



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년09월21일
(11) 등록번호 10-1658002
(24) 등록일자 2016년09월09일

(51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01)
(52) CPC특허분류
G06F 17/3082 (2013.01)
G06F 17/30843 (2013.01)
(21) 출원번호 10-2015-0177008
(22) 출원일자 2015년12월11일
심사청구일자 2015년12월11일
(56) 선행기술조사문헌
KR1020150022088 A
KR1020080112975 A
KR1020000007558 A
KR1020000038290 A

(73) 특허권자
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
남종호
서울특별시 마포구 마포대로 195 307동 1501호,
(마포래미안 푸르지오)
최지수
경기도 안양시 동안구 흥안대로223번길 47 110동
605호 (호계동, 샘마을대우아파트)
최기석
경기도 고양시 일산서구 강선로 71 709동 601호
(주엽동, 강선마을7단지아파트)
(74) 대리인
이지연

전체 청구항 수 : 총 15 항

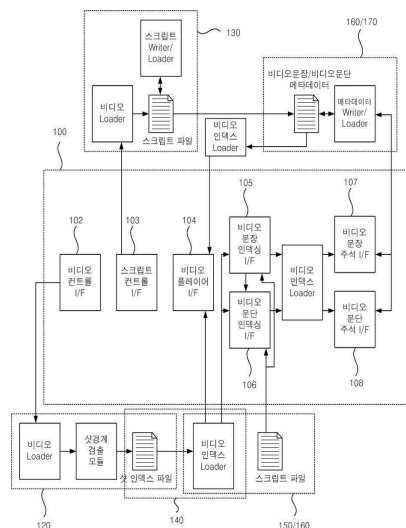
심사관 : 박승철

(54) 발명의 명칭 동영상 주석 시스템 및 주석 방법

(57) 요약

본 발명은 동영상 주석 시스템에 관한 것이다. 상기 동영상 주석 시스템은 새로운 비디오 인덱싱 단위로 비디오 문장(Video Sentence)과 비디오 문단(Video Paragraph)을 정의하고, 상기 비디오 문장 및 비디오 문단의 단위로 인덱싱을 수행한다. 상기 비디오 문장은 한 명의 화자가 하나 혹은 여러 샷을 통해 한 문장 단위의 음성 설명을 하는 비디오 구간을 의미하며, 상기 비디오 문단은 하나 이상의 비디오 문장으로 이루어진 여러 명의 화자가 하나의 소주제를 두고 말하는 구간이며 다른 주제로 변경되기 전까지를 경계로 한다. 본 발명에 따른 인덱싱 방법은 계층적인 구조를 가지고 있어 의학적인 설명 및 소주제에 맞추어 한 편의 에피소드를 주제별 검색을 효과적으로 할 수 있도록 한다.

대표도 - 도7



(52) CPC특허분류

G06F 17/30849 (2013.01)

G06F 17/30858 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 R0101-15-0097

부처명 미래창조과학부

연구관리전문기관 정보통신기술진흥센터

연구사업명 정보통신·방송 기술개발 및 표준화 사업

연구과제명 클라우드 기반 개방형 소셜 방송미디어 콘텐츠 융합 생성, 편집 및 재생을 위한 미디어 제작 및 전송 시스템 개발

기 여 율 1/1

주관기관 (주)썬멀티미디어

연구기간 2013.03.01 ~ 2016.02.29

공지예외적용 : 있음

명세서

청구범위

청구항 1

동영상을 로딩하고 상기 동영상에 대한 샷 인덱스 파일을 생성하는 비디오 제어부;

상기 동영상의 음성 데이터를 분석하여 텍스트 데이터로 구성된 스크립트 파일을 생성하는 스크립트 제어부;

상기 샷 인덱스 파일과 스크립트 파일을 이용하여 비디오 문장을 형성하고 인덱싱하는 비디오 문장 인덱싱부;

상기 비디오 문장들을 그룹핑하여 비디오 문단을 형성하고 비디오 문단을 인덱싱하는 비디오 문단 인덱싱부;

를 구비하여, 동영상을 비디오 문장(Video Sentence) 및 비디오 문단(Video Paragraph)으로 나누어 비디오 문장 단위 및 비디오 문단 단위로 인덱싱하는 것을 특징으로 하며,

상기 비디오 문장은 단일의 화자에 의한 음성 설명에서 단일의 문장을 이루는 하나 또는 둘 이상의 샷으로 구성되는 비디오 구간으로 설정되며,

상기 비디오 문단은 공통적인 주제에 대한 하나 또는 둘 이상의 비디오 문장들로 구성되는 비디오 구간으로 설정되는 것을 특징으로 하는 동영상 주석 시스템.

청구항 2

제1항에 있어서, 상기 동영상 주석 시스템은

상기 샷 인덱스 파일을 이용하여 샷 시퀀스(Shot Sequence)에 따른 샷들을 재생하는 비디오 재생부;

상기 인덱싱된 비디오 문장들에 대하여 주석을 기입하고 비디오 문장에 대한 메타데이터를 생성하는 비디오 문장 주석부;

상기 비디오 문단들에 대하여 주석을 기입하고 비디오 문단에 대한 메타데이터를 생성하는 비디오 문단 주석부;

를 더 구비하여, 동영상을 비디오 문장과 비디오 문단으로 인덱싱하고 주석을 기입하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 3

제1항에 있어서, 상기 주석은 동영상의 콘텐츠의 특화된 정보를 고려하여 설계된 택소노미(taxonomy)를 기반으로 한 구조를 갖는 주석 데이터로 구성된 것을 특징으로 하는 동영상 주석 시스템.

청구항 4

제1항에 있어서, 상기 비디오 제어부는 인덱싱하고자 하는 동영상을 로딩하고, 상기 로딩된 동영상에 대하여 샷 경계를 검출(Shot Boundary Detection)하고, 샷 경계 검출을 통해 상기 동영상을 샷 단위로 세그멘테이션하며, 상기 동영상에 대한 샷 인덱스 파일을 생성하여 저장하는 것을 특징으로 하며,

상기 샷 인덱스 파일은 각 샷의 경계에 대한 시각 정보 및 프레임 정보를 포함하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 5

제1항에 있어서, 상기 스크립트 제어부는 상기 동영상에 대한 자막 파일을 불러와서 상기 스크립트 파일에 저장하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 6

제1항에 있어서, 상기 비디오 문장 인덱싱부는

스크립트 파일의 문장의 단위를 기준으로 하여 샷들을 그룹핑하여 비디오 문장을 설정하고, 상기 설정된 비디오

문장들을 인덱싱하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 7

제1항에 있어서, 상기 비디오 문단 인덱싱부는

상기 비디오 문장들을 공통적인 주제에 따라 그룹핑하여 비디오 문단을 설정하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 8

제2항에 있어서, 상기 비디오 문장 주석부는 각 비디오 문장에 대하여 비디오 문장 타입, 특화 장면 주석 정보 및 특화 프레임 주석 정보를 기입하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 9

제2항에 있어서, 상기 비디오 문단 주석부는 각 비디오 문단에 대한 비디오 문단 타입과 비디오 문단 키워드 정보를 기입하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 10

제2항에 있어서, 상기 동영상 주석 시스템은 사용자가 상기 동영상 주석 시스템의 각 부에 접속할 수 있도록 하는 사용자 인터페이스부를 더 구비하고,

상기 사용자 인터페이스부는

상기 비디오 제어부에 접속하는 비디오 컨트롤 인터페이스부;

상기 스크립트 제어부에 접속하는 스크립트 컨트롤 인터페이스부;

상기 비디오 재생부에 접속하는 비디오 플레이어 인터페이스부;

상기 비디오 문장 인덱싱부에 접속하는 비디오 문장 인덱싱 인터페이스부;

상기 비디오 문단 인덱싱부에 접속하는 비디오 문단 인덱싱 인터페이스부;

상기 비디오 문장 주석부에 접속하는 비디오 문장 주석 인터페이스부;

상기 비디오 문단 주석부에 접속하는 비디오 문단 주석 인터페이스부;

를 구비하는 것을 특징으로 하는 동영상 주석 시스템.

청구항 11

(a) 동영상을 로딩하고 상기 동영상에 대한 샷 인덱스 파일을 생성하는 단계;

(b) 상기 동영상의 음성 데이터를 분석하여 텍스트 데이터로 구성된 스크립트 파일을 생성하는 단계;

(c) 상기 샷 인덱스 파일을 이용하여 샷 시퀀스(Shot Sequence)에 따른 샷들을 재생하는 단계;

(d) 상기 샷 인덱스 파일과 스크립트 파일을 이용하여 비디오 문장을 형성하고 인덱싱하는 단계;

(e) 상기 비디오 문장들을 그룹핑하여 비디오 문단을 형성하고 비디오 문단을 인덱싱하는 단계;

를 구비하여, 동영상을 비디오 문장과 비디오 문단으로 인덱싱하는 것을 특징으로 하는 동영상 주석 방법.

청구항 12

제11항에 있어서, 상기 동영상 주석 방법은,

(f) 상기 인덱싱된 비디오 문장들에 대하여 주석을 기입하고 비디오 문장에 대한 메타데이터를 생성하는 단계;

(g) 상기 비디오 문단들에 대하여 주석을 기입하고 비디오 문단에 대한 메타데이터를 생성하는 단계;

를 더 구비하여, 동영상을 비디오 문장과 비디오 문단으로 인덱싱하고 주석을 기입하는 것을 특징으로 하는 동영상 주석 방법.

청구항 13

제11항에 있어서, 상기 (a) 단계는 인덱싱하고자 하는 동영상상을 로딩하고, 상기 로딩된 동영상상에 대하여 샷 경계를 검출(Shot Boundary Detection)하고, 샷 경계 검출을 통해 상기 동영상상을 샷 단위로 세그멘테이션하며, 상기 동영상상에 대한 샷 인덱스 파일을 생성하여 저장하는 것을 특징으로 하며,

상기 샷 인덱스 파일은 각 샷의 경계에 대한 시각 정보 및 프레임 정보를 포함하는 것을 특징으로 하는 동영상 주석 방법.

청구항 14

제11항에 있어서, 상기 (d) 단계는 스크립트 파일의 문장의 단위를 기준으로 하여 샷들을 그룹핑하여 비디오 문장을 설정하고, 상기 설정된 비디오 문장들을 인덱싱하는 것을 특징으로 하는 동영상 주석 방법.

청구항 15

제11항에 있어서, 상기 (e) 단계는 상기 비디오 문장들을 공통적인 주제에 따라 그룹핑하여 비디오 문단을 설정하는 것을 특징으로 하는 동영상 주석 방법.

발명의 설명

기술 분야

[0001] 본 발명은 동영상 주석 시스템에 관한 것으로서, 더욱 구체적으로는, 비디오 문장 및 비디오 문단을 동영상의 새로운 인덱싱 단위로 설정하고, 동영상을 새로운 인덱싱 단위인 비디오 문장 및 비디오 문단으로 인덱싱하고 주석을 기입하여 각 주제별로 효과적으로 검색할 수 있도록 하는 동영상 주석 시스템에 관한 것이다.

배경 기술

[0002] 최근 방송의 다채널화로 인해 방송 콘텐츠의 종류와 양이 방대해지면서 사용자들의 방송 콘텐츠에 대한 접근이 증가하고 있다. 특히, 시청자들의 흥미를 위한 주제를 가지고 있거나 정보 획득을 위한 목적을 가지고 있는 등 요구사항에 맞춘 콘텐츠가 지속적으로 발생하면서 방송 콘텐츠에 대한 관심과 발전은 나날이 증가할 것으로 전망된다. 이런 사용자들의 움직임은 디지털 기기의 보급과 지원하는 소프트웨어가 큰 역할을 하고 있다. '수동적 방송 시청' 방법인 방송 편성 일정에 따라 TV에 앉아서 프로그램을 보는 것이 아니라 '능동적 방송 시청' 방법으로 스마트폰, 태블릿, PC 등의 기기를 통해 제공되는 어플리케이션이나 웹 페이지를 통하여 원하는 시간에 원하는 부분을 시청할 수 있다. 이처럼 시청자의 요구사항을 잘 소화한 높은 질의 방송 콘텐츠가 누적되면서 빅 데이터를 형성하게 되고, 디지털 기기를 이용한 효과적인 소프트웨어가 방송 콘텐츠들을 분석하고 가공하여 사용자들이 '능동적'으로 정보를 획득할 수 있게 지원한다면 방송 콘텐츠에 대한 앞으로의 소비 전망은 더욱 밝을 것으로 기대된다.

[0003] 한편, 방송 콘텐츠에는 다양한 장르가 있는데, 방송 콘텐츠 중 다큐멘터리는 정보 전달을 목적으로 하는 방송 장르이다. 다큐멘터리는 실제 일어난 사건이나 과학적으로 실험한 과정 및 결과를 시청자들에게 전달하면서 가능한 사실적인 정보 전달을 목적으로 한다. 다큐멘터리 내에서도 주제에 따라 다양한 하위 장르를 갖는다. 휴먼 다큐멘터리의 경우 사람들의 사는 모습이나 사건 등을 주제로 하며, 동물 다큐멘터리의 경우 동물들의 주거지, 음식, 행동 등을 보여줌으로써 동물에 대한 정보를 전달한다. 각기 주제와 목적에 따라 구성이 달라지지만 다큐멘터리들의 주목적은 시청자들에게 유익한 정보를 제공하는 것에 있다. 예를 들어 역사 다큐멘터리는 문헌에 담겨있는 내용이나 역사학자들의 인터뷰 같은 신뢰성 있는 정보를 통해 시청자들에게 간접적인 경험이나 지식을 전달하는 역할을 하고 있다. 이처럼 다큐멘터리 장르가 시청자들에게 의도하는 바를 잘 정리하여 필요한 때에 활용할 수 있다면 특별한 지식 데이터 아카이브가 될 수 있다. 하지만 현재 다큐멘터리에서 정보를 얻는 방법에는 문제점이 있다. 필요한 정보를 얻기 위해 콘텐츠를 처음부터 재생되는 흐름에 따라 수동적인 시청을 해야 하는 단점이 있다. 즉, 영상 한편의 에피소드(Episode)를 제목만으로 검색은 가능하지만 해당 에피소드의 세부 내용들은 인덱싱(Indexing), 주석(Annotation) 그리고 분류되어 있지 않아 검색이 힘들다. 지속적인 인기를 통해 수백 회 방송되고 있는 다큐멘터리 프로그램도 있고, 단편 특집으로 방송되었지만 상당히 유용한 정보를 담고 있는 다큐멘터리 프로그램이 있음에도 불구하고 정보검색이 힘들다는 것은 인덱싱, 주석 그리고 분류를 통해 정

보에 대한 체계적인 라이브러리화가 되어 있지 않음을 의미한다.

- [0004] 다큐멘터리 장르는 객관적이고 사실성 있는 정보를 전달한다는 점에서 유용한 콘텐츠이다. 하지만 한 편의 에피소드 안에 있는 정보를 체계적으로 분류하고 있지 않아 정보를 효과적으로 찾아보기 힘든 문제점을 가지고 있다.
- [0005] 이하, 제안되고 있는 다양한 동영상 주식 시스템들 중 몇가지에 대하여 구체적으로 살펴본다.
- [0006] 'IPTV 환경에서의 방송콘텐츠 온톨로지 구축' 논문 (김형주, 김종덕 and 이중식, IPTV 환경에서의 방송콘텐츠 온톨로지 구축, 방송통신위원회 조사연구지원사업 보고서, pp. 1-113, 방송통신위원회 & 한국전파진흥원, 2008.)은 방송 콘텐츠의 라이브러리화를 강조하면서 하나의 프로그램을 소단위화시키고, 사용자들의 검색 환경에 맞는 주식 정보를 기입하여 정리되어 있는 방송 콘텐츠 검색 시스템을 위한 주식 모델을 설계하고 있다. 이 논문에서는 다양한 방송 콘텐츠 중에서도 동물을 주제로 한 다큐멘터리를 주제로 하여 접근하였고, 주 설계 방법을 크게 인덱싱과 주식 모델 방법 두 가지로 설명할 수 있다. 우선, 인덱싱 단위는 챕터(Chapter)와 씬(Scene)을 사용하였다. 해당 논문에서 언급하는 씬은 보통 30~60초의 길이를 가지며 길게는 300초가 되는 형태를 가진다. 씬은 보는 사람에 따라서 주관적이지만 해당 주식 시스템의 행위자는 콘텐츠 관리자가 되기 때문에 프로그램을 구성한 사람으로서 주관적이지만 보다 전문적인 구성을 기대할 수 있다.
- [0007] 해당 논문에 따른 주식 모델 설계 방법은 주식 정보가 되는 필드는 뉴스 태깅(Tagging)에 활용되는 육하원칙 중 가장 핵심적인 정보만을 간추린 물체(objects), 행위(actions), 장소(sites)를 활용하였다. 이 세 가지 정보는 주어(subject), 동사(Verb), 목적어(object)가 되어 자연스러운 문장을 이룰 수 있다. 예를 들어, 동물 다큐멘터리에 등장하는 '사자'는 물체가 되고, '식사한다'는 행위가 된다. 그리고 '정글 속, 나무 위'는 장소가 된다. 이렇게 하나의 문장이 되는 정보를 씬은 가지고 있고 사용자는 이 세 가지 정보를 통해 씬을 찾는 시나리오를 가지고 있다. 이 중 물체, 행위 필드는 시소러스(Thesaurus)를 조성하였다. 물체 필드에 시소러스를 활용한 이유는 콘텐츠 관리자가 미리 분류된 코드를 사용하여 빠르게 주식할 수 있고, 같은 의미지만 텍스트가 다른 정보 입력을 방지하고 통일된 주식이 가능하며, 생물학 분류를 가져왔기 때문에 전문적이기 때문이다. 물체 시소러스는 해당 시스템에서 최대 6Depth를 가지며 사용자의 수준에 맞추어 전문가를 대상으로 할 경우 보다 세분화된 시소러스로 변경하고, 일반 사용자일 경우 Depth를 낮춘 시소러스를 사용하는 융통성을 가질 수 있다. 행위 필드 역시 2Depth의 시소러스를 갖는다. 세부 뜻은 다르지만 유사한 행위간의 그룹을 형성하여 사용자가 검색을 할 때 계층적으로 찾을 수 있는 방법을 사용하였다. 각각의 시소러스에 해당하는 데이터들은 서로에게 해당하는 관계(Relation)를 통해 온톨로지(Ontology)를 형성한다. 즉, 여러 개의 주어, 동사, 목적어가 되는 데이터 간에 다차원적인 네트워크 구조를 형성한다. 이런 온톨로지 구조를 통해 사용자 성향에 따른 추천 및 관련 데이터를 제공할 수도 있다.
- [0008] 'Semantic Annotation and Retrieval of Documentary Media Objects]' 논문(D. Kanellopoulos, "Semantic Annotation and Retrieval of Documentary Media Objects," *The Electronic Library*, Vol. 30, No. 5, pp. 721-747, 2012.)은 다큐멘터리의 세부 장르를 구별하지 않고 효과적인 검색을 위해서 어떤 시멘틱 요소를 주식하였고, 다큐멘터리 콘텐츠를 기존의 씬이 아닌 새로운 인덱스를 사용하였는지를 보여 주고 있다.
- [0009] 인덱싱의 특징으로 새 인덱스 Documentary, DocumentaryClip, DocumentaryPiece를 사용하였다. 각 인덱스에 들어가는 주식 필드가 다르고 그에 따른 이벤트에 따라서 주식자의 주관에 따른 경계를 가지며 계층적인 구조를 통해 전체 다큐멘터리에 대한 메타데이터 가독성을 가진다. 최소 인덱스 단위인 DocumentaryPiece는 다큐멘터리에 등장하는 Subject 혹은 Documentarist의 Action 집합이 된다. 유사한 이벤트의 Action 리스트들이 하나가 되어 DocumentaryPiece가 된다. DocumentaryClip은 DocumentaryPiece의 집합으로 DocumentaryPiece에 등장하는 Documentarist의 성격에 대한 정보를 가지며 전후 DocumentaryClip에 대한 문맥적 의미, 장르적 특징 등 해당 계층에 대한 정보를 갖는다. 최상위 인덱스로 Documentary는 DocumentrayClip의 집합으로 해당 콘텐츠에 대한 타이틀, 포맷, 시간 등의 정보를 갖는다. 이처럼 계층적인 구조를 통해 각 계층에 맞는 정보를 분류하였고 한 편의 다큐멘터리에 대한 표면적이고 세부적인 살펴보기가 가능한 탐색에 대한 장점을 가지고 있다.
- [0010] 상기 논문은 새 인덱스를 사용하고 시멘틱 정보를 표현하는 데 있어 보다 자유롭고 더 많은 정보를 기입하기 위해 메타데이터는 MPEG-7 표준을 따르지 않았다. 하지만 메타데이터는 해당 시스템만의 DTD를 정의하여 XML Scheme를 따른다. 비록 주식자의 입장에서 주식 데이터를 기록하기는 편할 수 있으나 표준화하지 못해 다른 콘텐츠나 다른 검색 시스템에 적용하는 데 있어 확장성이 문제가 되어 해당 논문에서도 표준화를 향후 해결 과제로 하고 있다. 그리고 온톨로지를 통해 메타데이터 간의 관계를 정의하여 보다 관련 있는 정보를 사용자에게 반환하는 연구를 필요로 언급하고 있다.

- [0011] 'Ontologies for the Metadata Annotation of Stories' 논문(V. Lombardo and A. Pizzo, "Ontologies for the Metadata Annotation of Stories," *Proceedings of the Digital Heritage International Congress*, pp. 153-160, 2013.)은 스토리(Story)가 있는 장르의 콘텐츠를 효과적으로 주석하여 디지털 유산화(Digital Heritage)하는 것에 목적을 두고 있다. 특히 텍스트로 원작이 있는 드라마, 오페라, 연극 등의 콘텐츠를 장면 별로 Object는 물론 시멘틱 정보까지 모두 담는 것을 목표로 하고 있다.
- [0012] 상기 논문에 따른 주석 시스템에서의 인덱싱 방법을 먼저 살펴보면 인덱스 단위는 이벤트가 이루어지는 씬(Scene)을 대상으로 한다. 스토리를 이루고 있는 장르의 콘텐츠이기 때문에 이벤트의 전환이 장소변경이나 대화 중인 인물 구성원들의 변경으로 다큐멘터리 장르보다 명확하게 이루어진다. 다음으로 주석 모델의 형태는 온톨로지를 사용하여 콘텐츠의 Object 및 Semantic 요소를 필드로 정의하였으며, 온톨로지의 각 필드들은 다른 필드와 관계(Relation)을 형성한다. 이를 통해 필드간의 의미와 관계를 알 수 있다. 그리고 등장하는 인물을 중심으로 이루어지기 때문에 다큐멘터리와 다르게 인물의 감정, 심리상태 등의 시멘틱 정보가 들어간다. 스토리 중심의 콘텐츠 정보를 효과적으로 표현하기 위해서 필드의 구성이 바뀌는 것을 알 수 있다.
- [0013] 상기 논문에서, 추가적으로 사용자에게 반환되는 정보의 시각화를 언급할 수 있다. 스토리는 시간의 변화가 중요한 만큼 사용자가 주석 정보를 시간의 흐름에 따라 이해할 수 있게 등장인물 별로 이벤트 Plans가 구성되어 있다. 이처럼 콘텐츠의 특성에 따라 그리고 사용자의 편의를 위해 검색에서 반환되는 정보의 구조 또한 중요함을 언급하고 있다.
- [0014] 이하, 계층적 인덱스 설계와 관련하여, 종래의 샷(Shot)과 씬(Scene)을 이용한 인덱싱 방법에 대하여 살펴본다.
- [0015] 의학 관련 방송콘텐츠인 <생로병사의 비밀>의 에피소드 한 편을 주요 장면이 있는 부분으로 나누어야 사용자들이 그 부분들을 검색하여 정보를 얻을 수 있다. 즉, 전체 영상에 대하여 인덱싱 과정을 우선 거쳐야 부분 영상에 대하여 주석 처리를 할 수 있고 해당 주석 데이터를 통해 검색을 할 수 있다. 비디오 콘텐츠에 대해서 기본적으로 샷을 최소 세그멘테이션(Segmentation) 단위로 사용한다.
- [0016] 비디오 콘텐츠에서 샷(Shot)은 일반적으로 카메라가 한 번의 촬영을 시작하여 멈추기 전까지를 의미한다. 때문에 샷은 화면전환의 경계를 나타낼 수 있으며 비디오 콘텐츠를 나누는 기본적인 단위가 될 수 있다. 하지만 이런 샷으로 <생로병사의 비밀> 콘텐츠를 인덱싱할 경우 몇 가지 단점이 발생한다.
- [0017] 도 1은 종래의 샷 인덱스 사용의 문제점을 설명하기 위하여 도시한 개념도이다. 도 1에서 설명하는 것과 같이 샷으로 나누게 될 경우 음성 데이터의 문맥 손실(loss of context)이 발생한다. ①과 같이 <생로병사의 비밀>에 선 나레이터를 통한 내용 전개가 자주 등장하는데 질병과 관련된 설명을 하거나 출연자에 대한 소개를 미리 시작할 때 여러 가지 샷을 활용하여 내용을 전달한다. 즉, 나레이터는 내용을 전달할 때 한 문장 정도의 길이가 되는 음성 설명을 진행하면서 여러 가지 샷을 사용하기 때문에 만약 샷 단위로 이런 장면은 인덱싱 할 경우 음성 정보가 분할되어 어떤 내용을 전달하는지 알기 어렵다. 또한, ②와 같이 주석자가 음성 정보를 무시하고 비디오 정보만을 통해 주석하였을 때 왜곡된 주석 데이터를 생성하게 된다. 따라서 <생로병사의 비밀> 콘텐츠는 비디오 정보는 물론 음성 정보 역시 중요하며 두 정보를 모두 담을 수 있는 단위로 인덱싱을 하여야 한다.
- [0018] 다음으로 씬 단위 인덱싱을 할 수가 있다. 씬(Scene)은 샷 보다는 큰 단위로 물리적으로는 연속된 샷의 모임이라 할 수 있으며 샷과 씬을 모두 사용하여 계층적으로 콘텐츠를 인덱싱하였을 경우 사용자들이 보다 효과적인 검색을 유도할 수 있다. 의미적으로 카메라들이 촬영하는 동일한 장소 혹은 이벤트의 변경을 기점으로 하나의 씬을 정의할 수 있다. 장소의 변경을 씬의 경계점으로 할 경우보다 정확할 수 있지만 이벤트의 변경을 경계점으로 할 경우 씬을 구분하는 사람에 따라 주관적일 수 있다.
- [0019] 도 2는 종래의 인덱싱 방법에 따라 이벤트 전환을 씬 경계로 하였을 경우의 문제점으로 설명하기 위하여 도시한 개념도이다. 도 2에 도시된 바와 같이, 이런 이벤트에 대한 씬 경계의 주관성 때문에 <생로병사의 비밀>을 서로 다른 주석자가 인덱싱할 경우 한 편의 에피소드도 다른 의미와 크기로 부분될 수 있다. 만약 장소로 씬 경계를 구분한다 해도 앞서 샷에서 설명하였던 것처럼 특정 장면은 장소를 넘나들고 음성 정보를 지원할 수 있는 시각적으로 통일되지 않은 정보를 연속적으로 사용하기 때문에 모호성이 생길 수 있다. 따라서 <생로병사의 비밀> 콘텐츠를 씬 단위로 인덱싱 할 때는 주석자 사이에 이해할 수 있는 규칙을 두어 사용자가 검색을 효과적으로 할 수 있는 특징을 살리는 방향으로 진행해야 한다.
- [0020] 따라서, 본 발명은 동영상의 음성 정보의 문맥 손실을 방지하고 주석자가 보다 명확한 내용을 주석할 수 있도록 하는 동영상 주석 시스템을 제안하고자 한다.

선행기술문헌

특허문헌

- [0021] (특허문헌 0001) 한국공개특허공보 제 10-2002-0009757호
(특허문헌 0002) 한국공개특허공보 제 10-2014-0051412호
(특허문헌 0003) 한국등록특허공보 제 10-1398700호

발명의 내용

해결하려는 과제

- [0022] 전술한 문제점을 해결하기 위한 본 발명의 목적은 콘텐츠들이 가진 정보를 효과적으로 인덱싱하고 주석할 수 있는 동영상 주석 시스템 및 동영상 주석 방법을 제공하는 것이다.
- [0023] 본 발명의 다른 목적은 동영상을 인덱싱하고 주석할 수 있는 새로운 인덱싱 단위로 비디오 문장 및 비디오 문단을 설정하고, 상기 비디오 문장 및 비디오 문단 단위로 인덱싱하고 주석을 기입할 수 있는 동영상 주석 시스템 및 동영상 주석 방법을 제공하는 것이다.

과제의 해결 수단

- [0024] 전술한 기술적 과제를 달성하기 위한 본 발명의 제1 특징에 따른 동영상 주석 시스템은, 동영상을 비디오 문장(Video Sentence) 및 비디오 문단(Video Paragraph)으로 나누어 비디오 문장 단위 및 비디오 문단 단위로 인덱싱하고, 각 비디오 문장 및 비디오 문단에 대하여 주석을 기입하는 것을 특징으로 하며, 상기 비디오 문장은 단일의 화자에 의한 음성 설명에서 단일의 문장을 이루는 하나 또는 둘 이상의 샷으로 구성되는 비디오 구간으로 설정되며, 상기 비디오 문단은 공통적인 주제에 대한 하나 또는 둘 이상의 비디오 문장들로 구성되는 비디오 구간으로 설정되며, 상기 주석은 동영상의 콘텐츠의 특화된 정보를 고려하여 설계된 택소노미(taxonomy)를 기반으로 한 구조를 갖는 주석 데이터로 구성된다.
- [0025] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 동영상 주석 시스템은, 동영상을 로딩하고 상기 동영상에 대한 샷 인덱스 파일을 생성하는 비디오 제어부; 상기 동영상의 음성 데이터를 분석하여 텍스트 데이터로 구성된 스크립트 파일을 생성하는 스크립트 제어부; 상기 샷 인덱스 파일을 이용하여 샷 시퀀스(Shot Sequence)에 따른 샷들을 재생하는 비디오 재생부; 상기 샷 인덱스 파일과 스크립트 파일을 이용하여 비디오 문장을 형성하고 인덱싱하는 비디오 문장 인덱싱부; 상기 비디오 문장들을 그룹핑하여 비디오 문단을 형성하고 비디오 문단을 인덱싱하는 비디오 문단 인덱싱부; 상기 인덱싱된 비디오 문장들에 대하여 주석을 기입하고 비디오 문장에 대한 메타데이터를 생성하는 비디오 문장 주석부; 상기 비디오 문단들에 대하여 주석을 기입하고 비디오 문단에 대한 메타데이터를 생성하는 비디오 문단 주석부; 를 구비하여, 동영상을 비디오 문장과 비디오 문단으로 인덱싱하고 주석을 기입하는 것이 바람직하다.
- [0026] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 비디오 제어부는 인덱싱하고자 하는 동영상을 로딩하고, 상기 로딩된 동영상에 대하여 샷 경계를 검출(Shot Boundary Detection)하고, 샷 경계 검출을 통해 상기 동영상을 샷 단위로 세그멘테이션하며, 상기 동영상에 대한 샷 인덱스 파일을 생성하여 저장하는 것을 특징으로 하며, 상기 샷 인덱스 파일은 각 샷의 경계에 대한 시각 정보 및 프레임 정보를 포함하는 것이 바람직하다.
- [0027] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 스크립트 제어부는 상기 동영상에 대한 자막 파일을 불러와서 상기 스크립트 파일에 저장하는 것이 바람직하다.
- [0028] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 비디오 문장 인덱싱부는 스크립트 파일의 문장의 단위를 기준으로 하여 샷들을 그룹핑하여 비디오 문장을 설정하고, 상기 설정된 비디오 문장들을 인덱싱하는 것이 바람직하다.
- [0029] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 비디오 문단 인덱싱부는 상기 비디오 문장들을 공통적인 주제에 따라 그룹핑하여 비디오 문단을 설정하는 것이 바람직하다.
- [0030] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 비디오 문장 주석부는 각 비디오 문장에 대하여 비

디오 문장 타입, 특화 장면 주석 정보 및 특화 프레임 주석 정보를 기입하는 것이 바람직하다.

- [0031] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 비디오 문단 주석부는 각 비디오 문단에 대한 비디오 문단 타입과 비디오 문단 키워드 정보를 기입하는 것이 바람직하다.
- [0032] 전술한 제1 특징에 따른 동영상 주석 시스템에 있어서, 상기 동영상 주석 시스템은 사용자가 상기 동영상 주석 시스템의 각 부에 접속할 수 있도록 하는 사용자 인터페이스부를 더 구비하고, 상기 사용자 인터페이스부는, 상기 비디오 제어부에 접속하는 비디오 컨트롤 인터페이스부; 상기 스크립트 제어부에 접속하는 스크립트 컨트롤 인터페이스부; 상기 비디오 재생부에 접속하는 비디오 플레이어 인터페이스부; 상기 비디오 문장 인덱싱부에 접속하는 비디오 문장 인덱싱 인터페이스부; 상기 비디오 문단 인덱싱부에 접속하는 비디오 문단 인덱싱 인터페이스부; 상기 비디오 문장 주석부에 접속하는 비디오 문장 주석 인터페이스부; 상기 비디오 문단 주석부에 접속하는 비디오 문단 주석 인터페이스부; 를 구비하는 것이 바람직하다.
- [0033] 본 발명의 제2 특징에 따른 동영상 주석 방법은, (a) 동영상을 로딩하고 상기 동영상에 대한 샷 인덱스 파일을 생성하는 단계; (b) 상기 동영상의 음성 데이터를 분석하여 텍스트 데이터로 구성된 스크립트 파일을 생성하는 단계; (c) 상기 샷 인덱스 파일을 이용하여 샷 시퀀스(Shot Sequence)에 따른 샷들을 재생하는 단계; (d) 상기 샷 인덱스 파일과 스크립트 파일을 이용하여 비디오 문장을 형성하고 인덱싱하는 단계; (e) 상기 비디오 문장들을 그룹핑하여 비디오 문단을 형성하고 비디오 문단을 인덱싱하는 단계; (f) 상기 인덱싱된 비디오 문장들에 대하여 주석을 기입하고 비디오 문장에 대한 메타데이터를 생성하는 단계; (g) 상기 비디오 문단들에 대하여 주석을 기입하고 비디오 문단에 대한 메타데이터를 생성하는 단계;를 구비하여, 동영상을 비디오 문장과 비디오 문단으로 인덱싱하고 주석을 기입한다.
- [0034] 전술한 제2 특징에 따른 동영상 주석 방법에 있어서, 상기 (a) 단계는 인덱싱하고자 하는 동영상을 로딩하고, 상기 로딩된 동영상에 대하여 샷 경계를 검출(Shot Boundary Detection)하고, 샷 경계 검출을 통해 상기 동영상을 샷 단위로 세그멘테이션하며, 상기 동영상에 대한 샷 인덱스 파일을 생성하여 저장하는 것을 특징으로 하며, 상기 샷 인덱스 파일은 각 샷의 경계에 대한 시각 정보 및 프레임 정보를 포함하는 것이 바람직하다.
- [0035] 전술한 제2 특징에 따른 동영상 주석 방법에 있어서, 상기 (d) 단계는 스크립트 파일의 문장의 단위를 기준으로 하여 샷들을 그룹핑하여 비디오 문장을 설정하고, 상기 설정된 비디오 문장들을 인덱싱하는 것이 바람직하다.
- [0036] 전술한 제2 특징에 따른 동영상 주석 방법에 있어서, 상기 (e) 단계는 상기 비디오 문장들을 공통적인 주제에 따라 그룹핑하여 비디오 문단을 설정하는 것이 바람직하다.

발명의 효과

- [0037] 본 발명에 따른 동영상 주석 시스템은 동영상 중 특히 다큐멘터리 영상에서 중요하다 할 수 있는 객관적인 설명의 주체인 MC(Master of Ceremonies)나 나레이터(Narrator)의 음성 정보를 문맥 손실(loss of context) 없이 인덱싱할 수 있고, 제안하는 인덱스 단위인 비디오 문장과 비디오 문단은 주석자(Annotator)에게도 있어 규칙 있는 주석 데이터를 부여할 수 있게 하고 인덱스 사이의 경계를 구분할 수 있기 때문에 주석 구조의 일관성을 제공한다.
- [0038] 또한, 본 발명에 따른 동영상 주석 시스템은, 콘텐츠 분석과 사용자 시나리오 분석을 통해 사용자가 원하는 정보에 맞추어 주석 모델을 구성하여 높은 접근성을 기대할 수 있다. 이때, 주석 모델링에 사용한 태크노미 구조는 주석 정보에 대한 명확한 범위를 제공할 뿐만 아니라 계층적 구조를 통해 주석자 및 사용자에게도 높은 가독성과 활용성을 제공한다. 이로 인해, 기존 '수동적 방송 시청'의 문제점을 해결하고 다양한 주석 정보를 활용한 콘텐츠 소비 방법을 통해 사용자의 접근을 증대할 수 있다.
- [0039] 음성 정보와 비디오 정보를 모두 가지고 있는 비디오 문장은 <생로병사의 비밀>의 에피소드들에서 몇 가지 반복되는 형태를 가지고 등장한다. 비디오 문장의 이런 형태를 서로간의 공통점으로 묶어 범주화할 수 있다면 분류가 쉽고 사용자는 찾기가 쉽다. 특히, 에피소드 한 편당 발생하는 비디오 문장의 수가 100개 이상 되기 때문에 범주화 되어 있지 않을 경우 사용자가 쉽게 살펴보기(overview) 하는 것이 어렵고, 원하는 정보를 찾는 것이 쉽지 않다. 하지만 비디오 문장은 각기 형태를 가지고 있고 분류가 가능하기 때문에 모든 비디오 문장에 주석 정보를 기입하여 검색 시 누락되는 인덱스없이 찾아 볼 수 있다. 비디오 문장이 각기 형태를 가질 수 있는 이유는 다음과 같다. 첫째로, <생로병사의 비밀>이 건강을 주제로 한 콘텐츠인 만큼 목적이 명확하여 내용 구성이 한정적이라는 것이다. 둘째로, 사용자 접근성을 고려하여 검색 시 활용 가치가 높은 내용을 중심으로 분류할 수 있기 때문에 범주의 한계를 지정 가능하고 때문에 비디오 문장을 형태에 따라 분류할 수 있다.

도면의 간단한 설명

- [0040] 도 1은 종래의 샷 인덱스 사용의 문제점을 설명하기 위하여 도시한 개념도이다.
- 도 2는 종래의 인덱싱 방법에 따라 이벤트 전환을 썬 경계로 하였을 경우의 문제점으로 설명하기 위하여 도시한 개념도이다.
- 도 3은 본 발명에 따른 비디오 문장의 구조를 도시한 모식도이다.
- 도 4는 본 발명에 따른 비디오 문장 인덱스를 사용하는 경우의 장점을 도시한 것이다.
- 도 5는 본 발명에 따른 비디오 문단의 구조를 도시한 모식도이다.
- 도 6은 본 발명에 따른 비디오 문장과 비디오 문단의 내용에 대한 일 예를 도시한 것이다.
- 도 7은 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템을 전체적으로 도시한 구성도이다.
- 도 8는 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템에 있어서, 사용자 인터페이스부를 설명하기 위하여, 사용자 인터페이스부의 일예에 따른 화면을 캡처한 사진이다.
- 도 9는 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템에 있어서, 사용자 인터페이스부의 비디오 문장 주석 인터페이스부를 설명하기 위하여, 비디오 문장 주석 인터페이스부의 일예에 따른 화면을 캡처한 것이다.
- 도 10은 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템에 있어서, 사용자 인터페이스부의 비디오 문단 주석 인터페이스부를 설명하기 위하여, 비디오 문단 주석 인터페이스부의 일예에 따른 화면을 캡처한 것이다.
- 도 11은 본 발명에 따라 정보 검색을 위한 주석 및 검색 시스템을 전체적으로 도시한 구성도이다.
- 도 12는 방송 콘텐츠 <생로병사의 비밀>에 대한 비디오 문장 타입 텍소노미를 예시적으로 도시한 그래프이며, 도 13은 비디오 문단 타입 텍소노미를 예시적으로 도시한 그래프이며, 도 14는 비디오 문단 키워드 텍소노미를 예시적으로 도시한 그래프이다.
- 도 15는 종래의 방송 콘텐츠 주석 방법과 본 발명에 따른 주석 방법과의 특징을 비교한 표이다.

발명을 실시하기 위한 구체적인 내용

- [0041] 본 발명에 따른 동영상 주석 시스템은 새로운 비디오 인덱싱 단위로 비디오 문장(Video Sentence)과 비디오 문단(Video Paragraph)을 정의하고, 상기 비디오 문장 및 비디오 문단의 단위로 인덱싱을 수행한다. 상기 비디오 문장은 한 명의 화자가 하나 혹은 여러 샷을 통해 한 문장 단위의 음성 설명을 하는 비디오 구간을 의미하며, 상기 비디오 문단은 하나 이상의 비디오 문장으로 이루어진 여러 명의 화자가 하나의 소주제를 두고 말하는 구간이며 다른 주제로 변경되기 전까지를 경계로 한다. 본 발명에 따른 인덱싱 방법은 계층적인 구조를 가지고 있어 의학적인 설명 및 소주제에 맞추어 한 편의 에피소드를 주제별 검색을 효과적으로 할 수 있도록 한다. 특히, 본 발명에 따른 동영상 주석 시스템은, 콘텐츠 사용자들의 정보 접근 의도와 콘텐츠에 특화된 정보를 고려하여 의학적인 텍소노미(Taxonomy)를 정의하고, 이런 텍소노미를 기반으로 한 주석을 하게 된다. 즉, 비디오 문장은 등장인물, 주제, 화면 내용에 맞추어서 여러 개의 비디오 문장 타입(Video Sentence Type)으로 분류하여 주석을 할 수 있도록 하며, 비디오 문단은 질병명 및 대상/개념 등에 해당하는 비디오 문단 타입(Video Paragraph Type)과 키워드(Video Paragraph Keyword) 주석을 할 수 있도록 한다. 그리고 비디오 문단 타입 및 키워드 구조는 주석자에 의해 계속 추가/확장할 수 있도록 한다.
- [0042] 또한, 본 발명에 따른 동영상 주석 시스템에 있어서, 주석 방법은 콘텐츠의 특징 분석을 기반으로 한 결과를 통해 사용자에게 보다 필요한 정보를 중점으로 설계하였으며 텍소노미(Taxonomy)를 사용하여 주석 정보를 분류 및 구조화한다. 본 발명은 이러한 인덱싱 및 주석 방법을 통해 콘텐츠에 특화된 새로운 주석 도구를 구현하고 주석 도구를 통해 생성된 메타데이터의 구조를 정의한다.
- [0043] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템의 구성 및 동작에 대하여 구체적으로 설명한다.
- [0044] 먼저, 본 발명에서 인덱싱 단위로 새로이 정의되는 비디오 문장 및 비디오 문단에 대하여 구체적으로 설명한다. 도 3은 본 발명에 따른 비디오 문장의 구조를 도시한 모식도이다. 도 3을 참조하면, 비디오 문장은 개념적으로 하나의 화자가 하나 혹은 여러 샷을 통해 한 문장 단위의 음성 설명을 포함하는 구간을 의미한다. 비디오 문장

의 경계는 다른 화자로 전환되는 시점을 경계로 하며, 같은 화자가 음성 설명을 이어갈 경우 다른 내용을 말하기 전까지를 경계로 한다.

- [0045] 도 4는 본 발명에 따른 비디오 문장 인덱스를 사용하는 경우의 장점을 도시한 것이다. 도 4를 참조하면, 비디오 문장을 인덱싱 단위로 설정하는 경우, 음성 정보의 보존이 가능하며 비디오 및 음성 정보를 동시에 반영할 수 있게 된다.
- [0046] 다음, 본 발명에 따른 비디오 문단은 썸의 단점으로 보완하고 비디오 문장과 계층적인 구조를 형성하여 효과적인 검색을 유도하기 위하여 정의된 것이다. 도 5는 본 발명에 따른 비디오 문단의 구조를 도시한 모식도이다. 도 5를 참조하면, 비디오 문단(Video Paragraph)은 동영상의 내용을 끊어서 구분한 한 토막으로, 연속된 비디오 문장들 혹은 하나의 비디오 문장으로 구성되며 해당 비디오 문장들이 가진 공통점을 통해 비디오 문단을 생성할 수 있다. 비디오 문단은 개념적으로 여러 명의 화자가 하나의 소주제를 두고 말하는 구간이며 다른 주제로 변경되기 전까지를 경계로 한다. 비디오 문단 인덱스의 경계는 비디오 문장의 내용과 같이 자주 반복되는 패턴과 효과적인 검색을 고려하여 소주제를 범주화하였고, 주석자는 이런 범주의 구조를 고려하여 주관적인 썸과 달리 보다 명확한 인덱싱을 할 수 있다.
- [0047] 한편, 상기 주석 모델은 동영상의 콘텐츠의 특화된 정보를 고려하여 설계된 택소노미(taxonomy)를 기반으로 한 구조를 갖는 주석 데이터로 구성된다.
- [0048] 도 6은 본 발명에 따른 비디오 문장과 비디오 문단의 내용에 대한 일 예를 도시한 것이다. 도 6을 참조하면, 의학 관련 다큐멘터리 방송 프로그램인 <생로병사의 비밀> 한 편의 에피소드를 소주제를 가진 여러 개의 비디오 문단으로 분류할 수 있고, 하나의 비디오 문단은 공통된 소주제를 가졌으나 각기 다른 내용 형식을 가진 비디오 문장으로 분류될 수 있다. 이처럼 에피소드를 비디오 문단과 비디오 문장으로 분류하여 큰 범주에서 작은 범주로 내용을 검색할 수 있는 구조를 갖출 수 있다. 비디오 문장이 지나치게 과분할(over segmentation) 되었을 때 비디오 문단은 비디오 문장들을 살펴보기(overview) 할 수 있게 돕기 때문에 검색에 대한 계층적 구조의 장점을 확인할 수 있다.
- [0049] 도 7은 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템을 전체적으로 도시한 구성도이다.
- [0050] 도 7을 참조하면, 본 발명에 따른 동영상 주석 시스템(10)은 사용자 인터페이스부(100), 비디오 제어부(120), 스크립트 제어부(130), 비디오 재생부(140), 비디오 문장 인덱싱부(150), 비디오 문단 인덱싱부(160), 비디오 문장 주석부(170) 및 비디오 문단 주석부(180)를 구비하여, 임의의 동영상을 비디오 문장과 비디오 문단으로 인덱싱하고 주석을 기입하는 것을 특징으로 한다. 이하, 상기 동영상 주석 시스템(10)을 구성하는 각 부의 동작에 대하여 보다 구체적으로 설명한다.
- [0051] 도 8는 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템에 있어서, 사용자 인터페이스부를 설명하기 위하여, 사용자 인터페이스부의 일예에 따른 화면을 캡처한 사진이다. 도 8을 참조하면, 상기 사용자 인터페이스부(100)는 사용자가 상기 동영상 주석 시스템의 각 부에 접속할 수 있도록 하는 것으로서, 상기 비디오 제어부에 접속하는 비디오 컨트롤 인터페이스부(102), 상기 스크립트 제어부에 접속하는 스크립트 컨트롤 인터페이스부(103), 상기 비디오 재생부에 접속하는 비디오 플레이어 인터페이스부(104), 상기 비디오 문장 인덱싱부에 접속하는 비디오 문장 인덱싱 인터페이스부(105), 상기 비디오 문단 인덱싱부에 접속하는 비디오 문단 인덱싱 인터페이스부(106), 상기 비디오 문장 주석부에 접속하는 비디오 문장 주석 인터페이스부(107) 및 상기 비디오 문단 주석부에 접속하는 비디오 문단 주석 인터페이스부(108)을 구비한다.
- [0052] 상기 비디오 제어부(120)는 인덱싱하고자 하는 동영상 파일을 로딩하는 비디오 로더(Video Loader), 및 샷 경계 검출 모듈을 구비한다. 상기 샷 경계 검출 모듈은 비디오 로더에 의해 로딩된 동영상 파일에 대하여 샷 경계를 검출(Shot Boundary Detection)하고, 샷 경계 검출을 통해 상기 동영상 파일을 샷 단위로 세그멘테이션하며, 상기 동영상을 구성한 각 샷들에 대한 샷 인덱스 파일(122)을 생성하여 저장한다.
- [0053] 상기 샷 인덱스 파일은 각 샷의 경계에 대한 시각 정보 및 프레임 정보를 포함한다.
- [0054] 상기 스크립트 제어부(130)는 PCM 모듈 및 스크립트 라이터/로더를 구비하여 스크립트 파일(132)을 생성하여 저장한다. 상기 PCM 모듈은 상기 동영상의 음성 데이터인 오디오 PCM 정보를 분석하여 텍스트 데이터로 구성된 스크립트 파일(132)을 생성하여 저장한다. 상기 스크립트 라이터/로더(132)는 상기 동영상에 대한 자막 파일을 불러와서 스크립트 파일에 저장할 수 있으며, 사용자 창을 통해 스크립트를 확인하고 수정할 수도 있다.

- [0055] 상기 비디오 재생부(140)는 상기 샷 인덱스 파일(122)을 이용하여 샷 시퀀스(Shot Sequence)에 따른 샷들을 재생한다.
- [0056] 상기 비디오 문장 인덱싱부(150)는 상기 샷 인덱스 파일과 스크립트 파일을 이용하여 비디오 문장을 형성하고 인덱싱한다. 즉, 상기 비디오 문장 인덱싱부(150)는 스크립트 파일의 문장의 단위를 기준으로 하여 샷들을 그룹핑하여 비디오 문장을 설정하고, 상기 설정된 비디오 문장들을 인덱싱한다.
- [0057] 상기 비디오 문단 인덱싱부(160)는 공통적인 주제에 따라 상기 비디오 문장들을 그룹핑하여 비디오 문단을 형성하고 비디오 문단을 인덱싱한다.
- [0058] 상기 비디오 문장 주석부(170)는 상기 인덱싱된 비디오 문장들에 대하여 주석을 기입하고 이를 토대로 하여 비디오 문장 메타데이터를 생성한다. 도 9는 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템에 있어서, 사용자 인터페이스부의 비디오 문장 주석 인터페이스부를 설명하기 위하여, 비디오 문장 주석 인터페이스부의 일예에 따른 화면을 캡처한 것이다. 도 9를 참조하면, 상기 비디오 문장 주석부가 각 비디오 문장에 대하여 기입하는 주석은 비디오 문장 타입, 특화 장면 주석 정보 및 특화 프레임 주석 정보를 포함할 수 있다.
- [0059] 상기 비디오 문단 주석부(180)는 상기 비디오 문단들에 대하여 주석을 기입하고 이를 토대로 하여 비디오 문단 메타데이터를 생성한다. 도 10은 본 발명의 바람직한 실시예에 따른 동영상 주석 시스템에 있어서, 사용자 인터페이스부의 비디오 문단 주석 인터페이스부를 설명하기 위하여, 비디오 문단 주석 인터페이스부의 일예에 따른 화면을 캡처한 것이다. 도 10을 참조하면, 상기 비디오 문단 주석부가 각 비디오 문단에 대하여 기입하는 주석은 각 비디오 문단에 대한 비디오 문단 타입과 비디오 문단 키워드 정보를 포함한다.
- [0060] 이하, 전술한 구성을 갖는 본 발명에 따른 동영상 주석 시스템을 의학 관련 방송 콘텐츠인 <생로병사의 비밀>에 적용하여 실제로 실시한 과정을 구체적으로 설명한다.
- [0061] 먼저, 방송 콘텐츠 <생로병사의 비밀>의 특징을 분석하면, 첫 번째 특징은 '① 방대한 에피소드 데이터'이며, 두 번째 특징은 '② 건강을 주제로 한 콘텐츠'라는 것이다. 세 번째 특징은 '③ 패턴이 있는 내용 구조'로서, 시청자의 연령대가 워낙 다양하고 건강 정보를 보다 습득하기 쉬운 형태로 전달하려는 방송 의도가 있기 때문에 일정한 구조의 내용이 반복되는 형태를 지닌다. 네 번째 특징은 '④ 다양하고 유용한 정보 구성'으로서, 사람이 육안으로 확인하기 힘든 신체 내부의 장기 장면이나 매우 작은 바이러스의 구조 및 작용 장면 등을 묘사하여 알기 쉽게 설명해 주는 그래픽 장면, 신뢰성을 주어 시청자로 하여금 믿을 수 있는 정보로 재 활용할 수 있게 해주는 출처가 분명한 통계 자료(그래프, 표 등), 전문가들의 인터뷰 장면을 들 수 있다.
- [0062] <생로병사의 비밀>에 대한 특징 분석한 다음, 사용자 및 시청자들의 콘텐츠 접근 의도를 파악해야 한다. 전술한 바와 같이, 콘텐츠의 특징 분석과 콘텐츠 접근 의도를 파악하고 콘텐츠 기술자(Contents Description) 모델링을 하고, 이를 바탕으로 구체적인 전체 시스템을 설명한다.
- [0063] 도 11은 본 발명에 따라 정보 검색을 위한 주석 및 검색 시스템을 전체적으로 도시한 구성도이다.
- [0064] 도 11을 참조하면, 먼저 ① Contents Database가 가진 <생로병사의 비밀> 에피소드(Episode)는 우선 주석 시스템에 입력된다. ② 에피소드는 샷 단위로 최초 세그멘테이션(Segmentation)이 이루어진다. 샷은 현 시스템에서 인덱싱을 위한 가장 작은 단위로 사용되며 검색 시스템에서 결과 반환에 사용되는 가장 작은 단위는 아니다. 따라서 본 발명에서 주장하는 동영상 검색 단위로 만들기 위해 샷들의 정보를 통해 ③ 비디오 문장과 비디오 문단의 새로 제안하는 단위로 그룹 과정을 수행한다. 의미 단위로 그룹된 비디오 문장과 비디오 문단은 ④ 주석자(Annotator)에 의해 주석 구조에 맞추어 정보가 입력된다. ⑤ 입력되어진 정보는 주석 도구의 메타데이터 가공 과정을 통하여 Annotation Database로 저장된다. Annotation Database의 정보는 Contents Database 정보와 동기화 되어 검색 시스템에서 함께 사용된다. ⑥ 검색 시스템에선 검색 질의를 입력받아 Annotation Database와 Contents Database에 접근하게 되고 Database의 반환 결과를 통해 ⑦ 사용자에게 정보를 보여주게 된다.
- [0065] <생로병사의 비밀> 에피소드 한 편을 주요 장면이 있는 부분으로 나누어야 사용자들이 그 부분들을 검색하여 정보를 얻을 수 있다. 즉, 전체 영상에 대하여 인덱싱 과정을 우선 거쳐야 부분 영상에 대하여 주석 처리를 할 수 있고 해당 주석 데이터를 통해 검색을 할 수 있다. 일반적으로, 비디오 콘텐츠에 대해서 기본적으로 샷을 최소 세그멘테이션(Segmentation) 단위로 사용한다. 하지만 <생로병사의 비밀> 에피소드의 검색 단위를 샷으로 하였을 경우 장면이 가진 의미를 이해하기가 어려움이 있어 해당 장에선 비디오 문장과 비디오 문단이라는 인덱싱 단위를 새롭게 정의하고 제안한다. 새로운 단위를 주장한 이유는 해당 콘텐츠를 보다 의미 단위로 효과적으로 나눌 수 있고 내용상 구조가 명확하고 새 단위 형태로 반복되는 패턴을 가지고 있다는 것으로 사용자는 비디오

문장과 비디오 문단 단위 검색을 통해 효과적으로 정보를 검색할 수 있다.

[0066] <생로병사의 비밀> 콘텐츠에 있어서, 샷으로 나누게 될 경우 음성 데이터의 문맥 손실(loss of context)이 발생한다. <생로병사의 비밀>에선 나레이터를 통한 내용 전개가 자주 등장하는데 질병과 관련된 설명을 하거나 출연자에 대한 소개를 미리 시작할 때 여러 가지 샷을 활용하여 내용을 전달한다. 즉, 나레이터는 내용을 전달할 때 한 문장 정도의 길이가 되는 음성 설명을 진행하면서 여러 가지 샷을 사용하기 때문에 만약 샷 단위로 이런 장면은 인덱싱할 경우 음성 정보가 분할되어 어떤 내용을 전달하는지 알기 어렵다. 또한, 주석자가 음성 정보를 무시하고 비디오 정보만을 통해 주석하였을 때 왜곡된 주석 데이터를 생성하게 된다. 따라서 <생로병사의 비밀> 콘텐츠는 비디오 정보는 물론 음성 정보 역시 중요하며 두 정보를 모두 담을 수 있는 단위로 인덱싱을 하여야 한다. 또한, <생로병사의 비밀> 콘텐츠를 썸 단위로 인덱싱할 때는 주석자 사이에 이해할 수 있는 규칙을 두어 사용자가 검색을 효과적으로 할 수 있는 특징을 살리는 방향으로 진행해야 한다.

[0067] 이러한 특징을 고려하여, 인덱스 단위를 비디오 문장 및 비디오 문단으로 구성하여, 음성 정보의 문맥 손실(loss of context)을 방지하고 주석자가 보다 명확한 내용을 주석할 수 있게 돕는다. 예를 들면, 인터뷰 장면의 경우 대부분 하나의 샷으로 이루어져 있고 인터뷰 대사만으로 충분히 장면의 의도를 알 수 있기 때문에 하나의 샷이 비디오 문장이 될 수도 있다. 도 3과 같이, 비디오 문장을 형성하고, 도 5와 같이 비디오 문단을 형성한다. <생로병사의 비밀> 한 편의 에피소드를 소주제를 가진 여러 개의 비디오 문단으로 분류할 수 있고, 하나의 비디오 문단은 공통된 소주제를 가졌으나 각기 다른 내용 형식을 가진 비디오 문장으로 분류될 수 있다. 이처럼 에피소드를 비디오 문단과 비디오 문장으로 분류하여 큰 범주에서 작은 범주로 내용을 검색할 수 있는 구조를 갖출 수 있다. 비디오 문장이 지나치게 과분할(over segmentation) 되었을 때 비디오 문단은 비디오 문장들을 살펴보기(overview) 할 수 있게 돕기 때문에 검색에 대한 계층적 구조의 장점을 확인할 수 있다.

[0068] <생로병사의 비밀> 콘텐츠의 특징을 반영하여 새로운 인덱스 단위인 비디오 문장과 비디오 문단을 정의 및 제안하여 에피소드를 나누었다. 이제 나누어진 인덱스를 보다 잘 검색할 수 있게 주석 과정을 통하여 정보를 기입해야 한다. 본 발명에 따른 동영상 주석 시스템은 비디오 혹은 오디오의 특징 추출을 통한 자동 주석 과정을 배제하고 주석자에 의한 수동적인 주석 과정을 전제로 하기 때문에 정보 입력 방법에는 ① 자유롭게 단어 혹은 문장을 key in하는 방법과 ② 주어진 범주에서 데이터를 선택하는 방법으로 간략화할 수 있다. 이 중 ①과 ②의 장점을 모두 반영하여 확장성이 있고 콘텐츠에 특화된 정보를 명확하게 찾을 수 있는 모델을 설계하였다. 또한 인덱스를 계층적으로 구성하고 주석 모델을 통해 각 인덱스를 분류하였을 경우 몇 가지 장점이 있다. 방대한 인덱스 데이터를 체계적으로 분류가 가능하며, 성격이 명확한 경우 분류가 쉽다. 그리고 모든 인덱스에 대한 최소 정보 주석이 가능하여 검색에서의 누락을 피할 수 있다. 이런 인덱스들을 범주화하기 위해 텍소노미를 사용하여 주석 시스템과 검색 시스템 구조의 중심을 형성하였고 텍소노미를 통해 사용자는 효과적인 검색을 할 수 있다. 그리고 비디오 문장에 나타나는 <생로병사의 비밀>에 특화된 유용한 정보가 있는데 구성한 텍소노미로 모든 정보를 담을 수 없는 경우가 있다. 이를 해결하고 유용한 정보에 대한 사용자 접근성 증가를 위해 특화된 주석 모델을 추가로 설계하였다.

[0069] 도 12는 방송 콘텐츠 <생로병사의 비밀>에 대한 비디오 문장 타입 텍소노미를 예시적으로 도시한 그래프이며, 도 13은 비디오 문단 타입 텍소노미를 예시적으로 도시한 그래프이며, 도 14는 비디오 문단 키워드 텍소노미를 예시적으로 도시한 그래프이다.

[0070] 다음, 비디오 문장과 비디오 문단에 있는 정보만으로 <생로병사의 비밀> 콘텐츠만 존재하는 특별한 장면의 특별한 정보를 담을 수는 없기 때문에, 우선 특별한 장면 중에 사용자들에게 검색 활용도가 높은 장면들을 선정하였다. 예를 들면, 전문의 인터뷰 장면, 그래픽 설명 장면, 생활속 건강 정보 장면 등이 여기에 포함된다. (a) 전문의 인터뷰 장면은 전문의 이름, 전문의 소속 병원, 전문의 소속 진료과 이렇게 3가지 정보를 주석하여 사용자들이 검색할 수 있는 모델을 설계하였다. (b) 그래픽 설명 장면은 어느 신체부위에 해당하는 장면인지 그리고 어떤 개념이나 대상이 영향을 미치는 장면인지가 자주 등장하기 때문에 신체부위 명칭과 개념 및 대상을 주석할 수 있는 모델을 설계하였다. 마지막으로 (c) 생활 속 건강 정보 장면은 앞서 설명하였듯 음식, 운동, 생활 습관을 주석할 수 있는 모델을 설계하였다. (a), (b), (c)의 주석 요소는 텍소노미로 지정하기에는 범위가 넓기 때문에 모두 주석자의 key in으로 이루어지며 key in된 목록은 저장을 통하여 주석자가 목록에서 다시 찾을 수 있게 설계하였다.

[0071] 다음, 특화 장면은 동영상을 통해 사용자들이 정보를 얻고 활용하는데 의미를 두고 있다. 동영상과 다르게 프레임을 활용하는 특화 방법이 있다. 특화 장면의 종류로는, (a) 의학 장치 출력 사진, (b) 그래픽 영상, (c) 통계 및 학술 자료 등이 포함되며, 사용자가 조사된 수치를 통해 누군가를 설득하거나 정보 전달용 자료로 사용할 경

우 효과적인 프레젠테이션 자료로 활용될 수 있다. (a) 의학 장치 출력 사진은 사용자의 사용의도를 반영하기 위해 출력에 사용된 의학 도구 명칭, 출력한 대상을 주석정보로 한다. (b) 그래픽 영상은 장면에서 드러난 신체 부위 그리고 개념 및 대상을 주석 정보로 한다. 마지막으로 (c) 통계 및 학술 자료는 자료의 주제, 출처, 조사 년도 그리고 형태(그래프, 표 등)를 주석 정보로 한다.

[0072] 지금까지 설계한 인덱스 구조와 주석 모델을 통해 필요한 정보를 정리하였다. 해당 설계 구조를 통해 구현될 주석 시스템의 반환 데이터인 메타데이터를 클래스 다이어그램으로 구성하였다. 우선 Episode 클래스를 시작으로 비디오 문단에 해당하는 여러 개의 VideoPara 클래스가 있고 VideoPara 클래스는 다시 여러 개의 비디오 문장에 해당하는 VideoSen 클래스를 갖는다. VideoPara 클래스는 비디오 문단 타입 텍소노미에 해당하는 Enumeration 클래스를 갖고, 비디오 문단 키워드 텍소노미에 해당하는 목록들과 각각의 Enumeration 클래스를 갖는다. VideoSen 클래스 역시 비디오 문장 타입 텍소노미에 해당하는 Enumeration 클래스를 갖고, 콘텐츠의 특화 장면(Char_) 및 프레임 주석 정보(Achieve_)에 대한 클래스를 갖는다.

[0073] 전문한 내용을 토대로 하여 동영상 주석 시스템을 구현하였으며, 도 8 내지 도 10은 사용자 인터페이스부의 화면, 비디오 문장 주석 인터페이스부, 비디오 문단 주석 인터페이스부의 일예에 따른 화면을 캡처한 것들이다.

[0074] 도 15는 종래의 방송 콘텐츠 주석 방법과 본 발명에 따른 주석 방법과의 특징을 비교한 표이다.

[0075] 도 15를 참조하면, 우선, 인덱스 활용 부분에서 새로운 인덱스는 설계하고 활용하였다. 기존 비디오 콘텐츠는 샷과 씬을 동영상 인덱스로 주로 활용하였지만 샷은 다큐멘터리에 등장하는 나레이터의 음성 정보 문맥 손실(loss of context)을 일으켜 장면에 대한 올바른 정보를 줄 수 없으며, 씬은 다큐멘터리 장르에서 경계를 기준으로 할 수 있는 규칙이 모호하기 때문에 주석자가 여러 명일 경우 인덱스 크기가 상이할 수 있다. ①은 다큐멘터리 장르 인덱싱을 위해 씬을 사용하고 있는데 나레이터의 음성 정보를 고려하지 않고 장면에 등장하는 물체의 행위만을 주석 정보로 활용하였기 때문에 다큐멘터리가 갖는 모든 정보를 활용하고 있지 않다. 하지만 ②는 다큐멘터리가 진행되는 내용을 고려하여 Documentarist와 Commentarist의 Actions 정보를 담은 새로운 인덱스 DocumentaryPiece를 주석하는 만큼 시멘틱 요소를 활용하고 있다. 본 발명에 따른 주석 방법도 <생로병사의 비밀> 콘텐츠에서 나레이터가 전달하는 정보의 중요성을 콘텐츠 분석을 통해 확인하였고 음성 정보를 통해 사용자들이 정보를 얻을 수 있도록 음성 정보의 시멘틱 전환을 경계로 하여 비디오 문장 단위를 제안하였다. 그리고 비디오 문장의 공통점을 활용하고 사용자들이 콘텐츠를 살펴보기(overview) 하면서 깊이 있게 탐색할 수 있도록 비디오 문장의 상위 인덱스인 비디오 문단을 정의하였다. 비디오 문단과 비디오 문장의 계층적인 구조를 통하여 사용자 입장에서 얼마나 구체적인 정보를 원하는지 혹은 정보의 특징에 따라서 얼마나 넓은 정보를 원하는지 접근이 용이하다.

[0076] 주석 모델링 과정에선 주석 정보 구조를 콘텐츠의 성격에 맞게 활용하였으며 확장가능하고 변화 가능한 방법을 활용하였다. ①의 주석 데이터를 정해진 필드 목록을 사용하는 방법은 주석자가 주어진 범위 안에서 빠르게 주석할 수 있고 지나치게 넓은 폭의 주석 정보가 발생하여 검색시 도태되어지는 정보가 생기는 문제를 방지할 수 있다. 반면, 단점으로는 필드의 목록이 다양하지 못할 경우 정보가 왜곡될 수 있으며 상세한 정보를 가질 수 없다는 면이 있다. 본 발명에 따른 주석 방법에서는 ①의 장점을 활용하고 단점을 보완하기 위해 일부 주석 구조에 정해진 주석 필드 목록을 사용하고 일부는 key in을 통해 확장할 수 있는 구조를 사용하였다. 본 발명에서 제안한 비디오 문장의 경우 콘텐츠에서 나올 수 있는 유형이 제한되어 있다. 그리고 지나치게 유형을 세분화할 경우 주석 비용이 증가하며 사용자가 검색하는데 일치하기 힘든 정보가 생길 수 있어 비디오 문장의 의미들을 공통점으로 묶어 계층적으로 표현하고 4Depth가 넘지 않는 텍소노미로 구성하였다. 구성된 텍소노미를 통해 주석자는 보다 제한된 필드에서 빠른 주석 처리가 가능하고 모든 비디오 문장에 대하여 누락없이 데이터가 발생하기 때문에 사용자는 정보를 효과적으로 찾을 수 있다. 반면 비디오 문단은 일부 제한적인 필드와 일부 확장 가능한 key in할 수 있는 필드를 설정하여 텍소노미를 구성하였다. 질병명과 등장하는 음식의 경우 콘텐츠에 맞춰 제한할 수 있는 정보라고 할 수 없다. 때문에 제한된 필드를 설정할 수는 없지만 <생로병사의 비밀>에서 중요한 주석 정보가 되는 요인이기 때문에 주석자에 의해 확장될 수 있는 구조를 갖도록 설정하였다. 이런 key in 형태의 주석 텍소노미는 일반적인 확장성뿐만 아니라 주석자 집단이 질병명의 의학적 분류를 하용하거나 혹은 음식 분류를 요리 협회와 같은 부분에서 활용한다면 특색에 맞게 전문적이거나 대중적으로 변화 가능하기 때문에 실용적이라 할 수 있다.

[0077] 메타데이터 스키마는 다른 연구들과 마찬가지로 메타데이터 표준을 따르지 않았다. 새로 제안한 인덱스를 보다 계층적으로 표현하여 XML 구조의 활용성을 높이고, 다른 연구에서 활용하고 있지 않은 <생로병사의 비밀> 콘텐츠만의 특화 정보를 주석 모델에 추가하였기 때문에 복잡해 질 수 있는 메타데이터의 구조를 단순화하기 위하여

표준화를 하지 않고 본 발명에 따른 시스템만의 XML DTD를 설정하였다. 하지만 다른 연구들이 시사하는 바와 마찬가지로 표준화가 되지 않으면 다른 콘텐츠 활용 시스템과 응용할 수 있는 메타데이터로 활용하기 어렵고 콘텐츠에 대한 추가적인 주석 정보가 생겼을 경우 혹은 기존 텍소노미를 활용한 주석 모델이 변경되었을 경우 시스템에 많은 비용이 발생할 수 있다.

[0078] 한편, 전술한 본 발명에 따른 동영상 주석 시스템은 컴퓨터 등에서 실행될 수 있는 프로그램으로 구현될 수 있으며, 따라서 본 발명에 따른 동영상 주석 시스템은 컴퓨터 등에 의해 관독되어 실행될 수 있는 프로그램으로서 컴퓨터에 의해 관독가능한 기록매체에 기록될 수 있다.

[0079] 이상에서 본 발명에 대하여 그 바람직한 실시예를 중심으로 설명하였으나, 이는 단지 예시일 뿐 본 발명을 한정하는 것이 아니며, 본 발명이 속하는 분야의 통상의 지식을 가진 자라면 본 발명의 본질적인 특성을 벗어나지 않는 범위에서 이상에 예시되지 않은 여러 가지의 변형과 응용이 가능함을 알 수 있을 것이다. 그리고, 이러한 변형과 응용에 관계된 차이점들은 첨부된 청구 범위에서 규정하는 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다.

산업상 이용가능성

[0080] 본 발명에 따른 동영상 주석 시스템은 시스템을 통해 생성된 구조적인 주석 정보를 바탕으로 하여 방송 콘텐츠 파일에 대한 주제별 검색을 효과적으로 지원할 수 있으므로, 동영상 검색 분야에 널리 사용될 수 있다.

부호의 설명

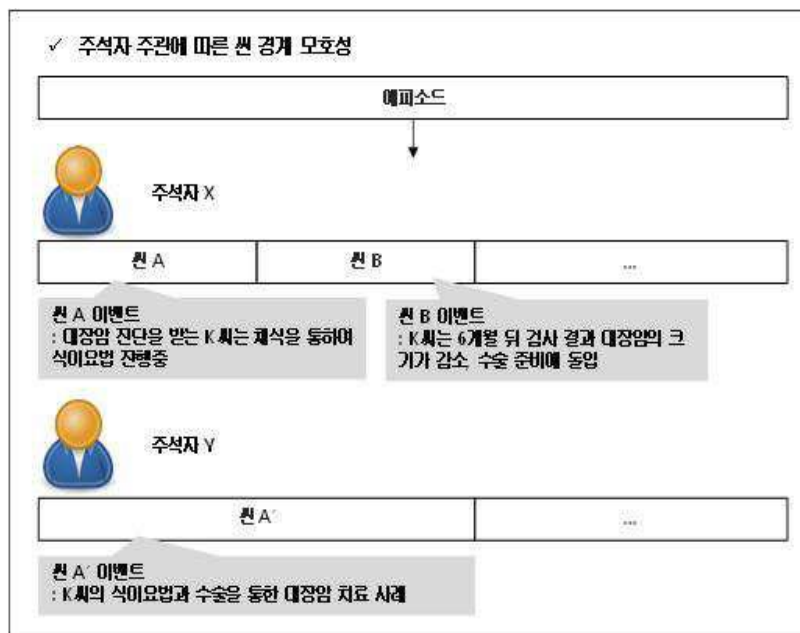
[0081] 10 : 동영상 주석 시스템
 100 : 사용자 인터페이스부
 120 : 비디오 제어부
 130 : 스크립트 제어부
 140 : 비디오 재생부
 150 : 비디오 문장 인덱싱부
 160 : 비디오 문단 인덱싱부
 170 : 비디오 문장 주석부
 180 : 비디오 문단 주석부

도면

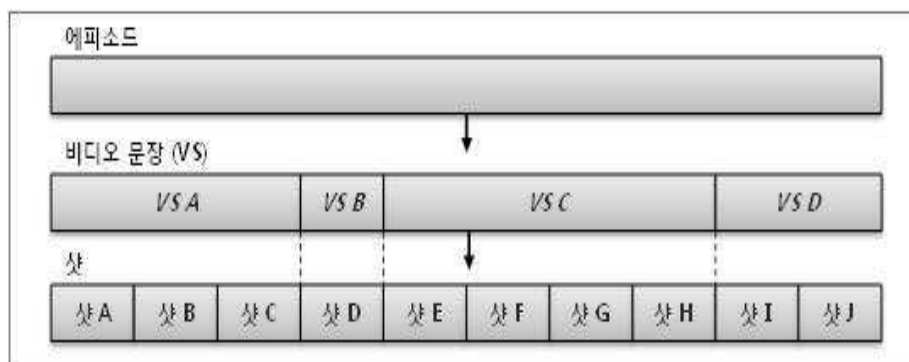
도면1



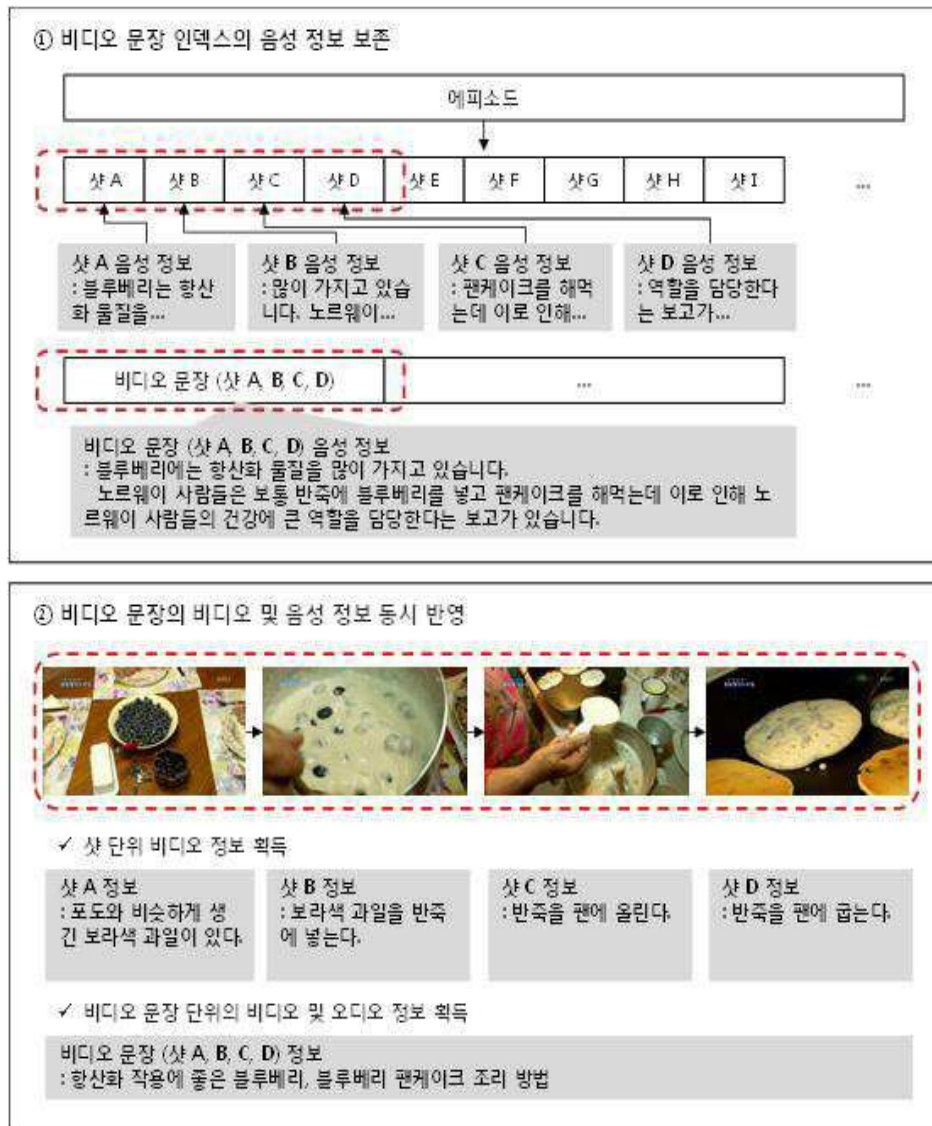
도면2



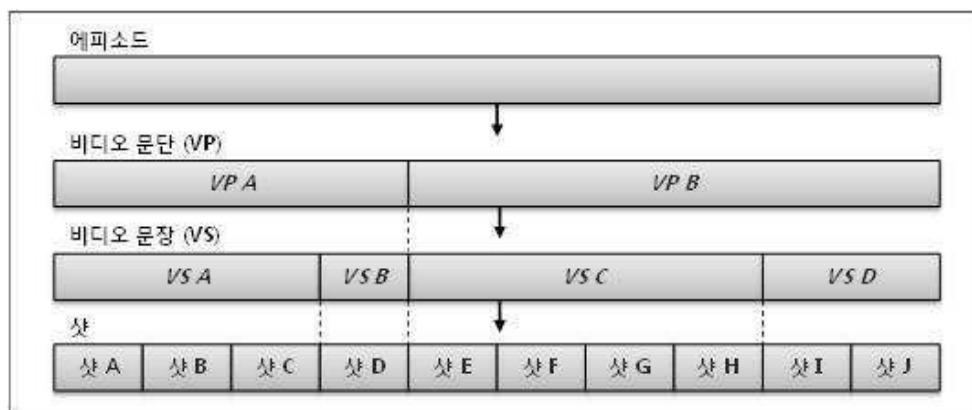
도면3



도면4



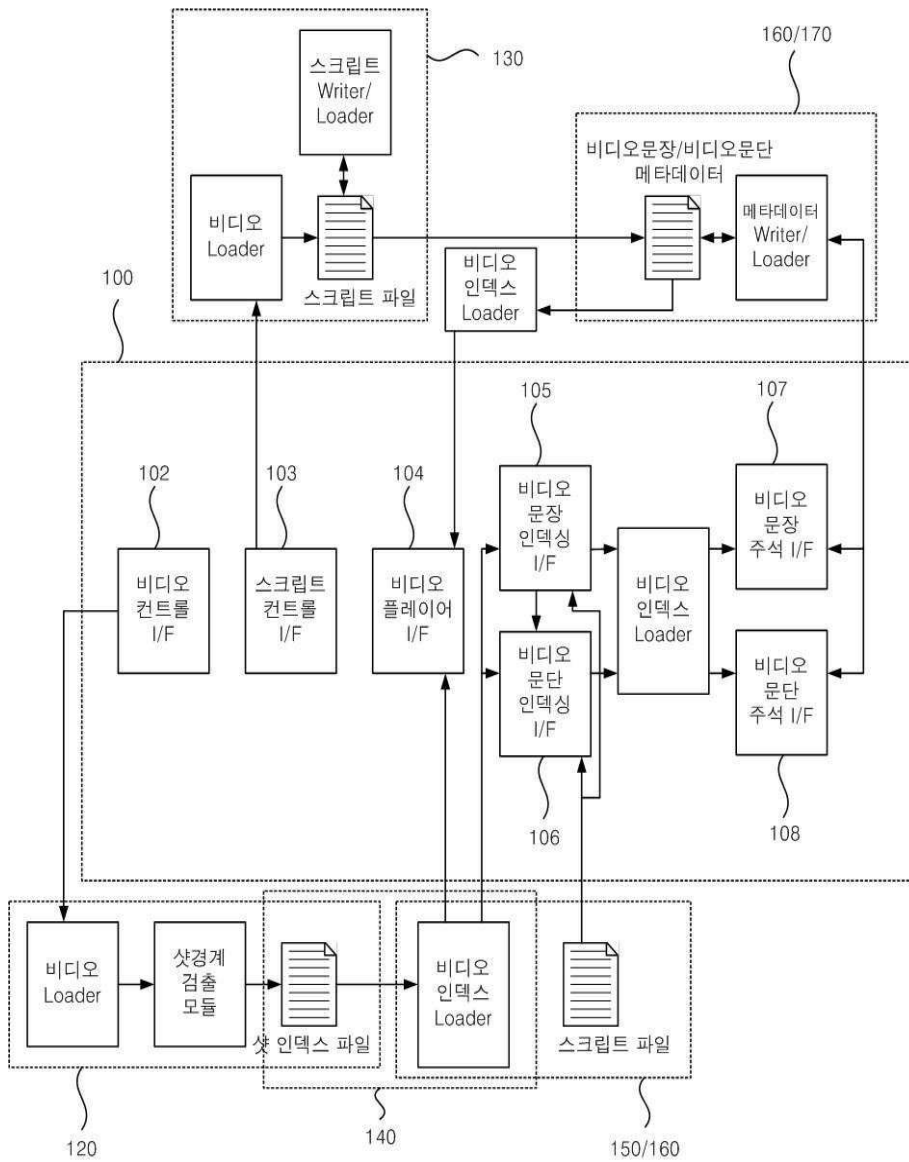
도면5



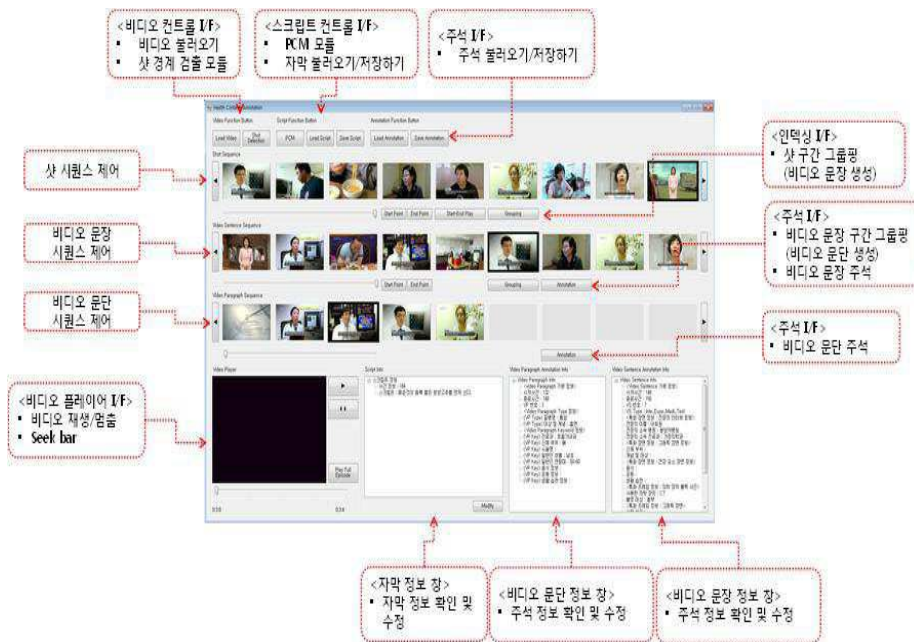
도면6

비디오 문장 번호	비디오 문장 키 프레임	비디오 문장 내용	비디오 문단 내용
#1		■ 나레이터의 환자 증 상 소개	■ 고연령자의 활성산소 관리 양호 사례
#2		■ 나레이터의 건강 긍정 요소 설명	
#3		■ 출연자의 긍정 습관 인터뷰	
#4		■ 나레이터의 출연자 검사 결과 언급	
#5		■ 나레이터의 최근 연 구 설명	■ 신약을 통한 활성산소 관리 양호 사례 및 전문 가의 견해
#6		■ 출연자의 증상 호전 상태 인터뷰	
#7		■ 나레이터의 학술 및 통계 자료 설명	
#8		■ 전문가의 치료 방법 인터뷰	

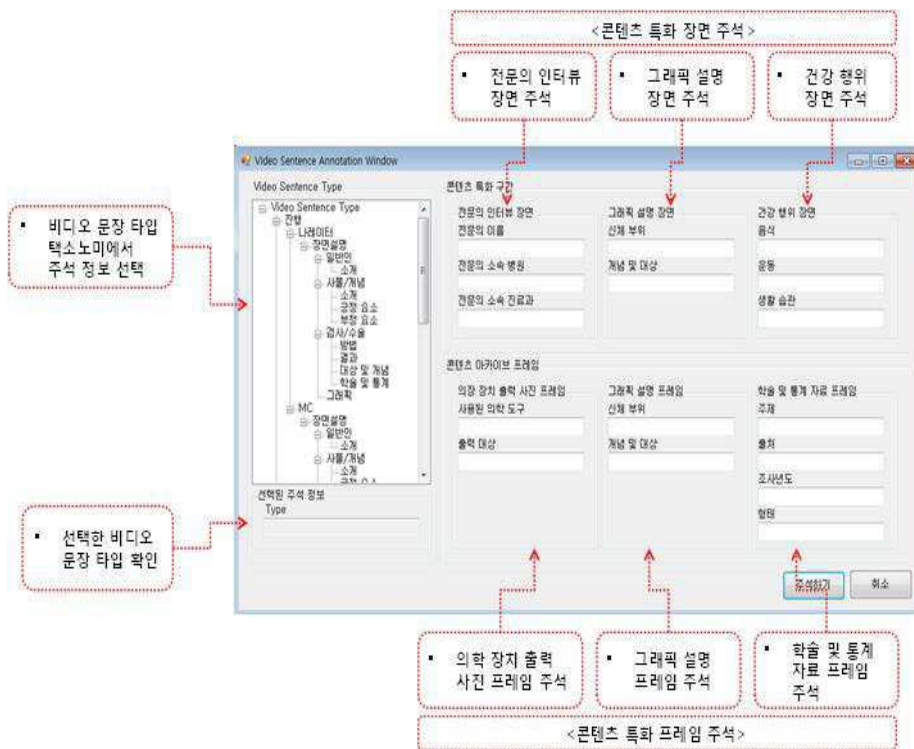
도면7



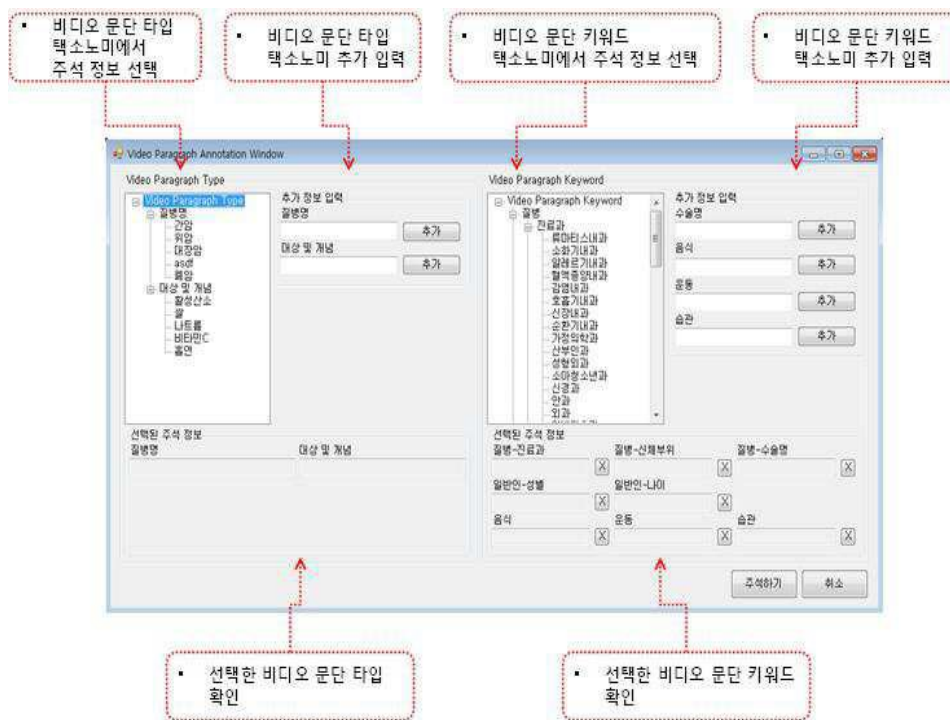
도면8



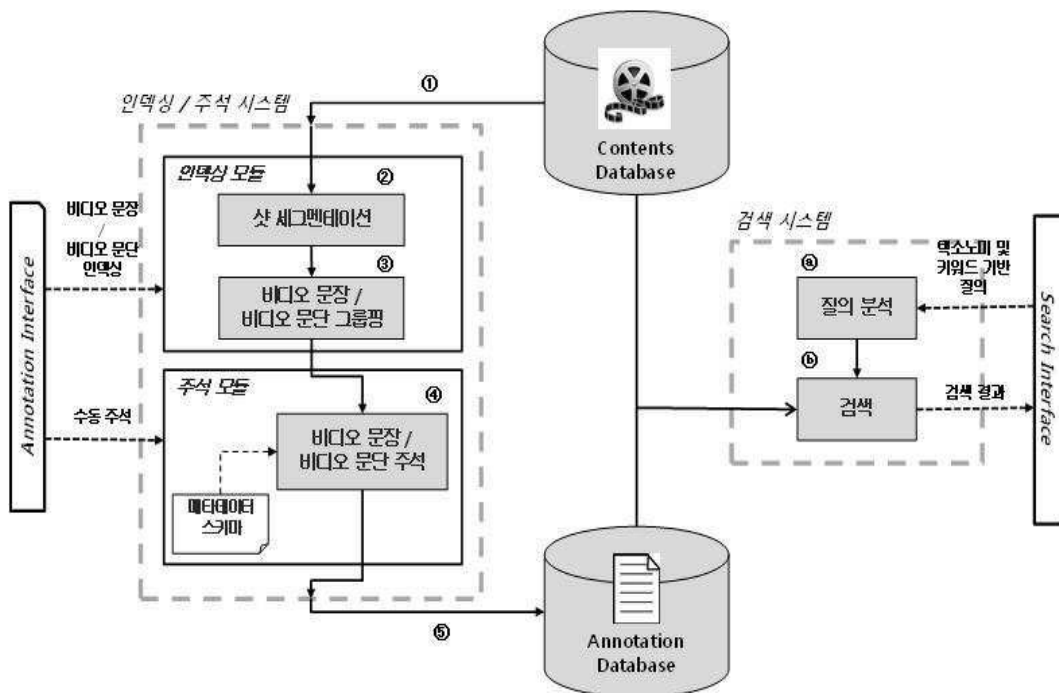
도면9



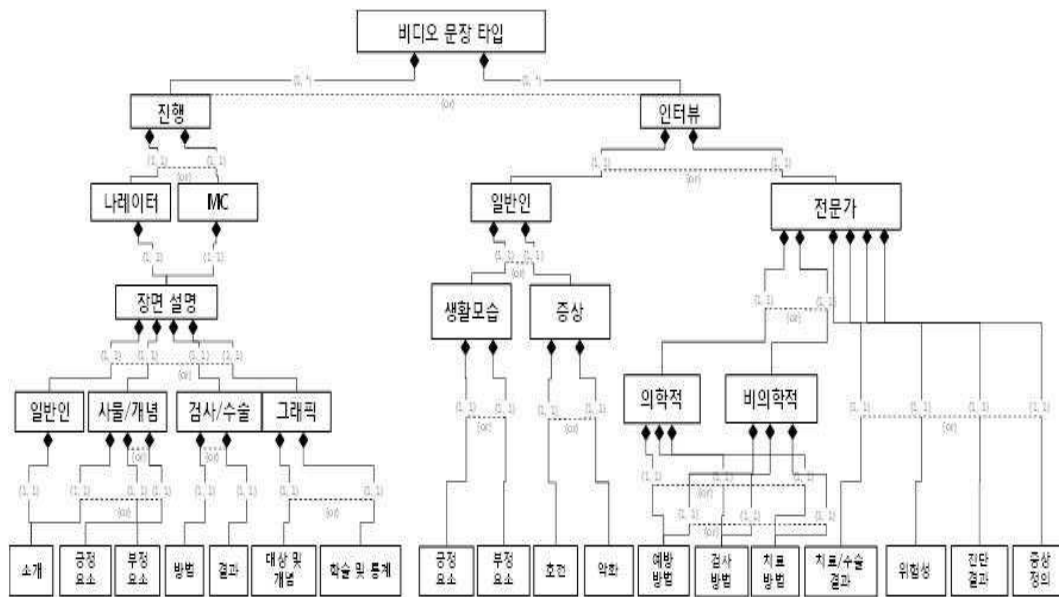
도면10



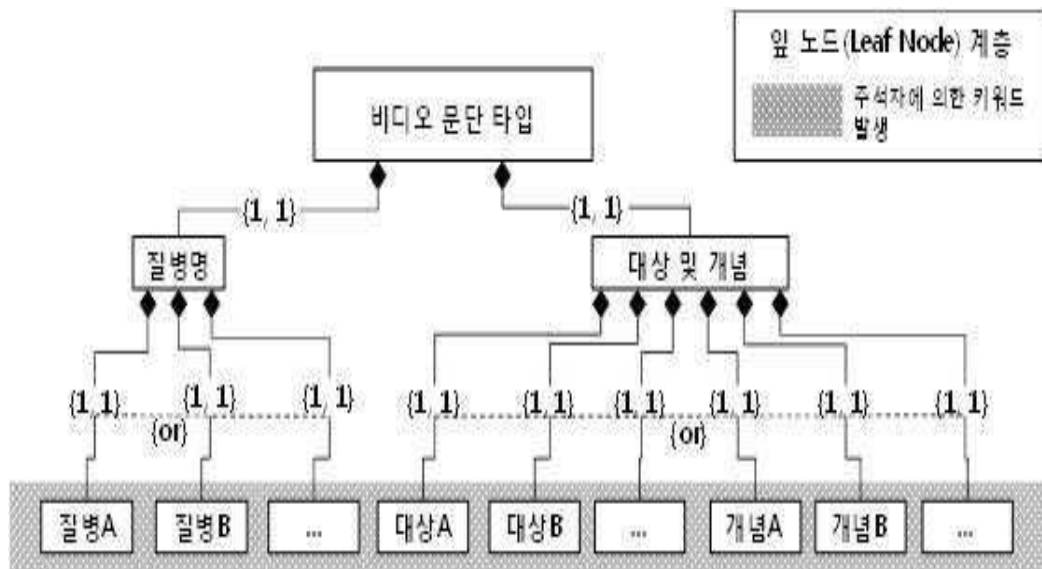
도면11



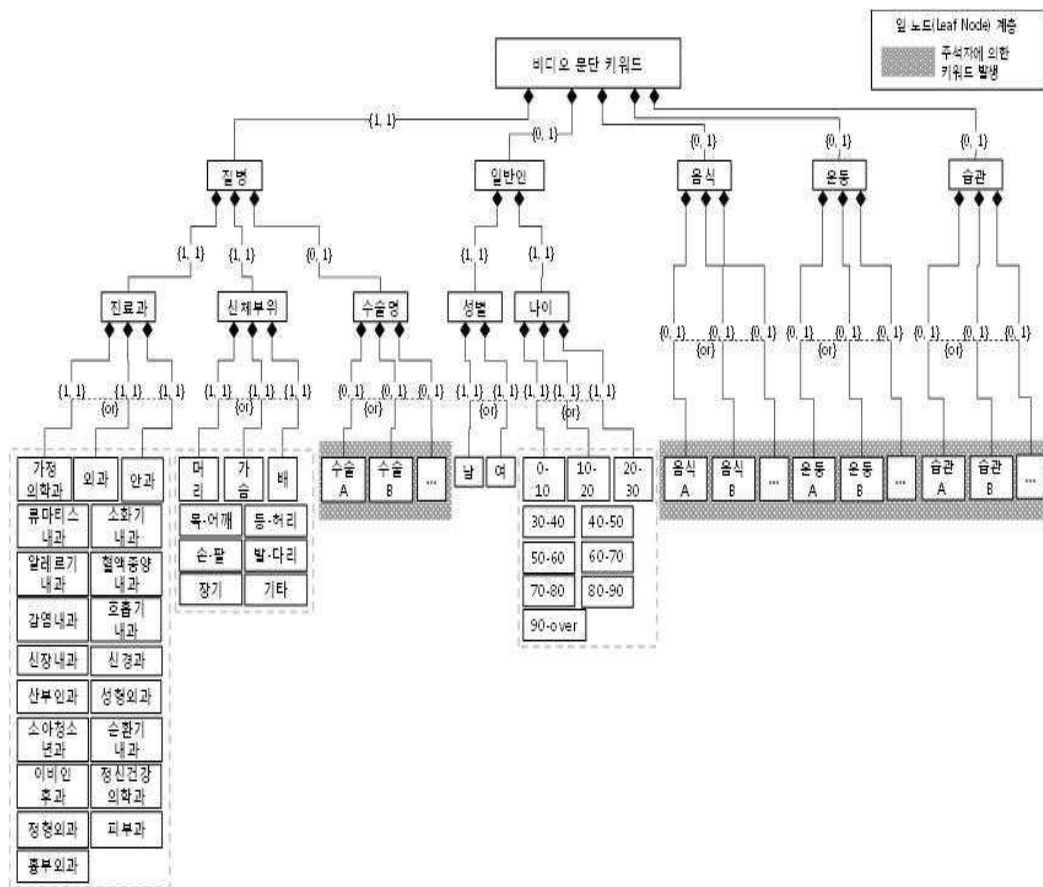
도면12



도면13



도면14



도면15

연구 논문	인덱스	주석 모델	메타데이터	대상 콘텐츠 특화 정보
① IPTV 환경에서의 방송콘텐츠 온톨 로지 구축 [9]	▪ 챕터(Chapter), 씬(Scene)	▪ 온톨로지 활용 ▪ 물체, 행위, 장소 필드로 주석 ▪ 주석 데이터는 정해진 필드 목록 을 사용	▪ 메타데이터 표준을 따르지 않음 ▪ 시스템 자체 XML DTD 표현	▪ 없음
② Semantic Annotation and Retrieval of Documentary Media Objects [13]	▪ Documentary /DocumentaryClip /DocumentaryPlace 새로운 인덱스 제안	▪ 각 인덱스마다 특화된 주석 필드 사용 ▪ 주석 데이터는 key in 방식을 통 해 자유롭게 사용	▪ 메타데이터 표준을 따르지 않음 ▪ 시스템 자체 XML DTD 표현	▪ 없음
③ Ontologies for the Metadata Annotation of Stories [14]	▪ 씬(Scene)	▪ 온톨로지 및 시소러스 활용 ▪ 인물 중심의 스토리 특징 주석 필드 사용 ▪ 주석 데이터는 key in 방식을 통 해 자유롭게 사용	▪ 메타데이터 표준을 따르지 않음 ▪ OWL 표현 (온톨로지 데이터)	▪ 없음
④ MPEG-7 Multimedia Description Scheme [17]	▪ MPEG7 Visual 및 Audio Descriptor에 따 른 구분 가능한 단위	▪ 각 주석 필드에 대한 정의 필요 (주석 정보에 대한 상세 스펙)	▪ 메타데이터 표준 ▪ XML 구조로 표현	▪ 메타데이터 정의 필요
본 발명	▪ 비디오 문장 (Video Sentence) / 비디오 문단 (Video Paragraph) 새로운 인덱스 제안	▪ 텍스트노미 활용 ▪ 각 인덱스마다 특화된 주석 필드 사용 ▪ 주석 데이터는 정해진 필드 + key in 방식을 모 두 활용	▪ 메타데이터 표준을 따르지 않음 ▪ 시스템 자체 XML DTD 표현	▪ 인덱스 주석 필드 이외에 특화 장면 주석 모델 활용