Automatic Semantic Rule Based Coordination for Content Extraction in Videos Using Fuzzy Ontology

P.Pavithra, Dr.N.Uma Maheswari

II-ME, Department Of CSE, PSNACET.Dindigul, Tamil Nadu, India.

Professor, PSNACET.Dindigul, Tamil Nadu, India

Abstract — A Tremendous increase in the use of video-based applications has revealed the need for extracting the content in videos. Video semantic event detection is essential for its video summarization and retrieval. Input video data and low-level features alone are not sufficient to fulfill being used to bridge the gap between low-level representative features and high-level semantic content. Here, an Intelligent Video semantic content extraction system is proposed that allows the user to query and retrieve cluster objects, events, and concepts that are extracted automatically and provides a textual representation for the extracted frames. Domain ontology based fuzzy video semantic content model that uses spatial/temporal relations is used in event and concept definitions. This meta-ontology definition provides a wide-domain applicable rule construction standard that allows the user to construct ontology for wide domain. The proposed framework provides crisp video output eliminating unnecessary the user's needs that is, a deeper understanding of the content at the semantic level is required. The present manual techniques, which are inefficient, subjective and costly in time and limit the querying capabilities, are contents and explanation of video frames for physically challenged users which reduces the time consumption for semantic extraction in lengthy videos.

Keywords- Intelligent Video semantic content model, Frameset representation, Cluster objects

I. INTRODUCTION

The rapid increase in the available amount of video data has caused an urgent need to develop intelligent methods to model and extract 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 meaningful manner. Users are mostly eager in querying and retrieving the video in terms of the video that contains. Therefore, raw video data and low-level features alone are not sufficient to fulfill the user requirements that is, a deeper understanding of the information at the semantic level is required. Hence 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.

Manual extraction approaches are tedious, subjective, and time consuming, which limit querying capabilities. The most of these studies propose techniques for specific event type extraction or work for specific cases and assumptions.

Many researchers have studied this from different aspects. A simple representation could relate the events with their low-level features (shape, color, etc.) using shots from videos, without any spatial or temporal relations. However, and effective use of spatiotemporal relations is crucial to achieve reliable recognition of

events. Employing domain ontologies facilitate use of applicable relations on a domain.

There are no studies using both spatial relations between objects, and temporal relations between events together in an ontology-based model to support automatic semantic content extraction.

A Video Event Recognition Language (VERL) that allows users to define the events without interacting with the low level processing is defined. VERL is intended to be a language for representing events for the purpose of designing ontology of the domain, and, Video Event Markup Language (VEML) is used to manually annotate VERL events in videos. This is accomplished through the development of an ontology-based semantic content model and semantic content extraction algorithms.

Our work differs from other semantic content extraction and representation studies in many ways and contributes to semantic video modeling and semantic content extraction research areas. First of all, an I-VSCM meta-ontology, a rule construction standard is proposed which is domain independent, to construct domain ontologies. Domain specific rule construction standard is proposed and to analyze the group of cluster objects a training data model is proposed.

II. INTELLIGENT VIDEO SEMANTIC CONTENT MODEL

The proposed semantic video content model and the use of special rules (without using ontology) are described.

2.1 Overview of the Model

Ontology provides many advantages and capabilities for content modeling. Yet, a great majority of the ontology based video content modeling studies propose domain specific ontology models limiting its use to a specific domain. Besides, generic ontology models provide solutions for multimedia structure representations. A wide-domain applicable video content model is proposed in order to design the semantic content in videos. I-VSCM is a well-defined meta-ontology for constructing domain ontologies. It is an alternative to the rule-based and domain-dependent extraction methods. Constructing rules for extraction is a tedious task and is not scalable. It eases the rule construction process and makes its use on larger video data possible. The Class logic given as:

$$I\text{-}VSCM = \begin{cases} \begin{array}{l} \text{Class} \rightarrow \{Ci\} \\[6pt] \text{Temporalrelation} \rightarrow \{\text{before, after,} \\ \text{Starts, finishes}\} \\ \text{Spatialrelation} \rightarrow \{\text{far, near,} \\ \text{Disjoint, above, below, left, right}\} \\[6pt] \text{Where (Ci, I-VSCM Class)} \end{array} \end{cases}$$

(1)

The rules that can be constructed via I-VSCM ontology can cover most of the event definitions for a wide variety of domains. However, there can be some exceptional situations that the ontology definitions cannot cover. To handle such cases, I-VSCM provides an additional domain specific rule construction standard. Hence, I-VSCM provides a solution that is applicable on a wide variety of domain videos. Objects, events, concepts, spatial and temporal relations are components of this generic ontology-based model. Similar generic models such which use objects and spatial and temporal relations for semantic content modeling neither use ontology in content representation nor support automatic content extraction.

The starting point is identifying what video contains and which components can be used to model the video content. Key frames are the elementary video units which are still images, extracted from original video data that best represent the content of shots in an abstract manner. Name, domain, frame rate, length, format are examples of general video attributes which form the metadata of video.

Object instances represent the relevance of the given Object Store Bound (OSB) to the object type. Spatial relation calculation spectrum fuzzy results and Spatial Relation Component instances are extracted with fuzzy membership values.

2.2 Ontology-Based Modeling

The linguistic part of I-VSCM 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.. I-VSCM is developed on an ontology-based structure where semantic content types and relations between these types are collected under I-VSCM Classes, I-VSCM Data Properties which associate classes with constants and I-VSCM Object Properties which are used to define relations between classes.

C-Logic is used for the formal representation of I-VSCM classes and operations of the semantic content extraction framework. C-Logic includes a representation framework for entities, their attributes, and classes using identities, labels, and types.

Below, the I-VSCM classes are introduced with their description, formal representation and important relation (property) descriptions.

2.2.1. Component

I-VSCM collects all of the semantic content under the class of Component. A component can have synonym names and similarity relations with other components. Component class has three categories.

They are: objects, concept and events.

2.2.2 Object

Objects correspond to existential entities. An object is the starting point of the composition. An object has a name,

low-level features, and composed-of relations. Student, instructor, board are examples of objects in learning video.

$$Object: \begin{cases} \text{[Name} \rightarrow \text{[string], lowlevelfeature} \rightarrow \text{\{Li\},} \\ \quad \text{Composed Of} \rightarrow \text{\{relation j\}]} \\ \qquad \text{Where (Li, Lowlevel feature)} \end{cases} \tag{2}$$

2.2.3 Event

Events are long-term temporal objects and object relation changes. They are described by using objects and spatial/temporal relations between objects. Relations between events and objects and/or their attributes indicate how events are inferred from objects and/or object attributes. Read the book, write in board are examples of events.

$$Event: \begin{cases} \text{[Name} \rightarrow \text{[string], eventdef} \rightarrow \text{\{EDi\},} \\ \quad \text{ObjectRole} \rightarrow \text{\{ORi\},} \\ \quad \text{Temporalcomponent} \rightarrow \text{\{TECj\}]} \\ \quad \text{Where (EDi, Event Definition)} \\ \quad \text{Where (ORi, Object Role)} \end{cases} \tag{3}$$

2.2.4 Concept

Concepts are general definitions that contain related events and objects in it. Each concept has a relation with its components that are used for its definition. Individual chapters are examples of concepts for the learning video.

$$Concept: \begin{cases} \text{[Name} \rightarrow \text{[string], concept component} \\ \qquad \rightarrow \text{\{CCi\}]} \\ \qquad \text{where} \quad \text{(CCi,} \quad \text{Concept} \\ \text{component)} \end{cases} \tag{4}$$

2.2.5 Spatial Relation

Spatial relations express the relative object positions between two objects such as above, inside, or far. The spatial relation types are grouped under three categories as topological, distance and positional spatial relations. The individuals of this class are utilized by the individuals of Spatial Relation Component class.

2.2.6 Temporal Relation

Temporal relations are used to order Spatial Changes or Events in Event Definitions. Allen's temporal relationships are used to express parallelism and mutual exclusion between components

2.2.7 Event Definition

An event can have several definitions where each definition describes the event with a certainty degree. Each event definition has a membership value for the event it defines that denotes the clarity of description.

2.2.8 Similarity

Similarity class is used to represent the relevance of a component to another component in a fuzzy manner. Whenever a component which has a similarity relation with another component is extracted, the semantically related component is automatically extracted by using similarity relation.

2.3 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 with a value μ, representing the certainty of the definition given in the body part to represent the definition in the head part where $0 < \mu < 1$.
The basic syntax of rules has parentheses and logical connectives in both body and head parts. Rule definitions are used for two different purposes. The first purpose is to lower the spatial relation computation cost.
Consequently, it can be stated that rule definitions strengthens the framework in terms of both semantic content representation and semantic content extraction.

2.4 Domain Ontology Construction with I-VSCM meta-ontology

I-VSCM is utilized as a Meta model to construct domain ontologies. Basically, domain specific semantic contents are defined as individuals of I-VSCM classes and properties

Algorithm 1: Domain Ontology Construction with I-VSCM meta-ontology
Require: I-VSCM
Ensure: Domain Ontology
1: define Objects, Events and Concept individuals.
2: define all possible spatial relations occurring within an Event.
3: define all possible Object models occurring within an Event.
4: use Spatial Relations and Motions to define Semantic Contents.
5: describe temporal relations between spatial components.
6: make Event definitions with Spatial Components and Temporal Spatial Concept Classes.
7: for all Events do
8: if an event can be defined with an event   definition then
9: define E in terms of Event Definitions.
10: end if
11: if an event can be defined with temporal relations between other events then

12: define Events in terms of Event Temporal Relations.
13: end if
14: end for
15: for all Components do
16: construct a relation with the Classes that can be placed in its meaning.
17: end for
18: define Semantics.

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 over 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. All of the mantic content is defined as I-VSCM class individuals in a similar manner.

Ontology in the sports domain is constructed using Protégé for demonstrating the validity of the proposed framework. A reasoning algorithm based on temporal logic is proposed for event detection in sports videos. The proposed framework supports flexible and managed execution of various application and domain independent video low-level analysis tasks.

The choice of algorithm employed for the detection of sequences and objects is directly dependent on its available characteristic features which directly depend on the domain that the sequences and objects involve. So this association should be considered based on video analysis knowledge and domain knowledge, and is useful for automatic and precise detection.

III. AUTOMATIC SEMANTIC CONTENT EXTRACTION FRAMEWORK

The Automatic Semantic Content Extraction Framework is illustrated in Fig. 1. The ultimate goal is to extract all of the semantic content existing in video instances. In order to achieve this goal, the automatic semantic content extraction framework takes Vi, ONTi, and Ri, where Vi is a video instance, ONTi is the domain ontology for domain Di which Vi belongs to, and Ri is the set of rules for domain Di.

$$
\text{Intelligent Video Content Model}
\begin{cases}
\text{Video} \rightarrow [\text{Vi}], \\
\text{Object} \rightarrow [\text{Oi}], \\
\text{Events} \rightarrow [\text{Ei}], \\
\text{Concepts} \rightarrow [\text{Ci}] \\
\text{Input} \rightarrow \{\text{Vi, Video Ontology, Rulei}\} \\
\text{Output} \rightarrow \{\text{Video Semantic content and text explanation}\}
\end{cases}
\tag{5}
$$

The output of the extraction process is a set of semantic contents, named VSCi and the extracted frameset explanation.

In order to meet the object extraction and classification need, a cluster based automatic activity detection training dataset is provided.
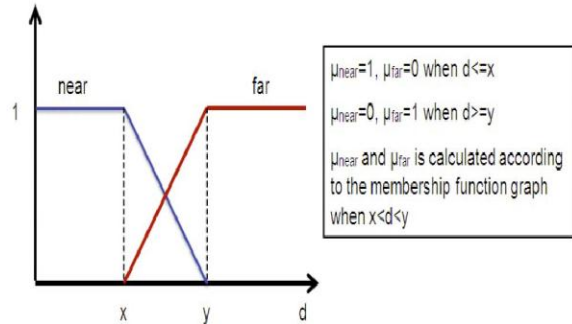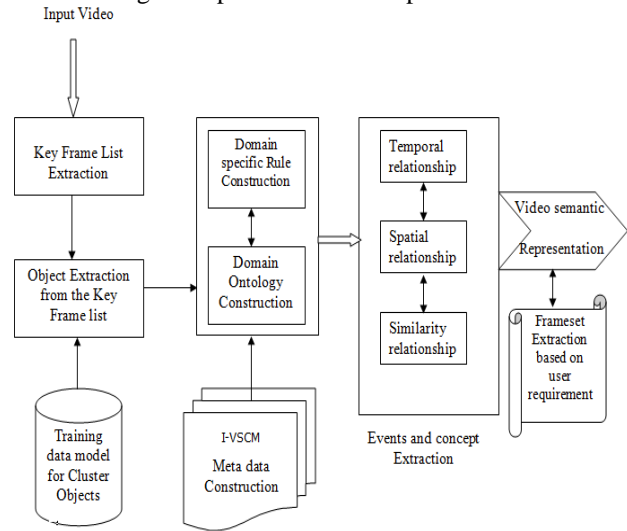

Fig 1.Graph for membership function

Fig.2. Automatic Semantic Content Extraction and its Textual representation

3.1 Object Extraction

Object extraction is one of most crucial components in the framework, since the objects play an important role in detecting the training data model.
The resultant objects are stored in the database. Objects are used as the input for the extraction process. However, the details of object extraction process are not presented in detail, considering that the object extraction process is mostly in the scope of computer vision and image analysis techniques.

During the object extraction process, for each representative key frame in the video, above-mentioned object extraction process is performed and a set of objects is extracted and classified.

The extracted objects are stored with frame number, membership function in Fig.1. and ObjectStoreBound

$$
\text{Membership function} =
\begin{cases}
[\mu => [\text{double}]] \\
\text{Where } 0 < \mu < 1
\end{cases}
\tag{6}
$$

$$
\text{ObjectStoreBound} =
\begin{cases}
[x => [\text{integer}], y => [\text{integer}], \\
\text{Width} => [\text{integer}], \\
\text{Length} => [\text{integer}],
\end{cases}
\tag{7}
$$

3.2 Spatial Relation Extraction

Object instances are represented with the MBR. There can be n object instance (as regions) represented with R in a frame F= {R0, R1….Rn}.For each R, the upper left-hand corner point represented with length and width of R are stored.

The area inside Ri is represented with Ri1 where the edges of Ri are represented with Ri2.

The topological relations are given by

$$\mu_{topology\,(Ri,Rk)} = \frac{(Ri^{\alpha} \cap Rk^{\alpha})}{Rk^{\alpha}}$$

$$\mu_{rightpos\,(Ri,Rk)} = \begin{cases} \sin(\phi+90)\,, & 0<\phi<90 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{leftpos\,(Ri,Rk)} = \begin{cases} \sin(\phi-90)\,, & 90<\phi<270 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{abovepos\,(Ri,Rk)} = \begin{cases} \sin(\phi)\,, & 0<\phi<180 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{belowpos\,(Ri,Rk)} = \begin{cases} \sin(\phi-180)\,, & 180<\phi<360 \\ 0 & \text{otherwise} \end{cases}$$

(8)

3.3 Temporal Relation Extraction

In the framework, temporal relations are utilized in order to add temporality to sequence spatial change or Events individuals in the definition of Event individuals.

TABLE I: Temporal Relation tabulation

Name	Notation	Definition
Before	B(x,y)	$x^+ < y^+$
After	A(x,y)	B(y,x)
Meets	M(x,y)	$x^+ = y^-$
Starts	S(x,y)	$x^- = y^- < x^+ = y^+$
Finishes	F(x,y)	$x^- = y^- > x^+ = y^+$

3.4 Event and Concept Extraction

Event instances are extracted after a sequence of automatic extraction processes. In the 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.

Algorithm2.Event and Concept Extraction Algorithm

Require: Domain Ontology, Object Instances, Event Instances
Ensure: Event Instances, Concept Instances
1: for all Class Component individuals in the ontology do
2: check is there are Objects or Events instances that satisfy the individual def.
3: end for
4: for all S individuals in the ontology do
5: extract C instances that satisfy the individual def.
6: end for
7: execute all rules defined for C individuals.

IV. VIDEO SEMANTIC REPRESENTATION BY FRAMESET EXTRACTION

In this framework a video semantic content is represented by the frameset extraction. This extraction is done based on the user's requirements. To cover the requirements of physically challenged users, the explanation of the corresponding frameset is provided automatically.. This also removes the unwanted content from the lengthy videos and extracts the required semantic information.

V. EXPERIMENTAL RESULTS

In this study, Open computer vision tool is used and I-VSCM is used to make rule definitions In this extension, the object individuals include a specification of the degree (a truth value between 0 and 1) of confidence with which an individual is an instance of a given class or property.

Objects are manually provided for the second step of the test. Obviously, this result is expected since the framework uses the objects in the key frames as the input and extracts events and concepts by using the objects and I-VSCM rules. When a missing or wrong classification of object instances occurs in the automatic object extraction process, then the success of event/concept extraction decreases. The key frames have been extracted and tested with multiple domains. The results have been very accurate compared to the previous techniques.

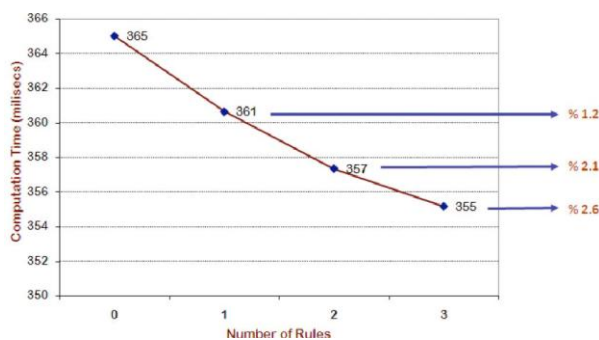The following graph has been simulated by the training data of the I-VSCM technique.

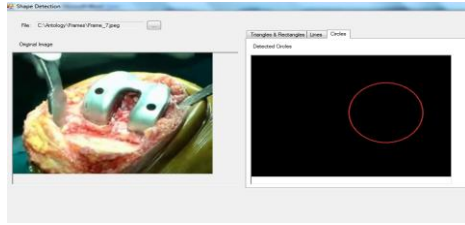Fig 3.Graph for the Automatic Events and Concept rules

Fig 4.Training data for object model

Fig 3.I-VSCM ontology based textual explanation

VI. CONCLUSION

The primary aim of this research is to develop a framework for an automatic semantic content extraction system for videos which can be utilized in various areas. Automatic Semantic Content Extraction Framework contributes in several ways to semantic video modeling and semantic content extraction research areas. First of all, the semantic content extraction process is done automatically. In addition, a generic ontology-based semantic metaontology model for videos is proposed.

An automatic cluster-based object extraction method is integrated to the proposed system to capture semantic content and explanation in textual format. The framework can improve the personalized querying a retrieval capabilities of user. It provides explanation for physically challenged users and consumes less time.

As a further study, one can improve the model and the extraction capabilities of the framework for spatial relation extraction by considering the viewing angle of camera and the motions in the depth dimension.

REFERENCES

[1] "Automatic Semantic Content Extraction in Videos Using Fuzzy Ontology and Rule Based Model" Yakup Yildirim, Adnan Yazici, Senior Member, IEEE, and Turgay Yilmaz , Student Member, IEEE,2013

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

[3] Y. Yildirim, "Automatic Semantic Content Extraction in Video Using a Spatio-Temporal Ontology Model," PhD dissertation, Computer Eng. Dept., METU, Turkey, 2009

[4] T. Yilmaz, "Object Extraction from Images/Videos Using a Genetic Algorithm Based Approach," master's thesis, Computer Eng. Dept., METU, Turkey, 2008.

[5] 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

[6] Y. Yildirim, T. Yilmaz, and A. Yazici, "Ontology-Supported Object and Event Extraction with a Genetic Algorithms Approach for Object Classification," Proc. Sixth ACM Int'l Conf. Image and Video Retrieval (CIVR '07), pp. 202-209, 2007

[7]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

[8]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.

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

[10] 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.

[11] 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.

[12] 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/, 2005