



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년07월15일

(11) 등록번호 10-2685135

(24) 등록일자 2024년07월10일

(51) 국제특허분류(Int. Cl.)

H04N 21/854 (2011.01) G06F 16/732 (2019.01)

G06F 16/738 (2019.01) G06F 16/783 (2019.01)

G06F 40/211 (2020.01) G06N 3/08 (2023.01)

(52) CPC특허분류

H04N 21/854 (2013.01)

G06F 16/7343 (2019.01)

(21) 출원번호 10-2022-0015488

(22) 출원일자 2022년02월07일

심사청구일자 2022년02월07일

(65) 공개번호 10-2023-0119398

(43) 공개일자 2023년08월16일

(56) 선행기술조사문헌

KR1020150116929 A*

KR1020200075114 A*

KR102155768 B1

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

피아스페이스 주식회사

대구광역시 달성군 구지면 국가산단서로71길 129

(72) 발명자

이성준

서울특별시 관악구 낙성대로 37

유현수

서울특별시 강동구 고덕로 360, 고덕아르테온 30
4동 2402호

(74) 대리인

특허법인스마트

전체 청구항 수 : 총 2 항

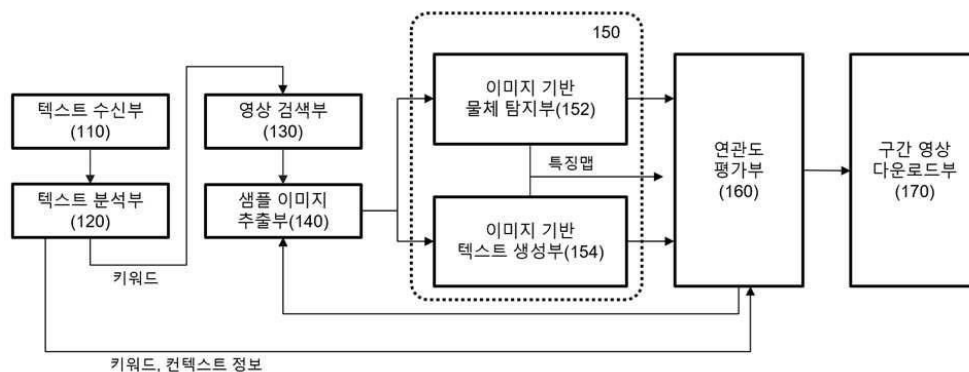
심사관 : 황은택

(54) 발명의 명칭 영상 편집 자동화 시스템

(57) 요약

본 발명은 영상 편집 자동화 시스템에 관한 것으로 보다 구체적으로는 일반 사용자가 동영상 콘텐츠 편집 또는 영상 편집을 보다 쉽고 편리하고 제작할 수 있도록 기능하는 영상 편집 자동화 시스템에 관한 것이다. 본 발명의 일양태에 따르면, 웹 상에서 영상 콘텐츠를 자동으로 검색하고 편집할 수 있는 영상 편집 자동화 시스템이 제공 되고, 이 시스템은, 사용자로부터 입력된 텍스트를 입력받는 텍스트 수신부, 입력된 텍스트를 분석하는 텍스트 분석부, 텍스트 분석부에서 분석된 키워드 결과에 기반하여 웹 상에서 영상을 검색하는 영상 검색부, 검색된 영상으로부터 메타데이터 및 샘플 이미지를 추출하는 샘플 이미지 추출부, 추출된 샘플 이미지를 분석하는 이미지 분석부, 분석된 이미지 정보에 기반하여 사용자 입력 텍스트와의 연관도를 분석하는 연관도 평가부 및 연관도 평가부의 평가 결과에 대응하여 해당하는 영상을 구간별로 다운로드 하는 구간 영상 다운로드부를 포함한다.

대표도



(52) CPC특허분류

G06F 16/738 (2019.01)

G06F 16/7837 (2019.01)

G06F 40/211 (2020.01)

G06N 3/08 (2023.01)

명세서

청구범위

청구항 1

사용자가 입력한 텍스트로 이루어진 문장, 파일 업로드를 통해 수신된 텍스트 파일, 키워드들의 집합, 또는 음성인식된 텍스트를 수신하는 텍스트 수신부(110),

입력된 텍스트 문장을 분석하여 문장별로 키워드를 추출하는 텍스트 분석부(120),

텍스트 분석부(120)는 딥러닝 기반 텍스트 마이닝 모델을 포함하고,

텍스트 분석부(120)에서 분석된 키워드 결과에 기반하여 웹 상에서 영상을 검색하는 영상 검색부(130);

검색된 영상으로부터 메타데이터 및 샘플 이미지를 추출하는 샘플 이미지 추출부(140);

추출된 샘플 이미지를 분석하는 이미지 분석부(150);

이미지 분석부(150)는 이미지 기반 물체 탐지부(152) 및 이미지 기반 텍스트 생성부(154)를 포함하고,

이미지 기반 물체 탐지부(152)는 이미지에 포함된 객체를 딥러닝 분석을 통해 객체의 종류, 크기, 숫자, 위치를 식별하는 딥러닝 기반 객체 감지 모듈을 포함하고,

이미지 기반 텍스트 생성부(154)는 이미지에 포함된 객체 또는 객체를 포함하는 상황을 딥러닝 분석을 통해 상황 정보 텍스트를 출력하는 딥러닝 기반 이미지 캡션 모듈을 포함하고,

분석된 이미지 정보에 기반하여 사용자 입력 텍스트와의 연관도를 분석하는 연관도 평가부(160),

연관도 평가부(160)는

- 이미지 기반 물체 탐지부(152)에서 의해 탐지된 물체의 종류, 숫자, 크기의 정보를 포함하는 제1 입력 정보;
- 이미지 기반 텍스트 생성부(154)에서 생성된 상황 정보 텍스트를 포함하는 제2 입력 정보;
- 이미지 기반 물체 탐지부(152)와 이미지 기반 텍스트 생성부(154)에 사용된 딥러닝 모델들의 중간 단계에서 얻어지는 이미지 특징맵(feature map)을 포함하는 제3 입력 정보;
- 텍스트 분석부(120)의 텍스트 마이닝 모델에서 추출된 키워드 정보를 포함하는 제4 입력 정보;
- 텍스트 분석부(120)의 컨텍스트 분석 모델에서 추출된 컨텍스트(문맥) 정보를 포함하는 제5 입력 정보를 이용하여 사용자가 입력한 텍스트 문장과 이미지 분석부(150)에서 분석된 분석 결과 사이에서의 연관도를 평가하도록 기능하고,

연관도 평가부(160)의 평가 결과에 대응하여 해당하는 영상을 구간별로 다운로드 하는 구간 영상 다운로드부(170)를 포함하는 것을 특징으로 하는 영상 편집 자동화 시스템.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

제1항에 있어서,

연관도 평가부(160)는

제1 입력 정보 내지 제3 입력 정보를 조합하여 비교 벡터 A를 생성하고; 제4 입력 정보 및 제5 입력 정보를 조합하여 비교 벡터 B를 생성하고; 비교 벡터 A와 비교 벡터 B 사이의 유사도 점수를 계산하도록 기능하고,

유사도 점수는 자연어 간 유사도를 평가하는 Word2Vec 모델 또는 이를 변형한 모델이 이용되는 것을 특징으로 하는 영상 편집 자동화 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 영상 편집 자동화 시스템에 관한 것으로 보다 구체적으로는 일반 사용자가 동영상 콘텐츠 편집 또는 영상 편집을 보다 쉽고 편리하고 제작할 수 있도록 기능하는 영상 편집 자동화 시스템에 관한 것이다.

배경 기술

[0003] 최근 IT 기술과 멀티미디어 기술의 비약적인 발전에 따라, 전문가 또는 업체가 아닌 일반 개인들도 직접 다양한 종류의 사진 및 동영상을 촬영하여 개인 디지털 콘텐츠를 제작하고, 이렇게 제작한 디지털 콘텐츠를 편집하여 각종 네트워크에 개설된 서버에 업로드함으로써 다른 이들에게 배포, 공유 및 판매 등의 행위가 가능하게 되었다.

[0005] 그러나, 이러한 기술적 발전에도 불구하고, 일반 개인들은 제작한 디지털 콘텐츠 들을 편집없이 원본을 그대로 공유하는 형태가 대부분이며, 이에 디지털 콘텐츠로서 상품가치를 부여하지는 못하는 실정이다. 대다수의 상업화된 디지털 콘텐츠는 편집관련 전문적인 지식을 가진 각 방송사나 콘텐츠 업체에 의해 생산 및 서비스되는 것이 일반적이며, 개인들은 단순히 시청자의 입장에서 제공되는 동영상을 수동적으로 감상하는 것이 현실이다.

[0007] 특히 일반인 사용자가 디지털 영상 콘텐츠를 제작함에 있어서, 디지털 영상 콘텐츠에 이용하고자 하는 영상 클립 등을 인터넷 등에 산재되어 있는 웹 페이지로부터 검색해 내고 이를 개별적으로 리뷰하여 원하는 영상을 추출해 내는 일은 많은 노력과 시간이 소요됨에 따라 일반인 사용자가 고품질의 디지털 영상 콘텐츠를 제작하는 것은 용이하지 못하는 문제점이 존재하게 된다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) 대한민국 공개특허 10-2020-0044435호(2020년 4월 29일)

발명의 내용

해결하려는 과제

[0010] 본 발명은 전술한 문제점에 기반하여 안출된 발명으로서 일반인 사용자 또는 개인 사용자가 보다 쉽고 편리하고 디지털 영상 콘텐츠를 제작할 수 있도록 사용자가 희망하는 주제 또는 사용자가 제작하고자 하는 디지털 콘텐츠의 스토리라인에 맞추어 인터넷 등에 산재되어 있는 영상 중 적합한 영상을 사용자에게 자동으로 제공할 수 있는 영상 편집 자동화 시스템을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0012] 전술한 과제를 해결하기 위해 본 발명의 일양태에 따르면, 웹 상에서 영상 콘텐츠를 자동으로 검색하고 편집할 수 있는 영상 편집 자동화 시스템이 제공되고, 이 시스템은, 사용자로부터 입력된 텍스트를 입력받는 텍스트 수신부, 입력된 텍스트를 분석하는 텍스트 분석부, 텍스트 분석부에서 분석된 키워드 결과에 기반하여 웹 상에서

영상을 검색하는 영상 검색부, 검색된 영상으로부터 메타데이터 및 샘플 이미지를 추출하는 샘플 이미지 추출부, 추출된 샘플 이미지를 분석하는 이미지 분석부, 분석된 이미지 정보에 기반하여 사용자 입력 텍스트와의 연관도를 분석하는 연관도 평가부 및 연관도 평가부의 평가 결과에 대응하여 해당하는 영상을 구간별로 다운로드 하는 구간 영상 다운로드부를 포함한다.

[0014] 전술한 양태에서 텍스트 수신부는 사용자가 입력한 텍스트로 이루어진 문장, 파일 업로드를 통해 수신된 텍스트 파일, 키워드들의 집합, 또는 음성인식된 텍스트를 수신하도록 구성되고, 텍스트 분석부는 입력된 텍스트 문장을 분석하여 문장별로 키워드를 추출하도록 구성되고, 텍스트 분석부는 딥러닝 기반 텍스트 마이닝 모델을 포함한다.

[0016] 또한 전술한 어느 하나의 양태에서 이미지 분석부는 이미지 기반 물체 탐지부 및 이미지 기반 텍스트 생성부를 포함하고, 이미지 기반 물체 탐지부는 이미지에 포함된 객체를 딥러닝 분석을 통해 객체의 종류, 크기, 숫자, 위치를 식별하는 딥러닝 기반 객체 감지 모듈을 포함하고, 이미지 기반 텍스트 생성부는 이미지에 포함된 객체 또는 객체를 포함하는 상황을 딥러닝 분석을 통해 상황 정보 텍스트를 출력하는 딥러닝 기반 이미지 캡션 모듈을 포함한다.

[0018] 또한 전술한 어느 하나의 양태에서 연관도 평가부는 텍스트 분석부에서 분석된 문맥(컨텍스트) 정보, 이미지 기반 물체 탐지부에서 분석된 이미지 감지 결과, 이미지 기반 텍스트 생성부에서 생성된 상황 정보 텍스트를 이용하여, 사용자가 입력한 텍스트 문장과 이미지 분석부에서 분석된 분석 결과 사이의 연관도를 평가하도록 기능한다.

[0020] 또한 전술한 어느 하나의 양태에서 연관도 평가부는

[0021] - 이미지 기반 물체 탐지부에서 의해 탐지된 물체의 종류, 숫자, 크기의 정보를 포함하는 제1 입력 정보;

[0022] - 이미지 기반 텍스트 생성부에서 생성된 상황 정보 텍스트를 포함하는 제2 입력 정보;

[0023] - 이미지 기반 물체 탐지부와 이미지 기반 텍스트 생성부에 사용된 딥러닝 모델들의 중간 단계에서 얻어지는 이미지 특징맵(feature map)을 포함하는 제3 입력 정보;

[0024] - 텍스트 분석부의 텍스트 마이닝 모델에서 추출된 키워드 정보를 포함하는 제4 입력 정보;

[0025] - 텍스트 분석부의 컨텍스트 분석 모델에서 추출된 컨텍스트(문맥) 정보를 포함하는 제5 입력 정보를 이용하여 사용자가 입력한 텍스트 문장과 이미지 분석부에서 분석된 분석 결과 사이의 연관도를 평가하도록 기능한다.

[0027] 또한 전술한 어느 하나의 양태에서 연관도 평가부는 제1 입력 정보 내지 제3 입력 정보를 조합하여 비교 벡터 A를 생성하고; 제4 입력 정보 및 제5 입력 정보를 조합하여 비교 벡터 B를 생성하고; 비교 벡터 A와 비교 벡터 B 사이의 유사도 점수를 계산하도록 기능하고, 여기서 유사도 점수는 자연어 간 유사도를 평가하는 Word2Vec 모델 또는 이를 변형한 모델이 이용될 수 있다.

발명의 효과

[0029] 본 발명의 실시예에 따르면 사용자는 웹 페이지 화면에 제공된 텍스트 입력창에 원하는 텍스트 문장 또는 검색어 세트를 단순히 입력하면 텍스트 문장 또는 검색어 세트에 적합한 구간 영상이 자동으로 사용자의 편집 작업 템플릿 상에 제공됨에 따라 사용자는 웹상에서 원하는 영상 정보를 취득하고자 수많은 영상을 검색하지 않고 보다 쉽고 편리하게 관련된 구간 영상을 취득할 수 있어 디지털 영상 콘텐츠 제작을 보다 빠르고 고품질로 제작할 수 있다.

도면의 간단한 설명

[0031] 도 1은 본 발명의 실시예에 따른 영상 편집 자동화 시스템의 일례를 나타낸 도면;

도 2는 본 발명에 따른 영상 편집 자동화 시스템의 웹 서버에서 사용자에게 제공되는 화면 구성의 일례를 나타내는 도면;

도 3은 본 발명에 따른 영상 편집 자동화 시스템의 웹 서버에서 사용자에게 제공되는 화면 구성의 다른 일례를 나타내는 도면;

도 4은 본 발명에 따른 영상 편집 자동화 시스템의 내부 구조를 기능별로 나타낸 도면; 및

도 5는 본 발명에 따른 영상 편집 자동화 시스템의 동작 흐름을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0032] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되는 실시예를 참조하면 명확해질 것이다. 그러나, 본 발명은 이하에서 개시되는 실시예에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이다.
- [0034] 본 명세서에서 본 실시예는 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이다. 그리고 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 따라서, 몇몇 실시예들에서, 잘 알려진 구성 요소, 잘 알려진 동작 및 잘 알려진 기술들은 본 발명이 모호하게 해석되는 것을 피하기 위하여 구체적으로 설명되지 않는다.
- [0036] 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다. 그리고, 본 명세서에서 사용된(언급된) 용어들은 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 또한, '포함(또는, 구비)한다'로 언급된 구성 요소 및 동작은 하나 이상의 다른 구성요소 및 동작의 존재 또는 추가를 배제하지 않는다.
- [0038] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 정의되어 있지 않은 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0040] 이하 첨부된 도면을 참조하여 본 발명의 실시예에 대해 설명하도록 한다.
- [0042] 도 1은 본 발명의 실시예에 따른 영상 편집 자동화 시스템을 나타낸 도면이다. 도 1에 도시된 바와 같이, 본 발명의 영상 편집 자동화 시스템(100)에는 네트워크를 통해 하나 이상의 사용자 단말기와 연결되며, 사용자 단말기의 요청에 따라, 웹 페이지상에서 선택가능한 메인메뉴 및 그 메인메뉴의 선택에 대응하여 결정되는 템플릿을 제공하며, 또한 템플릿을 이용한 디지털 콘텐츠를 생성하는 편집툴을 상기 사용자 단말기에 제공하도록 구성된다.
- [0044] 사용자는 사용자 단말을 이용하여 네트워크를 통해 영상 편집 자동화 시스템에 접속하고, 시스템에서 제공하는 툴을 이용하여 동영상を検査하고, 검색된 동영상을 편집하고, 디지털 콘텐츠를 생성하도록 하는 컴퓨팅 장치로서, 특히 웹 브라우저 기반으로 제공되는 영상 편집 서비스를 수행할 수 있는 Internet Explorer, Firefox 및 Chrome 등과 같은 웹 브라우저(Web Browser)가 실행가능한 휴대용 단말기, PC, 랩탑 또는 워크스테이션 등이 이용될 수 있다.
- [0046] 따라서, 사용자 단말기에는 연산장치, 저장장치, 입력장치 및 출력장치 등이 구비될 수 있으며, 웹 브라우저를 통해 편집툴을 실행하여 온라인 상에서 디지털 영상 콘텐츠 제작을 수행하게 된다.
- [0048] 여기서, 상기 네트워크는 WiFi 및 Zigbee와 같은 근거리 무선 통신망뿐만 아니라, CDMA, WCDMA, Wibro, 3G 및 LTE 와 같은 원거리 무선 통신망을 포함하는 개념으로서, 서로 동일하거나 별개의 망일 수 있다.
- [0050] 영상 편집 자동화 시스템(100)은 각 사용자 단말기로부터 입력되는 텍스트에 기반하여 인터넷 상에서 사진 및 동영상 등과 같은 소스를 자동으로 검색하고, 검색된 사진 및 동영상을 사용자 화면 인터페이스에 제공하여 사용자가 보다 쉽게 디지털 콘텐츠를 생성할 수 있도록 기능하는 고사양의 마이크로프로세서 및 대용량의 저장공간을 구비하는 컴퓨팅 장치, 예를 들면 웹 서버일 수도 있다.
- [0052] 도 2는 사용자가 영상 편집 자동화 시스템(100)에 웹 브라우저를 통해 접속하였을 때 사용자에게 제공되는 웹페이지 화면의 일례를 나타낸 도면이다. 도 2에 도시된 바와 같이 사용자의 웹 브라우저 화면의 일부분, 도시된 도면에서 화면의 좌측에 사용자로부터 텍스트(또는 문장)를 입력받기 위한 텍스트 입력창이 제공되고, 웹 브라우저 화면의 다른 부분, 도시된 도면에서 화면의 우측에는 사용자가 입력한 텍스트(또는 문장)에 대응하는 검색된 동영상 클립이 나열되는 검색 결과창(1520)이 제공된다.
- [0054] 이와 같은 웹 페이지 화면에서는 사용자는 텍스트 입력창(1501)에 검색을 희망하는 키워드 또는 문장을 입력하고 검색 버튼을 누르거나, 파일 업로드 기능을 이용하여 사용자 단말기에 저장된 텍스트 파일을 업로드 한 후 검색 버튼을 눌러 검색을 수행하면, 영상 편집 자동화 시스템(100)은 사용자로부터 입력된 문장을 분석하여 키워드를 추출하거나, 사용자로부터 입력된 키워드를 이용하여 인터넷 웹으로부터 영상을 검색하고 검색된 영상 중 연관도가 높은 순서로 영상을 사용자의 화면에 제공한다.

- [0056] 도 3은 사용자가 영상 편집 자동화 시스템(100)에 웹 브라우저를 통해 접속하였을 때 사용자에게 제공되는 웹페이지 화면의 다른 일례를 나타낸 도면이다. 도 3에 도시도니 바와 같이 사용자의 웹 브라우저 화면은 복수의 템플릿(Ta~Td)으로 구성되고 사용자는 제작하고자 하는 동영상 콘텐츠의 스토리라인에 맞게 동영상 클립을 검색할 수 있다.
- [0058] 예를 들면, 이와 같이 제공된 웹 페이지화면에서 각각의 템플릿 별로, 텍스트 입력창에 검색을 희망하는 키워드들 또는 문장을 입력하고 검색 버튼을 누르거나, 파일 업로드 버튼(미도시)을 이용하여 사용자 단말기에 저장된 텍스트 파일을 업로드 한 후 검색 버튼을 눌러 검색을 수행하면, 영상 편집 자동화 시스템(100)은 사용자로부터 입력된 문장을 분석하여 문장별로 키워드 또는 문맥(컨텍스트) 정보를 추출하거나, 사용자로부터 입력된 키워드의 집합으로부터 문맥(컨텍스트) 정보를 추출하고, 문장별로 추출된 키워드 정보 또는 컨텍스트 정보를 이용하여 인터넷 웹으로부터 영상을 검색하고 검색된 영상들 중 연관도가 높은 순서로 사용자의 화면에 제공한다.
- [0060] 도 4는 전술한 바와 같은 영상 편집 자동화 서비스를 제공하는 영상 편집 자동화 시스템의 내부 구조를 기능별로 도시한 블록도이고, 도 5는 영상 편집 자동화 시스템에서의 처리 절차를 개략적으로 도시한 블록도이다. 도 4 및 도 5를 참조하여 본 발명의 내부 구성 요소 및 내부 구성 요소에서 수행되는 기능에 대해 순차적으로 설명한다. 도시된 바와 같이 본 발명에 따른 영상 편집 자동화 시스템(100)은 사용자로부터 입력된 텍스트를 입력받는 텍스트 수신부(110), 입력된 텍스트를 분석하는 텍스트 분석부(120), 텍스트 분석부(120)에서 분석된 결과(키워드 및/또는 컨텍스트 정보)에 기반하여 웹으로부터 영상을 검색하는 영상 검색부(130); 검색된 영상으로부터 메타데이터 및 샘플 이미지를 추출하는 샘플 이미지 추출부(140); 추출된 샘플 이미지를 분석하는 이미지 분석부(150); 분석된 이미지 정보에 기반하여 사용자 입력 텍스트와의 연관도를 분석하는 연관도 평가부(160); 및 연관도 평가부(160)의 평가 결과에 대응하여 해당하는 영상을 구간별로 다운로드 하는 구간 영상 다운로드부(170)를 포함한다.
- [0062] 텍스트 수신부(110)는 도 2 및 도 3에 도시된 바와 같이 화면 구성에서 입력된 텍스트로 이루어진 문장, 파일 업로드를 통해 수신된 텍스트 파일, 또는 키워드 집합을 수신하도록 구성된다. 도시되지 않았지만 본 발명은 이에 한정되는 것은 아니고 사용자 단말로부터 음성을 입력받는 경우 음성인식 기능을 이용하여 사용자가 입력한 음성을 텍스트로 전환하고 전환된 텍스트가 텍스트 수신부(110)에서 수신되도록 구성될 수도 있다.
- [0064] 텍스트 수신부(110)로부터 수신된 텍스트로 이루어진 문장은 텍스트 분석부(120)로 전달되고, 텍스트 분석부(120)는 입력된 텍스트를 분석하여 주요 키워드를 추출하도록 구성된다. 이와 같은 키워드 추출을 위해 텍스트 분석부(120)는 인공지능기반의 텍스트 마이닝 모델을 포함할 수 있다.
- [0066] 텍스트 분석부(120)는 수신된 텍스트 문장으로부터 키워드를 추출하기 위해 텍스트 전처리를 수행한다. 텍스트의 경우 사람이 인지할 수 있는 국문, 영문, 숫자 등 자연 형태의 언어로 구성되어 있다. 텍스트 마이닝을 통한 키워드 제공을 위하여 입력된 텍스트는 컴퓨터가 이해할 수 있는 형태로 표현해야 하는데, 이에 따른 제반 기술을 자연어처리(Natural Language Processing, NLP)라 한다.
- [0068] 자연어는 컴퓨터 프로그램 언어와 같은 인공 언어와 구별된다. 인공 언어는 인간이 규칙을 정해 컴퓨터가 이해할 수 있도록 만들어진 언어이기 때문에, 누구나 그 규칙을 배우면 이해할 수 있게 되지만 자연어는 인간이 규칙을 정하긴 했지만, 오랜 시간이 흐르면서 관습적 규칙, 비형식적 규칙 등 다양한 변화가 가미됨으로 인해 규칙만으로 이해될 수 없는 특성이 있다. 자연어 처리 과정은 형태소 분석, 통사 분석, 의미 분석, 화용 분석의 4개 단계로 구분할 수 있으며, 각 단계는 특정이론을 통해 명확한 이론이 확립된 단계가 아니라 지속해서 이론이 제시되고 검증되는 발전단계에 있다.

[0070] - 문자열의 처리

전처리 방법	사유
특수문자(special character) 제거	문서의 의미와 무관한 특수문자 제거
각 기사내 줄 바꿈 문자("\r", "\n")는 공백으로 대체하여 기사당 하나의 라인(line)으로 변환	gensim의 word2vec 라이브러리 [20]를 사용하기 위해서는 하나의 문서가 하나의 라인으로 표현되어야 함
영어를 제외한 외국 문자 제거	영어를 제외한 외국어가 분류에 큰 영향을 미치지 못할 것으로 판단.

[0071]

[0073] - 자동 띄어쓰기 처리

[0074] 텍스트 분석에서 가장 기초적인 작업은 텍스트로부터 단어를 식별하고 추출하는 토큰화(Tokenization)라고 할 수 있다. 한국어는 '어절(語節)'로 토큰의 단위 기준으로 보고, 중국어나 일본어와 같이 어절 경계 표시가 없는 언어와는 달리, 어절과 어절 사이에 공백을 두어 띄어쓰기를 하도록 규정하였다. 한국어에 있어서 잘못된 띄어쓰기는 중의성(ambiguity)을 유발시키거나 텍스트분석에서 잡음(noise)을 일으켜 오히려 토큰화를 방해하며, 가독성을 떨어뜨린다. 이와 같이 한국어에서 띄어쓰기는 텍스트에 대한 사용자 가독성만큼이나 기계 가독성에도 영향을 주는 중요한 요소이다. 문장 내의 띄어쓰기 오류는 많은 문법적, 의미적 모호성을 일으키며, 때로는 형태소 분석을 불가능하게 만들기도 한다.

[0076] 본 발명에서는 사용자가 띄어쓰기를 고려하지 않고 텍스트를 입력하는 경우도 있는 것으로 가정하고, 자동 띄어쓰기를 적용하여 전처리를 수행한다.

[0078] - WPM(Word Piece Model) 처리

[0079] WPM은 음성 검색 시스템 구축을 위한 방법으로 언어에 대한 사전 지식 없이 혼잡도(perplexity)를 최소로 하여 어휘를 자동 생성하는 방법이다. 기존의 자연어 처리에서는 형태소 분석, 통사 분석, 의미 분석, 화용 분석 등 4가지로 진행되나, WPM은 국제 발음기호(the International Phonetic Alphabet, PA)기반의 발음 세트(Set)로 유닛을 코드화 시킨 후, 통계적 기법을 활용하여 사용 빈도수에 따라 조합을 하여 새로운 유닛을 생성한다. WPM의 장점은 언어에 독립적이며, 통계적인 방식을 사용하므로 특정 도메인 또는 아직 의미가 파악되지 않는 언어에도 적용할 수 있다. 본 발명에서는 WPM 알고리즘을 적용하여 명사와 조사를 분리키시고 조사를 제거하여 명사만을 추출하게 된다.

[0081] 또한 명사와 명사가 연속적으로 출현된 경우에는 합성 명사로서 사용가능한지 여부를 데이터베이스에 저장된 데이터세로부터 검색하고 매칭되는 결과가 발견될 때는 합성명사로서 간주하도록 처리된다.

[0083] 텍스트에 대한 전처리가 수행되고 난 후 전처리된 텍스트에 대해 노이즈 필터링을 수행하게 된다. 여기서 노이즈 필터링이라 함은 미리 정의되어 있는 불용어들에 대해 제외하는 처리를 의미한다. 이 후, 텍스트 문장으로부터 키워드 정보 및/또는 컨텍스트 정보를 추출하기 위해 다음과 같은 처리들이 더 이용될 수도 있다.

[0085] ° 형태소(단어) 분석: 의미의 최소단위인 형태소(morpheme)이나 단어에 대한 분석은 텍스트를 분석하는 가장 보편적인 방법으로, 문법적 규칙 혹은 확률에 의한 품사 태깅(part of speech tagging), 개체명 인식(named entity recognition), 철자 교정, 단어 식별(tokenization) 기법 등을 이용

[0086]

[0087] ° 핵심 어구 추출: 텍스트를 어휘적으로 상호 관련 있는 단어들로 나누는 청킹(chunking)을 통해 핵심 어구를 추출하거나, 개체명 인식, 관계 추출(relation extraction) 등에 활용한다.

[0088]

[0089] ° 벡터 공간 모델링: 벡터 공간 모델에서 문서(document)와 단어(term)를 각각 벡터(vector)와 차원(dimension)에 대응시켜 통계적 가중치를 구함. TF-IDF, 카이제곱 검정, 코사인 유사도 등을 통해 단어의 빈도수 분포에서 중요하지 않은 단어를 걸러내고, 문서의 유사도를 추출함

[0090]

- [0091] ° 동시출현 단어 분석: 단어들 사이의 의미상의 관계성을 파악하기 위해 일정한 문맥 내에서 두 단어가 동시 출현하는 빈도를 구한 후, 다양한 통계적 방법을 활용하여 유의미한 단어 쌍을 추출
- [0092]
- [0093] ° 토픽 모델링: 구조화되지 않은 대량의 텍스트로부터 숨겨져 있는 주제 구조를 발견하고 카테고리화 하기 위한 통계적 추론 알고리즘으로, LDA (Latent Dirichlet Allocation) 모델이 주로 활용됨.
- [0094]
- [0095] ° 의미연결망 분석: 키워드 동시출현 분석 기법에 기반하거나 그 한계를 넘어 단어 간의 의미 혹은 맥락상의 연결 관계를 정의하고, 해당 연결 관계를 시각화하거나 중요한 컨셉을 네트워크 속의 위상(centrality)에 따라 추출
- [0096]
- [0097] 텍스트 분석부(120)를 통해 문장별로 추출된 키워드 세트는 영상 검색부(130)으로 전달되고 영상 검색부(130)는 추출된 키워드를 기반으로 하여 웹크로잉 기술을 사용하여 웹 상에서 영상을 검색하고 검색된 영상의 리스트를 생성한다. 또한 영상 검색부(130)는 영상의 리스트 생성시 각각의 영상에 대한 웹페이지에 영상에 대한 메타데이터가 존재하는 경우 메타데이터를 더 저장하도록 동작한다. 메타 데이터는 영상의 제목, 제작자, 장르, 제작일자와 같은 정보를 포함하지만 본 발명이 전술한 메타 데이터 정보에 의해 한정되는 것은 아니다. 한편, 메타 데이터는 동영상 파일의 헤더 정보가 더 포함될 수도 있다.
- [0099] 영상 검색부(130)로부터 검색된 영상의 리스트가 작성되고 나면 샘플 이미지 추출부(140)는 각각의 영상의 랜덤한 구간에서 소량의 샘플 이미지들을 추출하고 다운로드를 수행한다. 이는 검색된 영상을 모두 다운로드 받는 경우에 비교하여 네트워크 트래픽의 부담을 감소시킬 뿐만 아니라 영상을 다운로드할 수 있는 저장공간의 부담을 감소시키기 위함이다.
- [0101] 한편 전술한 설명에서 영상 검색부(130)은 웹상에서 영상을 검색하는 것으로 설명되었지만 본 발명은 이에 한정되는 것은 아니고 jpg, jpeg, gif, bmp, png 등의 이미지 파일에 대한 검색을 수행하거나 영상 파일 검색에 추가하여 이미지 검색을 더 수행하도록 구성될 수도 있다.
- [0103] 샘플 이미지 추출부(140)에서 추출된 샘플 이미지는 이어서 이미지 분석부(150)로 제공된다. 이미지 분석부(150)는 이미지 기반 물체 탐지부(152) 및 이미지 기반 텍스트 생성부(154)를 통해 추출된 샘플 이미지에 대한 분석을 수행한다.
- [0105] 이미지 기반 물체 탐지부(152)는 이미지에 포함된 물체 또는 객체를 딥러닝 분석을 통해 물체의 종류를 식별하는 컴퓨터 비전 기술을 총칭한다. 이와 같은 맥락에서, 이미지 기반 물체 탐지부(152)는 객체 감지(Object detection) 및 객체 인식(Object recognition) 기술을 포함한다. 객체 감지(Object detection)와 객체 인식(Object recognition)은 서로 유사한 객체 식별 기술이지만, 실행 방식에 있어서 차이점이 있다. 객체 감지는 이미지에서 객체의 인스턴스를 찾아내는 프로세스로 딥러닝의 경우 객체 감지는 이미지에서 객체를 식별할 뿐만 아니라 위치까지 파악되는 객체 인식의 서브셋이다. 이를 통해 하나의 이미지에서 여러 객체를 식별하고 각 위치를 파악할 수 있다. 본 발명에서 이미지 기반 물체 탐지부는 이미지에서 객체를 식별할 뿐만 아니라 위치까지 파악하고 이미지 내에서 여러 객체를 식별할 수 있는 딥러닝 기반 또는 머신러닝 기반의 객체 감지 모듈이 이용되는 것이 바람직하다. 이러한 딥러닝 분석은 복수의 히든 레이어가 포함되는 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network) 등의 심층 신경망 방식이 있고, 이를 개선한 다양한 방식들이 계속하여 소개되고 있으며, 적용 분야에 따라 급격한 성능 향상과 다양한 분석 결과를 제공하고 있다.
- [0107] 이미지 기반 텍스트 생성부(154)는 추출된 이미지에 포함된 객체나 객체를 포함하는 상황을 딥러닝 분석을 통해 텍스트 정보를 출력한다. 이러한 딥러닝 분석은 딥러닝 기반의 객체 감지 모듈과 유사하게, 복수의 히든 레이어가 포함되는 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network) 등의 심층 신경망 방식이 이용될 수 있다. 따라서 이미지 기반 텍스트 생성부(154)는 영상 내 객체의 종류와 위치, 크기에 대한 기본적인 분석 결과는 물론이고 최근에는 객체의 심리 상태, 객체의 일부에 대한 상세한 설명, 상황에 대한 인지 결과 등을 제공할 수 있으며 이미지를 통해서 상당히 많은 상황 정보를 텍스트와 분석 정보(정확도, 가중치, 부가적인 분석 정보 등)로 제공할 수 있게 된다.
- [0109] 연관도 평가부(160)는 이전 설명한 텍스트 분석부(120)에서 분석된 컨텍스트 정보, 이미지 기반 물체 탐지부(152)에서 분석된 이미지 감지 결과, 이미지 기반 텍스트 생성부(154)에서 생성된 상황 정보 텍스트를

이용하여, 사용자가 입력한 텍스트 문장과 이미지 분석부(150)에서 분석된 분석 결과 사이의 연관도를 평가하도록 기능한다.

[0111] 구체적으로 연관도 평가부(160)는 다음과 같은 입력:

[0112] - 입력1: 이미지 기반 물체 탐지부(152)에서 의해 탐지된 물체의 종류, 숫자, 크기의 정보

[0113] - 입력2: 이미지 기반 텍스트 생성부(154)에서 생성된 상황 정보 텍스트

[0114] - 입력3: 이미지 기반 물체 탐지부(152)와 이미지 기반 텍스트 생성부(154)에 사용된 딥러닝 모델들의 중간 단계에서 얻어지는 이미지 특징맵(feature map)

[0115] - 입력4: 텍스트 분석부(120)의 텍스트 마이닝 모델에서 추출된 키워드 정보

[0116] - 입력5: 텍스트 분석부(120)의 컨텍스트 분석 모델에서 추출된 컨텍스트(문맥) 정보

[0117] 을 이용하여 연관도를 분석한다.

[0119] 연관도 평가부(160)는 입력 1 ~ 입력 3를 조합하여 비교 벡터 A를 생성하고, 또한 입력 4 ~ 입력 5를 조합하여 비교 벡터 B를 생성한 후, 비교 벡터 A와 비교 벡터 B 사이의 유사도 점수를 계산하도록 기능한다. 유사도 점수를 평가하는데 있어서 자연어 간 유사도를 평가하는 Word2Vec 모델 또는 이를 변형한 모델이 이용될 수 있다.

[0121] 연관도 평가부(160)는 연관도가 높은 것으로 평가된 이미지가 포함된 영상에 대해 더 많은 이미지 추출을 시도하도록 샘플 이미지 추출부(140)에 피드백을 제공하는 반면, 낮은 연관도의 이미지가 포함된 영상은 제외함으로써 이미지 분석에 소요되는 시간을 단축하고 최종 결과물의 품질을 높일 수 있도록 구성되는 것이 바람직하다. 이와 같이 본 발명에서는 연관도 평가부를 통해 검색된 영상으로부터 추출된 이미지와 사용자 입력 텍스트로부터 얻어진 정보를 연관도를 분석함으로써 영상에 태그정보가 없이 이미지로부터 직접 분석을 수행할 수 있다는 장점이 더 얻어질 수 있다.

[0123] 또한 연관도 평가부(160)는 각각의 영상에 대해 연관도 평가 점수가 부여되고 나면 구간 영상 다운로드부(170)는 연관도가 높은 영상에 대해 구간 영상을 다운로드하고, 다운로드된 구간 영상은 데이터베이스에 저장되는 한편 분석된 결과와 영상에 대한 매칭 정보가 인덱싱 데이터베이스에 함께 저장되어 영상의 재사용성을 높일 수 있다.

[0125] 구간 영상 다운로드부(170)로부터 구간 영상이 획득 또는 다운로드되고 나면, 다운로드된 구간 영상은 사용자 화면의 영상 편집용 템플릿에 제공되거나 선택된 템플릿의 타임라인에 자동으로 추가되고, 사용자는 웹상에서 원하는 영상 정보를 취득하고자 수많은 영상을 검색하지 않고 보다 쉽고 편리하게 관련된 구간 영상을 취득할 수 있고, 또한 사용자가 제작하고자 하는 디지털 콘텐츠의 스토리라인(또는 편집 템플릿)에 맞추어 구간 영상이 사용자에게 자동으로 제공되어 영상 제작에 도움을 줄 수 있다.

[0127] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0129] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로

연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0131] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0133] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

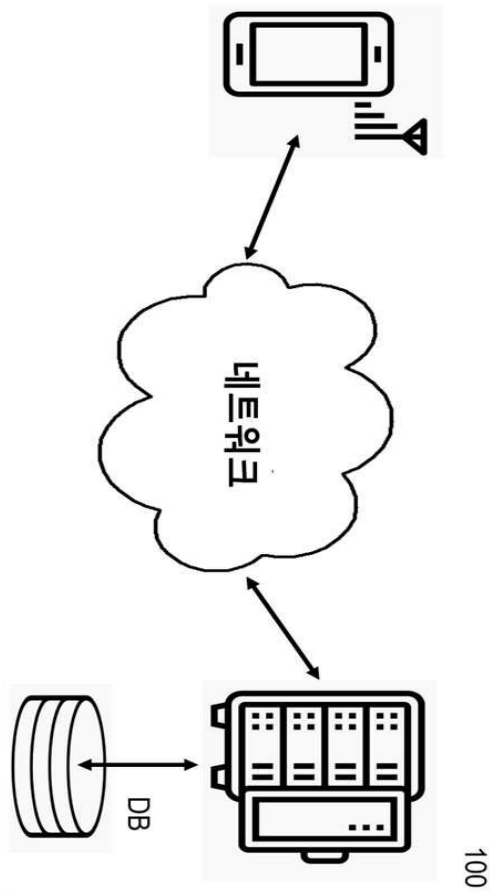
[0135] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속하는 것으로 해석되어야만 한다.

부호의 설명

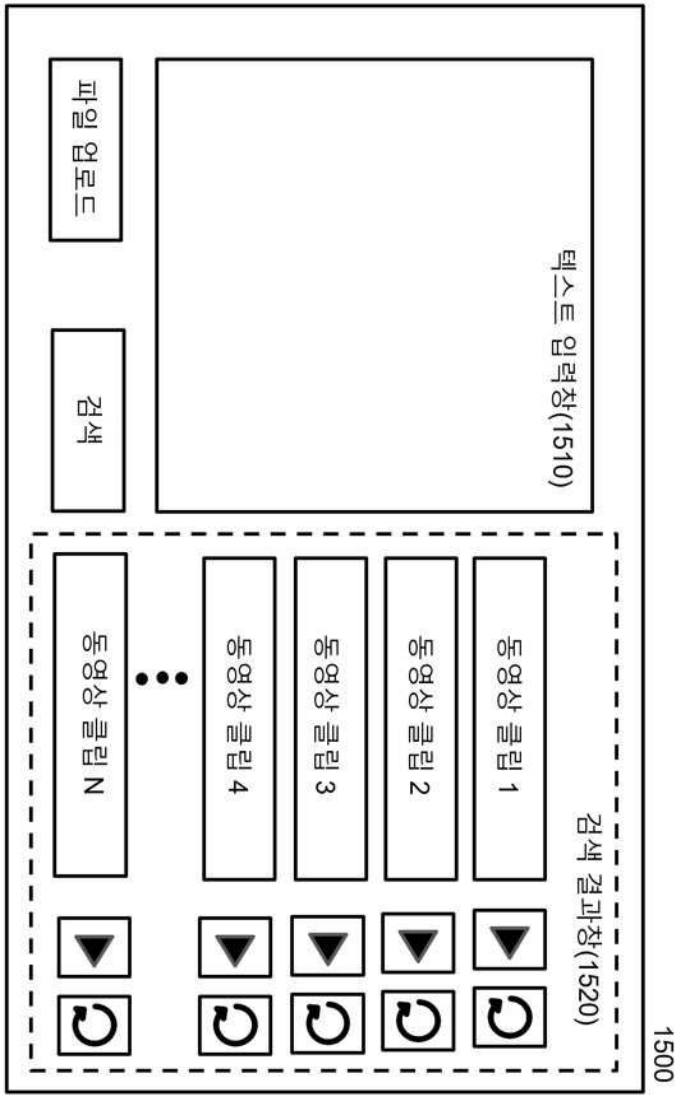
[0137]	100: 영상 편집 자동화 시스템	110: 텍스트 수신부
	120: 텍스트 분석부	130: 영상 검색부
	140: 샘플 이미지 추출부	150: 이미지 분석부
	160: 연관도 평가부	170: 구간 영상 다운로드부
	152: 이미지 기반 물체 탐지부	154: 이미지 기반 텍스트 생성부

도면

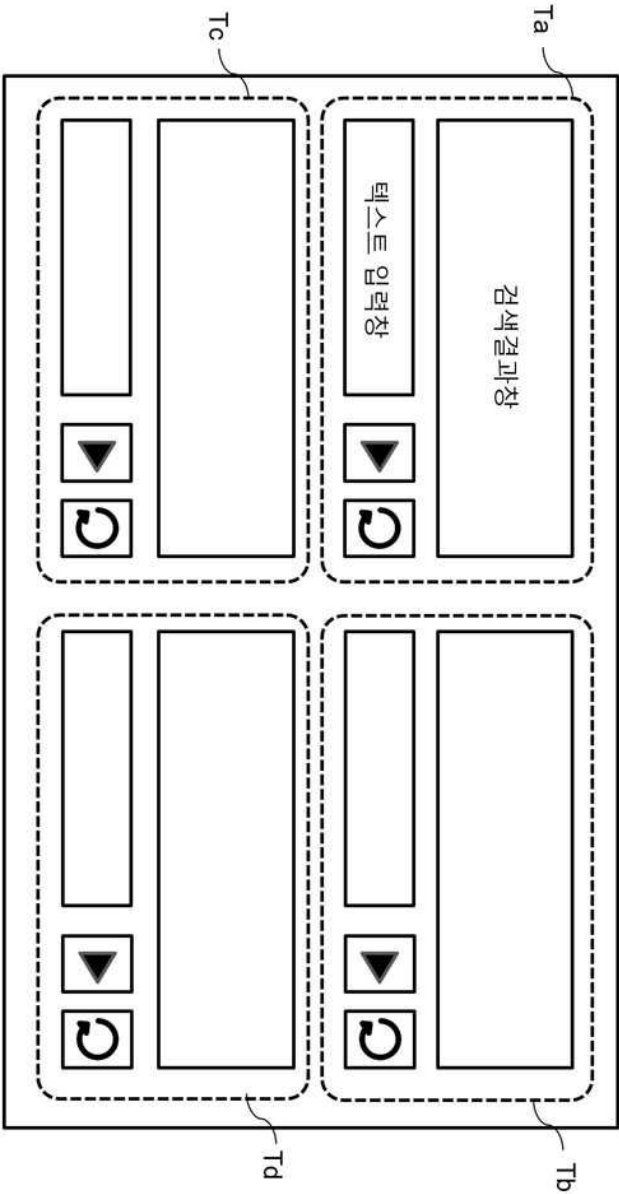
도면1



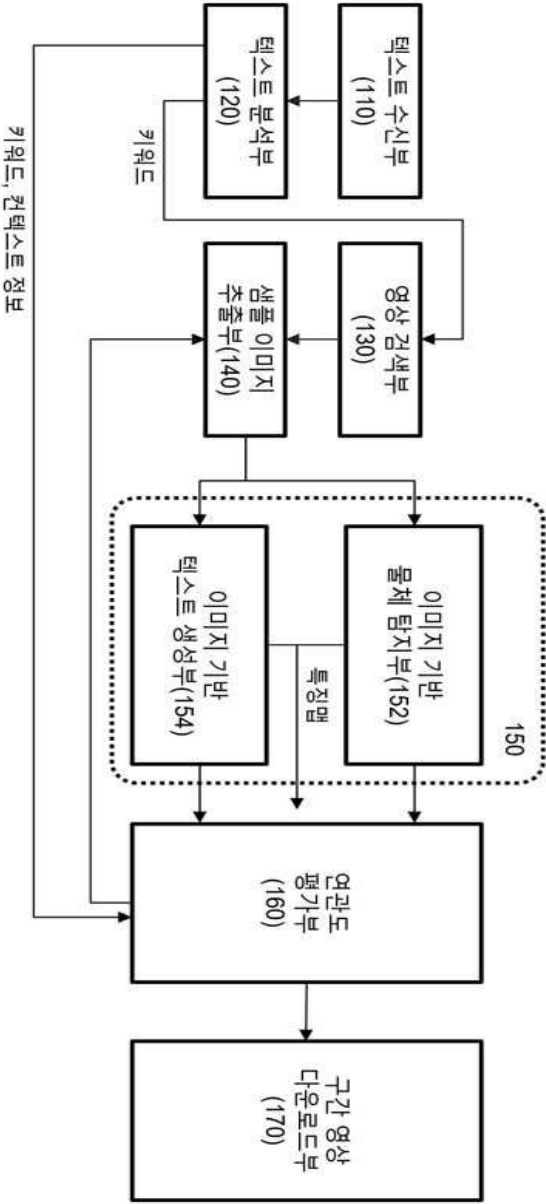
도면2



도면3



도면4



도면5

