Video Summarization using Adaptive Thresholding by Machine Learning for Distributed Cloud Storage

T. Naga Raja, Dr V.V. Venkata Ramana, A. Damodharam,

Abstract*: The growth in the video content-based communication and information management have motivated the applications dealing with the video contents to be placed on cloud-based data -storages, which are associated with datacentres. The applications ranging from social media to distance education and surveillance to business communication or any applications related to the e-governess are considering the video data as one of the most preferred mode of communication. The video data is enriched with audio and visual contents, which makes analysis or expressing information high convenient. Computerized video is an electronic portrayal of moving visual pictures as encoded advanced information. This is as opposed to simple video, which speaks to moving visual pictures with a simple sign. The advanced video contains a progression of computerized pictures showed in quick progression. The number of applications, as mentioned, dealing with video data is increasing and as a result a large amount of video content is generated every day. Hence, the complexity of retrieving the information from the video contents are also increasing. The bottleneck of the video retrieval process is for a lower sized segment of the video content can be retrieved in low time complexity and if the information to be preserved to a higher extend, then the retrieval time complexity increases. Thus, a good number of parallel researchers have introduced various methods for video content summarization and retrieval using summarization with efficient searching methods, but most of the parallel research outcomes are criticized for either higher time complexity or for higher information loss. This problem can be ideally solved by finding the highly accurate ratio of key information video frames from the total video content. Henceforth, this work, presents a novel machine learning method for identifying the key frames, not only based on the information available in the frame, also validating the key frames with the thresholds of the objects or changes in the frame. The work is again enhanced by considering the adaptive thresholding method for distributed and collaborative video information. The measures taken in the proposed algorithm produces a 98% accuracy for video information representation and nearly 99% reduction in the video frames, which results into nearly 99% reduction in the time complexity.

***Index Terms*: Key Frame, Adaptive Threshold, Segment Noise Reduction, Video Content Retrieval, Complexity Reduction.**

I. INTRODUCTION

The use of surveillance robots, distance education, video communication based medical diagnosis and business communication using video as a primary channel for communication have increased in the last few decades. The work by Y. Guo et al. [1] have demonstrated the use of video feeds as cyber security mechanism and the challenges of handling large amount of video information for various devices. The primary challenge is the complexity of the algorithms and the time required to extract complex information from the video data. The similar finding was reported by K. Simonyan et al. [2] for verification and dection of the actions from the video data feed. This work proposed a newer direction in the research by introducing a completely newfangled network to handle the storage and retrieval of the video data. This work was soon getting criticism from across the research community for higher cost of network upgradations. In the other hand, some of the work as the work by X. Hu et al. [3] have focused on the work by K. Simonyan et al. [2] and aimed to build more and more sophisticated networks for video data processing and communications. Nevertheless, the primary objective of these research attempts where to reduce the time complexity for the video data processing.The researchers were soon to identity the bottleneck of the time complexity and started working towards the selection of the key information from the video data as recommend by J. R. R. Uijlings et al. [4] with the recommendation of generic content-based frame extraction for video summarizations [Fig – 1].

Revised Manuscript Received on December 22, 2018.

T. Naga Raja,Research Scholar,JNTUH, Hyderabad, India
rajatnaga@yahoo.com .
Dr V.V.Venkata Ramana,Scientist – F, NIC,Vijayawada , India
v3ramana@gmail.com.
Prof A. Damodharam,Adviser to the Govt of Andhra Pradesh
Amaravathi, India,damodarama@rediffmail.com

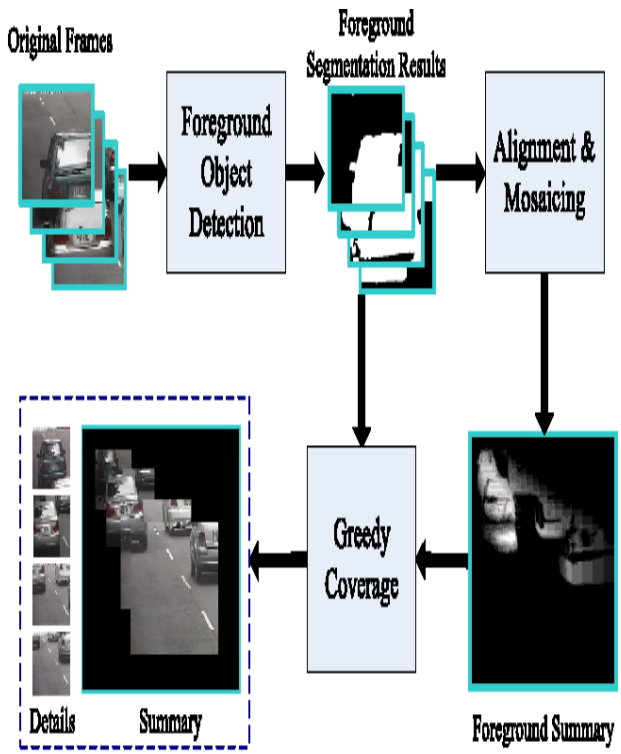

Fig. 1 Generic Video Summarization Process

The computation complexity was not the only challenge identified by the researchers, also the storage complexity is to be considered for this direction of the research. The work by X. Hu et al. [5] have suggested to migrate the storage to the distributed cloud storage segments. This solution was widely accepted by the practitioners as most of the applications dealing with the video contents either producing or consuming the video information are already on cloud [Fig – 2].

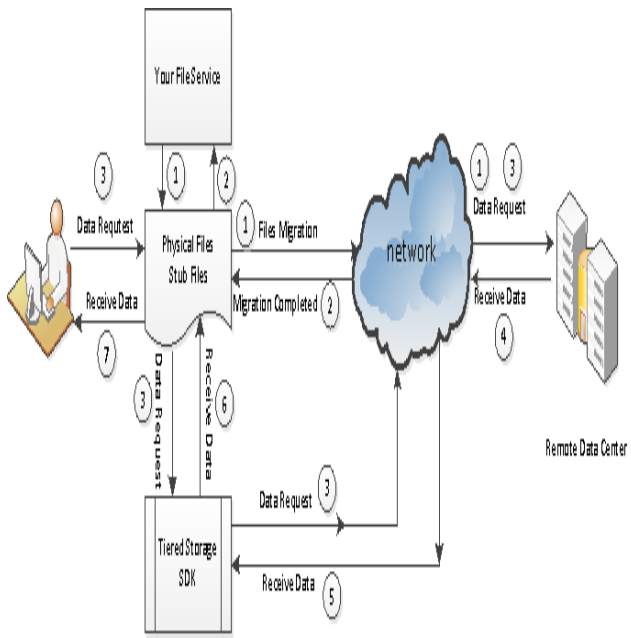

Fig. 2 Distributed Cloud-Based Storage for Video Contents

Henceforth, based on various research proposal and outcomes, this work identifies the following challenges to be solved for the improvements in the recent research outcomes.

- The discovery of the basic content elements of a multimedia document (equivalent to words in a text document) is a complex process identified in the outcomes of the parallel works.
- Thus, reducing the programming, space and time complexity is the demand from the present researches.
- Reducing the elements to the limited set of key content elements, which can be seen as multimedia keywords, is the second most important factor.
- Nevertheless, majority of the recent work outcomes have demonstrated the reduction of key words based on language recommendations rather the context recommendations. This leads to non-context aware reduction and further leads to incorrect extraction of keywords.
- Henceforth, the demand from the recent research is to identify the key words based on the context.

This work not only identifies the research challenges, but also provides solutions to these identified bottlenecks in the further sections of this work.

II. FUNDAMENTALS OF VIDEO SUMMARIZATION ON DISTRIBUTED STORAGE

In this section of the work, the fundamentals of video summarization for any distributed architecture similar to cloud are analysed. A distributed information store is an arrangement where data is put away on more than one hub, frequently in a reproduced design. It is normally explicitly used to allude to either a distributed database where clients store data on various hubs, or organization in which clients store data on various companion arrange hubs. Distributed cloud databases are typically non-relational databases that empower a brisk access to information over an enormous number of hubs. Some distributed storages uncover rich inquiry capacities while others are constrained to a key-esteem store semantics. Firstly, assuming that, the total video library or the video contents, V[], are distributed over multiple nodes of the cloud storage and each component or information, V_x, in the video library are part of the total video library. This relation can be formulated as,

$$V[] \leftarrow \sum_{i=1}^{n} V_i \qquad \text{(Eq.1)}$$

Where, n is the total number of video contents distributed over the cloud storages. These cloud storages are not independent and must be part of the instance set, I[], where each and every instance, I_x, can be identified uniquely. Hence, this relation can be presented as,

$$I[] \leftarrow \sum_{j=1}^{m} I_j \qquad \text{(Eq.2)}$$

Where, m denotes the total number of instances on a cloud-based datacentre. Further, each datacentre instance is a combination of computing capacity, C, storage capacity, S, memory capacity, M, and finally the network bandwidth, N. This combination can be represented as,

$$I_x \subseteq \oint(C, S, M, N) \qquad \text{(Eq.3)}$$

Nonetheless, each component in this collection is also a set of collaborated capacities as one instance can spread over multiple physical hardware infrastructures. Hence, Eq. 3 can be re-written as,

$$I_x \subseteq \oint(C[], S[], M[], N[]) \quad \text{(Eq.4)}$$

It is not a rare situation, where, the storage set, S[], is distributed over multiple instances and any instance, I_x, can be also mapped with many storage sets. This property can be visualized as,

$$I_x \rightarrow [S_x, S_y] \quad \text{(Eq.5)}$$

Or,

$$S_x : [I_x, I_y] \quad \text{(Eq.6)}$$

In either of the situations, the video components, V_x, can be distributed over multiple storage segments as,

$$V_x \rightarrow [S_x, S_y] \quad \text{(Eq.7)}$$

Further, for all video contents, the set of storage components must be retrieved. This can be represented as function, to which the input is the video component and the result is the storage sets,

$$\oint(V_x) = \|S[]\| \quad \text{(Eq.8)}$$

Once, all the components from each storage segment is retrieved, the duplicated contents must be reduced, and the summarized video can be reported. During this process, the identification of the duplicated information in the video data, can be identified using different methods. Nonetheless, the accuracy of this process will generate the most information video summery. The recent advancements and the challenges of this process is again identified in the next section of the work.

III. PARALLEL RESEARCH OUTCOMES

The fundamental process of video summarization is inferred in the previous section of this work. This accumulated knowledge will be the guardrail to realize the bottlenecks on the parallel research advancements in the domain of study.The prime focus of the parallel recent research attempts was to reduce the key frames with any information loss. Thus, the first attempt by D. Oneata et al. [6] aimed to detect the objects in the video data contents and based on the data available, perform selection of the key sections of the video data and produce the segments as key frames. This work gained a lot of attention as this was the first attempt to reduce the size of the video content in a very meaningful process. Nevertheless, this work was criticised by many researchers for increasing the time complexity to a greater extend for the content retrieval process. Moreover, this process was only applicable for the static video contents on a local storage. The similar approach was also carried out by M. Van den Bergh et al. [7] for the distributed contents. But, due to the lack of duplicate information detection techniques in that research, the work was soon started to get reattempted by various other researchers. During the object-based video key frame detection, the other challenges were noticed by the practitioners. During the video capture process, the object, the capture device or both can be in motion and can cause turbulence in the frames. The work by M. Jain et al. [8] have identified the action-based detection of objects during these situations. Further, this work was improvised by P. Weinzaepfel et al. [9] using frame of reference principle to identify the actual location of the objects in the frame with relative reference method. In the other hand, another group of researchers were aiming to reduce the time complexity with the help of high-performance computing methods. The work by J. Zhang et al. [10] have demonstrated the distribution of the computation load during such scenarios can reduce the energy consumption for the over all process of information retrieval. Using the distributed load computing, yet another challenge was observed. During the distribution of the work loads, the processing systems were adequately capable to understand the synchronization. But, during the summarization process, the work loads were bound to be failed. The solution to this problem was delivered in the work by S. Hare et al. [11]. This work was further improvised by H. Wang et al. [12] using the motion of the objects in the frames, which are extracted from the video data. Further, with the understanding of the parallel research outcomes and the limitations, this work formulates the actual research problem in the next section of this work.

IV. PROBLEM FORMULATION

With the detailed analysis of the parallel research outcomes, in this section of the work the problem is formulated with the help of mathematical lemma.

Lemma: During the key frame extraction process, considering the adaptive localized threshold improves the accuracy and reduces the time complexity of the key frame extraction process.

Proof: Assuming that the video content, V, is distributed in multiple segments of storage, S, with unique or replicable video contents, each identified as V_x. Hence, the relation between the video contents and the storage segments can be realized as,

$$V \leftarrow \sum_{x=1}^{n} V_x \quad \text{(Eq.9)}$$

And,

$$S \leftarrow \sum_{k=1}^{m} S_k \quad \text{(Eq.10)}$$

Where, n and m denote number of distributed video contents and number of storage segments respectively. Considering any two storage segments, S_x and S_y, contain two different video contents as V_i and V_j. Thus, the relation can be formulated as,

$$S_x : [V_i, V_j] \quad \text{(Eq.11)}$$

And,

$$S_y : [V_i, V_j] \quad \text{(Eq.12)}$$

Conversely,

$$V_i : [S_x, S_y] \quad \text{(Eq.13)}$$

And,

$$V_j : [S_x, S_y] \quad \text{(Eq.14)}$$

Further, the local threshold, TH_X, for each video content can be calculated as a function of Frames Per Second, FPS, Scan System method, SS, Aspect

Ratio, AR, Augmented Aspect Ratio, AAR and finally, Channel information as C. This can be formulated as,

$$\oint V_i[](FPS_i, SS_i, AR_i, AAR_i, C_i) \Rightarrow TH_{V_i} \quad \text{(Eq.15)}$$

And,

$$\oint V_j[](FPS_j, SS_j, AR_j, AAR_j, C_j) \Rightarrow TH_{V_i} \quad \text{(Eq.16)}$$

Hence, the video content based on key frames retrieved from each storage segment can be formulated as,

$$\oint TH_{V_i} :: S_x \rightarrow V_i[S_x]$$
$$\oint TH_{V_i} :: S_y \rightarrow V_i[S_y] \quad \text{(Eq.17)}$$

And,

$$\oint TH_{V_i} :: S_x \rightarrow V_j[S_x]$$
$$\oint TH_{V_i} :: S_y \rightarrow V_j[S_y] \quad \text{(Eq.18)}$$

Here, the accuracy of the key frame extraction or the accuracy of video summarization is completely depending on the threshold of the local video content. Thus, the extraction of key frame and the key frame counts will differ for each part of the video content and cannot generate an optimized set of key frames. Thus, the consideration of the adaptive threshold, TH_A, is introduced here.

The adaptive threshold can be calculated as,

$$TH_A = \frac{(\sum_{i=1}^{n} \oint TH_i) + (\sum_{i=1}^{n} \frac{\partial^2 TH_i}{\partial TH_i^2})}{\frac{\Delta}{\Delta \sum_{i=1}^{n} \oint TH(V_x)}(TH_i)} \quad \text{(Eq.19)}$$

Henceforth, the video key frame extraction will now have dependencies from each video distributed over each storage segments and the adaptive threshold will be considered as an overall threshold for a specific video content. This certainly improves the accuracy of the key frame extraction process.

$$\oint TH_{AV_i} :: S_x \rightarrow V_i[S_x]$$
$$\oint TH_{AV_i} :: S_y \rightarrow V_i[S_y] \quad \text{(Eq.20)}$$

And,

$$\oint TH_{AV_i} :: S_x \rightarrow V_j[S_x]$$
$$\oint TH_{AV_i} :: S_y \rightarrow V_j[S_y] \quad \text{(Eq.21)}$$

Also, regarding the time complexity, as each calculation of the threshold is now reduced to an adaptive threshold, hence the time taken for each extraction process will be reduced in each iteration of the extraction process. Elaborating this thought, assuming that each key frame extraction or calculating the threshold takes t time complexity. Hence for a total n number of video contents the total time complexity, T, can be represented as,

$$T = t^n \quad \text{(Eq.22)}$$

Conversely, the overall time complexity, T', will reduce for each iteration in case of adaptive thresholding method, as,

$$T' = (t-n)! \quad \text{(Eq.23)}$$

It is natural to realize that,

$$T' << T \quad \text{(Eq.24)}$$

Hence, this can be concluded that, the adaptive thresholding process can reduce the time complexity for key frame extraction process with reduced number of key frames and increase the accuracy of the extraction process. Furthermore, based on this principle, the proposed adaptive threshold-based algorithms is furnished in the next section of this work.

V. PROPOSED NOVEL ADAPTIVE THRESHOLDING ALGORITHM USING MACHINE LEARNING

After the formulation of the problem and confirmation to the fact that adaptive thresholding can increase the key frame extraction accuracy and reduce the time complexity, in this section of the work, the proposed set of algorithms are furnished with detailed discussion. For the success of any video content analysis framework or algorithm, it is the prime task to reduce the noise available in the video signal. Hence, firstly, the noise reduction algorithm is furnished here.

Algorithm - 1: **Segmented Noise Removal Processing using Deep Segmentation** Algorithm (SNRP-DS)
1. Accept all the video content as V[] 2. For each V[i] a. Accept the video signal in analog form as f(t):V[i] b. For each time interval t/Δt, i. Calculate the signal variance as Amp(t/Δt) = (f(t/Δt):V[i] ii. Revise the variance difference, AmpDiff(t/Δt) = (|Amp(t/Δt) - Amp(t+1/Δt)|)/(t/Δt) iii. Report the overall signal variance, Amp[] c. Reduce the noise, V[i] - Amp[] 3. Report the noise removed video content as V'[]

In sign preparing, noise is a general term for undesirable changes that a sign may endure amid catch, stockpiling, transmission, handling, or transformation. Now and again the word is additionally used to mean flag that are irregular and convey no valuable data; regardless of whether they are not meddling with different flag or may have been presented purposefully, as in solace noise. Noise decrease, the recuperation of the first sign from the noise-adulterated one, is a shared objective in the structure of sign preparing frameworks, particularly channels. As far as possible for noise expulsion are set by data hypothesis, to be specific the Nyquist– Shannon testing hypothesis.Secondly, the adaptive key frame extraction algorithm is furnished here.

Algorithm - 2: **Adaptive Threshold Calculation for Distributed Video Storage** Algorithm (ATC-DVS)

1. Accept the list of storage instances as S[]
2. For each instance S[i]
    a. Accept the list of noise reduced video content set as V'[]
        i. For each V'[i]
            1. Extract the video meta data as
                a. Frames Per Second, FPS
                b. Scan System, SS
                c. Aspect Ratio, AR
                d. Augmented Aspect Ratio, AAR
                e. Channel information, C
            2. Calculate the key frame extraction threshold as f(FPS, SS, AR, AAR, C)
            3. Scan V'[i] for each frame V'F[t]
            4. If f(V'F[t])>=f(V'[i]) and V'F[t](AAR)> V'[i](AAR)
            5. Then,
                a.V'[i]_KeyFrame[j] = V'F[t]
            6. Else,
                a.Discard V'F[t]
        ii. Report the total key frames for V'[i]
    b. Calculate the overall threshold as, f(S[i]V'[i])
    c. If f(S[i]V'F[t]))< f(S[i]V'[i])
    d. Then,
        i. Discard the key frames in V'[i]_KeyFrame[j]
    e. Else,
        i. Report the final key frames for V'[i]

In video pressure, a keyframe, otherwise called an "intra-frame", is a frame wherein a total picture is put away in the information stream. In video pressure, just changes that happen starting with one frame then onto the next are put away in the information stream, so as to extraordinarily lessen the measure of data that must be put away. This method exploits the way that most video sources have just little changes in the picture starting with one frame then onto the next. At whatever point an intense change to the picture happens, for example, when changing starting with one camera shot then onto the next, or at a scene change a keyframe must be made. The whole picture for the frame must be yield when the visual distinction between the two frames is great to the point that speaking to the new picture gradually from the past frame would require a greater number of information than reproducing the entire picture.Further, the results obtained from the proposed novel algorithms are discussed in the next section of the work.

VI. RESULTS AND DISCUSSION

The obtained results from the proposed algorithms are highly satisfactory and are discussed here. The results are divided into four different segments for discussions.

A. Storage Instance wise Video Content Threshold

Firstly, the storage segment or instance wise video content threshold retrieval results are discussed [Table – 1].

TABLE I
VIDEO CONTENT THRESHOLDS FOR EACH INSTANCES

Dataset	Storage Instance	Threshold
TRECVID Set - 1	inst_00000	42662.4
TRECVID Set - 1	inst_00001	24188.8
TRECVID Set - 1	inst_00002	79675
TRECVID Set - 1	inst_00003	61483.4
TRECVID Set - 1	inst_00004	52466.6
TRECVID Set - 1	inst_00005	73501
TRECVID Set - 1	inst_00006	35118.6
TRECVID Set - 1	inst_00007	48260.8
TRECVID Set - 1	inst_00008	177926.2
TRECVID Set - 1	inst_00009	183773.4
TRECVID Set - 1	inst_00010	42624
TRECVID Set - 1	inst_00011	26079.6
TRECVID Set - 1	inst_00012	81120
TRECVID Set - 1	inst_00013	61503.8
TRECVID Set - 1	inst_00014	52615.2
TRECVID Set - 1	inst_00015	73415.4
TRECVID Set - 1	inst_00016	35350.4
TRECVID Set - 1	inst_00017	48531.8
TRECVID Set - 2	inst_00000	50951.8
TRECVID Set - 2	inst_00001	55420.2
TRECVID Set - 2	inst_00002	66351.6
TRECVID Set - 2	inst_00003	63382.2
TRECVID Set - 2	inst_00004	62505.2
TRECVID Set - 2	inst_00005	51672
TRECVID Set - 2	inst_00006	51870
TRECVID Set - 2	inst_00007	62603.8
TRECVID Set - 2	inst_00008	59420
TRECVID Set - 2	inst_00009	55687.4
TRECVID Set - 2	inst_00010	64763.6
TRECVID Set - 2	inst_00011	60022.2
TRECVID Set - 2	inst_00012	87881.8
TRECVID Set - 2	inst_00013	65095.4
TRECVID Set - 2	inst_00014	56855.4
TRECVID Set - 2	inst_00015	62915
TRECVID Set - 2	inst_00016	56276.2
TRECVID Set - 2	inst_00017	77531.6
TRECVID Set - 2	inst_00018	51590.2
TRECVID Set - 2	inst_00019	51151.8
TRECVID Set - 2	inst_00020	49646.4
TRECVID Set - 2	inst_00021	64883
TRECVID Set - 2	inst_00022	51489.2
TRECVID Set - 2	inst_00023	65630
TRECVID Set - 2	inst_00024	53329.6
TRECVID Set - 2	inst_00025	56441.8
TRECVID Set - 2	inst_00026	54976.8
TRECVID Set - 2	inst_00027	54108.4
TRECVID Set - 2	inst_00028	51622.4
TRECVID Set - 2	inst_00029	61833.2
TRECVID Set - 2	inst_00030	63464.8
TRECVID Set - 2	inst_00031	51893.2
TRECVID Set - 2	inst_00032	50981.6
TRECVID Set - 2	inst_00033	53680.4
TRECVID Set - 2	inst_00034	53539.4
TRECVID Set - 2	inst_00035	74830.6
TRECVID Set - 2	inst_00036	81664.2
TRECVID Set - 2	inst_00037	56241.8
TRECVID Set - 2	inst_00038	55984.2
TRECVID Set - 2	inst_00039	53849.4
TRECVID Set - 2	inst_00040	52354.4

TRECVID Set - 2	inst_00041	50051.2
TRECVID Set - 2	inst_00042	50346.6
TRECVID Set - 2	inst_00043	50344.8
TRECVID Set - 2	inst_00044	49255.8
TRECVID Set - 2	inst_00045	57522
TRECVID Set - 2	inst_00046	51424.2
TRECVID Set - 2	inst_00047	52808.8
TRECVID Set - 2	inst_00048	71580.6
TRECVID Set - 2	inst_00049	71465.4
TRECVID Set - 2	inst_00050	70670.8
TRECVID Set - 2	inst_00051	81748
TRECVID Set - 2	inst_00052	70451.6
TRECVID Set - 2	inst_00053	72074.8
TRECVID Set - 2	inst_00054	82590.2
TRECVID Set - 2	inst_00055	83250.4
TRECVID Set - 2	inst_00056	54769.2
TRECVID Set - 2	inst_00057	54309.4
TRECVID Set - 2	inst_00058	53121.4
TRECVID Set - 2	inst_00059	49509.4
TRECVID Set - 2	inst_00060	73666
TRECVID Set - 2	inst_00061	62066.4
TRECVID Set - 2	inst_00062	61067.2
BBC Motion Gallery Set - 1	inst_00000	37164.6
BBC Motion Gallery Set - 1	inst_00001	52662.8
BBC Motion Gallery Set - 1	inst_00002	105018.6
BBC Motion Gallery Set - 1	inst_00003	27366.8
BBC Motion Gallery Set - 1	inst_00004	27631
BBC Motion Gallery Set - 1	inst_00005	25311.6
BBC Motion Gallery Set - 1	inst_00006	22445.8
BBC Motion Gallery Set - 1	inst_00007	20899
BBC Motion Gallery Set - 1	inst_00008	23406.4
BBC Motion Gallery Set - 1	inst_00009	25719
BBC Motion Gallery Set - 1	inst_00010	28732.6
BBC Motion Gallery Set - 1	inst_00011	23321.2
BBC Motion Gallery Set - 1	inst_00012	70366.8
BBC Motion Gallery Set - 2	inst_00000	111289.4
BBC Motion Gallery Set - 2	inst_00001	99606
BBC Motion Gallery Set - 2	inst_00002	56443.6
BBC Motion Gallery Set - 3	inst_00000	23434.6
BBC Motion Gallery Set - 3	inst_00001	26897.2

Hence, it is to be observed that the total video library is consisting of 5 different video datasets and spread over multiple storage instances.
The result is visualized graphically here [Fig – 3].

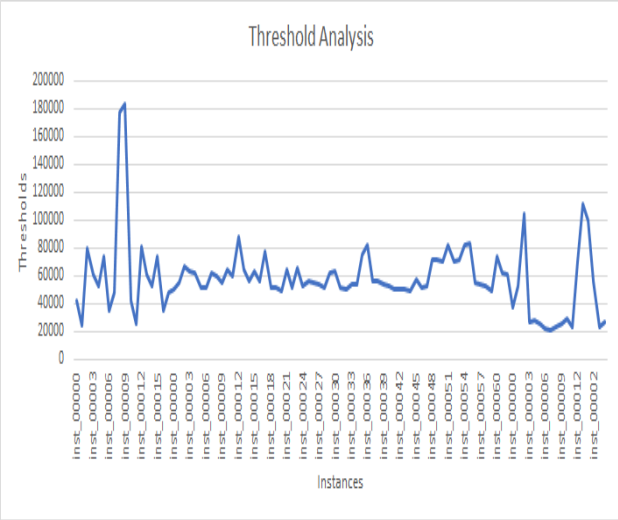

Fig. 3 Storage Instance wise Video Content Threshold Analysis

B. Key Frame Extraction Results and Comparative Analysis

Secondly, the key frame extraction results are elaborated here and compared with the standard method for key frame extraction [Table – 2].

TABLE II
KEY FRAME EXTRACTION RESULTS WITH COMPARATIVE ANALYSIS

Dataset	Total number of key frames (Legacy Method)	Extracted Key Frames (Proposed Method)	Reduction (%)
TRECVID Set - 1	968	18	98.14049587
TRECVID Set - 2	4037	63	98.43943522
BBC Motion Gallery Set - 1	3958	13	99.67155129
BBC Motion Gallery Set - 2	2516	3	99.88076312
BBC Motion Gallery Set - 3	3007	2	99.93348853

Here the improvements are to be noted as nearly 99% of reduction in terms of key frames.

The result is visualized graphically here [Fig – 4].

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

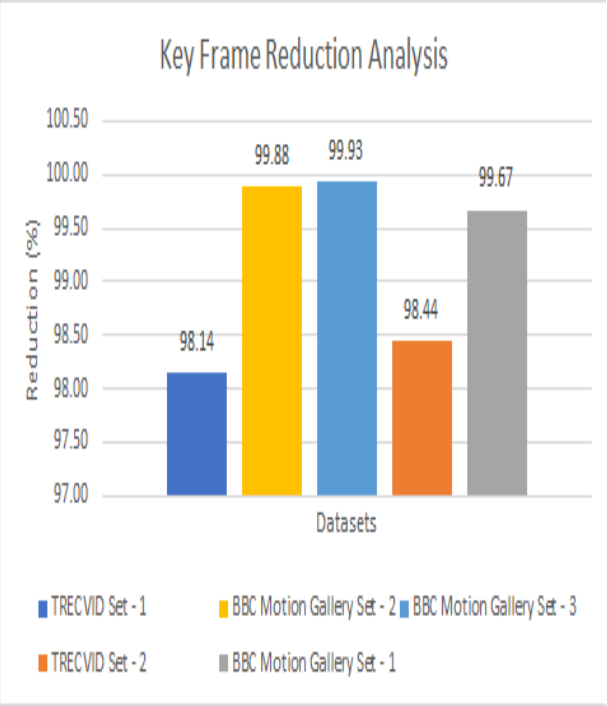

Fig. 4 Key Frame Reduction Analysis

C. Content Extraction Results and Comparative Analysis

Thirdly, the accuracy of the content extraction results is discussed here with the comparative analysis [Table – 3]. It is natural to realize that, the reduction in the key frame can cause loss of information in the video content summarization process or information retrieval from the video contents. Hence, this analysis is highly important to realize the benefits from key frame reduction.

TABLE III
CONTENT SUMMARIZATION / INFORMATION RETRIEVAL ACCURACY ANALYSIS

Dataset	Mean Threshold (Legacy Method)	Over All Threshold (Proposed Method)	Std Dev. Threshold	Accuracy (%)
TRECVID Set - 1	16276.35	173026.6642	107921.2578	96
TRECVID Set - 2	48999.36	290369.5432	94372.09229	99.07401
BBC Motion Gallery Set - 1	18627.27	102148.2045	27639.13416	99.5162
BBC Motion Gallery Set - 2	12797.56	80091.22189	28900.99196	98.74168
BBC Motion Gallery Set - 3	7462.01	43504.44994	13656.40633	99.16988

Here the improvements are to be noted as nearly 98% of accuracy in terms of content summarization process or information retrieval process.

The result is visualized graphically here [Fig – 5].

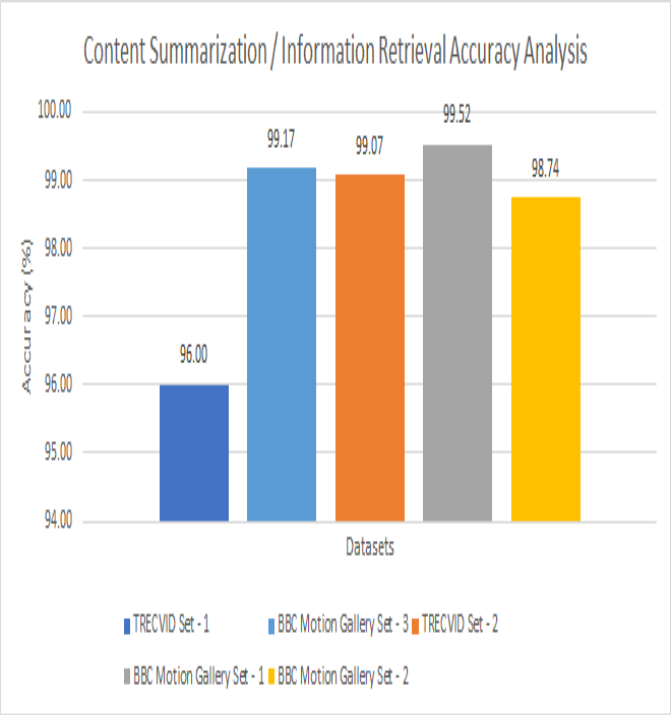

Fig. 5 Content Summarization / Information Retrieval Accuracy Analysis

D. Time Complexity Results and Comparative Analysis

Finally, the time complexity analysis results are discussed and compared with the legacy systems [Table – 4].

TABLE IV
TIME COMPLEXITY RESULTS WITH COMPARATIVE ANALYSIS

Dataset	Actual Time (Sec) (Legacy Method)	Reduced Time (Sec) (Proposed Method)	Improvements (%)
TRECVID Set - 1	40820.56	759.06	98.14049587
TRECVID Set - 2	203720.4767	3179.19	98.43943522
BBC Motion Gallery Set - 1	937312.2477	3078.59	99.67155129
BBC Motion Gallery Set - 2	1619658.227	1931.23	99.88076312
BBC Motion Gallery Set - 3	842997.415	560.69	99.93348853

Here the improvements are to be noted as nearly 99% of reduction in terms of time complexity for key frame extraction.

The result is visualized graphically here [Fig – 6].

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

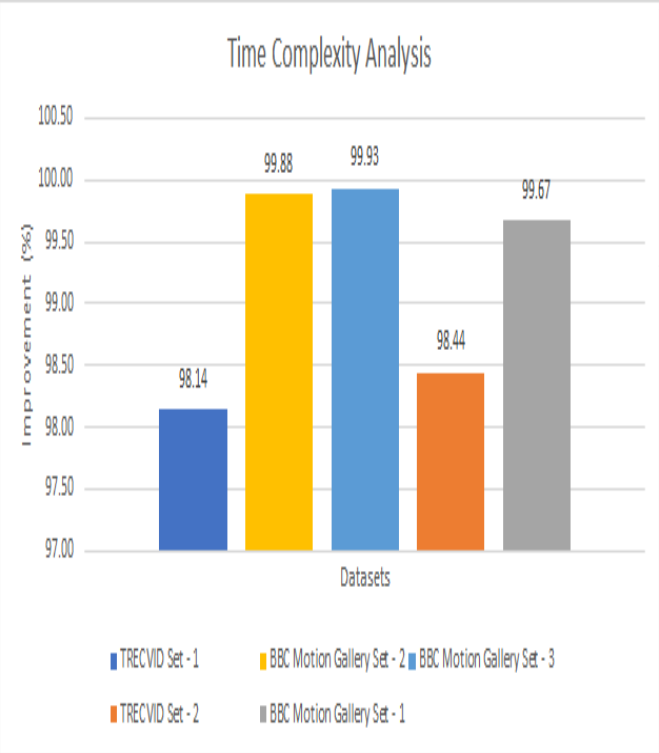

Fig. 6 Time Complexity Improvement Analysis

Henceforth, with the detailed analysis of the results, the work presents the final research conclusion in the next section of the work.

VII. CONCLUSION

Growth in the video content generation, use and storage have motivated researchers to carry out the multiple novel research attempts. The primary focus of the research is to maintain low time complexity and low storage complexity for the video contents during content storage, retrieval and processing for information summarization. After a long pursue of study in this field of research, the research community have realised that, the success of the reduction in time and storage complexity relies on the extraction of the key frames from the video contents. The extraction of the key frames ensures that the content from the video data is not lost and the length of the video data is also reduced. After analysing the fundamentals of the content retrieval and summarization on distributed cloud storage environments, this work formulates the actual problem of viable length key frames extraction issues during a replicated video data processing. Hence, this work proposes a novel adaptive thresholding based key frame extraction and reduction process, where each storage instance for replicated data will contribute to the global thresholding process and based on the global thresholding, the final key frames will be extracted as elaborated in the mathematically and syntactically in this work. The measures taken in the proposed algorithm produces a 98.50% accuracy for video information representation and nearly 99.21% reduction in the video frames, which results into nearly 99% reduction in the time complexity comparted to the legacy systems for making the world of video content-based computing a faster and preferable domain of usage.

REFERENCES

1. Y. Guo, X. Hu, B. Hu, J. Cheng, M. Zhou, R. Y. K. Kwok, "Mobile cyber physical systems: Current challenges and future networking applications", IEEE Access, vol. 6, pp. 12360-12368, Dec. 2017.
2. K. Simonyan, A. Zisserman, "Two-stream convolutional networks for action recognition in videos", Proc. NIPS, pp. 568-576, 2014.
3. X. Hu et al., "Emotion-aware cognitive system in multi-channel cognitive radio ad hoc networks", IEEE Commun. Mag., vol. 56, pp. 180-187, Apr. 2018.
4. J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, A. W. M. Smeulders, "Selective search for object recognition", Int. J. Comput. Vis., vol. 104, no. 2, pp. 154-171, Apr. 2013.
5. X. Hu, X. Li, E. C. H. Ngai, V. C. M. Leung, P. Kruchten, "Multidimensional context-aware social network architecture for mobile crowdsensing", IEEE Commun. Mag., vol. 52, no. 6, pp. 78-87, Jun. 2014.
6. D. Oneata, J. Revaud, J. Verbeek, C. Schmid, "Spatio-temporal object detection proposals", Proc. ECCV, pp. 737-752, 2014.
7. M. Van den Bergh, G. Roig, X. Boix, S. Manen, L. Van Gool, "Online video SEEDS for temporal window objectness", Proc. ICCV, pp. 290-294, Dec. 2013.
8. M. Jain, J. van Gemert, H. Jégou, P. Bouthemy, C. G. M. Snoek, "Action localization with tubelets from motion", Proc. CVPR, pp. 740-747, Jun. 2014.
9. P. Weinzaepfel, Z. Harchaoui, C. Schmid, "Learning to track for spatio-temporal action localization", Proc. ICCV, pp. 3164-3172, Dec. 2015.
10. J. Zhang et al., "Energy-latency trade-off for energyaware offloading in mobile edge computing networks", IEEE Internet Things J., vol. 4, no. 6, pp. 1-12, Dec. 2017.
11. S. Hare, A. Saffari, P. H. S. Torr, "Struck: Structured output tracking with kernels", Proc. ICCV, pp. 263-270, Nov. 2011.
12. H. Wang, A. Kläser, C. Schmid, C.-L. Liu, "Dense trajectories and motion boundary descriptors for action recognition", Int. J. Comput. Vis., vol. 103, no. 1, pp. 60-79, May 2013.

ABOUT THE AUTHORS

T. Naga Raja is a research scholar from Computer Science department in JNTU Hyderabad. He received M Tech (CSE) in 2004 from JNTU Ananthapur and BE (ECE) from Osmania university hyderabad.

Dr. V.V. Venkata Ramana received Ph.D. from Sri Krishna Devaraya University, Anantapur, A.P, India in 1997, while M.Phil and M.Sc. in OR & SQC from S.V. University, Tirupati, A.P, India, 1986 and 1983 respectively. He is working as a Scientist – F in National Informatics Centre, Vijaywada, Andhra Pradesh. He have 47 number of published papers in National and International Journals. and presented papers at different National and International conferences. He guided a number of research scholars in the field of OR & SQC, Computer Science and other related fields.

Prof. Avula Damodaram joined at JNTU, Hyderabad in the year 1989 and rendered his services to the University at various capacities. He is presently the Adviser to the Govt of Andhra Pradesh. He was Vice Chancelor to the SV University. Tirupati. And Director of UGC-Academic Staff College and Director, School of Continuing & Distance Education at JNTU, Hyderabad. He has guided 6 Ph.D. and 2 MS Scholars apart from myriad M.Tech projects. Currently guiding 9 scholars for Ph.D and 1 scholar for MS. He is on the editorial board of 2 International Journals and a number of Course materials. He successfully executed an AICTE research projects. He has been a UGC nominee for a number of expert and advisory committees of various Indian Universities. He has published 50 papers in national and International journals and also presented 45 papers at different National and International conferences. He has organized as many as 40 Workshops, He visited the Universities of Austria and the United Kingdom for presenting papers at International conferences. He was honoured as DISTINGUISHED ACADAMICIAN by Pentagram Research Centre, India, in January 2010, for his meritorious service.