Analysis and Detection of Content based Video Retrieval

Shivanand S Gornale
Professor, Department of Computer Science, Rani Channamma University,
Belagavi, Karnataka.
Email: shivanand_gornale@yahoo.com, shivanand1971@rcub.ac.in

Ashvini K Babaleshwar
Research Scholar, Department of Computer Science, Rani Channamma University,
Belagavi, Karnataka.
Email: ashw.babaleshwar@gmail.com

Pravin L Yannawar
Assistant Professor, Vision and Intelligent System Lab, Department of Computer Science & IT,
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, (MS) India.
Email: pravin@visbamu.in

Abstract—Content Based Video Retrieval (CBVR) System has been investigated over past decade it's rooted in many applications like developments and technologies. The demand for extraction of high level semantics contents as well as handling of low level contents in video retrieval systems are still in need. Hence it motivates and encourages many researchers to discover their knowledge across CBVR domain and contribute their work to make the system more effective and useful in developing the system application. This paper highlights comprehensive and extensive review of CBVR techniques for detection of region of interest in a given video. The experiment is carried out for the detection of ROI using ACF detector. The detection rate of ROI is observed competitive and satisfactory.

Index Terms—Content Based Video Retrieval, Content Base Image Retrieval, Key Frames, and Shot.

I. INTRODUCTION

FOR most video databases, browsing as a means of retrieval is impractical, and query-based searching is required. Queries are often expressed as keywords, requiring the videos to be tagged. In view of the ever increasing size of video databases, the assumption of an appropriate and complete set of tags might be invalid, and content-based video retrieval techniques become vital. Video is an example of multimedia information, which contains huge set of raw data, richer content information and made up with continuous still images. This makes the retrieval of videos more tedious and challenging. The content based video retrieval have attracted the users as the use of multimedia data has been increased drastically in day to day life compared to text based search. Text based search or retrieval is good when the database is small and limited. But when it comes to huge and larger database or data warehouse, the content based retrieval play a vital role. Hence making the system more robust and reliable as well as automatized is necessary in the domain of CBVR system. The CBVR basically retrieve a related video from the database based on visual contents like colour, texture, edge, shape and motion. The steps involved in the CBVR system can be considered as follows along with the system architecture presented in Fig.1.

1. Inputting video and composition analysis of video.
2. Key frame identification and Object segmentation.
3. Determination of prominent discriminative features with respect to Region of Interest extraction process.
4. Construction of video database with respect to nature of application with segmentation and key frame scheme.
5. Querying database for comparing & extracting information content from database.
6. Retrieving the videos from the large multimedia database based on query videos.

The paper is organized as follows. The literature review is presented in the section II. The experimental analysis and results are presented in section III. Finally the conclusion of the paper is summarized in the section IV followed by references.
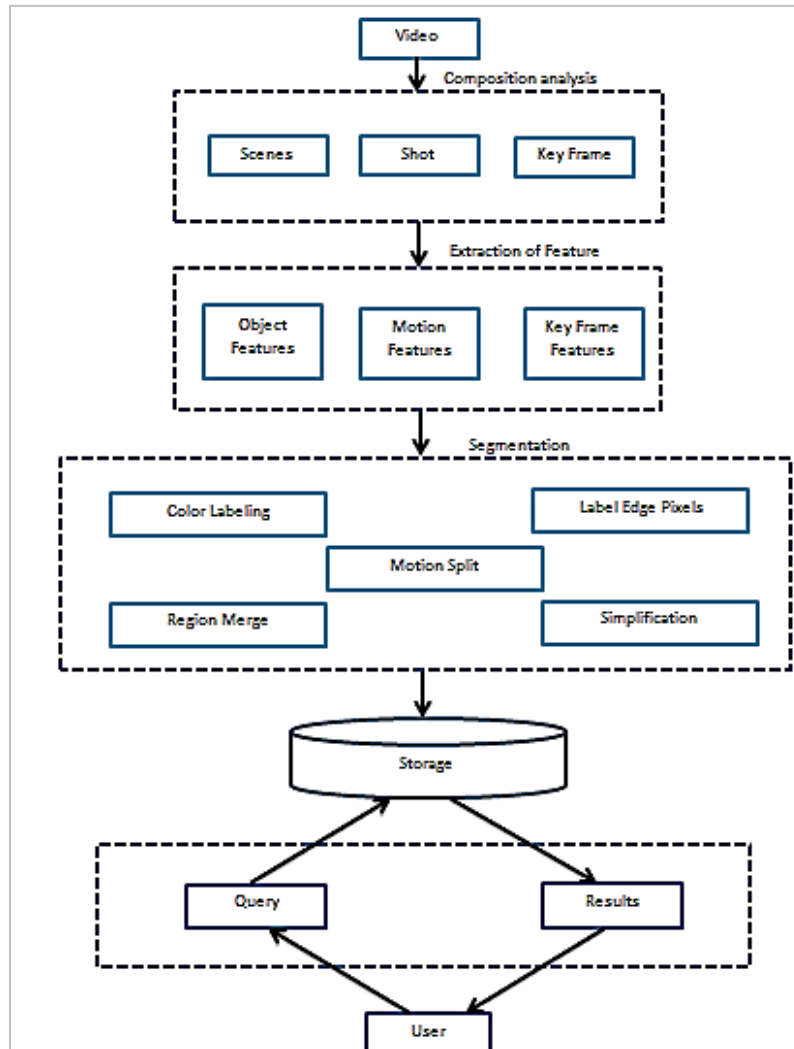
Fig.1 System Architecture

II. RELATED WORK

In this paper we summarize the contribution of various researchers as fallows.

Dipika H Patel [1] uses local feature descriptor and detector known as Speeded-Up Robust Feature for the retrieval of similar videos. Methods like Hessian matrix and distribution matrix are also been intended for the better performance of the system. Author found SURF worth when compared to previous methods such as SIFT, PCA-SIFT, GLOH.

H.B.Kekre et al. [2] have presented different methods and procedures used in the expansion of CBVR system. Techniques like invariance of color correlation, shot segmentation, key frame extraction, feature extraction, video structure analysis, object features, motion features, and Threshold-based, Edge-based, Region-based, scene detection for video database management system are few existing techniques that were mentioned and discussed in the review work [3, 10, 11, 15, 22, 42, 49, 51, 106].

B.V.Patel et al. [4] uses a technique named Principle Component Analysis for multimedia retrieval. Euclidean distance metric is been used for similarity measurement of videos present in the record and to increase the query performance rate multidimensional indexing is used in the system. Authors adopted two measures for the reduction of semantic gap named Automatic Metadata generation to the media and Relevance feedback for understanding and learning of the semantic context of a query operation.

Robles Oscar D. et al. [5] presents a CBVR system using Wavelet-Based Signature. The technique computes a multi-resolution representation using the Haar transform. The signature was specified using the standard description language XML2. The experiments were performed on the database consisting of 62 videos with 817 shots.

Zhong D et al. [6] uses a unified scheme for video object retrieval and segmentation on region-based approach. The model facilitates the spatial-temporal similarity searching of objects at several stages. The outcomes obtained were good in the performance of segmenting the objects.

P.M.Kamde et al. [7] used a texture descriptor approach for key frame extraction, video indexing, and edge detection algorithms for video retrieval. Experiment consists of different domain of videos with the combination of various features in different ways. The

result concludes that entropy based key frame extraction algorithm performs better when the image background is distinguished from the objects.

Milan R.Shetake et al. [8] uses a methodology for CBIR as well as for CBVR based on Support Vector Machine classifier. Authors concentrated on methods like HSV color feature, Gabion filter, Eccentricity, motion features, K-tree for the retrieval of videos from the data base. In [72] multi-resolution wavelet transformation feature extraction approach are used, images were resized and interpolated by utilizing multiclass SVM for image retrieval. Correctness for the Corel database is observed to be 62% and accuracy for ground truth database is 59.09%.

H.B.Kekre et al. [76] have proposed kekre transformations technique which is similar to Discrete Cosine, Walsh, Haar wavelet. Basically Kekre, Slant, Discrete Sine and Discrete Hartley usually reduce the size of feature vector using factorial coefficients by transformed frames. Similarity measures like Mean Square Error (MSE), precision and recall are performed. Average precision and recall using Haar Transformation is observed as 0.5702 accuracy was noted to be 57.02% with reduction of time complexity by 2048 times as compared to full feature set.

Meera R.Nath [92] have proposed a system that provides the video matching based on multiclass SVM classifier and RGB color components. Authors introduce two functions in multiclass Support Vector Machine. First one trains SVM. Second function, classifies the new data item among trained function. The system achieved fast and efficient video retrieval from the database.

Madhav Gitteet et al. [9] present a methodology that implements video mining from multimedia warehousing. The system involves two stages, constructing the warehouse and the retrieval of video from its multimedia warehouse. Feature like Local Color Histogram ,Global Color Histogram, RGB Algorithms, Eccentricity algorithm is considered, K-Mean's clustering Algorithms, and B+ Tree for indexing were performed by SVM classifier using Euclidian Distance for similarity measure good results achieved.

Navdeep Kaur et al. [12] present a video retrieval system using frequency domain analysis with 2-D correlation algorithm. Experimental analysis shows the proposed technique gives better performance result when it is matched with existing underlying method of SURF implementation.

Jin Yonggang et al. [13] present the video retrieval system based on motion trajectory. Proposed method segments the moving flagged objects from the given video automatically and extracts the motion trajectory features by descriptor sets. Based on their similarity matching score from spatial model and spatio temporal model retrieval of videos was carried out. In experimental analysis they had considered 83 moving objects with different speed in their database. Authors conclude conservative results were obtained while using MPEG-7 motion trajectory descriptor.

Truong et al. [14] has given an article that delivers an efficient classification of video abstraction by identifying the different approaches in two prevailing forms, the video key frame set and the video skimming with some effective evaluation ideas. The authors found the methods efficient for their dataset.

Jana et al. [103] has proposed unsupervised clustering of frames from video using local SURF features and K-Means sequential clustering algorithm. The performance took advantage as elusion requirement of shot detection and need for segmentation and the amount of clusters formed is more in numbers when the threshold value is lesser in comparison to the different method.

A. Dayan [16] proposes a method for CBVR based on shape features and motion trajectory features, for the efficient retrieval of videos Gabor filter response is utilized with Curvature Scale Space (CSS), Multi-spectral representation methods. Authors have experimented by using 3D Curvature Scale Space (3D-CSS) and Multi Spector Temporal CSS (MST-CSS) for the representation of video objects. Overall the results obtained are conservative.

Xiang Ma et al. [18] use a compressed representation of multi-object motion trajectories. The approach signifies three algorithms as Geometrical Multiple-Trajectory Indexing and Retrieval, Unfolded MIR, Concentrated MIR are applied for representation of motion patterns. Authors have used Global and Segmented Multiple Trajectory Tensors. Experimental results shows that CMIR algorithm gives good precision recall metric with less query processing time then other algorithms.

Benmokhtar Rachid et al. [19] presents a fusion of various low levels visual and shape descriptor using shape and motion features. The proposed method shows the improvement from 96% to 100% average precision rate.

Chao et al [17] have anticipated CBIR system with JPEG query pictures to check the recital of image comparison with different detector and descriptor. Authors concluded that, use of Hessian Affine detector gives good and robust performance in the presence of strong JPEG compression. [20] Color and texture features, and features extraction using color histogram and Gabor filter. [23] Multi-Scale technique to improve the excellence of natural images and also tried to discard the need of explicit scale selection and edge tracking. [24] Presents a technique that utilize the perceptual HIS (hue, intensity, saturation) color space for making algorithm more. The pixels in the image were divided into chromatic and achromatic regions using classification method to make the algorithm work effectively. [26] Adaptive Bayesian color segmentation method for segmentation of color information. Gibbs random field model is used for mapping. [37] Also use multi-scale image segmentation with integration of Region detection and edge counter. The homogeneity scale is used for images at different scale levels. Gestalt analyses were applied for integrated detection of edge and region count. The system requires 10 seconds to retrieve a 512 X 512 image on Sun SPARC20 workstation. [94] Use color

image retrieval using region-growing algorithm, with invariant shading and highlights. Methods like PCA and Euclidean distance matric is applied for gaining good and effective results.

Yasira Beevi C P et al. [21] presents the video segmentation using different modes to identify change detection, background registration techniques and real time adaptive threshold techniques respectively for video segmentation. Statistical-testing algorithm and noise-robust threshold method are used to calculate threshold noise. The generated segmentation results show less computation convolution and high efficiency as compare to other algorithms.

Haitao Jiang et al. [22] have proposed taxonomy for classification of Scene Change Detection with different categories like uncompressed full-image sequences, procedures working openly on the compressed video, and methods based on unambiguous models. Authors have matched some current SCD algorithms by gaining 90% for abrupt scene changes and 80% for gradual scene changes.

Brandt et al. [93] use frame-based scene change detection algorithm for compressed videotape. Motion vector histogram is used for each frame with different measures that are analyzed using DCT values and motion information. The average processing time for each frame ranges between 0.6 and 6.5ms for the test video stream.

Wu Chuan et al. [25] present a CBVR system based on motion features. The system has concentrated on background and foreground motion. Methods like 6-parameter affine model and global motion estimation algorithm are used. Methods like Global motion, object trajectory with MPEG-7 parametric motion, and motion trajectory descriptors are also discussed. Experimental outcomes show that the anticipated system is reliable and effective.

Liu Tianming et al. [27] use a combination of motion-based temporal segmentation and color-based shot detection for extraction of key frames. The system uses automatic perceived motion patterns for determination of the number of key frames and their location in a given video. The algorithm measures visual content complexity of a shot by action events and limits the key frames to abstract a video by action events give good percentage of result up to 77% to 86 % for different videos.

Guozhu Liu et al. [44] have proposed improved histogram matching method for segmentation and key frames extraction. To select key frames different technique are used like shod activity, macro-block, and based on motion analysis. Authors classify frames as I-frames, P-frames, B-frames and use features for each sub-lens. The average performance of algorithm was 99.48% with average fidelity of 0.76%, and thus extracted good and accurate key frames.

Benjamin. D et al. [102] presents a video object segmentation system based on weakly supervised method. Objects motion segmentation and local indications to methods were used for tracking and detecting of objects. Different set of database has been used for experimentation like YouTube, SegTrack, EgoMotion,

and FBMS. The results obtained using YouTube was 3% better than other sate of art methods where as 16% good performance on other datasets has been obtained.

Mohanty et al. [28] proposed a system for video indexing and searching based on semantic key words. The system retrieves essential frames from videos and further correlating the image characteristics with their semantic features. The SVM classifier is been trained on color, structure and texture features to classify the videos in the database.

Huang et al. [29] presents a method for integrated key frame extraction for video. They have made use of OpenCV tool to select key frames based on shot boundary and Visual content frames. For similarity, comparison of histogram, Chi-Square method was applied. Experiments show that algorithm requires a less time and space complexity.

Li Zhao et al. [50] introduced a metric to calculate the distance between query image and a shot based on the concept of NFL. The feature trajectory of a video is considered as key frames and the lines that pass through the points are considered as shot in a video. The distance between adjacent key frames is calculated using Euclidean distance measure. The color histogram is used as features in the experimentation. Moreover, their results show that breakpoints based key frame extraction +NFL method gives batter performance when compared to NN, WL and traditional NFL method.

Azra Nasreen et al. [95] presented a shot segmentation based on Edge Change Ratio approach. For the detection of edge Sobel method is used and for the selection of key frames adaptive threshold method is been applied. The system obtained efficiency and robust results.

Zhong Qu et al. [96] present a system for key frame extraction and video summarization using pixel and histogram based methods. Method like Histogram intersection, non-uniform portioning and weighting are used for video shot segmentation. Each of these shot considered with a key frame depending on calculation of image entropy of every frame in HSV color space. The experimental result shows that the algorithm gives an improvement in shot detection at certain degree.

Kalpana Takre et al. [30] use motion feature, quantized color feature and edge density feature for retrieval of video objects respectively. Shot segmentation is computed by Bi-orthogonal wavelet transformation. In addition, for similarity measurement Latent Semantic Indexing is considered to retrieve video from database. The system has obtained efficient result with 0.85 precision rate and 0.89 recall rate.

Janko calic et al. [31] has proposed a key-frame summarization based on dynamic programming solution that gives a comic-like representation of analyzed videos. K-means clustering algorithm was applied to calculate eigenvalues. TRECVID 2006 evaluation content used to conduct experiments. The summary of one hour video of 250 shots where browsed easily as a result. Werner Bailer et al. [43] also presents a framework for video summarization with automatic feature extraction model and web based retrieval. The framework consists of

homogenous approach to signify semantic objects, and low-level features, which is independent of several components and is fully rely on the MPEG-7 standard as information base. Algorithms like color histogram, MPEG-7 edge histogram descriptor, scene boundary indicator were described. The results were demonstrated on XML and XSLT web pages, the three most relevant key frames were displayed to user, and then navigated through them for the desired video summary.

Tjondronegoro et al. [52] presents a framework for video summarization to automatically detect and mark the key events and key players in broadcasted sports video. The algorithm extracts key events and key actors on sentiments and players popularity. As a result, authors have successfully detected the important actors and key events.

T.N.Shanmugam et al [32] worked on video retrieval system with various methods. Features like motion, edge, color and texture were extracted with the help of FFT, L2 distance function, geometric approach, HSV conversion and Gabor Wavelets. Shot segmentation is computed using 2-D correlation coefficient technique. Further Kullback distance is used for similarity measurement. Authors concluded that system exhibits a satisfactory precision which proves the system is effective. Similarly [33] have also worked on integrated multimedia retrieval system using low level and motion features, with good results.

Santhosh Kumar K. et al. [97] presents a CBVR System using low and high level features. Techniques like Mean, Standard deviation, and Entropy are used to select the feature vector. The similarity measurement is processed using Euclidean Distance method. The obtained results have reached up to 60% in average and the use of multi-feature approach gave increment about 18 to 16 % of precision and recall respectively.

Philippe Salembier et al. [34] have proposed segmentation based on video coding that allows the content-based manipulation of objects using spatiotemporal segmentation. The system involves a global optimization for partitioning each region. The system propose MPEG-4 framework standard. The algorithm works efficiently by facilitating to track and manipulate objects.

Chu-Hong et at. [35] Worked on video similarity detection based on two signatures, Pyramid Density Histogram and Nearest Feature Trajectory technique. To calculate the similarity measure color histogram is been considered as low level features to generate the two signatures coarse (PDH) and fine (NFT). Authors concluded that the NFT technique works better matched to conventional approach. The approach gives 90% of precision rate and 85% of recall rate.

Yun Zhai et al. [36] presented a video segmentation system based on visual and text information for news videos. The videos are represented using Shot Connectivity Graph. Authors made use of Automatic Speech Recognition for text extraction. The dataset used in the experiment were taken from NIST dataset. The merging technique of text and visual reduce the average

false rate to 2-3 and increase in precision call up to 5% to 10%.

M. Cooper et al. [38] worked on temporal media segmentation using supervised classification. The system builds intermediate features using pairwise similarity via standard low-level features. Experiments were carried on TRECVID 2004 protocol. The results obtained were good number of segmentations.

N. Peyrard et al. [39] presented a methodology for multimedia segmentation using low-level motion information. Segmentation is carried using statically motion activity model of a video. Algorithms like Gibbsian modeling distribution and kullback-Leibler distance measures are used for merging. The experiments were conceded out for real large hour videos of sports and movie, resulting better performance.

Hari Sundaram et al. [40] present an algorithm for video scene segmentation. The system consider thoroughly constant piece of audio-visual data and fuse the subsequent segments using a nearest neighbor algorithm that is further refined using a time-alignment distribution derived from the ground truth. The segmentation algorithm achieved an accuracy of 20/24 scenes 84% and 18/24 75% using the refinement.

Z. Gu et al. [73] presents an Energy Minimization Segmentation algorithm for scene segmentation. Authors introduce two energy items Content (global) and Context (local). To find initial estimation of scene label for every shot Expectation Maximization algorithm is used and Iterated Conditional Modes algorithm for global optimization. The result the EMS algorithm demonstrates that it works effectively with various domains of videos.

Lin Tong et al. [101] present a framework for video scene extraction using shot grouping. The segmentation was carried using pseudo-object-based shot correlation analysis using dominant color in individual frames. The method performed effectively and efficient on experiments tested.

S.S.Gornale et al. [104] have proposed the work on automatic detection of signage's from a video streams. The content of the work was concentrated on traffic signage's. This signage's are then classified using SVM and KNN on LBP feature vector. The system obtained 89 % for KNN and 90% for SVM classifier.

Chandra Mani Sharma et al. [41] have proposed a methodology for offline multiclass object classification. Haar-like features, Stage-wise Additive Modeling, and Multiclass Exponential boosting strategy have been used. The detection accuracy of method is 98.30% processing with 25 frames per second.

Chenliang Xu et al. [45] presents a framework for hierarchical video segmentation with data streams using Markovian assumption and instantiate the streaming by graph based streaming hierarchical approach. Authors conclude that the method give fully unsupervised online categorized video segmentation that can handle any randomly lengthy video. The experimental results for the approach gave better performance.

Hatim G. Zaini et al. [46] has discussed the growth of the multimedia and its retrieval techniques and the future

trends. Authors have adopted threshold method for segmentation and extraction of still images. The results obtained were good with 13% to 19% increment as compared to traditional system that concentrates on color.

Kin-Wai Sze et al. [47] uses an optimal representation for video shot segmentation based on global statistics. The system construct Temporal Maximum Occurrence Frame as optimum demonstration of all frames, and further with k-TMOF and k-TMOF pixels with large probability values. The experiments were concentrated on 427 shots taken from MPEG-7 dataset. Results show that both the k-TMOF and k-p TMOF achieves the good performance.

Margarita Osadchy et al. [48] uses anti-face method for the detection of actions in a video at different velocities tested on rotating objects and visual speech recognition. The database contains 20 objects taken from COIL database. Authors conclude that the result using anti-sequence method is robust for varying events with 4.8% of false rate and the method is good to handle rotation invariants.

L Zelnik Manor et al. [88] presents extraction of video sequence based on their behavioral content and statistical distance method. The system is able to handle a wide range of dynamic events as it is nonparametric. Authors use the measure for different multimedia uses, including event-based detection, video segmentation/indexing, and long term video clustering. Authors concluded that accuracy of event-based distance measure is lesser and sophisticated for parametric models.

Zhai Yun et al. [53] proposed a framework for temporal scene segmentation for different videos. The video scene boundaries are calculated using Markov chain Monte Carlo technique. The merging of two scenes and the splitting of an existing scene were done using diffusion and jump method**. The experiment is concentrated on home videos and movies and obtained high accuracy result.

John R. Smith [54] worked on internet video retrieval system. Spatial and temporal resolutions building blocks are used to construct the video graph to represent the view elements. Wavelet filter bank with different filter in space and time are used to gain progressive retrieval session with uniform video graph. The results obtained were satisfactory.

Tabii Youness et al. [55] worked on shot detection approach. The system used motion features for two dimensional videos. The dataset consists of different domain of videos like sport, news etc. The result obtained is effective with average precision rate of 91.80% and 86.15% recall rate for sports video.

John S. Boreczky et al. [56] presents a video shot boundary detection using motion features. The system use Diamond Search algorithm to extract the motion vector. The experiment is carried on 2004 shot boundaries and around 200 scene boundaries. The result obtained 30% of scene boundaries marked as gradual transitions and 10% of nonsense shot boundaries marked as gradual transitions.

Jesus Bescos et al. [87] present shot change detection

for online videos with MPEG-2 standard. The total number of abrupt transitions obtained by system is 0.99 recall and 0.95 precision rate, whereas for gradual transition 0.87 recall and 0.66 precision rate.

David Asatryan et al. [99] have proposed CBVR system and indexing based on structural similarity method. The system gives a concept of shot segmentation analysis and interpretation using graphical and numerical methods. Gradient field based algorithm is used for shot detection using edge features. Authors make use of two parameters of the hypothetical frame to determine the key frame that calculates arithmetic mean of corresponding parameters in the current shot. The method obtained accurate shots.

H. Lu et al. [68] have proposed shot boundary detection based on hypothesis testing and clustering. The K-Means clustering algorithm is used for video scene group whereas scene segmentation is carried using hypothesis testing method to identify the shot boundaries at every level with different scene cluster number. The database consists of 8710 frames with 46 shot boundaries. As a result authors calculated 39 abrupt changes and 7 gradual changes for tennis video.

Nikita Sao et al. [91] presents video shot boundary detection by introducing the concept of graph theoretic, which gives a dominant set from set of nodes in a graph depending on like and unlike groups. Edge weight between nodes are calculated by using pair wise dissimilarities of frames with the complete information of graph and its discontinuity is dignified and compared with threshold value. Authors have conclude by marking that the nodal analysis work well with successful results in the area of shot boundary detection.

L. Bai et al. [67] presents a shot boundary detection system. The system detects gradual transitions and abrupt cuts. Frame difference was calculated using color information using HSV components. The system results obtained 0.89%, 0.75%, 0.81% precision rates and 0.92%, 0.75%, 0.73% of recall rate for cuts, fades and dissolve respectively.

Mohini Deokar et al. [58] have discussed different techniques of shot boundary detection depending on video contents and change in the video contents. Methods like shot boundary detection (pixel comparison, transformation-based, histogram-based difference, edge, motion, statistical based), Shot change detection, Shot Classification Algorithm (Spatial feature, temporal domain of continuity metric) were discussed.

Sanket Shinde et al. [59] have proposed video detection approach base on SIFT feature. Then system is suitably robust to detect transformed versions of original videos by pointing the copied location using different photometric and geometric transformations. Experimentation gained original video from reference video identified through query video.

Ran Dubin et al. [61] present an algorithm for video quality representation classification on YouTube videos. The solution is tested on the flash player called Safari browser with both online and offline network traffic over HTTPS. Authors have achieved classification accuracy of

97.18% in fixed mode and 97.14% in automatic quality representation switching mode. The algorithm estimates the user buffer layout level after each segment download with an average error of 0.035 seconds. The method exhibits about 8.95% average classification rate.

Ho et al. [62] use video retrieval system based on coarse to fine. The scheme proposed consists of two level coarse searches and a fine search. Shot-level motion and color distribution are computed as spatial-temporal for shot matching. The causality relation is used to improve the search quality. Authors concluded that the system result is satisfactory for retrieval of video with low complexity as matched to the traditional single coarse system. Deepak C R et al. [60] have also use spatial-temporal features for matching of videos. The result proves that proposed method has high precision and recall with average precision of 85% and 76% recall rates obtained.

Yasuo Ariki et al. [63] present a methodology for scene detection from live videos as well as recognition of speech. The system gives 13% of improvement in word accuracy for continuous speech using corpus integration, name classes and pronunciation modification whereas acoustic model gained 20% improvement at word accuracy and about 15% improvement at keyword extraction rate using the supervised and unsupervised adaptation.

Yusuf Aytar et al. [64] present an approach for retrieving video based on semantic word similarity and visual content of video. Authors have used TRECVID 06 and 07 for feature extraction test and search test respectively. The system obtained 81% better then text based query approach.

Bashir et al. [65] use a motion trajectory and temporal ordering sub-trajectories for retrieval of videos from the database. PCA is used to represent the sub-trajectories and further for indexing and retrieval. Authors considered the experiments on two components accuracy and efficiency which were measured on retrieval effectiveness using precision-recall matric and efficiency is computed by the time taken by indexing and retrieval process. As a result authors observed that PCA based system gives a marked improvement in precision compared to recall values.

Hamdy K. Elminir et al. [66] presents a methodology for video segmentation and key frame selection based on adaptive threshold method. The video representation is done using both low level features and high level semantic object annotation. The experimental results obtained are 13% to 19 % more when compared to traditional system that uses color feature for video retrieval.

Guironnet et al. [69] have proposed a methodology for video summarization. The summarization is carried on camera motion that selects frames according to the series and the magnitude of camera motions. The experimental result shows good efficiency of the system for motion-based summary.

S. Padmakala et al. [70] present a CBVR system based on optimal key frame extraction and video segmentation.

The segmentation is carried using LBP features, color features and the optimal frame is obtained using OFR. These features are compared using feature weighted distance. If the matched distance is less than the threshold the video is retrieved from the data base. Authors observed that the experiments were good and efficient.

A Paul et al. [71] have presented a system for video search and indexing based on SVM. State transition correction technique and transition quality estimation is used in the system. Energy efficiency is measured by Dynamic power Management technique. The result obtained by SVM achieved precision rate about 83.83% and the query result of indexing graph give 80% accuracy.

Krishnamoorthy R et al. [74] have proposed video shot detection. Object Level Frame Compression method is used for shot detection through orthogonal polynomial in a compressed domain. Where in to represent the background and adapt the background orthogonal polynomials transform coefficients is been used. The results obtained by the system are 98.3%.

B.V. Patel et al. [77] present a video retrieval system based on Edge detection, Entropy, Black and White color features. The database creation is carried using GLC matrix. To retrieve the videos Black and white feature points are calculated on edges using Prewitt edge detection. The results were good and efficient.

Poonam V Patil et al. [100] presents a CBVR system using key frame extraction. The system uses auto-dual threshold method to eliminate duplicate frames, and shot boundary detection for video segmentation. Color, edge and shape features were extracted using LCH and GCH and average RGB algorithm respectively. SVD method for similarity matching and video sequence matching is also been used. The elimination rate for near duplicate frame extraction number reached 30 % to 40% and elimination of redundant video frames using SVD method achieved accuracy up to 35% to 36%. [75] Have also performed for video sequence matching based on color correlation approach with good result.

Chiranjoy Chattopadhyay et al. [78] present a methodology for sketch based CBVR system that combines shape and motion trajectory features. The video contents were represented using a method known as Multi-Spector Temporal-Curvature Scale Space. Authors concluded that the results of retrieval profound to correctness of hand-drawn sketches.

Xingxiao Wu et al. [80] proposed a system for action recognition based on multilevel feature and Latent Structured SVM (LSSVM). Gaussian mixture model (GMM) is used for representing local spatio temporal contexts. The compatibility between multilevel action feature and action labels LSSVM method is used. The experiments were carried on UCF sports, YOUTUBE and UCF50 database. As a result they have obtained 86.91% for SVM and 88.04% for LSSVM.

Michele Merler et al. [81] present a video event representation based on semantic model vectors. The system covers huge number of semantic detectors like scenes, objects, people, and various image types. The model shows the best performance on single feature with

correctness of 0.392 on the TRECVID MED10 dataset. The properties make this semantic representation appropriate for large-scale video modeling, classification, and retrieval. Further the experiments are carried with different feature granularities and fusion types which were observed that the performance boosted with 0.46 overall MAP scores.

Wang X et al. [82] presents a video event classification system based on Independent Component Analysis mixture feature functions. The experiments shows that the method as better discriminative power than other HMM-based methods for low activity videos. In addition, system determines the benefit of using ICA mixture features over Gaussian mixture for non-Gaussian features.

J Fan [83] presents a system for video database indexing and retrieval. The videos are classified using hierarchical semantics sensitive technique. The classification imperious is defined using Expectation Maximization algorithm. The approach has achieved 0.9% of precision as well as recall rate with less time complexity.

G Xu et al. [84] presented a methodology for video semantic analysis using Hidden Markov Model. The method composes semantics in different granularities that are plotted into hierarchical model space. Authors also present a two-motion representation scheme- Energy Redistribution and Motion Filters, which is robust to different motion vector sources.

J Tang et al. [85] have proposed a novel multi-label method named correlative LNP for video annotation. The method incorporates the semantic concept correlation into graph-based semi-supervised learning to improve the performance of the system. Authors concluded that the system works gives less time complexity.

Sarah Joy et al. [86] present a system for video forgery detection, which detects and recognizes forged video sequence from an original database. The system proposes color correlation algorithm to overcome problems of grouping the characters and eliminate false positives the performance evolution MUSCLE VCD is used. The method achieved 89.70% precision rate and 82.30% recall rate.

The Thepade's et al. [90] proposed a methodology for video retrieval system based on color feature extraction method. Thepade's sorted Ternary Block Truncation Coding (TSTBTC) is been used for feature extraction method. Absolute difference measurement is used to match the videos from the database. The database consists of 500 videos with 10 categories. Authors have made use of RGB color space including KLUV, YIQ, YUV, YCgCb, YCbCr and XYZ. Authors concluded that the TSTBTC gives better result as compared to traditional BTC method.

C. Papageorgiou et al. [89] presents a framework for object detection in cluttered scenes .The framework make use of Wavelet representation for an object class derived from statistical analysis and SVM classifier features like color, texture and patterns are used for extraction. The results detect faces from the minimal size of 19 x 19 to 5 times size by scaling the novel image from 0.2 to 1.0 times its original size.

S.B.Mahalakshmi et al. [98] presents the multimodal system for effective CBVR system with capacity to work with text, image, and audio, embedded text in signal, embedded text in image for video searching, video surveillance, and clustering. Method like HCT (Hierarchical Clustering Tree) is used for clustering and similarity matching by feature extraction is done using SIFT (Scale Invariant Feature Transform) descriptor based on color, texture and shape. The proposed system bridges the semantic gap between the low- level visual features and high-level features.

D. Asha et al. [105] uses the multiple feature extraction method for video retrieval. The Experiment was concentrated on color, texture features and motion vectors. The proposed system has obtained 80% precision and recall rate for online retrieval of data. The observation made in the experimentation is that the multiple feature CBVR system performance good than the single feature CBVR system. [79] Use motion features to extract direction and intensity of the video, obtaining good results.

Some of the challenges and issues raised by researchers during their course on research and development are - Faster extraction of low level features, Motion feature analysis, Affective computing based video retrieval and video matching, Multiple object tracking, a hybrid strategy of video objects, the possibility of taking multiple user inputs at beginning of segmentation, to develop an interactive user interface based on visually interpreting the data using a selected measure to assist the selection process, the single video shot should be permitted to be classified into the related multiple visual concept nodes as it consist of multiple semantic visual concepts, to generate the classification rules by integrating the unlabeled video clips with the limited labeled video clips, and development of High-dimensional data visualization is necessary for the effectiveness of the semantic video classifier. In addition, the CBVRS is used in video archiving and medical applications. From this analysis, it seems clear that there is a need for CBVR in different domains.

The results proposed in the literature are found to be competitive and efficient. Hence there is a scope for developing more significant and robust algorithm for CBVR, which will compete to the earlier results of Semantic Retrieval, Motion Feature Analysis, Hierarchical Analysis of Video Contents, and Combination of Perception with Video Retrieval. This review will be helpful for the budding researchers understanding the domain knowledge in video analysis more comprehensively. Table 1 Provides some information and the results obtained by different method and techniques used by the authors.

Table1. Analysis and findings obtained by different methods.

Author	Methods/ techniques	Database	Conclusions/ Findings
Dipika H Patel 2015	Video retrieval using enhance feature extraction SURF, C-SURF	Model data library OpenCV library	The experimentation showed 540 frames were extracted using OpenCV library. With the frame rate of 15 fps.
Mei Huang et al 2013	An integrated scheme for video key frame extraction	Microsoft Visual Studio 2010 and OpenCV2.4.3	The experimental results have gained 0.80 precision rates whereas 0.94 recall rate for football video.
Kalpana Thakre et al 2011	CBVR system based on dominant feature	Database in MPEG-2 format.	The system obtained efficiency with 0.85 precision rate and 0.88 recall rate.
Janko Calic et al 2007	Video summarization based on dynamic programming	TRECVID 2006	The system gave good time complexity with the overall summary extraction of one hour video giving 250 shots
Chu-Hong et al 2003	Two-phase scheme for video similarity detection	Original data space	The methodology used have obtained an average of 90% recall with 50% average precision rate using FDPH
M. Cooper et al 2007	temporal media segmentation using supervised classification	Database TRECVID	The experimental results are efficient with the 0.92% of precision rate and 0.87% recall rate.
Chandra Mani Sharma et al 2012	Video Object Classification Scheme using Offline Feature Extraction and Machine Learning	standard datasets like CalTek101, MIT-CMU datasets	The method obtained detection accuracy of 98.30% and it processes 25 fps.
Guozhu Liu et al 2009	MPEG Video stream	MPEG video stream	The experimental results show that good fidelity and compression ratio with performance of 99.48% for compression and 0.76% for fidelity.
Tabii Youness et al. 2014	Shot boundary detection using motion activities using variance 2D	Own database	The algorithm have obtained 91.80% precision rate and 86.15% recall rate for sports video
Deepak C R et al. 2013	Cluster overlapping method	Own database	The average precision of 85% and 76% recall were obtained.
Ran Dubin et al. 2016	encrypted HTTP adaptive streaming	Safari (Flash player) browser and online network traffic over HTTP	System can independently classify the real time video with 97.18% average accuracy.
Yasuo Ariki et al. 2003	sophisticated speech recognition	System constructed on continuous speech database	The language model adaptation achieved 13% improvement. The exemplary reached 20% improvement at word accuracy and around 15% improvement at key word extraction rate and almost 90% recall and precision were achieved for the scene detection.
Hamdy K. Elminir et al 2012	Adaptive threshold method	Feature database	The experimental results show that multi features improved both precision and recall rates with 13% to 19 % compared to traditional system.
Ela Yildizer et al 2012	Multiple SVM classifier	Corel database, ground truth database	The Classification accuracy with Corel database is 62% whereas accuracy for ground truth database is found 59.09%
Z-J Zha et al 2012	Statistical active learning	TRECVID-07 dataset	The method progress the video indexing performance compared to the state-of-the-art by 35.8%, 35.8%, 24.2%, 17.5%, and 15.5% compared to DUAL, LapDOD, SVsM*active*, and SVM*active+rdd*
Krishnamoorthy R et al 2012	OLFC for video shot detection with orthogonal polynomials model	MPEG video	The system achieves an average of 98% recall and 99% precision rate.
Dr. H.B.Kekre et al 2015	Factorial coefficients of orthogonal wavelet transforms	Image database	The system shows that wavelet transform generated using Kekre transform matrix gives best performance with 0.048% fractional coefficient
Zakarya Droueche et al. 2012	motion histogram and Combination of segmentation and tracking.	MPEG-4 AVC/H.264 stream	The system is applied to a dataset of 69; the retrieval competence is greater than 69%.
Michele Merler et al. 2012	Semantic Model Vectors for Complex Video Event Recognition	TRECVID MED10 dataset	The technique reached the precision of 0.392 on the TRECVID MED10 dataset, and 0.46 overall MAP scores obtained from a late fusion of static and dynamic features.
Sarah Joy et al 2014	Invariance of color correlation	MUSCLE VCD benchmark	The system performed 89.70% precision rate and 82.30% recall rate

Santhosh kumar et al 2015	Low and high level semantic gap	Feature database	The experimental results show that both precision and recall reached up to 60% in average. And the use of multi-feature approach gave increment about 18 to 16 % of precision and recall respectively.
Poonam Patil et al 2015	Video retrieval by extracting key frames	Feature library	The method for elimination of near duplicate frame extraction number reached 30 % to 40% and elimination of redundant video frames using SVD method achieved accuracy up to 35% to 36%.
Benjamin Drayer et al 2016	Object detection, Tracking, and Motion Segmentation	YouTube, SegTrackv2, egoMotion, and FBMS	The method used to in experiment have shown good results for YouTube, SgeTrackv2, egoMotion authors conclude that FBMS performance do not achieve the sate-of-art performance due to the missing annotation of static object.
Jana Selvaganesan et al 2016	Unsupervised Feature Based key-Frame Extraction Towards Face Recognition	Honda/UCSD dataset1	Authors observed that the number of clusters formed is more when the threshold is less.
Shivanand S. Gornale et al 2018	Detection of ROI	Self-trained database	The method used in the experiment have shown good results, KNN with 89% accuracy whereas SVM with 90% accuracy
D. Asha et al 2018	Multiple feature CBVR	Google.video.URL http://Video.goole.com /	Their experiments obtained 80% of precision and recall rate. Authors concluded that the multiple feature CBVR system performance is better than the single feature CBVR system.

III. EXPERIMENTAL ANALYSIS

The experiment is carried out on our own dataset of 70 video streams captured from a 13-15 mega pixel mobile camera in .mp4 file formats. The video streams are captured in different durations like 30 second, 1 minute, 6 minutes, and 10 minutes of video with sampling rate of 30fps. The algorithm steps involved in the detection of ROI using Aggregated Channel Features (ACF) detector is as follows and the samples of the results are shown in the Fig2.

Input: Video with .mp4 format.
Output: Video detecting the region of interest.
Step1: Read an input video.
Step 2: Extraction of frames
Step3: Detection of ROI using ground-truth values.
Step4: With the help ground-truth values the object is being trained using ACF detector.
Step5: If the value of the object trained is greater than the considered threshold value then detect it as the valued object of interest.
Step6: Otherwise move to next frame.

As a result we observed that the system detects sign boards accurately using ground truth values with different time complexity for different duration of videos. The detection rate is good and satisfactory. It has been also noticed that the frame rate plays a vital role in the detection of the ROI for presented algorithm.
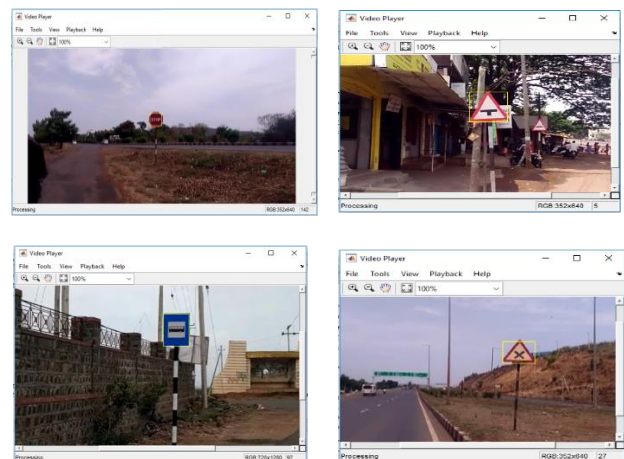

Fig.2. Detected ROI

IV. CONCLUSION

The comprehensive review of content based video retrieval system has been presented in this paper. As per the experimental analysis it is observed that the frame rate plays a vital role in the detection of the region of interest in a given video depending on the duration of video. The overall detection rate is good and satisfactory. Further the experiment will be carried out on retrieval and recognition of the sign boards.

REFERENCES

[1] Dipika H Patel, "Content-Based Video Retrieval using Enhanced Feature Extraction", *International Journal of Computer Applications* (0975-8887), (vol119),pp.4-8, 2015.

[2] Dr.H.B.Kekre, Dr. Dhirendra Mishra, Ms. P. R. Rege, "Survey on Recent Technique in Content Based Video Retrieval", *International Journal of Engineering and Technical Research (IJETR)*, (vol3), pp.69-73,2015.

[3] Weiming Hu, Nianhua Xie, Li Li, Xianglin Zeng, and Stephen Maybank , "A Survey on Visual Content-Based Video Indexing and Retrieval",*IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews* ,pp.797-819, 2011.

[4] B V Patel and B B Meshram, "Contented Based Video Retrieval System", *International Journal of UbiComp(IJU)*, (vol3), pp.13-3,2012.

[5] Robles, Oscar D, Pablo Toharia, Angel Rodríguez, and Luis Pastor. "Towards a content-based video retrieval system using wavelet-based signatures." *In 7th IASTED International Conference on Computer Graphics and Imaging-CGIM*, pp. 344-349. 2004.

[6] Zhong, D. and Chang, S.F., 1999. An integrated approach for content-based video object segmentation and retrieval. IEEE Transactions on Circuits and Systems for Video Technology, 9(8), pp.1259-1268.

[7] P.M.Kamde, Sankirti Shiravale, S. P. Algur et.al, "Entropy Supported Video Indexing For Content Based Video Retrieval", *International Journal of Computer Applications* (0975 – 8887), (vol62), pp.1-6, 2013.

[8] Milan R.Shetake. Sanjay. B. Waikar, "Content Based Image and Video Retrieval", *Proceedings of 33th IRF International Conference*, pp.34-37, 2015.

[9] Madhav Gitte, Harshal Bawaskar, Sourabh Sethi, Ajinkya Shinde, "Content BasedVideo Retrieval System", *IJRET: International Journal of Researchmin Engineering and Tecnology*, (vol3), pp.430-435,2014.

[10] Muhammad Nabeel Asghar, Fiaz Hussain, Rob Manton,"Video Indexing: Survey", *International Journal of computer an Information Technology*,(vol3),pp.148-169,2014.

[11] P.Geetha and Vasumathi Narayanan, "A survey of Content-Based Video Retrieval", *Journal of Computer Science*, (vol4), pp.474-486, 2008.

[12] Navdeep Kaur, Mandeep Singh,"Content-Based Video Retrieval with Frequency domain Analysis using 2- D Correlation Algorithm", *International Journal of Advanced Research in Computer Science and Software Engineering* ,(Vol4), pp.388-393,2014.

[13] Jin, Yonggang, and Farzin Mokhtarian. "Efficient Video Retrieval by Motion Trajectory." *In BMVC*, pp. 1-10. 2004.

[14] Truong, Ba Tu, and Svetha Venkatesh. "Video abstraction: A systematic review and classification." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 3, no. 1 (2007): 3.

[15] Avinash N Bhute and B B Meshram, "System Analysis and Design for Multimedia Retrieval Systems", *The International Journal of Multimedia and Its Applications (IJMA)*, (vol5), pp.25-44, 2013.

[16] Dyana, A. "Video Object Representation for Content-based Video Retrieval (CBVR)." *PhD diss., Indian Institute Of Technology, Madras*, 2010.

[17] Chao, Jianshu, Anas Al-Nuaimi, Georg Schroth, and Eckehard Steinbach. "Performance comparison of various feature detector-descriptor combinations for content-based image retrieval with JPEG-encoded query images." *In Multimedia Signal Processing (MMSP), 2013 IEEE 15th International Workshop on*, pp. 029-034. IEEE, 2013.

[18] Xiang Ma, Faisal Bashir, Ashfaq A. Khokhar, and Dan Schonfeld, "Event Analysis Based on Multiple Interactive Motion Trajectories", *IEEE Transaction on Circuits and System for Video Tecnology, University of Illinois at Chicago, Chicago*, IL 60607, U.S.A .pp.1-10, 2008.

[19] Benmokhtar, Rachid, Benoit Huet, Sid-Ahmed Berrani, and Patrick Lechat. "Video shots key-frames indexing and retrieval through pattern analysis and fusion techniques." *In Information Fusion, 2007 10th International Conference on*, pp. 1-6. IEEE, 2007.

[20] Nitin Jain, and Dr. S. S. Salankar, "Color & Texture Feature Extraction for Content Based Image Retrieval", *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)*, pp. 53-58,2014.

[21] Yasira Beevi C P et.al.2009, "An efficient Video Segmentation Algorithm with real time Adaptive Threshold Technique ", *International Journal of Signal Processing, Image Processing and Pattern Recognition* (vol2), pp.13-27

[22] Haitao Jiang, Abdelsalam (Sumi) Helal, Ahmed K. Elmagarmid, Anupam Joshi, "Scene change detection techniques for video database systems", ©Springer-Verlag, pp.186-195, 1998.

[23] Baris Sumengen, et.al. "Multi-Scale Edge Detection and Image Segmentation", *IEEE Signal processing Conference*, pp.1-4, 2005

[24] N Ikonomakis, K. N. Plataniots and A.N. Venetsanopoulos,"A Region Based Color Image Segmentation Scheme", *Proc.SPIE3653, Visual Communications and Image Processing* , (vol3653), 1999.

[25] Wu, Chuan, Yuwen He, Li Zhao, and Yuzhuo Zhong. "Motion feature extraction scheme for content-based video retrieval." *In Electronic Imaging 2002*, pp. 296-305. International Society for Optics and Photonics, 2001.

[26] Saber, Eli, A. Murat Tekalp, and Gozde Bozdagi. "Fusion of color and edge information for improved segmentation and edge linking." *Image and Vision Computing* 15.10 (1997): 769-780.

[27] Liu, Tianming, Hong-Jiang Zhang, and Feihu Qi. "A novel video key-frame-extraction algorithm based on perceived motion energy model." *IEEE transactions on circuits and systems for video technology* 13.10 (2003): 1006-1013.

[28] Mohanty, Ambika Ashirvad, Bipul Vaibhav, and Amit Sethi. "A frame-based decision pooling method for video classification." 2013 *Annual IEEE India Conference (INDICON). IEEE*, 2013.

[29] Huang, Mei, et al. "An Integrated Scheme for Video Key Frame Extraction."2nd *International Symposium on Computer, Communication, Control and Automation. Atlantis Press*, 2013.

[30] Kalpana Takre, Archana Rajurkar and Ramchandra Manthalkar "An Effective CBVR System Based On Motion, Quantized Color And Edge Density Features", *International Journal of Computer Science & Information Technology* , (vol3), pp.78-92, 2011.

[31] Janko calic, David P. Gibson, and Neill W. Campbell"Efficient Layout of Comic-Like Video Summaries", *IEEE Transactions On Circuits And Systems For Video Technology*, (vol17), pp.931-936, 2007.

[32] T.N.Shanmugam, and Priya Rajendran "An Enhanced Content-Based Video Retrieval System Based on Query

Clip", *International Journal of Research and Reviews in Applied Sciences*, (vol1), pp.236-253, 2009.

[33] Hong Jiang Zhang „Jianhua Wu, Di Zhong and Stephen W. Smoliar "An Integrated System For Content-Based Video Retrieval And Browsing",*Institute of System Science, National university of Singapore Pattern recoganition.* (vol30), pp.643-658, 1997.

[34] Philippe Salembier, Ferran Marques,Montse Pardas, Josep Ramon Morros, Isabelle Corset, Sylvie Jennin, Lionel Bouchard, Fernand Meyer, and Beatriz Marcotegui."Segmentation Based Video Coding System Allowing the Manipulation of Objects", *IEEE Transactions on Circuits and Systems for Video Technology*, pp.60-74, 1997.

[35] Chu-Hong Hoi, Wei Wang, and Michael R. Lyu, 2003, "A Novel Scheme for Video Similarity Detection", Department of Computer Science and Engineering, the Chinese University of Hong Kong ©Springer-Verlag Berlin Heidelberg. Pp.373–382, 2003.

[36] Yun Zhai, Alper Yilmaz, and Mubarak Shah, "Story Segmentation in News Videos using Visual and Text Cues", *School of Computer Science,University of Central Florida ©Springer-Verlag Berlin Heidelberg* , pp. 92–102, 2005.

[37] Mark Tabb, and Narendra Ahuja, "Multiscale Image Segmentation by Integrated Edge and Region Detection", IEEE Transactions On Image Processing, pp.642-655,1997

[38] M. Cooper, T. Liu, and E. Rieffel, "Video segmentation via temporal pattern classification," *Multimedia, IEEE Transactions on*, pp. 610–618, 2007.

[39] N. Peyrard, P Bouthemy, "Content–Based Video Segmentation using Statistical Motion Model", Multimedia Tools and Applications, 2005-Springer. *IRISA/INRIA, Campus universities de Beaulieu 35042 Rennes Cedex*, France, pp.527-536.

[40] Hari Sundaram and Shih-Fu Chang, "Video Scene Segmentation Using Video and Audio Features" *Dept. Of Electrical Engineering, Columbia University*, New York, New York 10027.

[41] Chandra Mani Sharma, Alok Kumar Singh Kushwaha, Rakesh Roshan, Rabins Porwal and Ashish Khare,"Intelligent Video Object Classification Scheme using Offline Feature Extraction and Machine Learning based Approach", *IJCSI International Journal of Computer Science Issues*, (vol9), pp.247-256, 2012.

[42] Shilpa Kamdi and R.K.Krishna, "Image Segmentation and Region Growing Algorithm", *International Journal of Computer Technology and Electronics Engineering (IJCTEE)* (vol2), pp.103-107.

[43] Werner Bailer, Harald Mayer, Helmat Neuschmied, Werner Haar, Mathias Lux, Werner Klieber, "Content-Based Video Retrieval and Summarization using MPEG-7", *Proc. Internet Imaging V*, San Jose, CA, USA, pp. 1-12.Janauary 2004.

[44] Guozhu Liu, and Junming Zhao, "Key Frame Extraction from MPEG Video Stream", *Proceedings of the Second Symposium International Computer Science and Computational Technology(ISCSCT '09)* Huangshan, P. R. China, pp.7-11, 2009.

[45] Chenliang Xu, Caiming Xiongy and Jason J. Corso, "Streaming Hierarchical Video Segmentation" pp.1-14, 2012.

[46] Hatim G. Zaini, and T. Frag, "Multi feature content based Video Retrieval Using High Level Semantic Concepts", *IPASJ International Journal of Computer Science (IIJCS),* (vol2), pp.15-27, 2014.

[47] Kin-Wai Sze, Kin-Man Lam, and Guoping Qiu, "A New Key Frame Representation for Video Segment Retrieval",*IEEE Transactions On Circuits And Systems For Video Technology*, (vol 15), 2005.

[48] Margarita Osadchy and Deniel Keren, "A Rejection-Based Method for Event Detection in Video", *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 14, No. 4.Pp.534-541.April 2004

[49] Aasif Ansari and Muzammil H Mohammed, "Content based Video Retrieval Systems - Methods, Techniques, Trends and Challenges" *International Journal of Computer Applications*, (vol112), pp.13-22, 2015.

[50] Li Zhao, Wei Qi, Stan Z. Li, Shi-Qiang Yang, H. J. Zhang, "Key-frame Extraction and Shot Retrieval Using Nearest Feature Line (NFL)", *Multimedia '00 proceedings of the 2000 ACM workshops on Multimedia*, USA@ 2000, pp.217-220.

[51] Laxmikant S. Kate, M.M.Waghmare, "A Survey on Content based Video Retrieval Using Speech and Text information", *International Journal of Science and Research (IJSR)* (vol3), pp.1152-1154, 2014.

[52] Tjondronegoro, Dian, Xiaohui Tao, Johannes Sasongko, and Cher Han Lau. "Multi-modal summarization of key events and top players in sports tournament videos." *In Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pp. 471-478. IEEE, 2011.

[53] Zhai, Yun, and Mubarak Shah. "Video scene segmentation using Markov chain Monte Carlo." IEEE Transactions on Multimedia 8.4 (2006): 686-697.

[54] John R. Smith, "VideoZoom Spatio-Temporal Video Browser", *IEEE Transactions on Multimedia*, pp.157-171, 1999.

[55] Tabii Youness, Sadiq Abdelalim, "Shot Boundary Detection In Videos Sequences Using Motion Activities", Advances in Multimedia - An International Journal (AMIJ), (vol5), pp.1-7, 2014.

[56] John S. Boreczky, Lawrence A. Rowe, "Comparison of video shot boundary detection techniques", Journal of Electronic Imaging, (vol5), pp.122–128, 1996.

[57] Jay K. Hackett, Mubarak Shah, "Segmentation Using *Intensity and Range Data", Optical Engineering,* (vol28), pp. 667-671, 1989.

[58] Mohini Deokar, Ruhi kabra, "Video Shot Detection Techniques Brief Overview", *International Journal of Engineering Research and General Science* (vol2), pp.817-820, 2014.

[59] Sanket Shinde and Girija Chiddarwar, "An Enhanced Strategy for Efficient Content Based Video Copy Detection", *International Journal of Computer Science and Information Technologies*, (Vol6), pp.3062-3066, 2015.

[60] Deepak C R, Sreehari S, Gokul M, Anuvind B, "Content Based Video Retrieval Using Cluster Overlapping", *International Journal of Computational Engineering Research*, (vol03), pp.104-108, 2013.

[61] Dubin, Ran, et al. "Real Time Video Quality Representation Classification of Encrypted HTTP Adaptive Video Streaming-the Case of Safari." arXiv preprint arXiv: 1602.00489, 2016.

[62] Ho, Yu-Hsuan, et al. "Fast coarse-to-fine video retrieval via shot-level statistics." *Visual Communications and Image Processing 2005. International Society for Optics and Photonics*, 2005.

[63] Yasuo Ariki, Masahito Kumano, Kiyoshi Tsukada, "Highlight Scene Extraction in Real Time from Baseball Live Video", Proceedings of the 5th ACM SIGMM

International workshop on Multimedia information retrieval - MIR '03, pp.209, 2003

[64] Yusuf Aytar, Mubarak Shah, Jiebo Luo, "Utilizing Semantic Word Similarity Measures for Video Retrieval" in *Computer Vision and Pattern Recognition, CVPR 2008. IEEE Conference*, pp. 1 – 8, 2008.

[65] F. I. Bashir, A. A. Khokhar, and D. Schonfeld, "Real-time motion trajectory-based indexing and retrieval of video sequences," *Multimedia, IEEE Transactions on*, pp.58–65, 2007

[66] Hamdy K. Elminir, Mohamed Abu ElSoud, Sahar F. Sabbeh, Aya Gamal "Multi feature content based video retrieval using high level semantic concept *IJCSI International Journal of Computer Science Issues*, (vol9), pp.254-260, 2012.

[67] L. Bai, S.-Y. Lao, H.-T. Liu, and J. Bu, "Video shot boundary detection using petri-net," in *Machine Learning and Cybernetics,International Conference on*, (vol5), pp. 3047–3051, 2008

[68] H. Lu, Y.-P. Tan, X. Xue, and L. Wu, "Shot boundary detection using unsupervised clustering and hypothesis testing," in *Communications, Circuits and Systems ICCCAS* 2004, International Conference on, (vol2), pp. 932–936, 2004.

[69] Guironnet, Mickaël, et al. "Video summarization based on camera motion and a subjective evaluation method." *EURASIP Journal on Image and Video Processing* 2007.1 (2007): 1-12.

[70] Padmakala, S., G. S. AnandhaMala, and M. Shalini. "An effective content based video retrieval utilizing texture, color and optimal key frame features."*Image Information Processing (ICIIP)*, 2011 International Conference on. IEEE, 2011.

[71] A. Paul, B.-W. Chen, K. Bharanitharan, and J.-F. Wang, "Video search and indexing with reinforcement agent for interactive multimedia services," *ACM Trans. Embed. Comput. Syst*, (vol12) pp. 25:1–25:16, 2013.

[72] Yildizer, Ela, Ali Metin Balci, Mohammad Hassan, and Reda Alhajj. "Efficient content-based image retrieval using multiple support vector machines ensemble." *Expert Systems with Applications* 39, pp.2385-2396, 2012.

[73] Z. Gu, T. Mei, X.-S. Hua, X. Wu, and S. Li, "EMS: Energy Minimization Based Video Scene Segmentation," *Multimedia and Expo, 2007 IEEE International Conference on*, pp. 520–523, 2007.

[74] Krishnamoorthy, R., & Braveen, M. "Object level frame comparison for video shot detection with orthogonal polynomials model" *In Emerging Trends in Science, Engineering and Technology (INCOSET),* 2012 International Conference on pp. 235-243. IEEE 2012.

[75] Yanqiang Lei, WeiqiLuo, Yuangen Wang, Jiwu Huang, "Video Sequence Matching Based On The Invariance Of Color Correlation", *IEEE Transactions On Circuits And Systems For Video Technology*, (vol22), 2012.

[76] Dr. H.B. Kekre, Dr. Sudeep D. Thepade and Saurabh Gupta, "Content Based Video Retrieval in Transformed *Domain using Fractional Coefficients", International Journal of Image Processing (IJIP)*, (vol7), 201.3

[77] B. V. Patel, A. V. Deorankar, B. B. Meshram, "Content Based Video Retrieval using Entropy, Edge Detection, Black and White Color Features", *Proc. 2nd International Conference on Computer Engineering and Technology (ICCET)*, (vol6), pp.V6-272 - V6-276, 2010.

[78] Chattopadhyay, Chiranjoy, and Sukhendu Das. "A motion-sketch based video retrieval using MST-CSS representation." *In Multimedia (ISM), 2012 IEEE International Symposium on*, pp. 376-377. IEEE, 2012.

[79] Droueche, Zakarya, Mathieu Lamard, Guy Cazuguel, Gwénolé Quellec, Christian Roux, and Béatrice Cochener. "Motion-based video retrieval with application to computer-assisted retinal surgery." *In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4962-4965. IEEE, 2012.

[80] Xingxiao Wu, Dong Xu, LixinDuan, JieboLuo, "Action Recognition Using MultilevelFeatures and Latent Structural SVM", *IEEE Transactions on Circuits and Systems for Video Technology*, (vol23), pp.1422-1431, 2013.

[81] Michele Merler, Bert Huang,LexingXie, Gang Hua, Apostol Natsev, "Semantic Model Vectors for Complex Video Event Recognition", *IEEE Transactions on Multimedia*, vol. 14, pp.88-101, 2012.

[82] Wang, X. and Zhang, X.P., 2013. An ICA mixture hidden conditional random field model for video event classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(1), pp.46-59.

[83] J. Fan, A. K. Elmagarmid, X. Zhu, W. G. Aref, and L. Wu, "Classview: hierarchical video shot classification, indexing, and accessing", Multimedia, IEEE Transactions , pp. 70–86, 2004.

[84] G. Xu, Y.F. Ma, H.-J. Zhang and S.Q. Yang, "An HMM-based framework for video semantic analysis", *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1422–1433, 2005.

[85] J. Tang, X.-S. Hua, M. Wang, Z. Gu, G.-J. Qi, and X. Wu. "Correlative linear neighborhood propagation for video annotation." *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics: a publication of the IEEE Systems, Man, and Cybernetics Society*, pp. 409–16, 2009.

[86] Sarah Joy, Lisha Kurian,"Video Forgery Detection Using Invariance of Color Correlation", *International Journals of Computer Science and Mobile Computing*, (vol3), pp.99 – 105,2014

[87] J. Bescos, "Real-time shot change detection over online mpeg-2 video," *IEEE Trans. on Circuits and Systems for Video Technology*, (vol14), pp.475–484, 2004.

[88] L. Zelnik-Manor and M. Irani, "Event-based video analysis", *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*'01, (vol2), pp.123–130, 2001.

[89] C. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," *in Proc. IEEE Int. Conf. on Computer Vision*, pp. 555-562, 1998.

[90] Thepade, Sudeep D, Krishnasagar S. Subhedarpage, and Ankur A. Mali. "Performance rise in Content Based Video retrieval using multi-level Thepade's sorted ternary Block Truncation Coding with intermediate block videos and even-odd videos." *In Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on*, pp. 962-966. IEEE, 2013..

[91] Nikita Sao & Ravi Mishra, "Video Shot Boundary Detection Based On Nodal Analysis of Graph Theoritic Approach", *International Journal of Management, Information Technology and Engineering (BEST: IJMITE)*, (vol2), pp.15-24, 2014.

[92] Meera R. Nath, "Video Sequence Matching Based On Multiclass SVM Classifiers and RGB Component Relations", *International Journal of Science and Research (IJSR)*, (vol3), pp.1938-1944, 2014.

[93] Brandt, Jens, Jens Trotzky, and Lars Wolf. "Fast Frame-Based Scene Change Detection in the Compressed Domain for MPEG-4 Video." In 2008 The Second International Conference on Next Generation Mobile

Applications, Services, and Technologies, pp. 514-520. IEEE, 2008.

[94] Wesolkowski, Slawo, and Paul W. Fieguth. "Color image segmentation using vector angle-based region growing." *9th Congress of the International Color Association. International Society for Optics and Photonics*, 2002.

[95] Azra Nasreen, Dr Shobha G, "Key Frame Extraction using Edge Change Ratio for Shot Segmentation", *International Journal of Advanced Research in Computer and Communication Engineering,* (vol2), pp.4421-4423, 2013.

[96] Zhong Qu, Lidan Lin, Tengfei Gao and Yongkun Wang, "An Improved Keyframe Extraction Method Based on HSV Colour Space", *Journal Of Software,* (vol8), pp.1751-1758, 2013

[97] Santhosh kumar K and Mallikarjuna Lingam.K, "Content Based Video Retrieval Using Low and High Level Semantic Gap", *International Journal for Research in Emerging Science and Technology*, (vol2), pp.149-153, 2015.

[98] S. B. Mahalakshmi, Capt. Dr.Santhosh Baboo, "Efficient Video Feature Extraction and Retrieval on Multimodal Search", *International Journal of Advanced Research in Computer and Communication Engineering*, (vol4), pp.355-357, 2015.

[99] David Asatryan, Manuk Zakaryan, "Novel Approach To Content-Based Video Indexing And Retrieval By Using A Measure Of Structural Similarity Of Frames", *International Journal, Information Content and Processing*, (vol2), pp.71-81, 2015.

[100] Miss. Poonam V. Patil, Prof. S. V. Bodake, "Video Retrieval by Extracting Key Frames in CBVR System", *International Journal of Innovative Research in Science, Engineering and Technology*, (vol4), pp.11920-11926, 2015.

[101] Lin, Tong, and Hong-Jiang Zhang. "Automatic video scene extraction by shot grouping." *In Pattern Recognition, 2000. Proceedings. 15th International Conference on*, pp.39-42. IEEE, 2000.

[102] Benjamin Drayer and Thomas Brox, "Object detection, Tracking, and Motion Segmentation for Object-level Video Segmentation", *Computer Vision and Pattern Recognition*, pp:1-17 2016.

[103] Jana selvaganesan and Kannan Natarajan, "Unsupervised Feature Based Key-Frame extraction Towards Face Recognition", *The International Arab Journal of Information Technology*, (vol13),pp: 777-783, 2016.

[104] Shivanand S Gornale, Ashvini K Babaleshwar, Pravin L Yannawar," Detection and Classification of Signage's from Random Mobile Videos Using Local Binary Patterns", *International Journal of Image, Graphics and Signal Processing(IJIGSP)*, Vol.10, No.2, pp. 52-59, 2018.

[105] D.Asha, Madhavee Lata, V.S.K. Reddy, "Content based video retrieval system using Multiple Features", *International Journal of pure and Applied Mathematics*, (vol. 118), pp.287-294, 2018.

[106] Samabia Tehsin, Asif Masood and Sumaira Kausar, "Survey of Region-Based Text Extraction Techniques for Efficient Indexing of Image/Video Retrieval", *I.J.Image, Graphics and Signal Processing*, No.12, pp.53-64, 2014.

Authors' Profiles

Dr. Shivanand S. Gornale has obtained M. Sc. in Computer Science. M.Phil. in Computer Science., Ph.D. in Computer Science from Savitribai Phule Pune University , Pune Maharashtra, India in 2009 under the guidance of Dr. K V Kale and has been recognized as research guide for PhD in Computer Science and Engineering from Rani Channamma University, Belagavi and Jain University Bangalore. He has published more than 85 research papers in various National and Inter-national Journals and conferences. He is a Fellow of IETE New Delhi, Life Member of CSI, Life Member of Indian Unit of Pattern Recognition and Artificial Intelligence (IPRA), Member of Indian Association for Research in Computer Science (IARCS), Member of International Association of Computer Science and Information Technology (IACS&IT) Singapore, Member of International Association for Engineers', Hong Kong, Member of Computer Science Teachers' Association, USA, Life Member of Indian Science Congress Association, Kolkata-India. Presently he is working as Professor and Chairman, Department of Computer Science, Rani Channamma University, Belagavi – Karnataka, India. His research area of interest is Digital Image Processing, Pattern Recognition, Computer Vision and Machine Learning, Video Retrieval and Biometric analysis.

Ashvini K. Babaleshwar completed M.Sc in Computer Science from Rani Channamma University Belagavi, Karnataka, India in 2013. She is pursuing Ph.D in Computer Science at Rani Channamma University Belagavi, Karnataka, India. Her area of research includes Image Processing and Pattern recognition, Content based Image and Video retrieval.

Pravin L. Yannawar is Assistant Professor in Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS). He has completed M. Sc. (2001), SET (Computer Science & application) Ph. D (2011). He is a Member of IETE & life member of IAEng, CSTA, IUPRAI and IACSIT. He is recipient of DST Fast Track Young Scientist research project. His area of research includes Vision, Intelligent Systems, AVSR, OCR, Image Processing and pattern recognition.