Journal of Theoretical and Applied Information Technology
20th June 2014. Vol. 64 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ONTOLOGY AND RULE-BASED MODEL FOR EXTRACTING SEMANTIC CONTENT IN VIDEOS

[1]KARTHIK GANESH.R, [2]Dr.R.KANTHAVEL, [3]ALEXANDRIYA CELIN.V

Research Scholar, Information and Communication Engineering, Anna University of Technology,Tvl.
Professor &Head, Dept. of ECE Vellammal Engineering College, Chennai
PG Scholar, Computer Science and Engineering, SCAD College of Engineering & Technology, Tvl.
E-mail: [1]Karthiganesh2k4@gmail.com, [2]r_kanthavel@yahoo.com,
[3]alexandriyavincent@gmail.com

ABSTRACT

As a large amount of video based applications becomes widely available, there is a need to extract the semantic data in videos. Manual techniques, which are inefficient and subjective and costly in time, bound querying capabilities and bridge gap between low-level representative feature and high level semantic content. For the deeper understanding apart from the raw data and low-level features, the content at semantic level is required. The existing periodicity mining was based on audio retrieval which cannot be applied to the video. Optical flow algorithm is used to track the direction of the spatial movements between two image frames. Ontology provides a vast domain applicable rule production set that allows the user to construct the domain. Rule based model use the spatial/temporal relations to extract the definitions. The moving regions are extracted by the Graph-Based Visual Saliency (GPVS) algorithm and it detects the temporal locations. The phase correlation method is a well-known image mining technique in multimedia data which has main application in coordination of frames.

Keywords: *Spatial Movement, Ontology, Spatial/Temporal Relation, Rule Based Model, Phase Correlation.*

1. INTRODUCTION

The increase in the available amount of video data has caused a need to develop the model and extract the content in video. A multimedia database system should be able to extract semantic contents from the multimedia data, and it stores the index semantic contents efficiently. The user has to retrieve the content of video data in an efficient and semantically significant manner. There are three levels of video content which are unprocessed video data, low-level features and semantic content. Raw video data contains elementary physical video element mutually with some common video attributes such as layout, length, and frame rate. Low-level features are categorized by audio, text, and visual features such as surface, color allocation, outline, faction, etc.

The semantic content contains high-level concepts such as objects, events with values. The first two levels on which substance modeling and extraction approach are based on routinely extracted data, which is represented by the low-level content of a video, but they hardly offer semantics which is more appropriate for users. The main issue in semantic content extraction is the representation of the semantic content. A simple representation could relate the events with their low-level features using shot from videos, without any spatial or temporal relations. An effectual use of spatiotemporal relations is to recognize the events with their actions.

Video event recognition allows the user to classify the events without interacting with the low level features and values that are defined. The shortage of low-level processing and using manual explanation are the drawbacks. The necessity for objects, events and concepts during the extraction process, a wide-domain related ontology-based fuzzy model that uses objects and spatial/temporal relations in event and rule definitions are developed. A meta-ontology for domain ontology provides an independent rule production standard. It is also potential to give additional rule definitions for defining some particular situations and for speeding up the extraction process. Rule based model has number of rules that has individual values for the domain class. The optical flow algorithm used to find the direction of the object

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and find the magnitude flow with respect to the spatial changes of objects. The temporal relations are specified by the saliency features.

2. RELATED WORK

Many interesting techniques of periodicity pattern mining were proposed to detect various kind of periodicities namely character, series and segment periodicity and values. The Periodicity Mining techniques are audio based which is more appropriate for numerical and alpha numerical data for deriving periodic patterns.

The ability to handle the faultily occurring periodicities is narrow to assured techniques at the cost of poor memory management and restricted variety of periodicity detection [3]. Only few methods are flexible to noise, but those techniques possess larger response time. However, it is very tricky to extract semantic content directly from raw video data. This is because video is a temporal series of frames without a direct relation to its semantic content [2]. The Video Semantic Content model retrieves [1] the objects but the motion of the objects are not considered. The event recognition methods described in [4] are based on a heuristic method that could not switch multiple-actor events.

In [5], scenario events are modeled from outline and trajectory features using a hierarchical activity representation extended from [4]. Studies such as Bil Video [7], [8], extended-AVIS [9], multi View [10] and class View [11] propose methods using spatial/temporal relations but do not have ontology-based models for semantic content illustration. Manual extraction approaches are monotonous, subjective, and time overriding [6], which limit querying capabilities.

The optical flow computation use combined local-global method to find the variation methods [15] in the image sequence. The Hierarchical Temporal Memory [HTM] comprises a hierarchical tree structure that exploits enhanced spatiotemporal modules to memorize objects appearing in various orientations [16] which has the limitation in temporal event sequence with locations.

3. PROPOSED SYSTEM

The optical flow algorithm calculates the spatial movement direction and magnitude of the object. Saliency features may detect the temporal locations with the variation point of code book. The rule based model is used to define the type of the video by extracting the objects and its event. It specifies the center of the tracked object with respect to the variation points to classify the events.

3.1. Phase Correlation

Phase correlation is the method of detecting the peak of the correlation function by comparing the cross correlation ideals. It is used to evaluate two functions and coordinate the frames with phase differences. Cross Correlation is a numerical apparatus for finding repeating patterns in interrupted signals by analyzing the degree of similarity between them.

3.2. Preface to Phase Correlation

The phase correlation method is a well- known image mining technique in multimedia data which has main application in coordination of frames. This method relies on estimation the most of the phase-only correlation (POC) function, with sub pixel accuracy, which denotes cross-spectrum between two descriptions. The coordinates are maximum correspond to the translation between the two images. Phase correlation utilizes a fast frequency transform method to estimate the relative movement between two similar images corrupted by independent Gaussian noise. Let $f_1(a,b)$ and $f_2(a,b)$ be two functions that are absolutely integrable over the interval[14].

$$f_2(a,b)=f_1(a-a_0,y-y_0).$$

Fourier shift property states that,

$$\hat{f}_2(u,v) = \hat{f}_1(u,v)\exp(-i(ux_0+vy_0)).$$

When dealing with images, $f_1$and f_2 are specified with finite size discredited arrays.

3.3. Architecture For Extracting The Content From Video

The input video signal received as a dataset. The frames are separated from the input object. The spatial movements are used to find the location and direction and flow of the input object. The temporal sequence model traces the identity of the object by means of the saliency features. The object instances are the moving objects in the video. Event classification, classify the reduced frames.

Journal of Theoretical and Applied Information Technology
20th June 2014. Vol. 64 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

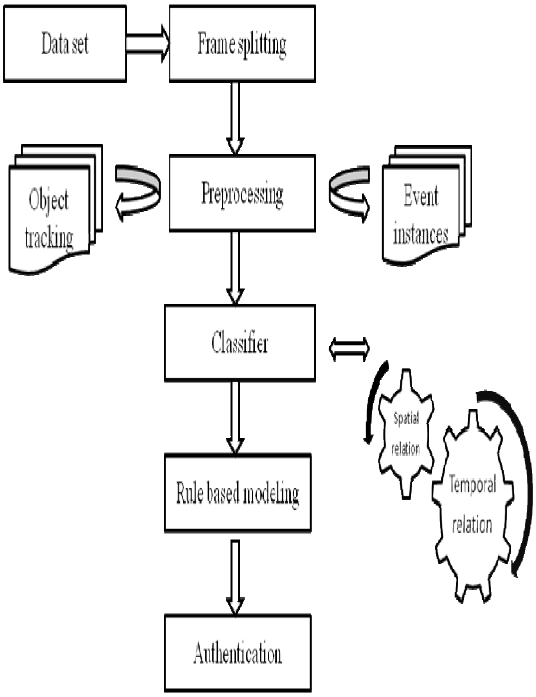


Fig: Architecture For Extracting The Content From Video

Rule based model track the whole object with the observation of variation points and it reduce the number of frames from the video. Event occurrences are extracted after a series of automatic extraction processes. Each extraction process resulting instances of a semantic content type defined as an character in the domain ontology. It describes the whole event extraction process. The event may classify from the rules designed and finally the type of object is authenticated and analyzed.

3.4. Rule Based Modeling

Additional rules are utilized to extend the modeling capabilities. Each rule has body part and head part where body part contains any number of domain class or property individuals and head part contains only one individual character with a value µ, representing the conviction of the definition given in the body part to represent the definition in the head part where O≤µ≤1. The fundamental rule has parentheses and logical connectives.

Additional rule definitions are hard to define situations as a natural component of ontology. The main purpose of rule definition is to lower the spatial relation computation cost. Functional exploitation is as follows: Inverse spatial relations and spatial relations that can be described in terms of other spatial relatives can be expressed with rule definitions. In the spatial relation mining process, these rules can be utilized to extract the content represented with the head part of the rule definition automatically.

3.5. Ontology Construction with the Components

Objects are belongs to the existential entities. An object is the origin point of the composition. The relevant object position between two objects is specified by the spatial relation. Each object has the subclass of spatial relation class, which represents the relation between the object individuals [1]. The spatial movements in the object specify the spatial changes.

Object: {
Low level feature $\Rightarrow$ $\{L_a\}$
Composed Of $\Rightarrow$ $\{COR_b\}$
Where,
COR (Composed Of Relation)
}

Events are long-term temporal objects and object relation values. They are described by using objects and spatial/ temporal relations between objects. Temporal Event Component is used to define the relation between the event individuals [1] from the event class. Event has number of definition with the membership values of individuals.

Event: {
EventDef $\Rightarrow$ $\{ED_a\}$
objectRole $\Rightarrow$ $\{OR_b\}$
temporalEventComp $\Rightarrow$ $\{TEC_c\}$
Where,
TEC(Temporal Event Component)
}

3.6 Algorithm for Detecting the Spatial Movements

The Optical Flow algorithm calculate approximately the direction and speed of object movement from individual image to another or from one video frame to a different using either the Horn-Schunck technique. To compute the optical flow between two images, you must resolve the following optical flow constraint equation:

$$\mathbf{I_x u + I_y v + I_t = 0}$$

In this equation, the following values are represented: I_x, I_y and I_z are the spatiotemporal

Journal of Theoretical and Applied Information Technology
20th June 2014. Vol. 64 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

image intensity derivatives. u and v are considered as horizontal optical flow and vertical optical flow. The Optical Flow block uses an iterative process to calculate the optical flow between two images or two video frames. Use the Stop iterative solution parameter to control when the iterative process stops. In Horn-Schunck technique, by presume that the optical flow is smooth over the complete image, the Horn-Schunck method computes an estimation of the velocity field and this technique separate the original image into slighter sections and assumes a constant velocity in each section.

Input: The periodic pattern of Input sequence (I) with default regularization parameter (α=15) and stopping threshold (ϵ =0.001).

Output: The direction and magnitude of spatial movement in the extracted frames are plotted from the input periodic video sequence.

Algorithm

1. Images with 3D array that contains a sequence of images.

2.Calculate the optical flow of a sequence of images. Initialize the parameter values [13] as mentioned below

$$u \leftarrow 0$$
$$y \leftarrow 0$$
$$n \leftarrow 0$$

3. While n<N and stopping-criterion> ϵ do

$$u \leftarrow \hat{u} - I_x \frac{I_x\hat{u} + I_y\hat{v} + I_t}{\alpha^2 + I_x^2 + I_y^2}$$

$$v \leftarrow \hat{u} - I_y \frac{I_x\hat{u} + I_y\hat{v} + I_t}{\alpha^2 + I_x^2 + I_y^2}$$

4. Compute $I_{x,}$ $I_{y,}$ I_t for each extracted frames (k=1: n-1).

3.7 Algorithm for Detecting the Temporal Locations

Graph Based Visual Saliency (GPVS) algorithm is used to represent the temporal locations in the saliency map, it's calculated from the direction of the tracked object. The code book is used to specify the variation points.

Input: The output of spatial movement direction of the periodic input sequence.

Output: Temporal locations of the objects are tracked by using variation points.

Algorithm

1. The input images undergo a saliency detection procedure by exploiting the Graph-Based Visual Saliency algorithm (GBVS).
2. The information rich regions coordinates of the event are extracted.
3. Hierarchical Temporal Model (HMT) specially designed with the detected salient locations of the events.
4. Each level consists of ad joint nodes, the number of which decreases ascending in the hierarchy.
5. Level 0 corresponds to the images presented to the network. The input images are also divided into patches of n by n pixels.
6. The codes are varied depending upon the variation method.
7. The nodes follow the same algorithmic procedures independent of the spatial or correlation level.

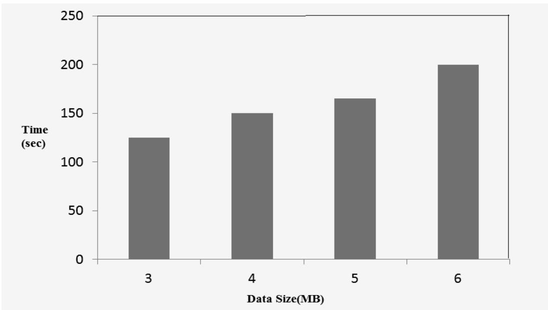
4. COMPARATIVE ANALYSIS

The analysis and knowledge representation with rules was extremely valuable for determining modularity and regularity in the video. The temporal and the spatial relations are used in this mining method. In order to estimate the consequence of additional rule, set of rules are defined in the rule based model**.** The temporal saliency algorithm is not possible to identify whether salient motion occurs before or after the central frame at which it is detected. So saliency map algorithm was proposed to capture the temporal sequences with the increased time performance. When compared to other methods the rule based modeling proposed for increasing the efficiency in tracking the objects.

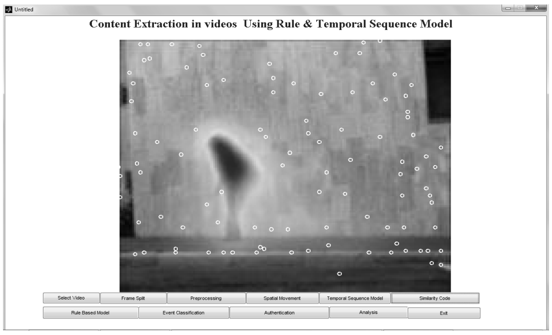
Journal of Theoretical and Applied Information Technology
20th June 2014. Vol. 64 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

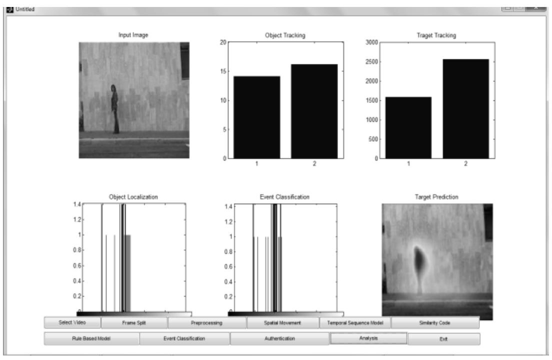
Time Performance:



Temporal and Rule based modeling:



Analysis:



5. CONCLUSION

The main objective of this paper is to extract the objects and events from the videos which can be utilized in more video applications. Rule based model contributes in several ways to extract semantic content in examine areas. The semantic content extraction process is done automatically. In addition, the semantic content representation capability and extraction achievement are improved by adding fuzziness in class, relation, and rule definitions in the multimedia data. Rule based model is restricted for some domains which are so complex that tens of thousands of rules would be needed to represent the huge number of feasible situations. In every component of the structure, ontology-based modeling and mining capabilities are used. The experiment results clearly show the success of the developed system. As a further study many number of objects will be extracted from the multimedia data.

REFERENCES:

[1] YakupYildirim, Adnan Yazici, Turgay Yilmaz, "Automatic Semantic Content Extraction in Videos using a Fuzzy Ontology and Rule Based Model" IEEE *Transactions On Knowledge And Data Engineering* January 2013 vol.25, no.1.

[2] M. Petkovic and W. Jonker, "An Overview of Data Models for Content-Based Video Retrieval," *Proc. Int'l Conf. Advances in Infrastructure for E-Business*, Computer Science Department, University of Twente, Aug 2000.

[3] Kanthavel.R, Karthik Ganesh.R, Jency Premalatha.M, "Correlation based Effective Periodic Pattern Extraction from Multimedia Data", *International Conference on Recent Trends in Computational Methods, Communication and Controls (ICON3C 2012)* data mining, 2012.

[4] G.G. Medioni, Hongeng, and R. Nevatia, "Event Detection and Analysis from Video Streams," *IEEE Transactions On Pattern Analysis And Machine Intelligence*, Aug 2001 vol.23, no.8, pp.873-889.

[5] S. Hongeng, R. Nevatia, and F. Bre´mond, "Video-Based Event Recognition: activity representation and probabilistic recognition methods" *Computer Vision and Image Understanding*, 13 Aug 2004, vol.9, pp. 129-162.

[6] M. Petkovic and W. Jonker, "Content-Based Video Retrieval Detection and Recognition of Events in Video", *Proc. IEEE Int'l Workshop,* 2001, pp. 75-82.

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

[8] T. Sevilmis, M. Bastan, U. Gudukbay, and O. Ulusoy, "Automatic Detection of Salient Objects and Spatial Relations in Videos for a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

video database system" *Image and Vision Computing*, Bilkent University 3 Jan 2008 vol. 26, no. 10, pp. 1384-1396.

[9] M. Kopru, N.K. Cicekli, and A. Yazici, "Spatio-Temporal Querying in Video Databases," *Information and Sciences*, Middle East Technical University, Ankara vol. 160, nos. 1-4, pp. 131-152, 2004.

[10] J. Fan, W. Aref, A. Elmagarmid, and X. Zhu, "Multiview: Multilevel Video Content Representation and Retrieval," *Journal of Electronic Imaging*, Oct 2001 vol. 10, no. 4, pp. 895-908.

[11] J. Fan, A.K. Elmagarmid, X. Zhu, and L. Wu, "Classview: Hierarchical Video Shot Classification, Indexing, and Accessing," *IEEE Transaction on Multimedia*, Feb 2004, vol. 6, no. 1, pp. 70-86.

[12] L. Bai, S.Y. Lao, G. Jones, and A.F. Smeaton, "Video Semantic Content Analysis Based on Ontology" *International Machine Vision and Image Processing Conference*, 2007, pp. 117-124.

[13] Enric Meinhardt-Llopis, Javier S´anchez, Daniel Kondermann, "Horn–Schunck Optical Flow with a Multi-Scale Strategy," *Image Processing*, 2013, vol. 20, pp. 151–172.

[14] H.Foroosh J. Zerubia and M. Berthod "Extension of Phase Correlation to Subpixel Registration" *IEEE Trans on image processing*, 3 march 2002 vol. 11, no.

[15] Andr´es Bruhn, Joachim Weickert, Christian Feddern, Timo Kohlberger, and Christoph Schn¨orr "Real-Time Optic Flow Computation with Variational Method" CAIP 2003, LNCS 2756, pp. 222–229, 2003.

[16] Ioannis Kostavelis, Lazaros Nalpantidis, and Antonios Gasteratos "Object Recognition Using Saliency Maps and HTM Learning" *IEEE Transaction, Management Engineering*, 2012.