ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Automatic Semantic Content Extraction from Video Using Ontology and Rule Based Model

Vijaykumar Modi
II-Year M.Tech (CNE)
Department of CSE
Dayananda Sagar College of Engg.
Bangalore-78

Prof. Rashmi.S.R
Assistant Professor
Department of CSE
Dayananda Sagar College of Engg.
Bangalore-78

Abstract

In the Modern era use of video-based applications has made the need for extracting the useful content in videos. Raw data and low-level features alone are not sufficient to fulfill the user's needs that is, a deeper understanding of the content at the semantic level is required. Currently available manual techniques, which are inefficient, subjective and costly in time and limit the querying capabilities. Semi automatic methods are used to bridge the gap between low-level representative features and high-level semantic content. Here, we propose a semantic content extraction system that allows the user to query and retrieve objects, events, and concepts that are extracted automatically. The proposed framework has been fully implemented and tested on three different domains. We have obtained satisfactory precision and recall rates for object, event and concept extraction.

Keywords—Ontology, Semantic Content extraction, Rule-Based.

INTRODUCTION

The rapid increment in the accessible measure of video information has made a dire need create intelligent techniques to model and concentrate the video content. Typical applications in which modeling and extracting video content are crucial include surveillance, video-on-demand systems, intrusion detection, border monitoring, sport events, criminal investigation systems, and many others. The ultimate goal is to enable users to retrieve some desired content from massive amounts of video data in an efficient and semantically meaningful manner.

There are basically three levels of video content which are raw video data, low-level features and semantic content. First, raw video data consist of elementary physical video units together with some general video attributes such as format, length, and frame rate. Second, low-level features are characterized by audio, text, and visual features such as texture, color distribution, shape, motion, etc. Third, semantic content contains high-level concepts such as objects and events. The first two levels on which content modeling and extraction approaches are based use automatically extracted data, which represent the low-level content of a video, but they hardly provide semantics which is much more appropriate for users. Users are for the most part intrigued by querying and recovering the video regarding what the video contains. In this way, raw video information and low-level components alone are not adequate to satisfy the users need that is, a more deeper knowledge of the data at the semantic level is required in numerous video-based applications.

However, it is very difficult to extract semantic content directly from raw video data. This is because video is a temporal sequence of frames without a direct relation to its semantic content [2]. Therefore, many different representations using different sets of data such as audio, visual features, objects, events, time, motion, and spatial relations are partially or fully used to model and extract the semantic content. No matter which type of data set is used, the process of extracting semantic content is complex and requires domain knowledge or user interaction.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

OVERVIEW OF MODEL

The major parts of the model are discussed below.

A. Ontology-Based Modeling

The VISCOM contains classes and relations between these classes. Some of the classes represent semantic content types such as Object and Event while others are used in the automatic semantic content extraction process [1]. Relations defined in VISCOM give ability to model events and concepts related with other objects and events. VISCOM is developed on an ontology-based structure where semantic content types and relations between these types are collected under VISCOM Classes, VISCOM Data Properties which associate classes with constants and VISCOM Object Properties which are used to define relations between classes. In addition, there are some domain independent class individuals.

B. Rule-Based Modeling

Additional rules are utilized to extend the modeling capabilities. Each rule has two parts as body and head where body part contains any number of domain class or property individuals and head part contains only one individual.

C. Domain Ontology Construction

A Algorithm 1 presents the steps followed to construct a domain ontology . For the evaluation purposes, we have constructed an Office Surveillance Ontology, a Basketball Ontology and a Football Ontology by using Protégé. A small portion of the basketball ontology is illustrated in Fig. for Rebound event, as an example.

Algorithm 1. Ontology Construction with VISCOM

- define O, E and C individuals.
- define all possible SR’s occuring within an E.
- define all possible OM’s occuring within an E.
- use SR’s and M’s to define SC’s.
- describe temporal relations between SC’s as TSCC’s.
- make EDs with SC’s, SR’s and TSCC’s.
- for all E’s do
- if an event can be defined with an event def then
- define E in terms of ED’s.
- end if
- if an event can be defined with temporal relations between other events then
- define E’s in terms of ETR’s.
- end if
- end for
- for all C’s do
- construct a relation with the C that can be placed in its meaning.
- end for
- define S’s.

In accordance with the Algorithm 1, ontology construction starts with defining Rebound as the Event individual, and Hoop, Ball, Player and Basket as the Object individuals. Next step is to define all Spatial Relation Component individuals that happen during a Rebound event such as Ball Above Player, Player Below Basket and Ball Far from Hoop. Then, the sequence of the Spatial Relation Component individuals are defined as Spatial Change individuals such as Jump to Ball and Hit Hoop. One or more Spatial Change individuals may be used to create different Event Definition individuals. In the Rebound example, two Event Definition individuals are defined. Rebound Definition 1 has two Spatial Change individuals (Hit Hoop and Jump to Ball) in its definition which have a temporal relation with each other, while Rebound Definition 2 has only one Spatial Change individual (Jump for Rebound). Each event definition uses different spatial and temporal relations between objects in order to define the event. The ontology developer always has a chance to add a new definition that will cover cases where existing definitions are not sufficient enough. Also he/she has an opportunity to add new individual definitions, modify, or delete them at any time.

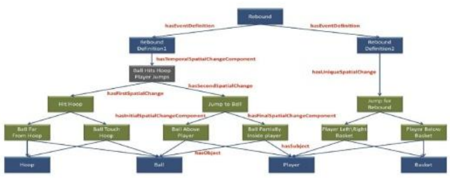


Fig.1 Reboud Event Representation

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

PROPOSED FRAMEWORK

In this section, we present our framework for Automatic Semantic Content Extraction from Video. Fig 2 gives an overview of the proposed Semantic content framework. Given an input video, we first covert video into the frame. Then, frames are used to detection of objects .Based on the local feature and spatio temporal relations event is detected in accordance with the ontology and rules.

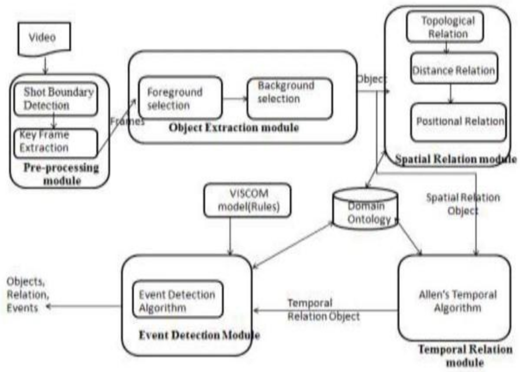


Figure 2: Overview of proposed System

Pre-Processing Module:

In Pre-Processing Module Video is fed as input to the system is converted into the frames for future processing.

Object Extraction Module:

Most Object extraction is one of most crucial components in the framework, since the objects are used as the input for the extraction process. object extraction techniques use training data to learn object definitions, which are usually shape, color, and texture features. These definitions are mostly the same across different domains

Spatial Relation Extraction Module:

Every spatial relation extraction is stored as a Spatial Relation Component instance which contains the frame number, object instances, type of the spatial relation,

Event Extraction Module:

Event instances are extracted after a sequence of automatic extraction processes. Each extraction process outputs instances of a semantic content type defined as an individual in the domain ontology. Algorithm 2 describes the whole event extraction process. In addition, relations between the extraction processes are illustrated in Fig. 3.

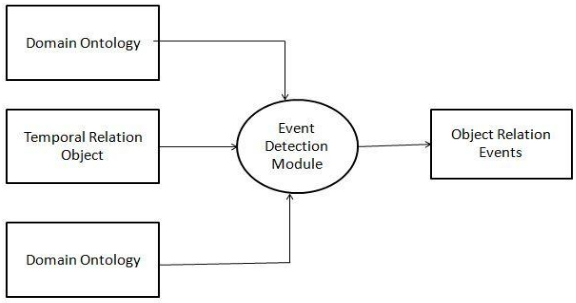


Fig.3 Event Extraction Process

Algorithm 2:

- for all SRC individuals in the ontology do
- extract SRC instances that satisfy the individual def.
- execute SR rule def
- end for
- for all SMC individuals in the ontology do
- extract SMC instances that satisfy the individual def.
- end for
- for all SC individuals in the ontology do
- check if there are SRC or SMC instances that satisfy the individual def.
- end for
- for all TSC individuals in the ontology do
- extract SC instances that satisfy the individual def.
- end for
- for all ED individuals in the ontology do
- check if there are SC, SR or TSC instances that satisfy the individual def.
- end for
- for all E individuals in the ontology do
- check if there are ED instances that satisfy the individual def.
- end for
- for all Event individuals which have Temporal Event Component individuals do

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

- extract Event instances that satisfy the individual def.
- end for
- for all S individuals in the ontology do
- extract E instances that satisfy the individual def.
- end for
- execute all rules defined for E individuals to extract Additional Events.

Concept Extraction Module:

In concept extraction process, Concept Component individuals and extracted object, event, and concept instances are used. Concept Component individuals relate objects, events, and concepts with concepts. When an object or event that is used in the definition of a concept is extracted, the related concept instance is automatically extracted with the relevance degree given in its definition. In addition, Similarity individuals are utilized in order to extract more concepts from the extracted components. The last step in the concept extraction process is executing concept rule definitions. k-nearest neighbor's algorithm (k-NN) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space.

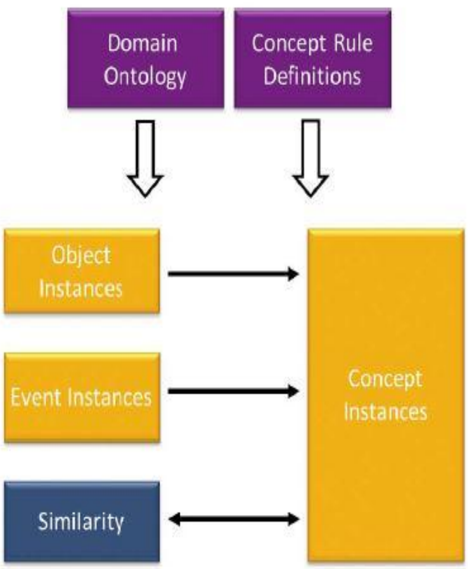


Fig 5. Concept Extraction Process

Algorithm

- for all CC individuals in the ontology do
- check is there are O or E instances that satisfy the individual def.
- end for
- for all S individuals in the ontology do
- extract C instances that satisfy the individual def
- end for
- execute all rules defined for C individuals

Example on Ontology Construction:

Office Activity

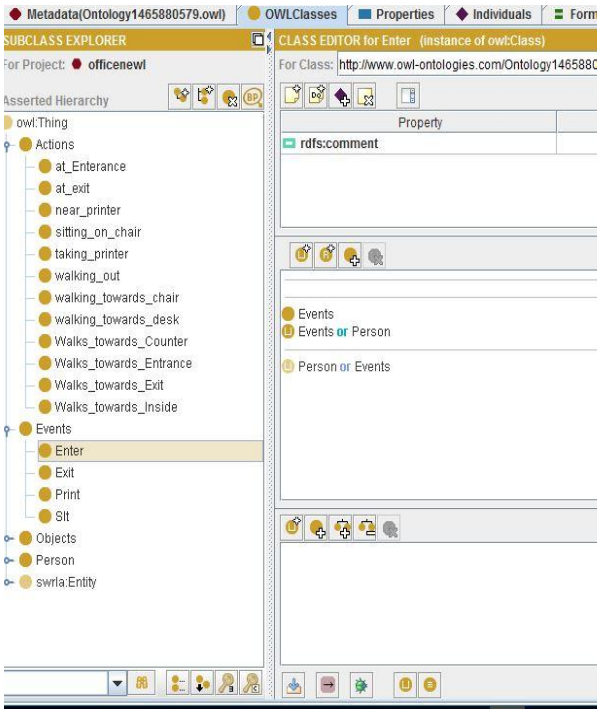
Step 1: Collect the Domain Information needed for the construction of the ontology.

Step2: Collect the activity's, and find the relationship between the Activities and Objects.

Activities

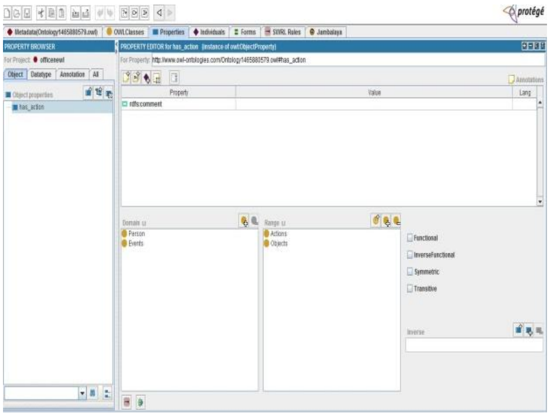
- Enter
- Exit
- Print
- Sit

Step 3: Create the OWL classes Defining Actions , Events , Objects, Person

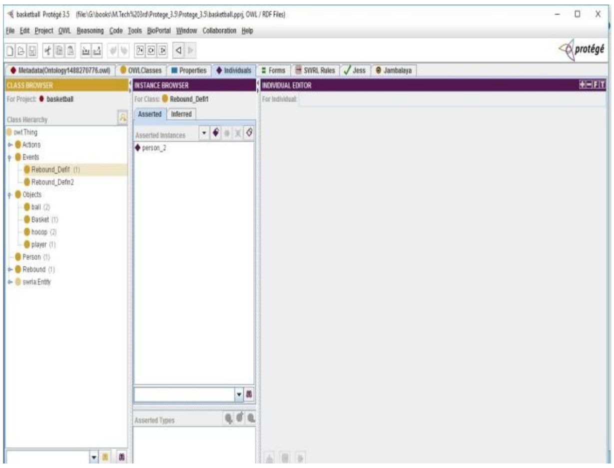


ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

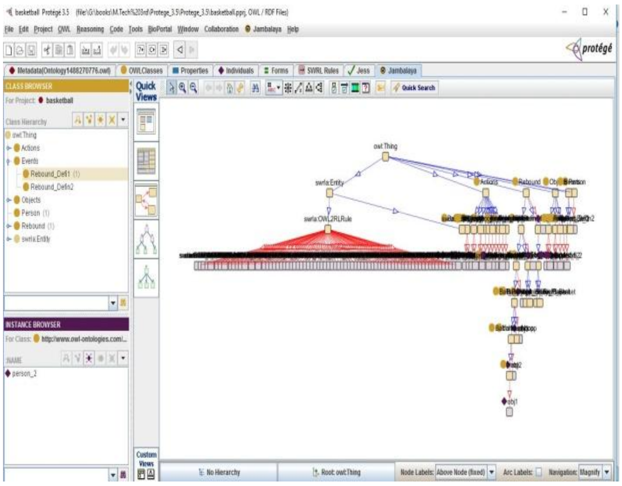
Step 4 : Define the class Property ,object , and Domain Range



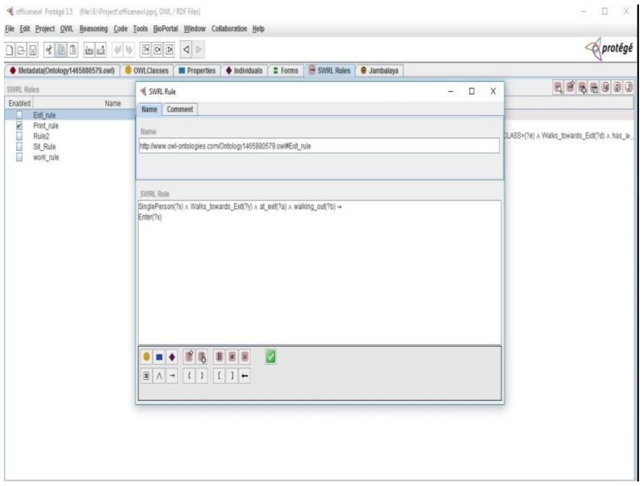
Step 5: Create Instances for each classes in individual tab office individual



Step 6 : Open Jambalaya Tab to view the ontology construction .

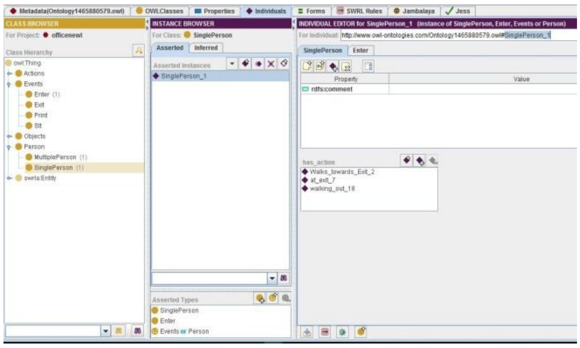


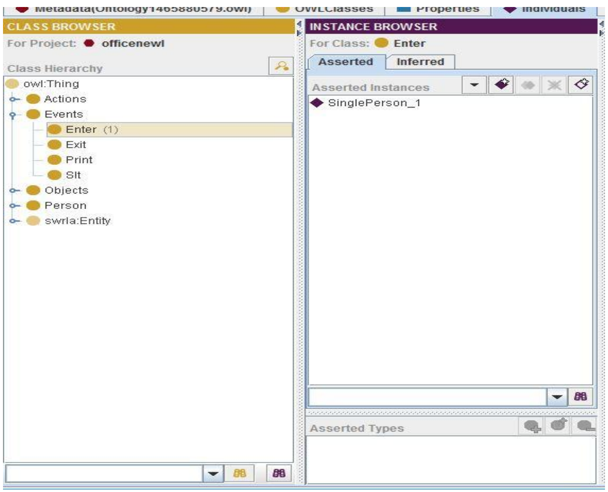
Step7 : Open SWRL tab to write the rules for actions



Step 8 : To see the working

- Create instances of SinglePerson that has properties mentioned in rules
- run the SWRLJess tab
- New instances are going to be added in the events classes





ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Experimental Results:

Precision, Recall, BDA values are taken are taken for performance analysis. Precision and Recall is calculated based on the following relation.

$$Prec_{int} = \frac{\tau_{mb} \cap \tau_{db}}{\tau_{db}}$$

$$Rec_{int} = \frac{\tau_{mb} \cap \tau_{db}}{\tau_{mb}}$$

$$BDA = \frac{\tau_{mb} \cap \tau_{db}}{max(\tau_{mb}, \tau_{db})},$$

where ţdb and ţmb are the automatically detected event

Name	Precision	Recall	BDA
Rebound	50	50	62.50
Jump Ball	100	100	97.23
Free Throw	100	100	95.00
Attack	100	100	94.65

Figure 4 .Precision and Recall Value
Figure 1: precision and recall values for different actions
Figure 4 shows the precision, recall, BDA values for basketball domain videos.

CONCLUSION

The primary aim of this project is to develop a framework for an automatic semantic content extraction system for videos which can be utilized in various areas, such as surveillance, sport events, and news video applications. The novel idea here is to utilize domain ontologies generated with a domain-independent ontology-based semantic content metaontology model and a set of special rule definitions.

REFERENCES

[1] L Yakup Yildirim, Adnan Yazici, Turgay Yilmaz, "Automatic Semantic Content Extraction in Videos Using a Fuzzy Ontology and Rule-Based Model". IEEE transactions on knowledge and data engineering, vol. 25, no. 1, January 2013

[2] M. Petkovic and W. Jonker, "An Overview of Data Models and Query Languages for Content-Based Video Retrieval," Proc. Int'l Conf. Advances in Infrastructure for E-Business, Science, and Education on the Internet, Aug. 2000.

[3] M. Petkovic and W. Jonker, "Content-Based Video Retrieval by Integrating Spatio-Temporal and Stochastic Recognition of Events," Proc. IEEE Int'l Workshop Detection and Recognition of Events in Video, pp. 75-82, 2001

[4] G.G. Medioni, I. Cohen, F. Bre´mond, S. Hongeng, and R. Nevatia, "Event Detection and Analysis from Video Streams," IEEE Trans. Pattern Analysis Machine Intelligence, vol. 23, no. 8, pp. 873-889, Aug. 2001.

[5] S. Hongeng, R. Nevatia, and F. Bre´mond, "Video-Based Event Recognition: Activity Representation and Probabilistic Recognition Methods," Computer Vision and Image Understanding, vol. 96, no. 2, pp. 129-162, 2004.

[6] A. Hakeem and M. Shah, "Multiple Agent Event Detection and Representation in Videos," Proc. 20th Nat'l Conf. Artificial Intelligence (AAAI), pp. 89-94, 2005.

[7] M.E. Do¨nderler, E. Saykol, U. Arslan, O ¨ . Ulusoy, and U. Gu¨du¨ kbay, "Bilvideo: Design and Implementation of a Video Database Management System," Multimedia Tools Applications, vol. 27, no. 1, pp. 79-104, 2005.

[8] T. Sevilmis, M. Bastan, U. Gu¨du¨ kbay, and O ¨ . Ulusoy, "Automatic Detection of Salient Objects and Spatial Relations in Videos for a Video Database System," Image Vision Computing, vol. 26, no. 10, pp. 1384-1396, 2008.

[9] M. Ko¨pru¨ lu¨, N.K. Cicekli, and A. Yazici, "Spatio-Temporal Querying in Video Databases," Information Sciences, vol. 160, nos. 1-4, pp. 131-152, 2004.

[10] J. Fan, W. Aref, A. Elmagarmid, M. Hacid, M. Marzouk, and X. Zhu, "Multiview: Multilevel Video

Content Representation and Retrieval," J. Electronic Imaging, vol. 10, no. 4, pp. 895-908, 2001.

[11] J.F. Allen, "Maintaining Knowledge about Temporal Intervals," Comm. ACM, vol. 26, no. 11, pp. 832-843, 1983.

[12] M.J. Egenhofer and J.R. Herring, "A Mathematical Framework for the Definition of Topological Relationships," Proc. Fourth Int'l Symp. Spatial Data Handling, pp. 803-813, 1990.

[13] M. Vazirgiannis, "Uncertainty Handling in Spatial Relationships," SAC '00: Proc. ACM Symp. Applied Computing, pp. 494- 500, 2000.

[14] P.-W. Huang and C.-H. Lee, "Image Database Design Based on 9D-SPA Representation for Spatial Relations," IEEE Trans. Knowledge and Data Eng., vol. 16, no. 12, pp. 1486-1496, Dec. 2004.

[15] I. Horrocks, P.F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean, "Swrl: A Semantic Web Rule Language," technical report, W3C, http://www.w3.org/Submission/SWRL/, 2004.

[16]"Prote´ge´ Ontology Editor," http://protege.stanford.edu/, 2012.

[17] "Jena: A Semantic Web Framework," http://www.hpl.hp.com/semweb/, 2012.

[18] C. Xu, J. Wang, K. Wan, Y. Li, and L. Duan, "Live Sports Event Detection Based on Broadcast Video and Web-Casting Text," MULTIMEDIA '06: Proc. 14th Ann. ACM Int'l Conf. Multimedia, pp. 221-230, 2006.

[19] Y. Zhang, C. Xu, Y. Rui, J. Wang, and H. Lu, "Semantic Event Extraction from Basketball Games Using Multi-Modal Analysis," Proc. IEEE Int'l Conf. Multimedia and Expo (ICME '07), pp. 2190- 2193, 2007.

[20] L.S. Davis, S. Fejes, D. Harwood, Y. Yacoob, I. Haratoglu, and M.J. Black, "Visual Surveillance of Human Activity," Proc. Third Asian Conf. Computer Vision (ACCV), vol. 2, pp. 267-274, 1998.