ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SPECTRAL CLUSTER BASED TEMPORAL FEATURE EXTRACTION AND B-TREE INDEXING FOR VIDEO RETRIEVAL

[1] RENUKADEVI .S , [2] Dr.S.MURUGAPPAN
[1] Assistant Professor, KSIT,
Department of Computer Science, India
[2]Associate Professor, Tamil Nadu Open University,
Department of Computer Science, India
E-mail: [1]renu_selvam2000@yahoo.co.in, [2]drmryes@gmail.com

ABSTRACT

Storage and retrieval of video data is considered as a simple and straight forwarded task but considered to be trivial when retrieval of information from video data is concerned. Recently many research works has been developed for video indexing and retrieval. But, there is a need for effective video indexing and retrieval method. In order to overcome such limitation, Spectral Cluster based temporal feature extraction and B Tree indexing (SC-BT) model is proposed in this paper. The SC-BT model is designed to achieve higher video retrieval rate and to reduce the video retrieval time. At first, SC-BT model used spectral clustering algorithm to extracts video features form the collection of video frames and clustering the video clips in the data set. After that, SC-BT model used B tree indexing technique to index the clustered video clips in N-dimensional space with their features with the objective of improving the true positive rate of video retrieval and reducing the video retrieval time. Finally, SC-BT model effectively extracts more similar detected video clips based on user query by evaluating the features observed using co-visibility graph through spectral clustering are recomputed in all iteration. The performance of SC-BT model is evaluated with sports repositories data set by using parameters such as spectral clustering time, spectral clustering accuracy, true positive rate of video retrieval and video retrieval time. The experimental results show that SC-BT model is able to improve the true positive rate of video retrieval rate by 12% and also reduces the video retrieval time by 45% when compared to state-of-the-art works.

Keywords: *Video, Indexing, Retrieval, Spectral Clustering, Video Frames, B tree indexing, User query*

1. INTRODUCTION

Recently, there has been a significant improvement in video retrieval process. An efficient video retrieval provides competent communication solutions for several applications like video surveillance, educational purpose, monitoring terrorism and so on at an early stage. Different video retrieval mechanisms are designed to solve the difficulties of video retrieval process, however the rate and the time at which the retrieval takes place has to be addressed. At present, a lot of research works has been designed for video retrieval. For example, an automatic shot based keyframe extraction was introduced in [1] for video indexing and retrieval that comprised of three phases, shot boundary detection, keyframe extraction and video indexing and retrieval. But, Automatic keyframe extraction is not implemented to improve the retrieval performance.

Segment Based Indexing Technique was implemented in [2] for image retrieval from the given video where RGB feature eliminates redundant frames in video and segment based technique is employed to indexing the individual frame. Segment Based Indexing Technique minimized bandwidth and decreased delays the video through the network of searching and reviewing. However, RGB feature increased time complexity by more number of frames accumulate extra space in the database. The Content-based Video Indexing and Retrieval using the Correspondence-Latent Dirichlet Allocation (corr-LDA) probabilistic framework was designed in [3]. The Corr-LDA framework employ the concept-level matching presented by corr-LDA to construct correspondences between text and multimedia with the intention of retrieving content with higher accuracy. However, a combination of tag-based query search and content-based semantic linkages was remained unaddressed.

The comprehensive literature review of text extraction in images and video and text based image and video retrieval was presented in [4]. A novel method was developed in [5] to map a surveillance video to a temporal profile for indexing and searching. Though computation cost and data size was reduced extensively. But, efforts to support parallel search to improve scalability of video searching is remained unaddressed. A compressed domain video indexing method was designed in [6] for video indexing and retrieval where indexing is depends on the histogram of the positions of the blocks that are employed in motion compensation of the coded video.

Shrinkage Optimized Direction information Assessment (SODA) was introduced in [7] for multimodal video indexing and retrieval by using both audio and video features. The SODA enhances the precision rate of video retrieval and recall performance. However, Single modality provides not enough information for precise video indexing. Speaker keyword spotting system was planned in [8] to support instructors with indexing their lecture videos with the aiming at helping low-vision students to more easily locate topics in the videos. The keyword spotting system achieved average Precision performance. Though, Query expansion techniques are not used to expand the speed of queries.

An optimization model called as Video Search Engine Optimization (VSEO) was developed in [9] that improved ranking of video in search engine by assigning key attributes during video searching and retrieval. The VSEO significantly improves the video ranking in search engine. But, VSEO is not implemented on further video search engines. Vector of Locally Aggregated Descriptors (VLAD) was presented in [10] to provide particular query video, with aim of retrieved relevant videos from the database. Though, VLAD is inappropriate to the STIP features because it is extracted for the entire video.

Based on the aforementioned techniques and methods presented, in this work we propose a novel framework called as Spectral Cluster based temporal feature extraction and B Tree indexing (SC-BT) model for effectively performing video indexing and video retrieval process.

The rest of the paper organized as follows. In Section 2, a summary of different video retrieval techniques are explained. In Section 3, the proposed SC-BT model is described with the help of neat architecture diagram. In Section 4, simulation environment is presented with exhaustive analysis of results explained in Section 5. In Section 6, the concluding remarks are presented.

2. RELATED WORKS

Bag-of-systems (BoS) was presented in [11] where BoS Tree integrates the expressiveness of a large BoS codebook that enables efficient mapping of videos to the BoS codebook. BoS Tree achieves better accuracy to the direct-indexed large BoS during computational reducing. However, the computational cost is increased. Audio-Visual-Based Query by Example Video Retrieval was presented in [12] where audio features are firstly employed for rough retrieval to narrow the scope of objective videos in database and then the visual features are applied to purify retrieval. Finally, the system presents the similar videos which are similar to a user-provided example.

An Efficient Method for Automatic Video Annotation and Retrieval in Visual Sensor Networks was implemented in [13] to solve the issues of pattern classification in video annotation. The accurate expressions of spatial orders and video sequences that are related with concepts and sub concepts remained unsolved. A new approach was developed in [14] for medical multimedia document retrieval supporting cross-type medical multimedia retrieval and providing users' retrieval intent. But, the retrieval speed is not considered. Content based video retrieval system was planned in [15] with the help of entropy based shot detection method where entropy feature is employed for indexing the video with the combination of RGB color values and video metadata.

Content Based Video Retrieval Using Cluster Overlapping was presented in [16] which utilized dissimilar spatio-temporal features of the video and clustering methods to efficiently retrieve videos from the database while a video clip is given as a query. A novel algorithm for content-based video indexing and retrieval using key-frames texture, edge, and motion features was designed in [17] for content-based video indexing and retrieval.

Content Based Video Retrieval with Frequency Domain Analysis Using 2-D Correlation Algorithm was developed in [18] to discover all the objects from video that matched with the user's query image.

A Latent Semantic Indexing (LSI) technique was implemented in [19] based on Singular Value Decomposition and fusion of visual features such as color and edge for video indexing and retrieval. Video information retrieval by using sequential indexing technique algorithm was designed in [20] to improve the system performance of video retrieval based on the user input query and to reduce the searching time. Based on the above mentioned methods and techniques, the following proposed work is designed to provide an appropriate solution to solve the existing issues.

2. SPECTRAL CLUSTER BASED TEMPORAL FEATURE EXTRACTION AND B TREE INDEXING (SC-BT) MODEL

The Spectral Cluster based temporal feature extraction and B Tree indexing (SC-BT) model design a very specific way of searching video clips based on the user query. The key objective of SC-BT model is to efficiently retrieving the video contents based on user query and to reduce the video retrieval time in a significant manner.

3. SPECTRAL CLUSTER BASED TEMPORAL FEATURE EXTRACTION AND B TREE INDEXING (SC-BT) MODEL

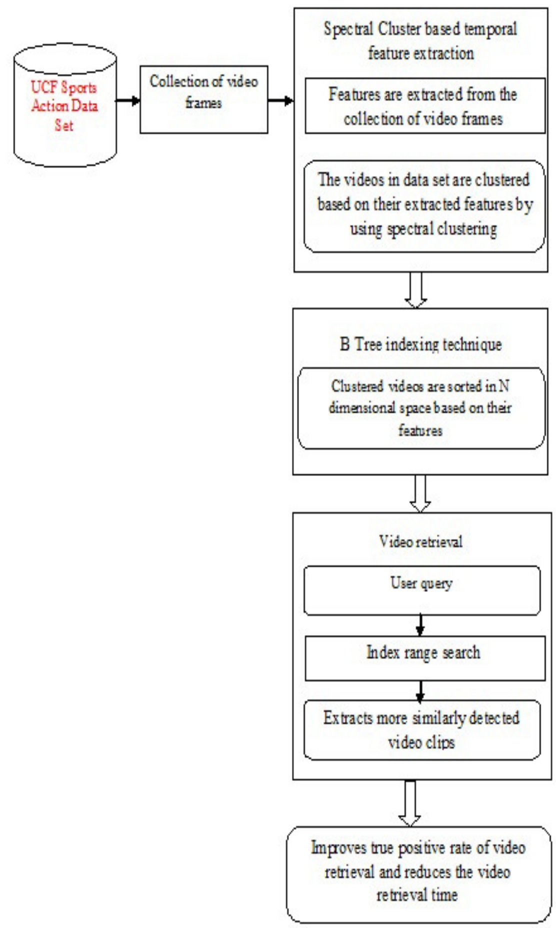

Figure 1 Overall Architecture Diagram of Spectral Cluster Based Temporal Feature Extraction and B Tree Indexing (SC-BT) Model

The SC-BT model takes the UCF Sports Action Data Set as input for performing video indexing and video retrieval process. In SC-BT model, UCF Sports Action Data Set contains the collection of videos. The videos in UCF Sports Action Data Set comprises of numerous video frames. With the support of these video frames, SC-BT model extracts their features i.e. informative characteristics with aiming at clustering the video clips by using the spectral clustering algorithm. After performing the clustering process, the clustered videos are indexed based on their features by using the B tree indexing techniques. The overall architecture diagram of Spectral Cluster based temporal feature extraction and B Tree indexing (SC-BT) model is shown Figure 1.

As shown in Figure 1, proposed SC-BT model initially performs feature extraction process where the features of video frames are efficiently extracted for clustering the video clips in data set.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

These features contain the static features in video frames, object features, motion features, etc. Then, the collections of videos in data set are effectively clustered based on their extracted features with aiming at improving the accuracy of video clustering and reducing the clustering time of videos with the help of spectral clustering algorithm. After that, with the support of classified video clips, SC-BT model build a semantic video index. The semantic index together with the high-dimensional index of video feature vectors constitutes the total index for video sequences that are stored in given data set with aiming at improving the true positive rate of video retrieval and reducing the video time. Finally, SC-BT model performs video retrieval process where the similarly detected video clips are effectively retrieved for the given user query. The detailed explanation about SC-BT model is described in following sections.

3.1 Spectral Cluster Based Temporal Feature Extraction for Clustering the Videos

In SC-BT model, Spectral Cluster Based Temporal Feature Extraction based on the concept of locality of features which described the set of features that have been seen observed concurrently from the video. Feature extraction is the process of extracting features (informative characteristics) from a frame in video. Feature is anything of interesting in a video such as face, car, tree, shapes etc. With the aim of capturing this property, an auxiliary graph is constructed during the scene exploration which is called as Co-visibility Graph (CovGraph). In this graph, features are characterized as nodes and those features that have been observed in the same frame are associated by edges. The CovGraph is incrementally constructed after features have been gained from each video frame. The sets of nodes and edges are updated by including those features that were not detected in previous frames. Attributes of previously perceived nodes and edges are also updated. The sample co-visibility graph is constructed for video features is shown in below,

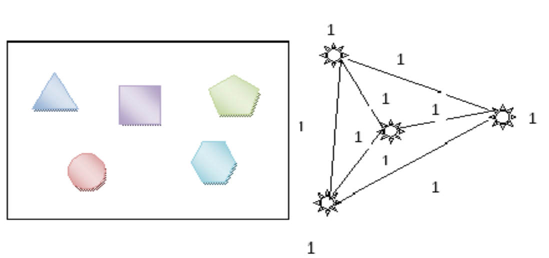

Figure 2 a) Video frame Figure 2 b) Co-visibility Graph

Figure 2 a) is an example video frame taken and Figure 2 b) shows the example co-visibility graph constructed for that video frame where each node is labeled with the number of observations of the corresponding video feature. Edges signify the locality of features and their values correspond to the number of times they have been observed concurrently. The construction of co-visibility graph will require the set of features discovered in each video frame and the connection between features of consecutive frames. Once this auxiliary graph is created, the edge weights must be measured in order to define the similarity matrix. For each node in the graph, num_i is the number of observations of the feature $Feature_i$ up to time t. If two features $Feature_i^t$ and $Feature_j^t$ has been observed in the same frame at time t, an edges is created between them or if it exists, i.e. both features have been previously observed simultaneously, the edge value is increased. In order to determine an average value associated to the locality of video features, the co-visibility rate is described which mathematically formulated as below,

$$CovRate_{ij} = \frac{avg_{ij}}{\min(n_i, n_j)} \quad (1)$$

From the equation (1), the edge value is divided by the minimum number of observations of the corresponding video features. As it can be seen, the minimum possible value of the $CovRate$ is zero while two features have not been observed at the same time and the maximum value is one while two features always have been observed concurrently at least one of them. Thus the $CovRate$ is within the range [0, 1].

The $CovRate$ represents the similarity function which computes the locality of video features in the CovGraph. Once this function is exploited to calculate the edge weights, the resulting adjacency or similarity matrix W can be employed to determine the degree of relationship between the last perceived frame and the set of previously processed video frames. This degree of relationship will be described by the weights of the edges which constitute the minimum cut of the graph, represented by the weighted similarity matrix W of CovRate measures. This matrix W is symmetric, non-negative and block diagonal. The bipartition of the graph will be conducted by using a spectral clustering algorithm.

Once a similarity measure has been defined by using equation (1), the weighted adjacency matrix of the graph can be used to find a minimum partition. The aim is to split the graph in order to determine dissimilar clusters in the features detected, i.e. to discover a change in appearance. By using the spectral clustering, we can concurrently reduce the dissociation between groups and improves the association within the groups.

Let consider $G = (N, E)$ be the weighted graph with set of node $N = \{node_1, \dots node_n\}$ and $W = (w_{ij})_{i,j=1,\dots,n}$ the weighted similarity matrix of G. The degree of a node $node_i \in N$ is mathematically formulated as below,

$$degree_i = \sum_{j=1}^{n} w_{ij} \qquad (2)$$

After that, we can define the degree matrix D as the diagonal matrix with the degrees $d_1, d_2, \dots d_n$ on the diagonal. The Laplacian matrix computed for clustering the video features is formulated as,

$$L = D - W \qquad (3)$$

Thus, the normalized Laplacian matrix is formulated as follows,

$$L = I - D^{-1}W \qquad (4)$$

The goal of spectral clustering is to separate video features into different clusters where video features in the same cluster are similar to each other. We can construct a graph from the similarity matrix, where the vertexes represent the video features, and the edge weights represent similarities between video features. Given a weighted graph G, we want to find a cut of the graph, such that the cut will be minimized. In SC-BT model, we denote $\Delta = \{A_1, A_2, \dots, A_n\}$ as a clustering result, where A_i includes all the video features that belong to cluster i.

As said by spectral graph theory, there can be numerous dissimilarity objective functions for cluster analysis for example MNCut, RatioCut, NCut, and MinCut. In SC-BT model, we focus on MNCut for clustering the videos where the objective is to achieve a rational minimum cut. Given a graph with similarity matrix W, where w_{ij} designates the i, jth entry of W. Followed by the problem of minimizing MNCut for video clustering is formulated as,

$$MN(\Delta) = \sum_{i=1}^{k} \frac{cut(A_i, \overline{A_i})}{vol(A_i)} \qquad (5)$$

Form the equation (5), $\overline{A_i}$ refers to the complement of A and $vol(A_i)$ denotes the total weights of A_i. Given the weighted similarity matrix W of the G and the degree matrix D, the spectral clustering algorithm is designed for clustering the video clips is explained in below Figure 3.

As shown in Figure 3, spectral clustering algorithm is designed for clustering the video features comprises of seven steps as follows. Initially, spectral clustering algorithm evaluates the similarity matrix for given video features. Then, it computes the Laplacian matrix and normalized Laplacian matrix of video features for clustering the videos contents. After that, spectral clustering algorithm computes the first K eigenvectors of Laplacian matrix and which is denoted as U. Subsequently, spectral clustering algorithm considers the rows of U as video features and employs k-means to group them into K clusters. Afterward, it assign video feature $Feature_i$ to cluster A_j if and only if row i of the matrix U was assigned to cluster A_j. Finally, spectral clustering algorithm returns the clustering results of videos with their features. This in turn improves the spectral clustering accuracy of videos and also reduces the spectral clustering time in a significant manner.

3.2. B-Tree Indexing

After clustering the video clips, B-tree indexing is employed for storing the videos in sorted order for features. In SC-BT model, the B-Tree Indexing based on the probabilistic videos features that captures all constituents of video and performs ordered list with aiming at reducing video retrieval time. The tuple elements on clustered videos help to capture the video features for performing indexing operation. The B-tree indexing operation supports update operations by using clustered video clips as input vector.

// **SPECTRAL CLUSTERING ALGORITHM FOR CLUSTERING THE VIDEOS**

INPUT: SET OF VIDEOS '$video_i = video_1, video_2, .. video_n$' SET OF VIDEOS FEATURES '$Feature_i = Feature_1, Feature_2, .. Feature_n$', CLUSTER NUMBER K.

OUTPUT: IMPROVED SPECTRAL CLUSTERING ACCURACY (I.E. CLUSTERING OF VIDEOS BASED ON THEIR FEATURES)

BEGIN:

STEP 1:COMPUTE SIMILARITY MATRIX 'W'

STEP 2: COMPUTE LAPLACIAN MATRIX 'L' BY USING (3)

STEP 3: COMPUTE NORMALIZED LAPLACIAN MATRIX L BY USING (4)

STEP 4: COMPUTE THE FIRST K EIGENVECTORS OF L, REPRESENT AS U

STEP 5: CONSIDER THE ROWS OF U AS VIDEO FEATURES, AND USE K-MEANS TO CLUSTER THEM INTO K CLUSTERS.

STEP 6: ASSIGN $Feature_i$ TO CLUSTER A_j IF AND ONLY IF ROW i OF THE MATRIX U WAS ASSIGNED TO CLUSTER A_j

STEP 7: RETURN CLUSTERING RESULTS

END

Figure 3 Spectral Clustering Algorithms for Clustering the Videos

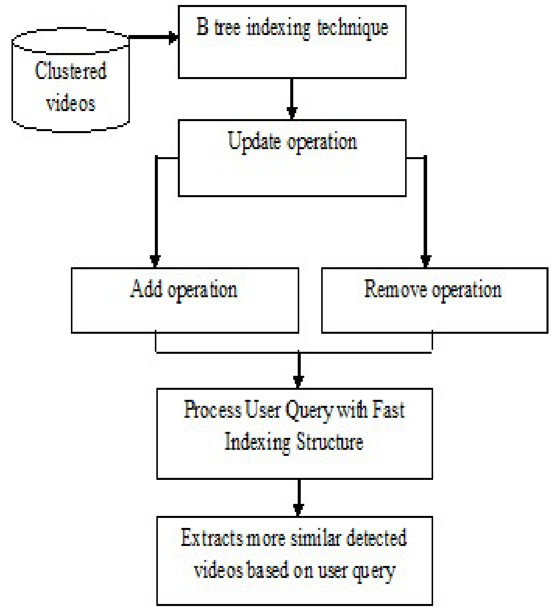

Figure 4 B-Tree Indexing For Video Retrieval

The B-tree indexing for video retrieval process is shown in below Figure 4.

B-tree is a tree structure that sorts the features of video contents and permit effective searching operation. The searching process is carried out using self-balancing the features of video content in SC-BT model. B-tree indexing is optimized where clustered video clips are indexed to n-dimensional blocks with their features. SC-BT model extensively improves indexing operation on primary set where by allows the users to manipulate indexes and achieve higher video retrieval rate. The main advantage of B-tree indexing is that the nodes on the overall tree are connected, in order that the overall clustered features of video content is scanned with linear pass. The scanning through linear pass reduces the average retrieval time of video in an effective manner.

In SC-BT model, B-tree index frequently access random nodes that lie closer to the root for performing indexing. Random based index reduces retrieval time of video and perform indexing more quickly on clustered video clips. The B-tree indexing technique builds root, internal and leaf points. The degree of the elements is analyzed and performs ordered list in SC-BT model. The ordered list performs addition and removal operation, where addition is performed by initially searching down the tree of video features. The removal operation is performed when video features are unrelated to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

indexing structure. The adding up and removal procedure on clustered video features generated from a set of video is described as,

// **ADDING OPERATION**

STEP 1: IF THE DIMENSIONAL SPACE N IS NOT FULL, THEN ADD THE VIDEO CLIPS BASED ON THEIR FEATURES

STEP 2: ALLOCATE NEW LEAF STRUCTURE, AND THEN ADD THE NEWLY CLUSTERED VIDEO FEATURES

$$Adding\ operation = N[add\ new(v1, v2, v3..vn)] \quad (6)$$

STEP 3: IF THE ROOT POINT SPLITS, THEN THE ROOT NODE HAS ONE KEY AND TWO INTERMEDIATE NODES

STEP 4: REPEAT STEP UNTIL FOUND NO ITERATION

// **REMOVAL OPERATION**

STEP 5: ROOT NODE ON B-TREE WHERE IT BELONGS

STEP 6: REMOVE THE UNRELATED LEAF STRUCTURE AS

$$Removal\ operation = N[remove\ old(v1, v2, v3..vn)$$

STEP 7: IF THE LEAF NODE REMOVED, THEN THE LEAF NODE HAS TWO INTERMEDIATE NODES

STEP 8: REPEAT STEP UNTIL FOUND NO ITERATION

If the clustered videos are not placed on dimensional space, then videos are added into the relevant field based on their visual features. Similarly, the unrelated video and their features are removed from dimensional space using removal operations. SC-BT model transfers the block after searching of user information and achieves precise results. This in turn improves the video retrieval rate of SC-BT model in an effective manner.

3.3 Video Retrieval

Let us consider a set of user range query 'query' to search the result from N dimensional space. The query series of Random B-tree node is analyzed and performs processing operation in SC-BT model. Starting from the lower bound of the B tree, SC-BT model follows adjacent links to perform pruning process. The pruning process still removes unwanted video features and extracts only the user requested query result in SC-BT model. The retrieval process of B tree is carried out as follows,

// **RETRIEVAL PROCESS IN B TREE**

BEGIN

STEP 1: N DIMENSIONAL CLUSTERED VIDEOS GENERATED FROM SET OF VIDEOS BASED ON THEIR FEATURES

STEP 2: PERFORM VIDEO RETRIEVAL PROCESS BASED ON USER QUERY '$query$' ON 'N'

STEP 3: WHILE N=N.RIGHT AND N.LOW<I DO

STEP 4: OUTPUT RETRIEVED SIMILARLY DETECTED VIDEOS TO THE USER

END

The first optimized result using B-tree indexing creates search with lower bound of the tree and retrieves the similar detected videos accordingly. The result is uniformly distributed among the users, thereby reducing the video retrieval time in an effective manner. Random B-tree indexing attains relative adaptivity on pruning which in turn improves robustness result on retrieving the video clips. Therefore, SC-BT model achieves higher video retrieval rate and also reduces the video retrieval time in an effective manner.**4.EQUATIONS**

When numbering equations, enclose numbers in parentheses and place flush with right-hand margin of the column. Equations must be typed, not inserted.
(If nonstandard fonts are used its better to put equations as images instead of text)
Example:

$$Net_j = w_0 + \sum_{i=1}^{n} x_i w_{ij} \quad (1)$$

4. EXPERIMENTAL SETTING

The Spectral Cluster based temporal feature extraction and B Tree indexing (SC-BT) model is implemented in Java Language using UCF Sports Action Data Set. The UCF Sports dataset comprises of a collection of actions gathered from a variety of sports that are usually featured on broadcast television channels for instance the BBC and ESPN. The video sequences were obtained from a broad range of stock footage websites containing BBC Motion gallery and GettyImages. The dataset comprises a total of 150 sequences with the resolution of 720 x 480. The collection characterizes a natural pool of actions featured in an extensive range of scenes and viewpoints. The UCF Sports Action Data Set contains 10 different actions of videos.

The performance of SC-BT model is compared against with the existing two methods namely Automatic Shot based Keyframe Extraction [1] and Segment Based Indexing Technique [2] respectively. The experimental evaluation of SC-BT model is conducted on various factors such as such as spectral clustering accuracy, spectral clustering time, true positive rate of video retrieval and video retrieval time.

5. RESULTS AND DISCUSSIONS

The efficiency of SC-BT model is compared against with exiting two methods namely Automatic Shot based Keyframe Extraction [1] and Segment Based Indexing Technique [2]. The performance of SC-BT model is evaluated along with the following metrics with the help of tables and graphs.

5.1 Measurement of Spectral Clustering Accuracy

In SC-BT model, spectral clustering accuracy refers the clustering accuracy of videos based on their features. Therefore, it defined as the ratio of number of correctly clustered videos based on their features to the total number of videos taken. The spectral clustering accuracy is measured in terms of percentages (%) and mathematically formulated as follows,

$$\text{spectral clustering accuracy} = \frac{\text{number of correctly clustered videos based on their features}}{\text{total number of videos taken}} * 100 \quad (8)$$

From the equation (9), the clustering accuracy of videos is obtained. When the spectral clustering accuracy is higher, the method is said to be more efficient.

Table 1 shows the spectral clustering accuracy result of three methods based on different number of videos taken in the range of 10-70. SC-BT model considers the framework with different number of videos for conducting experimental works by using Java Language.

From the table value, it is clear that the spectral clustering accuracy using proposed SC-BT model is higher when compared to other existing methods [1],[2]. Based on the above table, the graph is plotted in below Figure 5.

Table 1 Tabulation for Spectral Clustering Accuracy

Number of Videos	Spectral Clustering Accuracy (%)		
	Automatic Shot based Keyframe Extraction	Segment Based Indexing Technique	SC-BT Model
10	65.02	72.14	81.26
20	70.14	75.18	83.14
30	72.05	76.59	85.69
40	73.69	78.15	88.47
50	75.18	81.69	91.08
60	79.58	86.24	94.57
70	86.12	93.24	98.36

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

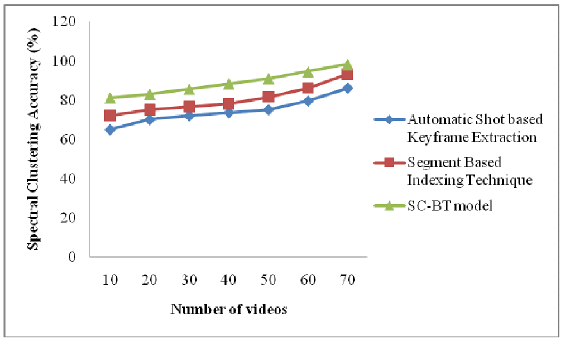

Figure 5 Measurement of Spectral Clustering Accuracy

Figure 5 illustrates the impact of spectral clustering accuracy versus different number of videos using three methods namely Automatic Shot based Keyframe Extraction [1], Segment Based Indexing Technique [2], SC-BT model. As exposed in figure, the proposed SC-BT model provides better spectral clustering accuracy as compared to other existing methods [1], [2]. Further, while increasing the number of videos, the spectral clustering accuracy is also gets increased by using all three methods. But, comparatively the spectral clustering accuracy using SC-BT model is higher. This is due to the application of spectral clustering algorithm in SC-BT model where it efficiently classifies the videos based o their different visual features. This in turn helps in improving the spectral clustering accuracy in an effective manner. As a result, SC-BT model improves the spectral clustering accuracy by 16% when compared to Automatic Shot based Key frame Extraction [1] and 10% when compared to Segment Based Indexing Technique [2] respectively.

5.2 Measurement of spectral Clustering time

In SC-BT model, spectral clustering time measures the amount of time taken for clustering the videos based on their features. The spectral clustering time is measured in terms of milliseconds (ms) and mathematically as follows,

$$spectral\ clustering\ time = n * time(clusering\ the\ one\ video) \qquad (9)$$

From the equation (10), spectral clustering time is obtained where n represents the number of videos taken. When the spectral clustering time is lower, the method is said to be more efficient.

Table 2 Tabulation for Spectral Clustering Time

Number of Videos	Spectral Clustering Time (ms)		
	Automatic Shot based Keyframe Extraction	Segment Based Indexing Technique	SC-BT model
10	23.1	18.3	11.6
20	39.6	31.6	24.1
30	52.8	45.9	37.9
40	66.1	58.4	50.4
50	79.3	71.3	63.8
60	92.7	84.2	77.5
70	99.8	96.5	89.2

The result analysis of spectral clustering time using three methods based on different number of videos taken in the range of 10-70 is demonstrated in Table 2. From the table value, it is illustrative that the spectral clustering time using proposed SC-BT model is lower when compared to other existing methods [1], [2]. Based on the above table, the graph is drawn in below Figure 6.

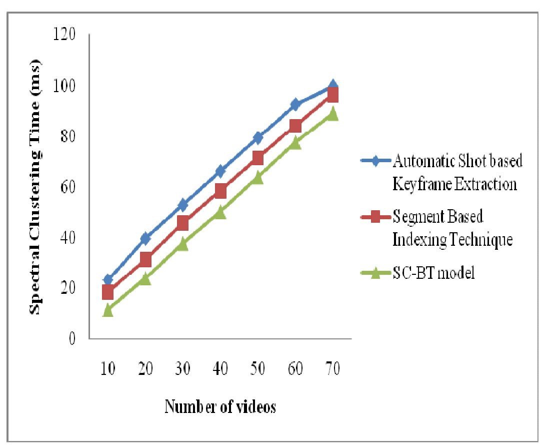

Figure 6 Measurement of Spectral Clustering Time

Figure 6 shows the impact of spectral clustering time versus different number of videos using three methods namely Automatic Shot based Keyframe Extraction [1], Segment Based Indexing Technique [2], SC-BT model. As revealed in figure, the proposed SC-BT model provides better spectral clustering time as compared to other existing methods [1], [2]. Further, while increasing the number of videos, the spectral clustering time is also gets increased by using all three methods. But, comparatively the spectral clustering time using SC-BT model is lower. This is because of the application of spectral clustering algorithm in SC-BT model where it evaluates the similarity matrix and Laplacian matrix for clustering the features of video frame. This in turn assists in reducing the spectral clustering time in a significant manner. As a result, SC-BT model reduces the spectral clustering time by 41% when compared to Automatic Shot based Keyframe Extraction [1] and 22% when compared to Segment Based Indexing Technique [2] respectively.

5.3 Measurement of True Positive Rate of Video Retrieval

In SC-BT model, video retrieval true positive rate is defined as the ratio of number of correctly retrieved videos based on user query to the total number of videos taken. The true positive rate of video retrieval is measured in terms of percentages (%) and mathematically formulated as follows,

$$\text{true positive rate} = \frac{\text{number of correctly retrieved videos based on user query}}{\text{total number of videos taken}} * 100 \qquad (10)$$

From the equation (11), true positive rate of video retrieval is obtained. When the true positive rate of video retrieval is higher, the method is said to be more efficient.

Table 3 Tabulation for True Positive Rate of Video Retrieval

Number of Videos	True Positive Rate of Video Retrieval (%)		
	Automatic Shot based Keyframe Extraction	Segment Based Indexing Technique	SC-BT Model
10	69.26	78.15	85.69
20	71.69	80.15	87.59
30	73.26	82.69	89.23
40	76.91	83.65	90.15
50	78.69	85.79	92.41
60	80.66	87.69	94.36
70	85.69	93.44	99.25

Table 3 explains the comparative result analysis of true positive rate of video retrieval using three methods based on different number of videos taken in the range of 10-70. From the table value, it is descriptive that the true positive rate of video retrieval using proposed SC-BT model is higher when compared to other existing methods [1], [2]. Based on the above table, the graph is plotted in below Figure 7.

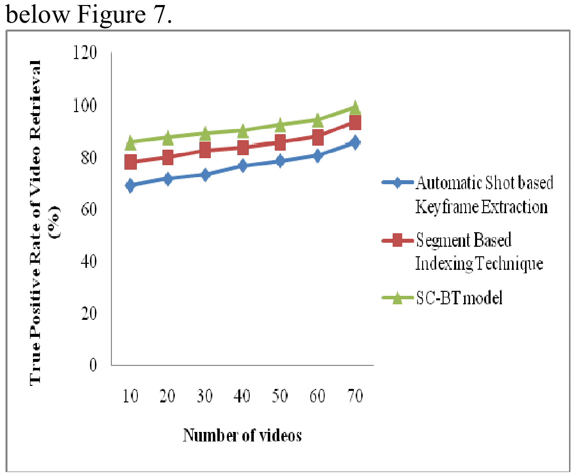

Figure 7 Measurement of True Positive Rate of Video Retrieval

The impact of true positive rate of video retrieval based on different of videos taken is represented in Figure 7. As shown in figure, the proposed SC-BT model provides better true positive rate of video retrieval when compared to other existing methods [1], [2]. In addition, while increasing the number of videos, the true positive rate of video retrieval is also gets increased by using all three methods. But, comparatively the true positive rate of video retrieval using SC-BT model is higher. This is due to the application of spectral clustering algorithm and B-tree indexing in SC-BT model. The spectral clustering algorithm used in SC-BT model efficiently classifies the videos in UCF Sports Action Data Set on their visual features and then B-tree indexing used for storing the clustered video in sorted ordered of video features. This in turn supports in improving the true positive rate of video retrieval in an efficient manner. Therefore, SC-BT model improves the true positive rate of video retrieval by 16% when compared to Automatic Shot based Keyframe Extraction [1] and 7% when compared to Segment Based Indexing Technique [2] respectively.

5.4 Measurement of Video Retrieval Time

In SC-BT model, video retrieval time refers the amount of time taken for retrieving the more similarly detected videos for the given user query. The video retrieval time is measured in terms of milliseconds (ms). When the video retrieval time is lower, the method is said to be more efficient.

Table 4 depicts the comparative result analysis of video retrieval time using three methods based on different number of videos sizes taken in the range of 113.6-936.2. From the table value, it is expressive that the video retrieval time using proposed SC-BT model is lower as compared to other existing methods [1], [2]. Based on the above table, the graph is drawn in below Figure 8.

Table 4 Tabulation for Video Retrieval Time

Size of Video (MB)	Video Retrieval Time (ms)		
	Automatic Shot based Keyframe Extraction	Segment Based Indexing Technique	SC-BT Model
113.6	16.2	11.4	8.6
323.7	18.6	15.9	11.3
349.5	22.9	21.6	15.2
454.5	27.5	24.7	18.6
635.2	34.6	27.3	21.9
905.3	37.9	30.2	24.7
936.2	41.3	34.6	27.6

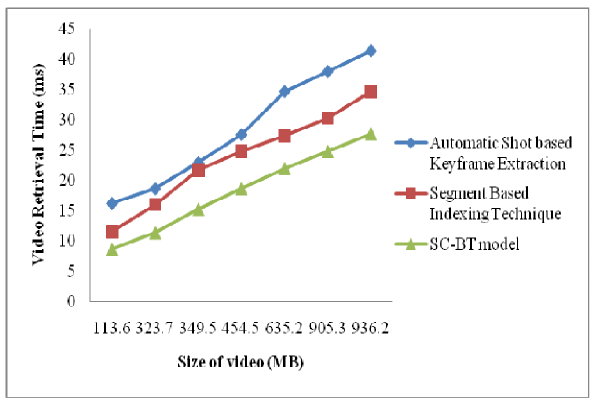

Figure 8 Measurement of Video Retrieval Time

Figure 8 describes the impact of video retrieval time versus different number of video sizes by using three methods. As shown in figure, the proposed SC-BT model provides better video retrieval time when compared to other existing methods [1], [2]. Besides, while increasing the number of videos, the video retrieval time is also gets increased by using all three methods. But, comparatively the video retrieval time using SC-BT model is lower. This is due to the application of B-tree indexing in SC-BT model. With the help of B-tree indexing, SC-BT model effectively retrieves more similar detected videos based on user query.

This in turn significantly reduces the video retrieval time. As a result, SC-BT model reduces the video retrieval time by 59% when compared to Automatic Shot based Keyframe Extraction [1] and 32% when compared to Segment Based Indexing Technique [2] respectively.

6. CONCLUSION

In this work, an efficient novel framework is designed called as Spectral Cluster based temporal feature extraction and B Tree indexing (SC-BT) model for effectively retrieving video clips based on user query. The main objective of SC-BT model is to achieve higher true positive rate of video retrieval and to lessen the retrieval time of videos. The objective of SC-BT model is attained by using spectral clustering algorithm and B tree indexing technique. Initially, SC-BT model used spectral clustering algorithm for extracting the video features form the collection of video frames and clustering the video clips in the data set. Next, SC-BT model used B tree indexing technique for indexing the clustered video clips in N-dimensional space with their features which in turn improves the true positive rate of video retrieval and also reduces the video retrieval time in an effective manner. Finally, SC-BT model efficiently retrieves more similarly detected video clips from the N-dimensional space based on user query. The performance of SC-BT model is measured in terms of spectral clustering time, spectral clustering accuracy, true positive rate of video retrieval and video retrieval time by using UCF Sports Action Data Set and compared with two exiting methods. With the experiments conducted for SC-BT model, it is observed that the true positive rate of video retrieval rate provides more accurate results as compared to state-of-the-art works. The experimental results demonstrate that SC-BT model provides better performance with an improvement of true positive rate of video retrieval rate by 12% and reduces the video retrieval time by 45% when compared to state-of-the-art works.

REFERENCES

[1] T.G. Lakshmi Priya, S. Domnic," Shot based keyframe extraction for ecological video indexing and retrieval", Ecological Informatics, Elsevier, Volume 23, September 2014, Pages 107–117.

[2] D.Saravanan, "Segment Based Indexing Technique for Video Data File", Procedia Computer Science, Elsevier, Volume 87, 2016, Pages 12–17.

[3] Rahul Radhakrishnan Iyer, Sanjeel Parekh, Vikas Mohandoss, Anush Ramsurat, Bhiksha Raj, Rita Singh, "Content-based Video Indexing and Retrieval Using Corr-LDA", Computer Science, Information Retrieval, Pages 1-7, 2016.

[4] Avinash N Bhute, B.B. Meshram, "Text Based Approach for Indexing and Retrieval of Image and Video: A Review", Advances in Vision Computing: An International Journal (AVC) Volume 1, Issue 1, Pages 27-38, March 2014.

[5] Saeid Bagheri, Jiang Yu Zheng, Shivank Sinha, "Temporal mapping of surveillance video for indexing and summarization", Computer Vision and Image Understanding, Elsevier, Volume 144, March 2016, Pages 237–257.

[6] Farahnaz Akrami, Farzad Zargari, "An efficient compressed domain video indexing method", Multimedia Tools and Applications, Springer, September 2014, Volume 72, Issue 1, pp 705–721.

[7] Xu Chen, Alfred O. Hero, III, Silvio Savarese, "Multimodal Video Indexing and Retrieval Using Directed Information", IEEE Transactions on Multimedia, Volume: 14, Issue: 1, Pages 3 – 16, Feb. 2012.

[8] Manish Kanadjea, Zachary Millera, Anurag Agarwal, Roger Gaborskia, Richard Zanibbia, Stephanie Ludi, "Assisted keyword indexing for lecture videos using unsupervised keyword spotting", Pattern Recognition Letters, Elsevier, Volume 71, 1 February 2016, Pages 8–15.

[9] Krishna Choudharia,Vinod K Bhallab," Video Search Engine Optimization Using Keyword and Feature Analysis", Procedia Computer Science, Elsevier, Volume 58, 2015, Pages 691-697.

[10] Mopuri K Reddy, Sahil Arora, R. Venkatesh Babu," Spatio-Temporal feature based VLAD for efficient Video Retrieval", IEEE 2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), , Pages 1 – 4, Dec 2013.

[11] Adeel Mumtaz, Emanuele Coviello, Gert R.G. Lanckriet, and Antoni B. Chan, "A Scalable and Accurate Descriptor for Dynamic Textures Using Bag of System Trees", IEEE Transactions On Pattern Analysis And Machine Intelligence, Volume 37, Issue 4, April 2015.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[12] Sujuan Hou and Shangbo Zhou, "Audio-Visual-Based Query by Example Video Retrieval", Mathematical Problems in Engineering, Hindawi Publishing Corporation, Volume 2013 (2013), Article ID 972438, 8 pages.

[13] Jiangfan Feng and Wenwen Zhou, "An Efficient Method for Automatic Video Annotation and Retrieval in Visual Sensor Networks", Hindawi Publishing Corporation, International Journal of Distributed Sensor Networks, Volume 2014, Article ID 832512, 8 pages.

[14] Kehua Guo and Shigeng Zhang, "A Semantic Medical Multimedia Retrieval Approach Using Ontology Information Hiding", Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine, Volume 2013, Article ID 407917, 7 pages.

[15] P. M. Kamde, Sankirti Shiravale, S. P. Algur, "Entropy Supported Video Indexing for Content based Video Retrieval", International Journal of Computer Applications (0975 – 8887), Volume 62, Issue 17, Pages 1-6, January 2013.

[16] Deepak C R , Sreehari S , Gokul M , Anuvind B, "Content Based Video Retrieval Using Cluster Overlapping", International Journal of Computational Engineering Research, Volume 03, Issue 5, Pages 104-108, 2013.

[17] M.Ravinder and T.Venugopal, "Content-Based Video Indexing and Retrieval using Key frames Texture, Edge and Motion Features", International Journal of Current Engineering and Technology, Volume 6, Issue 2, Pages 672-676, April 2016.

[18] Navdeep Kaur, Mandeep Singh, "Content Based Video Retrieval with Frequency Domain Analysis Using 2-D Correlation Algorithm", Volume 4, Issue 9, Pages 388-393, September 2014.

[19] Rashmi M, Roshan Fernandes, "Video Retrieval Using Fusion of Visual Features and Latent Semantic Indexing", International Journal of Innovative Research in Computer and Communication Engineering, Volume 2, Issue 5, Pages 4272-4277, May 2014.

[20] D.Saravanan, V.Somasundaram, "Matrix Based Sequential Indexing Technique for Video Data Mining", Journal of Theoretical and Applied Information Technology, Volume 67, Issue 3, Pages 725-731, 2014.