Temporal Ontology for Video Data Modelling

Behrang QasemiZadeh[a 1], Mohammadreza Kangavari[a], and S. Hadi Rastgou Haghi [a]
[a] *Iran University of Science and Technology, Computer Faculty, Narmak, Tehran, IRAN*

Semantic video retrieval has emerged in the last decade as one of the most important features of pervasive multimedia systems. Although the initial excitement led to a hype-boosted research effort dedicated to achieve this functionality, the outcome was limited to few specialized systems based on the query-by-example retrieval model. This paper introduces a novel framework for temporal semantic video modeling. The proposed system uses an ontological infrastructure, which is called temporal ontology to deliver an abstract model of a video, which is augmented with temporal information. Temporal ontology then can be used to answer temporal queries such "Events BEFORE Event A" etc.

Keywords: Ontology, Time Ontology, Temporal Ontology, Video Modelling, Annotation

1 INTRODUCTION

Driven by development of high-capacity storage devices and the ubiquity of digital media in global networked environments, image and video retrieval has emerged in the last decade as one of the most important features of pervasive multimedia systems. Although the initial excitement led to a hype-boosted research effort dedicated to achieve this functionality, the outcome was limited to few specialized systems based on the query-by-example retrieval model. It is only recently that the research community has focused on the challenging problem caused by the gap between the information that can be extracted automatically from visual data and the interpretation that the same data has for a user in a given situation: the semantic gap. [1]

Because machine understanding of the video data is still an unsolved research problem, text annotations are usually used to describe the content of video data according to the annotator's understanding and the purpose of that video data. [2] In general, computer vision techniques may aid in answering the question "what is in the video?" but cannot answer questions such as "what is happening in the video?" or "what is the video trying to tell us?". For example, the background information of a video stream cannot be obtained directly from the video but needs to be annotated. [2][3]

Former video data models such as Informedia [4], VideoText [5], VideoSTAR [6], whether they use the video annotation layering (stratification) approach [5], [6] or the keyword-based annotation approach [4] to represent video semantics, fail to model semantic relationships among the concepts expressed in the video. The importance of capturing video semantic associations lies in the fact that it can greatly improve the effectiveness of video querying by providing knowledge-based query processing.

In addition to these words, Time is a key dimension of our information space, with many applications standing to benefit from exploiting it. Applications such as information extraction, question answering, summarization, visualization, and developments in the Semantic Web can all benefit from analysis and interpretation along the temporal dimension. For example, in summarizing a story in terms of a timeline, a system may have to extract and chronologically order events in which a particular person participated. [7] The issues of time are more important in applications like video data modeling where the video data is a temporal document.

On the other hands, nowadays, we can see the usefulness of ontologies as a content theory [8] in wide variety of applications. Important aspects of ontologies, which have been of major concern, are as follows:

- Ontologies are used to describe a specific domain
- The terms and relations are clearly defined in that domain
- There is a mechanism to organize the terms (commonly a hierarchical structure is used as well as IS−A or HAS−A relationships)
- There is an agreement between users of an ontology in such a way the meaning of the terms is used consistently
- Ontologies encode an implicit knowledge (semantic relations) in their structure.

we will describe how these aspects are used in our application.

In this paper, a general framework for unified temporal representation of video semantic is introduced and ontologies are used to archive this goal. The proposed framework takes video data as an input and it performs automatic temporal video annotation. Then, temporal annotations are organized in a new

[1] Corresponding Author: Behrang QasemiZadeh, qasemizadeh@digitalclone.net

ontological infrastructure, which is called temporal ontology to deliver semantics of video integrated with temporal information to the user. In the proposed framework, ontologies are served to represent domain knowledge as well as representative model of input video data. This paper is organized as follows: next section describes related works in short. The intended meaning of ontologies is described in section 3. In section 4 issues about time modeling and temporal reasoning are discussed. In section 5, the proposed framework is described. Experimental results are the subject of section 6. Conclusion and future works are discussed in section 7.

2 RELATED WORKS

One of the old challenges in artificial intelligence and more specific in the field of "Information Retrieval" is the semantic retrieval of information. Although classical information retrieval systems achieve to some success for semantic retrieval of structured data but therefore, there are few enhancements for semi or unstructured data like video data. Several systems have been proposed to solve this problem [9][10][5][6][4] but they have their special strength and weakness. This is come back to the nature of these data.

In [2] a video data model is proposed based on conceptual graphs. It utilizes the conceptual graph knowledge representation formalism to capture the semantic associations among the concepts described in text annotations of video data. Although there is no explicit description of modeling underlying temporal information of video data. In addition, there is no framework for temporal reasoning about the conceptual model, which is introduced. Besides, an approach for knowledge assisted semantic analysis and annotation of video content, based on an ontology infrastructure is presented in [9].

Simultaneously, there are several systems try to prepare automatic annotation for videos. [1][11] But actually, they do not specify how these annotations can be used for semantic reasoning about their underlying video data. In addition, most of these frameworks are domain specific and they do not show how they can be used for different kind of videos. Moreover, they do not specify how they models temporal issues in their prepared annotation.

In most of these works, proposed frameworks do not support temporal relations. Actually, there is no concern about the nature of video as a sequence of temporal data. In addition to these works, as we have mentioned before there are several works which are intended to prepare a model for video data but in most of them, the lack of underlying basic logic for time is seen.

3 TIME AND TIME ONTOLOGY

Discussion about time is rooted in philosophy and studies about these topic results in Common fundamental ontological issues, which are listed below:

-Primitive Time Entities
-Time Topology
-Temporal Relationships
-Boundedness
-Time Structure
-Temporal Metrics

Fortunately the problems computer engineers are faced with, do not ask for metaphysical answers on the very nature of time but they need a set of pragmatic guidelines, which could assist designers and programmers in the realization of architectures and applications. Temporal Reasoning is the major of concern in AI applications. [12]

Temporal Reasoning consists of formalizing the notion of time and providing means to represent and reason about the temporal aspects of knowledge Hence a Temporal Reasoning framework should provide:

-An extension to the language for representing the temporal aspects of the knowledge.

-A Temporal Reasoning System. A method for reasoning about the assertions, which are formed using the extended language, which allows one to determine the truth of any temporal logical assertion. [13]

In this domain, ontologies can be used as content theories to provide fundamental ontological issues about time. A time ontology aims to develop a representative ontology of time that expresses temporal concepts and properties. Time axiomatization can be declared by using time ontologies. This specifies what sort of object is going to be taken as the primitive to represent time, and imposes a set of constraints on temporal relations. Since now, there has been some efforts toward a common framework for this axiomatization e.g. [14].

DAML-TIME [7] [15] and SOUPA time [16] are well known examples of time ontologies. In [17], sub-ontology of time is introduced. The purpose of this entry sub-ontology of time is to provide quick access to the essential vocabulary in OWL for the basic temporal concepts and relations. It is believed that this ontology should be able to help to describe most of the temporal properties of real world services, since they usually only require basic topological relations, and information about durations, dates and times.

In our application, a time ontology is fitted in the proposed framework to describe desired abstraction of time according to the user point of view and the application.

5 THE PROPOSED FRAMEWORK

Figure 1 shows the block diagram of the proposed framework. User may define domain knowledge, which includes time ontology and domain ontology. User can use provided interfaces to define ontologies; also it is possible that this knowledge imported from other existed ontologies. In fact, temporal elements of domain ontology could be related to one of the primitives, which are described in time ontology.

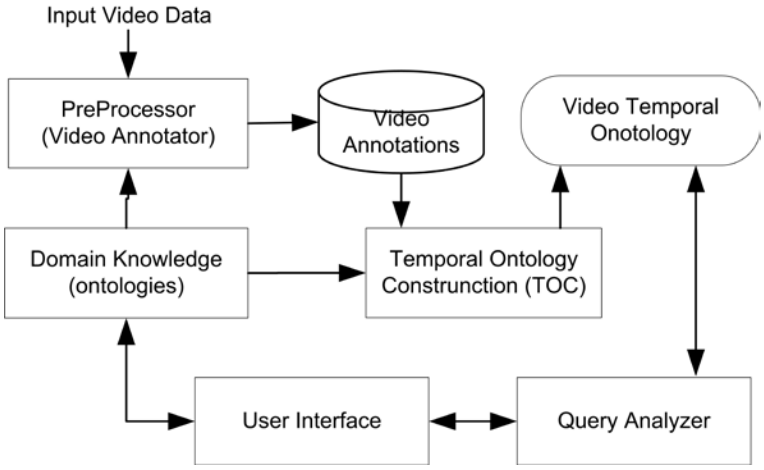

Fig. 1. Block diagram of the proposed system.

The task of Preprocessor is to provide video annotations in terms of domain knowledge elements. For this reason, preprocessor uses meta-rules. These rules are made of domain knowledge primitives as its constituents, logical operators and video analysis tools. In other words, user designed a pipeline from the provided elements to provide proper annotation from a video. In addition, it performs temporal information for each output annotation according to the specified primitives in domain knowledge. More detail description about annotation process can be found in [18].

TOC module provided Temporal Ontology from the Video Annotations. In literature, usually there is no difference between the terms Time Ontology and Temporal Ontology but in our work, we distinguish between the terms Time Ontology and Temporal Ontology. In the proposed framework, Temporal Ontology is defined as follow: an ontology, which is used to represent the body of knowledge of a certain example (here a certain video data) in a domain which is augmented with temporal primitives, which are defined in time ontology. As an example, consider an ontology, which is described using RDF. RDF describes the ontology in terms of triples. Here the temporal ontology will be described using Temporal RDF, which is introduced in [19]. In [19] there is no explicit assertion about time. we made this assertion using a Time Ontology. In this way, the framework of temporal reasoning will be fully described. In the current version of the proposed system, temporal ontology is constructed using the structure of domain knowledge without any additional process. However, it is possible to utilize other algorithms to construct temporal ontology depending on the application of the system.

To demonstrate the advantages of framework, we have added a simple query analyzer. User can query about events, concepts or anything, which is described in domain ontology in addition to temporal information. for example he/she can ask about "All events which occurred before the event B". Here, query analyzer is provided with a temporal reasoner. In addition, the User Interface, provide a Time-Line representation of videos. Note that domain knowledge is defined for a domain of interest e.g. Soccer match

videos, but temporal ontology is made for each input instance.

6 CONCLUSION AND FUTURE WORKS

Meaning is not a datum that is present in the image or video and that can be computed and decoded prior to the query process. It is rather a complex instantiation of static and dynamic elements emerging from relations within the system: database record itself, temporal context, user's circumstances, etc. [1] During the last few years there has been increasing research effort put toward automatic generation of links between low-level features and high-level concepts.

In this paper, a new framework for automatic video abstraction independent of video data was proposed. The proposed framework provides a system for automatic video annotation according to the defined domain knowledge and it constructs a temporal ontology from these annotations as an abstract model of an input video, which can be used to represent meaning of videos using temporal context and relation between entities in domain knowledge.

In short, the advantages of the proposed framework are as follows: suggesting a unified platform to define domain knowledge including temporal information, promising coordination between annotation process and modeling, temporal semantic representation, and providing a platform for temporal reasoning. In addition, the output of the system is an ontological infrastructure which can be used in data-mining application, from simple query analysis to video summarization and temporal relation mining.

REFERENCES

[1] J. Calic, and E. Izquierdo. A Rule-Based Video Annotation System, Andres Dorado, IEEE Transactions On Circuits And Systems For Video Technology, VOL. 14, NO. 5, MAY 2004.

[2] F. Kokkoras, H. Jiang, I. Vlahavas, A.K. Elmagarmid, E.N. Houstis, and W.G. Aref, Smart Videotext: A Video Data Model Based on Conceptual Graphs, Multimedia Systems, Springer-Verlag, 2002.

[3] H. Jiang, D. Montesi, and A.K. Elmagarmid, Integrated Video And Text for Content-Based Access to Video Databases , Multimedia Tools , 1999.

[4] Informedia Digital Video Library Project Web Site, URL: http://www.informedia.cs.cmu.edu, Carnegie Mellon University, USA.

[5] H. Jiang, D. Montesi and A.K. Elmagarmid. VideoText Database Systems, in the Proceedings of the 4th IEEE International Conference on Multimedia Computing and Systems, pp. 334–351, 1997.

[6] R. Hjelsvold, Videostar - A Database For Video Information Sharing, PhD Thesis, Norwegian Institute of Technology, November 1995.

[7] J. R. Hobbs and F. Pan, An Ontology of Time for the Semantic Web**,** ACM Transactions on Asian Language Information Processing, Vol. 3, No. 1, March 2004, Pages 66-85.

[8] B. Chandrasekaran, JR Josephson, and VR Benjamins**,** What are Ontologies and Why Do We Need Them? IEEE Intelligent Systems, 14(1):20-26, Jan/Feb 1999.

[9] S. Dasiopoulou, V. K. Papastathis, V. Mezaris, I. Kompatsiaris, and M. G. Strintzis, An Ontology Framework For Knowledge-Assisted Semantic Video Analysis and Annotation, Proc. 4th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot 2004) at the 3rd International Semantic Web Conference (ISWC 2004), November 2004, Hiroshima, Japan.

[10] H. C. Shih and Huang C. L., A Semantic Network Modeling for Understanding Baseball Video, ICASSP, 2003, Hong-Kong.

[11] N. Nitta, Semantic Content Analysis Of Broadcasted Sports Videos With Intermodal Collaboration, PhD Thesis, Osaka University, 2003.

[12] I. Mani, J. Pustejovsky, B. Sundheim: Introduction To The Special Issue on Temporal Information Processing, ACM Trans. Asian Lang. Inf. Process. 3(1): 1-10 (2004)

[13] L. Vila, A Survey of Temporal Reasoning in Artificial Intelligence, AI Communications 7(1), 4--128, 1994.

[14] J. Ma, B. Knight, A Reified Temporal Logic, Comput. J. 39(9), 800-807, 1996.

[15] DAML-Time Homepage, http://www.cs.rochester.edu/~ferguson/daml/ .

[16] H. Chen, F. Perich, T. W. Finin, A. Joshi: SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications. MobiQuitous 2004: 258-267.

[17] F. Pan and J. R. Hobbs, Time in OWL-S, In Proceedings of AAAI Spring Symposium Series on Semantic Web Services, 2004.

[18] S. H. Rastgu Haghi , B. QasemiZadeh, M. Kangavari, A Novel Approach for Video Temporal Annotation, I International Conference On Multidisciplinary Information Sciences And Technologies, Spain, 2006.

[19] C. Gutiérrez, C. A. Hurtado, A. A. Vaisman, Temporal RDF, ESWC 2005: 93-107.