A Survey Paper on Flexible and Improved Method for Automatic Semantic Content Extraction in Videos.

Mrs. Prajakta Chaudhari
G.S.Moze College of Engineering, Balewadi, Pune-45.
University Of Pune, Pune, India.

Prof. Ratnaraj Kumar
G.S.Moze College of Engineering, Balewadi, pune-45.
University Of Pune, Pune, India.

Abstract

Today , video based applications are used so widely. So it is necessary to extract video content. Now user wants deeper understanding of video content at semantic level. The content of video data at a semantic level is out of the scope of the standards because low-level features and raw data of a video hardly provide semantics which is much more appropriate for users. Manual techniques take too much time to extract semantic content. In this paper, we introduce semantic content extraction system which extracts objects, events, and concepts automatically.

Keywords– Semantic content extraction, video content modeling, ontology.

1. Introduction

Now days, there is rapid increase in the available amount of video data and it has caused an urgent need to develop intelligent methods to model and extract the video content. There are so many applications in which modelling and extracting video content are crucial. These applications are 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 video content at semantic level.

Video content can be extracted in three levels which are raw video data, low-level features and semantic content. First, raw video data consist of some general video attributes such as format, length, and frame rate. Second, low-level features are audio, text, and visual features. Third, semantic content contains high-level concepts such as objects and events. Users are mostly interested in retrieving the video in terms of what the video actually contains. The content of video data at a semantic level is out of the scope of the standards because low-level features and raw data of a video hardly provide semantics which is much more appropriate for users.

There are many methods which are making use of manual semantic content extraction methods. But manual semantic content extraction methods have some limitations like they are tedious, subjective, and time consuming. To solve such problem, it is necessary to fix the semantic gap between the low level feature and the high level feature [1seminar]. For this purpose, multimedia retrieval system with ontology concept is attempted.

Ontology is a formal, explicit specification of domain knowledge: it consists of concepts, concept properties, and relationships between concepts and is typically represented using linguistic terms, and has been used in many fields as a knowledge management and representation approach. [12] present a semantic content analysis framework based on a domain ontology. Domain Ontology is used to define semantic events with a temporal description logic where event extraction is done manually and event descriptions only use temporal information.

In [3], simple periodic events are recognized where the success of event extraction is highly dependent on robustness of tracking. In [5], the event recognition methods are based on a heuristic method. Those methods could not handle multiple-actor events. Event definitions are made through predefined object motions and their temporal behaviour.

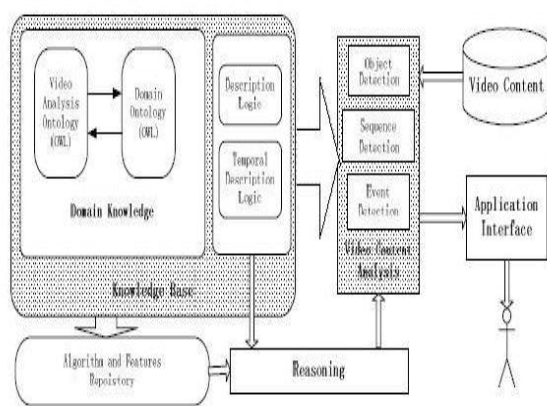## 2. Framework for Video Semantic Content Analysis based on Ontology

Figure 1. Framework

In[4],proposed video semantic content analysis framework is shown in Fig.1. Framework consists of:

-Video analysis Ontology: Knowledge for video analysis is collected and Video analysis ontology is constructed. It gives key elements in Video content analysis and helps to support detection process of corresponding domain specific semantic contents.

-Domain Ontology: Domain ontology consists of semantic concepts of the examined domain area. Also it has qualitative attributes of the semantic content, low-level features and video processing algorithms which determined by the semantic content of video to be detected and its low level features. ;

-Description Logic: DL is used to describe how video processing methods and low-level features should be applied according to different semantic content, aiming at the detection of special semantic objects and sequences corresponding to the high-level semantic concepts defined in the ontology.;

-Temporal Description Logic: TDL model focuses temporal relationships and define semantically important events in the domain.;

-The OWL language is used for knowledge representation for video analysis ontology and domain ontology.

- Reasoning based DL and TDL can carry out object, sequence and event detection automatically. Based on this framework, video semantic content analysis depends on the knowledgebase of the system. This framework can easily be applied to different domains provided that the knowledge base is enriched with the respective domain ontology.

3. Limitations of Existing Methods

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.

4. Proposed Solution

In this we are presenting the extended method for efficient extraction of video contents. The main aim of this is to present the annotation based automatic semantic content extraction framework for video applications. For this purpose the ontology based semantic model as well as semantic content extraction algorithms are proposed. At first, we are presenting a Meta ontology, a rule construction standard which is domain independent, to construct domain ontologies. Secondly, success of the automatic semantic content extraction framework is improved by handling fuzziness in class and relation definitions in the model and in rule definitions. Hence we are further introducing the use of fuzzy rules for proposed framework. In addition to this we want this framework more flexible and dynamic to address the different viewing of camera angles. Hence we are adding the functionality to 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.

5. Related Work

In [2], as amounts of publicly available video data grow the need to query this data efficiently becomes significant. Consequently content-based retrieval of video data turns out to be a challenging and important problem. The specific aspect of inferring semantics automatically from raw video data is addressed. In particular, a new video data model is introduced that supports the integrated use of two different approaches for mapping low-level features to high-level concepts.

Firstly, the model is extended with a rule-based approach that supports spatio-temporal formalization of high-level concepts, and then with a stochastic approach. Furthermore, results on real tennis video data are presented, demonstrating the validity of both approaches, as well us advantages of their integrated use.

In [4], shown that there is rapid increase in the available amount of video data which is creating a growing demand for efficient methods for understanding and managing it at the semantic level. New multimedia standards, such as MPEG-4 and MPEG-7, provide the basic functionalities in order to manipulate and transmit objects and metadata. But importantly, most of the content of video data at a semantic level is out of the scope of the standards. In this paper,a video semantic content analysis framework based on ontology is presented. Domain ontology is used to define high level semantic concepts and their relations in the context of the examined domain. And low-level features (e.g. visual and aural) and video content analysis algorithms are integrated into the ontology to enrich video semantic analysis. OWL is used for the ontology description. Rules in Description Logic are defined to describe how features and algorithms for video analysis should be applied according to different perception content and low-level features. Temporal Description Logic is used to describe the semantic events, and a reasoning algorithm is proposed for events detection. The proposed framework is demonstrated in a soccer video domain and shows promising results.

In [6], present a new representation and recognition method for human activities. An activity is considered to be composed of action threads, each thread being executed by a single actor. A single-thread action is represented by a stochastic finite automaton of event states, which are recognized from the characteristics of the trajectory and shape of moving blob of the actor using Bayesian methods. A multi-agent event is composed of several action threads related by temporal constraints. Multi-agent events are recognized by propagating the constraints and likelihood of event threads in a temporal logic network. We present results on real-world data and performance characterization on perturbed data

6. Conclusions

The aim 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 idea behind this is to utilize domain ontologies. A domain ontology is used to define high level semantic concepts and their relations in the context of the examined domain. Low-level features (e.g. visual and aural) and video content analysis algorithms are integrated into the ontology to enrich video semantic analysis. These domain ontologies are generated with a domain independent ontology-based semantic content metaontology model and set of special rule definitions to extract semantic content automatically.

7. REFERENCES

[1] Yakup Yildirim,Adnan Yazici, Turgay Yilmaz, "Automatic Semantic Content Extraction in Videos Using a Fuzzy Ontology and Rule-Based Model," IEEE Transations knowl. Data Eng.25(1): 47-61(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 Spatiotemporal and Stochastic Recognition of Events," Proc. IEEE Int'l Workshop Detection and Recognition of Events in Video, pp. 75-82, 2001.

[4] L. Bai, S.Y. Lao, G. Jones, and A.F. Smeaton, "Video Semantic Content Analysis Based on Ontology," IMVIP '07: Proc. 11th Int'l Machine Vision and Image Processing Conf., pp. 117-124, 2007.

[5] 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.

[6] 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.

[7] 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.