International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 7 (2015) pp. 16219-16236

http://www.ripublication.com

Content Based Video Retrieval (CBVR) in Multimedia Video Sequences Using Adaptive Fuzzy C-Means (AFCM) Clustering

Manjunath R.
Research Scholar, Tumkur University
Department of Computer Science & Engineering, City Engineering College
Doddakalasandra, Off Kankapura Road, Bangalore-560061
manjunathr0779@gmail.com

S. Balaji
Centre for Emerging Technologies, Jain Global Campus, Jain University
Jakkasandra P.O., Kankapura Taluk, Ramanagara Dist, Bangalore-562112
drsbalaji@gmail.com

Abstract

Content Based Video Retrieval *(*CBVR*)* system concerns with the application of video segmentation technique for the retrieval of videos in multimedia. The existing clustering based methods efficiently cluster the data but some drawbacks are still present in the existing clustering based retrieval systems. The clustering based retrieval methods result in over classification and inaccurate final decision. Hence, the existing clustering methods degrade the retrieval performance. To overcome the drawbacks in the existing techniques we propose a new CBVR technique using Adaptive Fuzzy C-Means (AFCM) clustering method. The proposed technique comprises of three stages: feature extraction, clustering and video retrieval. Texture and color are the two features used in the feature extraction process. Texture feature is extracted by applying wavelet method and the color feature is extracted through color histogram. Then the extracted features are clustered using adaptive fuzzy-C-means clustering method. In retrieval phase, the distance between the query image feature and the cluster centroid values are computed. The videos that have less distance with the clustered indexing number aree retrieved. Thus, the videos can be retrieved more accurately by achieving higher retrieval rate compared to the conventional techniques. The results are analyzed to demonstrate the superior performance of the proposed AFCM content based video retrieval technique.

Keywords: *Adaptive Fuzzy C-Means, Content Based Video Retrieval, Clustering, LGXP, Haar Wavelet, Color and Texture Feature*

\

1. Introduction

Multimedia is defined as the combination of more than one media [10]. The term multimedia refers to the data such as text, image, video, audio, graphical, relational and categorical data [1]. They are classified as two types: static and dynamic media. Text, graphics and images are categorized as static media while objects like animation, music, audio, speech, and video are categorized as dynamic media [8]. Multimedia data contains enormous amount of information [12]. Data mining techniques navigate through large amounts of databases and extracts specific and required data, which get translated and transcribed into useful and predictive information as per needs [11]. Multimedia is more complex - as the sequence progresses, the concept being studied may change as well [7]. Understanding and representing changes in the mining process is necessary to mine multimedia data [4].

Data in the multimedia databases are semi-structured or unstructured [15]. Unstructured data is simply a bit stream. This kind of data is not broken down into smaller logical structures and is not typically interpreted by the database [6]. Data mining is the extraction of hidden predictive information from large database which is almost unstructured or semi-structured data [10]. The ever improvements in the technology have posed new challenges to data mining, the various challenges being data formats, data from disparate locations, advances in computation, networking resources, research and scientific fields, ever growing business challenges etc [14]. The list of their applications spans from distance-learning, digital libraries, and home entertainment to fine arts, fundamental and applied science and research [20]. The methods of multimedia data mining are better known as multimedia information analysis and retrieval [17]. Multimedia retrieval is a technique of searching for texts, images or videos in large databases [18]. By a suitable matching technique the similarity between the two sets of multimedia data, which can either be images or videos can be found. The parameters of such a technique are: (i) level of abstraction of features, (ii) distance measures, and (iii) normalization of features [5]. The features which have been commonly used are color, shape, textures and spatial distribution [19].

The textual information is usually non-existent or incomplete with the emergence of massive multimedia databases. To address these problems, content-based multimedia information retrieval is used [3]. Effective and efficient video retrieval primarily lies in two aspects: (i) audio and (ii) video retrieval both of which comprise two methods, namely, indexing and searching phases [16]. In video, a typical model-based coding is first analyzed to estimate local and global motion. Then the video is synthesized using the estimated parameters. Based on the difference between the real video and synthesized video, the model parameters are updated and finally coded for transmission. This is essentially the analysis followed by synthesis, followed by model update and followed by coding [9].

The rest of this paper is organized as follows. In Section 2, brief description about the recent related research is given. Section 3 presents proposed CBVR using adaptive fuzzy C-means technique. Section 4 shows empirical evaluations of our experimental results and further discussion on the results, and Section 5 gives the concluding remarks.

2. Recent Related Research

Ja-Hwung Su *et al.* [21] have proposed an innovative method to achieve the high quality of content-based video retrieval by discovering the temporal patterns in the video contents. On the basis of the discovered temporal patterns, an efficient indexing technique and an effective sequence matching technique were integrated to reduce the computation cost and to raise the retrieval accuracy, respectively. Experimental results have shown that their approach was very promising in enhancing content-based video retrieval in terms of efficiency and effectiveness.

Morand *et al.* [22] have proposed a scalable indexing of video content by objects. A method for scalable moving object extraction was designed. Using the wavelet data, it relies on the combination of robust global motion estimation with morphological color segmentation at a low spatial resolution. It was then refined using the scalable order of data. A descriptor was built only on the objects extracted. This descriptor was based on multi-scale histograms of wavelet coefficients of objects. The experiment have given promising results compared with Scale Invariant Feature Transform (SIFT) features extracted on segmented object.

Barbara Andre *et al.* [23] have proposed a content-based video retrieval method that uses an expert-annotated database. A local dense multi-scale description was proposed to keep the proper level of invariance; they introduce a video-mosaicing technique that provides large field-of-view mosaic images. To remove outliers, retrieval was followed by a geometrical approach that captures a statistical description of the spatial relationships between the local features. To evaluate the retrieval, they perform a simple nearest neighbor's classification with leave-one-patient-out cross-validation. From the results of binary and multi-class classification, they have shown that their approach outperforms, with statistical significance, several state-of-the art methods.

Huanbo Luan *et al.* [24] have introduced an effective interactive video retrieval system named Vision Go. It jointly explores human and computer to accomplish video retrieval with high effectiveness and efficiency. It assists the interactive video retrieval process in different aspects: (1) it maximizes the interaction efficiency between human and computer by providing a user interface that supports highly effective user annotation and an intuitive visualization of retrieval results; (2) it employs a multiple feedback technique that assists users in choosing proper method to enhance relevance feedback performance; and (3) it facilitates users to assess the retrieval results of motion-related queries by using motion-icons instead of static key frames. Experimental videos have been shown to demonstrate the effectiveness of the Vision Go system.

Jaesik Choi *et al.* [25] have proposed a video matching called Spatio-Temporal Pyramid Matching (STPM). Considering features of objects in 2D space and time, STPM recursively divides a video clip into a 3D spatio-temporal pyramidal space and compares the features in different resolutions. In order to improve the retrieval performance, they consider both static and dynamic features of objects. They also provide a sufficient condition in which the matching can get the additional benefit from temporal information. The experimental results have shown that STPM performed better than the other video matching methods.

The Problem Statement

Our review of the literature expose several drawbacks in the existing techniques and emphasis the need for better techniques for multimedia data mining. The proposed research is outlined in the problem statement below:

Multimedia information systems are increasingly important with the advent of broadband networks, high powered workstations, and compression standards. Since visual media requires large amounts of storage and processing, there is a need to efficiently index, store, and retrieve the visual information from multimedia database. Video has become an important element of multimedia computing and communication environments, with applications as varied as broadcasting, education, publishing and military intelligence. Hence, knowledge discovery from the massive amount of multimedia data, so-called multimedia mining, has been the focus of attention over the past few years. In multimedia data mining, the mining or retrieval of videos has been brought to researchers' attention for a long time. Various research methods exist to retrieve the videos from the database. Among them, majority of the research methods were utilizing clustering methods in the features or image clustering process. The clustering based methods efficiently cluster the data, but some drawbacks still exist in the existing clustering based retrieval systems. The clustering based retrieval methods result in over classification and inaccurate final decision.

To minimize the drawbacks in the existing technique, we propose a new CBVR technique using adaptive fuzzy C-means clustering method. The proposed technique comprises three stages: feature extraction, clustering and video retrieval. In the proposed technique, the videos can be retrieved more accurately by achieving higher retrieval rate than the existing techniques. The results reveal superior performance of the proposed AFCM based content based video retrieval as compared to other methods.

3. The Proposed CBVR Using Adaptive Fuzzy C-Means

We propose a new content based video retrieval technique using Adaptive fuzzy C-means clustering method. This technique retrieves more related videos by using the algorithm for the query video. The proposed technique comprises of three stages: feature extraction, clustering and video retrieval. The architecture of the proposed method is shown in Figure1.

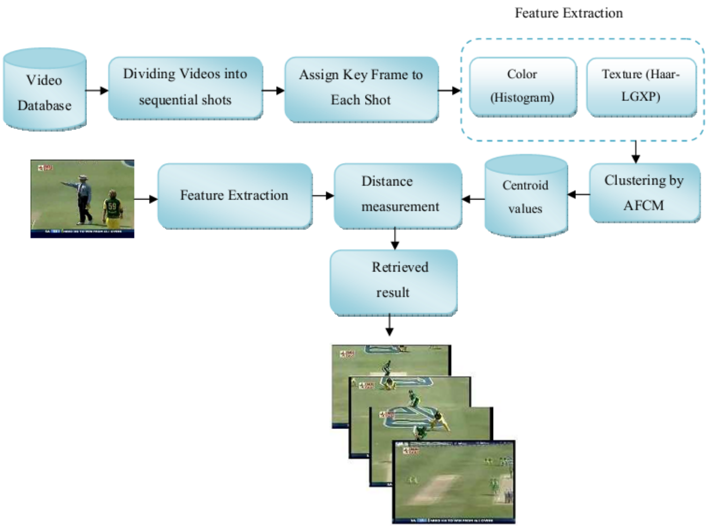


Figure 1: Architecture of the Proposed CBVR-AFCM Technique

Let us consider the video database $VD = \{v_1, v_2, \cdots v_i\}$, where i is the total number of videos in the database. The query video database is defined as $QD = \{q_1, q_2, \cdots q_j\}$, j represents the total number of query videos in the query database. Initially the video and query database videos v_i and q_j are divided into a set of sequential shots. Thus, the divided shots from the v_i and q_j videos are stated as,

$$v_i = \{s_1^i, s_2^i, s_3^i, \cdots s_n^i\} \tag{1}$$

$$q_j = \{s_1^j, s_2^j, s_3^j, \cdots s_m^j\} \tag{2}$$

In Equ. (1) and (2), n and m represent the number shots in the videos from the video and query database. To reduce the complexity and time, the given shots are pruned by eliminating some regions from the shots around the object. During pruning, the shots s_n^i and s_m^j are divided into 8x8 regions. The central 36 regions are kept and others are eliminated. After the shot segmentation, the feature extraction process is employed on each shot and thus the extracted features are exploited in the retrieval process. The three stages of the proposed CBVR-AFCM technique is further explained below.

3.1. Feature Extraction

In the proposed video retrieval technique, we utilize texture and color features in which the texture features are extracted by applying Haar wavelet transform and the color features are extracted through color histogram.

- Extract the color feature and texture feature from each shot.
- It is obvious that, if more number of features is used to represent the data, then the retrieval accuracy will be high.
- However, the feature vector dimension increases with increasing number of features; hence, there is a tradeoff between the retrieval accuracy and the complexity.
- So, it is essential to have minimal features representing the videos, compactly.

3.1.1. Texture Extraction Using LGXP

Here the video shots s_n^i and s_m^j are subjected to texture feature extraction. First, the Haar wavelet transform is applied and then the texture features are extracted from the wavelet band by exploiting the local Gabor XOR patterns (LGXP) method.

Haar Wavelet Transform:

Haar wavelet is the simplest type of wavelet which is normally employed for signal and image smoothening by considering its "energy compaction" properties, namely, large values are likely to become larger and small values smaller. One distinctive feature that the Haar transform contains is that it lends itself easily to simple hand calculations, memory efficient and also it is exactly reversible without the edge effects. In the Haar wavelet transform, LL band is the most significant band in which texture can be extracted more efficiently. Thus, using the Haar wavelet transform, LL band is obtained from the video shots s_n^i and s_m^j .

The Haar Wavelet is discontinuous and resembles a step function, for a function f , HWT is defined as:

$$f \to a^L / d^L \tag{3}$$

$$a^{L} = (a_1, a_2 \ldots\ldots a_{N/2}) \tag{4}$$

$$d^{L} = (d_1, d_2 \ldots\ldots d_{N/2}) \tag{5}$$

where L is the decomposition level, a is the approximation sub band and d is the detailed sub band.

$$a_m = \frac{f_{2m} + f_{2m-1}}{\sqrt{2}} \qquad \text{for m} = 1, 2, \ldots\ldots, N/2 \tag{6}$$

$$d_m = \frac{f_{2m} - f_{2m-1}}{\sqrt{2}} \qquad \text{for m} = 1, 2, \ldots\ldots, N/2 \tag{7}$$

For example, if f $= \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8\}$ is a time signal of length 8, HWT decomposes f into an approximation sub-band containing the low frequencies and detailed sub-band containing high frequencies:

$$\text{Low} = \text{a} = \{f_2 + f_1, f_4 + f_3, f_6 + f_5, f_8 + f_7\} / \sqrt{2} \tag{8}$$

$$\text{High} = \text{d} = \{f_2 - f_1, f_4 - f_3, f_6 - f_5, f_8 - f_7\} / \sqrt{2} \tag{9}$$

Here, we apply a one level Haar wavelet on video shots to each row and each column of the resulting image. The resulting image is decomposed into four sub-bands LL, HL, LH and HH band (L= Low and H= High). The LL sub-band contains the approximation of the original image while the other sub-band contains missing details. Hence, LL band is the most significant band in which texture can be extracted more efficiently.

LL	HL
LH	HH

Figure 2: Four Sub-bands of One-level Haar Wavelet

The portioned video shots which are obtained from the one-level wavelet transform is defined as $w(s_n^i)$ and $w(s_m^j)$. Thus, the obtained video shots LL band is exploited in the texture feature extraction process.

3.2.1. Local Gabor XOR Patterns (LGXP)

After obtaining $w(s_n^i)$ and $w(s_m^j)$, texture features $t(w(s_n^i))$, $t(w(s_m^j))$ are extracted by exploiting the LGXP method. Here, the RGB color images are converted into gray-scale images. Then, preprocessing is done by using the Gabor filter.

In LGXP, frames or shots $w(s_n^i)$ and $w(s_m^j)$ are divided into 3*3 blocks based on the size. Then compare every pixel value of the block with its neighbors from left to right within the block as shown in Figure 3. If its value is greater than the neighbor pixel value then replace it as 1 else as 0. For every block repeat the same process. Then convert binary into decimal value. Replace the original mid pixel value with the new computed value.

$$0,\ if \quad <= neighbour$$
$$1,\ else \tag{10}$$

$$p = \{\ \}$$

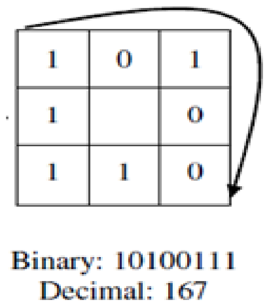


Figure 3: Blocks and Pixel Values in the Images

We obtain the decimal values for each pixel in the block. Then compute the histogram for each pixel value in the block. The texture features $t(w(s_n^i))$,$t(w(s_m^j))$ are extracted from the video shots based on the count value.

3.1.2. Color features

Color is one of the most widely used visual features in multimedia context and video retrieval. Here, color features $c(s_n^i)$ and $c(s_m^j)$ are extracted by applying histogram on the video shots s_n^i and s_m^j . The histogram provides a compact summarization of the distribution of data in an image or video.

Histogram

Color histogram h (s_n^i , s_m^j) =(h_k (s_n^i , s_m^j .) k=1,...,K) is a K-dimensional vector such that each component h_k (s_n^i , s_m^j) represents the relative number of pixels of color C_k in the image, that is, the fraction of pixels that are most similar to the corresponding color. To build the color histogram, the image colors should be transformed to an appropriate color space and quantized according to a particular codebook of the size K.

$$p_{(x)}i = \frac{n_i}{n},\ 0 \le i < L \tag{1}$$

In Equ. (11), L being the total number of gray-levels in the image, n_i is the number of occurrences of gray-level *I, n* being the total number of pixels in the image, and $p_x(i)$ is the image histogram for pixel value i.

Computationally, the color histogram is formed by discrediting the colors within an image and counting the number of pixels of each color. Color descriptors of video can be global and local. Global descriptors specify the overall color content of the image but with no information about the spatial distribution of these colors. Local descriptors relate to particular image regions and in conjunction with geometric properties of these latter, describe also the spatial arrangement of the colors. By comparing histogram signatures of two videos and matching the color content of one video with the other, the color feature $c(s_n^i)$,$c(s_m^j)$ can be extracted.

The extracted texture $t(w(s_m^j))$ and color feature $c(s_m^j)$ obtained from the query videos were stored in the feature database. Similarly, the texture $t(w(s_n^i))$ and color feature ($c(s_n^i)$) extracted from the video database (VD) are subjected to clustering using adaptive fuzzy C- means.

3.2. Clustering by Adaptive Fuzzy C- Means

Here, the extracted texture feature ($t(w(s_n^i))$) and color features ($c(s_n^i)$) are clustered by using the adaptive fuzzy-C-means clustering method. The fuzzy C-means algorithm is commonly used for clustering where the performance of the FCM depends on the selection of initial cluster center or membership value. The FCM algorithm starts with a set of initial cluster centers (or) arbitrary membership values.

The FCM algorithm assigns pixels to each category by using fuzzy memberships.

$$J_m = \sum_{i=1}^{I}\sum_{j=1}^{J}(\mu_{ij})^m \left\| x_i - z_j \right\|^2$$

(12)

In Equ. (12), x_i represents the texture and color features $t(w(s_n^i))$, $c(s_n^i)$ extracted from the video database, z_j is the j th cluster centre and m is the constant value.

The membership function represents the probability that a pixel belongs to a specific cluster. In the FCM algorithm, the probability is dependent on the distance between the pixel and each individual cluster center in the feature domain. The membership functions and cluster centers are updated by the equations (13) and (15).

$$u_{ij} = \frac{1}{\sum_{k=1}^{J}\left(\frac{\left\| x_i - z_j \right\|}{\left\| x_i - z_k \right\|}\right)^{\frac{2}{m-1}}} \tag{13}$$

Repeat the algorithm until the coefficients' change between two iterations is no more than ξ , for the given sensitivity threshold.

$$\max_{ij}\left\| U_{ij}^{(k)} - U_{ij}^{(k+1)} \right\| < \xi \tag{14}$$

In equation (14), ξ is a termination criterion between 0 and 1, whereas k are the iteration steps. The clusters centroid values is computed by using the equation (15)

$$z_j = \frac{\sum_{i=1}^{I} u_{ij}^m . x_i}{\sum_{i=1}^{I} u_{ij}^m} \tag{15}$$

To enhance the performance of the fuzzy-C-means clustering method, adaptiveness is invoked by measuring the clustering effectiveness (α) and absolute density (β). On the basis of these two, we set two thresholds to ensure the clustering being good.

3.2.1. Clustering Effectiveness

The number of points for each cluster is different based on the C-means characteristics. The Clustering effectiveness is based on the number of shots in a cluster. If many clusters contain very few shots at a clustering iteration, the clustering result is not good. The threshold for effective clustering is defined as:

$$\alpha = \frac{|\varpi|}{|\rho|} \tag{16}$$

In equ (16), α is the threshold for effective clustering, ϖ denotes the set of all shots and ρ denotes the set of all clusters.

3.2.2. Absolute Density

Absolute density is the density of the absolute clusters. In contrast to effective clustering, the large average distance among the absolute clusters represents the good dispersion and distinction. According to this notion, the absolute density is defined as:

$$\beta = \bar{\delta} - 0.7 * \sigma_\delta \tag{17}$$

In equ (17), β is the threshold for absolute density, 0.7 is the constant value, δ is the set of distances between two clusters, $\bar{\delta}$ is the average of δ and σ_δ is the standard deviation of δ. Hence, if 30% of the distances in δ cannot exceed β, the quality of clustering is bad. On the basis of the above, we set two thresholds to ensure the clustering being good. If it satisfies the above two conditions the clustering is good else the clustering method should be repeated.

3.4. Video Retrieval

Here, the most similar video clips are retrieved based on their distance with the cluster centroid value z_i and the query feature ($t(w(s_m^j))$), ($c(s_m^j)$). Extract the video from the cluster which has minimum distance.

3.4.1. Distance Measurement

Distance Measurement is the important stage in video retrieval. A query frame is given to a system which retrieves similar videos from the video database.

$$f_q = (t(w(s_m^j)), c(s_m^j)) \tag{18}$$

The query video q feature is represented as f_q which is given in Equ. (18). Similarly, each s_n^i in the video database is represented with the feature vector $F_q = F_{q_1}, F_{q_2}, \cdots F_{q_n}$, $n = 1,2,\cdots |V_i|$ and their cluster centroid values z_j. Our main objective is to retrieve videos by measuring the distance between the query video features (($t(ws_m^j)$, $c(s_m^j)$) which is obtained from the feature extraction process and the cluster centroid value z_j for the videos in the database. In order to retrieve the video a Euclidean distance measure is computed. The process of Euclidean distance measure is described below,

$$\left(\frac{1}{E}\sum_{e=1}^{E}\left(z_{je} - f_{q_e}\right)\right) \tag{19}$$

where E denotes the length of the feature ($t(w(s_m^j))$), $c(s_m^j)$) and centroid values z_j. By exploiting the Equ. (19), the relevant images are extracted which have the minimum value of RS. By following the aforementioned process, the videos similar to query video are successfully retrieved from the video database.

4. Experimental Results and Discussion

The proposed CBVR using adaptive fuzzy C-means technique was implemented in the working platform of MATLAB (Version 7.12) with machine configuration as follows:
Processor: Intel Core i5
OS: Windows 7
CPU Speed: 3.20 GHz
RAM: 4GB

We use CBVR using adaptive fuzzy C-means algorithm for the purpose of multimedia video retrieval. The videos from database and the query videos are subjected to feature extraction process, and then clustered by using adaptive fuzzy C-means. The next stage is to find the distance between the query video feature and the cluster centroid values. The videos which have less distance with the clustered indexing number are retrieved. The results are analyzed to demonstrate the superior performance of the proposed AFCM content based video retrieval technique with the existing techniques.

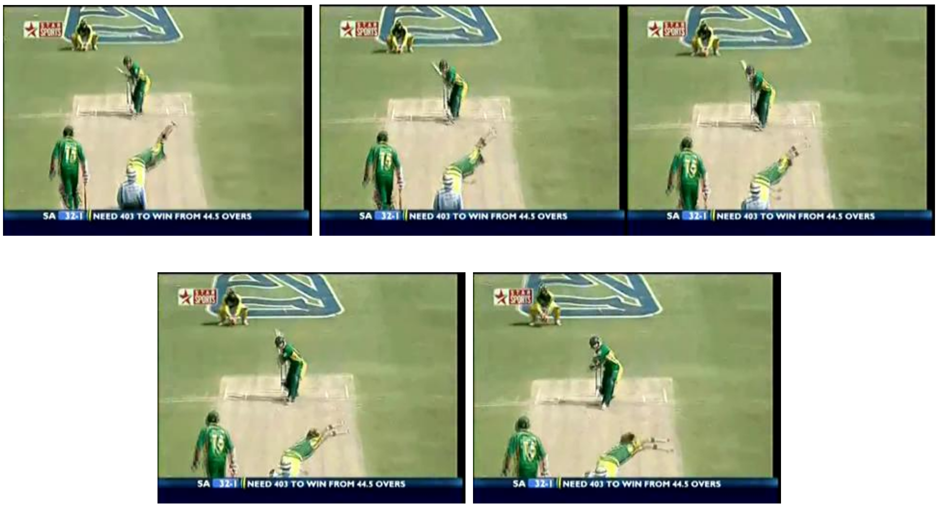


Figure 4: Sample Videos from the Database

(i) (ii) (iii)

Figure 5: Sample Query Video (i) Query Video 1 (ii) Query Video 2 and (iii) Query Video 3

In Figure 4 and Figure 5, the sample database and query videos are shown. Using Haar wavelet transform and color histogram the texture and color features are extracted from both the sample videos. The database sample video features are clustered by the AFCM clustering and then estimated the cluster centroid values. Afterward, the videos are retrieved which have the minimum distance with the query video feature values. The retrieved video results for the query videos are shown in Figure 6.

(i)

(ii)

(iii)

Figure 6: Retrieved Videos for (i) Query video 1 (ii) Query video 2 and (iii) Query video 3

Our proposed CBVR method performance is analyzed by invoking various performance measures as described in the following section.

4.1. Performance Analysis

The results are analyzed to demonstrate the performance of the proposed AFCM content based video retrieval technique compared to the existing techniques. The performance is evaluated by two quantitative performance metrics, namely,

- Average Retrieval Rate
- Average Precision Rate

4.1.1. Average Retrieval Rate (ARR)

Average retrieval rate can be evaluated by using the below equation (28)

$$ARR = \frac{1}{N_q} \sum_{q=1}^{N_q} RR(q) \leq 1 \tag{28}$$

where, N_q - number of queries that are used for verifying the retrieval performance

RR - retrieval rate of a single query image.

The RR can be calculated by the following Equation,

$$RR(q) = \frac{N_r(q)}{N_{dr}(q)} \leq 1 \tag{29}$$

- $N_{dr}(q)$ Denotes the number of relevant images in the database D of the query q
- $N_r(q)$ Denotes the number of retrieved relevant images of the query q .

The high ARR value indicates good performance of video retrieval whereas low value indicates a bad retrieval rate.

4.1.2. Average Precision Rate (APR)

$$AP = \frac{1}{r}\sum_{k=1}^{r} p_k \tag{30}$$

- Average Precision Rate can be evaluated by using the term given above
- AP represents the mean of precision values of all relevant images

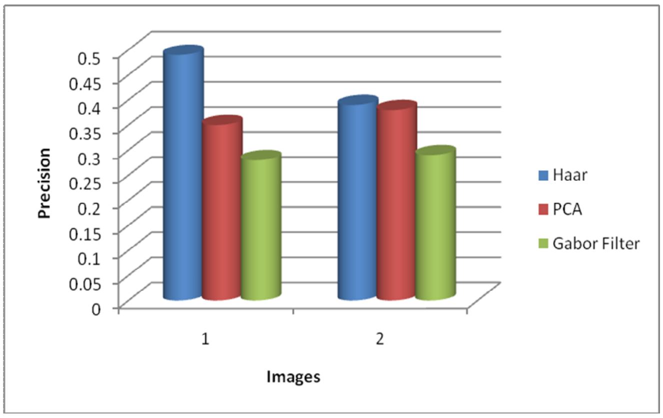
4.2. Precision and Recall Values

The precision and recall values are calculated by using the following equations,

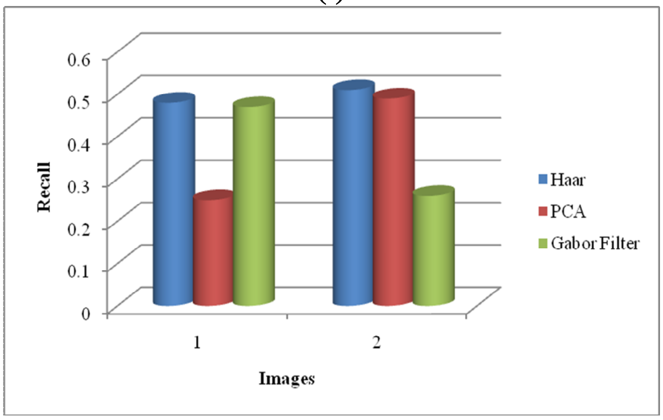
$$p = \frac{N_{rr}}{N_r} \tag{31}$$

$$R = \frac{N_{rr}}{N_d} \tag{32}$$

- N_{rr} denotes the number of relevant videos retrieved,
- N_r represents the total of retrieved videos
- N_d is the total number of videos in the database



(i)



(ii)

Figure 7: Different Feature Extraction Methods - (i) Precision (ii) Recall Performance

Discussion: From the Figure 7, the different feature extraction methods retrieval results are analyzed in terms of their precision and recall rates. The graph shows that the Haar technique has given high precision and recall rates than the other feature extraction methods.

The different feature extraction methods with proposed AFCM, FCM and C-means retrieval performance is evaluated by two quantitative performance metrics namely ARR and APR are given in Table 1.

Table 1: Quantitative Performance metrics of Proposed AFCM and Exiting Clustering Techniques

Methods	ARR	APR
K-Means-Haar	0.45	0.32
K-Means-PCA	0.35	0.26
K-Means-Gabor Filter	0.28	0.37
FCM-Haar	0.48	0.50
FCM-PCA	0.25	0.49
FCM -Gabor Filter	0.47	0.16
AFCM -Haar	0.79	0.69
AFCM -PCA	0.68	0.58
AFCM - Gabor Filter	0.78	0.66

From the Table 1, the high value of ARR and APR shows that our proposed Haar-AFCM based retrieval techniques has given the high retrieval rates than the other C-means and FCM based retrieval techniques. Our proposed method has given 79% and 69% value of ARR and APR values. The results prove that our proposed AFCM content based video retrieval technique more precisely and efficiently retrieves the videos by achieving higher retrieval rate compared to the existing clustering techniques.

5. Conclusion

In this paper, we addressed the problem of video retrieval in multimedia. We proposed a new CBVR technique using adaptive fuzzy C-Means (AFCM) clustering method in order to overcome the shortcomings in the existing methods. We compared the performance of our technique with the existing clustering techniques. The comparison shows that our AFCM content based video retrieval technique retrieved the videos more accurately than the existing clustering methods. Our technique retrieves the videos more precisely and efficiently by achieving higher retrieval rate than the other existing clustering techniques.

References

[1] Fabio Faria, Jefersson Dos Santos, Anderson Rocha and Ricardo Da Torres, “A Framework for Selection and Fusion of Pattern Classifiers in Multimedia Recognition”, Pattern Recognition Letters, Vol. 20, No. 7, pp. 1-13, 2013.

[2] Abhinav Srivastava, Shamik Sural and Majumdar, “Database Intrusion Detection Using Weighted Sequence Mining”, Journal of Computers, Vol. 1, No. 4, pp. 8-17, 2006.
[3] Guang-Tong Zhou, Kai Ming Ting, Fei Tony Liu and Yilong Yin, “Relevance Feature Mapping for Content-Based Multimedia Information Retrieval”, Pattern Recognition, Vol. 45, No. 4, pp. 1707–1720, 2012.
[4] Manda Jaya Sindhu, Madhavi Latha, Samson Deva Kumar and Suresh Angadi, “Multimedia Retrieval Using Web Mining”, International Journal of Recent Technology and Engineering, Vol. 2, No. 1, pp. 106-108, 2013.
[5] Farham Mohamed, Nordin Rahman, Yuzarimi Lazim and Saiful Bahri Mohamed, “Managing Multimedia Data: A Temporal-Based Approach”, International Journal of Multimedia and Ubiquitous Engineering, Vol. 7, No. 4, pp. 73-86, 2012.
[6] Manjunath, Ravindra, Hegadi and Ravikumar, “A Survey on Multimedia Data Mining and its Relevance Today”, International Journal of Computer Science and Network Security, Vol. 10, No. 11, pp. 165-170, 2010.
[7] Varun Kumar, "Intelligent Data Mining: Data Mining Powered By Artificial Intelligence", International Journal of Computer Science and Information Technology, Vol. 3, No. 1, pp. 44-47, 2006.
[8] Chih-Ming Chen and Hui-Ping Wang, "Using Emotion Recognition Technology to Assess the Effects of Different Multimedia materials on Learning Emotion and Performance", Information Science Research, Vol. 33, No. 44, pp. 244 –255, 2011.
[9] Mohamed Hussain Tawaris and Farooqui Wasee muddin, “Analysis of Data Mining Concepts in Higher Education with Needs to Najran University”, International Journal of English and Education, Vol. 2, No. 2, pp. 590-597, 2013.
[10] Vishwadeepak Singh, Baghela and Tripathi, “A Survey on Association Rules in Case of Multimedia Data Mining”, International Journal of Computer Science and Technology, Vol. 3, No. 1, pp. 649-652, 2012.
[11] Mehta Neel, "Predictive Data Mining and Discovering Hidden Values of Data Warehouse ", ARPN Journal of Systems and Software, Vol. 1, No. 1, pp. 1-5, 2011.
[12] Nordin Rahman, Yuzarimi Lazim and Farham Mohamed, “Applying Rough Set Theory In Multimedia Data Classification”, International Journal on New Computer Architectures and Their Applications, Vol. 1, No. 3, pp. 683-693, 2011.
[13] Markus Sched, Gerhard Widmer, Peter Knees and Tim Pohle, “A Music Information System Automatically Generated Via Web Content mining Techniques”, Information Processing and Management, Vol. 47, No. 3, pp. 426–439, 2011.
[14] Sangeeta Goele and Nisha Chanana, "Data Mining Trend in Past, Current and Future", International Journal of Computing & Business Research, Vol. 2, No. 2, pp. 1-8, 2013.
[15] Pravin Kamde, Siddu and Algur, “A Survey On Web Multimedia Mining”, The International Journal of Multimedia & Its Applications, Vol. 3, No. 3, pp. 72-84, 2011.
[16] Ja-Hwung Su, Yu-Ting Huang, Hsin-Ho Yeh and Vincent Tseng, “Effective Content-Based Video Retrieval Using Pattern Indexing and Matching Techniques”, Expert Systems with Applications: An International Journal, Vol. 37, No. 7, pp. 1-33, 2010.
[17] Jeroen Famaey, FredeIterbeke, Tim Wauters and Filip De Turck, “Towards A Predictive Cache Replacement Strategy for Multimedia Content”, Journal of Network and Computer Applications, Vol. 36, No. 3, pp. 219–227, 2013.

[18] Thanga Ramya and Rangarajan, “Knowledge Based Methods For Video Data Retrieval”, International Journal of Computer Science & Information Technology, Vol. 3, No. 5, pp. 165-172, 2011.
[19] Ankush Mittal, "An Overview of Multimedia Content-Based Retrieval Strategies", Journal of Informatica, Vol. 30, No. 3, pp. 347–356, 2006.
[20] Rosaci and Sarne, “Recommending Multimedia Web Services in a Multi-Device Environment”, Information Systems, Vol. 38, No. 2, pp. 198–212, 2007.
[21] Ja-Hwung Su, Yu-Ting Huang, Hsin-Ho Yeh and Vincent Tseng, “Effective Content-Based Video Retrieval Using Pattern-Indexing and Matching Techniques”, Expert Systems With Applications, Vol. 37, No. 7, pp. 5068–5085, 2010.
[22] Morand, Benois-Pineau, Domenger, Zepeda, Kijak and Guillemot, “Scalable object-based video retrieval in HD video databases”, Signal Processing: Image Communication, Vol. 25, No. 1, pp. 450–465, 2010.
[23] Barbara, Tom Vercauteren, Anna Buchner, Michael Wallace and Nicholas, “A smart atlas for end microscopy using automated video retrieval”, Medical Image Analysis, Vol. 15, No. 1, pp. 460–476, 2011.
[24] Huanbo Luan, Yan-Tao Zheng, Meng Wang and Tat-Seng Chua, “VisionGo: Towards video retrieval with joint exploration of human and computer”, Information Sciences, Vol. 181, No. 19, pp. 4197–4213, 2011.
[25] Jaesik Choi, Ziyu Wang, Sang-Chul Lee and Won Jeon, “A spatio-temporal pyramid matching for video retrieval”, Computer Vision and Image Understanding, Vol. 117, No. 6, pp. 660–669, 2013.

Authors biography

Manjunath R. received his bachelor's degree in Computer Science & Engineering from University of Kuvempu, Shimoga, Karnataka, a Master's degree in Computer Science & Engineering from Visvesvaraya Technological University, Belguam, Karnataka and a Master of Business Administration in HR from Indra Gandhi National Open University, New Delhi, India. Currently, he is an Associate Professor in the Department of Computer Science & Engineering, City Engineering College, Visvesvaraya Technological Technology. His areas of interest include data mining, data science and analytics, multimedia, cloud and grid computing, image processing, software engineering and business intelligence.

S.Balaji received a Bachelor's degree in Mathematics from Madras University, a Master's degree in Applied Mathematics from Anna University, Chennai and a Ph.D. degree in Computer Science and Engineering from Indian Institute of Science, Bangalore. He has served Indian Space Research Organization for more than two decades before getting into academics. He is currently a Professor at Centre for Emerging Technologies, Jain Global Campus, Jain University. His research interests include fault-tolerant computing, real-time systems, mobile computing, image processing and computer vision, data science and analytics, energy efficient computing, wireless sensor networks, ground penetrating RADAR applications, biomolecular modelling and simulation, bioinformatics, and MEMS based systems development for mission and safety critical applications.