ISSN:2320-0790

COMPUSOFT
An International Journal of Advanced Computer Technology

Efficient Semantic Video Data Content Extraction Using Fuzzy Ontology

[1]A.Rajeswary, [2] Mr. V.Gunalan

[1]PG, M.Tech. (Department of CSE), [2]Assistant Professor, Dept of CSE
Bharathiyar College of Engineering & Technology
Thiruvettakudy, Kariakal-609609

Abstract: The use of video-based applications has revealed the need for extracting the content in videos. Currently manual techniques which are inefficient subjective and costly in time and limit the querying capabilities are being 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. We introduce an ontology-based fuzzy video semantic content model that uses spatial/temporal relations in event and concept definitions. In addition to domain ontology we use additional rule definitions (without using ontology). 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, modelling, video semantic

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 modelling 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, colour distribution, shape, motion, etc. Third semantic content contains high-level concepts such as objects and events. The first two levels on which content modelling 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 mostly interested in

Querying and retrieving the video in terms of what the video contains. Therefore, raw video data and low level features alone are not sufficient to fulfil the user's need. Deeper understanding of the information at the semantic level is required in many 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. 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. There are many research works in this area. Most of them use manual semantic content extraction methods .Manual extraction approaches are tedious, subjective, and time consuming which limit querying capabilities. Besides, the studies that perform automatic or semiautomatic extraction do not provide a satisfying solution. Although there are several studies employing different methodologies such as object detection and tracking, multimodality and spatiotemporal derivatives, the most of these studies propose techniques for specific event type extraction or work for specific cases and assumptions. In simple periodic events are recognized where the success of event extraction is highly dependent on robustness of tracking. The event recognition methods described in are based on a heuristic method that could not handle multiple-actor events. Event definitions are made through predefined object motions and their temporal behaviour. The shortcoming of this study is its dependence on motion detection. In, scenario events are modelled from shape and trajectory features using a hierarchical activity representation extended propose a method to detect events in terms of a temporally related chain of directly measurable and highly correlated low level actions (sub events) by using only temporal relations. Another key issue in semantic content extraction is the representation of the semantic content. Many researchers have studied this from different aspects. A simple representation could relate the events with their low-level features (shape, colour, etc.) using shots from videos, without any spatial or temporal relations. However, an effective use of spatiotemporal relations is crucial to achieve reliable recognition of events. Employing domain ontology facilitates use of applicable relations on a domain. There are no studies using both spatial relation between objects, and temporal relations between events together in an ontology-based model to support automatic semantic content extraction. Studies such as extended-AVIS, multiview and class View [1propose methods using spatial/temporal relations but do not have ontology-based models for semantic content representation. present a semantic content analysis framework based on a domain ontology that is used to define semantic events with a temporal description logic where event extraction is done manually and event descriptions only use temporal information. propose an ontology model using spatiotemporal relations to extract complex events where the extraction process is manual. In each linguistic concept in the domain ontology is associated with a corresponding visual concept with only temporal relations for soccer videos. Define an event ontology that allows natural representation of complex spatiotemporal events in terms of simpler sub events. 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 Mark up Language (VEML) is used to manually annotate VERL events in videos. The lack of low-level processing and using manual annotation are the drawbacks of this study present a systematic approach to address the problem of designing ontology for visual activity recognition. The general ontology design principles are adapted to the specific domain of human activity ontology's using spatial/temporal relations between contextual entities. However, most of the contextual entities which are utilized as critical entities in spatial and temporal relations must be manually provided for activity recognition provide a detailed survey of the existing approaches for semantic content representation and extraction. Considering the above-mentioned needs for content based retrieval and the related studies in the literature, methodologies are required for automatic semantic content extraction applicable in wide-domain videos In this study, a new Automatic Semantic Content Extraction Framework (ASCEF) for videos is proposed for bridging the gap between low-level representative features and high-level semantic content in terms of object, event concept, spatial and temporal relation extraction. In order to address the modelling need for objects, events and concepts during the extraction process, a wide-domain applicable ontology-based fuzzy Video Semantic Content Model (VISCOM) that uses objects and spatial/temporal relations in event and concept definitions is developed. VISCOM is Meta ontology for domain ontologies and provides a domain-independent rule construction standard. It is also possible to give additional rule definitions (without using ontology) for defining some special situations and for speeding up the extraction process. ASCEF performs the extraction process by using these meta ontology- based and additional rule definitions, making ASCEF wide-domain applicable. In the automatic event and concept extraction process, objects, events, domain ontologies, and rule definitions are used. The extraction process starts with object extraction. Specifically, a semiautomatic Genetic Algorithm-based object extraction approach is used for the object extraction and classification needs of this study. For each representative frame, objects and spatial relations between objects are extracted. Then, objects extracted from consecutive representative frames are processed to extract temporal relations, which is an important step in the semantic content extraction process. In these steps, spatial and temporal relations among objects and events are extracted automatically allowing and using the uncertainty in relation definitions. Event extraction process uses objects, spatial relations between objects and temporal relations between events. Similarly, objects and events are used in concept extraction process. This study proposes an automatic semantic content extraction framework. 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 modelling and semantic content extraction research areas. First of all, we propose a meta ontology, a rule construction standard which is domain independent, to construct domain ontologies. Domain ontologies are enriched by including additional rule definitions. The 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. A domain-independent application for the proposed system has been fully implemented and tested. As a proof of wide-domain applicability, experiments have been conducted for event and concept extraction for basketball, football, and office surveillance videos. Satisfactory precision and recall rates in terms of object, event, and concept extraction are obtained by the proposed framework. Our results show that the system can be used in practical applications. Our earlier work can be found in the organization of the paper is as follows. The proposed video semantic content model is described in detail. The automatic semantic content extraction system is performed experiments and the performance evaluations of the system are given.

Video Semantic Content Model

In this section, the proposed semantic video content model and the use of special rules (without using ontology) are described in detail.

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. In this study, we propose a wide-domain applicable video content model in order to model the semantic content in videos. VISCOM 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. Without any standard on rule construction, different domains can have different rules with different syntax. In addition to the complexity of handling such difference, each rule structure can have weaknesses. Besides, VISCOM provides standardized rule construction ability with the help of its metaontology. It eases the rule construction process and makes its use on larger video data possible. The rules that can be constructed via VISCOM 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, VISCOM provides an additional rule based modelling capability without using ontology. Hence, VISCOM 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 as which use objects and spatial and temporal relations for semantic content modelling neither use ontology in content representation nor support automatic content extraction. To the best of our knowledge, there is no domain-independent video semantic content model which uses both spatial and temporal relations between objects and which also supports automatic semantic content extraction as our model does.

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. Each video instance, Vi 2 Video Database, is represented as where is the set of key frames of n is an attribute of video metadata that represents the domain of the video instance is the set of all possible domains.

Ontology-Based Modelling:

The linguistic part of 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. Relations defined in VISCOM give ability to model events and concepts related with other objects and events. VISCOM is developed on ontology-based structurewhere 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. C-Logic is used for the formal representation of VISCOM 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. VISCOM is represented with the following C-logic formulation where the predicate in (Entitiy, Class) is used to mean “an entity is defined as an individual of a class” in the formal representation of classes.

All of the VISCOM classes and relations are given in Fig. 1.Red colour arrows represent is-a relation, blue-colour arrows represent has-a relations. Below, the VISCOM classes are introduced with their description, formal representation and important relation(property) descriptions.

Component:

VISCOM 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 subclasses as Objects, Events, and Concepts and is represented as Component :

type)fOi; Ej; Ckg; sim) fSmg;
synname) ½string_
_ _
where
indðOi; ObjectÞ; indðEj;EventÞ;
indðCk;ConceptÞ; indðSm; SimilarityÞ;
at most one of i; j; k >0;
8>>>>>><
>>>>>>:
ð2Þ
where property has Similar Context is used to associate similar components in a fuzzy manner when there is a similar component in the ontology with a component that is supposed to be extracted.

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. Basketball player, referee, ball and hoop are examples of objects for the basketball domain
Object :
name) ½string_; lowLevelFeature) fLig;
composedOf)fCORjg
_ _
where
indðLi; LowLevelFeatureÞ;
indðCORj;ComposedOfRelationÞ;
8>>>><
>>>>:
ð3Þ
where property has Composed Of Object Relation is used to define concept inclusion, membership, and structural object relations such as part of, member of, substance of, is a, and composed of. It has a relevance degree and a reference to an "object composed-of group" individual in its definition
.

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. In addition, temporal event relations can also be used in event definitions. An event has a name, a definition in terms of temporal event relations or spatial/temporal object relations, and role definitions of the objects taking part in the event. Jump ball, rebound, and free throw are examples of events for the basketball domain Event
name) ½string_; eventDef) fEDig;

where property has Temporal Event Component is used to define temporal relations between events which are used in the definition of other events. Has Event Definition is utilized to associate events with event definitions. An event can be expressed with more than one event definition.

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. Attack and defence are examples of concepts for the basketball domain.

½ name) ½string_; conceptComp) fCCig _
where
indðCCi;ConceptComponentÞ; i>0;
8<
:
ð5Þ
where property hasConceptComponentis used to define the relation that exists in concept's meaning.

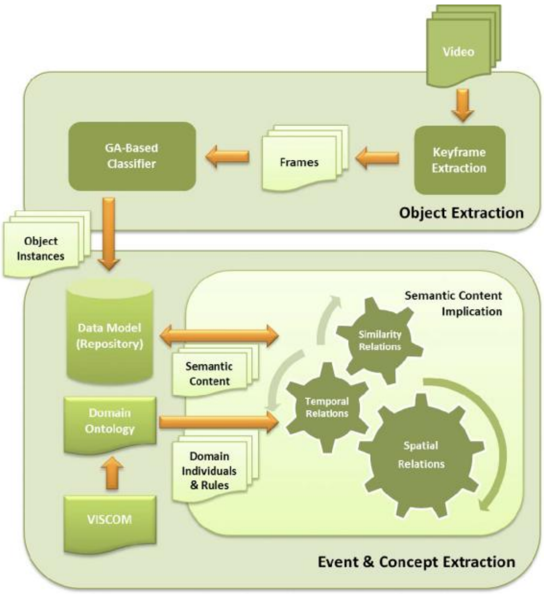
Spatial Relation:

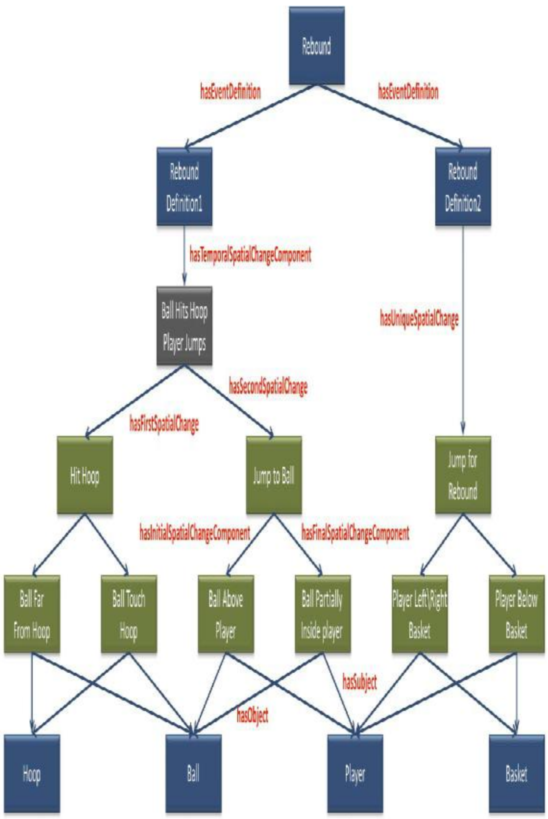
Spatial relations express the relative object positions between two objects such as above, inside, or far. The spatial relation types are grouped under three categories as topological, distance and positional spatial relations. The individuals of this class are utilized by the individuals of Spatial Relation (Component class.
Spatial Relation Component
Spatial Relation Component) class is used to represent spatial relations between object individuals. It takes two object individuals and at most one spatial relation individual from each subclass of Spatial Relation class. This class is utilized in spatial change and event definition modelling. It is possible to define imprecise relations by specifying the membership value for the spatial relation individual used in its definition. For the basketball domain, Player under Hoop is an example of Spatial Relation Component class individuals Spatial Change

Spatial Change:

class is utilized to express spatial relation changes between objects or spatial movements of objects in order to model events. Spatial regions representing objects have spatial relations between each other. These relations change in time. This information is utilized in event definitions. Temporal relations between spatial changes are also used when more than one spatial change is needed for definition. This concept is explained under Temporal Relations and Event.

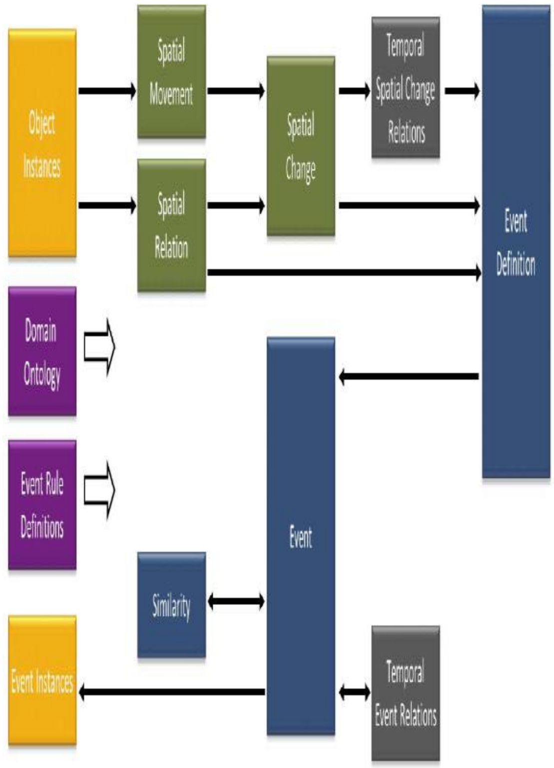
Automatic Semantic Content Extraction Framework:





Rebound Event Representation:

Event Extraction Process:



Concept Extraction Process:

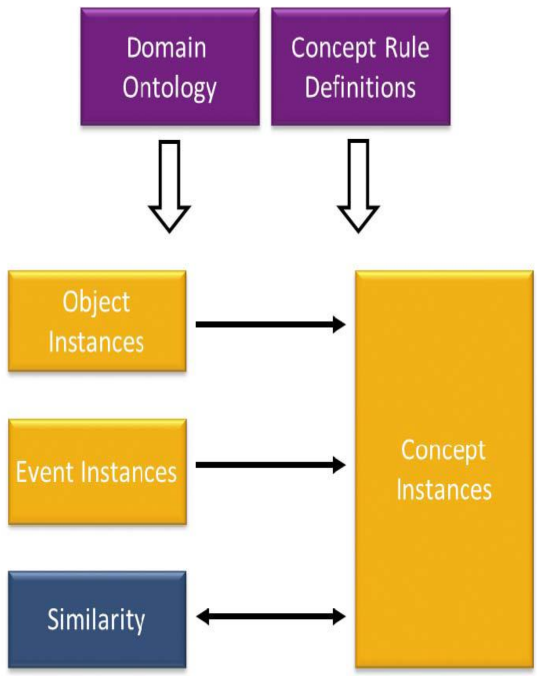
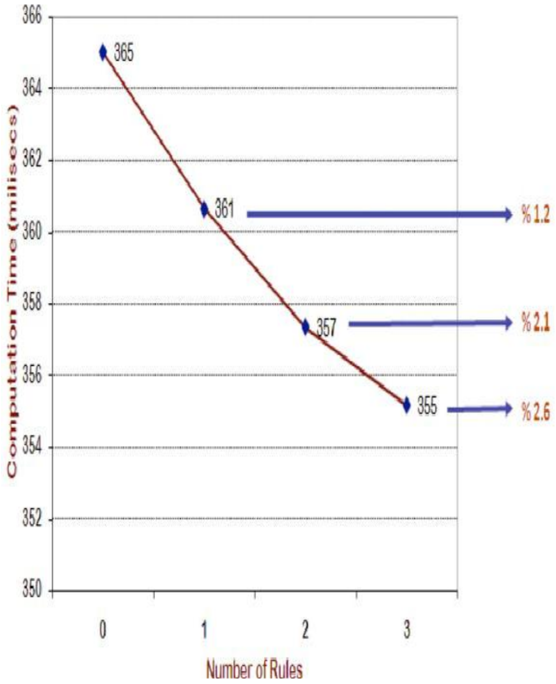


Table Relation:



Reference:

[1] 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, .

[2] 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,
.

[3] 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,

[4] S. Hongeng, R. Nevatia, and F. Bre´mond, "Video-Based Event Recognition: Activity Representation and Probabilistic Recognition Methods," Computer Vision and Image Understanding.

[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,

[6] M. Ko¨pru¨ lu¨, N.K. Cicekli, and A. Yazici, "Spatio-TemporalQuerying in Video Databases," sss