수 있다. 공통 음향 데이터 셋 중의 일부는 음향의 특정 화자 정보 또는 특정 언어 정보에 대응하는 음향 데이터들의 집합일 수 있다. 전자 장치(201)는 입력된 텍스트에 기초하여 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택할 수 있고, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여 제 1 음향 신호를 발생시킬 수 있다. 공통 음향 데이터 셋 중의 일부 중 적어도 일부는 음향 신호의 요소에 대응하는 음향 데이터를 나타내는 것으로서, 음향 신호의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응할 수 있다. 예를 들어 도 5를 참조하면, 전자 장치(201)가 획득한 텍스트인 "go"를 음향 신호로 변환하기 위하여 전자 장치(201)는 제 1 결정 트리(510)에 포함된 "g"에 대한 경로(인덱스 A4까지의 경로)와, 제 1 결정 트리(510)에 포함된 "o"에 대한 경로(인덱스 An-1까지의 경로)를 선택할 수 있고, 선택된 적어도 하나의 제 1 경로에 대응하는 적어도 하나의 음향 데이터(선택된 인덱스가 지시하는 음향 데이터)를 공통 음향 데이터 셋으로부터 선택할 수 있다. 전자 장치(201)는 상기 공통 음향 데이터 셋 중 선택된 적어도 하나의 음향 데이터를 로드하고, 상기 로드된 음향 데이터에 기초하여 제 1 음향 신호를 발생시킬 수 있다. 전자 장치(201)는 스피커(282)를 통하여 상기 제 1 음향 신호를 출력할 수 있다. 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 입력된 텍스트 문장을 음소 단위로 분석하거나, 상기 음소를 분할한 세분 음소 단위로 분석할 수 있다. 전자 장치(201)는 각 음소 또는 각 세분 음소 단위에 대한 음향 데이터를 선택할 수 있고, 상기 선택된 음향 데이터들을 합성하여 텍스트 전체에 대한 합성음을 생성할 수 있다. 전자 장치(201)는 스피커(282)를 통하여 텍스트 전체에 대한 합성음을 출력할 수 있다.

[0081] 전자 장치(201)는 411 동작에서, 텍스트가 변환될 음향과 관련된 정보가 제 2 정보인 경우, 복수의 제 2 경로들 중 적어도 하나를 선택할 수 있다. 제 2 정보는 제 1정보와 상이한 정보로서, 음향의 언어 정보 또는 음향의 화자 정보 중 적어도 하나를 포함할 수 있다. 예를 들어 도 5를 참조하면, 선택된 정보가 남성 목소리의 한국어에 대한 정보이고, 선택된 정보에 대응하는 제 2 결정 트리(520)가 존재하는 경우, 결정 트리의 적어도 하나의 인덱스는 공통 음향 데이터 셋을 구성하는 적어도 하나의 음향 데이터를 지시할 수 있다. 본 발명의 다양한 실시예들에 따른 복수의 제 2 경로들은 공통 음향 데이터 셋 중의 일부를 지시할 수 있다. 예를 들어 도 5를 참조하면, 제 2 결정 트리(520)의 하나의 경로(인덱스 B1까지의 경로)는 공통 음향 데이터 셋(500)의 음향 데이터(S4)를 지시할 수 있고, 다른 하나의 경로(인덱스 B2까지의 경로)는 공통 음향 데이터 셋(500)의 음향 데이터(S5)를 지시할 수 있다.

[0082] 전자 장치(201)는 413 동작에서, 선택된 적어도 하나의 제 2 경로에 기초하여 제 2 음향 신호를 발생시킬 수 있다. 전자 장치(201)는 선택된 적어도 하나의 제 2 경로에 기초하여 공통 음향 데이터 셋 중의 일부(409 동작에서 제 1 경로에 기초하여 로드된 음향 데이터) 또는 다른 일부를 로드하고, 상기 로드된 일부 또는 다른 일부에 기초하여 제 2 음향 신호를 발생시킬 수 있다. 예를 들어 도 5를 참조하면, 제 1 결정 트리(510)의 하나의 경로(인덱스 A4까지의 경로) 및 제 2 결정 트리(520)의 하나의 경로(인덱스 B2까지의 경로)는 동일한 음향 데이터(S5)를 지시할 수 있다. 공통 음향 데이터 셋 중의 일부 또는 다른 일부는 음향의 특정 화자 정보 또는 특정 언어 정보에 대응하는 음향 데이터들의 집합일 수 있다. 전자 장치(201)는 입력된 텍스트에 기초하여 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택할 수 있고, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여 제 2 음향 신호를 발생시킬 수 있다. 공통 음향 데이터 셋 중의 일부 중 적어도 일부는 음향 신호의 요소에 대응하는 음향 데이터를 나타내는 것으로서, 음향 신호의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응할 수 있다. 전자 장치(201)는 상기 공통 음향 데이터 셋 중 선택된 적어도 하나의 음향 데이터를 로드하고, 상기 로드된 음향 데이터에 기초하여 제 2 음향 신호를 발생시킬 수

있다. 전자 장치(201)는 스피커(282)를 통하여 상기 제 2 음향 신호를 출력할 수 있다. 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 입력된 텍스트 문장을 음소 단위로 분석하거나, 상기 음소를 분할한 세분 음소 단위로 분석할 수 있다. 전자 장치(201)는 각 음소 또는 각 세분 음소 단위에 대한 음향 데이터를 선택할 수 있고, 상기 선택된 음향 데이터들을 합성하여 텍스트 전체에 대한 합성음을 생성할 수 있다. 전자 장치(201)는 스피커(282)를 통하여 텍스트 전체에 대한 합성음을 출력할 수 있다.

[0083] 도 6은 본 발명의 다양한 실시예들에 따른 전자 장치(201)가 공통 음향 데이터를 생성하는 동작을 나타내는 순서도이다.

[0084] 전자 장치(201)는 601 동작에서, 음향과 관련된 제 1 정보에 대응하는 제 1 음향 데이터 셋, 및 상기 음향과 관련된 제 2 정보에 대응하는 제 2 음향 데이터 셋을 획득할 수 있다. 제 1 정보 또는 제 2 정보는 상기 음향의 언어 정보 또는 화자 정보를 포함할 수 있다. 예를 들어 도 7a를 참조하면, 전자 장치(201)는 여성 목소리의 영어(제 1 정보)에 대응하는 음향 데이터들의 집합인 제 1 음향 데이터 셋(710), 및 남성 목소리의 한국어(제 2 정보)에 대응하는 음향 데이터들의 집합인 제 2 음향 데이터 셋(720)을 획득할 수 있다.

[0085] 601 동작에서는 제 1 음향 데이터 셋과 제 2 음향 데이터 셋으로 공통음향 데이터를 구성하는 방법을 설명하고 있으나, 그 이상의 음향 데이터 셋을 획득할 수 있다. 복수의 음향 데이터 셋을 획득하고, 상기 복수의 음향 데이터 셋에 대하여 603 이하의 과정을 수행할 수 있다.

[0086] 전자 장치(201)는 603 동작에서, 제 1 음향 데이터 셋의 적어도 일부, 및/또는 제 2 음향 데이터 셋의 적어도 일부의 유사도를 판단할 수 있다. 전자 장치(201)는 음향 데이터 셋의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나의 유사도를 판단할 수 있다. 예를 들어, 전자 장치(201)는 유사도 판단을 위하여 벡터양자화(Vector Quantization)를 통하여 음향 데이터 셋의 적어도 일부에 해당하는 음향 데이터를 벡터화할 수 있다. 전자 장치(201)는 음향 신호의 스펙트럼, 피치, 또는 노이즈 중 적어도 하나를 벡터화할 수 있고, 벡터화된 값을 이용하여 상기 유사도를 판단할 수 있다. 예를 들어 도 7a를 참조하면, 전자 장치(201)는 제 1 음향 데이터 셋(710)의 적어도 일부 및/또는 제 2 음향 데이터 셋(720)의 적어도 일부를 수집한 전체 음향 데이터 셋(701)을 획득할 수 있다. 전자 장치(201)는 전체 음향 데이터 셋(701)의 음향 데이터 A2(711), 및 전체 음향 데이터 셋(701)의 음향 데이터 B3(721)의 유사도를 판단할 수 있다. 전자 장치(201)는 유사도 판단을 위하여 음향 데이터 A2(711)의 스펙트럼(712)을 벡터화하여 벡터 값(713)을 획득하고 음향 데이터 B3(721)의 스펙트럼(722)을 벡터화하여 벡터 값(723)을 획득할 수 있다. 전자 장치(201)는 A2의 음성 벡터 값(521), 및 B3의 음성 벡터 값(522)을 비교하여 음향 데이터의 유사도를 판단할 수 있다. 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 K-means 알고리즘, Fuzzy 알고리즘, GMM 알고리즘, Lloyd 알고리즘 등을 수행하여 제 1 음향 데이터 셋의 적어도 일부 및/또는 제 2 음향 데이터 셋의 적어도 일부의 유사도를 판단할 수 있다. 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 제 1 음향 데이터 셋(710)의 적어도 일부 및 제 2 음향 데이터 셋(720)의 적어도 일부를 수집한 전체 음향 데이터 셋(701)을 획득할 수 있고, (1) 전체 음향 데이터 셋(701) 중 제 1 음향 데이터 셋(710)의 음향 데이터와 제 2 음향 데이터 셋(720)의 음향 데이터간의 유사도를 판단하거나 (2) 전체 음향 데이터 셋(701) 중 제 1 음향 데이터 셋(710)의 음향 데이터간의 유사도를 판단하거나, 또는 (3) 전체 음향 데이터 셋(701) 중 제 2 음향 데이터 셋(720)의 음향 데이터간의 유사도를 판단할 수 있다.

[0087] 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 적어도 하나의 음향 데이터 셋을 수집한 전체 음향 데이터 셋을 획득할 수 있고, 상기 전체 음향 데이터 셋을 복수개의 음향 데이터들을 포함하는 소정 개수의 군집으로 구분할 수 있다. 예를 들어 도 7b <730>을 참조하면, 전자 장치(201)는 적어도 하나의 음향 데이터 셋을 수집한 전체 음향 데이터 셋(710)으로부터 대표 음향 데이터(731, 732, 733)를 랜덤하게 선택할 수 있다. <740>을 참조하면, 전자 장치(201)는 각 음향 데이터들에 대한 대표 음향 데이터(731, 732, 733)의 평균 거리를 기준으로 군집(741, 742, 743)을 구분할 수 있다. <750>을 참조하면, 전자 장치(201)는 각 음향 데이터들과 대표 음향 데이터(731, 732, 733)들의 유사도를 판단하여 유사도가 높은 대표 음향 데이터로 각 음향 데이터를 구분할 수 있다. <760>을 참조하면, 전자 장치(201)는 구분된 음향 데이터들로 군집을 재조정할 수 있다. 전자 장치(201)는 <730> 내지 <760> 과정을 반복하여 유사도가 높은 음향 데이터들끼리 군집을 형성하는 클러스터링(clustering) 알고리즘을 수행할 수 있다. 전자 장치(201)는 605 동작에서, 상기 유사도 판단에 기초하여 제 1 음향 데이터 셋의 일부, 및 제 2 음향 데이터 셋의 적어도 일부와 관련된 공통 음향 데이터 셋을 생성할 수 있다. 전자 장치(201)는 상기 유사도가 선택된 쓰레쉬 홀드(threshold) 값 이상인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부, 및 상기 제 2 음향 데이터 셋의 적어도 일부 모두에 대응하는 제 1 파라미터를 결정할 수 있고, 상기 유사도가 상기 쓰레쉬 홀드 값 미만인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부에 대응하는 제 2 파라미터, 및 상기 제 2 음향 데이터 셋의 적어도 일부에 대응하는 제 3 파라미터를 결정할 수 있다. 제 1 파

라미터, 제 2 파라미터, 또는 제 3 파라미터는 상기 음향의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응할 수 있다. 예를 들어 도 7a를 참조하면, 전체 음향 데이터 셋(701)의 음향 데이터 A2(711)의 스펙트럼(712), 및 전체 음향 데이터 셋(720)의 음향 데이터 B3(721)의 스펙트럼(722)의 유사도가 쓰레쉬 홀드 값 이상인 경우, 전자 장치(201)는 상기 음향 데이터 A2(711)의 스펙트럼(712), 및 음향 데이터 B3(721)의 스펙트럼(722) 모두에 대응하는 음향 데이터 S1(530a)의 스펙트럼을 생성할 수 있다. 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 전체 음향 데이터 셋(701)의 음향 데이터 A2(711)의 스펙트럼(712), 및 전체 음향 데이터 셋(720)의 음향 데이터 B3(721)의 스펙트럼(722)의 유사도가 쓰레쉬 홀드 값 이상인 경우, 상기 음향 데이터 A2(711)의 스펙트럼(712) 또는 음향 데이터 B3(721)의 스펙트럼(722) 중 하나를 공통 음향 데이터 셋(500)의 음향 데이터 S1(501)로 결정할 수 있다.

[0088] 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 전체 음향 데이터 셋(701)의 음향 데이터 A2(711)의 스펙트럼, 및 전체 음향 데이터 셋(721)의 음향 데이터 B3(721)의 스펙트럼의 유사도가 쓰레쉬 홀드 값 미만인 경우, 상기 음향 데이터 A2(711)의 스펙트럼에 대응하는 음향 데이터 S2(502)의 스펙트럼을 생성할 수 있고, 음향 데이터 B3(721)의 스펙트럼에 대응하는 음향 데이터 S3(503)의 스펙트럼을 생성할 수 있다. 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 전체 음향 데이터 셋(701)의 음향 데이터 A2(711)의 스펙트럼, 및 전체 음향 데이터 셋(721)의 음향 데이터 B3(721)의 스펙트럼의 유사도가 쓰레쉬 홀드 값 미만인 경우, 상기 음향 데이터 A2(711)의 스펙트럼을 음향 데이터 S2(502)의 스펙트럼으로 결정할 수 있고, 음향 데이터 B3(721)의 스펙트럼을 음향 데이터 S3(503)의 스펙트럼으로 결정할 수 있다. 본 발명의 다양한 실시예들에 따른 전자 장치(201)는 공통 음향 데이터 셋의 음향 데이터들 간에 음질 저하를 일으키지 않는 수준의 쓰레쉬 홀드 값을 설정할 수 있고, 상기 쓰레쉬 홀드 값에 기초하여 공통 음향 데이터 셋의 음향 데이터들을 클러스터링(clustering)할 수 있다. 전자 장치(201)는 K-means 알고리즘, Fuzzy 알고리즘, GMM 알고리즘, Lloyd 알고리즘 등을 수행하여 쓰레쉬 홀드 값 이상의 유사도를 갖는 음향 데이터들을 판단할 수 있고, 상기 음향 데이터들을 대표하는 공통 음향 데이터를 결정할 수 있다. 전자 장치(201)는 쓰레쉬 홀드 값 미만의 유사도를 갖는 음향 데이터들을 판단할 수 있고, 상기 음향 데이터들 각각에 대응하는 공통 음향 데이터들을 결정할 수 있다.

[0089] 도 8은 본 발명의 다양한 실시예들에 따른 전자 장치(201)가 공통 음향 데이터 셋을 생성하고 특정 음향 데이터 셋의 복수의 경로들을 상기 공통 음향 데이터 셋에 매칭하는 동작을 나타내는 도면이다.

[0090] 도 8을 참조하면, 전자 장치(201)는 적어도 하나의 음향 데이터 셋을 이용하여 공통 음향 데이터 셋(500)(SCCAD, Super-Clustered Common Acoustic Data)을 생성할 수 있다. 전자 장치(201)는 각 음향 데이터 셋을 수집한 전체 음향 데이터 셋의 음향 데이터들의 유사도를 판단할 수 있다. 상기 음향 데이터들의 유사도 판단은 음향의 스펙트럼, 피치, 또는 노이즈 중 적어도 하나를 비교하여 수행될 수 있다. 전자 장치(201)는 음향 데이터들의 유사도가 선택된 쓰레쉬 홀드 값 이상인 경우, 전자 장치(201)는 상기 음향 데이터들의 모두에 대응하는 파라미터를 결정하고, 유사도가 상기 쓰레쉬 홀드 값 미만인 경우, 전자 장치(201)는 각 음향 데이터에 대응하는 파라미터를 결정할 수 있다. 예를 들어 도 7a를 참조하면, 전자 장치(201)는 전체 음향 데이터 셋(701)의 음향 데이터 A3와 전체 음향 데이터 셋(701)의 음향 데이터 B2의 유사도를 판단하여 상기 유사도가 쓰레쉬 홀드 값 이상인 경우, 전자 장치(201)는 음향 데이터 A3 및 음향 데이터 B2 모두에 대응하는 제 1 파라미터를 결정할 수 있고, 상기 유사도가 쓰레쉬 홀드 값 미만인 경우, 음향 데이터 A3에 대응하는 제 2 파라미터 및 음향 데이터 B2에 대응하는 제 3 파라미터를 결정할 수 있다. 전자 장치(201)는 제 1 파라미터, 제 2 파라미터, 또는 제 3 파라미터에 기초하여 공통 음향 데이터 셋(500)의 음향 데이터를 생성할 수 있다.

[0091] 전자 장치(201)는 기존의 음향 모델 이외에 새로운 음향 모델을 추가적으로 획득할 수 있고, 새로이 획득한 상기 음향 모델은 결정 트리 및 상기 결정 트리와 매칭되는 음향 데이터 셋을 포함할 수 있다. 전자 장치(201)는 새로운 음향 모델을 획득하면, 상기 음향 모델의 결정 트리를 공통 음향 데이터 셋에 새로이 매칭시킬 수 있다. 예를 들어 도 8을 참조하면, 전자 장치(201)가 P 결정 트리(726) 및 P 음향 데이터를 포함하는 P 음향 모델을 획득하고, 상기 P 결정 트리(726)는 복수의 경로들(인덱스 P1, P2, P3, P4까지의 경로)로 구성된 경우, 전자 장치(201)는 상기 P 결정 트리(726)의 인덱스 P1(801)이 지시하는 P 음향 데이터 셋의 음향 데이터를 확인할 수 있다. 전자 장치(201)는 상기 P1(801)이 원래 지시하던 음향 데이터와 유사도가 가장 높은 음향 데이터를 공통 음향 데이터 셋(500)에서 검색할 수 있고, 상기 P 결정 트리(726)의 인덱스 P1(801)을 공통 음향 데이터의 음향 데이터를 지시하는 인덱스 S8(811)로 교체할 수 있다. 마찬가지로, 전자 장치(201)는 상기 P 결정 트리(726)의 인덱스 P2(802)를 공통 음향 데이터의 음향 데이터를 지시하는 인덱스 S21(812)로 교체할 수 있고, 상기 P 결정 트리(726)의 인덱스 P3(803)를 공통 음향 데이터의 음향 데이터를 지시하는 인덱스 S3(813)로 교체할 수 있고, 상기 P 결정 트리(726)의 인덱스 P4(804)를 공통 음향 데이터의 음향 데이터를 지시하는 인덱스 S30(814)로 교

체할 수 있다. P 결정 트리(726)의 각각의 인덱스들은 원래 지시하던 음향 데이터와 유사도가 가장 높은 음향 데이터(공통 음향 데이터 셋의 음향 데이터)를 지시하는 인덱스들로 교체될 수 있다.

- [0092] 도 9는 본 발명의 다양한 실시예들에 따른 제 1 전자 장치의 블록도 및 제 2 전자 장치의 블록도이다.
- [0093] 도 9를 참조하면, 제 1 전자 장치(901)는 프로세서(910), 메모리(920), 입력 장치(930), 통신 모듈(940)을 포함할 수 있다. 제2 전자 장치(902)는 프로세서(950), 메모리(960), 통신 모듈(970)을 포함할 수 있다. 본 발명의 다양한 실시예들에 따른 제 1 전자 장치(901) 및 제 2 전자 장치(902)는 도 9에는 도시되지 않았지만 도 2 에 도시된 전자 장치(201)의 구성 요소들을 모두 포함할 수 있다.
- [0094] 본 발명의 다양한 실시예들에 따른 제 1 전자 장치(901)의 프로세서(910)는 도 2의 전자 장치(201)의 프로세서(210)의 기능을 수행할 수 있다. 프로세서(910)는 텍스트 분석부(911), 링커부(912), 합성음 생성부(913)를 포함할 수 있다.
- [0095] 텍스트 분석부(911)는 전자 장치(901)가 획득한 적어도 하나의 텍스트를 분석할 수 있고, 상기 획득한 텍스트가 변환될 음향과 관련된 정보를 선택할 수 있다. 예를 들어, 텍스트 분석부(911)는 텍스트를 분석하여 상기 텍스트를 한국어로 재생할 것인지 남성 목소리로 재생할 것인지에 관한 정보를 선택할 수 있다.
- [0096] 링커부(912)는 선택된 정보가 제 1 정보인지 또는 제 2 정보인지를 판단할 수 있다. 링커부(912)는 선택된 정보에 대응하는 결정 트리(decision tree)를 확인할 수 있다. 링커부(912)는 텍스트가 변환될 음향과 관련된 정보가 제 1 정보인 경우, 결정 트리에 포함된 복수의 제 1 경로들 중 적어도 하나를 선택할 수 있다. 링커부(912)는 선택된 적어도 하나의 제 1 경로에 기초하여 공통 음향 데이터 셋 중의 일부를 로드할 수 있다. 링커부(912)는 텍스트가 변환될 음향과 관련된 정보가 제 2 정보인 경우, 결정 트리에 포함된 복수의 제 2 경로들 중 적어도 하나를 선택할 수 있다. 링커부(912)는 선택된 적어도 하나의 제 2 경로에 기초하여, 공통 음향 데이터 셋 중의 일부 또는 다른 일부를 로드할 수 있다. 합성음 생성부(913)는 선택된 적어도 하나의 제 1 경로에 기초하여 제 1 음향 신호를 발생시킬 수 있다. 합성음 생성부(913)는 입력된 텍스트에 기초하여 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택할 수 있고, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여 제 1 음향 신호를 발생시킬 수 있다. 합성음 생성부(913)는 스피커(282)를 통하여 상기 제 1 음향 신호를 출력할 수 있다. 합성음 생성부(913)는 링커부(912)에 의해 선택된 복수개의 제 1 경로에 기초하여 복수개의 공통 음향 데이터들을 로드할 수 있고, 문장 단위로 음향을 출력하기 위하여 로드된 상기 음향 데이터들을 합성한 후 상기 합성된 음향 데이터들을 출력할 수 있다.
- [0097] 합성음 생성부(913)는 선택된 적어도 하나의 제 2 경로에 기초하여 제 2 음향 신호를 발생시킬 수 있다. 합성음 생성부(913)는 입력된 텍스트에 기초하여 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택할 수 있고, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여 제 2 음향 신호를 발생시킬 수 있다. 합성음 생성부(913)는 스피커(282)를 통하여 상기 제 2 음향 신호를 출력할 수 있다. 합성음 생성부(913)는 링커부(912)에 의해 선택된 복수개의 제 2 경로에 기초하여 복수개의 공통 음향 데이터들을 로드할 수 있고, 문장 단위로 음향을 출력하기 위하여 로드된 상기 음향 데이터들을 합성한 후 상기 합성된 음향 데이터들을 출력할 수 있다.
- [0098] 본 발명의 다양한 실시예들에 따른 제 1 전자 장치(901)의 메모리(920)는 실행 시에, 상기 프로세서(910)가, 적어도 하나의 텍스트를 획득하고, 상기 획득된 텍스트가 변환될 음향과 관련된 정보를 선택하고, 상기 선택된 정보가 제 1 정보인 경우, 복수의 제 1 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 1 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 일부를 로드하고, 상기 로드된 일부에 기초하여 제 1 음향 신호를 발생시키고, 상기 선택된 정보가 제 2 정보인 경우, 복수의 제 2 경로들 중 적어도 하나를 선택하고, 상기 선택된 적어도 하나의 제 2 경로에 기초하여, 상기 공통 음향 데이터 셋 중의 상기 일부 또는 다른 일부를 로드하고, 상기 로드된 일부 또는 다른 일부에 기초하여 제 2 음향 신호를 발생시키도록 하는 인스트럭션들을 저장할 수 있다.
- [0099] 본 발명의 다양한 실시예들에 따른 메모리(920)는, 실행 시에, 상기 프로세서(910)가, 사용자로부터 상기 적어도 하나의 텍스트를 획득하거나 외부 장치로부터 상기 적어도 하나의 텍스트를 포함하는 문자 메시지를 수신하도록 하는 인스트럭션들을 저장할 수 있다.
- [0100] 본 발명의 다양한 실시예들에 따른 메모리(920)는, 실행 시에, 상기 프로세서(910)가, 상기 입력된 텍스트에 기초하여, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부를 선택하고, 상기 공통 음향 데이터 셋 중의 일부 중 적어도 일부에 더 기초하여, 상기 제 1 음향 신호 또는 상기 제 2 음향 신호를 발생시키도록 하는 인스트럭

선들을 저장할 수 있다.

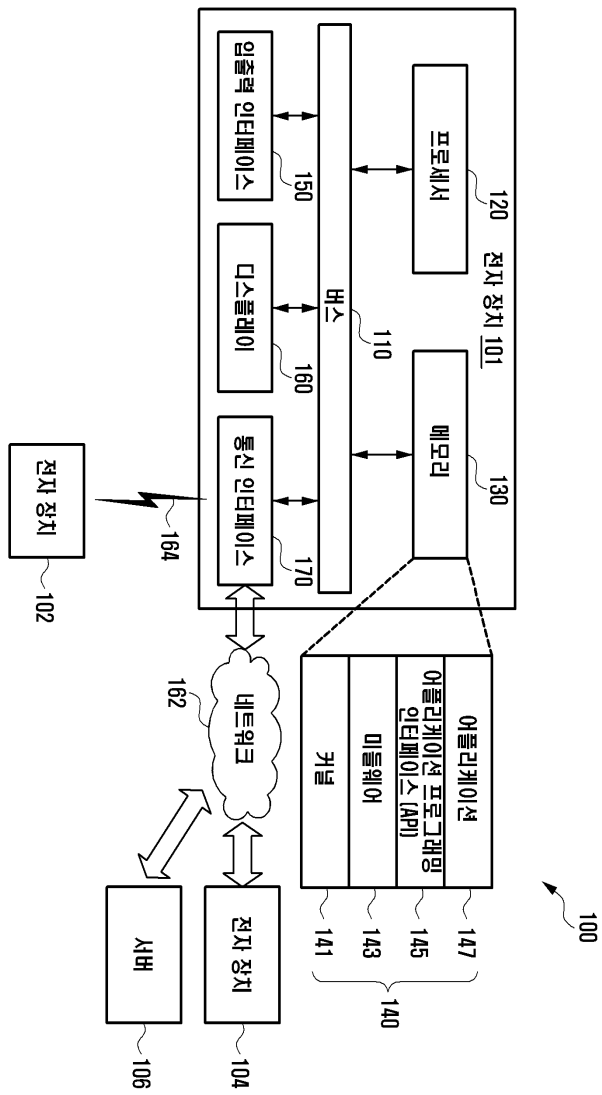
- [0101] 본 발명의 다양한 실시예들에 따른 메모리(920)는 공통 음향 데이터 셋 및 적어도 하나의 결정 트리에 관한 정보를 저장할 수 있다.
- [0102] 본 발명의 다양한 실시예들에 따른 제 1 전자 장치(901)의 입력 장치(930)는 도 2의 전자 장치(201)의 입력 장치(250)의 기능을 수행할 수 있다. 입력 장치(250)는 사용자로부터 음향으로 변환될 적어도 하나의 텍스트를 획득할 수 있다.
- [0103] 본 발명의 다양한 실시예들에 따른 제 1 전자 장치(901)의 통신 모듈(940)은 도 2의 전자 장치(201)의 통신 모듈(220)의 기능을 수행할 수 있다. 통신 모듈(940)은 결정 트리에 관한 정보 및/ 또는 공통 음향 데이터 셋에 관한 정보를 요청하는 요청 메시지를 제 2 전자 장치(902)로 전송할 수 있고, 제 2 전자 장치(902)로부터 결정 트리에 관한 정보 및/ 또는 공통 음향 데이터 셋에 관한 정보를 수신할 수 있다.
- [0104] 본 발명의 다양한 실시예들에 따른 제 2 전자 장치(902)는 공통 음향 데이터 셋을 생성하고, 상기 공통 음향 데이터 셋을 제공하는 서버 역할을 수행할 수 있다.
- [0105] 본 발명의 다양한 실시예들에 따른 제 2 전자 장치(902)의 프로세서(950)는 도 2의 전자 장치(201)의 프로세서(210)의 기능을 수행할 수 있다. 프로세서(950)는 공통 음향 데이터 셋 생성부(951), 인덱스 매칭부(952)를 포함할 수 있다.
- [0106] 본 발명의 다양한 실시예들에 따른 공통 음향 데이터 셋 생성부(951)는 음향과 관련된 제 1 정보에 대응하는 제 1 음향 데이터 셋, 및 상기 음향과 관련된 제 2 정보에 대응하는 제 2 음향 데이터 셋을 획득할 수 있다. 공통 음향 데이터 셋 생성부(951)는 제 1 음향 데이터 셋 및 제 2 음향 데이터 셋 뿐만 아니라 복수개의 음향 데이터 셋을 획득하여 이하의 동작을 수행할 수 있다. 공통 음향 데이터 셋 생성부(951)는 제 1 음향 데이터 셋의 적어도 일부, 및/ 또는 제 2 음향 데이터 셋의 적어도 일부의 유사도를 판단할 수 있다. 공통 음향 데이터 셋 생성부(951)는 상기 유사도 판단에 기초하여 제 1 음향 데이터 셋의 일부, 및 제 2 음향 데이터 셋의 적어도 일부와 관련된 공통 음향 데이터 셋을 생성할 수 있다. 공통 음향 데이터 셋 생성부(951)는 상기 유사도가 선택된 쓰레쉬 홀드(threshold) 값 이상인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부, 및 상기 제 2 음향 데이터 셋의 적어도 일부 모두에 대응하는 제 1 파라미터를 결정할 수 있고, 상기 유사도가 상기 쓰레쉬 홀드 값 미만인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부에 대응하는 제 2 파라미터, 및 상기 제 2 음향 데이터 셋의 적어도 일부에 대응하는 제 3 파라미터를 결정할 수 있다. 제 1 파라미터, 제 2 파라미터, 또는 제 3 파라미터는 상기 음향의 적어도 일부의 스펙트럼(spectrum), 피치(pitch), 또는 노이즈 중 적어도 하나에 대응할 수 있다.
- [0107] 본 발명의 다양한 실시예들에 따른 인덱스 매칭부(952)는 새로운 음향 모델을 획득하면, 상기 음향 모델의 결정 트리를 공통 음향 데이터 셋에 새로이 매칭시킬 수 있다. 새로이 획득한 음향 모델은 결정 트리 및 상기 결정 트리가 지시하는 음향 데이터 셋을 포함할 수 있다. 인덱스 매칭부(952)는 새로이 획득한 음향 모델에 포함된 음향 데이터 셋과 공통 음향 데이터 셋의 유사도를 판단할 수 있고, 새로이 획득한 음향 모델의 결정 트리가 공통 음향 데이터 셋의 데이터들(상기 새로이 획득한 음향 데이터 셋과 유사도가 가장 높은 데이터들)을 지시하도록 인덱스를 교체할 수 있다.
- [0108] 본 발명의 다양한 실시예들에 따른 제 2 전자 장치(902)의 메모리(960)는 도 2의 전자 장치(201)의 메모리(230)의 기능을 수행할 수 있다. 메모리(960)는, 실행 시에, 상기 프로세서(950)가, 음향과 관련된 제 1 정보에 대응하는 제 1 음향 데이터 셋, 및/ 또는 상기 음향과 관련된 제 2 정보에 대응하는 제 2 음향 데이터 셋을 획득하고, 상기 제 1 음향 데이터 셋의 적어도 일부, 및/ 또는 상기 제 2 음향 데이터 셋의 적어도 일부의 유사도를 판단하고, 상기 판단에 기초하여, 상기 제 1 음향 데이터 셋의 적어도 일부, 및/ 또는 상기 제 2 음향 데이터 셋의 적어도 일부와 관련된 공통 음향 데이터 셋을 생성하도록 하는 인스트럭션들을 저장할 수 있다.
- [0109] 본 발명의 다양한 실시예들에 따른 메모리(960)는, 실행 시에, 상기 프로세서(950)가, 상기 판단에 기초하여, 상기 유사도가 선택된 쓰레쉬 홀드(threshold) 값 이상인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부, 및 상기 제 2 음향 데이터 셋의 적어도 일부 모두에 대응하는 제 1 파라미터를 결정하고, 상기 유사도가 상기 쓰레쉬 홀드 값 미만인 경우, 상기 제 1 음향 데이터 셋의 적어도 일부에 대응하는 제 2 파라미터, 및 상기 제 2 음향 데이터 셋의 적어도 일부에 대응하는 제 3 파라미터를 결정하고, 상기 제 1 파라미터, 상기 제 2 파라미터, 또는 상기 제 3 파라미터에 기초하여 상기 공통 음향 데이터 셋을 생성하도록 하는 인스트럭션들을 저장할 수 있다.
- [0110] 본 발명의 다양한 실시예들에 따른 메모리(960)는 공통 음향 데이터 셋, 적어도 하나의 결정 트리에 관한 정보,

및 상기 결정 트리의 인덱스가 지시하는 적어도 하나의 음향 데이터 셋을 저장할 수 있다.

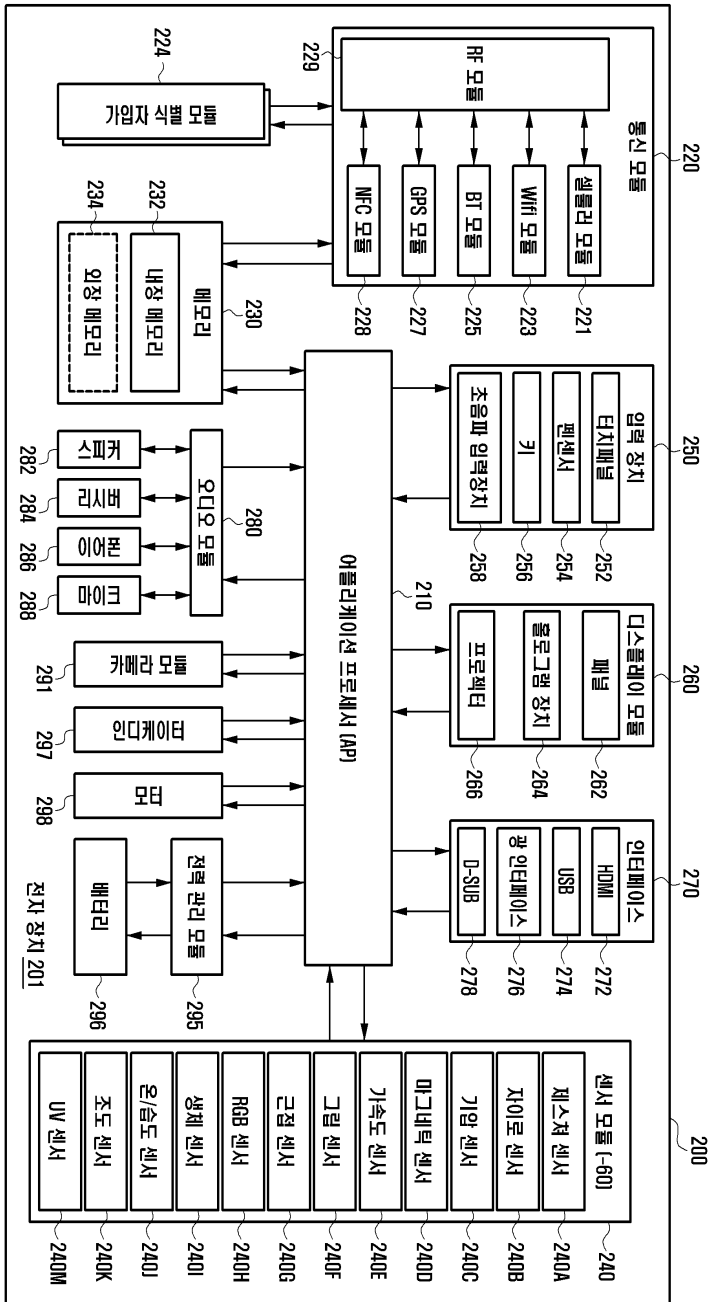
- [0111] 본 발명의 다양한 실시예들에 따른 제 2 전자 장치(902)의 통신 모듈(970)은 도 2의 전자 장치(201)의 통신 모듈(220)의 기능을 수행할 수 있다. 통신 모듈(940)은 결정 트리에 관한 정보 및/ 또는 공통 음향 데이터 셋에 관한 정보를 요청하는 요청 메시지를 제 1 전자 장치(901)로부터 수신할 수 있고, 결정 트리에 관한 정보 및/ 또는 공통 음향 데이터 셋에 관한 정보를 제 1 전자 장치(901)에 전송할 수 있다.
- [0112] 본 문서에서 사용된 용어 "모듈"은, 예를 들면, 하드웨어, 소프트웨어 또는 펌웨어(firmware) 중 하나 또는 둘 이상의 조합을 포함하는 단위(unit)를 의미할 수 있다. "모듈"은, 예를 들면, 유닛(unit), 로직(logic), 논리 블록(logical block), 부품(component), 또는 회로(circuit) 등의 용어와 바꾸어 사용(interchangeably use)될 수 있다. "모듈"은, 일체로 구성된 부품의 최소 단위 또는 그 일부가 될 수 있다. "모듈"은 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수도 있다. "모듈"은 기계적으로 또는 전자적으로 구현될 수 있다. 예를 들면, "모듈"은, 알려졌거나 앞으로 개발될, 어떤 동작들을 수행하는 ASIC(application-specific integrated circuit) 칩, FPGAs(field-programmable gate arrays) 또는 프로그램 가능 논리 장치(programmable-logic device) 중 적어도 하나를 포함할 수 있다.
- [0113] 다양한 실시예에 따른 장치(예: 모듈들 또는 그 기능들) 또는 방법(예: 동작들)의 적어도 일부는, 예컨대, 프로그램 모듈의 형태로 컴퓨터로 읽을 수 있는 저장매체(computer-readable storage media)에 저장된 명령어로 구현될 수 있다. 상기 명령어가 프로세서(예: 프로세서(120))에 의해 실행될 경우, 상기 하나 이상의 프로세서가 상기 명령어에 해당하는 기능을 수행할 수 있다. 컴퓨터로 읽을 수 있는 저장매체는, 예를 들면, 메모리(130)가 될 수 있다.
- [0114] 컴퓨터로 판독 가능한 기록 매체는, 하드디스크, 플로피디스크, 마그네틱 매체(magnetic media)(예: 자기테이프), 광기록 매체(optical media)(예: CD-ROM(compact disc read only memory), DVD(digital versatile disc), 자기-광 매체(magneto-optical media)(예: 플롭티컬 디스크(floptical disk)), 하드웨어 장치(예: ROM(read only memory), RAM(random access memory), 또는 플래시 메모리 등) 등을 포함할 수 있다. 또한, 프로그램 명령에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함할 수 있다. 상술한 하드웨어 장치는 다양한 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지다.
- [0115] 다양한 실시예에 따른 모듈 또는 프로그램 모듈은 전술한 구성요소들 중 적어도 하나 이상을 포함하거나, 일부가 생략되거나, 또는 추가적인 다른 구성요소를 더 포함할 수 있다. 다양한 실시예에 따른 모듈, 프로그램 모듈 또는 다른 구성요소에 의해 수행되는 동작들은 순차적, 병렬적, 반복적 또는 휴리스틱(heuristic)한 방법으로 실행될 수 있다. 또한, 일부 동작은 다른 순서로 실행되거나, 생략되거나, 또는 다른 동작이 추가될 수 있다. 그리고 본 문서에 개시된 실시예는 개시된, 기술 내용의 설명 및 이해를 위해 제시된 것이며, 본 문서에서 기재된 기술의 범위를 한정하는 것은 아니다. 따라서, 본 문서의 범위는, 본 문서의 기술적 사상에 근거한 모든 변경 또는 다양한 다른 실시예를 포함하는 것으로 해석되어야 한다.

도면

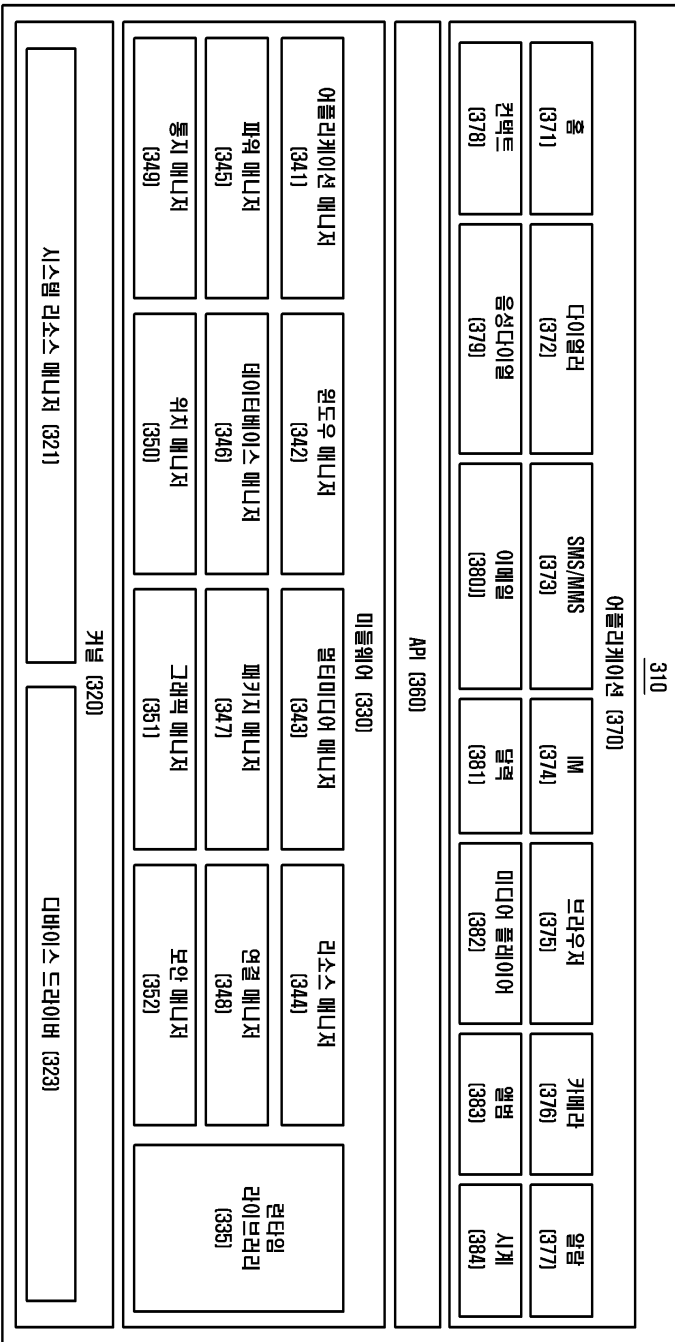
도면1



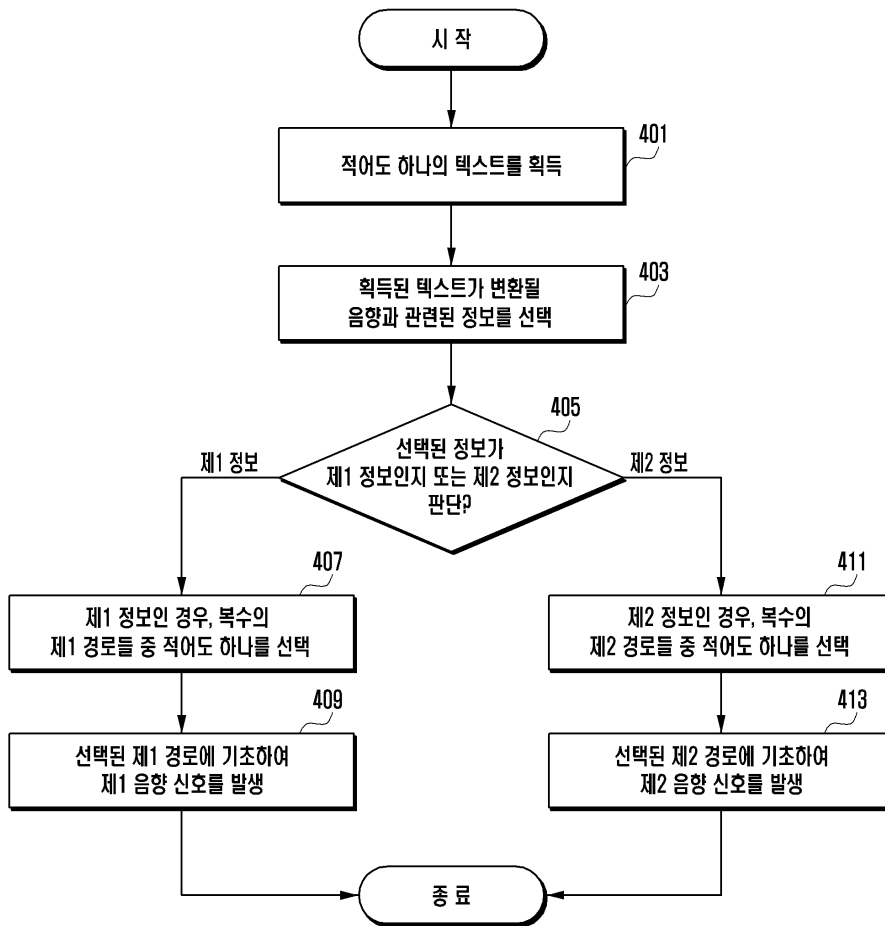
도면2



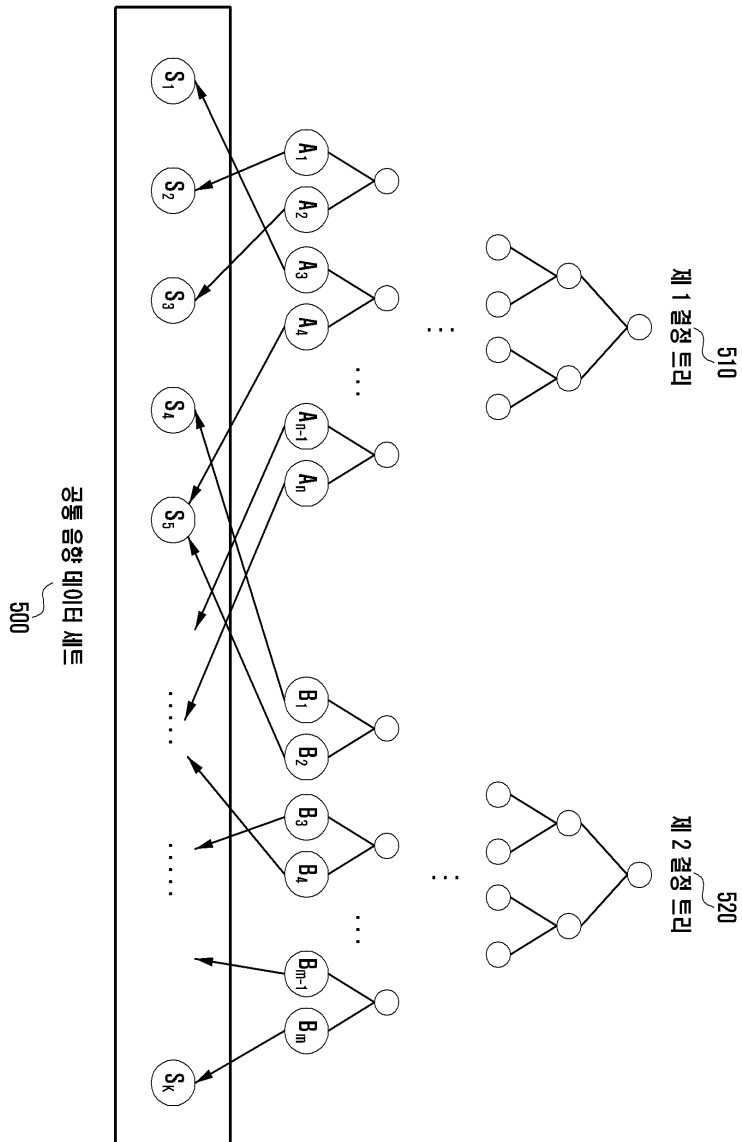
도면3



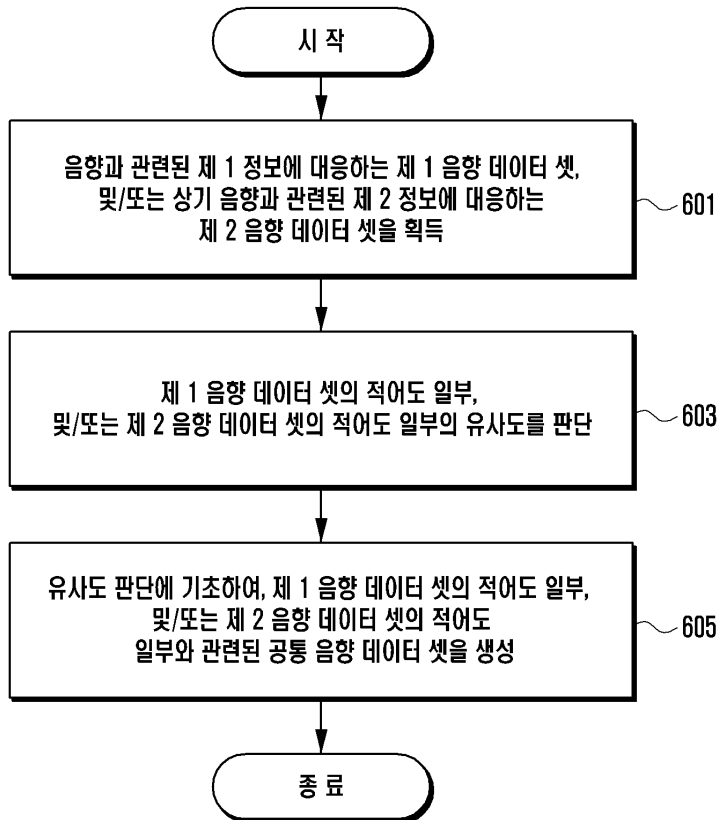
도면4



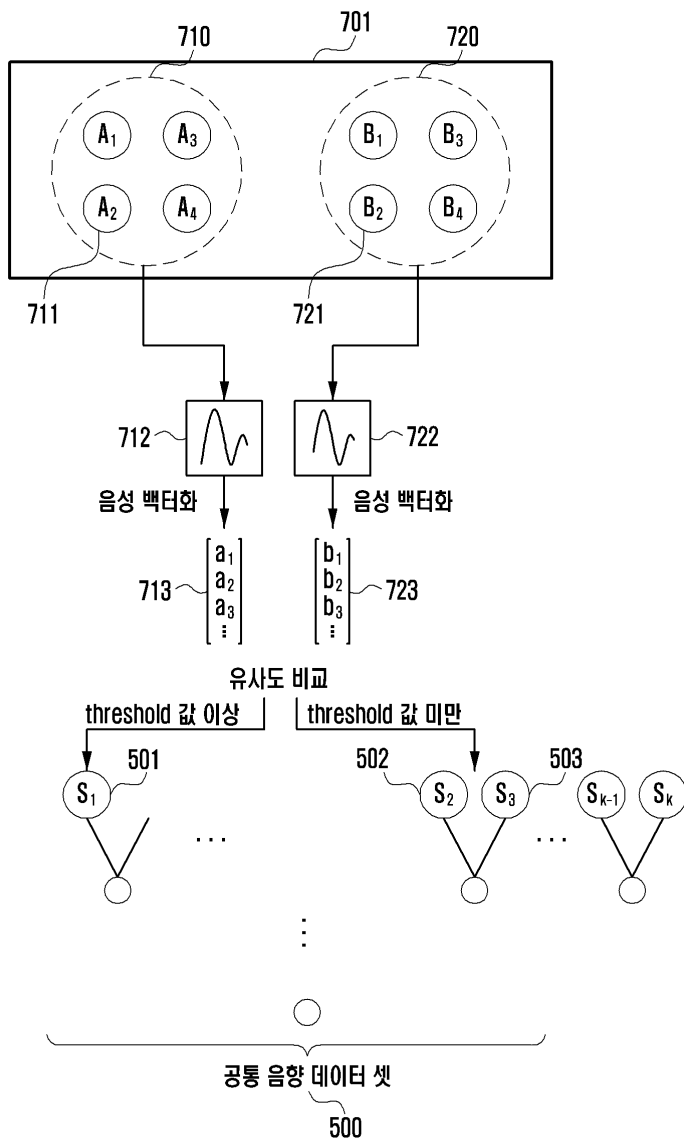
도면5



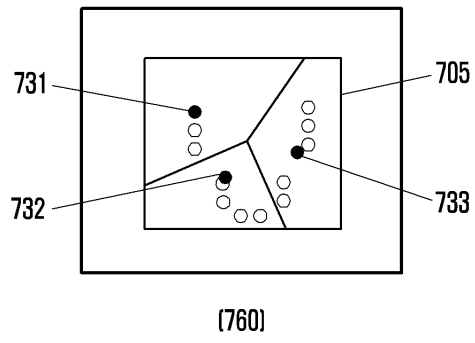
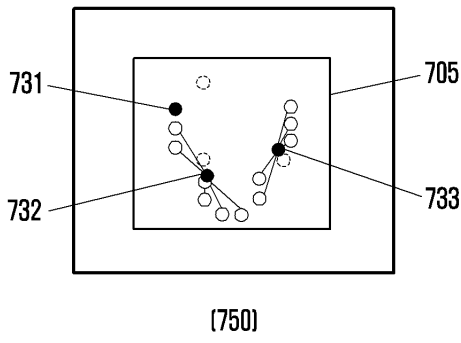
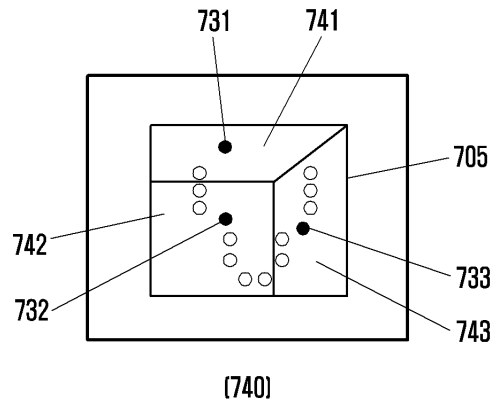
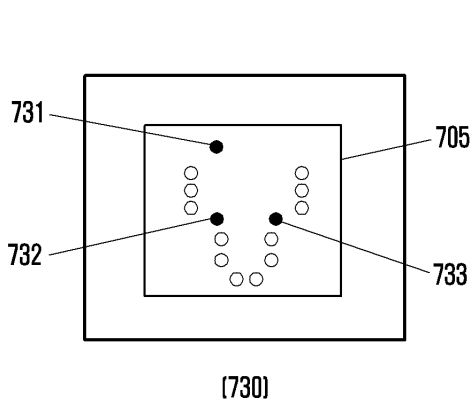
도면6



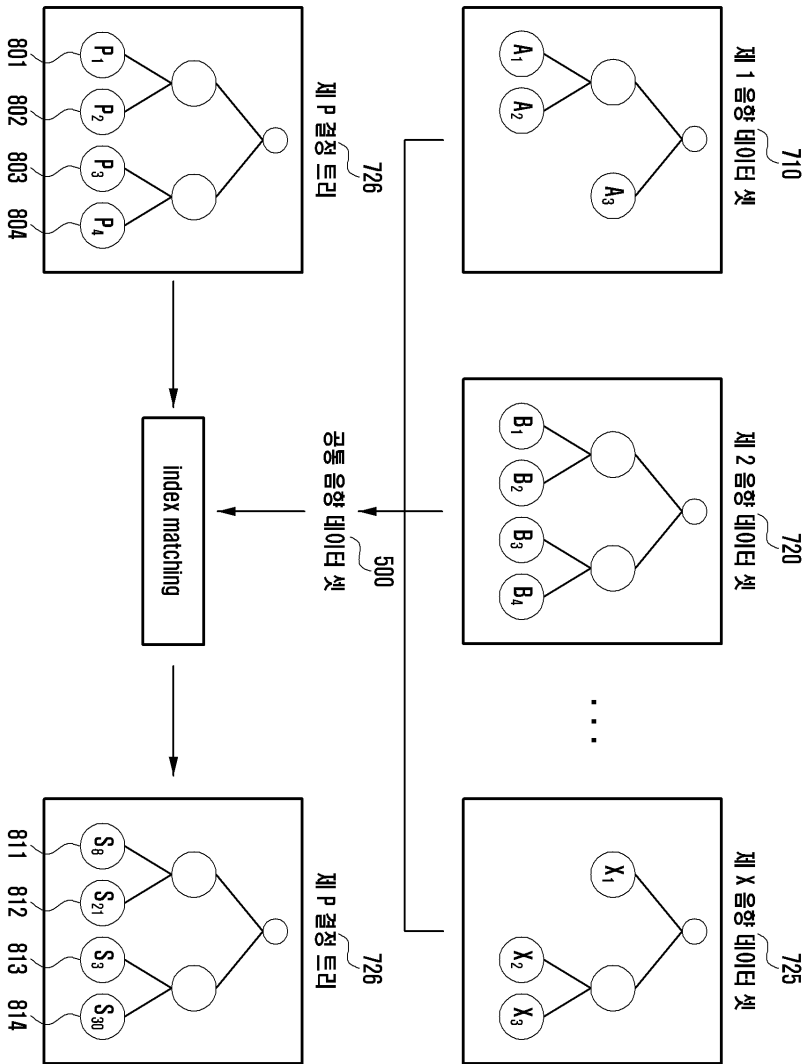
도면7a



도면7b



도면8



도면9

