



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0138993  
(43) 공개일자 2020년12월11일

(51) 국제특허분류(Int. Cl.)

G10L 13/02 (2006.01) G10L 13/08 (2006.01)

G10L 15/06 (2006.01) G10L 25/63 (2013.01)

(52) CPC특허분류

G10L 13/02 (2013.01)

G10L 13/08 (2013.01)

(21) 출원번호 10-2019-0065463

(22) 출원일자 2019년06월03일

심사청구일자 없음

(71) 출원인

한국전자통신연구원

대전광역시 유성구 가정로 218 (가정동)

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

장인선

대전광역시 유성구 배울2로 78, 607동 1101호(관평동, 대덕테크노밸리6단지아파트)

강홍구

서울특별시 서대문구 연희로32길 48, 108동 702호(연희동, 연희동성원아파트)

(뒷면에 계속)

(74) 대리인

팬코리아특허법인

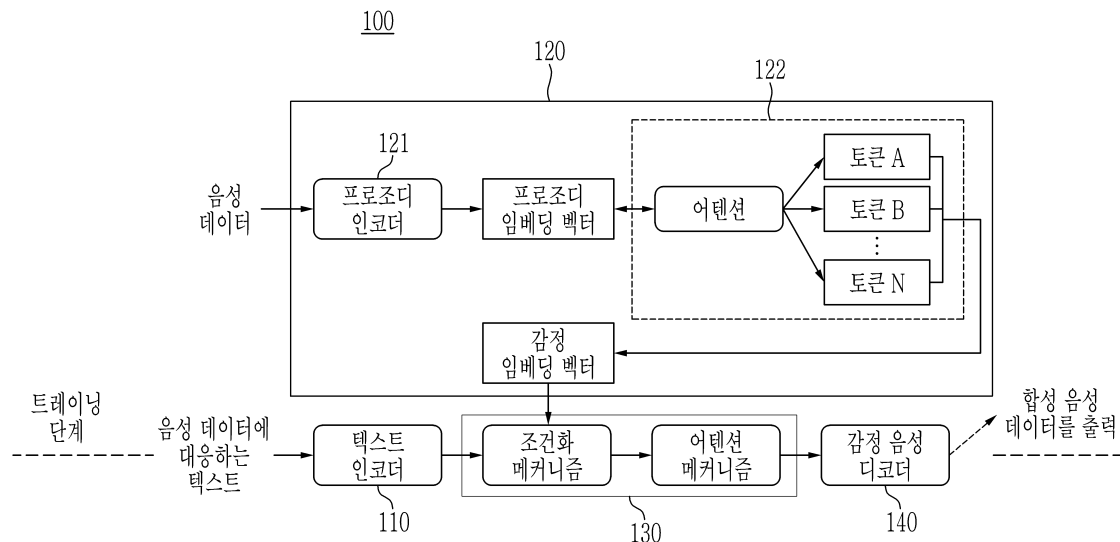
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 감정 토큰을 이용한 감정 음성 합성 방법 및 장치

(57) 요약

데이터베이스에 저장된 음성 데이터를 이용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계, 텍스트를 인코딩하는 단계, 복수의 감정 토큰의 트레이닝 결과를 바탕으로 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계, 그리고 감정 임베딩 벡터 및 인코딩된 텍스트를 합성하는 단계를 통해 텍스트에 감정 음성을 합성하는 방법 및 장치가 제공된다.

대표도 - 도1



(52) CPC특허분류

**G10L 15/063** (2013.01)

**G10L 25/63** (2013.01)

(72) 발명자

**권오성**

서울특별시 동대문구 정릉천동로 36, 107동 604호  
(용두동, 래미안허브리즈)

**안충현**

대전광역시 유성구 대덕대로541번길 68, 103동 30  
6호 (도룡동, 대덕연구원현대아파트)

**서정일**

대전광역시 유성구 반석서로 109, 704동 1704호 (반석동, 반석마을7단지아파트)

**양승준**

대전광역시 유성구 지족로 240, 504동 505호 (지족동, 노은해랑숲마을5단지아파트)

**최지훈**

대전광역시 유성구 은구비남로 56, 903동 1902호(노은동, 열매마을 아파트 9단지)

이 발명을 지원한 국가연구개발사업

과제고유번호

2015-0-00860

부처명

과학기술정보통신부

과제관리(전문)기관명

정보통신기획평가원

연구사업명

ETRI 연구개발지원사업

연구과제명

시청각장애인 방송접근권 향상을 위한 디지털자막·음성해설 서비스 기술 개발

기 여 율

1/1

과제수행기관명

한국전자통신연구원

연구기간

2018.03.01~2018.12.31

## 명세서

### 청구범위

#### 청구항 1

텍스트에 감정 음성을 합성하는 방법으로서,

데이터베이스에 저장된 음성 데이터를 이용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계,

텍스트를 인코딩하는 단계,

상기 복수의 감정 토큰의 트레이닝 결과를 바탕으로 상기 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계, 그리고

상기 감정 임베딩 벡터 및 상기 인코딩된 텍스트를 합성하는 단계

를 포함하는 감정 음성 합성 방법.

#### 청구항 2

제1항에서,

상기 복수의 감정 토큰을 트레이닝하는 단계는,

상기 음성 데이터로부터 생성된 프로조디 임베딩 벡터 및 상기 복수의 감정 토큰을 사용하여 후보 감정 임베딩 벡터를 생성하는 단계,

상기 음성 데이터의 텍스트에 상기 후보 감정 임베딩 벡터를 합성하여 합성 음성 데이터를 출력하는 단계, 그리고

상기 합성 음성 데이터가 상기 음성 데이터와 동일한지 여부를 결정하는 단계

를 포함하는, 감정 음성 합성 방법.

#### 청구항 3

제2항에서,

상기 데이터베이스에는 복수의 음성 데이터가 감정별로 카테고리화되어 있고,

상기 후보 감정 임베딩 벡터를 생성하는 단계는,

상기 데이터베이스의 제1 감정 카테고리에 속한 제1 음성 데이터를 프로조디 임베딩 벡터로 인코딩하는 단계,

상기 어텐션 메커니즘을 바탕으로 상기 프로조디 임베딩 벡터 및 상기 복수의 감정 토큰 사이의 유사도를 학습하는 단계,

상기 프로조디 임베딩 벡터에 대한 상기 복수의 감정 토큰의 가중치를 결정하는 단계, 그리고

상기 가중치의 가중치 행렬 및 상기 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 상기 후보 감정 임베딩 벡터를 생성하는 단계

를 포함하는, 감정 음성 합성 방법.

#### 청구항 4

제3항에서,

상기 합성 음성 데이터가 상기 음성 데이터와 동일하면, 상기 제1 감정 카테고리에 속한 다른 음성 데이터를 사용하여 상기 어텐션 메커니즘을 바탕으로 상기 복수의 감정 토큰을 트레이닝하는 단계

를 더 포함하는 감정 음성 합성 방법.

#### 청구항 5

제3항에서,

상기 합성 음성 데이터가 상기 음성 데이터와 동일하면, 상기 제1 감정 카테고리나 다른 감정에 대응하는 제2 감정 카테고리에 속한 제2 음성 데이터를 사용하여 상기 어텐션 메커니즘을 바탕으로 상기 복수의 감정 토큰을 트레이닝하는 단계

를 더 포함하는 감정 음성 합성 방법.

#### 청구항 6

제2항에서,

상기 합성 음성 데이터가 상기 음성 데이터와 동일하지 않으면, 트레이닝 프로세스를 업데이트하고, 상기 음성 데이터로부터 프로조디 임베딩 벡터를 다시 생성하여 상기 트레이닝 프로세스를 다시 수행하는 단계

를 더 포함하는 감정 음성 합성 방법.

#### 청구항 7

제1항에서,

상기 복수의 감정 토큰의 트레이닝 결과를 바탕으로 상기 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계는,

상기 텍스트에 합성하려는 감정으로 발성된 참조 음성 데이터를 사용하여 상기 감정 임베딩 벡터를 생성하는 단계

를 포함하는, 감정 음성 합성 방법.

#### 청구항 8

제2항에서,

상기 복수의 감정 토큰의 트레이닝 결과를 바탕으로 상기 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계는,

상기 데이터베이스에서 상기 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계, 그리고

결정된 가중치 행렬 및 상기 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 감정 임베딩 벡터를 생성하는 단계를 포함하는, 감정 음성 합성 방법.

#### 청구항 9

제8항에서,

상기 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계는,

텍스트에 합성하려는 감정에 대응하는, 상기 데이터베이스의 가중치 행렬 집합의 통계치를 바탕으로 상기 가중치 행렬을 결정하는 단계

를 포함하는, 감정 음성 합성 방법.

#### 청구항 10

제1항에서,

상기 감정 임베딩 벡터 및 상기 인코딩된 텍스트를 합성하는 단계는,

상기 감정 임베딩 벡터를 컨디셔닝 벡터로서 사용하여 상기 인코딩된 텍스트에 상기 감정 임베딩 벡터를 합성하

는 단계

를 포함하는, 감정 음성 합성 방법.

#### 청구항 11

텍스트에 감정 음성을 합성하는 장치로서,

프로세서, 메모리, 데이터베이스, 및 입출력 인터페이스를 포함하고, 상기 프로세서는 상기 메모리에 저장된 프로그램을 실행하여,

상기 데이터베이스에 저장된 음성 데이터를 이용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계,

상기 입출력 인터페이스로 입력된 텍스트를 인코딩하는 단계,

상기 복수의 감정 토큰의 트레이닝 결과를 바탕으로 상기 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계, 그리고

인코딩된 텍스트에 상기 감정 임베딩 벡터를 합성하는 단계

를 수행하는 감정 음성 합성 장치.

#### 청구항 12

제11항에서,

상기 프로세서는 상기 복수의 감정 토큰을 트레이닝하는 단계를 수행할 때,

상기 음성 데이터로부터 생성된 프로조디 임베딩 벡터 및 상기 복수의 감정 토큰을 사용하여 후보 감정 임베딩 벡터를 생성하는 단계,

상기 음성 데이터의 텍스트에 상기 후보 감정 임베딩 벡터를 합성하여 합성 음성 데이터를 출력하는 단계, 그리고

상기 합성 음성 데이터가 상기 음성 데이터와 동일한지 여부를 결정하는 단계

를 수행하는, 감정 음성 합성 장치.

#### 청구항 13

제12항에서,

상기 데이터베이스에는 복수의 음성 데이터가 감정별로 카테고리화되어 있고,

상기 프로세서는 상기 후보 감정 임베딩 벡터를 생성하는 단계를 수행할 때,

상기 데이터베이스의 제1 감정 카테고리에 속한 제1 음성 데이터를 프로조디 임베딩 벡터로 인코딩하는 단계,

상기 어텐션 메커니즘을 바탕으로 상기 프로조디 임베딩 벡터 및 상기 복수의 감정 토큰 사이의 유사도를 학습하는 단계,

상기 프로조디 임베딩 벡터에 대한 상기 복수의 감정 토큰의 가중치를 결정하는 단계, 그리고

상기 가중치의 가중치 행렬 및 상기 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 상기 후보 감정 임베딩 벡터를 생성하는 단계

를 수행하는, 감정 음성 합성 장치.

#### 청구항 14

제13항에서,

상기 프로세서는 상기 프로그램을 실행하여,

상기 합성 음성 데이터가 상기 음성 데이터와 동일하면, 상기 제1 감정 카테고리에 속한 다른 음성 데이터를 사

용하여 상기 어텐션 메커니즘을 바탕으로 상기 복수의 감정 토큰을 트레이닝하는 단계를 더 수행하는, 감정 음성 합성 장치.

#### 청구항 15

제13항에서,

상기 프로세서는 상기 프로그램을 실행하여,

상기 합성 음성 데이터가 상기 음성 데이터와 동일하면, 상기 제1 감정 카테고리나 다른 감정에 대응하는 제2 감정 카테고리에 속한 제2 음성 데이터를 사용하여 상기 어텐션 메커니즘을 바탕으로 상기 복수의 감정 토큰을 트레이닝하는 단계

를 더 수행하는, 감정 음성 합성 장치.

#### 청구항 16

제12항에서,

상기 프로세서는 상기 프로그램을 실행하여,

상기 합성 음성 데이터가 상기 음성 데이터와 동일하지 않으면, 상기 감정 음성 합성 장치를 업데이트하고, 상기 음성 데이터로부터 프로조디 임베딩 벡터를 다시 생성하여 트레이닝 프로세스를 다시 수행하는 단계

를 수행하는, 감정 음성 합성 장치.

#### 청구항 17

제11항에서,

상기 프로그램은 상기 복수의 감정 토큰의 트레이닝 결과를 바탕으로 상기 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계를 수행할 때,

상기 텍스트에 합성하려는 감정으로 발생된 참조 음성 데이터를 사용하여 상기 감정 임베딩 벡터를 생성하는 단계

를 수행하는, 감정 음성 합성 장치.

#### 청구항 18

제12항에서,

상기 프로세서는 상기 복수의 감정 토큰의 트레이닝 결과를 바탕으로 상기 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계를 수행할 때,

상기 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계, 그리고

결정된 가중치 행렬 및 상기 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 감정 임베딩 벡터를 생성하는 단계

를 수행하는, 감정 음성 합성 장치.

#### 청구항 19

제18항에서,

상기 프로세서는 상기 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계를 수행할 때,

텍스트에 합성하려는 감정에 대응하는, 상기 데이터베이스의 가중치 행렬 집합의 통계치를 바탕으로 상기 가중치 행렬을 결정하는 단계

를 수행하는, 감정 음성 합성 장치.

#### 청구항 20

제11항에서,

상기 프로세서는 상기 감정 임베딩 벡터 및 상기 인코딩된 텍스트를 합성하는 단계를 수행할 때,

상기 감정 임베딩 벡터를 컨디셔닝 벡터로서 사용하여 상기 인코딩된 텍스트에 상기 감정 임베딩 벡터를 합성하는 단계

를 수행하는, 감정 음성 합성 장치.

## 발명의 설명

## 기술 분야

[0001] 본 기재는 감정 토큰에 대해 기계 학습을 수행하여 텍스트에 감정 음성을 합성하는 장치 및 방법에 관한 것이다.

## 배경 기술

[0002] 음성 합성 시스템(text-to-speech; TTS)은 텍스트로부터 사람의 발성과 같은 자연스러운 음성을 출력하는 시스템이다. 종단 간 음성 합성 시스템(end-to-end TTS system)은 심층 신경망을 통해 텍스트로부터 숨겨진 특징 표현(hidden feature representation)을 추출하고, 추출된 특징 벡터를 활용하여 시간 영역의 음성 신호(time-domain speech signal)를 생성하는 시스템이다. 기존의 통계적 파라메트릭 음성 합성(statistical parametric speech synthesis; SPSS) 방식은 텍스트 분석, 음향 모델링 등 복잡한 모델링 과정의 파이프 라인을 갖고 있으며, 이로부터 불가피하게 발생하는 오차는 합성되는 음성의 품질을 저하시킬 수 있다. 하지만 종단간 음성 합성 시스템은 하나의 단일한 신경망 기반의 구조를 사용하여, 분야 전문성(domain expertise) 없이, 텍스트와 오디오 데이터를 포함하는 대용량 학습용 데이터베이스를 기반으로 기계학습을 수행하고, 기계학습의 결과를 바탕으로 고품질의 합성음을 제공할 수 있다.

[0003] 신경망에 기반하는 음성 합성 시스템을 통해 감정을 나타내는 음성을 표현하려는 연구가 진행 중이다. 하지만, 음성 합성 시스템의 훈련에 사용되는 데이터베이스를 여러 감정을 표현하는 다양한 음성 데이터로 구성하더라도 텍스트로부터 생성된 음성에 감정을 표현하기 위한 음성 내 운율(prosody)을 합성하는 것은 어려운 일이다. 감정 음성을 합성하기 위해 여러 감정을 서로 다른 레이블 벡터를 사용하여 나타내고, 감정에 대한 정보를 음성의 합성 과정에 주입하는 방법이 시도되었다. 각 감정을 표현하는 음성에 따라 음향 정보 특성에 차이가 있으므로, 서로 다른 감정에 따라 특정되는 벡터를 이용하여 감정 표현과 관련된 정보가 음성의 합성 과정에 주입될 수 있다.

## 발명의 내용

### 해결하려는 과제

[0004] 한 실시예는, 텍스트에 감정 음성을 합성하는 방법을 제공한다.

[0005] 다른 실시예는, 텍스트에 감정 음성을 합성하는 다른 방법을 제공한다.

[0006] 또 다른 실시예는, 텍스트에 감정 음성을 합성하는 장치를 제공한다.

### 과제의 해결 수단

[0007] 한 실시예에 따르면 텍스트에 감정 음성을 합성하는 방법이 제공된다. 상기 감정 음성 합성 방법은, 데이터베이스에 저장된 음성 데이터를 이용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계, 텍스트를 인코딩하는 단계, 복수의 감정 토큰의 트레이닝 결과를 바탕으로 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계, 그리고 감정 임베딩 벡터 및 인코딩된 텍스트를 합성하는 단계를 포함한다.

[0008] 상기 감정 음성 합성 방법에서 복수의 감정 토큰을 트레이닝하는 단계는, 음성 데이터로부터 생성된 프로조디 임베딩 벡터 및 복수의 감정 토큰을 사용하여 후보 감정 임베딩 벡터를 생성하는 단계, 음성 데이터의 텍스트에 후보 감정 임베딩 벡터를 합성하여 합성 음성 데이터를 출력하는 단계, 그리고 합성 음성 데이터가 음성 데이터와 동일한지 여부를 결정하는 단계를 포함할 수 있다.

[0009] 상기 감정 음성 합성 방법에서 데이터베이스에는 복수의 음성 데이터가 감정별로 카테고리화되어 있고, 후보 감

정 임베딩 벡터를 생성하는 단계는, 데이터베이스의 제1 감정 카테고리에 속한 제1 음성 데이터를 프로조디 임베딩 벡터로 인코딩하는 단계, 어텐션 메커니즘을 바탕으로 프로조디 임베딩 벡터 및 복수의 감정 토큰 사이의 유사도를 학습하는 단계, 프로조디 임베딩 벡터에 대한 복수의 감정 토큰의 가중치를 결정하는 단계, 그리고 가중치의 가중치 행렬 및 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 후보 감정 임베딩 벡터를 생성하는 단계를 포함할 수 있다.

- [0010] 상기 감정 음성 합성 방법은, 합성 음성 데이터가 음성 데이터와 동일하면, 제1 감정 카테고리에 속한 다른 음성 데이터를 사용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계를 더 포함할 수 있다.
- [0011] 상기 감정 음성 합성 방법은, 합성 음성 데이터가 음성 데이터와 동일하면, 제1 감정 카테고리와 다른 감정에 대응하는 제2 감정 카테고리에 속한 제2 음성 데이터를 사용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계를 더 포함할 수 있다.
- [0012] 상기 감정 음성 합성 방법에서 합성 음성 데이터가 음성 데이터와 동일하지 않으면, 트레이닝 프로세스를 업데이트하고, 음성 데이터로부터 프로조디 임베딩 벡터를 다시 생성하여 트레이닝 프로세스를 다시 수행하는 단계를 더 포함할 수 있다.
- [0013] 상기 감정 음성 합성 방법에서 복수의 감정 토큰의 트레이닝 결과를 바탕으로 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계는, 텍스트에 합성하려는 감정으로 발생된 참조 음성 데이터를 사용하여 감정 임베딩 벡터를 생성하는 단계를 포함할 수 있다.
- [0014] 상기 감정 음성 합성 방법에서 복수의 감정 토큰의 트레이닝 결과를 바탕으로 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계는, 데이터베이스에서 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계, 그리고 결정된 가중치 행렬 및 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 감정 임베딩 벡터를 생성하는 단계를 포함할 수 있다.
- [0015] 상기 감정 음성 합성 방법에서 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계는, 텍스트에 합성하려는 감정에 대응하는, 데이터베이스의 가중치 행렬 집합의 통계치를 바탕으로 가중치 행렬을 결정하는 단계를 포함할 수 있다.
- [0016] 상기 감정 음성 합성 방법에서 감정 임베딩 벡터 및 인코딩된 텍스트를 합성하는 단계는, 감정 임베딩 벡터를 컨디셔닝 벡터로서 사용하여 인코딩된 텍스트에 감정 임베딩 벡터를 합성하는 단계를 포함할 수 있다.
- [0017] 다른 실시예에 따르면, 텍스트에 감정 음성을 합성하는 장치가 제공된다. 상기 감정 음성 합성 장치는, 프로세서, 메모리, 데이터베이스, 및 입출력 인터페이스를 포함하고, 프로세서는 메모리에 저장된 프로그램을 실행하여, 데이터베이스에 저장된 음성 데이터를 이용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계, 입출력 인터페이스로 입력된 텍스트를 인코딩하는 단계, 복수의 감정 토큰의 트레이닝 결과를 바탕으로 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계, 그리고 인코딩된 텍스트에 감정 임베딩 벡터를 합성하는 단계를 수행한다.
- [0018] 상기 감정 음성 합성 장치에서 프로세서는 복수의 감정 토큰을 트레이닝하는 단계를 수행할 때, 음성 데이터로부터 생성된 프로조디 임베딩 벡터 및 복수의 감정 토큰을 사용하여 후보 감정 임베딩 벡터를 생성하는 단계, 음성 데이터의 텍스트에 후보 감정 임베딩 벡터를 합성하여 합성 음성 데이터를 출력하는 단계, 그리고 합성 음성 데이터가 음성 데이터와 동일한지 여부를 결정하는 단계를 수행할 수 있다.
- [0019] 상기 감정 음성 합성 장치에서 데이터베이스에는 복수의 음성 데이터가 감정별로 카테고리화되어 있고, 프로세서는 후보 감정 임베딩 벡터를 생성하는 단계를 수행할 때, 데이터베이스의 제1 감정 카테고리에 속한 제1 음성 데이터를 프로조디 임베딩 벡터로 인코딩하는 단계, 어텐션 메커니즘을 바탕으로 프로조디 임베딩 벡터 및 복수의 감정 토큰 사이의 유사도를 학습하는 단계, 프로조디 임베딩 벡터에 대한 복수의 감정 토큰의 가중치를 결정하는 단계, 그리고 가중치의 가중치 행렬 및 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 후보 감정 임베딩 벡터를 생성하는 단계를 수행할 수 있다.
- [0020] 상기 감정 음성 합성 장치에서 프로세서는 프로그램을 실행하여, 합성 음성 데이터가 음성 데이터와 동일하면, 제1 감정 카테고리에 속한 다른 음성 데이터를 사용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계를 더 수행할 수 있다.
- [0021] 상기 감정 음성 합성 장치에서 프로세서는 프로그램을 실행하여, 합성 음성 데이터가 음성 데이터와 동일하면,



제1 감정 카테고리와의 다른 감정에 대응하는 제2 감정 카테고리에 속한 제2 음성 데이터를 사용하여 어텐션 메커니즘을 바탕으로 복수의 감정 토큰을 트레이닝하는 단계를 더 수행할 수 있다.

[0022] 상기 감정 음성 합성 장치에서 프로세서는 프로그램을 실행하여, 합성 음성 데이터가 음성 데이터와 동일하지 않으면, 감정 음성 합성 장치를 업데이트하고, 음성 데이터로부터 프로조디 임베딩 벡터를 다시 생성하여 트레이닝 프로세스를 다시 수행하는 단계를 더 수행할 수 있다.

[0023] 상기 감정 음성 합성 장치에서 프로그램은 복수의 감정 토큰의 트레이닝 결과를 바탕으로 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계를 수행할 때, 텍스트에 합성하려는 감정으로 발성된 참조 음성 데이터를 사용하여 감정 임베딩 벡터를 생성하는 단계를 수행할 수 있다.

[0024] 상기 감정 음성 합성 장치에서 프로세서는 복수의 감정 토큰의 트레이닝 결과를 바탕으로 텍스트에 합성하려는 감정에 대응하는 감정 임베딩 벡터를 결정하는 단계를 수행할 때, 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계, 그리고 결정된 가중치 행렬 및 복수의 감정 토큰의 감정 토큰 행렬을 바탕으로 감정 임베딩 벡터를 생성하는 단계를 수행할 수 있다.

[0025] 상기 감정 음성 합성 장치에서 프로세서는 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정하는 단계를 수행할 때, 텍스트에 합성하려는 감정에 대응하는, 데이터베이스의 가중치 행렬 집합의 통계치를 바탕으로 가중치 행렬을 결정하는 단계를 더 수행할 수 있다.

[0026] 상기 감정 음성 합성 장치에서 프로세서는 감정 임베딩 벡터 및 인코딩된 텍스트를 합성하는 단계를 수행할 때, 감정 임베딩 벡터를 컨디셔닝 벡터로서 사용하여 인코딩된 텍스트에 감정 임베딩 벡터를 합성하는 단계를 더 수행할 수 있다.

### 발명의 효과

[0027] 텍스트에 합성하려는 감정이 감정 토큰의 트레이닝을 통해 모델링됨으로써, 여러 가지 감정 음성이 텍스트에 효과적으로 합성될 수 있다. 또한 감정별로 카테고리화된 음성 데이터를 바탕으로 감정 토큰의 트레이닝이 수행되고, 감정 모델링의 정확도가 향상된다. 또한, 감정 토큰의 트레이닝 결과로 생성되는 가중치 행렬이 데이터베이스 내에 카테고리화됨으로써, 기계 학습을 위한 데이터베이스의 관리가 용이하게 될 수 있다.

### 도면의 간단한 설명

- [0028] 도 1은 한 실시예에 따른 감정 음성 합성 장치의 트레이닝 단계를 나타낸 개념도이다.  
 도 2는 한 실시예에 따른 감정 음성 합성 장치의 트레이닝 방법을 나타낸 흐름도이다.  
 도 3은 한 실시예에 따른 감정 음성 합성 장치의 추론 단계를 나타낸 개념도이다.  
 도 4는 한 실시예에 따른 감정 음성 합성 장치의 감정 음성 합성 방법을 나타낸 흐름도이다.  
 도 5는 다른 실시예에 따른 감정 음성 합성 장치의 감정 음성 합성 방법을 나타낸 흐름도이다.  
 도 6은 다른 실시예에 따른 감정 음성 합성 장치를 나타낸 블록도이다.

### 발명을 실시하기 위한 구체적인 내용

[0029] 아래에서는 첨부한 도면을 참고로 하여 본 기재의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 기재는 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 기재를 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0030] 도 1은 한 실시예에 따른 감정 음성 합성 장치의 트레이닝 방법을 나타낸 개념도이고, 도 2는 한 실시예에 따른 감정 음성 합성 장치의 트레이닝 방법을 나타낸 흐름도이다.

[0031] 도 1을 참조하면, 한 실시예에 따른 감정 음성 합성 장치(100)는 텍스트 인코더(110), 감정 토큰 학습부(120), 감정 음성 결정부(130), 및 감정 음성 디코더(140)를 포함한다. 한 실시예에 따른 감정 음성 합성 장치(100)는 트레이닝 단계에서, 데이터베이스에 저장된 음성 데이터를 사용하여, 훈련에 사용된 음성 데이터와 동일한 오디오가 출력될 수 있도록 감정 토큰을 훈련시킨다. 한 실시예에 따른 감정 음성 합성 장치(100)는 트레이닝의 결

과를 검증하기 위한 트레이닝 제어부(150)를 더 포함할 수 있다. 또한 한 실시예에 따른 감정 음성 합성 장치(100)는 추론 단계에서, 텍스트에 미리 결정된 감정의 유형에 대응하는 감정 음성을 결정하고, 텍스트에 결정된 감정 음성을 합성한다. 예를 들어, 텍스트 "이리와"가 감정 음성 합성 장치(100)에 입력되고, "이리와"에 합성될 감정의 유형이 '분노'라면, 감정 음성 합성 장치(100)는 '분노'를 나타내는 감정 음성을 "이리와"에 합성한다. 이때 감정 음성의 발성 스타일(발화자의 성별 및 연령), 발성 언어의 종류, 방언의 사용 여부 등도 미리 결정되어 감정 음성 합성 장치(100)에 입력될 수 있다.

[0032] 도 2를 참조하면, 감정 음성 합성 장치(100)는 데이터베이스에 저장되어 있는 음성 데이터를 사용하여 감정 토큰 학습부(120) 및 감정 음성 디코더(140)를 트레이닝한다. 데이터베이스에는 음성 데이터가 각 감정별로 카테고리화되어 있다. 감정 음성 합성 장치(100)는 감정 카테고리에 포함된 음성 데이터를 사용하여 해당 카테고리의 감정을 모델링할 수 있다. 도 2에서, 감정 토큰 학습부(120)는 감정 모델링을 위해 프로조디 인코더(prosody encoder)(121) 및 감정 토큰 레이어(122)를 포함한다. 프로조디 인코더(121)는 데이터베이스의 오디오 데이터를 프로조디 임베딩 벡터로 변환하고, 프로조디 임베딩 벡터를 감정 토큰 레이어(122)에 입력한다. 감정 토큰 레이어(122)는 어텐션 메커니즘을 이용하여 감정 토큰의 가중치를 결정함으로써 감정 토큰을 트레이닝한다. 즉, 감정 토큰 학습부(120)는 프로조디 임베딩 벡터에 포함된 어쿠스틱 정보로부터 오디오 데이터의 감정 정보를 모델링할 수 있다. 이때 프로조디 임베딩 벡터의 어쿠스틱 정보는 오디오 데이터의 프로조디 정보를 포함한다.

[0033] 트레이닝 단계에서, 감정 토큰 학습부(120)의 프로조디 인코더(121)는 데이터베이스의 카테고리화된 음성 데이터로부터 프로조디 임베딩 벡터를 생성한다(S110). 데이터베이스의 각 감정 카테고리는 감정 카테고리에 해당하는 감정으로 발성된 음성 데이터를 포함한다. 즉, 사람의 감정인, 중립(neutral), 분노(anger), 경멸(contempt), 혐오(disgust), 두려움(fear), 행복(happiness), 슬픔(sadness), 놀람(surprise) 등의 감정으로 발성된 음성 데이터가 데이터베이스에 카테고리화 되어 있다. 데이터베이스의 음성 데이터는 단 시간 푸리에 변환(short-time Fourier transform)을 통해 스펙트로그램으로 변환되어 프로조디 인코더(121)로 입력될 수 있다. 또한 사람의 음성 데이터는 일반적으로, 녹음 환경에 의한 영향, 사람의 신체 상태 등 녹음 과정에서 발생할 수 있는 노이즈를 포함하고 있기 때문에 필터링 및 정규화가 필요하다. 그리고 한 실시예에 따른 감정 음성 합성 장치(100)의 학습 효율을 높이기 위해 녹음된 음성의 앞뒤에 존재하는 묵음도 미리 삭제될 필요가 있다. 따라서 음성 데이터는 필터링, 정규화, 및 묵음 삭제 등의 전처리가 수행된 후 감정 토큰 학습부(120)에 입력될 수 있다.

[0034] 감정 토큰 학습부(120)의 감정 토큰 레이어(122)는 어텐션 메커니즘을 바탕으로 음성 데이터의 프로조디 임베딩 벡터를 사용하여 후보 감정 임베딩 벡터를 생성한다(S120). 예를 들어, 감정 토큰 학습부(120)는 카테고리 '기쁨'에 속한 음성 데이터로부터 '기쁨'에 관한 프로조디 임베딩 벡터가 생성되면, 어텐션 메커니즘을 바탕으로 '기쁨'에 관한 프로조디 임베딩 벡터를 사용하여 '기쁨'을 모델링하는 후보 감정 임베딩 벡터를 생성할 수 있다.

[0035] 후보 감정 임베딩 벡터는, 어텐션(attention) 메커니즘을 사용하여 결정되는, 프로조디 임베딩 벡터 및 감정 토큰 사이의 유사도에 기반할 수 있다. 감정 토큰은 사람의 감정을 나타내기 위한 일종의 기저이고, 복수의 감정 토큰의 기여도에 따라 하나의 감정이 모델링될 수 있다. 예를 들어, 감정 토큰 학습부(120)는 프로조디 임베딩 벡터 및 감정 토큰 사이의 유사도를 바탕으로 프로조디 임베딩 벡터에 대한 각 감정 토큰을 가중치를 결정한다. 각 감정 토큰의 가중치는, 음성 데이터에 담긴 감정에 대한 각 감정 토큰의 기여도를 나타낼 수 있다. 예를 들어, 데이터베이스의 '기쁨' 카테고리의 음성 데이터가 어린 아이의 들뜬 목소리일 때, 어린 아이의 들뜬 목소리의 프로조디 임베딩 벡터와 복수의 감정 토큰 간의 유사도를 바탕으로 복수의 감정 토큰의 가중치가 결정된다. 복수의 감정 토큰의 가중치는 가중치 행렬을 구성할 수 있고, 감정 토큰 학습부(120)는 복수의 감정 토큰의 감정 토큰 행렬과 가중치 행렬의 행렬곱 연산을 수행하여 후보 감정 임베딩 벡터를 생성할 수 있다.

[0036] 감정 음성 디코더(140)는 후보 감정 임베딩 벡터를 음성 데이터에 대응하는 텍스트에 합성하여 합성 음성 데이터를 출력하고, 트레이닝 제어부(150)는 후보 감정 임베딩 벡터를 바탕으로 합성된 합성 음성 데이터와 트레이닝에 사용된 음성 데이터의 동일 여부(또는 동일 정도)를 판단하여 트레이닝 결과를 검증한다(S140). 트레이닝에 사용된 음성 데이터의 텍스트는, 음성 데이터와 함께 데이터베이스에 저장되어 있을 수도 있고, 또는 음성 데이터의 음성 인식 결과로부터 트레이닝의 결과 검증을 위해 별도로 입력될 수도 있다.

[0037] 즉, 감정 음성 합성 장치(100)는 후보 감정 임베딩 벡터가 트레이닝에 사용된 음성 데이터의 텍스트에 합성되어 출력되는 합성 음성 데이터가, 트레이닝에 사용된 음성 데이터와 동일한지 여부를 확인하여, 감정 임베딩 벡터를 통해 목표 감정이 적절히 모델링되었는지 여부를 판단할 수 있다. 트레이닝 제어부(150)는, 미리 결정된 오

차 범위에 따라, 후보 감정 임베딩 벡터 및 트레이닝에 사용된 음성 데이터의 텍스트 간의 합성이, 트레이닝에 사용된 음성 데이터와 동일한지 여부를 판단할 수 있다. 즉, 트레이닝 제어부(150)는 합성 음성 데이터와 트레이닝에 사용된 음성 데이터 간의 오차가 미리 결정된 오차 범위 내에 있다면, 합성 음성 데이터가 트레이닝에 사용된 음성 데이터와 동일한 것으로 판단할 수 있다.

[0038] 한 실시예에 따르면, 트레이닝의 목표는, 트레이닝에 사용된 음성 데이터와 동일한 오디오 신호가 출력되도록 감정 임베딩 벡터를 결정하는 것이다. 트레이닝 제어부(150)는 합성 음성 데이터가 트레이닝에 사용된 음성 데이터와 동일하지 않은 것으로 판단하면, 감정 음성 합성 장치(100)를 업데이트하고(예를 들어, 프로조디 인코더의 업데이트, 감정 토큰의 업데이트, 감정 음성 디코더의 업데이트 등 트레이닝 프로세스의 업데이트를 포함함) 동일한 음성 데이터의 프로조디 임베딩 벡터를 다시 생성하여 트레이닝을 재수행한다(S150). 하지만 트레이닝 제어부(150)는 합성 음성 데이터가 트레이닝에 사용된 음성 데이터와 동일한 것으로 판단하면, 동일한 감정 카테고리 내의 다른 음성 데이터 또는 다른 감정 카테고리 내의 음성 데이터를 사용하여 트레이닝을 수행한다(S160).

[0039] 그리고 트레이닝 제어부(150)는 생성된 가중치 행렬을 트레이닝된 감정 또는 트레이닝에 사용된 감정 카테고리와 연결지어 데이터베이스에 저장한다. 트레이닝된 하나의 감정에 대응하는 가중치 행렬 집합은 복수의 가중치 행렬을 포함한다. 가중치 행렬 집합에 포함된 가중치 행렬의 수량은 트레이닝된 감정 카테고리에 포함된 음성 데이터의 수량과 같거나 음성 데이터의 수량보다 많을 수 있다. 추론 단계에서, 텍스트에 합성하려는 감정의 감정 카테고리에 대응하는 가중치 행렬 집합 내에서 하나의 가중치 행렬이 선택되거나, 또는 가중치 행렬 집합의 통계치로부터 가중치 행렬이 결정될 수 있다. 아래에서는 실시예에 따른 감정 음성 합성 장치(100)의 감정 음성 합성 방법을 설명한다.

[0040] 도 3은 한 실시예에 따른 감정 음성 합성 장치의 추론 단계를 나타낸 개념도이고, 도 4는 한 실시예에 따른 감정 음성 합성 장치의 감정 음성 합성 방법을 나타낸 흐름도이다.

[0041] 추론 단계에서, 감정 음성 합성 장치(100)는 감정 임베딩 벡터를 사용하여 텍스트에 감정 음성을 합성한다. 추론 단계에서, 참조 음성 데이터로부터 생성된 감정 임베딩 벡터가 사용되거나 또는 미리 결정된 감정 카테고리에 대응하는 감정 임베딩 벡터가 사용될 수 있다. 도 4는 참조 음성 데이터로부터 생성되는 감정 임베딩 벡터를 사용하여 텍스트에 감정 음성을 합성하는 방법을 설명하고, 도 5는 미리 결정된 감정 카테고리에 대응하는 감정 임베딩 벡터를 사용하여 텍스트에 감정 음성을 합성하는 방법을 설명한다.

[0042] 도 4를 참조하면, 텍스트 인코더(110)는 감정 음성 합성 장치(100)로 입력되는, 감정을 합성하려는 텍스트 또는 텍스트 시퀀스를 인코딩한다(S210). 한 실시예에 따른 텍스트 인코더(110)는 입력된 텍스트로부터 높은 차원의 언어적인 특징 표현(higher-level linguistic feature representation)을 추출할 수 있다. 텍스트 또는 텍스트 시퀀스는 사용자에게 의해 개별적으로 감정 음성 합성 장치(100)에 입력될 수도 있고, 또는 음성 인식 장치, 자막 제공 장치 등의 장치에 의해 자동적으로 감정 음성 합성 장치(100)에 입력될 수도 있다. 또는 화면 해설 장치가 사용되는 경우, 화면의 해설을 위한, 미리 결정된 텍스트 또는 텍스트 시퀀스가 감정 음성 합성 장치(100)에 입력될 수 있다. 텍스트 시퀀스가 자막 제공 장치에 의해 입력될 때, 텍스트 시퀀스의 각 텍스트에는 서로 다른 감정 및 서로 다른 발성 스타일의 감정 음성이 합성될 수 있다.

[0043] 다음, 프로조디 인코더(121)가 참조 음성 데이터를 프로조디 임베딩 벡터로 인코딩한다(S220). 참조 음성 데이터는 텍스트에 합성하려는 감정으로 발성된 음성 데이터이다. 따라서, 감정 음성 합성 장치(100)는 입력 텍스트에 참조 음성 데이터가 나타내는 감정을 나타내는 감정 임베딩 벡터를 합성할 수 있다. 이후, 감정 토큰 레이어(122)는 어텐션 메커니즘을 사용하여 결정되는, 참조 음성 데이터의 프로조디 임베딩 벡터 및 감정 토큰 사이의 유사도를 바탕으로 감정 토큰을 트레이닝하고, 감정 임베딩 벡터를 생성한다(S230). 한 실시예에 따른 감정 토큰 학습부(120)는 이미 트레이닝 단계를 통해 훈련되었기 때문에, 참조 음성 데이터가 나타내는 감정을 모델링하는 감정 임베딩 벡터는 기훈련된 감정 토큰 학습부(120)에 의해 정교하게 생성될 수 있다.

[0044] 마지막으로, 감정 음성 디코더(140)는 참조 음성 데이터에 대응하는 감정 임베딩 벡터를 컨디셔닝 벡터로서 사용하여(즉, 조건화 메커니즘) 인코딩된 텍스트에 감정 음성을 합성한다(S240).

[0045] 도 5는 다른 실시예에 따른 감정 음성 합성 장치의 감정 음성 합성 방법을 나타낸 흐름도이다.

[0046] 도 5를 참조하면, 텍스트 인코더(110)는 감정 음성 합성 장치(100)로 입력되는, 감정을 합성하려는 텍스트 또는 텍스트 시퀀스를 인코딩한다(S310). 그리고 감정 음성 결정부(130)는 텍스트에 합성하려는 감정에 대응하는, 감정 토큰의 가중치 행렬을 결정한다(S320). 이때 텍스트에 합성하려는 감정은 미리 결정되고 감정 음성 합성 장

치(100)에 입력될 수 있다. 그리고 텍스트에 합성하려는 감정에 의해 표현될 수 있는 발성 스타일(발화자의 성별 및 연령), 발성 언어의 종류, 방언의 사용 여부 등도 미리 결정되어 감정 음성 합성 장치(100)에 입력될 수 있다. 감정 음성 결정부(130)는 텍스트에 합성하려는 감정에 대응하는, 데이터베이스의 가중치 행렬 집합 내에서 하나의 가중치 행렬을 선택하거나, 또는 가중치 행렬 집합 내의 가중치 행렬들의 통계치로부터 가중치 행렬을 결정할 수 있다. 데이터베이스의 가중치 행렬 집합은 각 감정 카테고리에 대응하고, 감정 토큰 학습부(120)에 의해 데이터베이스 내에 미리 저장된다. 감정 음성 결정부(130)는 가중치 행렬 집합 내에서 무작위로 하나의 가중치 행렬을 선택할 수 있다. 또는, 감정 음성 결정부(130)는 가중치 행렬 집합 내의 가중치 행렬 간의 평균 또는 표준편차를 바탕으로 가중치 행렬 집합 내에서 가중치 행렬을 결정할 수 있다.

[0047] 다음, 감정 음성 결정부(130)는 결정된 가중치 행렬 및 감정 토큰 행렬을 바탕으로 감정 임베딩 벡터를 생성한다(S330). 마지막으로, 감정 음성 디코더(140)는 합성하려는 감정에 대응하는 감정 임베딩 벡터를 컨디셔닝 벡터로서 사용하여 인코딩된 텍스트에 감정 음성을 합성한다(S340).

[0048] 위에서 설명한 대로, 텍스트에 합성하려는 감정이 감정 토큰의 트레이닝을 통해 모델링됨으로써, 여러 가지 감정 음성이 텍스트에 효과적으로 합성될 수 있다. 또한 감정별로 카테고리화된 음성 데이터를 바탕으로 감정 토큰의 트레이닝이 수행되고, 감정 모델링의 정확도가 향상된다. 또한, 감정 토큰의 트레이닝 결과로 생성되는 가중치 행렬이 데이터베이스 내에 카테고리화됨으로써, 기계 학습을 위한 데이터베이스의 관리가 용이하게 될 수 있다.

[0049] 도 6은 다른 실시예에 따른 감정 음성 합성 장치를 나타낸 블록도이다.

[0050] 한 실시예에 따른 종단간 음성 합성 장치는, 컴퓨터 시스템, 예를 들어 컴퓨터 판독 가능 매체로 구현될 수 있다. 도 6을 참조하면, 컴퓨터 시스템(600)은, 버스(670)를 통해 통신하는 프로세서(610), 메모리(630), 입력 인터페이스(650), 출력 인터페이스(660), 및 데이터베이스(640) 중 적어도 하나를 포함할 수 있다. 컴퓨터 시스템(600)은 또한 네트워크에 결합된 통신 장치(620)를 포함할 수 있다. 프로세서(610)는 중앙 처리 장치(central processing unit, CPU)이거나, 또는 메모리(630) 또는 데이터베이스(640)에 저장된 명령을 실행하는 반도체 장치일 수 있다. 메모리(630) 및 데이터베이스(640)는 다양한 형태의 휘발성 또는 비휘발성 저장 매체를 포함할 수 있다. 예를 들어, 메모리는 ROM(read only memory) 및 RAM(random access memory)를 포함할 수 있다. 본 기재의 실시예에서 메모리는 프로세서의 내부 또는 외부에 위치할 수 있고, 메모리는 이미 알려진 다양한 수단을 통해 프로세서와 연결될 수 있다. 메모리는 다양한 형태의 휘발성 또는 비휘발성 저장 매체이며, 예를 들어, 메모리는 읽기 전용 메모리(read-only memory, ROM) 또는 랜덤 액세스 메모리(random access memory, RAM)를 포함할 수 있다.

[0051] 따라서, 본 발명의 실시예는 컴퓨터에 구현된 방법으로서 구현되거나, 컴퓨터 실행 가능 명령이 저장된 비일시적 컴퓨터 판독 가능 매체로서 구현될 수 있다. 한 실시예에서, 프로세서에 의해 실행될 때, 컴퓨터 판독 가능 명령은 본 기재의 적어도 하나의 양상에 따른 방법을 수행할 수 있다.

[0052] 통신 장치(620)는 유선 신호 또는 무선 신호를 송신 또는 수신할 수 있다.

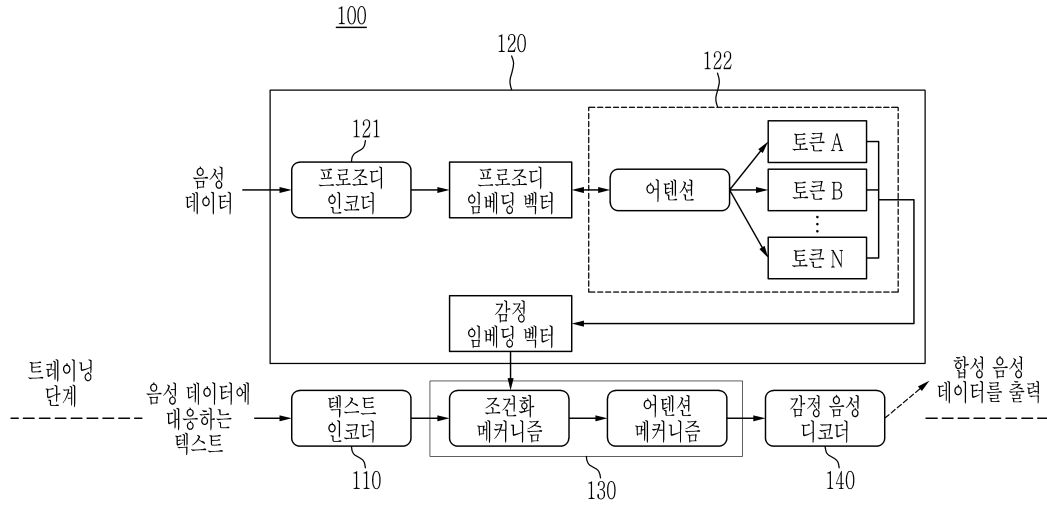
[0053] 한편, 본 발명의 실시예는 지금까지 설명한 장치 및/또는 방법을 통해서만 구현되는 것은 아니며, 본 발명의 실시예의 구성에 대응하는 기능을 실현하는 프로그램 또는 그 프로그램이 기록된 기록 매체를 통해 구현될 수도 있으며, 이러한 구현은 상술한 실시예의 기재로부터 본 발명이 속하는 기술 분야의 통상의 기술자라면 쉽게 구현할 수 있는 것이다. 구체적으로, 본 발명의 실시예에 따른 방법(예, 네트워크 관리 방법, 데이터 전송 방법, 전송 스케줄 생성 방법 등)은 다양한 컴퓨터 수단을 통해 수행될 수 있는 프로그램 명령 형태로 구현되어, 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은, 본 발명의 실시예를 위해 특별히 설계되어 구성된 것이거나, 컴퓨터 소프트웨어 분야의 통상의 기술자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체는 프로그램 명령을 저장하고 수행하도록 구성된 하드웨어 장치를 포함할 수 있다. 예를 들어, 컴퓨터 판독 가능 기록 매체는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광 기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 롬(ROM), 램(RAM), 플래시 메모리 등일 수 있다. 프로그램 명령은 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라, 인터프리터 등을 통해 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함할 수 있다.

[0054] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고

다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

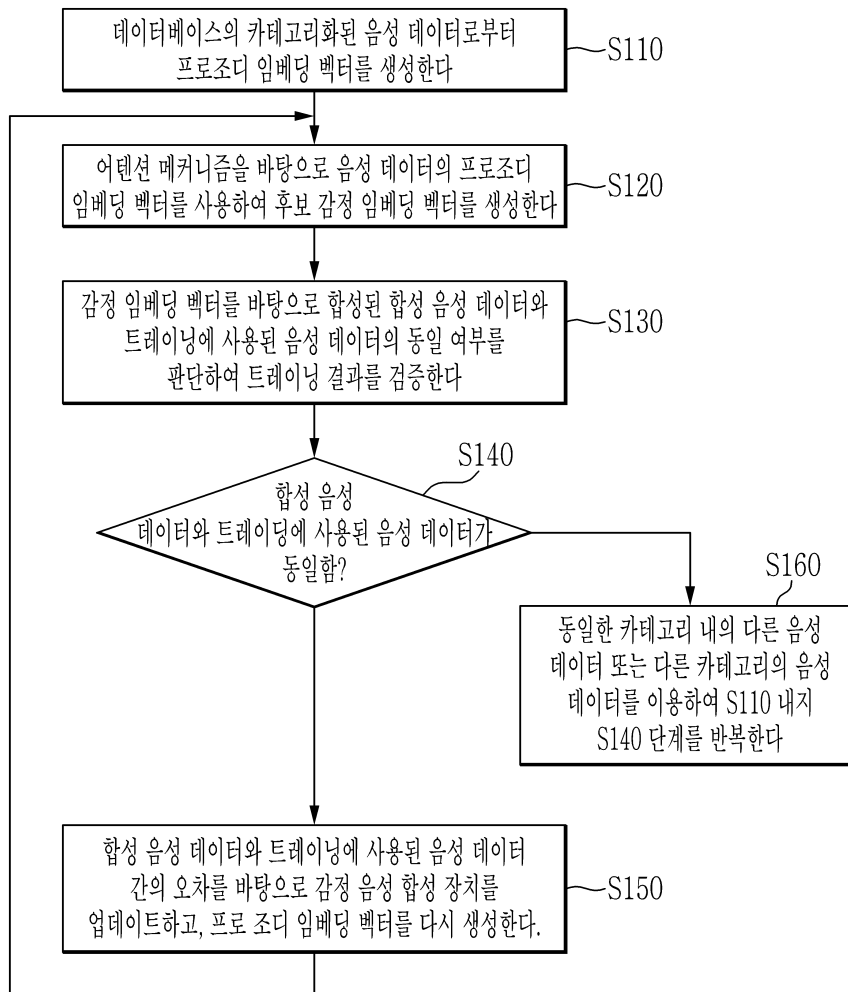
## 도면

### 도면1

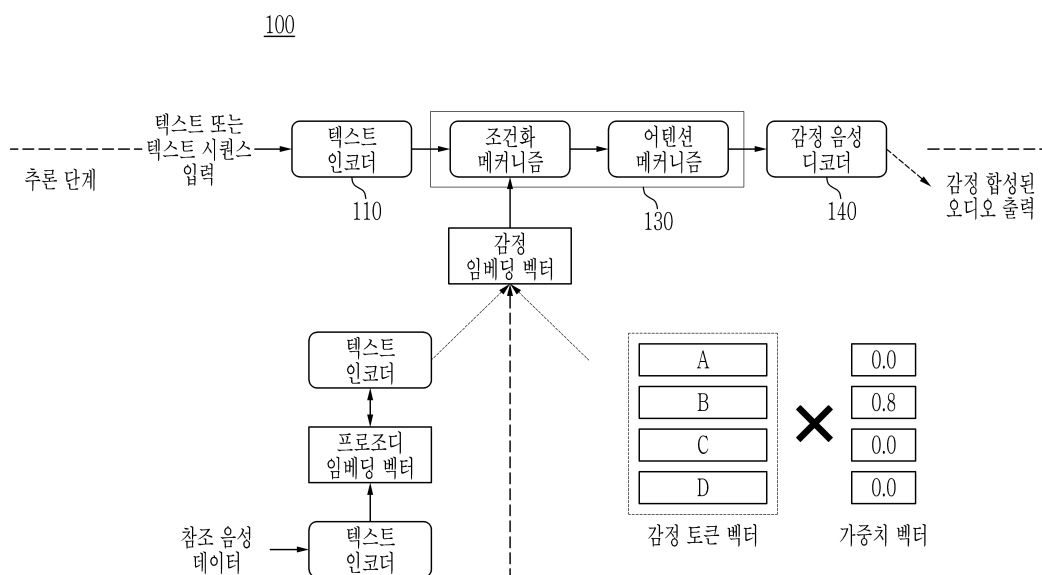




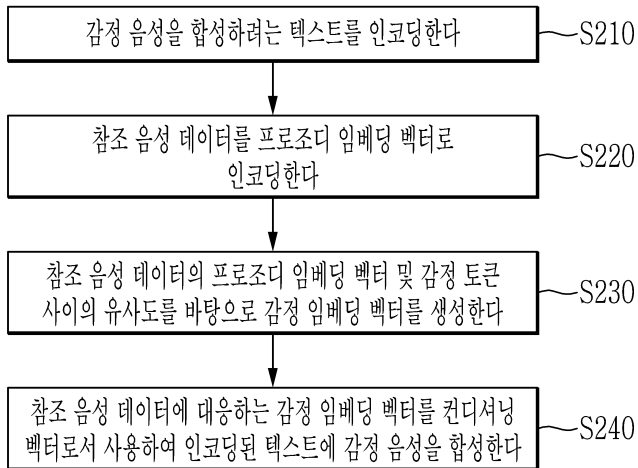
도면2



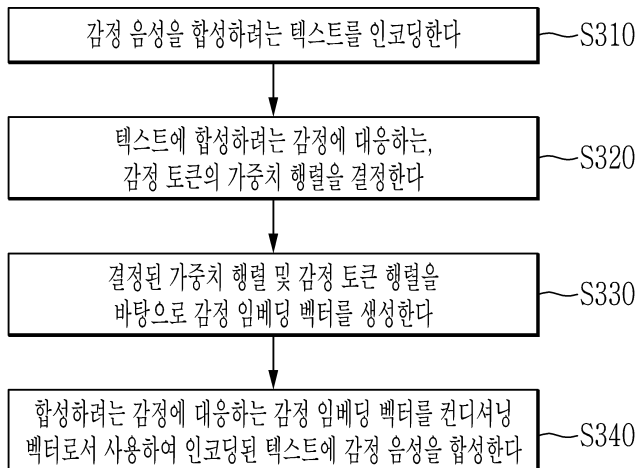
도면3



도면4



도면5



도면6

