OPEN ACCESS at journalijcir.com

Research Article

ISSN: 2395 -5775

# GRADIENT BOOST ENSEMBLE CLASSIFIER FOR UNSTRUCTURED MULTIMEDIA INFORMATION RETRIEVAL

Vijayan K[1] and Chandra Sekar C[2]

**Abstract**

Retrieval of videos from large database using video queries plays a significant role for a lot of applications. Few researches works have been designed for retrieving relevant videos from large database using classification techniques. But, the classification performance of was not effectual for achieving higher precision and recall for video retrieval. In order to solve this limitation, Gradient Boost Ensemble Classification (GBEC) technique is proposed. The GBEC technique at first takes video query as input. Then, GBEC technique used Independent Component Analysis Model in order to extracts the visual features such as shape, color, texture in videos for efficient classification. After visual features extraction, GBEC technique applied Gradient Boost Ensemble Classifier in order to classify the videos in a given dataset as similar or dissimilar based on video query which resulting enhanced classification accuracy. Finally, the classified similar videos' are retrieved based on video query which in turn helps for increasing precision and recall of video retrieval with minimum time. The performance of GBEC technique is measured in terms of metrics such as classification accuracy, time complexity, Precision and recall with aid of three datasets as compared to state-of-the-art works.

***Key Words:*** Gradient Boost Ensemble Classifier, Independent Component Analysis Model, query video, retrieval, visual features

## INTRODUCTION

Multimedia information retrieval is required to organize multimedia information such as video, image and to help people in finding multimedia resources easily and quickly. One of considerable application of multimedia information retrieval is video mining. Besides, visual content based video retrieval has been increasingly employed to retrieve required videos from a large collection on the basis of features that are extracted from the videos. A lot of research works have been intended for performing video retrieval process by using their visual content. But, performance of conventional work is not effective. Therefore, there is a requirement for efficient video retrieval system to search relevant video contents from a large dataset based on user query video.

A Content Based Video Shot Classification was introduced in [1] based on support vector machine with objective of enhancing the performance of content based video classification with higher speed of data retrieval operations. This strategy attains an improved classification performance. But, retrieval time was higher. A Large-Scale Video Retrieval was designed in [2] with objective of enhancing scalability of video retrieval and attaining memory-efficient retrieval. However, precision and recall rate of video retrieval was poor.

An automated video indexing and video search approach was designed in [3] with aid of Optical Character Recognition (OCR) technology and Automatic Speech Recognition (ASR) for content-based video browsing and search. However, the amount of time taken for retrieving relevant videos was more. A review of different techniques designed for visual content-based video indexing and retrieval was presented in [4] for video structure analysis.

A content-based recommender system was developed in [5] to extract video contents with higher accuracy. However, true positive rate of video retrieval was poor. A survey of different video retrieval technique developed for human action analysis was presented in [6] to improve performance of content based information retrieval system. Another Content Based Video Retrieval was developed in [7] with application of frequency domain analysis and 2-D correlation algorithm to find out all the objects from video that matched with user's query image for efficient retrieval. However, precision and recall of video retrieval process was remained unsolved.

A Latent Semantic Indexing (LSI) technique was intended in [8] with aid of singular value decomposition and visual features to provide better video retrieving with minimum time. But, the performance of LSI technique was not effectual. Sequential indexing algorithm was introduced in [9] to increase the system performance of video retrieval and minimize video searching time. Sequential indexing algorithm retrieves the more number of similar video according to the user input query.

Histogram clustering technique was presented in [10] that used similarity parameter to group the video objects for efficient video content retrieval. But, time complexity of video retrieval was more.

In order to solve the above mentioned existing issues, the major contribution of GBEC Technique is formulated as,

- To achieve higher precision, recall for retrieving videos from a large database, GBEC technique is designed. GSTC Technique used Independent Component analysis model and Gradient Boost Ensemble Classifier for improving the performance of video retrieval in data mining.

[1]Department of Computer Science Manomaniam Sundaranar University
[2]Department of Computer Science Periyar University
**Correspondence and Reprint Requests:** Vijayan K

**Received:** August,15 2017 | **Accepted:** September, 17 2017 | **Published Online:** October, 28 2017

**Conflict of interest:** None declared | **Source of funding:** Nil

- To extract visual features such as shape, color, texture in videos, Independent Component analysis model is used in GSTC Technique and also estimates separation matrix that separate linearly mixed sources for efficient visual features extraction.
- To retrieve the similar videos in a dataset based on input query video, Gradient Boost Ensemble Classifier combines results of all base classifier into a strong classifier for attaining higher classification performance for video retrieval with minimum time complexity.

The rest of the paper structure is arranged as follows. Section 2 describes the process of GBEC technique with neat architecture diagram. In Section 3 the simulation environment is presented and the results are explained in section 4. Section 5 describes the related works. Finally, the conclusion of the paper is presented in section 6.

## Gradient Boost Ensemble Classification Based Video Retrieval

The GBEC technique is designed in order to achieve higher precision, recall with minimum time complexity for retrieving similar videos form a large database through performing classification. The GBEC Technique is efficient for video retrieval process. The overall Architecture diagram of GBEC technique is shown in below Figure 1.

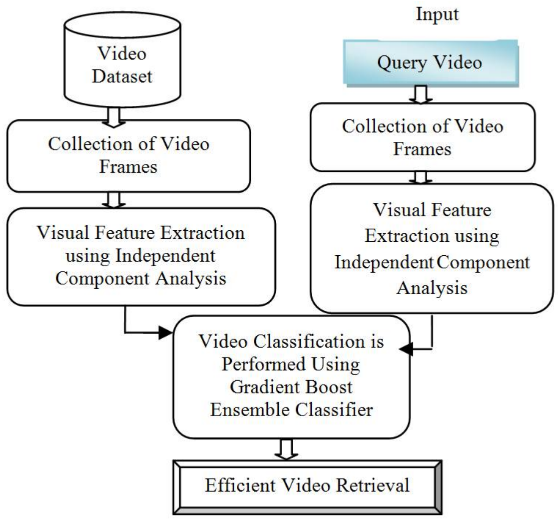


**Figure 1** Architecture Diagram of GBEC for Efficient Video Retrieval

As shown in Figure 1, GBEC technique at first takes query video as input and query video is partitioned into number of video frames for performing feature extraction process. Then, GBEC technique carried outs visual feature extraction process for both query video and videos in dataset with application of Independent Component Analysis (ICA). Finally, Gradient Boost Ensemble Classifier is applied for efficiently retrieving videos relevant to user query through classification. This helps to improve the precision and recall of video retrieval process with minimum time complexity.

## Independent Component Analysis for Visual Feature Extraction

The GBEC technique uses ICA Model to carry out visual feature extraction process during video retrieval. The frames in video are used for identifying visual features for example shape, color, texture. The visual feature extraction is significant for video retrieval.

ICA is a linear non-orthogonal transform that separates the independent source signals (i.e. video) from their linear mixtures without knowing the mixture matrix. ICA is an extension of classical principal component analysis. The principal component analysis is optimal in terms of reconstruction error in Euclidean space. The features extracted by principal component analysis are uncorrelated. However, ICA de-correlates the data and also lessens higher-order statistical dependence of data which helps for efficient feature extraction. The process of visual feature extraction using ICA Model is shown in below Figure 2.

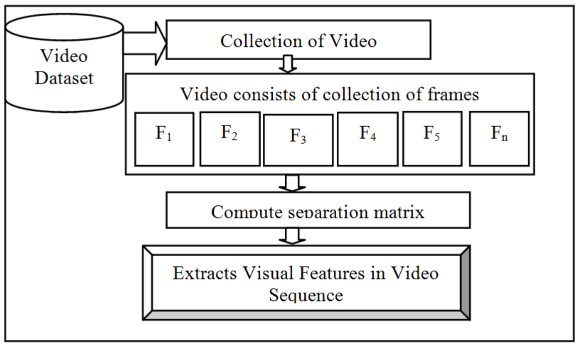


**Figure 2** Process of ICA Model for Visual Feature Extraction

In Figure 2, the video comprises of multiple video frames $F_1, F_2, .. F_n$. For each frame in video, ICA Model is applied for visual feature extraction and also computes the separation matrix to efficiently extract the visual features of frames in videos.

A training set of N videos in dataset is represented as $X = [X_1, X_2, X_3, .. X_n]^T$. Thus, ICA model is considered as follows,

$$X = AS \quad (1)$$

From equation (1), row vector $X$ is a zero-mean vector and is the concatenation of the column vectors representing the localized frame matrixes in a video sequence. The row vector S is the ICA basis vector and $A$ is a weight matrix to be determined. The key objective of ICA is to represent video frame with independent component bases and their interrelated co-efficient (i.e. visual features). ICA is performed on principal Eigen-vectors of the covariance matrix of input video sequence. After that, the separation matrix $W$ is computed by,

$$B = WX = WAX \quad (2)$$

$$W = A^{-1} \quad (3)$$

From equation (2) and (3), $B$ is the estimation of statistical independent base vector i.e. $B = (B_1, B_2, .., B_N)^T$ and each row represents a statistically independent basis vectors. The visual feature vector of the video frames to be identified is then projected onto the subspace, i.e. the linear combination of the set of base vectors to represent. Let $f$ be identified visual feature vector of the video frames which is extracted using,

$$f = a_1 B_1 + a_2 B_2 + \cdots . + a_N B_N \quad (4)$$

From equation (4), $B_1, B_2, .., B_N$ denotes N basis vector where $a_1, a_{2,..} a_n$ represents projection coefficients (i.e. extracted visual features of video sequence such as shape, color, texture). The algorithmic process of ICA for Visual Feature Extraction is shown in below,

**Input:** Training set of N videos in dataset
**Output:** Extraction of visual features in videos
**Step 1:Begin**
**Step 2:** **For** each videos in dataset
**Step 3:** **For** each frames of video
**Step 4:** Compute separation matrix using (2) and (3)
**Step 5:** Extract visual features in video frame using (4)
**Step 6:** **End for**
**Step 7:** **End for**
**Step 8:End**

***Algorithm 1 ICA based Visual Feature Extraction***

ICA based Visual Feature Extraction initially takes video dataset as input in algorithm 1. For each frame of video in dataset, ICA Model evaluates separation matrix for separating the linearly mixed sources and also projects visual feature vector of the video frames onto the subspace. Thus, ICA Model efficiently extracts the visual feature of videos.

### *Gradient Boost Ensemble Classifier*

After extracting the visual features in video dataset, gradient boost ensemble classifier classifies the videos in dataset into a different class based on their extracted visual features such as shape, color, textures. The gradient boost ensemble classifier designed in GBEC technique is a machine learning technique for solving classification and regression problems. The gradient boost ensemble classifier constructs a prediction model in the structure of an ensemble of weak decision trees prediction. The gradient boost ensemble classifier is an algorithm based on an ensemble of decision tree similar to random forests.

The gradient boost ensemble classifier categorizes the videos in a given into a similar or dissimilar class based on their extracted visual features. The classification of similar or dissimilar videos based on input query video. The gradient boost ensemble classifier is an ensemble of base classifier (i.e. random decision tree). The gradient boost ensemble classifier combines all base classifier into a strong classifier to present the final results for video classification. Figure 3 illustrates the black diagram of gradient boost ensemble classifier for classifying videos into a similar or dissimilar class according to their extracted visual features such as shape, color, texture etc.

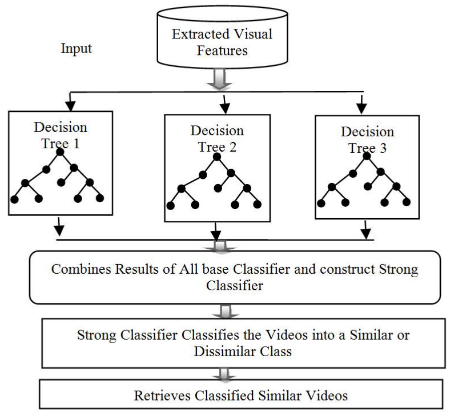


**Figure 3** Process of Gradient Boost Ensemble Classifier for Video Classification

Initially gradient boost ensemble classifier takes extracted visual features as input. For each visual features of video, then gradient boost ensemble classifier constructs decision trees. In gradient boost ensemble classifier, decision tree is a structure which comprises a root node, branches, and leaf nodes where each inner node represents a test on a visual features and the branch node indicates that a result value. The leaf node in the tree denotes a class label. The top node is referred to a root node of decision tree. After that, the constructed decision trees perform classification through comparing visual features of video with visual features of input query video. The results of decision trees are considered as base or weak classifiers in gradient boost ensemble classifier. Next, gradient boost ensemble classifier measures the loss function of all base classifiers to make strong classifier that efficiently classifies videos in a given dataset as similar or dissimilar classes. Thus, gradient boost ensemble classifier achieves classification accuracy for video retrieval with minimum time. At last, GBEC technique retrieves classified similar videos as relevant videos related to input query video which in turn increases the precision and recall of video retrieval in a significant manner.

In GBEC technique, the gradient boost ensemble classifier algorithm employs the prediction models and faults (i.e. loss function of decision tree classification) discovered by gradients for efficient video retrieval. The prediction model in gradient boost ensemble classifier considers the training set $(x_1, y_1), (x_2, y_2) \ldots . (x_i, y_i)$ which the process is to fit a model to loss function $f_{Loss}(x)$. Thus, the prediction output of the gradient boost ensemble classifier is formulated as,

$$y_i = f_{Loss}(x) + h(x_i) \quad (5)$$

From equation (5), $f_{Loss}(x)$ represents predicted loss function whereas $h(x)$ denotes the decision tree classification outputs. In gradient boost ensemble classifier, the loss function (i.e. error rate of decision tree classification) is determined as the differentiation between actual and predicted value. As a result, the loss function is calculated by,

$$f_{Loss}(x) = (y_i - h(x_i))^2 \quad (6)$$

From equation (6), $(y_i)$ denotes actual value and $h(x_i)$ indicates a predicted value. Here $h(x_i)$ represents the decision tree classifier results in which $x_i$ refers to the number of visual features such as shape, color, texture to classify video into a similar or dissimilar for efficient retrieval. The weighted sum functions of all decision tree classifier (i.e. strong classifier) is mathematically expressed as,

$$h(x_i) = h_1(x) + h_1(x) + \cdots + h_n(x) \quad (7)$$

Afterward, the gradient boost ensemble classifier fit a base classifier $h(x_i)$ to pseudo-residuals with training visual features of videos in dataset. Therefore, pseudo-residuals (PR) function is mathematically formulated as below,

$$PR = -\left[\frac{\partial(y_i, f_{Loss}(x_i))}{\partial f_{Loss}(x_i)}\right] \; where \; i = 1,2,3.n \quad (8)$$

From equation (8), the PR is determined. Then, weak classifier trains on the remaining errors (i.e. pseudo-residuals). For each iteration, the PR is determined and a weak learner is fitted to these pseudo-residuals which help to minimize the overall error of the strong learner. Next, fit a decision tree to an input training set (i.e. extracted visual features of videos) is performed. The output of the first predictive classifier $h_1(x)$ with the extracted visual features of video is obtained by,

$$h(x_i) = \sum_{i=1}^{n}\{x_i, (y_i - f_{Loss}(x_i)\} \quad (9)$$

$$h_1(x) = (x_1, (y_1 - f_{Loss}(x_1)) \quad (10)$$

In the same way, the output of the second weak classifier $h_2(x)$ is formulated as,

$$h_2(x) = (x_2, (y_2 - f_{Loss}(x_2)) \quad (11)$$

Thus, the output of the last weak classifier $h_n(x)$ is obtained using below,

$$h_n(x) = (x_n, (y_n - f_{Loss}(x_n)) \quad (12)$$

From equation (10), (11), (12), outputs of all weak classifiers $h_1(x) + h_1(x) + \cdots + h_n(x)$ are combined in order to create a strong classifier for categorizing videos into a similar or dissimilar. As a result, the predictor's classifier function $h(x)$ in gradient boost ensemble classifier reduces the total loss function $f_{Loss}(x)$ of video classification significantly. The best gradient output is discovered as strong classifier for classifying the videos as similar or dissimilar based on input video query. Therefore, the best gradient descent step-size ($\rho_{best}$ ) is find out using below,

$$\rho_{best} = arg\min_{\rho} \sum_{i=1}^{n}[y, f_{Loss-1}(x_i) + \rho h(x_i)] \quad (13)$$

At last, gradient boost ensemble classifier algorithm updates the model to classify videos as similar or dissimilar, $y_i = \sum_{i=1}^{n} f_{Loss-1}(x_i) + \rho_{best} h(x_i)$ (14)

From equation (14), $y_i$ indicates the strong classifier output for effectual video classification and retrieval. Hence, an ensemble of weak classifier is used to make strong classifiers and performs the better classification until the accuracy of model is achieved. The strong classifier output is $y_i \in \{0,1\}$ to classify the videos based on the query video. Here, 1 denotes the video is similar to input query video whereas 0 denotes the video is dissimilar to input query video. The algorithmic process of gradient boost ensemble classifier is explained in below,

**Input** : Number of visual features, training sets
**Output :** Increased classification accuracy with minimum time complxity
**Step 1**: **Begin**
**Step 2**: **For** each visual features of video in a dataset
**Step 3**: Construct decision tree classifier
**Step 3**: Compute the loss function using (6)
**Step 4**: Compute the pseudo-residuals using (7)
**Step 5**: Fit a base classifier $h(x_i)$ to pseudo-residuals with training visual features using (8)
**Step 6:** Find out the best gradient descent step-size using (12)
**Step 7**: Update the model as strong classifier and provides classification results '$y_i$'
**Step 8:** **If** ($y_i$ results is '1' ) **then**
**Step 9:** Video is classified as similar based on query video
**Step 10**: **else**
**Step 11:** Video is classified as dissimilar
**Step 12**: **End if**
**Step 13:** **End for**
**Step 14**: **End**

### *Algorithm 2 Gradient Boost Ensemble Classifier for Video Retrieval*

Initially builds decision tree classifier using visual features of video. After that, the error rate of base classifier is estimated to improve video classification performance. Then the base classifier is fit to pseudo-residuals with training visual features. Subsequently, the best gradient descent step size is computed to get the strong classifier. The strong classifier efficiently classifies the videos in given dataset as similar or dissimilar based on visual features of query video. Finally, Gradient Boost Ensemble Classifier helps to enhance the accuracy of video classification with minimum time.

### *Experimental Setting*

GBEC technique is implemented in Java Language using three data sets such as VIRAT Video Dataset [21] and UCF Sports Action Data Set [22] and INRIA Holidays dataset [23] for conducting experimental work. The VIRAT Video Dataset employed in GBEC technique includes numerous videos recorded from 11 scenes (1080p or 720p). The UCF Sports Action Data Set used in GBEC technique comprises of lots of videos with a collection of actions grouped from different sports and a total of 150 sequences with the resolution of 720 x 480. Besides, INRIA Holidays dataset utilized in GBEC technique contains 500 image groups where each image group represents a dissimilar scene or object. For carried outing the experimental process, GBEC technique takes 100 videos and 100 images from above three datasets.

**Table 1** a) Tabulation of Classification Accuracy for Video Retrieval

| Number of videos | Classification Accuracy (%) | | | | | |
|---|---|---|---|---|---|---|
| | VIRAT Video Dataset | | | UCF Sports Action Data Set | | |
| | Content Based Video Shot Classification | Large-Scale Video Retrieval | GBEC technique | Content Based Video Shot Classification | Large-Scale Video Retrieval | GBEC technique |
| 10 | 60.25 | 69.23 | 83.53 | 62.26 | 75.56 | 85.45 |
| 20 | 63.36 | 70.56 | 85.64 | 63.35 | 77.36 | 86.12 |
| 30 | 63.47 | 72.45 | 86.12 | 65.74 | 78.15 | 87.95 |
| 40 | 64.69 | 73.12 | 86.98 | 66.25 | 78.78 | 88.53 |
| 50 | 66.23 | 73.96 | 88.12 | 68.69 | 80.65 | 89.26 |
| 60 | 66.90 | 75.32 | 88.90 | 69.64 | 80.34 | 90.88 |
| 70 | 67.14 | 77.58 | 90.13 | 70.32 | 83.71 | 93.98 |
| 80 | 68.36 | 78.62 | 91.89 | 71.12 | 84.65 | 95.12 |
| 90 | 70.87 | 79.82 | 94.35 | 73.87 | 87.76 | 95.90 |
| 100 | 72.68 | 82.75 | 95.16 | 74.38 | 88.25 | 96.13 |

**Table 1 b** Tabulation of Classification Accuracy for Image Retrieval

| Number of Images | Classification Accuracy (%) | | |
|---|---|---|---|
| | Content Based Video Shot Classification | Large-Scale Video Retrieval | GBEC technique |
| 10 | 63.54 | 72.35 | 87.32 |
| 20 | 64.58 | 73.58 | 88.87 |
| 30 | 65.98 | 74.69 | 89.54 |
| 40 | 66.47 | 76.12 | 90.12 |
| 50 | 67.69 | 77.65 | 91.47 |
| 60 | 68.87 | 78.96 | 92.56 |
| 70 | 69.84 | 79.51 | 93.65 |
| 80 | 72.54 | 81.25 | 94.75 |
| 90 | 74.56 | 82.54 | 95.87 |
| 100 | 75.98 | 83.69 | 97.86 |

## RESULTS AND DISCUSSION

In this section, the result of GBEC techniques is analyzed and compared with existing [1] and [2]. The effectiveness of GBEC techniques is evaluated along with the following metrics with the assist of tables and graphs.

### Measurement of Classification Accuracy

The Classification Accuracy (CA) measures the ratio of number of correctly classified videos to the total number of videos. The CA is measured in percentages (%) and formulated as,

$$CA = \frac{number\ of\ correctly\ classified\ videos}{total\ number\ of\ videos} * 100 \quad (15)$$

From equation (15), the classification accuracy of videos is estimated with respect to dissimilar number of input videos. When the classification accuracy is higher, the method is said to be more efficient.

The classification accuracy result is obtained for both video and image retrieval using three methods in Table 1 a) and b).

The GBEC technique considers different number of videos and images in the range of 10-100 for conducting the experimentalworks. While increasing the number of videos or images for experimental work, the classification accuracy is also gets increased using all three methods. But comparatively, classification accuracy using GBEC techniques is higher. This is because of application of gradient boost ensemble classifier in GBEC techniques. The gradient boost ensemble classifier combines the results of all base classifier into a strong classifier for efficiently classifying the videos in dataset into a similar or dissimilar. Therefore, proposed GBEC technique improves the classification accuracy of video retrieval using VIRAT Video Dataset by 34 % and 18% as compared to existing [1] and [2]. Further, proposed GBEC technique increase the classification accuracy of video retrieval using UCF Sports Action Data Set by 33 % and 12%. Besides, proposed GBEC technique increases the classification accuracy of image retrieval using INRIA Holidays dataset by 34 % and 18% as compared to existing [1] and [2].

### Measurement of Time Complexity

Time Complexity (TC) determines the amount of time taken for retrieving similar videos form a given a dataset. TC is measured in milliseconds (ms) and expressed as,

$$TC = N * Time(Retrieving\ one\ video) \quad (16)$$

From equation (16), $N$ represents the number of videos taken. When the time complexity is lower, the method is said to be more effective.

Figure 4 a) and b) describes the result analysis of time complexity for video and image retrieval using three methods with respect to varying number of videos and images in the range of 10-100.

While increasing the number of videos or images for carried outing experimental process, the time complexity is also gets increased using all three methods. But comparatively, time complexity using GBEC techniques is lower. This is owing to application of gradient boost ensemble classifier in GBEC techniques where it unites the results of all base classifier into a strong classifier for categorizing the videos into a similar or

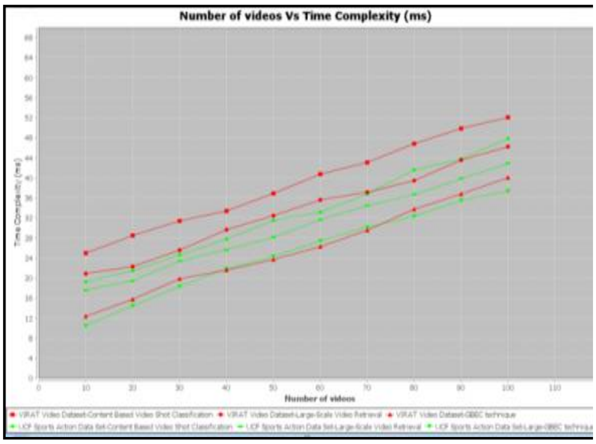


**Figure 4 a** Measurement of Time Complexity for Video Retrieval

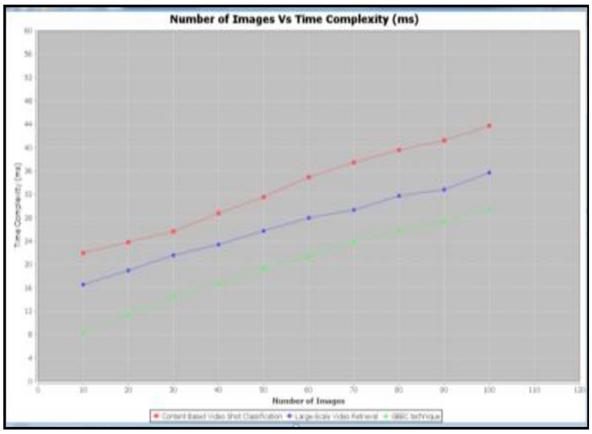


**Figure 4 b** Measurement of Time Complexity for Image Retrieval

dissimilar based on query video. This helps for GBEC techniques to efficiently retrieve the similar videos related to query video with minimum time. As a result, proposed GBEC technique minimizes the time complexity of video retrieval using VIRAT Video Dataset by 35 % and 24% as compared to existing [1] and [2]. Moreover, proposed GBEC technique lessens the time complexity of video retrieval using UCF Sports Action Data Set by 25 % and 18%. As well, proposed GBEC technique reduces the time complexity of image retrieval using INRIA Holidays dataset by 41 % and 27% as compared to existing [1] and [2].

### Measurement of Precision

Precision defines the ratio of number of similar videos retrieved based on query video to total number of videos. The precision is measured in percentages (%) and expressed as,

$$Precision\ (P) = \frac{Number\ of\ similar\ videos\ retrieved}{Total\ number\ of\ videos} * 100 \quad (17)$$

From equation (17), the precision rate of video retrieval is determined with respect to diverse number of videos. When the precision value is higher, the method is said to be more effectual.

The precision result is obtained for both video and image retrieval using three methods is depicted in Table 2 a) and b). Further, while increasing the number of videos or images for conducting experimental work, the precision is also gets increased using all three methods.

**Table 2 a** Tabulation of Precision for Video Retrieval

| Number of videos | Precision (%) | | | | | |
|---|---|---|---|---|---|---|
| | VIRAT Video Dataset | | | UCF Sports Action Data Set | | |
| | Content Based Video Shot Classification | Large-Scale Video Retrieval | GBEC technique | Content Based Video Shot Classification | Large-Scale Video Retrieval | GBEC technique |
| 10 | 61.87 | 69.87 | 85.52 | 70.12 | 80.12 | 89.36 |
| 20 | 62.56 | 71.54 | 86.71 | 71.56 | 81. 25 | 90.15 |
| 30 | 63.47 | 72.65 | 87.45 | 72.69 | 82.63 | 91.25 |
| 40 | 64.89 | 73.96 | 88.51 | 73.58 | 83.79 | 92.74 |
| 50 | 65.71 | 74.56 | 89.65 | 74.32 | 84.79 | 93.65 |
| 60 | 66.98 | 75.96 | 90.12 | 75.96 | 85.45 | 94.12 |
| 70 | 67.19 | 76.58 | 92.78 | 76.10 | 86.65 | 94.92 |
| 80 | 68.32 | 79.68 | 93.65 | 77.45 | 87.15 | 95.68 |
| 90 | 69.45 | 81.56 | 94.72 | 78.45 | 88.26 | 96.23 |
| 100 | 71.58 | 82.45 | 96.56 | 79.63 | 89.69 | 97.12 |

**Table 2 b** Tabulation of Precision for Image Retrieval

| Number of Images | Precision (%) | | |
|---|---|---|---|
| | Content Based Video Shot Classification | Large-Scale Video Retrieval | GBEC technique |
| 10 | 74.26 | 85.68 | 90.23 |
| 20 | 75.84 | 86.96 | 90.89 |
| 30 | 77.96 | 87.49 | 91.54 |
| 40 | 78.96 | 88.96 | 92.14 |
| 50 | 79.65 | 89.47 | 92.91 |
| 60 | 80.14 | 91.78 | 93.56 |
| 70 | 81.56 | 91.74 | 94.23 |
| 80 | 82.69 | 92.82 | 95.01 |
| 90 | 83.96 | 93.24 | 95.76 |
| 100 | 84.69 | 93.65 | 96.98 |

But comparatively, precision using GBEC techniques is higher. This is due to application of gradient boost ensemble classifier in GBEC techniques where it compares the visual features of videos in a given dataset with visual features query video for video classification. This supports for categorizing the videos into a similar or dissimilar based on query video. Therefore, GBEC techniques to significantly retrieve the similar videos interrelated to query video. Thus, proposed GBEC technique enhances the precision of video retrieval using VIRAT Video Dataset by 37 % and 19% as compared to existing [1] and [2]. Furthermore, proposed GBEC technique increases the precision of video retrieval using UCF Sports Action Data Set by 25 % and 10%. Also, proposed GBEC technique improves the precision of image retrieval using INRIA Holidays dataset by 17 % and 4% as compared to existing [1] and [2].

### Measurement of Recall

Recall defines the ratio of number of correctly retrieved similar videos based on query video to total number of videos. The recall is measured in percentages (%) and expressed as,

$$Recall = \frac{Number\ of\ correctly\ retrieved\ similar\ videos}{Total\ number\ of\ videos} * 100 \qquad (18)$$

From equation (18), the recall of video retrieval is estimated with respect to different number of videos. When the precision value is higher, the method is said to be more effective.

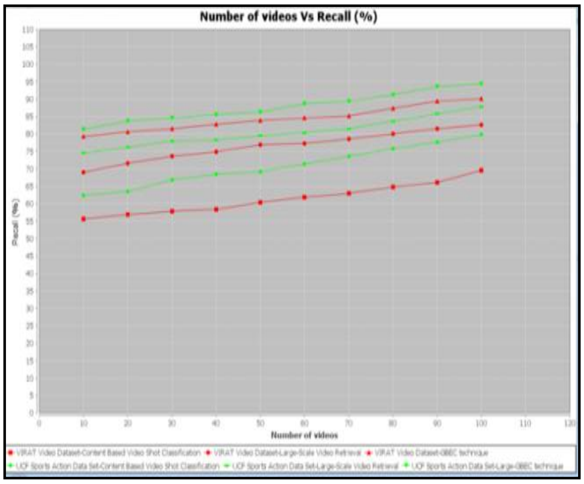


**Figure 5 a** Measurement of Recall for Video Retrieval

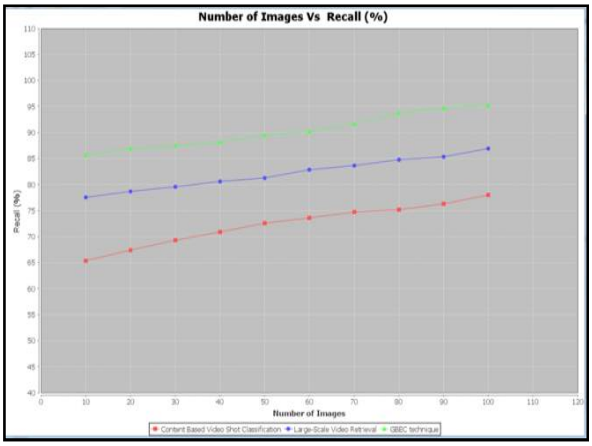


**Figure 5 b** Measurement of Recall for Video Retrieval

Figure 5 a) and b) shows the performance of recall results for video and image retrieval using three methods with respect to different number of videos and images in the range of 10-100.

Moreover, while increasing the number of videos or images for performing experimental work, the recall is also gets increased using all three methods. But comparatively, recall using GBEC techniques is higher. This is because of usage of gradient boost ensemble classifier in GBEC techniques where it employs the visual features for efficient video classification and retrieval. This assists for GBEC techniques to considerably retrieving the more similar videos based on query video. Hence, proposed GBEC technique improves the recall of video retrieval using VIRAT Video Dataset by 37 % and 10% as compared to existing [1] and [2]. Besides, proposed GBEC technique increases the recall of video retrieval using UCF Sports Action Data Set by 24 % and 9%. In addition, proposed GBEC technique improves the recall of image retrieval using INRIA Holidays dataset by 25 % and 10% as compared to existing [1] and [2].

### Related Works

A Discrete wavelet transforms and sparse representation was employed in [11] to improve the accuracy of content based video retrieval. However, the recall of video retrieval was poor. A sample-based hierarchical adaptive K-means clustering method was presented in [12] for carry outing large-scale video retrieval with higher accuracy and minimum time complexity. However, the precision of video retrieval was remained unsolved.

Shrinkage optimized directed information assessment (SODA) approach was introduced in [13] with application of directed information for multimodal video indexing and retrieval. But, video retrieval rate was poor. A novel approach was designed in [14] by integrating visual features to enhance performance of video retrieval. But, accuracy and recall of video retrieval process was lower.

A survey of diverse techniques developed for content based video retrieval systems and their improvements, confronts was studied in [15]. Multi feature content based Video Retrieval was intended in [16] with application of high level semantic concepts to retrieve required video from database. However, efficiency of video retrieval was not enough.

An Enhanced video retrieval and classification of video database using multiple frames was designed in [17] to reduce the computational cost of video retrieval process. But, classification accuracy of video was remained unsolved. Color feature based Video Retrieval System was presented in [18] to mine relevant videos by using most dominant color of key frame in a video. However, performance of color feature based video retrieval system was not effective.

A Latent topics based relevance feedback scheme was introduced in [19] for performing effectual video retrieval. But, computational time taken for video retrieval was higher. A Spatio-Temporal Pyramid Matching (STPM) was presented in [20] for extracting most relevant video from video database and increase the precision rate of video retrieval. But, recall results of video retrieval was lower.

## CONCLUSION

The GBEC technique is designed with aiming at improving the performance of visual content based video retrieval. The GBEC technique initially takes video query as input. After that, GBEC technique applied ICA Model for extracting visual features in videos. The ICA Model evaluates the separation matrix for each video frame which helps for significantly extracting the visual features such as shape, color, texture for video classification. After visual feature extraction process, GBEC technique employed Gradient Boost Ensemble Classifier for efficiently retrieving similar videos related to query video from a large dataset. The Gradient Boost Ensemble Classifier combines the results of all base classifier into a strong classifier for classifying the videos as similar or dissimilar based on video query which improves classification accuracy for video retrieval. Finally, GBEC technique retrieves the classified similar videos' as relevant videos related video query which resulting in increased precision and recall of video retrieval. The effectiveness of GBEC technique is measured in terms of classification accuracy, time complexity, Precision and recall by using three datasets and compared against with state of the art works. With the experiments carried out for GBEC technique, it is expressive that the recall of video retrieval affords more precise results and minimization of time complexity as compared to state-of-the-art works.

## References


1. Chiranjoy Chattopadhyay, Amit Kumar Maurya, "Genre-specific modeling of visual features for efficient content based video shot classification and retrieval", Springer, Volume 2, Issue 4, Pages 289-297, November 2013
2. Andre Araujo and Bernd Girod, "Large-Scale Video Retrieval Using Image Queries", IEEE Transactions On Circuits And Systems For Video Technology, Volume PP, Issue 99, Pages 1-14, 2017
3. Haojin Yang and Christoph Meinel, "Content Based Lecture Video Retrieval Using Speech and Video Text Information", IEEE Transactions on Learning Technologies, Volume 7, Issue 2, Pages 142 - 154, April-June 2014
4. Weiming Hu, Nianhua Xie, Li Li, Xianglin Zeng, Stephen Maybank, "A Survey on Visual Content-Based Video Indexing and Retrieval", IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), Volume 41, Issue 6, Pages 797 - 819, 2011
5. Yashar Deldjoo, Mehdi Elahi1, Paolo Cremonesi1, Franca Garzotto1, Pietro Piazzolla, Massimo Quadrana, "Content-Based Video Recommendation System Based on Stylistic Visual Features", *Journal on Data Semantics, Springer*, Volume 5, Issue 2, Pages 99–113, June 2016
6. Mohsen Ramezani, Farzin Yaghmaee, "A review on human action analysis in videos for retrieval applications", Artificial Intelligence Review, Springer, Volume 46, Issue 4, Pages 485-514, December 2016
7. 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* Volume 4, Issue 9, Pages 388-393, September 2014
8. 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
9. 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
10. D.Saravanan, Vaithyasubramanian, K.N. Jothi Vengatesh, "Video Content Retrieval Using Historgram Clustering Technique", Procedia Computer Science, Elsevier, Volume 50, Pages 560- 565, 2015.
11. Sajad Mohamadzadeh and Hassan Farsi, "Content Based Video Retrieval Based On HDWT and Sparse Representation", Image Analysis and Strereology, Volume 35, Issue 2, Pages 67-80, 2016
12. Kaiyang Liao, Guizhong Liu, Li Xiao, Chaoteng Liu, "A sample-based hierarchical adaptive K-means clustering method for large-scale video retrieval", Knowledge-Based Systems, Elsevier, Volume 49, Pages 123–133, 2013
13. Xu Chen, Alfred O. Hero, and Silvio Savarese, "Multimodal Video Indexing and Retrieval Using Directed Information", IEEE Transactions on Multimedia, Volume 14, Issue 1, Pages 3 - 16, February 2012

14. Stefanos Vrochidis, Ioannis Kompatsiaris, and Ioannis Patras, "Utilizing Implicit User Feedback to Improve Interactive Video Retrieval", Hindawi Publishing Corporation, Advances in Multimedia, Volume 2011, Article ID 310762, Pages 1-18, 2011
15. Aasif Ansari, Muzammil H Mohammed, "Content based Video Retrieval Systems - Methods, Techniques, Trends and Challenges", *International Journal of Computer Applications,* Volume 112, Issue 7, Pages 13-22, February 2015
16. Hatim G. Zaini and T. Frag, "Multi feature content based Video Retrieval Using High Level Semantic Concepts", International Journal of Computer Science, Volume 2, Issue 9, Pages 15-27, September 2014
17. Mohd.Aasif Ansari, Hemlatavasishtha, "Enhanced video retrieval and classification of video database using multiple frames based on MBTC Features", *International Journal of Advances in Electronics and Computer Science*, Volume-2, Issue-10, Pages 31-38, 2015
18. Tejaswi Potluri, Gnaneswararao Nitta, "Content Based Video Retrieval Using Dominant Color of the Truncated Blocks of Frame", *Journal of Theoretical and Applied Information Technology*, Volume 85, Issue 2, Pages 165-171, 2016
19. RubenFernandez-Beltran, FilibertoPla, "Latent topics-based relevance feedback for video retrieval", Pattern Recognition, Elsevier, Volume 51, Pages 72-84, March 2016
20. JaesikChoia, ZiyuWangbSang-ChulLeeb, Won J.Jeonc, "A spatio-temporal pyramid matching for video retrieval", Computer Vision and Image Understanding, Elsevier, Volume 117, Issue 6, Pages 660-669, June 2013
21. VIRAT Video Dataset: http://www.viratdata.org/
22. UCF Sports Action Data Set: http://crcv.ucf.edu/data/UCF_Sports_Action.php
23. INRIA Holidays dataset: http://lear.inrialpes.fr/~jegou/data.php


*******