

비디오 이벤트 검색을 위한 온톨로지 기반 자동 주석 시스템¹⁾

임수민^{○*} 정진우* 주영도** 이동호*

한양대학교 컴퓨터공학과*, 강남대학교 컴퓨터미디어공학부**

{shannon[○], jwjeong, dhlee72}@hanyang.ac.kr*, ydjoo@kangnam.ac.kr**

OLiVE: Ontology-based approach for annotating video event

Sumin Lim^{○*} Jin-Woo Jeong* Joo, Young Do** Dong-Ho Lee*

Dept. of Computer Science & Engineering Hanyang University*,

Division of Computer Science and Media Engineering, Kangnam University**

{shannon[○], jwjeong, dhlee72}@hanyang.ac.kr*, ydjoo@kangnam.ac.kr**

1. 서론

최근 저장장치의 대용량화와 더불어 네트워크 대역폭에 대한 제한이 감소하면서 방대한 양의 비디오 콘텐츠를 쉽게 접할 수 있게 되었다. 이에 따라, 사용자의 다양한 질의들을 효과적으로 처리하기 위하여 대용량 멀티미디어 콘텐츠를 다양한 수준으로 자세하게 주석처리 하려는 연구들이 진행되고 있다. 그러나 기존의 자동 주석 시스템들은 객체 중심의 주석을 생성하기 때문에 비디오에서 객체간 발생할 수 있는 이벤트에 대한 검색에 한계가 있다. 객체를 중심으로 일어나는 이벤트에 대한 주석은 사용자로 하여금 보다 상세하고 효과적인 검색을 가능하게 한다.

대부분의 사용자들은 원하는 비디오를 찾기 위하여 객체 개념, 이벤트와 같은 고수준의 의미 정보를 표현하는 키워드를 이용한 검색을 시도한다. 예를 들어, "고양이", "호랑이", "구름", "말"과 같이, 비디오에 존재하는 객체들에 관한 정보를 나타내는 키워드를 이용하여 검색할 수도 있으며, 이벤트를 나타내는 키워드들을 이용한 "사자의 사냥이 담긴 영화" 또는 "빌딩이 무너지는 장면이 담긴 영화", "재난 영화"와 같은 질의가 있다. 본 연구는 이벤트를 이용한 질의를 가능케 하고, 사용자에게 보다 고수준의 비디오 검색을 제공하는 비디오 주석 시스템인 OLiVE (Ontology-based approach for annotating video event)를 제안한다. 시스템의 구조는 그림 1과 같으며 프로세스는 다음과 같다.

2. 시스템 구조

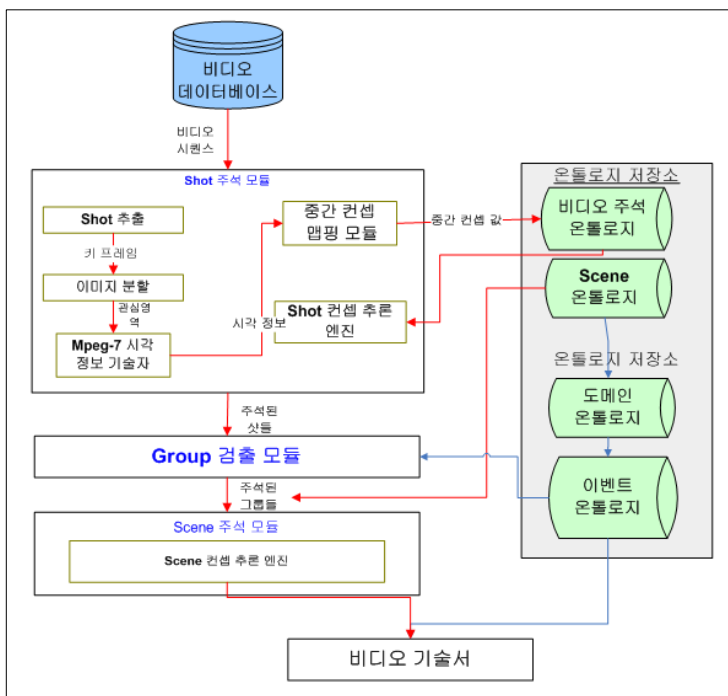


그림 1 제안하는 시스템의 구조도

OLiVE 시스템은 본 연구의 선행연구인 [1]의 객체 컨셉 추출과정을 바탕으로 객체의 움직임에 대한 정보를 추출하는 모듈을 추가함으로써, 객체의 모션과 패턴들을 분석하여 이벤트 추론에 이용한다. 또한, Scene 온톨로지와 도메인 온톨로지에 의미 추론 규칙을 적용하여 해당 Scene의 이벤트를 추출한다. 각각의 Shot, Group, Scene 수준 주석 생성 모듈에 의하여 추출된 정보는 각각의 Shot, Group, Scene 주석 온톨로지에 기술된다.

각 모듈별로 처리하는 과정은 다음과 같다. 첫째로, 그림 1에서 보이는 Shot 주석 모듈은 비디오 Shot에 존재하는 객체들의 컨셉 추출이며 비디오의 Shot으로부터 키 프레임이 추출되면, 시스템은 MPEG-7 시각 기술자들을 이용하여 해당 프레임에 존재하는 주요 객체의 저 수준 시각 특징 벡터를 추출한다. 이러한 저 수준 시각 특징 벡터는 중간-컨셉 맵핑 모듈을 통하여 해당되는 중간 컨셉 값으로 맵핑된다. 이 과정에서 시스템은 트레이닝 집합으로부터 각각의 객체를 인식하기 위한 의미 추론 규칙을 생성한다. 예를 들어, 컨셉 '치타'를 위한 의미 추론 규칙은 치타 객체의 대표적인 색상, 무늬, 모양에 해당하는 중간 컨셉 값들을 정의하고, 키 프레임

에 존재하는 객체가 추론규칙을 만족한다면, 해당 객체의 컨셉은 ‘치타’ 컨셉으로 결정된다. 두 번째부터는 [1]의 시스템을 이용하여 얻어진 객체를 기준으로 OliVE시스템에서 Group 수준으로 처리하는 과정을 수행한다. 객체와 연관된 Group의 추출 과정을 수행하는데 이벤트는 하나의 Group 안에서 두 개 이상의 객체들이 발생시키는 사건으로서, 객체들 간의 움직임 및 관계들을 바탕으로 추론된다. OLIVE는 객체 컨셉 추출 모듈을 통하여 추출한 각 Shot

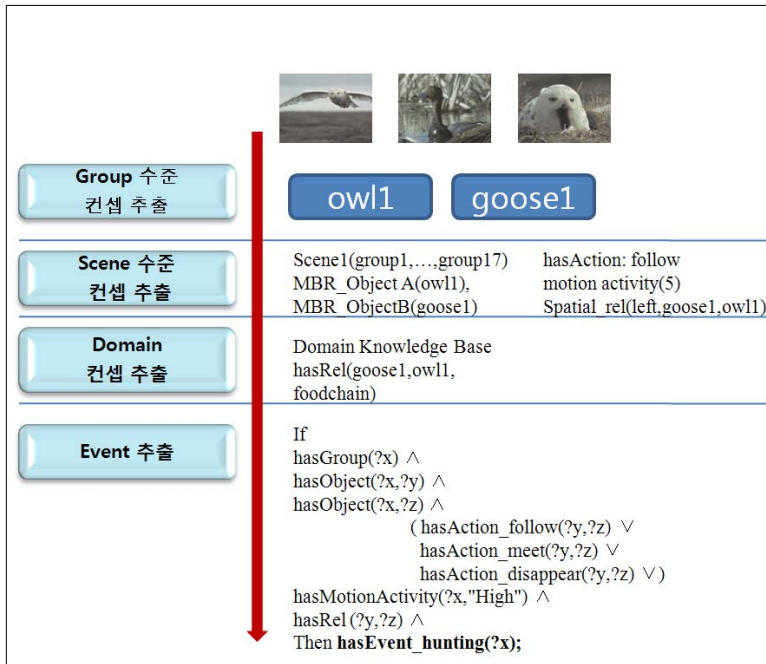


그림 2 이벤트 추출 과정

간의 거리의 변화 속도 및 방향을 알 수 있으며 시간적인 정보도 얻을 수 있게 된다. 예를 들어 보면, 그림 2에서 owl과 goose라는 객체의 MBR 영역의 시공간적 움직임을 바탕으로 follow, meet, disappear와 같은 액션을 추출한다. 시간적 관계의 경우 객체의 움직임을 바탕으로 시간적인 정보를 고려하여 객체의 모션과 동작 패턴 정보를 얻을 수 있다. 또한 실제 세계에서 두 객체들의 관계는 도메인 온톨로지인 동물 온톨로지로부터 소비자, 생산자의 관계에 있는 먹이사슬(Food chain)관계를 얻는다.

위 과정을 통해 객체의 시각 중간 컨셉을 바탕으로 Scene 온톨로지에 모션, 시간, 공간적 정보가 기술되며, 객체의 도메인 정보(Domain knowledge base)를 바탕으로 Domain 온톨로지에 객체의 의미적 관계가 기술된다. 이에 이벤트 온톨로지는 추론엔진을 사용하여 각 Scene에서 최종적으로 얻을 수 있는 이벤트를 추출해 낸다. 그리고 이벤트 온톨로지는 이벤트와 이벤트 사이에서 상위 이벤트 컨셉을 추론할 수 있는 정보를 기술 한다. 본 연구에서는 Scene 온톨로지로부터 얻은 객체간 정보를 바탕으로 이벤트 추론을 위한 규칙 정의 언어로 부침개 언어[2]를 사용한다. 그림 2에서 Event 추출 과정은 의미 추론 규칙으로 객체의 시간 및 공간적 관계 속성을 이용하여 이벤트 컨셉을 추론한 예이다. 이렇게 추출된 이벤트를 바탕으로 자동주석 시스템인 OLIVE는 최종적으로 RDF 형태의 주석을 생성하게 된다.

3. 결론

비디오의 이벤트는 객체 중심으로 일어나는 비디오를 묘사할 수 있는 가장 핵심적인 요소로 비디오에서 가장 중요한 부분이라고 할 수 있다. 이에 OLIVE 시스템은 결과적으로 행동의 주체인 객체에 대한 주석뿐만 아니라 객체 중심의 이벤트에 대한 주석을 생성하기 때문에 기존에 제안된 주석 시스템 보다 더욱 자세하고 정확한 의미들로 Scene 에 대한 주석을 생성한다. 본 논문에서는 OLIVE시스템을 구축하여 객체 중심의 관계 추론을 바탕으로 일어날 수 있는 사건에 대한 이벤트 검색이 가능 하도록 이벤트 온톨로지를 구축하였다. Mpeg-7 기술자와 객체들의 움직임을 담은 Scene 온톨로지를 통하여 구축된 이벤트 온톨로지는 도메인에 상관없이 다른 도메인 온톨로지와 함께 사용 될 수 있는 장점을 가진다.

[1] Jin-Woo Jeong, Kyung-Wook Park, Jeong Ho Lee, Young Shik Moon, Sung Han Park, and Dong-Ho Lee, "OLYVIA : Ontology-based Automatic Video Annotation & Summarization System using Semantic Inference Rules", IEEE International Conference on Semantics, 2007

[2] Minsu Jang, Joo-Chan Sohn, Bossam:An extended rule engine for OWL inferencing, RuleML 2004, LNCS 3323, pp.128-138,2004

1) 본 연구는 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2006-521-D00457)