Prediction of Rivaled-Unrivaled Scene in Video Using SCA Algorithm

Mr. Ajay B Gadicha
Research Scholar, Department of Computer Science & Engg
P.R.Pote (Patil) College of Engineering & Management,
ajjugadicha@gmail.com

Dr. M.V.Sarode
Professor and Head, Department of Computer Engg
Government Polytechnic, Yavatmal.

Dr. V.M.Thakare
Professor and Head, P.G. Department of Computer Science
Sant Gadge Baba Amravati University, Amravati (MH)

Abstract: - ***Video scene classification and segmentation are fundamental steps for multimedia retrieval, indexing and browsing. In this paper, a robust scene classification and segmentation approach based on Support Vector Machine (SVM) is presented, which extracts both frame and scene features and analyzes their inter-relations to identify and classify video scenes. This system works on content from a diverse range of genres by allowing sets of features to be combined and compared automatically with comparative use of thresholds. With the temporal behaviors of different scene classes, SVM classifier can effectively classify pre-segmented video clips into one of the predefined scene classes. After identifying scene classes, the scene change boundary can be easily detected. Researchers have actively developed wonderful strategies to wise video processing management, collectively with shot transition detection, key frame extraction, video retrieval, and lots of others methods related to video processing. This paper merely focuses on Rivaled-Unrivaled Scene in Video Using SCA Algorithm which can takes input from the dataset and produce output is the trained algorithm which can accurately classify the input scenes also it generate the scene consequence more accurately with respective to other peer technique.***

Index:- SCA,SVM,SBD

I. INTRODUCTION:-

The rapid growth of multimedia technology has caused an exponential increase of multimedia digital data in recent years. There are an increasing number of audio visual materials available publicly (e.g., broadcast news, dramas, movies, sports video). Opportunities in viewing privately produced videos in public are not uncommon (e.g., YouTube). Explosion in the amount of multimedia data is causing a serious management problem, which needs to be addressed. In video processing, identification of high level semantic concepts has been a hot issue in recent years, further leading to studies such as video classification, summarization and retrieval [7, 13, 14]. Combination of text, audio and image analysis techniques, so-called multimodal processing, has also been investigated widely

Video analysis is a exceedingly research area and currently there is an enormous interest in analysis at various levels of complexity, ranging from optical flow and dynamic texture analysis to high level analysis in terms of actions, activities and localization of particular events in videos. While the target application in this paper is dynamic scene recognition, at the same time this paper contributes a principled, well-founded suite of representations and algorithms with potential to benefit space time analysis at all levels of abstraction.

Classifying scenes (such as mountains, forests, and offices) is not an easy task owing to their variability, ambiguity, and the wide range of illumination and scale conditions that may apply. Two basic strategies can be found in the literature. The first uses low-level features such as colour, texture, power spectrum, etc. These approaches consider the scene as an individual object [16, 17] and are normally used to classify only a small number of scene categories (indoor versus outdoor, city versus landscape etc...). The second strategy uses an intermediate representations before classifying scenes [3, 11, 19], and has been applied to cases where there are a larger number of scene categories. In this paper we introduce a new scene classification algorithm based on a combination of unsupervised Neuro-optimization technique which is constructive for turn out accurate scene result from video

A. SHOT BOUNDARY DETECTION:

Algorithm 1: Shot Frontiers Detection

1: Let **M (f_i)** be the fi[th] frame in video sequence, where
F $_i$ = 1, 2, ….. , *Fg* *(Fg* refer to the whole variety of Video sequence)

2: Segregation of respective frame into blocks with x rows and y columns, and **L (m, n, fi)** stands for the block at (m, n) in the fi frame.

3: Evaluate the x^2 histogram matching oddity between the analogous blocks between successive frames in video sequence. **G (m, n, fi) and G (m, n, fi +1)** stand for the histogram of blocks at (m, n) inside the fith and (fi +1)th body respectively. Block's calculation is computed with the following equation:

$$D'(fi, fi+1, m, n,) = \sum_{1}^{L-1} [G(i,j,fi) - G(i,j,fi+1)]2 \Big/ G(i,j,fi)$$

4: Computing x2 histogram difference between two consecutive frames:

$$D'(fi, fi+1) = \sum_{1}^{x} \sum_{1}^{y} \text{Wmn } D'(\text{fi}, \text{fi}+1, \text{m}, \text{n})$$

where wij refers to the influence of block at (m, n) ;

5: Computing threshold automatically: computing the mean and standard variance of x2 histogram differentiation over the whole video sequence[7]. The following formulaes are used for Calculating the MD anSTD as follows:

$$MD = \sum_{fi=1}^{fv-1} \frac{D\ (fi, fi+1)}{fv-1}$$

$$STD = \sqrt{\sum_{fi=1}^{fv-1} \frac{(D\ (fi, fi+1) - MD)^2}{fv-1}}$$

6: Shot boundary detection
Let the threshold Td=Mean Deviation +b* STD
if D′ (i,i +1) ≥ T , the fi [th] frame is the end frame of previous shot, and the f(i +1)[th] frame is the final frame of subsequently shot.

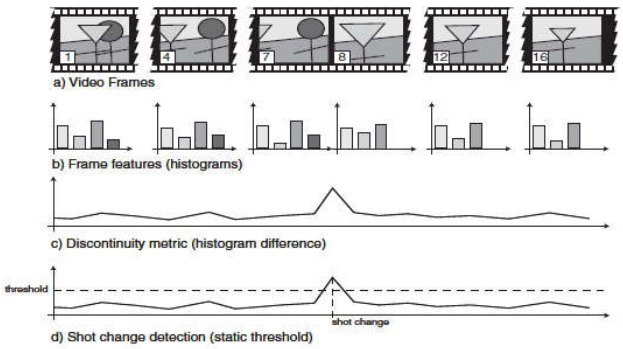


Fig 1: Shot Detection Mechanism [1]

Final shot detection: Shots may be identified as long sequences of frames instead of short once, usually, the short shot should last for 1 to 2.5 s. For the cause of fluency frame rate is at the least 22-25 fps. So, a shot consists of at the least a minimum quantity of 30 to 45 frames. In our test, video sequences are down sampled at 10 fps to enhance simulation velocity. On this situation, the shortest shot should comprise 9 to 14 frames. In this case we are considering 12th frame is chosen for our experimental purpose.

Reference Frame (RF): It is defined as the preliminary frame of any video;

$$GF = \sum_1^n \quad \begin{matrix} All\ frames\ in\ video\ sequence\ ; \\ except\ RF\ frames \end{matrix}$$

Where GF stands for General Frames

Dynamic Shot (DS) and Static Shot (SS): the shot can be declared as dynamic shot,

$$D''(s) = \begin{cases} \text{DS} & ;\text{if its max(i)} > MD \\ \text{SS} & \text{otherwise} \end{cases}$$

Algorithm 2: Key frame extraction

1: Calculate the differentiation between RF frames and GF frame with the above algorithm 1:

$$\boldsymbol{D}k(1,k) = \sum_1^x \sum_1^y \text{wij D fib}(1,\text{k},\text{m},\text{n}),\text{k} = 1,2,3,4 \ldots\ldots \text{Fc}$$

Where 'Dfib' is the differnce between the frame with their block differnce measures.

2. Trace the highest difference within a shot:

$\max(i) = \{Dc\ (1,k)\}_{max}$, k=2,3,4,.....Fcn.

3. Calculate "ShotType" according to the relationship between max(i) and MD: StaticShot(0) or DynamicShot:

$$\text{SHOTu} = \begin{cases} 1 & if\ \max(i) > MD \\ 0 & otherwise \end{cases}$$

4. Determining the position of key frame:

if SHOTu=0, with regards to odd numbers of shot,then the middle frames is choosen as key frame otherwise i.e.in even case the frame in between the two frame in the center of shot is treted as key frame.If SHOTu=1, the frame having MD is declared as key frame.

Based on some statistical information or reviews each researcher experiencing an evolution from small databases to large imaginary datasets, and now to Digital Library. A virtual library is a library wherein resources are to be had in machine-readable layout as opposed to print handy by way of computer systems. Digital Library is essential to Content based totally Retrieval (CBR). The CBR is a repository permits content-based retrieval. It contains digital textual content, sound, track, image, video, and so forth. A big leap forward from traditional database search which is largely based on simple attributes as a result content material based totally mechanism additionally serve as powerful surfing tool similar to the present day net search like Google. [4]

Therefore to find the exact end result through user question or through textual content based totally retrieval mechanism proposed machine specializes in Content-based totally Video Retrieval (CBVR) structures seem like a natural extension (or merge) of Content-based totally Image Retrieval (CBIR) structures. However, there are a number of things which can be not noted while managing frames which must be treated while the usage of shots. These factors are more often than not

associated with the temporal records available from video report. Visual file processing operations are important for routinely extracting an in depth description of a document. Feature extraction targets at characterizing a list of homes (called characteristic vector or record signature) for every aspect (pixel, frame place, frame, and sequence) of a video report.

Video indexing have to be analogous to text document indexing to facilitate speedy and accurate content get admission to video information, discern 1 indicates how video structuring have to segment a video report into pictures and scenes and once more extract key-frames or key sequences as index entries for scenes or stories. The paper specifically focus on content material-based totally video retrieval is growing technologies to routinely parse video, audio, and text to identify significant composition structure and to extract and constitute content attributes of any video resources.

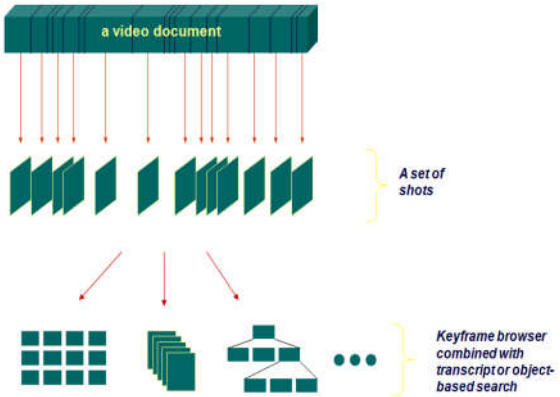


Figure 2: Structuring of Video

I.B. **Video Segmentation:**

Segmentation of Video is the first step towards content based video search analysis hence the entire video is broken up into particular scenes, afterwards those scene is converted in shots again shots are further divided in frames at the end. [10]

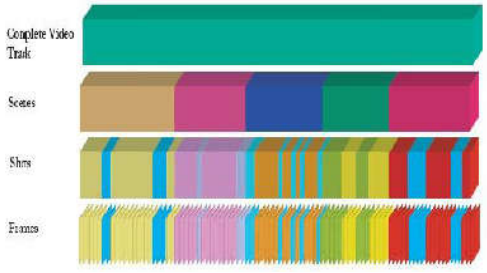


Figure 3: Video Segmentation

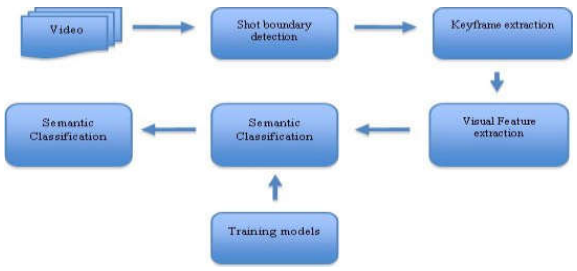


Figure 4: Video Content Analysis

From above diagram again checked it's far obvious that the video is classified as according to keyframes detection, extraction of functions is observed in addition semantic mechanism is applied for indexing motive on video as it's miles based totally different parameters. the afterward Semantic class is performed after visual features extraction at this degree first this system will train the various important keyframes which are saved inside the database through tagging with some bag of textual content phrases (bow).Moreover, this could be once more test through performing at evaluation degree for identifying exceptional phrases.

II. LITERATURE SURVEY:-

Costas Cotsaces et al.[1] this review work is dedicated to shot boundary detection and some gleam of video generalization & summarization. The various fundamental issues are provided and its countermeasures identified and calculated thoroughly.

B.H.Shekar , K P Uma [2] proposed the detection of shot using Gabor transformed first order method. This technique is applied on the datasets of TRECVID 2001 and compared the result.

Z. Cernekova et al. [3] keyframes are extracted and compressed video has formed the use of synthetic Genius strategies such so precision and take into account. The video is finished as input to the algorithm then framing is finished to clear up overseas frames current inner the video. Each body is converted into gray after limit the era day or computational attempt.

X. Zeng et al. [4] proposed approach for key frames have extracted the use of the dominant set of clustering for calculating similarity matrix

A Hnajlica et al. [5] first calculate seen settlement amongst each and each pair of frames according to pick function location but that saved amongst a matrix. All estimates is executed based totally on low-stage capabilities sure as an awful lot coloration, form yet texture.

Alan F. Smeaton [6] as the video is available in digital form hence it is tedious to detect SBD automatically therefore this paper gives the idea over the research based on seven years of TRECVID activities and outperforms the 57diffrent research groups to generates various best merits of SBD

J. Yuan et al. [7] Proposed that viably chooses the most precise and clearest picture for a grouping of vehicle image which begins tallying while a motional vehicle is going into the observation zone and closures whilst it takes to the air. Contrasted and special techniques, it has extended the adequacy and exactness for keyframe extraction of path vehicle statement video and accomplishes more powerful stress of video investigative data for route vehicle reconnaissance.

.

N. Babaguchi et al.[8] Proposed method is in mild of a shifting window of progressive casings that slides over the entire casing association (shot). The arrangement of edges included into each window is tried for content material homogeneity utilizing a fitting unimodality take a look at. In this way, each window is described as unimodal or no longer and the brink association of each non-unimodal window is splitted into (possibly unimodal) sections. Along these traces, each video shot is sectioned into unimodal fragments and the key-outlines are figured because the delegate outlines (medoids) of every unimodal portion.

Chi Zhang, Weiqiang Wang [9] proposed method first detects well-known shot obstacles based totally on the idea of Fisher criterion, and then classifies them into categories, reduce and slow transition (GT), through an SVM classifier. Further computation is finished for the GT shot boundaries to increase difficult boundary locations between frames into the transition c language including all of the transitional frames

ZHAN Chaohui DUAN Xiaohui [10] progressed shifting item detection set of rules primarily based on frame distinction and side detection. This approach no longer simplest keeps the small calculation from frame distinction technique and the impregnability of mild from area detection technique, however additionally improves in noise restraining..

Ijya Chugh, Ridhima Gupta[13] show the belief concerning the a variety of you will be able to techniques in location of remarkable stock going from keyframes to program surge. Equally there were numerous disturbing situations crocked apart numerous researches in addition to initiatives that one may also nicely meet approximately build a association that is capable of counter broadcast define descent in conjunction with endure credit score , this one essay ardour attempt that one may additionally examine powers that be inside a too treasured addiction.

A.B.Gadicha, Dr. M.V.Sarode, Dr. V.M.Thakare[11] proposed the basics methods to generate summaries are static and dynamic.it present extraordinary techniques for every mode inside the literature and describe a few capabilities used for producing video summaries

Kyong-Cheol Ko et al.[12] proposing new algorithm for detection of shot autonatically specially in flashlight events.here it uses long formula to compress and enhance the frame diffrence.

Ijya Chugh et al. [13] present a quick and successful video synopsis technique that is actualized in the packed area. Our four-advance proposed technique depends on a basic yet intense descriptor and a scene-recognition strategy, to recognize slow and unexpected advances with awesome accuracy. A saliency-based refinement procedure is utilized to maintain a strategic distance from repetition and speak to video content with as few key-outlines as could reasonably be expected

III. PROPOSED WORK:

From above survey it is observed that, video is basically decomposed in to shots , again shots are categorized into frames further frames are taken input as far as video processing or video information retraival is concerned.these all techniques works on principles of pixel by pixel mechenism the proposed work is dedicated to genearate efficient shot frontier detection using saliency map and feature extraction,edge map mechanism and key frame extartion from various categories of video using KNN classification.

A. Key frame extraction using total block difference:

For implementing the "key frame extraction using total block diffrence mechanism", our strategy is to extract frames from a video. Firstly, image frames are extracted from the target video and is stored in a particular directory. A function is created where each frame is then converted to its corresponding gray scale image. For every iteration, two consecutive gray scale images are taken and their Histogram difference is calculated. The sum of the elements of that histogram is then calculated and returned. The mean and standard deviation is calculated and threshold is computed using the values of this mean and standard deviation obtained.

B. Object Segmentation:

This methodology plays very vital role to generate the shot by doing object segmentation approach. In this approach the proposed system is capable of applying the saliency map techniques to find the object in the video. Then in the second part, proposed system is concentrating on SVM for creating different frames from entire video and generating most valuable key frames out of it for mining purpose.

C. Semantic Mechanism:

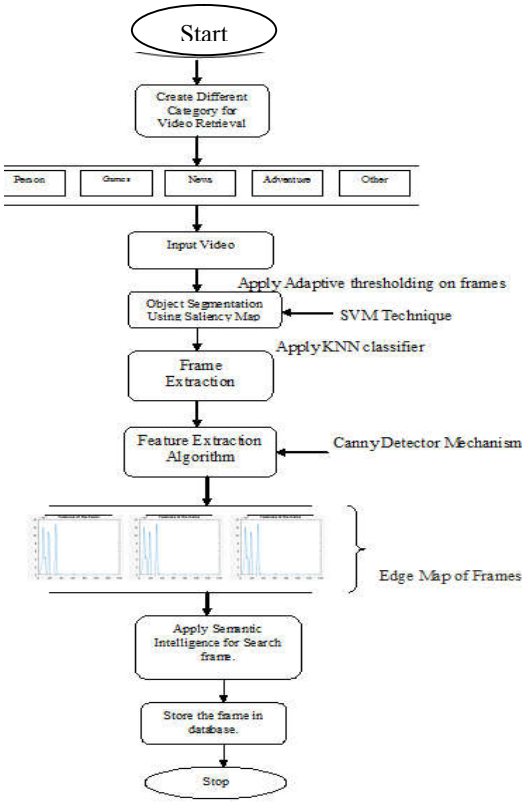


Figure 5: Semantic Mechanism

From figure 5, for enhancing semantic intelligence in video summarization , first a fall the proposed mechanism accepts input as video sequence then apply adaptive threshold or shot frontiers algorithm [Algorithm 1 stated above] to get the total number of shots in the entire video . Afterwards total number of shots is utilized for extraction of important key frames.

D. Canny Edge Detector :

The key frames are given to a feature extraction algorithm, which works as follows, Color map or extended histogram map is obtained by plotting the quantized color levels on X axis and the number of pixels matching the quantized color level on the Y axis. The obtained graph describes the color variation of the image and thus is used to describe the image during classification stage. The color map resembles to the gray level histogram of the image with one minor difference, that the color map quantizes the R, G and B components of the image before counting them, while the histogram directly counts the pixels belonging to a particular gray level and plots them. This ensures that all the color components of the image are taken into consideration by the descriptor.

While color map describes the color of the image, the extended edge map describes the edge variation in the image. To find the edge map, the image is first converted into binary, and then canny detector is applied to it. The original RGB image is quantized same as in the color map. The locations of the edges are observed, and the probability of occurrence edge on a particular quantized image level is plotted against the quantized pixels in order to evaluate the edge map of the image. The edge map is used to define the shape variation in the image and is a very useful and distinctive feature for any image classification system. These two features combined together can describe the image in terms of colour and shape.

Algorithm 3: Scene Classification Algorithm (SCA)

Input : Video Sequence
Output: Resulatant frame exact match with stored BOW
1: Input database is trained with semantic text or BOW
2: for each frame fi ← calculate words present in query;
3: Calculate frequency of each trained dataset(SF);
4: Calculate TF=N_{oe} /Total Words
5: Select ← Entry max(TF)

IV. RESULTS& DISCUSSION:

The prposed system is proficient to pull out the meaningful key frames from the video.this works professionally without compromising on the grounds of data integrity. The essence of our algorithm is that it brings into the use of histogram of each converted gray scale image through saliency map mechanism from which threshold is computed and based on this threshold key frames are selected. The proposed system has been divided into three parameetrs which performs various tasks at each level and are interrelated to each other.

1:Performance Evaluation:

The input video containing number plate of car had around 64 frames but for generating smooth and efficient result let us consider first 12 frames of an video. It extracts each frame from .avi format video; some of the frames are shown in Figure given as:

Table 1: Perfomrance Evaluation

Sr. No	Video Type	Precision (%)	Recall (%)	Our Proposed Algorithm (%)
1	Cartoons	80	93	98.22
2	News	91	89	93.33
3	Cricket	93	90	95.2
4	Football	89	90	94.11
5	Nature	90	89	91.45

2 Object Segmentaion:

This section the proposed idea will detect the object in gray scale from video using saliency map afterwards it calculates the Histograms of orignal frames to calculate the total difference between them.

3 Evaluate Semantic Analysis:

In this section the proposed system identifies the total time taken by each frame while processing meanwhile it also calculates the features of each frame.

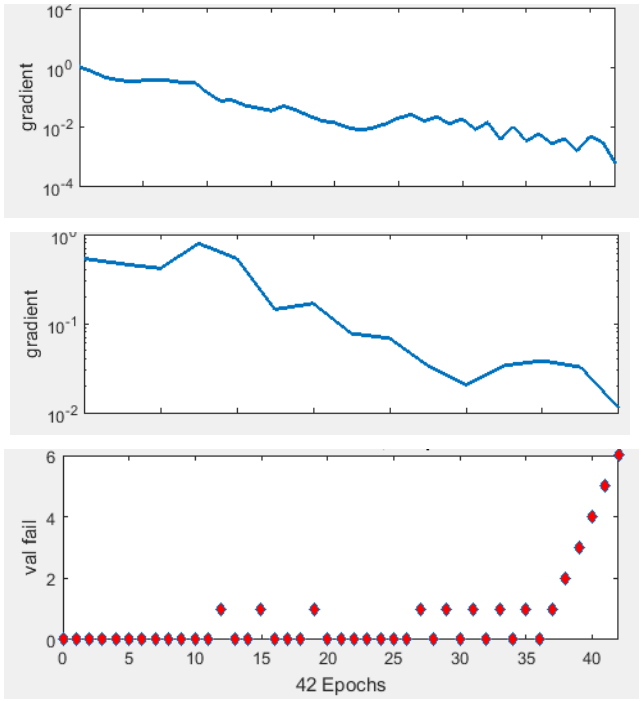
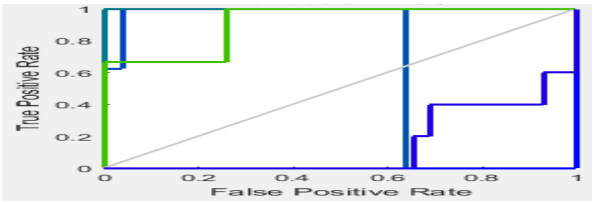


Fig 4: Semantic Values with Positive & Negative Values.



V . Conclusion:

Current research work is dedicated to detection of shot frontier detection using adaptive thresholding technique so that it is useful to identify the video claearly. Secondly key frame extraction from any .avi video with total block diffrence algorithm to get the ideal frames which are desirable for auxiliary processing.While applying the feature extratcion system appliy the the sailency map and canny detector which play vital role to generate the error diffusion hostogram for more detail study afterwards this paper yeilds to semantic analysis from database if the BOW(bag of words) are match with save database entry then we can conclude the semantic intelligance performs in significant way.