(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2020 년 4 월 23 일 (23.04.2020)

WIPO | PCT

(10) 국제공개번호
WO 2020/080606 A1

(51) 국제특허분류:
H04N 21/234 (2011.01) *H04N 21/235* (2011.01)

(21) 국제출원번호: PCT/KR2018/016854

(22) 국제출원일: 2018 년 12 월 28 일 (28.12.2018)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2018-0124863 2018 년 10 월 19 일 (19.10.2018) KR

(71) 출원인: 인하대학교 산학협력단 (**INHA-INDUSTRY PARTNERSHIP INSTITUTE**) [KR/KR]; 22212 인천시 남구 인하로 100, Incheon (KR).

(72) 발명자: 조근식 (**JO, Geun Sik**); 21996 인천시 연수구 해돋이로84번길 10 601동 1302호., Incheon (KR).

(74) 대리인: 양성보 (**YANG, Sungbo**); 06099 서울시 강남구 선릉로125길 14 삼성빌딩 2층 피앤티특허법률사무소, Seoul (KR).

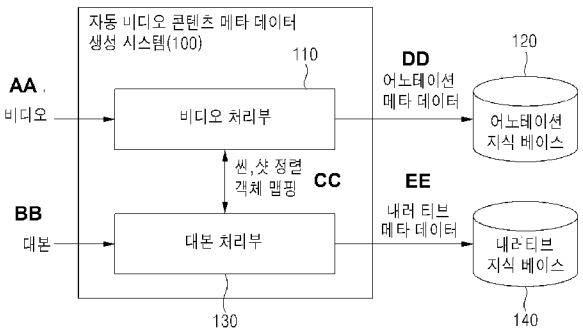
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM,

(54) Title: METHOD AND SYSTEM FOR AUTOMATICALLY GENERATING VIDEO CONTENT-INTEGRATED METADATA USING VIDEO METADATA AND SCRIPT DATA

(54) 발명의 명칭: 비디오 메타데이터와 스크립트 데이터를 활용한 비디오 컨텐츠 통합 메타데이터 자동 생성 방법 및 시스템

[도1]



100 ... Automatic video content metadata generation system
110 ... Video processor
120 ... Annotation knowledge base
130 ... Script processor
140 ... Narrative knowledge base
AA ... Video
BB ... Script
CC ... Scene and shot arrangement, Object mapping
DD ... Annotation metadata
EE ... Narrative metadata

(57) Abstract: Presented are a method and system for automatically generating metadata of video content. The system for automatically generating metadata of video content, presented in the present invention, comprises: a video processor for performing a preprocessing process of receiving a video, and recognizing and arranging shots and scenes of the video so as to be interconnected with a script processor, and analyzing video content; the script processor for performing a preprocessing process of receiving a script, and recognizing and arranging the script and scenes so as to be interconnected with the video processor, and analyzing script content; an annotation knowledge base for storing annotation metadata generated by the video processor; and a narrative knowledge base for storing narrative

WO 2020/080606 A1

[다음 쪽 계속]

ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

metadata generated by the script processor, wherein the annotation knowledge base and the narrative knowledge base share episodes, sequences, scenes, and shots and complementarily use the shared episodes, sequences, scenes, and shots to automatically generate metadata.

(57) 요약서: 비디오 콘텐츠의 메타데이터 자동 생성 방법 및 시스템이 제시된다. 본 발명에서 제안하는 비디오 콘텐츠의 메타데이터 자동 생성 시스템은 대본 처리부와 상호 연결하기 위해 비디오를 입력 받아 비디오의 샷 및 씬을 인식하여 정렬하는 전처리 과정을 수행하고, 비디오 콘텐츠를 분석하는 비디오 처리부, 비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하여 정렬하는 전처리 과정을 수행하고, 대본 콘텐츠를 분석하는 대본 처리부, 상기 비디오 처리부에서 생성된 어노테이션 메타데이터를 저장하는 어노테이션 지식 베이스 및 상기 대본 처리부에서 생성된 내러티브 메타데이터를 저장하는 내러티브 지식 베이스를 포함하고, 어노테이션 지식 베이스와 내러티브 지식 베이스는 에피소드, 시퀀스, 씬 및 샷을 공유함으로써 상호 보완적으로 사용하여 메타데이터를 자동으로 생성한다.

명세서

발명의 명칭: 비디오 메타데이터와 스크립트 데이터를 활용한 비디오 컨텐츠 통합 메타데이터 자동 생성 방법 및 시스템

기술분야

[1] 본 발명은 비디오(Video)와 비디오의 대본(Script)으로부터 비디오 콘텐츠의 다양한 정보를 기술하고, 스토리 이해를 위한 메타데이터(Metadata)를 자동으로 생성하는 방법과 그 시스템에 관한 것이다.

[2] 본 발명은 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터육성지원사업의 연구 결과로부터 도출된 것이다.
[과제관리번호:IITP-2017-0-01642, 과제명: 인공지능을 활용한 콘텐츠 창작 기술]

배경기술

[3] 비디오 콘텐츠의 메타데이터(Metadata)는 비디오 자체의 정보를 기술하는 메타데이터와, 비디오와 관련된 정보를 기술하는 메타데이터로 구분할 수 있다. 예를 들어 MPEG 비디오 규격의 DASH 기술 중 MPD(Media Presentation Description)와 같이 미디어 전송에 적합한 데이터 모델로 비디오의 세그먼트(Segment)정보를 계층적으로 분리하여 표현하는 메타데이터와 MPEG-4 LASeR 또는 MPEG-7와 같이 비디오, 음성과 함께 XML형태로 비디오 콘텐츠와 관련된 부가 정보를 기술하는 객체의 위치, 크기, 등장시간, 관련정보 URL 등을 어노테이션(Annotation) 할 수 있는 메타데이터로 나눌 수 있다.

[4] 이와 같이 비디오의 부가 정보를 기술하는 메타데이터를 사용하는 대표적인 예로, 인터랙티브 비디오(Interactive Video)가 있다. 인터랙티브 비디오는 비디오와 상호 작용 가능한 객체 등을 주로 비디오 위에 증강하여 비디오 시청자가 해당 객체를 클릭하였을 때 비디오와 관련된 부가 정보를 바로 얻고, 쉽게 공유할 수 있는 새로운 형태의 비디오이다. 이와 같은 인터랙티브 비디오를 만들기 위해 시청자가 관심있어 할 만한 비디오와 관련된 정보를 선택하고 그와 관련된 정보를 가져와 비디오의 특정 시점에 다양한 정보를 어노테이션(Annotation)하여 비디오의 메타데이터를 생성한다.

[5] 그러나 이와 같은 비디오 콘텐츠의 메타데이터를 생성하는 데는 많은 시간과 비용이 필요하다. 또한 메타데이터로 사용될 비디오 콘텐츠와 관련된 의미 있는 정보는 비디오만을 이용해 생성하기 어렵다. 예를 들어 비디오에서 등장인물 또는 비디오의 스토리에 중요한 객체가 인식되었다고 해도, 이와 같은 인물, 객체의 정확한 명칭을 지정하거나 인물과 인물, 사건과 사건의 관계를 유추하거나 하는 복잡한 활용은 어렵다. 이를 해결하기 위해 비디오에서 메타데이터를 생성하는 것과 함께 대본을 이용해 메타데이터를 생성하고 비디오 순서와 맞춰주어 다양한 정보를 담는 메타데이터를 자동으로 생성하는 방법이 필요하다.

발명의 상세한 설명

기술적 과제

[6] 본 발명이 이루고자 하는 기술적 과제는 비디오 콘텐츠와 관련된 다양한 정보를 갖는 메타데이터를 자동으로 생성하기 위해, 비디오로부터 다양한 정보를 추출하는 기술, 대본으로부터 의미 있는 정보를 추출하는 기술, 그리고 비디오와 대본의 시간 순서를 정렬하여 메타데이터의 사용성을 높이는 방법과 그 시스템을 제공하는데 있다.

과제 해결 수단

[7] 일 측면에 있어서, 본 발명에서 제안하는 비디오 콘텐츠의 메타데이터 자동 생성 시스템은 대본 처리부와 상호 연결하기 위해 비디오를 입력 받아 비디오의 샷 및 씬을 인식하여 정렬하는 전처리 과정을 수행하고, 비디오 콘텐츠를 분석하는 비디오 처리부, 비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하여 정렬하는 전처리 과정을 수행하고, 대본 콘텐츠를 분석하는 대본 처리부, 상기 비디오 처리부에서 생성된 어노테이션 메타데이터를 저장하는 어노테이션 지식 베이스 및 상기 대본 처리부에서 생성된 내러티브 메타데이터를 저장하는 내러티브 지식 베이스를 포함하고, 어노테이션 지식 베이스와 내러티브 지식 베이스는 에피소드, 시퀀스, 씬 및 샷을 공유함으로써 상호 보완적으로 사용하여 메타데이터를 자동으로 생성한다.

[8] 비디오 처리부는 입력된 비디오의 매 프레임 간의 유사도를 측정하여 유사도가 미리 정해진 기준 이상으로 급히 변경되는 프레임 전후로 비디오의 샷으로 검출하여 구분하고, 구분된 비디오의 모든 샷에서 시, 공간이 바뀌는 시점을 인식하여 비디오의 씬을 검출한다.

[9] 대본 처리부는 반 구조화된 대본 서식을 분석하고, 대본으로부터 인물, 발화, 설명문, 장소, 시간을 구분하기 위해 다양한 형태의 대본 구조 양식과 입력된 대본 구조를 비교하여 적합한 대본 양식을 선택하고, 구조화(Structured) 하며, 해당 양식을 입력된 대본이 따랐는지 검증하고, 잘못 입력된 양식이 구조화 결과로 변형되었을 경우 보정하며, 대본의 구조로부터 대본의 씬을 검출하여 정보를 메타데이터의 계층구조로 지정한다.

[10] 비디오와 대본의 씬 순서를 일치시키기 위해 비디오의 한 씬의 발화들을 음성인식 기술로 자막을 만들고 대본의 발화와 문자열 비교 알고리즘을 이용해 1차 정렬한 후, 다이나믹 프로그래밍 기법을 이용하여 비디오와 대본의 씬 순서를 정렬한다.

[11] 대본을 비디오의 샷 단위까지 일치시키기 위해 비디오의 씬 안에 속한 샷 안의 시간 및 음성 텍스트 정보를 대본 콘텐츠 분석 과정 전달해 이후에 처리한다.

[12] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 비디오 콘텐츠의 메타데이터 자동 생성 방법은 대본 처리부와 상호 연결하기 위해 비디오를 입력 받아

비디오의 샷 및 씬을 인식하는 단계, 비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하는 단계, 인식된 비디오 씬 및 대본 씬을 정렬하는 전처리 과정을 수행하는 단계, 전처리된 비디오 콘텐츠를 분석하는 단계, 전처리된 대본 콘텐츠를 분석하는 단계, 분석된 비디오 콘텐츠로부터 생성된 어노테이션 메타데이터를 저장하는 단계 및 분석된 대본 콘텐츠로부터 생성된 내러티브 메타데이터를 저장하는 단계를 포함하고, 어노테이션 지식 베이스와 내러티브 지식 베이스는 에피소드, 시퀀스, 씬 및 샷을 공유함으로써 상호 보완적으로 사용하여 메타데이터를 자동으로 생성한다.

[13] 대본 처리부와 상호 연결하기 위해 비디오를 입력 받아 비디오의 샷 및 씬을 인식하는 단계는 인물 및 객체 인식을 통해 한 샷 안에서 샷의 첫번째, 마지막, 및 한 샷 내에서 프레임 간의 유사도가 급격히 바뀌는 한 개 이상의 프레임 들을 키프레임으로 지정하고, 키프레임 이미지에서 인물의 외형, 얼굴과 객체의 외형 및 카테고리를 인식하여 메타데이터로 저장한다.

[14] 비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하는 단계는, 대본 샷 추출단계에서 전처리 단계에서 구조화된 대본의 계층 구조와 부가 정보를 입력으로 하여 씬 단위로 나뉘어진 대본과, 비디오로의 샷, 씬 정보와 비디오의 음성 자막을 입력으로 받아 한 비디오 안의 발화와 대본과 비교, 정렬하여 대본을 샷 단위로 나눈다.

[15] 전처리된 비디오 콘텐츠를 분석하는 단계는 인물과 객체가 인식되면, 프레임 캡셔닝 단계에서 인물과 인물, 인물과 객체, 객체와 객체의 모든 상대 위치를 계산하고, 장면을 기술하고, 장면이 기술되면, 샷 캡셔닝단계에서 인물의 포즈와 상대 위치를 이용해 행동을 인식하고, 씬 캡셔닝단계에서 인식된 행동의 패턴을 기반으로 인물의 활동을 구분하여 스토리 중심적 구조로 배정해 저장하고, 다양한 정보를 비디오로 추출하여 어노테이션 메타데이터로 저장한다.

[16] 전처리된 대본 콘텐츠를 분석하는 단계는 대본이 비디오와 같이 샷 단위로 나눈 후 등장인물, 공간, 객체 분석을 통해 비디오 안에 등장하는 등장인물, 공간, 객체를 모두 각각 클러스터링(Clustering)하여 등장하는 인물, 공간, 객체의 카테고리를 파악하고, 등장인물의 경우 등장인물과 등장인물이 한 샷에 같이 등장한 횟수를 카운트해 인물 간의 관계를 그리고, 객체의 이름을 인식한 후, 등장인물, 배경, 공간을 모두 시맨틱웹과 엔티티로 연결시키고, 다이얼로그 분석단계에서는 모든 발화를 분석해 형태소(POS), 개체명(Named-Entity) 감정(Emotion), 목적(Intent), 화행(Speech Act), 어투(Tone), 존칭(Honorific)을 인식, 모든 정보를 각 발화의 메타데이터로 저장하고, 플롯 분석 단계에서 분석된 발화의 화행과 발화를 분석해 대화를 사건 단위로 나누고, 사건간의 관계를 분석해 대본의 기승전결을 파악하여 내러티브 메타데이터로써 저장한다.

발명의 효과

[17] 본 발명의 실시예들에 따르면 자동으로 비디오 콘텐츠와 관련된 정보와 비디오 콘텐츠의 구조, 그리고 스토리를 이해할 수 있는 비디오 콘텐츠의 다양한 정보를 담는 메타데이터를 생성해 메타데이터를 활용할 수 있는 여러 경우에 활용이 가능하다. 예를 들어 메타데이터의 활용은 만들어진 메타데이터에서 나뉘어진 비디오 구조를 통해 특정 시간에 특정 정보를 어노테이션 하는 인터랙티브 비디오와 같이 비디오 콘텐츠와 관련된 정보를 쉽고 빠르게 얻고자 할 때 메타데이터를 활용할 수 있으며, 비디오 콘텐츠와 관련된 정보를 검색(Information Retrieval) 하는 경우, 비디오 콘텐츠를 요약하여 요약된 비디오를 만드는 경우, 비디오의 등장인물의 발화 특성을 살린 인공지능 챗봇(Chatbot)을 말들 때, 장면의 정보를 기반으로 대화의 컨텍스트(Context)를 인지하는 챗봇을 만들 때와 같이 다양한 용도로 활용할 수 있다.

도면의 간단한 설명

[18] 도 1은 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 시스템 구조이다.

[19] 도 2는 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법에서 전처리 과정을 설명하기 위한 시스템 구조이다.

[20] 도 3은 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법 중 비디오를 분석하고 메타데이터를 생성하는 흐름도이다.

[21] 도 4는 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법 중 대본을 분석하고 메타데이터를 생성하는 흐름도이다.

[22] 도 5는 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법을 설명하기 위한 단계별 메타데이터 축적 흐름도이다.

[23] 도 6은 본 발명의 일 실시예에 따른 생성한 비디오 콘텐츠의 메타데이터를 이용한 활용 예이다.

[24] 도 7은 본 발명의 또 다른 일 실시예에 따른 생성한 비디오 콘텐츠의 메타데이터를 이용한 활용 예이다.

[25] 도 8은 본 발명의 또 다른 일 실시예에 따른 생성한 비디오 콘텐츠의 메타데이터를 이용한 활용 예이다.

발명의 실시를 위한 형태

[26] 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

[27]

[28] 도 1은 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법을 설명하기 위한 간략한 시스템 구조이다.

[29] 도 1의 자동 비디오 콘텐츠 메타데이터 생성 시스템(100)은 비디오와 비디오의 대본을 입력으로 받아 어노테이션 메타데이터와 내러티브 메타데이터를 출력으로 하는 시스템이며, 비디오와 대본을 각각 처리하는 비디오 처리부(110)과 대본 처리부(130)를 포함한다. 비디오 처리부(110)와 대본

처리부(130)를 서로 연결하기 위해 비디오와 대본의 씬을 정렬하고, 샷을 인식해 정렬하는 과정을 전처리로 하며, 대본과 비디오에서 각각 인식된 객체를 씬 순서에 맞게 정렬한다. 어노테이션 메타데이터는 어노테이션 지식 베이스(120)에 저장되어 추론이 가능하고 URI로 쉽게 메타데이터에 접근 가능한 온톨로지 기반의 데이터저장 공간이다. 내러티브 메타데이터는 내러티브 지식 베이스(130)에 저장되며 어노테이션 지식 베이스와 같은 역할을 한다. 어노테이션 지식 베이스(120)와 내러티브 지식 베이스(130)는 에피소드, 시퀀스, 씬 및 샷 등을 공유함으로써 상호 보완적으로 사용하여 메타데이터를 자동으로 생성한다.

[30] 비디오 처리부(110)는 입력된 비디오의 매 프레임 간의 유사도를 측정하여 유사도가 미리 정해진 기준 이상으로 급히 변경되는 프레임 전후로 비디오의 샷으로 검출하여 구분하고, 구분된 비디오의 모든 샷에서 시, 공간이 바뀌는 시점을 인식하여 비디오의 씬을 검출한다.

[31] 대본 처리부(130)는 반 구조화된 대본 서식을 분석하고, 대본으로부터 인물, 발화, 설명문, 장소, 시간을 구분하기 위해 다양한 형태의 대본 구조 양식과 입력된 대본 구조를 비교하여 적합한 대본 양식을 선택하고, 구조화(Structured) 한다. 그리고, 해당 양식을 입력된 대본이 따랐는지 검증하고, 잘못 입력된 양식이 구조화 결과로 변형되었을 경우 보정하며, 대본의 구조로부터 대본의 씬을 검출하여 정보를 메타데이터의 계층구조로 지정한다.

[32] 이후, 비디오와 대본의 씬 순서를 일치시키기 위해 비디오의 한 씬의 발화들을 음성인식 기술로 자막을 만들고 대본의 발화와 문자열 비교 알고리즘을 이용해 1차 정렬한 후, 다이나믹 프로그래밍 기법을 이용하여 비디오와 대본의 씬 순서를 정렬한다. 이때, 대본을 비디오의 샷 단위까지 일치시키기 위해 비디오의 씬 안에 속한 샷 안의 시간 및 음성 텍스트 정보를 대본 콘텐츠 분석 과정 전달해 이후에 처리한다.

[33]

[34] 도 2는 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법에서 전처리 과정을 설명하기 위한 시스템 구조이다.

[35] 제안하는 비디오 콘텐츠의 메타데이터 자동 생성 방법은 대본 처리부와 상호 연결하기 위해 비디오를 입력 받아 비디오의 샷 및 씬을 인식하는 단계, 비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하는 단계, 인식된 비디오 씬 및 대본 씬을 정렬하는 전처리 과정을 수행하는 단계, 전처리된 비디오 콘텐츠를 분석하는 단계, 전처리된 대본 콘텐츠를 분석하는 단계, 분석된 비디오 콘텐츠로부터 생성된 어노테이션 메타데이터를 저장하는 단계 및 분석된 대본 콘텐츠로부터 생성된 내러티브 메타데이터를 저장하는 단계를 포함한다. 도 2 내지 도 4를 참조하여 비디오 콘텐츠의 메타데이터 자동 생성 방법을 더욱 상세히 설명한다.

[36] 도 2를 참조하면, 샷 및 씬 구분단계(210), 대본 검증 및 씬 검출(230), 비디오 씬

및 대본 씬 정렬(240)은 본 방법의 전처리 부분에 해당하며, 비디오와 대본으로부터 다양한 정보를 추출하는 비디오 콘텐츠 분석 단계(220)와 대본 콘텐츠 분석 단계(250) 이전에 비디오의 샷, 씬을 검출하고, 대본의 씬을 구분하는 단계이다. 먼저 샷 및 씬 구분단계(210)은 입력된 비디오의 매 프레임(Frame) 간의 유사도를 측정하여 유사도가 급히 변경되는 프레임 전후로 비디오의 샷으로 검출(Shot Boundary Detection)하여 구분한다. 이렇게 구분된 비디오의 모든 샷에서 시, 공간이 바뀌는 시점을 인식하여 비디오의 씬을 검출한다. 대본 검증 및 씬 검출 단계에서(230)에서는 반 구조화된(Semi-Structured) 대본 서식을 분석, 대본으로부터 인물, 발화, 설명문, 장소, 시간을 구분하기 위해 다양한 형태의 대본 구조 양식과 입력된 대본 구조를 비교하여 적합한 대본 양식을 선택하고 처리 쉽도록 구조화(Structured) 하며, 해당 양식을 입력된 대본이 따랐는지 검증하고, 잘못 입력된 양식, 예를 들어 인물명-발화 순으로 나와야 하는 대본이 구조화 결과로, 인물명-인물명과 같이 대본의 구조가 변형되었을 경우 적합하게 보정한다. 또한 대본의 구조로부터 대본의 씬을 검출하여 그 정보를 메타데이터의 계층구조(Hierarchy)로 하여 지정한다. 이와 같이 비디오의 샷, 씬과 대본의 씬 정보를 추출한 후 비디오 씬 및 대본 씬 정렬(240)에서 비디오와 대본의 씬 순서를 정렬시킨다. 씬 순서를 일치시키기 위해 비디오의 한 씬의 발화들을 음성인식(Speech To Text) 기술로 자막을 만들고 대본의 발화와 문자열 비교 알고리즘을 이용해 1차 정렬한 후, 다이나믹 프로그래밍(Dynamic Programming)기법을 이용해 비디오와 대본의 씬 순서를 정렬한다. 이 때 비디오의 샷과 대본의 순서는 일치시키기 어려운데, 그 이유는 대본은 비디오의 단위인 샷 단위가 아닌 씬 단위로 나뉘어져 있기 때문에, 샷 단위까지 일치시키기 위해 비디오의 씬 안에 속한 샷 안의 시간 및 음성 텍스트 정보(260)를 대본 콘텐츠 분석 과정 전달해 이후에 처리한다.

[37]

[38] 도 3은 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법 중 비디오를 분석하고 메타데이터를 생성하는 흐름도이다.

[39] 이전의 전처리 과정을 통해 비디오의 샷, 씬 계층구조(270)을 통해 넘어오면, 먼저 인물 및 객체 인식(310)을 통해 한 샷 안에서 샷의 첫번째, 마지막, 그리고 한 샷 내에서 프레임(Frame) 간의 유사도가 급격히 바뀌는 한 개 이상의 프레임 들을 키프레임으로 지정하고, 이 키프레임 이미지에서 인물의 외형(Contour), 얼굴과 객체의 외형 및 그 카테고리를 인식하여 메타데이터로 저장한다. 인물과 객체가 인식되면, 프레임 캡셔닝(320) 단계에서 인물의 위치만 인식하는 것이 아닌 인물과 인물, 인물과 객체, 객체와 객체의 모든 상대 위치를 계산하고, 인물의 경우 인물의 팔, 다리, 몸의 위치와 모습을 이용해 포즈(Pose)를 인식하여 장면을 기술한다. 이와 같은 정보가 기술되면, 샷 캡셔닝(330)단계에서 인물의 포즈와 상대 위치를 이용해 행동(Action)을 인식하고, 씬 캡셔닝(340) 단계에서

인식된 행동의 패턴을 기반으로 인물의 활동(Activity)을 구분하여 각각 사건(Event)이란 스토리 중심적 구조로 배정해 저장한다. 이와 같이 다양한 정보를 비디오로 추출하여 어노테이션 메타데이터로 저장한다.

[40]

[41] 도 4는 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법 중 대본을 분석하고 메타데이터를 생성하는 흐름도이다.

[42] 대본 샷 추출(410)단계에서는 두개의 입력을 받는데 전처리 단계에서 구조화된 대본의 계층 구조와 부가 정보(280)를 입력으로 하여 씬 단위로 나뉘어진 대본과, 비디오로의 샷, 씬 정보(260)와 비디오의 음성 자막을 입력으로 받아 한 비디오 안의 발화와 대본과 비교, 정렬하여 대본을 샷 단위로 나누는 단계이다. 대본이 비디오와 같이 샷 단위로 나눈 후 등장인물, 공간, 객체 분석(420)을 통해 비디오 안에 등장하는 등장인물, 공간, 객체를 모두 각각 클러스터링(Clustering)하여 등장하는 인물, 공간, 객체의 카테고리를 파악하고, 등장인물의 경우 등장인물과 등장인물이 한 샷에 같이 등장한 횟수를 카운트해 인물 간의 관계를 그리고, 객체의 이름을 인식한 후, 등장인물, 배경, 공간을 모두 시맨틱웹과 엔티티로 Entity Linking 시켜 더 많은 부가 정보를 갖고 올 수 있게끔 한다. 다이얼로그 분석(430)단계에서는 모든 발화를 분석해 형태소(POS), 개체명(Named-Entity) 감정(Emotion), 목적(Intent), 화행(Speech Act), 어투(Tone), 존칭(Honorific)을 인식, 모든 정보를 각 발화의 메타데이터로 저장한다. 다음으로 플롯 분석(440)에서는 분석된 발화의 화행과 발화를 분석해 대화를 사건(Event) 단위로 나누고, 여러 사건간의 관계를 분석해 대본의 기승전결을 파악하여 메타데이터로 저장한다. 이렇게 생성된 메타데이터는 내러티브 메타데이터로써 저장한다.

[43]

[44] 도 5는 본 발명의 일 실시예에 따른 비디오 콘텐츠의 메타데이터 자동 생성 방법을 설명하기 위한 단계별 메타데이터 축적 흐름도이다.

[45] 먼저 비디오와 대본을 입력(510) 받아, 대본을 구조화(520)하고 비디오와 대본의 계층 구조를 분석해 씬, 샷의 정보를 추가한다. 이후 비디오와 대본안의 다양한 정보를 인식하고 추출하여 콘텐츠(530) 정보를 추가한다. 그 다음 비디오와 대본의 스토리를 분석하기 위해 행동, 발화의 특성들을 분석해 사건(540) 정보를 생성해 추가하고 마지막으로 메타데이터(550)를 생성해 지식베이스에 추가한다.

[46]

[47] 도 6은 본 발명의 일 실시예에 따른 생성한 비디오 콘텐츠의 메타데이터를 이용한 활용 예이다.

[48] 내러티브 메타데이터의 정보를 이용해 비디오 콘텐츠의 스토리에서 주요한 장면들(610)을 추출해 아래에 썸네일의 슬라이딩바 형태로 보여준다. 사용자가 관심있는 장면을 직접 클릭하여 해당 장면을 재생할 수 있으며, 화면 위의

검색창(620)에서 직접 검색할 수 있다. 검색은 예약어가 아닌 자연어(Natural Language)로 입력하면, 자동으로 형태소 분석, 개체명 인식을 통해 등장인물명, 장소, 행동과 같은 주요 단어를 추출해 자동으로 SPARQL Query를 생성해 내러티브 지식베이스와 어노테이션 지식베이스의 SPARQL Endpoint로 지식을 검색하여 해당하는 장면의 비디오를 가져와 화면에 보여준다. 비디오 콘텐츠와 관련된 정보를 상호작용가능한 증강객체로써 어노테이션 하며, 인식된 인물이나 객체의 외각선과 인물의 명칭을 (630)을 화면 위에 그려 클릭 가능한 객체임을 표시한다. 인물을 클릭할 경우 인물의 부가정보와 인물과 인물의 관계도, 인물과 관련된 씬, 사건 등의 정보를 보여주며, 객체를 클릭할 경우 객체가 스토리에서 어떠한 의미를 갖는지를 보여주고, 객체가 만약 인터넷에서 구입 가능하다면, 구입할 수 있는 추가 객체(650)를 띄워 바로 구입 가능하도록 한다.

[49]

[50] 도 7은 본 발명의 또 다른 일 실시예에 따른 생성한 비디오 콘텐츠의 메타데이터를 이용한 활용 예이다.

[51] N-Screen 환경에서 TV 스크린(710)은 비디오 재생프로그램으로 어노테이션 메타데이터와 내러티브 메타데이터를 생성하기 위해 사용한 비디오를 재생한다. 비디오가 재생되며 모바일 디바이스(720)에서도 동시에 같은 비디오를 재생하며, TV 스크린의 비디오와 모바일 디바이스의 비디오를 소켓 통신을 이용해 실시간으로 재생시간을 맞추어 동기화(730)한다. 이렇게 동기화된 비디오는 모바일 디바이스에 비디오와 같이 보여줄 비디오와 관련된 스토리 정보를 보여줄 때 역시 비디오 시간과 동기화 하여 보여준다. 비디오 위의 얼굴 바운딩 박스(740)은 비디오에 등장하는 인물 중 해당 장면의 스토리에서 중요한 인물의 얼굴을 어노테이션 메타데이터의 데이터를 이용해 표시한다. 내러티브 메타데이터에 있는 비디오의 스토리 정보를 요약해 인물 관계(750)과 장면 스토리 요약(760)과 같은 비디오와 관련된 부가정보를 보여준다. 인물 관계는 해당 장면의 스토리에서 매우 중요한 인물의 인물 관계도를 보여주며, 비디오의 스토리를 쉽게 파악하고, 등장인물의 관계를 파악하기 위해 보여준다. 장면 스토리 요약은 해당 장면을 내러티브 메타데이터의 데이터를 이용해 요약하여 보여준다. 이와 같은 부가 정보를 통해 비디오를 보면서 사용자가 미처 파악하지 못한 추가 정보를 TV에서 재생중인 비디오를 멈추고 인터넷에서 검색할 필요 없이 모바일 디바이스에서 직관적으로 보여줘 스토리를 쉽게 이해하도록 한다.

[52]

[53] 도 8은 본 발명의 또 다른 일 실시예에 따른 생성한 비디오 콘텐츠의 메타데이터를 이용한 활용 예이다.

[54] 어노테이션 메타데이터 생성시 키프레임 이미지 위의 객체를 인식, 객체의 등장 시간과, 객체의 위치를 인식해 메타데이터에 저장한다. 비디오와 비디오

안의 객체와 관련된 정보를 가져와 비디오 위에 정보를 증강하기 위해 메타데이터에 저장된 객체의 등장 시간과 위치, 크기, 객체의 정보를 가져온다. 그리고 키프레임 사이의 프레임 위에 객체와 관련된 정보를 어노테이션 하기 위해 보간(interpolation)을 해 중간 값인 특정 시간의 객체 위치와 크기를 생성해 모든 프레임에 어노테이션 한다. 객체의 정보는 인식된 객체의 비디오와 객체의 이미지로 유사한 쇼핑몰의 제품 광고를 가져와 보여주거나, 내러티브 메타데이터에서 현재 장면과 관련된 객체, 예를 들어 선물과 같은 객체의 정보를 가져와 보여준다. 또한 제품 광고는 사용자에 따라 다른 정보를 추천하여 보여준다.

[55]

[56] 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, FPA(field programmable 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 처리 요소(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.

[57] 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 저장 매체 또는 장치에 구체화(embody)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.

[58] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 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 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.

[59] 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.

[60] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

[61]

청구범위

[청구항 1] 대본 처리부와 상호 연결하기 위해 비디오를 입력 받아 비디오의 샷 및 씬을 인식하여 정렬하는 전처리 과정을 수행하고, 비디오 콘텐츠를 분석하는 비디오 처리부;
비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하여 정렬하는 전처리 과정을 수행하고, 대본 콘텐츠를 분석하는 대본 처리부;
상기 비디오 처리부에서 생성된 어노테이션 메타데이터를 저장하는 어노테이션 지식 베이스; 및
상기 대본 처리부에서 생성된 내러티브 메타데이터를 저장하는 내러티브 지식 베이스
를 포함하고,
어노테이션 지식 베이스와 내러티브 지식 베이스는,
에피소드, 시퀀스, 씬 및 샷을 공유함으로써 상호 보완적으로 사용하여 메타데이터를 자동으로 생성하는
메타데이터 자동 생성 시스템.

[청구항 2] 제1항에 있어서,
비디오 처리부는,
입력된 비디오의 매 프레임 간의 유사도를 측정하여 유사도가 미리 정해진 기준 이상으로 급히 변경되는 프레임 전후로 비디오의 샷으로 검출하여 구분하고, 구분된 비디오의 모든 샷에서 시, 공간이 바뀌는 시점을 인식하여 비디오의 씬을 검출하는
메타데이터 자동 생성 시스템.

[청구항 3] 제1항에 있어서,
대본 처리부는,
반 구조화된 대본 서식을 분석하고, 대본으로부터 인물, 발화, 설명문, 장소, 시간을 구분하기 위해 다양한 형태의 대본 구조 양식과 입력된 대본 구조를 비교하여 적합한 대본 양식을 선택하고, 구조화(Structured) 하며, 해당 양식을 입력된 대본이 따랐는지 검증하고, 잘못 입력된 양식이 구조화 결과로 변형되었을 경우 보정하며, 대본의 구조로부터 대본의 씬을 검출하여 정보를 메타데이터의 계층구조로 지정하는
메타데이터 자동 생성 시스템.

[청구항 4] 제1항에 있어서,
비디오와 대본의 씬 순서를 일치시키기 위해 비디오의 한 씬의 발화들을 음성인식 기술로 자막을 만들고 대본의 발화와 문자열 비교 알고리즘을 이용해 1차 정렬한 후, 다이나믹 프로그래밍 기법을 이용하여 비디오와 대본의 씬 순서를 정렬하는

메타데이터 자동 생성 시스템.

[청구항 5] 제4항에 있어서,
대본을 비디오의 샷 단위까지 일치시키기 위해 비디오의 씬 안에 속한 샷 안의 시간 및 음성 텍스트 정보를 대본 콘텐츠 분석 과정 전달해 이후에 처리하는
메타데이터 자동 생성 시스템.

[청구항 6] 대본 처리부와 상호 연결하기 위해 비디오를 입력 받아 비디오의 샷 및 씬을 인식하는 단계;
비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하는 단계;
인식된 비디오 씬 및 대본 씬을 정렬하는 전처리 과정을 수행하는 단계;
전처리된 비디오 콘텐츠를 분석하는 단계;
전처리된 대본 콘텐츠를 분석하는 단계;
분석된 비디오 콘텐츠로부터 생성된 어노테이션 메타데이터를 어노테이션 지식 베이스에 저장하는 단계; 및
분석된 대본 콘텐츠로부터 생성된 내러티브 메타데이터를 내러티브 지식 베이스에 저장하는 단계
를 포함하고,
어노테이션 지식 베이스와 내러티브 지식 베이스는 에피소드, 시퀀스, 씬 및 샷을 공유함으로써 상호 보완적으로 사용하여 메타데이터를 자동으로 생성하는
메타데이터 자동 생성 방법.

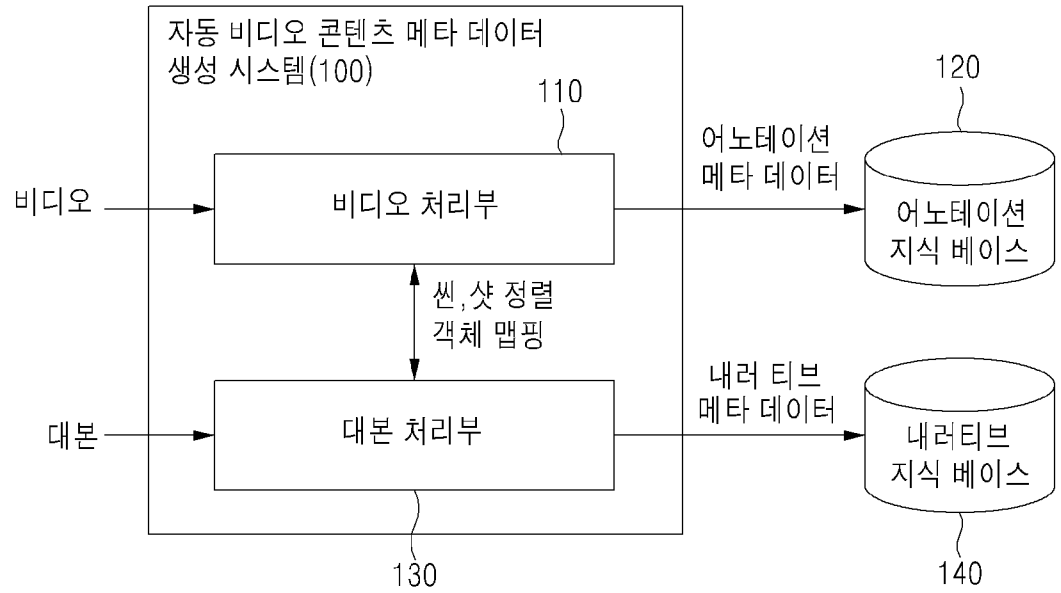
[청구항 7] 제6항에 있어서,
대본 처리부와 상호 연결하기 위해 비디오를 입력 받아 비디오의 샷 및 씬을 인식하는 단계는,
인물 및 객체 인식을 통해 한 샷 안에서 샷의 첫번째, 마지막, 및 한 샷 내에서 프레임 간의 유사도가 급격히 바뀌는 한 개 이상의 프레임 들을 키프레임으로 지정하고, 키프레임 이미지에서 인물의 외형, 얼굴과 객체의 외형 및 카테고리를 인식하여 메타데이터로 저장하는
메타데이터 자동 생성 방법.

[청구항 8] 제6항에 있어서,
비디오 처리부와 상호 연결하기 위해 대본을 입력 받아 대본 및 씬을 인식하는 단계는,
대본 샷 추출단계에서 전처리 단계에서 구조화된 대본의 계층 구조와 부가 정보를 입력으로 하여 씬 단위로 나뉘어진 대본과, 비디오로의 샷, 씬 정보와 비디오의 음성 자막을 입력으로 받아 한 비디오 안의 발화와 대본과 비교, 정렬하여 대본을 샷 단위로 나누는
메타데이터 자동 생성 방법.

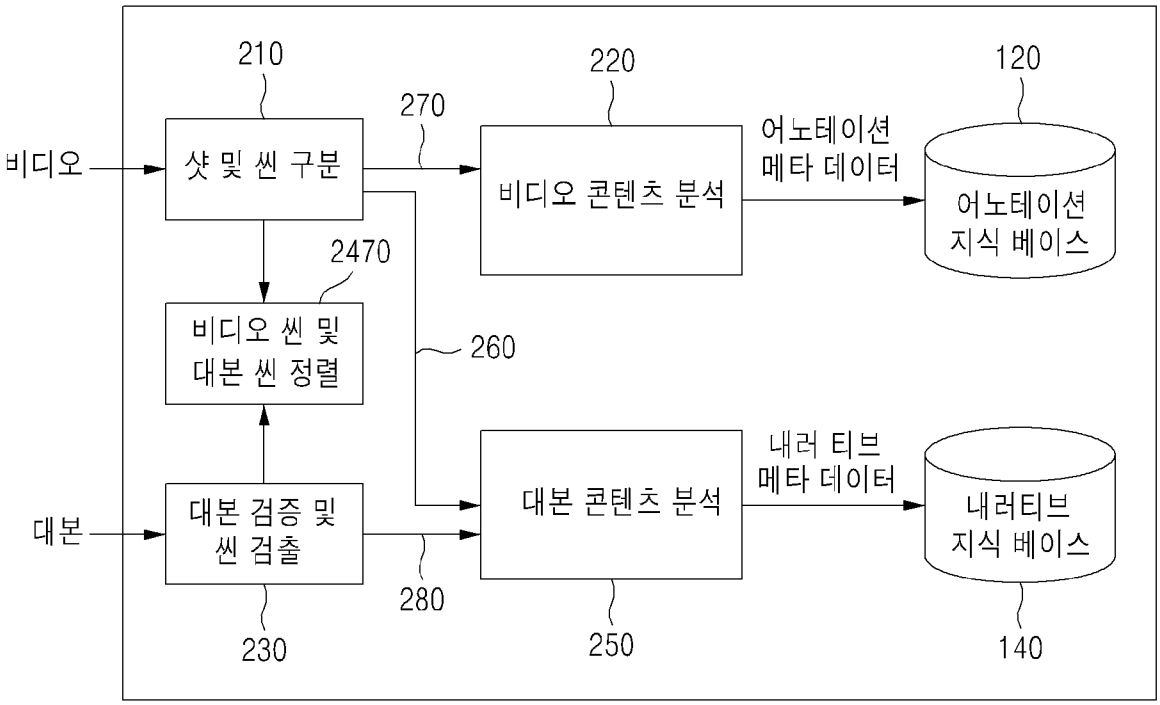
[청구항 9] 제7항에 있어서,
전처리된 비디오 콘텐츠를 분석하는 단계는,
인물과 객체가 인식되면, 프레임 캡셔닝 단계에서 인물과 인물, 인물과 객체, 객체와 객체의 모든 상대 위치를 계산하고, 장면을 기술하고,
장면이 기술되면, 샷 캡셔닝단계에서 인물의 포즈와 상대 위치를 이용해 행동을 인식하고,
씬 캡셔닝단계에서 인식된 행동의 패턴을 기반으로 인물의 활동을 구분하여 스토리 중심적 구조로 배정해 저장하고, 다양한 정보를 비디오로 추출하여 어노테이션 메타데이터로 저장하는
메타데이터 자동 생성 방법.

[청구항 10] 제8항에 있어서,
전처리된 대본 콘텐츠를 분석하는 단계는,
대본이 비디오와 같이 샷 단위로 나눈 후 등장인물, 공간, 객체 분석을 통해 비디오 안에 등장하는 등장인물, 공간, 객체를 모두 각각 클러스터링(Clustering)하여 등장하는 인물, 공간, 객체의 카테고리를 파악하고, 등장인물의 경우 등장인물과 등장인물이 한 샷에 같이 등장한 횟수를 카운트해 인물 간의 관계를 그리고, 객체의 이름을 인식한 후,
등장인물, 배경, 공간을 모두 시맨틱웹과 엔티티로 연결시키고,
다이얼로그 분석단계에서는 모든 발화를 분석해 형태소(POS), 개체명(Named-Entity) 감정(Emotion), 목적(Intent), 화행(Speech Act), 어투(Tone), 존칭(Honorific)을 인식, 모든 정보를 각 발화의 메타데이터로 저장하고,
플롯 분석 단계에서 분석된 발화의 화행과 발화를 분석해 대화를 사건 단위로 나누고, 사건간의 관계를 분석해 대본의 기승전결을 파악하여 내러티브 메타데이터로써 저장하는
메타데이터 자동 생성 시스템.

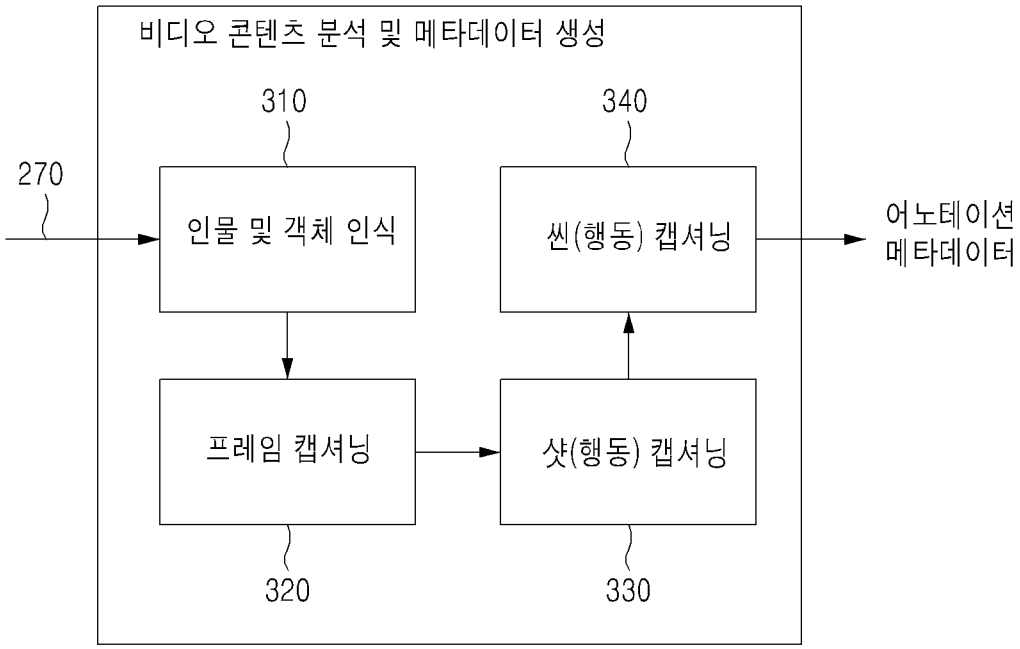
[도1]


자동 비디오 콘텐츠 메타 데이터
생성 시스템(100)
110
120
비디오
비디오 처리부
어노테이션
메타 데이터
어노테이션
지식 베이스
씬,샷 정렬
객체 맵핑
대본
대본 처리부
내러 티브
메타 데이터
내러티브
지식 베이스
130
140

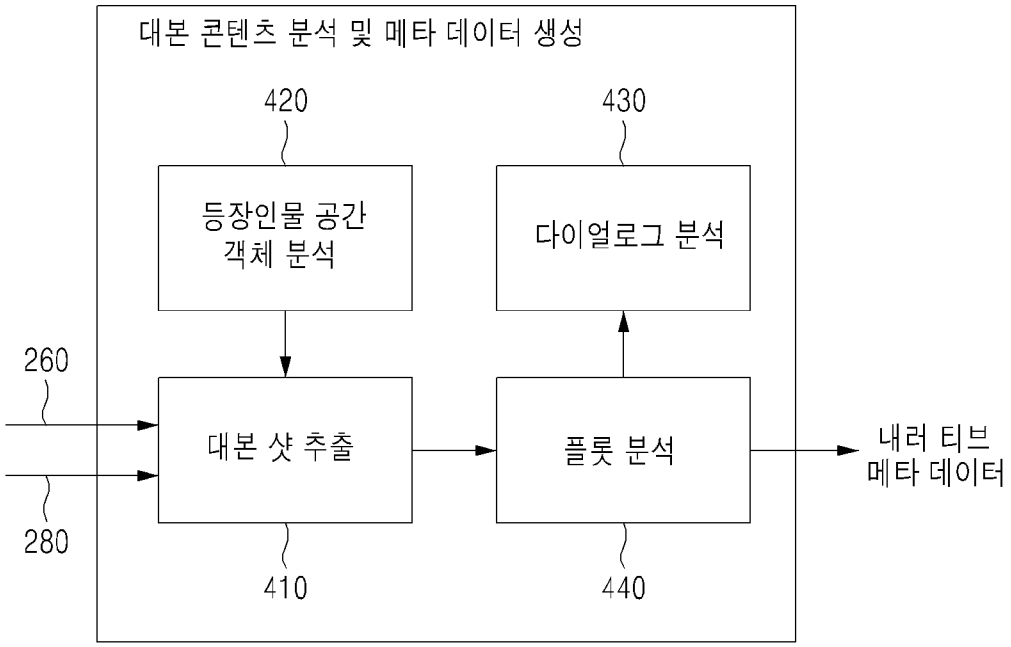
[도2]


210
220
120
270
비디오
샷 및 씬 구분
비디오 콘텐츠 분석
어노테이션
메타 데이터
어노테이션
지식 베이스
2470
비디오 씬 및
대본 씬 정렬
260
내러 티브
메타 데이터
대본
대본 검증 및
씬 검출
대본 콘텐츠 분석
내러티브
지식 베이스
280
230
250
140

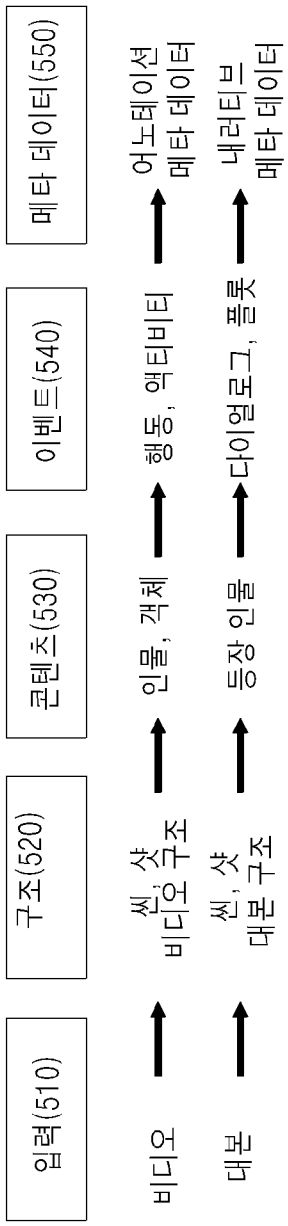
[도3]


비디오 콘텐츠 분석 및 메타데이터 생성
310
340
270
인물 및 객체 인식
씬(행동) 캡셔닝
어노테이션
메타데이터
프레임 캡셔닝
샷(행동) 캡셔닝
320
330

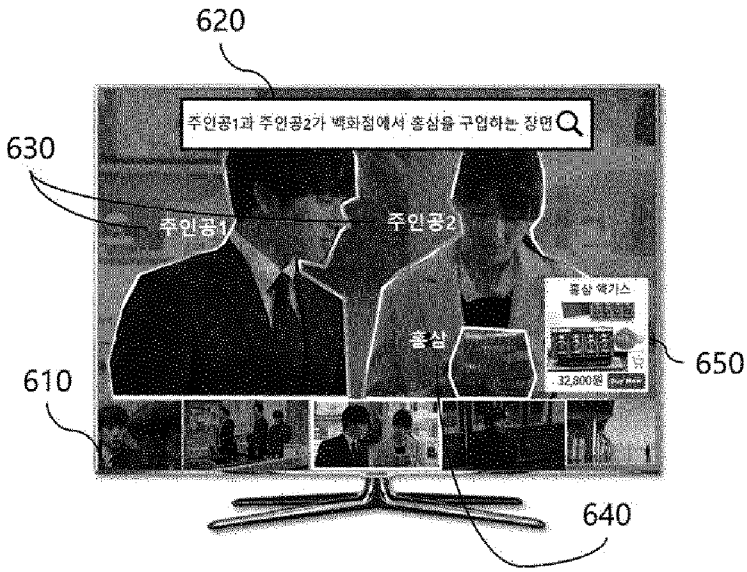
[도4]


대본 콘텐츠 분석 및 메타 데이터 생성
420
430
등장인물 공간
객체 분석
다이얼로그 분석
260
대본 샷 추출
플롯 분석
내러 티브
메타 데이터
280
410
440

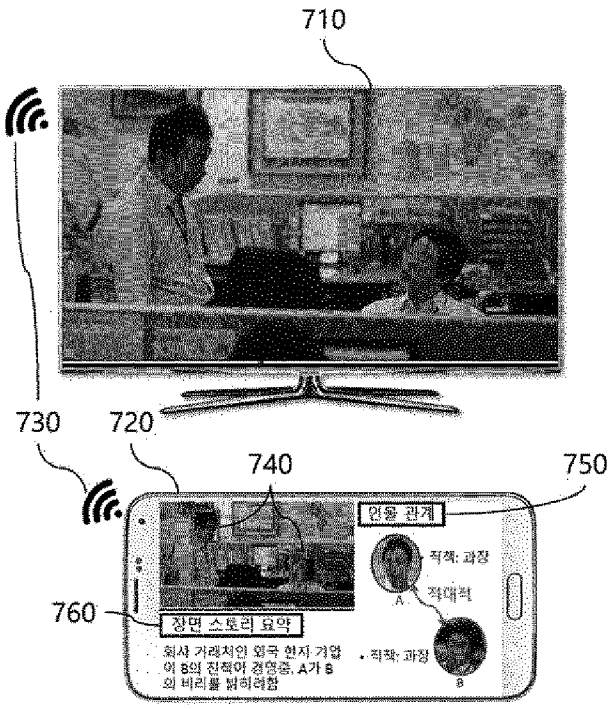
[도5]


입력(510)
구조(520)
콘텐츠(530)
이벤트(540)
메타 데이터(550)
비디오
대본
씬, 샷
비디오 구조
씬, 샷
대본 구조
인물, 객체
등장 인물
행동, 액티비티
다이얼로그, 플롯
어노테이션
메타 데이터
내러티브
메타 데이터

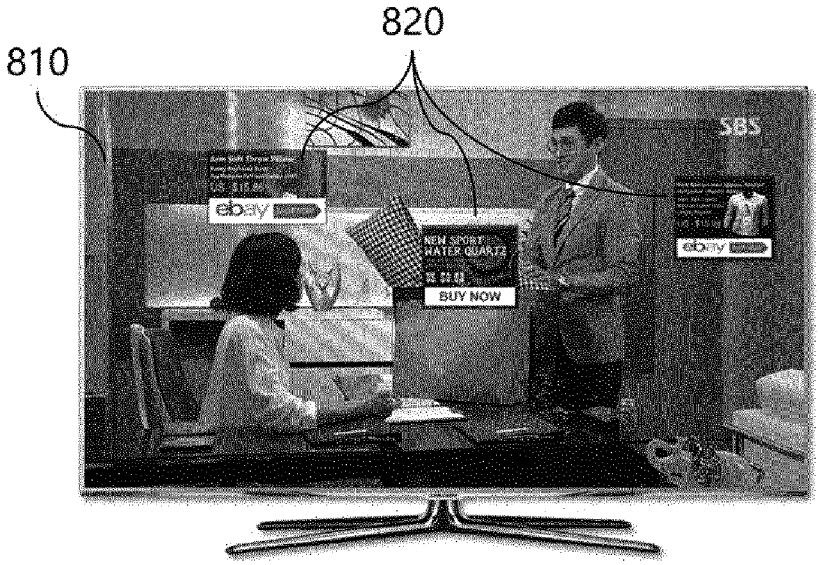
[도6]


620
주인공1과 주인공2가 백화점에서 홍삼을 구입하는 장면
630
주인공1
주인공2
홍삼
610
650
640

[도7]


710
730
720
740
750
인물 관계
적책: 과장
적대적
760
장면 스토리 요약
회사 거래처인 외국 현지 기업
이 B의 진책이 경영권, A가 B
의 비리를 밝히려함
• 적책: 과장

[도8]


820
810
SBS
ebay
NEW SPORT
WATER QUARTZ
BUY NOW

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2018/016854

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/234(2011.01)i, H04N 21/235(2011.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N 21/234; G06F 17/30; G06F 3/06; G06Q 50/00; H04N 21/235; H04N 21/43; H04N 21/4402; H04N 21/84; H04N 5/445

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models: IPC as above
Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: metadata, video shot, video scene, video processor, script processor, annotation knowledge base, narrative knowledge base

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2018-0087969 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 03 August 2018 See paragraphs [0034]-[0040]; and claim 1.	1-2,4-9
A		3,10
Y	KR 10-2008-0112975 A (SUH, Jong Hoon) 26 December 2008 See paragraph [0024]; claims 1, 3; and figure 1.	1-2,4-9
A	KR 10-2012-0050016 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 18 May 2012 See paragraphs [0041]-[0053]; claim 1; and figures 2-4.	1-10
A	KR 10-2017-0062722 A (INHA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION) 08 June 2017 See paragraphs [0020]-[0028]; and figure 1.	1-10
A	KR 10-2010-0123204 A (INHA UNIVERSITY RESEARCH AND BUSINESS FOUNDATION) 24 November 2010 See paragraphs [0029]-[0049]; and figure 1.	1-10

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 JULY 2019 (15.07.2019)	**16 JULY 2019 (16.07.2019)**

Name and mailing address of the ISA/KR
Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea
Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2018/016854

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2018-0087969 A	03/08/2018	US 2018-0213289 A1	26/07/2018
KR 10-2008-0112975 A	26/12/2008	None	
KR 10-2012-0050016 A	18/05/2012	KR 10-1301133 B1	12/09/2013
KR 10-2017-0062722 A	08/06/2017	KR 10-1817402 B1	10/01/2018
KR 10-2010-0123204 A	24/11/2010	KR 10-1072347 B1	12/10/2011

국 제 조 사 보 고 서

국제출원번호
PCT/KR2018/016854

A. 발명이 속하는 기술분류(국제특허분류(IPC))

H04N 21/234(2011.01)i, H04N 21/235(2011.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

H04N 21/234; G06F 17/30; G06F 3/06; G06Q 50/00; H04N 21/235; H04N 21/43; H04N 21/4402; H04N 21/84; H04N 5/445

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC
일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 메타데이터(metadata), 비디오 샷(video shot), 비디오 씬(video scene), 비디오 처리부(video processor), 대본 처리부(script processor), 어노테이션 지식 베이스(annotation knowledge base), 내러티브 지식 베이스(narrative knowledge base)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2018-0087969 A (한국전자통신연구원) 2018.08.03 단락 [0034]-[0040]; 및 청구항 1 참조.	1-2,4-9
A		3,10
Y	KR 10-2008-0112975 A (서종훈) 2008.12.26 단락 [0024]; 청구항 1, 3; 및 도면 1 참조.	1-2,4-9
A	KR 10-2012-0050016 A (한국전자통신연구원) 2012.05.18 단락 [0041]-[0053]; 청구항 1; 및 도면 2-4 참조.	1-10
A	KR 10-2017-0062722 A (인하대학교 산학협력단) 2017.06.08 단락 [0020]-[0028]; 및 도면 1 참조.	1-10
A	KR 10-2010-0123204 A (인하대학교 산학협력단) 2010.11.24 단락 [0029]-[0049]; 및 도면 1 참조.	1-10

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.

☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

"A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

"E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

"L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

"O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

"P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

"T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

"X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

"Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

"&" 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일	국제조사보고서 발송일
2019년 07월 15일 (15.07.2019)	2019년 07월 16일 (16.07.2019)

ISA/KR의 명칭 및 우편주소
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)
팩스 번호 +82-42-481-8578

심사관
안정환
전화번호 +82-42-481-8633

서식 PCT/ISA/210 (두 번째 용지) (2015년 1월)

국 제 조 사 보 고 서
대응특허에 관한 정보

국제출원번호
PCT/KR2018/016854

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2018-0087969 A	2018/08/03	US 2018-0213289 A1	2018/07/26
KR 10-2008-0112975 A	2008/12/26	없음	
KR 10-2012-0050016 A	2012/05/18	KR 10-1301133 B1	2013/09/12
KR 10-2017-0062722 A	2017/06/08	KR 10-1817402 B1	2018/01/10
KR 10-2010-0123204 A	2010/11/24	KR 10-1072347 B1	2011/10/12