International Journal of Research

International Journal of Research (IJR)

e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 05, May 2015

Available at http://internationaljournalofresearch.org

RGB and Histogram Based Searching Algorithm for Videos

Suyog Gupta[1], Himanshu Gupta[2]

[1]Department of Computer Science and Engineering
[2]Department of Information Technology
[1,2]Uttar Pradesh Technical University, India
E – Mail Address
er.suyogupta@gmail.com
shaktihimanshu@gmail.com

Abstract-

Advances in the media and entertainment industries, including streaming audio and digital TV, present new challenges for managing and accessing large audio-visual collections. Current content management systems support retrieval using low-level features, such as motion, color, and texture. However, low-level features often have little meaning for naive users, who much prefer to identify content using high-level semantics or concepts. This creates a gap between systems and their users that must be bridged for these systems to be used effectively. To this end, in this paper, we first present a knowledge-based video indexing and content management framework for domain specific videos. We will provide a solution to explore video knowledge by mining associations from video data. The explicit definitions and evaluation measures (e.g., temporal support and confidence) for video associations are proposed by integrating the distinct feature of video data. Our approach uses video processing techniques to find visual cues introduces multilevel sequential association mining to explore associations among the visual cues, classifies the associations by assigning each of them with a class label, and uses their appearances in the video to construct video indices. Our experimental results demonstrate the performance of the proposed approach.

Keywords-

Histogram; image; video; multimedia; mining

I. INTRODUCTION

The amount of audio-visual data currently accessible is staggering; everyday, documents, presentations, homemade videos, motion pictures and television programs augment this ever-expanding pool of information. Recently, Americans consumed information for about 1.3 trillion hours, an average of almost 12 hours per day. Consumption totaled 3.6 zettabytes and 10,845 trillion words, corresponding to 100,500 words and 34 gigabytes for an average person on an averageday. A zettabyte is 10 to the 21st power bytes, a million million gigabytes. These estimates are from an analysis of more than 20 different sources of information, from very old (newspapers and books) to very new (portable computer games, satellite radio, and Internet video).With digital technology becoming inexpensive and popular, there has been a tremendous increase in the availability of this audio-visual information through cable and the Internet. In particular, services such as video on demand allow the end users to interactively search for content of their interest. However, to be useful, such a service requires an intuitive organization of data available. Although some of the data is labeled at the time of production, an enormous portion remains un-indexed.

Detailed annotation is required so that users can quickly locate clips of interest without having to gothrough entire databases. With appropriate indexing, the users could extracting relevant content and navigate effectively in large amounts of available data. Thus, there is great incentive for developing automated techniques for indexing and

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 05, May 2015
Available at http://internationaljournalofresearch.org

organizing audio-visual data, and for developing efficient tools for browsing and retrieving contents of interest as well as security systems. Digital video is a rich medium compared to text material. It is usually accompanied by other information sources such as speech, music and closed captions. Therefore, it is important to fuse this heterogeneous information intelligently to fulfill the users' search queries.

Conventionally, the data is often indexed and retrieved by directly matching homogeneous types of data. Multimedia data, however, also contains important information related to the interaction between heterogeneous types of data, such as video and sound, a fact confirmed through human experience.

We often observe that a scene may not evoke the same response of horror or sympathy, if the accompanying sound is muted. Conventional methods fail to utilize these relationships since heterogonous data types cannot be compared directly. The challenge is to develop sophisticated techniques that fully utilize the rich source of information contained in multimedia data.

II. SEMANTIC INTERPRETATION OF VIDEOS

We believe that the categorization of videos can beachieved
by exploring the concepts and meaningsofthe videos. This task requires bridging the gap between low-level contents and high level concepts. Once a relationship is developed between the computable features of the video and its semantics, the user would be allowed to navigate through videos by ideas instead of the rigid approach of content matching. However, this relationship must follow the norms of human perception and abide by the rules that are most often adhered to by the creators (directors) of these videos. These rules are generally known as *Film Grammar* in video production literature. Likeany natural language, this grammar also has several dialects, but is fortunately, more or less universal.

The interpretation of concepts using this grammar first requires the extraction of appropriate features. Secondly, these features or *symbols* need to be semiotically (symbolic as opposed to semantic) explored as in natural languages. However, the interpretation of these symbols must comply with the governing rules for video-making of a particular genre. An important aspect of this approach is to find a suitable mapping between low-level video features and their bottom-line semantics. These steps can be summarized as:

- Learn the video making techniques used by the directors. These techniques are also called *Film Grammar*.
- Learn the theories and practices of film aesthetics, such as the effect of color on the mood, the effect of music on the scene situation and the effect of post processing of the audio and video on human perception.
- Develop a model to integrate this information to explore concepts.
- Provide users with a facility to navigate through the audiovisual data in terms of concepts and ideas.

This framework is represented in Fig.1. Moreover, the video can be indexed as *Host-shots* and *Guest-shots*.

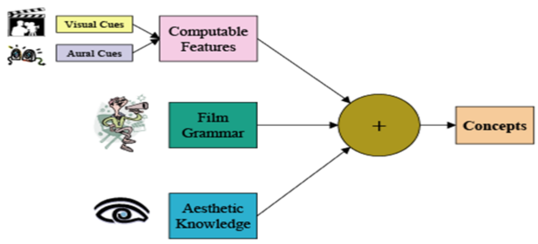

Fig.1: Semantic Interpretation of Videos

III.ALGORITHM

Organizations with large digital assets have needed to retrieve meaningful information from their digital collection. Applications such as digital libraries, video on demand systems, and interactive video application introduce new challenges in managing large collections of audio visual contents. To help users find and retrieve relevant video more effectively and to facilitate new and better ways of entertainment , advanced technology must be developed for indexing filtering , searching and mining the vast amount of videos. Motivated by these demands, many video research efforts have been made on exploring more efficient content management system. A simple framework is to partition continuous video frames into discrete physical shots and extract low level features from video shots to support activities like searching, indexing. Unfortunately, a single shot which is separate from its context has less capability conveying semantics.

In our approach, we exploited domain knowledge and used film grammar for video segmentation we were able to distinguish between the shots of host and guest by analyzing the shottransition. We also studied the cinematic principle used by the movie directors and mapped low level features, as the intensity histogram, to high level semantics, such as movie genre. Thus, we have provided an automatic method of video content annotation which is crucial for efficient media access.

To support the concept of video mining, we have proposed an algorithm using the concept of film grammar and the concept of video clustering and classification.For the sake of simplicity, we have classified the complete procedure in two major classes.

A. Primary Data Mining
B. Extended Data Mining

First we input our video data for the primary data mining, where the continuous video is divided into the frames and each frame is classified according to its difference to their previous frames. Then, the output of primary data mining enter as the input of extended data mining where more exploratory information is retrieved for better classification and to support high level searching. Fig.2 summarizes the steps of our algorithm.

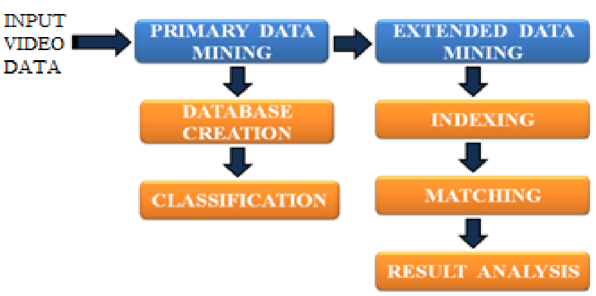

Fig.2: Introduction to Algorithm

Our input video should be in AVI format. "AVI" stands for the audio/visual interleave format and is a common animation format. We can get avi video through video cameras which are compatible with this format(Currently most of the video cameras are compatible with mp4 format due to less file size in comparison to AVI file).

The main motive of primary data mining is to convert the continuous video into discrete frames. These frames stored in a predefined location for further classification and analysis.

The process of primary data mining can be subdividedinto
twosteps:

- Database Creation
  Step1-Select the video
  Step2- Convert it into frames

Step3- Save these frames into JPEG format in predefined
location.
Step4- Commit the process and exit

Number of frames

[No. of frames=(Frame rate of video)*(Time duration)]

- Classification is done into two steps:
  1. Frame Classification
  2. Video Classification

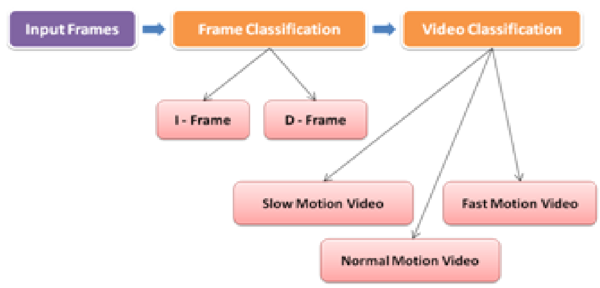

Fig.3: Classification

The frames are classified on the basis of their differences to their previous frames. For the classification of the frame we evaluate distance of the frames with their previous frames. This distance shows the difference with their previous frames. We can use the Bhattacharya Distance evaluation techniques.

Bhattacharya Distance

In statistics, the **Bhattacharya Distance** measures the similarity of two discrete probabilities distributions.Itis usuallyused to measure theseparabilityof classes in classification .for discrete probabilities distributions p and q over the same domain X; it is define as :

$$BC(p,q)=\sum\sqrt{p(x)q(x)}$$

The Bhattacharya coefficient is a divergence type measure, it can be seen as the scalar product of the two vectors (one for p and one for q) having as component the square root of the probability of the points x € X. it thereby lends itself to a geometric interpretation; the Bhattacharya coefficient is the cosine of the angle enclosed between these two vectors.On the basis of these distance, we can classify each frame on following two classes:

1. I-frame
2. D-frame

I-frame are also known as intra-frames. An I-frameis encoded as a single image, with on reference to any past or future frames. The first frame is always an I-frame. Number of I-frames found in the video plays important role to classify the videos. The information of I-frames are directly stored into database**.** The information of I-frame cannot retrieve from the information of other I frames.

D-frames are also known as dependent frames. A D-frame isencodedrelative to the past reference frames. The information of the D-frames can be retrieved through the information of the I-frames.Therefore, to store the information of the D-frames in database is always optional. The information of D-frames are also neglected while classification of videos.

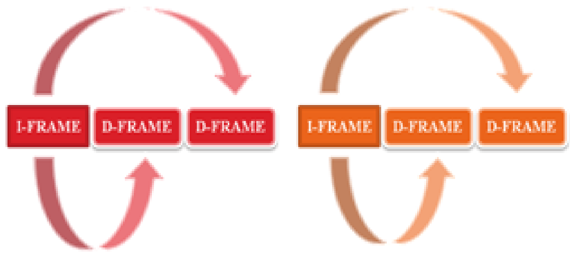

Fig.4: I-frames and D-frames

Video classification is done on the basis of number of I-frames are retrieved in frame classification. The ratio of number of I-frames and total frames signifies the change percentage in the given video. This ratio also signifies the notion of the video. To classify the video, we have to determine the thresholdsfor each class. Using this threshold, we can classify the video into these primary classes:

1. Slow Motion video
2. Normal Motion video
3. Fast Motion video

Slow motion videos are those videos where the number of I-framesisverylow. Underthisclass,videoslike Introductory Videos are classified where the scenes of the video do not change very rapidly.

In normal motion videos, the scenes of the videos change faster than slow motion video and slower than the fast motion video. The videos like Medical Videosareclassified under these classes, where the background does not change rapidly but the actions are changing frequently.

In fast motion videos, the scenes and action both changes rapidly. The videos like Action Movies are classified under these classes.

The output of the primary data mining is entered as input for extended data mining. The main motive of extended data mining is to explore more deep information so that the process of searching is database may become faster. Here, we use the conceptofindexing. The accuracy of searching result directly depends on the accuracy of indexing process.

Indexing extracts specific information from data and access data through it. Indexing can be classified into two classes:

1. Primary Key Indexing
2. Secondary Key Indexing

Primary key indexing is based on single attribute. No duplicates are allowed here. For supporting, good space utilization and good performance, the dynamic indexing is used along with primary key indexing. In dynamic indexing, the file grows or shrinks to adapt to the volume of data. The main methods of primary key indexing are:

1. B-Tress and variants(B+-Trees, B*-Trees)

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 05, May 2015
Available at http://internationaljournalofresearch.org

2. Hashing and variants(Linear hashing, Spiral etc)

Hashing is faster while B-Trees preserve the order of trees.

Secondary key indexing is mainly used in Multimedia. In secondary key indexing, signals are represented by feature vectors. Feature extraction computes Feature Vector from Signals. Then organizes the feature space so that it can answer on any attribute.

Our algorithm uses the concept of secondary data mining. The indexing of a frame can be done according to Fig.5:

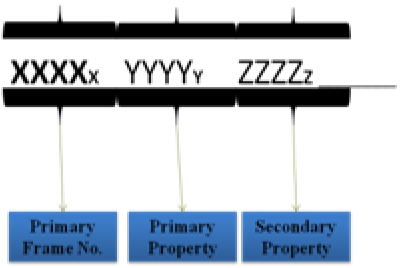

Fig.5: Indexing

The property and attribute selection is the case specific. For selection of attributes, we should have depth knowledge of the area to which it relates. For example in caseofFace Detection system, we should have knowledge about Biometric properties of the human faces.

Weuse tree to store the indexes, where the roots denote the I-frames and leaves denote D-frame. The height of the tree is always two.

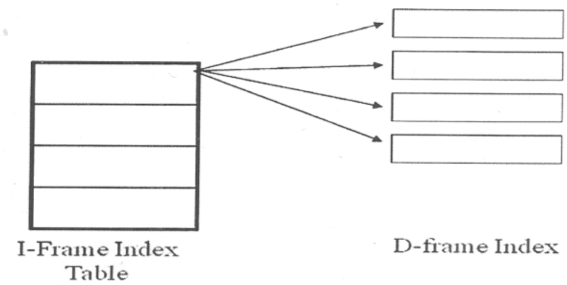

Fig.6: Data Structure for Indexing

For retrieval there are two step :

1. Hypothesis search through the index returns all qualifying documents plus some false alarms.
2. Verificationtheanswer is examined to eliminate false alarms.

The process of matching can be summarized in the following steps:

Step1-Evaluate the properties of the given object image.

Step2-Evaluate the thresholds for identity Index matching and dependent Index matching.

Step3-Search the frames in the given database.

Step4-Calculate the time slot of the found Frames.

Timeslot=[Primary Index Of frames/Total Frames]*Total Duration Of Video

IV. CONCLUSION

We have provided an automatic method of video content annotation which is crucial for efficient media access. We are successful to code Primary Video Data Mining algorithm which is working with high efficiency and the results obtained are meeting the accuracy. We have converted the 'avi' file into no of frames with JPEG format and maintained a database for future application like Matching, Searching. We are also able to distinguish that the particular video is whether slow motion, fast motion (action) or normal motion, but we are failed to code for extended Video Data Mining due to unavailability of Video tools in Matlab. We are currently trying to find out the other optional tools to overcome from this limitation but we are very glad to say that our Primary Classification algorithm is working with High Accuracy.

ACKNOWLEDGEMENT

Before we get into thick of the things we would like to add a few hearty words for the people who were part of this project in numerous ways, people who gave unending support right from the stage the project idea was conceived. It's our pleasure to take this opportunity to thank all those who help us directly or indirectly in preparation of this paper.

REFERENCES

[1]Perner, P. (2002). Data mining on multimedia data.Springer

[2] JungHwan Oh, JeongKyu Lee and Sae Hwang "Video Data Mining",University of Texas,USA.

[3]Till Quack, Vittorio Ferrari, and Luc Van Gool "VideoMining with Frequent Itemset Configurations".

[4]Data Mining: Foundations and Intelligent Paradigms. Springer

[5]JoseLuis PATINO, Etienne CORVEE, Francois BREMOND and Monique THONNAT "Data mining for activity extractionin video data"