Video Classification using SVM

Ashwini S. Mane, P. M. Kamde

Abstract— In today's world, information is indispensable in each and every activity. Precise retrieval of information according to the user's requirement is the dire need of the day. A Content Based Video Retrieval System, in a nutshell, aims at assisting a user to retrieve a video sequence target within a potentially large database. Content-based Video Retrieval Systems (CBVRS) are less common and is even now a research area. There are no existing systems of CBVRS in use because of various restrictions like video size, characteristics, high error rate etc. Search engines like Google etc use textual annotations to retrieve videos for the user which has very high error rate. Content based video retrieval has lots of applications in varied areas like medical sciences, news broadcasting, advertizing, video archiving and will surely revolutionize the field of information technology. A consequence of the growing consumer demand for visual information is sophisticated technology that is needed for representing, modeling, indexing and retrieving multimedia data. In particular, we need robust techniques to index/retrieve and compress visual information. Therefore this system Video Classification using SVM will play an important role in information retrieval and information storage.

Index Terms— Direct Frame Difference, FFT , Mean and Standard Deviation ,Support Vector Machine

I. INTRODUCTION

Video classification plays an important role in content based video retrieval. The number of systems are designed to classify videos according to their contents but each of these system differs by the techniques used to do the classification. In some systems video are classified into shots and scenes using shot boundary detection technique. There are number of methods used to detect the shot boundary such as Cut, Fading, dissolve, Wavelet Transform etc. Key frames will be extracted from shots by setting the time interval for image, for instance in a shot the frame appearing after every 5 seconds will be taken as a key frame. Thus there are number of methods used for key frames extraction but very less systems use the motion difference between two frames to extract key frames as used in our video classification system. Most of the systems use only visual features such as colour, text but these features some time may provide incorrect results therefore there is a need to extract more features and integrate them for robust classification.

For Classification number of systems use rule generation methodology ie system generates Fuzzy rules to define different classes but the rule generation methodology work efficiently when number of features extracted are more. In this case if Support vector machine is used then classification will be done efficiently and it will give correct results if number of classes are more.

Therefore in our system SVM is used. Video is the technology of electronically capturing, recording, processing, storing, transmitting, and reconstructing a sequence of still images representing scenes in motion to be very precise. Information Retrieval (IR) is the science of searching for documents and for information within documents. Similarly, Content-based Image Retrieval (CBIR), also known as query by image content (QBIC) is all about the searching of digital images in large databases. "Content based" means that the search will analyze the actual contents of the images. The term "Content" in this context might refer to colours, shapes, textures etc. Content-based Image Retrieval Systems work either through query by example technique or by using Semantic Retrieval. In query by example, a sample image is used for querying. The system retrieves images matching the query image based on characteristics like colour, shape, etc. Gaining accuracy in this method is difficult than in content based analysis. In Semantic Retrieval, the search uses textual information attached to the image to retrieve results. This method though easy is tedious and time consuming as attaching textual descriptions to all images is a very big process. Increased use of multimedia has resulted in paving the ways for the need of Concept based Video Retrieval Systems as well. However, though similar to CBIR, our system is more complex due to the presence of temporal information.

The main factors to be considered in our system are accuracy, speed and consistency. Concept based video retrieval uses the visual contents of a video such as colour, shape, texture, motion or audio to represent and index video. In our system, the visual contents of the videos are extracted and described by feature vectors. The feature vectors of the videos in the database form a feature database. To retrieve videos, users provide the retrieval system with example videos. The system then changes these examples into its internal representation of feature vectors. The similarities between the feature vectors of the query example and those in the database are then calculated and retrieval is performed with the aid of an indexing scheme.

II. SYSTEM ARCHITECTURE

Manuscript Received July, 2013.

Ashwini S. Mane, Computer Engineering Department ,Sinhgad College of Engineering, Pune University ,Pune ,India.

Prof. P.M.Kamde, Computer Engineering Department ,Sinhgad College of Engineering, Pune University ,Pune ,India.

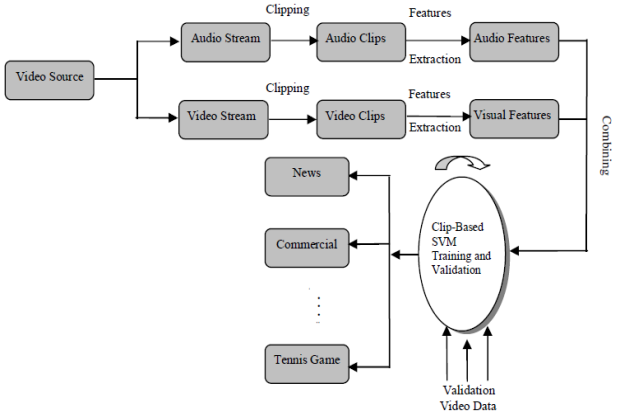


FIG.1 BLOCK DIAGRAM OF VIDEO CLASSIFICATION SYSTEM[8]

The main scope of the system is to classify the videos according to their classes. Prior to the Video classification, the user will have to create database for two purposes first database part is for training. In training phase features are extracted and then user will classify them by providing overlay tag to key frames. Training is the learning phase of the system. And other part of the database is used for testing phase, here key frames features tested and Support Vector Machine will classify them according to their class. In video classification system user will provide a binary image for the feature extraction. This feature extraction system will process the given image and it will extract the features like Motion, Audio and visual features. According to these values feature vectors are created. Lastly classification will take place. After the feature extraction of the given image those feature values will be compared with the trained database feature values which will decide the class of the objects recognized in the gray scale input image. As per the extracted features of the given input it will be classified into the respective classes.

III. ALGORITHMS

1 Algorithm for Training

Let Training Video Set = Vt

where ,

Vt ={V1,V2,V3,…, Vn }
t=0;

for i=1:n extract F ={Fi1,Fi2,..., Fik}
where Fik is set of frames from Vi

for j=2:k
D=|Fij − Fij − 1|;
if(D > TH)
Key[t++]= Fij;
end
end

for j=1: t
X(j)=Feature(Key[j]);
Tag(j)=USER GIVEN CLASS TAG FOR VIDEO KEY;
end
End

2 Algorithm for Testing

Let Test Video = Vin
t=0;
extract F={Fi1,Fi2...Fik}
where Fik is set of frames from Vin

for j=2:k
D=|Fij − Fij − 1|;

if(D > TH)
Key[t++]= Fij;

end
for j = 1 : t
X (j) =Feature (Key[j]);
Tag (j) =Classify (X(j),KNOWLEDGEBASE);
End

3. Algorithm for calculating Mean and Standard Deviation

1. Calculate the mean, x.
2. Write a table that subtracts the mean from each observed value.
3. Square each of the differences.
4. Add this column.
5. Divide by n -1 where n is the number of items in the sample This is the *variance*.
6. To get the *standard deviation* we take the square root of the variance.

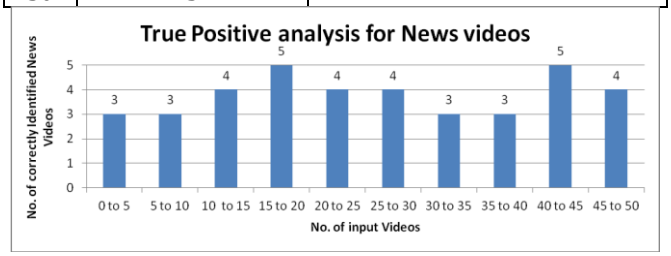
4. SVM Classification Algorithm

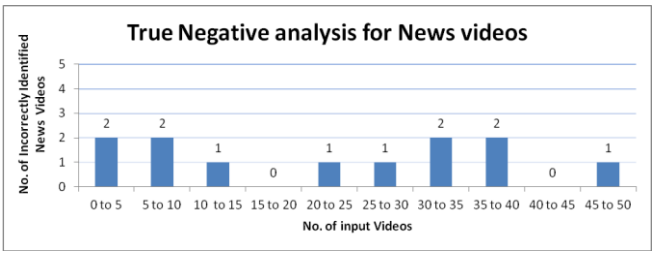
1. The input to the SVM is a vector which consists of Rmean, Gmean, Bmean, Rstd, Gstd, Bstd, FFTmean, FFTstd, class label
2. SVM multiplies this feature vector with weight function and compares this value with trained feature vector if it is similar to nearest value then class will be identified .

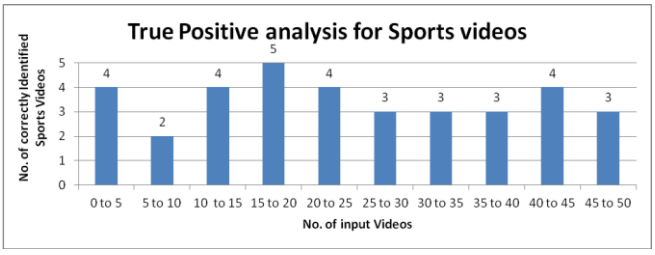
IV. RESULTS

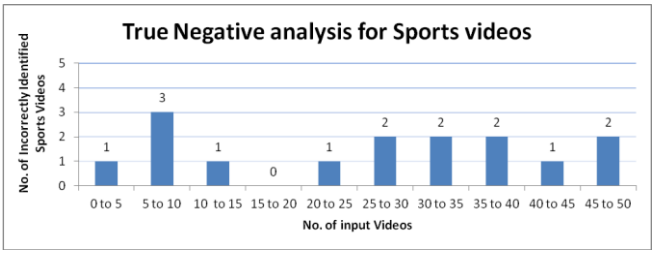
Data Set for News Video		
Sr. No.	**No. of Correctly Identified**	**No. of InCorrectly Identified**
5	3	2
10	3	2
15	4	1
20	5	0
25	4	1
30	4	1
35	3	2
40	3	2
45	5	0
50	4	1

Data Set for Sports Video		
Sr. No	**No. of Correctly Identified**	**No. of InCorrectly Identified**
5	4	1
10	2	3
15	4	1
20	5	0
25	4	1
30	3	2
35	3	2
40	3	2
45	4	1
50	3	2









V.CONCLUSION

This Video Classification System presents a new framework that integrates audio and visual features for content based classification of video using Support Vector Machine. The Video Classification System provides correct video classification as three features such as motion , audio and visual features are extracted.

This system will provide good results as compared to other system which uses only visual features and text extraction.

Support Vector Machine provides a good classification and uses "Winner Takes it All" algorithm to classify any classes.

REFERENCES

1. L. Chaisorn, et. al.,"The segmentation of news video into story units",in Proc. of IEEE Int. Conf.on Multimedia and Expo (ICME), vol. 1, pp.73 - 76, 2002.
2. M.R. Naphade et. al.,"A probabilistic framework for semantic video indexing, filtering, and retrieval",IEEE Transactions on Multimedia, vol. 3(1),pp. 141-151, 2001.
3. G. Xinbo et. al.,"A graph-theoretical clustering based anchorperson shot detection for news video indexing",in Proc IEEE Int. Conf On Computational Intelligence and Multimedia Applications (ICCIMA),pp 108 - 113, 2003.
4. Stefan Eickeler et. al.,"Content-based video indexing of tv broadcast news using hidden markov models", in Proc. IEEE ICASSP, 1999, pp. 2997-3000.
5. Weiming Hu et. al.,"A Survey on Visual Content-Based Video Indexing and Retrieval",ieee transactions on systems, man, and cybernetics part c, applications and reviews, vol. 41, no. 6, november 2011.
6. Vincent S. Tseng et. al,"Integrated Mining of Visual Features,Speech Features, and Frequent Patterns for Semantic Video Annotation",ieee transactions on multimedia, vol. 10, no. 2, february 2008.
7. Liuhong Liang et. al.,"Enhanced Shot Boundary Detection Using Video Text Information", IEEE Transactions on Consumer Electronics, Vol. 51,No. 2, MAY 2005.
8. Yingying Zhu et. al," Video Scene classification and Segmentation based on Support Vector Machine",978-1-4244-1821-3/08.
9. Wan Jianping et. al.,"News Video Story Segmentation Based on Nave Bayes Model",978-0-7695-3736-8/09 25.00 2009 IEEE.
10. X. Li et. al., "Content-Based News Video Mining",ADMA 2004, LNAI 3584, pp. 431438, 2005.Springer-Verlag Berlin Heidelberg 2005 .