INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

SURVEY ON RULE-BASED MODEL FOR SEMANTIC CONTENT EXTRACTION

J. Angelin Jenisha[1], Mrs. Rebecca Hudson[2]

[1]PG Full Time Student, CSE Department, Karunya University, Coimbatore

[2]Assistant Professor, CSE Department, Karunya University, Coimbatore

[1]erjenisha@gmail.com, [2]rebecca@karunya.edu

Abstract

Semantic contents are nothing but the object, event and concept instances that is present in the video. The problem is that the semantic contents cannot be easily extracted. The key issue in extracting the semantic content is its representation. In order to expand the modelling capabilities additional rules can be used. By using the rule-based model the spatial relation computation cost can be lowered and the usage of additional rules is also possible. The goal of this paper is to allow the users to retrieve some useful information from large amount of data in a semantically meaningful manner.

.

Keywords: Semantic content, Spatial relation, Modeling, Event, Raw data

1. Introduction

To model and extract the video content became a must because of the increase in the available amount of video. The video content consists of three levels. The first level consists of raw video data, the second level consists of low-level features and the third level consists of the semantic content. The semantic content extraction seems to be very difficult because video is an arrangement of frames in a order and there is no direct relation with the semantic content. The manual extraction process will be more time consuming and tedious. The querying capabilities will be limited in-case of the manual extraction process.

The Video Database System [1] has both the semantic query facilities and spatio-temporal facilities. The rule-based model is a popular technique that is used to handle the complex situation such as the representation problem. A rule-based system that makes use of knowledge-base is used in order to provide full support for spatio-temporal and similarity-based object-trajectory queries. Here the video clips are segmented into frames of shots with respect to the spatial relationship that is existing between the objects in the video frames. This rule-based approach to model spatial-temporal relation removes the need for computation of relation during the period of query processing which reduces the response time of the queries. Here the handling of semantic part is done by the object-relational database. The number of facts that need to be stored can be reduced by the usage of set of rules in the knowledge-base.

2. Methodology

The rule-based model is nothing but a modelling approach in which a set of rules will implies a model. By using those rules set a model can be formed. Video Semantic Content Model (VISCOM) is a metaontology that provides a independent rule construction standard. In order to speed up the extraction process and to handle some special situation there is a possibility to give additional rule definitions.

3. Video Content-Based Retrieval

Video content-based retrieval [2] provides modelling and querying capabilities. The video model consists of feature-based model and annotation-based model. Content-based retrieval of data does not work with some database indexing techniques that are based on B trees or hash-based model. The challenge is the need of appropriate query language that exceeds the power of traditional text-based query languages. The distinction between the structure and content of video can be made possible via the video modelling. The tool for managing the large volumes of video data starts with the construction of video structure. The modelling can be done either by feature-based modelling or semantic modelling. The main technique that is used to speed up the annotation process is visualisation. The drawback that is oriented with this approach is that a new model is needed to integrate the two techniques.

4. Layered Video Data Model

A Layered Video Data Model [3] provides a framework for mapping features to concepts. It provides flexibility in using different video processing techniques. In order for mapping low-level features to high-level concepts a new video data model which supports the integrated use of two different approaches. There are two main classes for categorization of video retrieval approaches. The first class is based on visual features such as shapes, texture, colour distribution which represent the low-level visual content. The second class contains annotation-based approaches which use keyword annotation to represent the high-level concepts. The major limitation that is oriented with these approaches are the process of searching is based only on the predefined attribute information and the manual annotation process is very tedious and time consuming process. The rule-based approach that is used here supports spatio-temporal formalization of high-level concepts.

5. VSAM Model

The VSAM model [4] deals with the observation of moving vehicles and humans. The event detection and analysis is based on the detection of moving objects and by estimating their speed. The difficulty with the motion detection is that in some case both the observer and few elements of the scene may be moving. The detection of regions that are moving is based on the current frame and the previous frame. To achieve robustness the behaviour inference is done on a buffered set of frames. By using the buffered set of frames a graphical representation of the moving objects is made. The limitation present here is stability

.

6. Activity Recognition Method

An Activity Recognition method [5] is composed of action threads and in which each thread is executed by a single actor. Multi-agent event is also possible which supports several action threads. The type of objects that are moving and their spatio-temporal interaction are analyzed and several challenges have to be addressed. In real video data the motion detection are often unstable due to the poor quality of video, shadows that are present and so on. Only based on the view point the interpretation of low-level features can be done. The execution of the same activity can be done by different actors that lead to the variety of temporal durations. The problem of redundancy may occur and also same motion pattern will be caused by other activities.

7. Video Event Graph

To learn the event structure from training videos, video event graph [6] is used. The video event graph consists of temporally correlated sub-events that automatically encode the event correlation graph. The event correlation graph specifies the number of occurrence of conditionally dependent sub-events. The interaction between the complex networks of agents is known as events. The visual representation of these events can be done by hierarchy of events and sub-events. The instance that is present in the event is the composition of directly measurable low-level action that will have a temporal order. The limitation that is present with these approaches is the ability to model some events that are complex and it contain multiple agents performing multiple actions simultaneously. The complex interactions can be made in the form of domain events, without knowing about the agents that are involved in the interaction and the length of the video is also not considered. The temporal structure of events can be represented as a Direct Acyclic Graph (DAG). The detection of events in a multi agent working environment is a bit harder when compared to single agent environment.

8. Video Database Management System

BilVideo that is the video database management system [7] provides support for queries on spatio-temporal, low-level features and semantic on the video data. BilVideo is domain independent and it can be used to support any kind of application video data. In order to handle the spatio-temporal queries a knowledge-base is used that consists of a fact-base and comprehensive set of rules. On the other hand, the queries in the semantic and low-level features are handled by an object-relational database. The knowledge-base consists of set of rules which reduces the number of facts that represent the relation which is to be stored for spatio-temporal querying of video data. Video clips are segmented into shots when the present set of relation between video objects changes. The BilVideo supports a finer granularity for query processing which is not dependent on the semantic segmentation the video clips. BilVideo is application independent because it supports any kind of application with spatio-temporal and semantic query requirements on video data. A visual query is formed in the BilVideo by using the collection of objects with certain conditions such as annotations and events. The feature database consists of semantic properties of videos that is used for category-based queries. These features are formed by video-annotator. The knowledge-base is used to handle spatio-temporal queries and the fact-based is done by fact-extractor.

9. Conclusion

The purpose of this survey is to compare the various rule-based methodologies that has been used. Additional rules can be added in the case of rule-based modelling which gives a credit to this approach. There are number of irregular situations that are not possible for representation which are defined in the ontology. The VISCOM is enriched with the additional rule definitions. The main usage of rule-based modelling is the ability to handle the complex situations

REFERENCES

[1] T. Sevilmis, M. Bastan, U. Gudukbay, 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.

[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. Bremond, 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. Bremond, "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. Donderler, E. Saykol, U. Arslan, O. Ulusoy, and U. Gudukbay, "Bilvideo: Design and Implementation of a Video Database Management System," Multimedia Tools Applications, vol. 27, no. 1, pp. 79-104, 2005.