ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

Improvement in Score of Classifier by Correlation between Concepts for Semantic Concept Detection

Priya Kadam[1], Nita Patil[2], Sudhir Sawarkar[3]

P.G. Student, Department of Computer Engineering, Datta Meghe college of Engineering College, Navi Mumbai, Maharashtra, India[1]

Assistant Professor, Department of Computer Engineering, Datta Meghe college of Engineering College, Navi Mumbai, Maharashtra, India[2]

Principal, Department of Computer Engineering, Datta Meghe college of Engineering College, Navi Mumbai, Maharashtra, India[3]

ABSTRACT: The multimedia storage is increasing day by day. Lots of videos available in the video warehouse are in unstructured format. As per user requirement, it is difficult to retrieve the relevant videos from such a huge video storage. Nowadays it is important to make such unstructured multimedia data easily available with flexibility. The low-level features are insufficient to explain the content properly at conceptual level. The semantic gap characterizes the difference between two descriptions of an object by different linguistic representations for instance languages or symbols. The semantic gap can be defined as "the difference in meaning between constructs formed within different representation systems". This “semantic gap” is basic problem in content-based multimedia retrieval system.

The main aim is to form semantic representations by extracting intermediate semantic levels (events, objects, locations, people, etc.) from low-level visual and audio features by using machine learning algorithms.

Here, we mentioned the description of our proposed system and its methodology for implementation of semantic concept detection. The main aim of this system is to improve the accuracy of concept detection and also to reduce the semantic gap. At initial step, we have performed the shot segmentation and key frame extraction. Then by selecting best key frame, we performed the feature extraction. For key frame extraction we used XM tool with help of which we can get the low level features of our key frames. Here, SVM performs the nonlinear classification so as to provide the best result. Moreover, correlation is used to find out the relationship between the concepts. The results obtained with precision, recall and F-score using mpeg features by SVM classifier on TRECVID dataset. The precision, recall and fscore using Mpeg features are 0.7494, 0.6243 and 0.7154 respectively. The Confusion matrix using Mpeg features and SVM classifier on TRECVID dataset is shows the average value for actual class and predicted class is 0.75 and 0.64 respectively.

The Confusion matrix using Correlation between the concepts on TRECVID dataset is shows the average value for precision and recall are 0.80 and 0.75 respectively.The result comparison shows that accuracy of the classifier is improved by using correlation between the concepts.

KEYWORDS: Shot detection; Key frame extraction; Support vector machine; Concept Detection; Concept Correlation.

I. INTRODUCTION

Lots of research is going on as amount of audiovisual content is increasing day by day and nowadays this is interest area for many researchers. The focus set mainly on the extraction of various low-level features i.e. audio, color, texture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

and shape properties of audiovisual content. Lot of new techniques such as neural networks, fuzzy logic systems and Support Vector Machines (SVM) used to find out high-level features from low level features.
Here, purpose is to develop a system for concept detection with better accuracy. General framework for video concept detection is shown in figure 1.Shot detection is the initial step in concept detection framework. It used to detect transition between successive shots. Key frame extraction gives the summary to retrieve video. A video shot is represented by a single image called as “key frame” and features are extracted locally and globally i.e. feature extraction is performed. High-level concept detectors are trained using a global key frame annotation. And Concept detection is performed.

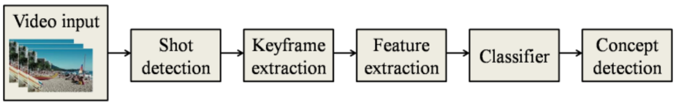

Fig.1. Semantic Concept detection framework

II. RELATED WORK

The huge research work is already done for analysis of video composition.
Alan Hanjalic et al. [1] developed a statistical shot-boundary detector. This detector fuses the range information and a priori information which gives an adaptive threshold as a result and provides optimal detection performance. Mr. Hattarge A.M. et al. [2] explains an adaptive approach that combines different features of video and calculates a unique feature which is used to identify the hard cuts and gradual transitions.Zeeshan Rasheed et al. [3] represented a high-level segmentation of videos into scenes. To cluster shots into scene, a weighted undirected graph also called as a shot similarity graph (SSG) is used. Janko Calic et al. [4] proposed novel key-frame extraction technique based on the difference metrics extracted directly from the MPEG compressed domain and discrete contour evolution. Pascal Kelm et al. [5] compared a key frame extraction approach against the key frame extraction of IBM Multimedia Analysis and Retrieval System. Andreas Girgensohn et al. [6] hierarchical clustering algorithm is used. To summarize videos and provide access points to key frames, the key frames used to recognize videos from each other.

Sylvie Jeannin et al. [7] in this paper, the main aim is to provide brief descriptors which are easy to extract and match. Also,how motion activity and motion trajectory both fulfil their aim is mentioned.Jens-Rainer Ohm et al. [8] mentioned a set of color descriptors which are able to capture the important aspects of color feature and allow color similarity computations. All these descriptors are compact and allow efficient description of color properties.Evaggelos Spyrou et al. [9] in his work implemented low level feature extraction, color features and texture features extraction. For color property, MPEG-7 dominant color descriptor is used. For the texture properties the actual MPEG-7 homogeneous texture descriptor is used. For region thesaurus, Subtractive clustering is used. After this a model vector is formed. The SVM are supervised learning models having associative learning algorithm which can perform data analysis for classification and regression analysis purpose.

Evaggelos Spyrou, et al. [10] in his work for concept detection used RSST segmentation algorithm for low level feature extraction. Moreover, for the training set of images a K-means clustering algorithm is used on the feature vectors of the regions. A model vector is formed for the representation of the semantics of a key frame which are based on the set of region types, a model vector is constructed. For each high level concept, a separate neural network based detector is trained.. Lin Lin et al. [11] here, in this work associative classification algorithm is proposed by using MCA for video concept detection. MCA is used to measure the correlation between various feature value pairs and classes to conclude the high-level concepts from the extracted low-level features. Nita S.Patil et al. [12] proposed a framework for multi-label semantic concept detection. In this framework classifiers trained on global and deep features.
Ken-Hao Lie et al. [13] in his paper to find out the relationship between targeted concepts and remaining other semantic concepts, MCA i.e. .multiple correspondence analysis is used. These relationships are utilized as transaction weight and to refine detection ranking scores.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

III. PROPOSED ALGORITHM

Figure 2 shows framework for proposed concept detection based on Mpeg features. The proposed method has following steps. Initial step is shot segmentation and key frame extraction from video. The next step is low level feature extraction to produce model vector for classification. The next step is classification using SVM. Next step is to find correlation between the concepts. Correlation between concepts is used to improve the detection scores of classifier.

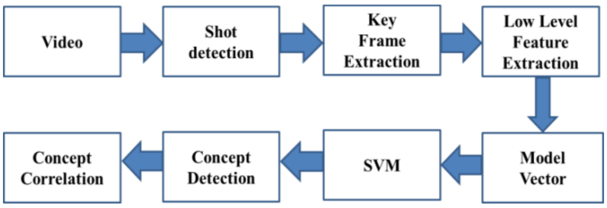

Fig. 2.Proposed Semantic Concept Detection using Mpeg features

A. Shot segmentation: The shot segmentation is the first step of key frame extraction. In this transition between the successive shots are detected. The video shot segmentation has two categories: one is compressed domain and another is uncompressed domain. Here for our proposed system we are working in compressed domain. In our system a color histogram method is used to segment the shots according to the frame difference. The Histogram-based method is the most common used method to calculate frame difference. Since color histograms do not relate spatial information with the pixels of a given color and only records the amount of color information, images with similar color histograms can have dramatically different appearances. To solve the problem, an improved histogram algorithm, x^2 histogram matching method is adopted. The color histogram difference $d(I_i, I_j)$ between two consecutive frames I_i and I_j can be calculated as follows:

$$d(i_{i,}i_j) = \sum_{k=1}^{n} \frac{(H_{ij}-H_{jk})2}{H_{jk}+H_{ik}},(H_{jk}\neq 0) \qquad (1)$$

Where H_i and H_j stand for the histogram of I_i and I_j, respectively.

A shot transition occurs when $d(I_i, I_j)$ is bigger than a given threshold. Selecting an appropriate threshold is the key to the method.

B. Key Frame Extraction: In key frame extraction, every image is compared with its consecutive image. Comparison takes place by using absdif ().In this function, images are converted from rgb to gray and then their histograms are calculated and compared. The absdif () returns some values which we store in array and which is used to calculate the threshold value. Again all frames are compared with its consecutive frames along with condition that wherever absdif () returns the value greater than threshold value that frames will be written in folder as keyframes.

Fig. 3. Extraction of shots and keyframes

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

C. Low-Level Feature Extraction: The MPEG-7 descriptors are selected to find out the low level features of each region. Following are the MPEG7 color descriptors which are extracted to find out the color characteristics:

1. **Dominant color descriptor**
   As the name indicates, dominant color descriptor allows specification of small number of dominant color values and its characteristics like variance and distribution over a region. The main aim is to provide a powerful, dense and instinctive representation of colors present in an image.

2. **Color structure descriptor**
   The color structure descriptor is based on color histogram. It uses small structuring window to find out the local color distribution. This descriptor is attached to the Hue-Min-Max-Difference color space .i.e. HMMD color space.

3. **Color layout descriptor**
   The Color layout descriptor holds the spatial layout of the dominant colors on a grid which is superimposed on an image or on region. Its representation is based on coefficient of DCT (Discrete cosine transform).This is a very compact and efficient descriptor and highly used in application of fast browsing and search application. It can be applied on still images as well as on video segments.

4. **Scalable color descriptor**
   The Scalable color descriptor is obtained from a color histogram. This histogram is defined in HSV i.e. Hue–Saturation-Value color space. It uses Haar transform coefficient encoding which allows scalable representation of description and scalability of feature extraction and matching procedure.

Following are the MPEG7 texture descriptors used to extract the texture properties :

1. **Homogeneous texture descriptor**
   The HTD gives quantitative characterization of texture for similarity based image to image matching. This descriptor is computed by first filtering the image with bank orientation. This descriptor is computed by first filtering the image with a bank orientation and scale sensitive filters .Then computes the mean and the standard deviation of output in frequency domain.

2. **Edge histogram descriptor**
   The edge histogram descriptor records the spatial distribution of edges. The distribution of edges is a good texture sign which is useful for image matching.

D. Experimentation Model (XM)**:** The purpose of an Experimentation Model (XM) within MPEG-7 is to specify and implement the feature extraction, encoding and decoding algorithms as well as search engines.

MPEG-7 Low Level Feature Extraction - Command Line Tool: It is a program to extract MPEG 7 low level color and texture descriptor from whole image using MPEG 7 low level feature extraction library. The tool uses OpenCV 2.The archives contains the executable (MPEG&Fex.exe), OpenCV2.0 DLLs ,Readme.txt(It explains how to use program),sample images and sample outputs. As an example, to extract Color Structure Descriptor (CSD) of size 64 from a set of images, just type MPEG7Fex.exe CSD 64 imageList.txt CSD.txt on the command line and the descriptors will be written to CSD.txt file, one line per image. Extraction will be faster if the image sizes are constant (e.g., video key frames).

E. Support vector machine: Classification of the given data is a common task in machine learning. In a classification problem, the learner approximates a function which can map a given vector data into one of the various class labels. In supervised setting, it is done by looking at a set of input-output examples of the function. The finite input-output example data which is used for learning the classification function is called the training data.

Support Vector Machines (SVM) is one of the successful supervised learning methods for concept detection. Support Vector Machines are trained so that the decision function would classify the unseen example data accurately. This ability to classify unseen example data accurately is referred to as generalization. High generalization capability is one of the main reasons for the success of SVMs. Given a set of d-dimensional vectors, a linear classifier tries to separate them with a d-dimensional hyperplane. There are many hyperplanes that might classify the data. If we define “margin”

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

as the distance between the nearest samples on both sides of the hyperplane, SVMs are designed to choose the hyperplane that has the largest margin between the two classes. If such a hyperplane exists, it is known as the maximum-margin hyperplane and the linear classifier it defines is known as a maximum margin classifier.In order to use much more general decision surfaces, the input data {x1,….xn}∈X is transformed into a high-dimensional feature space, using a non-linear map $\varphi: R^d \rightarrow F$. A. A linear classifier is then found in the transformed high-dimensional space. The only requirement of F is that dot product can be defined in that space. It can be an infinite-dimensional space and no assumptions are made on the dimensionality of F. For a given training data set, SVM is now constructed in F instead of R^d i.e., using the set of examples

F.Correlation between the concepts: Concept correlation refers to the relations among concepts within a video shot. By using this information it is possible to refine the predictions. Once features have been determined, correlation is used to find the matching of image features or to find image motion. Correlation between concepts is determined to improve the detection scores of classifier. Here, in this system to obtain the correlation between the concepts, Confusion matrix using Mpeg features is thoroughly studied. In that every predicted concept class of training data set is tried to match with actual concept class. And wherever the observation found True positive (TP) that value is added to actual concept class. This has positive effect on score of classifier. To find correlation, the performance is evaluated using the same performance evaluation metrics like precision, recall and f-score.

G. Dataset description**:** The work includes benchmark dataset TRECVID 2007 from video domain.TRECVID's 2007 Video dataset and ground-truth data: - The National Institute of Standards and Technology (NIST) is responsible for the annual Text Retrieval Conference (TREC) Video Retrieval Evaluation (TRECVID). Every year, it provides a test collection of video datasets along with a task list. It focuses its efforts to promote progress in video analysis and retrieval. It also provides ground-truth for researchers. Fig.4 shows sample images of 14 concepts used in this experiment.

Fig. 4.Sample images of 14 concepts used in experiment

H. Evaluation metrics**:** The benchmark metric for classifier evaluation includes classification precision, classification recall and F-measure. They are defined as:

Precision: Precision defines the ratio of the total number of the retrieved video or keyframes which matches the query to the sum of the relevant videos matching and not matching the query from the video database. The precision can be accurately expressed by equation (2).

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (2)$$

Recall: Recall defines the ratio of the total number of the retrieved video or keyframes which matches the query to the sum of the retrieved videos matching the query in the video database. Equation (3), defines the precise expression for the recall parameter.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (3)$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

F-measure: F-measure defines the harmonic mean, weighted average of the precision and the recall parameters. Since both measures are important, usually classifier with f measure is followed. F-measure is expressed by equation (4).

$$\text{F Score} = 2 \times \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

IV. SIMULATION RESULTS

1. **Feature set**: Features used are Color Layout Descriptor (CLD), Color Structure Descriptor (CSD), Dominant Color Descriptor (DCD), Edge Histogram Descriptor (EHD), Homogeneous Texture Descriptor (HTD), and Scalable Color Descriptor (SCD). For extraction MPEG-7 experimentation Model (XM) is used. Table I gives the dimensions of these features.

Table I. Mpeg Feature descriptors

Sr. No.	Mpeg Descriptor	Dimension
1	Color Layout Descriptor (CLD)	160
2	Color Structure Descriptor (CSD)	64
3	Dominant Color Descriptor (DCD)	32
4	Edge Histogram Descriptor (EHD)	80
5	Homogeneous Texture Descriptor (HTD)	102
6	Scalable Color Descriptor (SCD)	128
Total dimension size		566

2. Experimental Results and Analysis of Mpeg features

The performance is evaluated using the performance evaluation metrics like precision, recall and f-score. The results obtained are shown in table II with precision, recall and F-score using mpeg features by SVM classifier on TRECVID dataset. Figure 7shows the plot of precision, recall and F-Score, and Figure 8 shows the confusion matrix obtained using mpeg features on TRECVID dataset.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

Table II. Performance measures using Mpeg features

Class No./ Category	Precision	Recall	F-Score
Airplane	0.7143	0.6250	0.6667
Bus	0.4565	0.9545	0.6176
Charts	0.6111	0.8148	0.6984
Court	0.8649	0.4571	0.5981
Desert	0.9000	0.5000	0.6429
Explosion fire	0.8000	0.5517	0.6531
Flag-US	1.0000	0.1667	0.2857
Maps	0.5469	0.9211	0.6863
Mountain	0.9318	0.8367	0.8817
Natural Disaster	0.7273	0.5714	0.6400
Prisoner	1.0000	0.5833	0.7368
Snow	0.8393	0.9216	0.8785
Truck	0.8649	0.7619	0.8101
Weather	1.0000	0.0556	0.1053
Average Values	**0.7494**	**0.6243**	**0.7154**

3. Experimental Results for Correlation between the concepts

To obtain the correlation between the concepts, concept ontology and training dataset are thoroughly studied. For e.g. if we want to find the correlation between the desert and truck. Then we analyze the training dataset of truck with respect to desert concept. If both concept found in single keyframe then we marked it as "1" and hence in this way we can find the correlation between the concept.

	Airplane	Bus	Chart	Court	Desert	Exp-Fire	Flag US	Map	Mountain	Ndisaster	Prisnor	Snow	Truck	Wheather
Airplane	1	0	0	0	1	0	0	0	0	0	0	0	0	0
Bus	0	1	0	0	1	1	0	0	1	0	1	0	1	0
Chart	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Court	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Desert	1	1	0	0	1	0	0	0	1	0	0	0	1	0
Exp-Fire	0	1	0	0	0	1	0	0	0	0	0	0	0	0
Flag US	0	0	0	0	0	0	1	0	0	0	0	0	1	0
Map	0	0	1	0	0	0	0	1	0	0	0	0	0	1
Mountain	0	0	0	0	1	0	0	0	1	0	0	1	1	1
Ndisaster	0	0	0	0	0	0	0	0	0	1	0	1	1	0
Prisnor	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Snow	0	0	0	0	0	0	0	0	1	1	0	1	0	0
Truck	0	1	0	0	1	0	1	0	1	0	0	0	1	1
Wheather	0	0	0	0	0	0	0	1	1	0	0	0	1	1

Fig.5 Correlation table

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

Fig .6 shot76_106_RKF from Training dataset of TRECVID dataset

To improve score of classifier, correlation is used in our system. SVM classifier gives the top most matches. To improve the result we refine the result obtained by SVM classifier by selecting the second or third top most result. With the help of correlation table, it is possible to refine the result. Wherever we get the value "1" in correlation table that class will be taken into account for correlated class. Consider the image fig 6. This is the image taken from training dataset of TRECVID dataset. This image is classified as Desert image by SVM classifier. By using correlation table, this image is classified as Mountain and Bus image. The performance is evaluated using the same performance evaluation metrics like precision, recall and f-score. The results obtained are shown in table III with precision, recall and F-Score using correlation between concepts. Figure 9 shows the plot of precision, recall and F-Score and figure 10 shows the confusion matrix obtained using correlation between concepts.

Table III: Performance measures using correlation between the concepts

Class No./ Category	**Precision**	**Recall**	**F-Score**	**acc1**
Airplane	0.7143	0.6250	0.6667	97.8814
Bus	0.4565	0.9545	0.6176	94.3231
Charts	0.6111	0.8148	0.6984	95.8606
Court	0.8649	0.4571	0.5981	90.6725
Desert	1.0000	0.7778	0.8750	98.2533
Explosion fire	0.8000	0.5517	0.6531	96.2882
Flag-US	1.0000	0.5000	0.6667	99.3450
Maps	0.5738	0.9211	0.7071	93.6681
Mountain	0.9400	0.9592	0.9495	98.8739
Natural Disaster	0.7500	0.6429	0.6923	98.2533
Prisoner	1.0000	0.5833	0.7368	98.8864
Snow	0.9091	0.9804	0.9434	98.6813
Truck	0.9589	0.8333	0.8917	96.2719
Weather	1.0000	0.8889	0.9412	99.5604
Avg.Values	**0.7991**	**0.7493**	**0.7734**	**96.9157**

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

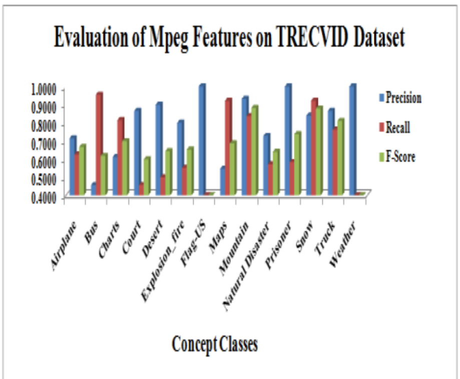

Fig.7. Evaluation of Mpeg features on TRECVID dataset

		Actual Class														
		Airplane	Bus	Charts	Court	Desert	Explosion Fire	Flag - US	Maps	Mountain	Natural Disaster	Prisnor	Snow	Truck	weather	
Predicted Class	Airplane	10 71.43 %	2 4.35 %	1 2.78 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	3 3.53 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.63 99.4
	Bus	0 0 %	21 45.65 %	1 2.78 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.95 99.0
	Charts	1 7.14 %	3 6.52 %	22 61.11 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	1 1.18 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.81 99.2
	Court	3 21.43 %	20 43.48 %	11 30.56 %	32 86.49 %	0 0 %	0 0 %	0 0 %	3 4.69 %	1 1.18 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.46 99.5
	Desert	0 0 %	0 0 %	0 0 %	1 2.7 %	18 90 %	0 0 %	0 0 %	6 9.4 %	8 9.4 %	0 0 %	0 0 %	1 1.8 %	2 2.7 %	0 0 %	0.50 99.5
	Explosion Fire	0 0 %	0 0 %	0 0 %	2 5.41 %	0 0 %	16 80 %	0 0 %	5 7.81 %	5 5.88 %	0 0 %	0 0 %	0 0 %	1 1.35 %	0 0 %	0.55 99.4
	Flag -US	0 0 %	0 0 %	0 0 %	2 5.41 %	0 0 %	0 0 %	1 100 %	0 0 %	1 1.18 %	0 0 %	0 0 %	0 0 %	2 2.7 %	0 0 %	0.17 99.8
	Maps	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	35 54.7 %	1 1.2 %	0 0 %	0 0 %	0 0 %	2 2.7 %	0 0 %	0.92 99.1
	Mountain	0 0 %	0 0 %	0 0 %	0 0 %	1 5 %	0 0 %	0 0 %	2 3.1 %	41 48.2 %	0 0 %	0 0 %	3 5.4 %	2 2.7 %	0 0 %	0.84 99.2
	Natural Disaster	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	1 5 %	0 0 %	0 0 %	4 4.71 %	8 72.73 %	0 0 %	1 1.79 %	0 0 %	0 0 %	0.57 99.4
	Prisnor	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	3 4.69 %	1 1.18 %	0 0 %	7 100 %	1 1.79 %	0 0 %	0 0 %	0.58 99.4
	Snow	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	1 1.56 %	3 3.53 %	0 0 %	0 0 %	47 83.93 %	0 0 %	0 0 %	0.92 99.1
	Truck	0 0 %	0 0 %	1 2.8 %	0 0 %	1 5 %	3 15 %	0 0 %	6 9.4 %	5 5.9 %	2 18.2 %	0 0 %	2 3.6 %	64 86.5 %	0 0 %	0.76 99.2
	weather	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	3 17.6 %	11 64.7 %	1 5.9 %	0 0 %	1 5.9 %	1 5.9 %	1 5.9 %	0.06 99.9
	Overall	71.4 28.6	45.7 54.3	61.1 38.9	86.5 13.5	90.0 10.0	80.0 20.0	100.0 0.0	54.7 45.3	48.2 51.8	72.7 27.3	100.0 0.0	83.9 16.1	86.5 13.5	100.0 0.0	0.62 0.75

Fig.8.Confusion matrix using Mpeg features and SVM classifier on TRECVID dataset

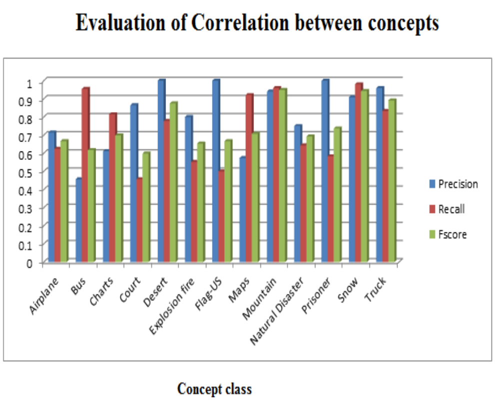

Fig.9. Evaluation of Correlation between concepts

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 9, September 2019

Predicted Class \ Actual Class	Airplane	Bus	Charts	Court	Desert	Explosion Fire	Flag - US	Maps	Mountain	Natural Disaster	Prisnor	Snow	Truck	weather	Overall
Airplane	**10** **71.43 %**	2 4.35 %	1 2.78 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	3 4.69 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.63 99.4
Bus	0 0 %	**21** **45.65 %**	1 2.78 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.95 99.0
Charts	1 7.14 %	3 6.52 %	**22** **61.11 %**	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	1 1.56 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.81 99.2
Court	3 21.43 %	20 43.48 %	11 30.56 %	**32** **86.49 %**	0 0 %	0 0 %	0 0 %	3 4.92 %	1 1.56 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.46 99.5
Desert	0 0 %	0 0 %	0 0 %	1 2.7 %	**28** 100 %	0 0 %	0 0 %	6 9.8 %	0 0 %	0 0 %	0 0 %	1 1.8 %	0 0 %	0 0 %	0.78 99.2
Explosion Fire	0 0 %	0 0 %	0 0 %	2 5.41 %	0 0 %	**16** **80 %**	0 0 %	5 8.2 %	5 7.81 %	0 0 %	0 0 %	0 0 %	1 1.37 %	0 0 %	0.55 99.4
Flag -US	0 0 %	0 0 %	0 0 %	2 5.41 %	0 0 %	0 0 %	**3** **100 %**	0 0 %	1 1.56 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.50 99.5
Maps	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	**35** **57.4 %**	1 1.6 %	0 0 %	0 0 %	0 0 %	2 2.7 %	0 0 %	0.92 99.1
Mountain	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	2 3.3 %	**47** **73.4 %**	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0.96 99.0
Natural Disaster	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	1 5 %	0 0 %	0 0 %	4 6.25 %	**9** **75 %**	0 0 %	0 0 %	0 0 %	0 0 %	0.64 99.4
Prisnor	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	3 4.92 %	1 1.56 %	0 0 %	**7** **100 %**	1 1.82 %	0 0 %	0 0 %	0.58 99.4
Snow	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	1 1.64 %	0 0 %	0 0 %	0 0 %	**50** **90.91 %**	0 0 %	0 0 %	0.98 99.0
Truck	0 0 %	0 0 %	1 2.8 %	0 0 %	0 0 %	3 15 %	0 0 %	6 9.8 %	0 0 %	2 16.7 %	0 0 %	2 3.6 %	**70** **95.9 %**	0 0 %	0.83 99.2
weather	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	0 0 %	1 5.9 %	0 0 %	1 5.9 %	0 0 %	**16** **94.1 %**	0.89 99.1
Overall	71.4 28.6	45.7 54.3	61.1 38.9	86.5 13.5	100.0 0.0	80.0 20.0	100.0 0.0	57.4 42.6	73.4 26.6	75.0 25.0	100.0 0.0	90.9 9.1	95.9 4.1	100.0 0.0	0.75 0.80

Fig. 10.Confusion matrix using correlation between the concepts on TRECVID dataset

V. CONCLUSION

The aim of our concept detection system is to improve the score of SVM classifier by using correlation between the concept. Here, working of our semantic concept detection system is explained along with its methodology and results obtained. At initial step, we have performed the shot segmentation and key frame extraction. For shot segmentation improved color histogram method is used. For Key frame extraction absolute frame difference image histogram of the consecutive frames is calculated. Then by selecting key frames, we performed the feature extraction. For feature extraction we used XM tool with help of which we can get the low level features of our key frames. Here, SVM performs the nonlinear classification so as to provide the best result. Moreover, correlation is used to find out the relationship between the concepts. The results evaluated with precision, recall and F-score using mpeg features by SVM classifier on TRECVID dataset. The results obtained by using MPEG feature extraction and results obtained by Correlation between the concepts also shown in result chapter. In this way we explaining the two layer concept detection. And it is also proved that semantic concept detection can be improved by using correlation between the concepts.

REFERENCES

1. Alan Hanjalic," Shot-Boundary Detection: Unraveled and Resolved?" IEEE transactions on circuits and systems for video technology, vol. 12, no. 2, February 2002.
2. Mr. Hattarge A.M., Prof. K.S. Thakre," Analysis and Review of Formal Approaches to Automatic Video Shot Boundary Detection", international Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 1, January 2014.
3. Zeeshan Rasheed and Mubarak Shah, "Detection and Representation of Scenes in Videos", IEEE transactions on multimedia, vol. 7, no. 6, December 2005.
4. Janko Calic and Ebroul Izquierdo," Efficient key-frame extraction and video analysis",2002.
5. Pascal Kelm, Sebastian Schmiedeke, and Thomas Sikora," Feature-based video key frame extraction for low quality video sequences",2009
6. Andreas Girgensohn John Boreczky Fx Palo Alto," Time-Constrained Keyframe Selection Technique", Multimedia Tools and Applications, 11, 347-358, 2000
7. Sylvie Jeannin and Ajay Divakaran ,"MPEG-7 Visual Motion Descriptors", IEEE transactions on circuits and systems for video technology, vol. 11, no. 6, June 2001.
8. Jens-Rainer Ohm,Leszek Cieplinski ,Heon Jun Kim ,Santhana Krishnamachari ,B. S. Manjunath ,Dean S. Messing ,Akio Yamada ," The MPEG-7 Color Descriptors".
9. Evaggelos Spyrou and Yannis Avrithis," High-Level Concept Detection in Video Using a Region Thesaurus", 2009.
10. Evaggelos Spyrou, Giorgos Tolias, and Yannis Avrithis, "Large Scale Concept Detection in Video Using a Region Thesaurus" Image, Video and Multimedia Systems Laboratory, School of Electrical and Computer Engineering National Technical University of Athens 9 Iroon Polytechniou Str., 157 80 Athens, Greece,2009.
11. Lin Lin, Mei-Ling Shyu, Guy Ravitz, Shu-Ching Chen" video semantic concept detection via associative classification"
12. Nita S.Patil, Sudhir Sawarkar ,"Semantic Concept Detection in Video Using Hybrid Model of CNN and SVM Classifiers", International Journal of Image Processing (IJIP), Volume (13) : Issue (2) : 2019
13. Ken-Hao Lie," Correlation based re-ranking for semantic concept detection"2014.