International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 8 Issue 12, December -2019, Impact Factor: 6.754

Content Based Video Retrieval Using Back propagation Neural Network and Radial Basis Function Network

Dr. Shahbaa I. Khaleel[1], Dr. Nidhal Hussein[2]

[1,2]Software Engineering Dept., College of Computer Science and Mathematics, Mosul University, Iraq

ABSTRACT

In this research we represent a new system content based video retrieval CBVR that retrieve a video shot which is matched or similar to the user query. After selecting the key frame from video shot, we must convert the key frame to image. For the color content extraction, a well-known and powerful techniques, Color Histogram, Color Moment and Color Correlogram are used in one method by combining all these features to produce efficient method called Combined Method. At search time, the query image is not compared with all the images in the database, but only with a small subset. Here we used M2k-means clustering algorithm, this algorithm is an efficient algorithm and produce efficient retrieval. Beside this, a nonlinear combination of features of the key frame is used and is based on Back propagation Neural Network BPNN and Radial Basis Function Network RBFN. It is very efficient and produces 100% retrieval of key frame. We conclude that color methods are useful in content based video retrieval. Combining color information usually improves the performance of the method; and it further improves the performance if the M2k-means clustering algorithm and nonlinear combiner (BPNN & RBFN) are used. These methods proved an efficient performance when used in content based video retrieval.

Keyword: video retrieval, key frame, color histogram, color moments, color correlogram, clustering algorithm, back propagation, RBF neural network.

HOW TO CITE THIS ARTICLE

Dr. Shahbaa I. Khaleel, Dr. Nidhal Hussein, "Content Based Video Retrieval Using Back propagation Neural Network and Radial Basis Function Network", International Journal of Enhanced Research in Science, Technology & Engineering, ISSN: 2319-7463, Vol. 8 Issue 12, December -2019.

1-INTRODUCTION

For content based video retrieval, there are currently two widely accepted query modes: Keyword-based and example-based. The example-based approach is necessary in the situation that users could not describe clearly what they exactly want by only using text. The example queries could be a video clip, a frame, an object or some low-leveled features such as color, texture and motion [1]. The standard representation for color has been the color histogram. Other color features include color moments and color correlogram [2,3]. The clustering algorithm are also applied[4] by using modified2 k-means (M2k-means) algorithm [5]. We used to determine nonlinear relationship between different features used the neural network that represent by back propagation neural network and radial basis function network model [2]. These three features extraction are applied on color image in this research, and combined all these features in one method called Combined Method; Moreover, the clustering algorithm are applied to all these methods the following section described all them.

2-COLOR FEATURES EXTRACTION

Color has been the most commonly-used feature type in Content Based Image Retrieval CBIR. Basic color features are easy to implement and usually yield reasonable and predictable results which can then be improved by including other types of features. The standard representation for color information has been the color histogram. The color histogram describes the distribution of different colors in an image in a simple and computationally efficient manner. Other commonly used color features include color moments and the color correlogram [2,3].

Color Histogram

One of the widely used methods for querying and retrieval by color content is color histograms. The histogram approach is commonly used in most of the existing systems supporting query-by-color content. The retrieval of similar images is based on the similarity between their respective color histogram. A common similarity metric is based on the Euclidean distance between the abstracted feature vectors that represent two images, and it is defined as: $d(Q,I) = \sqrt{\sum_{j=1}^{n} \left(h_j^Q - h_j^I\right)^2}$ Where Q and I represent the query image and one of the images in the image set, and h_j^Q and h_j^I represent the values of the feature vectors of these images respectively, i.e. color histogram feature. While the distance between intensity histogram features is calculated as follows: $d(Q,I) = \sqrt{\sum_R \sum_G \sum_B \left(h_{rgb}^Q - h_{rgb}^I\right)^2}$ Where h_{rgb}^Q , h_{rgb}^I represent the histogram of query image and one of the images in the image set. Note that a smaller distance reflects a closer similarity match [2,3].

Color Moment

Color moments have been successfully used in many retrieval systems, especially when the image contains just the object. The first order mean, the second variance or standard deviation and the third order skewness color moments have been proved to be efficient and effective in representing color distributions of images. Mathematically, the first three moments are defined as:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij} \qquad \sigma_i = \left[\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^2\right]^{1/2} \qquad S_i = \left[\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^3\right]^{1/3}$$

Where f_{ij} is the value of the i^{th} color component of the image pixel j, N is the number of pixels in the image, and μ , σ , s represent mean, standard deviation , and skewness respectively [2,3].

Color Correlogram

The Correlogram represents the distribution of the Co-occurrence matrix, but it is calculated using several different distances (d) instead of using a single distance; and it uses eight directions (angles). Let [D] denote a set of D fixed distances { $d_1, \ldots, d_D$ }, and f be an image consisting of M different colors $c_1, c_2, \ldots, c_M$. For a pixel $p(x,y) \in f$, let $f(p)$ expresses the pixel color and $f_c = \{p \mid f_{(p)} = c\}$. Therefore, $p \in f_c$ is equivalent to $p \in f, f_{(p)} = c$. For simplicity, $L_\infty - Norm$ is used to measure the distance between pixels. This measure is computed for two pixels $p_1(x_1,y_1)$, $p_2(x_2,y_2)$ as follows: $|p_1 - p_2| = \max \{|x_1 - x_2|, |y_1 - y_2|\}$

Assume that k is a specified distance and $i, j \in \{1, \ldots, M\}$. The correlogram of f is defined by: $\gamma(i,j,k) = \underset{p_1 p_2 \in f}{pr} \{p_2 \in f_{cj} \mid p_1 \in f_{ci}, |p_1 - p_2| = k\}$ where $\gamma(i, j, k)$ denotes the probability of finding pixels with color c_j at the distance k of the pixel with color c_i [2,3]. In this research we calculate five features from the correlogram that is :

$$Energy = \sum_{ci}\sum_{cj}\sum_{k}\left(\gamma_{cicj}^{k}\right)^2 \qquad Entropy = \sum_{ci}\sum_{cj}\sum_{k}\gamma_{cicj}^{k}\log\left(\gamma_{cicj}^{k}\right) \qquad Contrast = \sum_{ci}\sum_{cj}\sum_{k}(ci - cj)^2\gamma_{cicj}^{k}$$

$$Inverse-Difference-Moment = \sum_{ci}\sum_{cj}\sum_{k}\frac{\gamma_{cicj}^{k}}{|ci - cj|^2}, ci \neq cj \qquad Correlation = \frac{\sum_{ci}\sum_{cj}(cicj)\gamma_{cicj} - \mu_x\mu_y}{\sigma_x\sigma_y}$$

where means and standard deviations are defined as:

$$\mu_x = \sum_{ci} ci \sum_{cj}\gamma_{cicj} \qquad \mu_y = \sum_{cj} cj \sum_{ci}\gamma_{cicj} \qquad \sigma_x = \sum_{ci}(ci - \mu_x)^2\sum_{cj}\gamma_{cicj} \qquad \sigma_y = \sum_{cj}(ci - \mu_y)^2\sum_{ci}\gamma_{cicj}$$

2.4-THE COMBINED METHOD

To obtain effective retrieval, we combined the three features extraction :color histogram, color moment, and color correlogram. In order to obtain best retrieval, figure (1) explains the outline of the combined method. Since different features can generate different ranges of values of similarity, a normalization method should be applied to each similarity computation. We normalize each similarity by min/max normalization (linear scaling) method according to equation as follows:

$$N(I,I') = \frac{D(I,I') - \min(D(I,I'))}{\max(D(I,I')) - \min(D(I,I'))}$$

After normalizing similarity, the total similarity between the query and the image in the data collection is calculated via a weighted sum of the similarities provided by each of the features. The equation for combining the similarities of three image features is defined as follows:

$$D_{combine}(I, I') = W_1 N_{color}(I, I') + W_2 N_{moment}(I, I') + W_3 N_{corre\log ram}(I, I')$$

where $D_{combine}(I, I')$ is the weighted sum of similarities; $N_{color}(I, I')$ is the normalized similarity of color histogram; $N_{moment}(I, I')$ represents the color moment features; and $N_{corre\log ram}(I, I')$ represents the color correlogram features. W_1, W_2, and W_3 are weighting factors to adjust the relative importance of image features. We choose W_1 =1.0, W_2 =0.5, and W_3 =0.02 for our experiments in this research. The application of the combined algorithm in this work is very efficient compared with the previous techniques with precision and recall metrics and also with average precision.

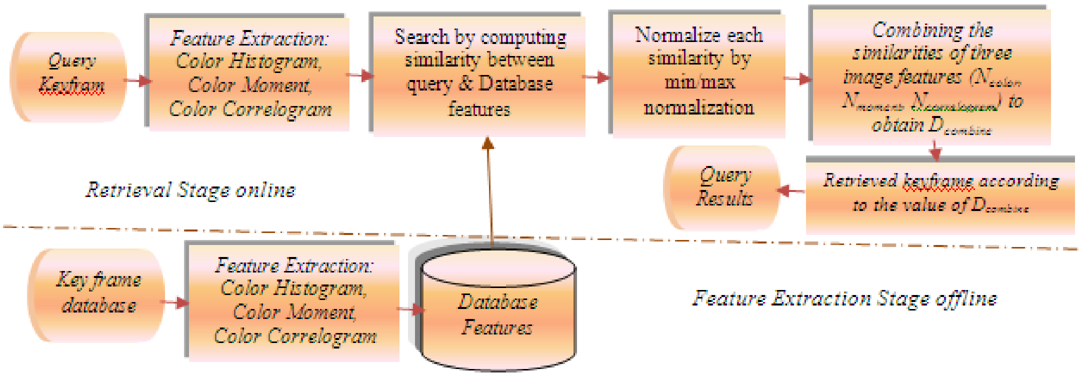

Figure 1: The Outline of Combined Method

3-CLUSTERING

In the k-means, the initial cluster assignment is random; different runs of the k-means clustering algorithm may not give the same final clustering solution; or when selected as the first k samples of the sample set the same as modified_1 k-means, these two states may not give the good solution. To deal with this, we need to get good starting points for the initial cluster assignment. This leads to develop a modified_1 k-means clustering algorithm where an additional step is used to provide the initial cluster centers and L_1 norm distance metric when computing the distance between two feature vectors. The initial centroids are selected in the following way: ***1)*** Given v d -dimensional feature vectors, divide the *d* dimensions to $p = \frac{d}{K}$. these subspaces are indexed by: $[1,2,3,\ldots, p],[p+1, p+2,\ldots, 2p],\ldots, [(k-1)p+1,(k-1)p+2,(k-1)p+3,\ldots, kp]$. ***2)*** In each subspace *j* f $[(j-1)p+1,\ldots, jp]$, associate a value $f_i^{\,j}$ for each feature vector f_i by : $f_i^{\,j} = \sum_{d=(j-1)p}^{jp} f_i(d)$. ***3)*** Choose the initial cluster centroids $\mu_1, \mu_2,\ldots, \mu_K$ by $\mu_j = \arg_{f_i} \min_{1<i<v} f_i^{\,j}$ [2]. In this research, we based on M_2K-means algorithm to cluster all the images in the database into classes.

4-THE NEURAL NETWORK BASED VIDEO RETRIEVAL

Content of an image can be expressed in terms of different features such as color, color moment or correlogram. Retrieval based on these features can be various by the way how to combine the feature values. All the preceeding approaches assume a linear relationship between different features. In this research, we work human-computer interaction system model of content based video retrieval with nonlinear combination of features of the keyframe. The approach Neural Network based video retrieval is based on Back propagation Neural Network BPNN, and Radial Basis Function Network Model RBFN. This system can be used to determine nonlinear relationship between different features in keyframe. The input to the neural network is the set of metric values of each pair of keyframe and the output is a number between 0 and 1 signifying similarity of keyframe based on various input features. This system offers a new method of combining image features. Using combined rather than individual features is especially efficient for generic keyframe databases, for which no single feature is outstanding. An experimental evaluation will demonstrate that this model can achieve both efficiency and flexibility on video retrieval combined features of keyframes.

5- VIDEO STRUCTURE

Video can be described by a hierarchy of scene and shots, where each entry in the hierarchy is composed of one or more entries at a lower level (e.g., a scene is composed of a sequence of shots and a shot is composed of a sequence of frames) [6]. Figure (2) shows the hierarchical display of video [7]. A shot is a sequence of consecutive frames taken from the same camera. Segmenting video into its shot structure is the fundamental procedure in analyzing video. Many video indexing or video browsing applications are implemented using only this structure. A collection of consecutive shots, which are related to each other by the same semantic content, are grouped together into one scene [8].

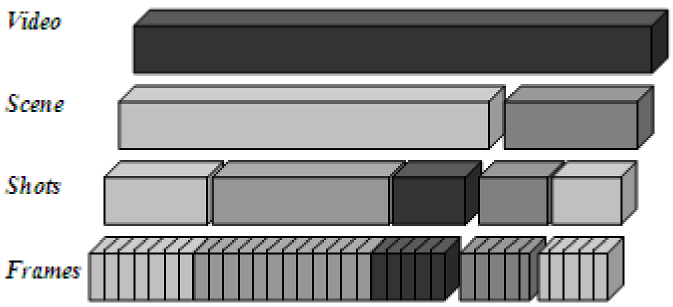

Figure 2: Video Structure

6- FEATURES FOR VIDEO RETRIEVAL

We define features of visual data as a set of attributes that can be extracted using image processing and computer vision techniques. This set includes, but is not limited to, shot boundaries, color characteristics of image frames (for example, color histogram, color moment, correlogram) as video features [9]. Hereafter, we discuss these features and present methods to compute them.

Shot Detection

For content based retrieval, retrievals based on the image features of video frames are more efficient and practical. By measuring the similarities among video frames, a hierarchical cinematic structure, including the shots, scenes of a video are retrieved. A shot is a sequence of frames which represents continuous action in time and space. The contents of the frames belonging to a shot are similar. Therefore, a shot change detection can be performed through similarity measurement of continuous frames [10,11]. Some of the existing methods for computing the content difference for consecutive video frames are the color histogram approaches which summarize the color distribution of a frame and computes the differences between it and the color distributions of its adjacent frames. When the difference exceeds a predefined threshold, a shot change is detected. In the pair wise pixel comparison approach, the values of the pixels are compared pixel by pixel. The sum of the differences of the values is computed. A shot change is detected when the sum exceeds a predefined threshold [9].

Histogram comparison: A gray level histogram of a frame i is defined to be a vector of length n $H_i(j)$, $j = 1,..., n$, where n is the number of gray levels and $H(j)$ is the number of pixels from the frame i with gray level j. Color histograms are defined in the same manner, or they could be defined as k-dimensional arrays where k is the dimension of the color space used, usually 3. This representation is equivalent to a vector of length $3n$ and the metrics for comparing histograms which are mentioned below could be applied in the same manner as gray level values. Another possible representation is to use k separate vectors, one vector for each color channel. The metric for comparing these histograms is usually the sum of metrics defined for each channel.

There are multiple ways to compute the difference between two histograms; the most efficient and simplest one is to compute the differences depending on the Euclidean distance:

$$Diff(i, i+1) = \sqrt{\sum_{j=1}^{n} [H_i(j) - H_{i+1}(j)]^2}$$

where $H_i(j)$ is the histogram value for the pixel value j in frame i. If that difference value $Diff(i, i+1)$ is greater than a threshold, a cut is declared [6,8].

Key Frame Detection

In video indexing and retrieval, representing every segmented shot of the processed video by one appropriate frame (called key-frame) or by a small set of key-frames, is a common useful early processing step. When considering fast video content visualization, the selection of one frame per shot, typically the median frame, could be sufficient to a video visual content redundancies. On the other hand, key-frames can also be used in the content-based video indexing stage to extract spatial descriptors to be attached to the shot and related to intensity, color, texture or shape, which enables to process a very small set of images while analyzing the whole shot content [12]. The method to select variable number of key frames depending upon the shot activity, each shot, S_i, is represented by a set of key frames, k_i, such that all key frames are distinct. Initially, the middle frame of the shot is selected and added to the set k_i (which is initially empty) as the first key frame. The reason for taking the middle frame instead of the first frame is to make sure that the frame is free from shot transition effects. Next, each frame within a shot is compared to every frame in the set k_i. If the frame differs from all previously chosen key frames by a fixed threshold, it is added in the key frame set; otherwise it is ignored. This algorithm of key frame detection can be summarized as:

Step1: Select middle frame as the first key frame

$$k_i \leftarrow \left\{ f^{\left\lfloor (a+b)/2 \right\rfloor} \right\}$$

where a, b represent the start and the end of the shot

Step2:

$$\begin{aligned} &for \quad j = a \quad to \quad b \\ &if \quad \max\left(S\left(f^{j}, f^{k}\right)\right) < Th \quad \forall \; f^{k} \in k_i \\ &Then \quad k_i \leftarrow k_i \cup \left\{f^{j}\right\} \end{aligned}$$

where Th is the minimum frame similarity threshold that declares two frames to be similar. Using this approach, multiple frames are selected for the shots which have higher dynamics and temporally changing visual contents. For less dynamic shots, fewer frames are selected. This method assures that every key frame is distinct and, therefore, prevents redundancy [9].

Key Frame Feature Extraction

Indexing multimedia information using the color feature is one of the most widely used of the visual features. The color feature captures the color content of images. Before that, we must convert the key frame to image, and the following features are extracted from these images frames, color histogram, color moment, and color correlogram [13]. Then, the pre-processing stage of multimedia retrieval (video retrieval) consists of the following steps: shot detection, selection of key frames, frame conversion (i.e. convert the frame to matrix values), and key frame (image) feature extraction that described above. All these steps are illustrated in figure (3).

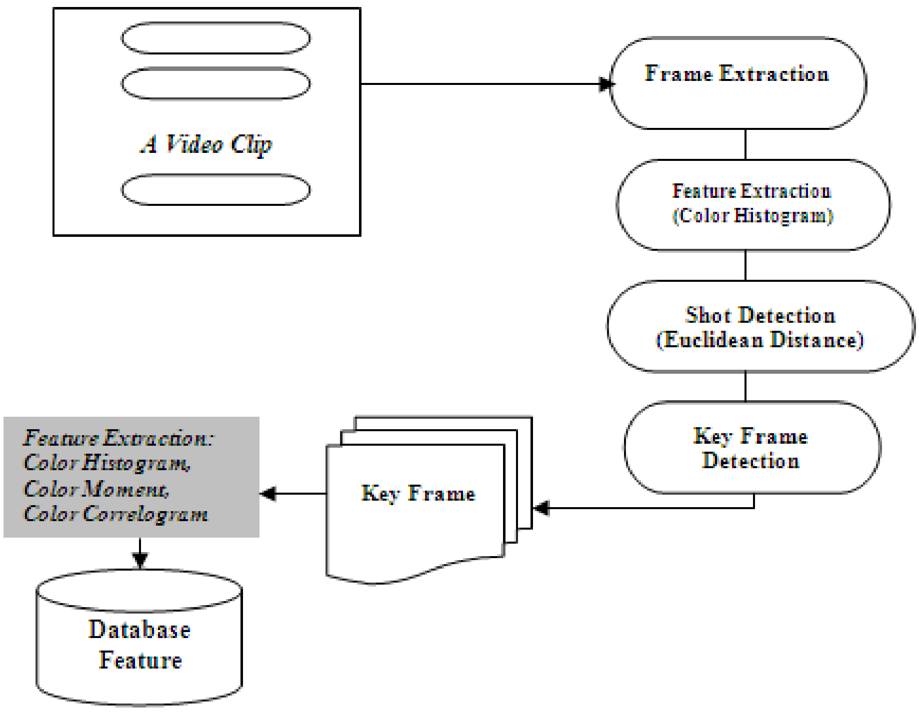

Figure 3: Pre-Processing Stage Steps of Video Indexing

7- SEARCH MODES

Searching can be done in two modes. The first mode is the "single image" search. In this mode, the sample image is a single image. The second is the "multiple image" mode where more than one image can be specified. The search process must start in the "single image" mode. First the same visual features from the sample image that were used in building the database are extracted. Then, these values are compared to the values of the features of the key frames in the database. This comparison generates a list of key frames, which match the sample image. This list display the key frames in order of relevance or similarity to the sample image. The key frame that is closest to the sample image will be displayed first. In the "multiple image" mode, multiple sample images are used, then the search process will retrieve more key frames from the database that are similar to those defined as samples[14].

8- Content-Based Video Retrieval System CBVR

In this research, we present a system CBVR that retrieves video shots which are matched or similar to the user query. In this approach, some representative key frames are extracted as an index. Users can query the videos using an example image, and the system compares the query image and the key frames to find possible results. Parsing the video into shots (i.e., shot change detection) is the first step in constructing such indexes for querying video data. Figure (4) shows all the steps of CBVR system.

1. Divide Video Content into Shots

In general, a video consists of a set of shots, and every shot consists of a set of frames. Here, in this research we use the Histogram Comparison to detect every shot in the video.

2. Extract Key Frames

In video indexing and retrieval, every shot of the video is represented by one appropriate frame, i.e. key frames that have spatial features to be attached to the shot. These key frames are obtained by using the key frame detection algorithm.

3. Feature Extraction of Key Frames

All key frames belonging to video database and query image example are extracted by three types of features: color histogram, color moment, and color correlogram, and then these feature vectors are stored in database feature and query feature vector respectively, These, then, are followed in the following steps.

4. Feature Vectors Clustering

We use modified2 k-means algorithm to cluster all feature vectors of database feature to reduce feature vectors and computational requirements, and obtaine centroid feature vectors. This operation gives an efficient and effective retrieval.

5. Sort the Cluster Centers

To speed up the search operation we must be sort all the distance between the query and cluster centers, and take the cluster center that has minimum distance with query, which represent the class that the query belong to it.

6. Search by Matching Features and Normalization

This step compares each query feature vector with the corresponding features of close centroid and returns the compared metric values. And these metric values are normalized.

7. Linear or Non-linear Combiner

In linear combiner, after normalizing similarity, the total similarity between the query and key frames is calculated via a weighted sum all the similarities provided by each of the features.

The non-linear combiner use BPNN or RBFN to combine the compared metric values; these metric values are used as input of BPNN or RBFN, and the output value of the network is the similarity between the query image and key frame of video.

International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 8 Issue 12, December -2019, Impact Factor: 6.754

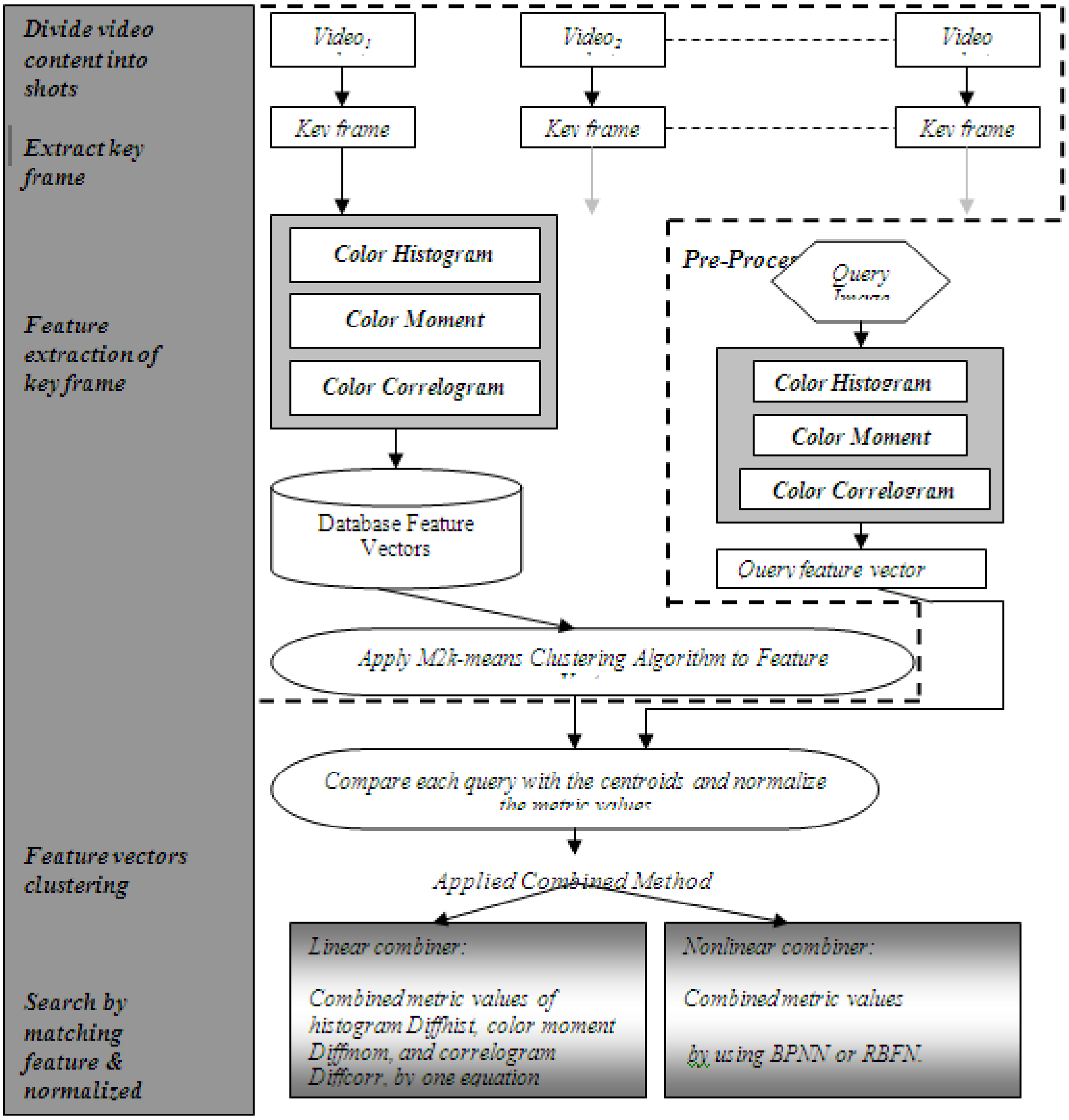

Figure 4: Steps of CBVR System

9- EXPERIMENTAL AND RESULTS

The system should search in the video database and cluster the relevant video segments such as video shots according to the content of the query example. Query by key frame or video example is a convenient and often effective way to search in video database. This research represent a new approach to support such searches. The main contribution of the approach is the consideration of both feature extraction and distance computation as a whole process. With a video shot represented by key-frames corresponding to feature points in a feature space, a new metric is defined to measure the distance between a query and a shot. In this approach, some representation key frames are extracted as an index. Users can query the videos using an example image, and the system compares the query image and the key frames to find possible results. Parsing the video into shots (i.e., shot change detection) is the first step in constructing such indexes for querying video data. Similarity measurement is based on comparison of the color histogram, that is used to detect shot.

The Results of Video Shot Retrieval

Here, in this research we represent a new system CBVR that retrieve a video shot which is matched or similar to the user query. After selecting the key frame from video shot, we must convert the key frame to image, and the following features are extracted from these images frames, Color Histogram, Color Moment and Color Correlogram. We use the linear combiner represented by combined method with M2k-means, and the nonlinear combiner by using BPNN & RBFN with M2k-means of online and batch training, and also the networks with bias and without bias node. We conclud that the nonlinear combiner is better than linear combiner as stated in table (1) and table(2) in both internal example and external with relevant & non-relevant. This can be clearly noticed in the following figurs (the videos was given from internet and select keyframes of all shots that used here in research and applied the research methods on thim).

Table 1: Comparison between Linear Combiner and Nonlinear Combiner (Back propagation Neural Network & Radial Basis Function Network) with M2k-means for internal example

Clip Name	Relevant Retrieved	Relevant No.	Retrieved No.	Recall	Precision	Method	Keyframe Type
Clip13	14	16	14	0.875000	1.000000	Combined	Internal
	16	16	16	1.000000	1.000000	Bp-bias-batch	Internal
	16	16	16	1.000000	1.000000	Bp-bias-online	Internal
	16	16	16	1.000000	1.000000	Bp-batch	Internal
	16	16	16	1.000000	1.000000	Bp-online	Internal
	16	16	16	1.000000	1.000000	RBF-bias-batch	Internal
	16	16	16	1.000000	1.000000	RBF-bias-online	Internal
	16	16	16	1.000000	1.000000	RBF-batch	Internal
	16	16	16	1.000000	1.000000	RBF-online	Internal
Clip20	1	3	1	0.333333	1.000000	Combined	Internal
	3	3	3	1.000000	1.000000	Bp-bias-batch	Internal
	3	3	3	1.000000	1.000000	Bp-bias-online	Internal
	3	3	3	1.000000	1.000000	Bp-batch	Internal
	3	3	3	1.000000	1.000000	Bp-online	Internal
	3	3	3	1.000000	1.000000	RBF-bias-batch	Internal
	3	3	3	1.000000	1.000000	RBF-bias-online	Internal
	3	3	3	1.000000	1.000000	RBF-batch	Internal
	3	3	3	1.000000	1.000000	RBF-online	Internal
Clip24	7	9	7	0.777778	1.000000	Combined	Internal
	9	9	9	1.000000	1.000000	Bp-bias-batch	Internal
	9	9	9	1.000000	1.000000	Bp-bias-online	Internal
	9	9	9	1.000000	1.000000	Bp-batch	Internal
	9	9	9	1.000000	1.000000	Bp-online	Internal
	9	9	9	1.000000	1.000000	RBF-bias-batch	Internal
	9	9	9	1.000000	1.000000	RBF-bias-online	Internal
	9	9	9	1.000000	1.000000	RBF-batch	Internal
	9	9	9	1.000000	1.000000	RBF-online	Internal
Clip33	5	6	5	0.833333	1.000000	Combined	Internal
	6	6	6	1.000000	1.000000	Bp-bias-batch	Internal
	6	6	6	1.000000	1.000000	Bp-bias-online	Internal
	6	6	6	1.000000	1.000000	Bp-batch	Internal
	6	6	6	1.000000	1.000000	Bp-online	Internal
	6	6	6	1.000000	1.000000	RBF-bias-batch	Internal
	6	6	6	1.000000	1.000000	RBF-bias-online	Internal
	6	6	6	1.000000	1.000000	RBF-batch	Internal
	6	6	6	1.000000	1.000000	RBF-online	Internal
Clip41	5	7	5	0.714286	1.000000	Combined	Internal
	7	7	7	1.000000	1.000000	Bp-bias-batch	Internal
	7	7	7	1.000000	1.000000	Bp-bias-online	Internal
	7	7	7	1.000000	1.000000	Bp-batch	Internal
	7	7	7	1.000000	1.000000	Bp-online	Internal
	7	7	7	1.000000	1.000000	RBF-bias-batch	Internal
	7	7	7	1.000000	1.000000	RBF-bias-online	Internal
	7	7	7	1.000000	1.000000	RBF-batch	Internal
	7	7	7	1.000000	1.000000	RBF-online	Internal

BPNN: Back propagation neural network.
RBFN: Radial Basis Function Network
Bp-batch: Back propagation with batch training
Bp-online: Back propagation with online training
Bp-bias-online: Back propagation with bias node online training.

Table 2: Comparison between Linear Combiner and Nonlinear Combiner (Back propagation Neural Network & Radial Basis Function Network) with M2k-means for external example

Clip Name	Relevant Retrieved	Relevant No.	Retrieved No.	Recall	Precision	Method	Key frame Type
Clip46	12	16	13	0.750000	0.923077	Combined	External –relevant
	16	16	16	1.000000	1.000000	Bp-bias-batch	External –relevant
	16	16	16	1.000000	1.000000	Bp-bias-online	External –relevant
	16	16	16	1.000000	1.000000	Bp-batch	External –relevant
	16	16	16	1.000000	1.000000	Bp-online	External –relevant
	16	16	16	1.000000	1.000000	RBF-bias-batch	External –relevant
	16	16	16	1.000000	1.000000	RBF-bias-online	External –relevant
	16	16	16	1.000000	1.000000	RBF-batch	External –relevant
	16	16	16	1.000000	1.000000	RBF-online	External –relevant
Clip57	8	9	8	0.888889	1.000000	Combined	External –relevant
	9	9	9	1.000000	1.000000	Bp-bias-batch	External –relevant
	9	9	9	1.000000	1.000000	Bp-bias-online	External –relevant
	9	9	9	1.000000	1.000000	Bp-batch	External –relevant
	9	9	9	1.000000	1.000000	Bp-online	External –relevant
	9	9	9	1.000000	1.000000	RBF-bias-batch	External –relevant
	9	9	9	1.000000	1.000000	RBF-bias-online	External –relevant
	9	9	9	1.000000	1.000000	RBF-batch	External –relevant
	9	9	9	1.000000	1.000000	RBF-online	External –relevant
Clip59	3	6	5	0.500000	0.600000	Com	External –relevant
	6	6	6	1.000000	1.000000	Bp-bias-batch	External –relevant
	6	6	6	1.000000	1.000000	Bp-bias-online	External –relevant
	6	6	6	1.000000	1.000000	Bp-batch	External –relevant
	6	6	6	1.000000	1.000000	Bp-online	External –relevant
	6	6	6	1.000000	1.000000	RBF-bias-batch	External –relevant
	6	6	6	1.000000	1.000000	RBF-bias-online	External –relevant
	6	6	6	1.000000	1.000000	RBF-batch	External –relevant
	6	6	6	1.000000	1.000000	RBF-online	External –relevant
Clip64	4	7	4	0.571429	1.000000	Com	External –relevant
	7	7	7	1.000000	1.000000	Bp-bias-batch	External –relevant
	7	7	7	1.000000	1.000000	Bp-bias-online	External –relevant
	7	7	7	1.000000	1.000000	Bp-batch	External –relevant
	7	7	7	1.000000	1.000000	Bp-online	External –relevant
	7	7	7	1.000000	1.000000	RBF-bias-batch	External –relevant
	7	7	7	1.000000	1.000000	RBF-bias-online	External –relevant
	7	7	7	1.000000	1.000000	RBF-batch	External –relevant
	7	7	7	1.000000	1.000000	RBF-online	External –relevant

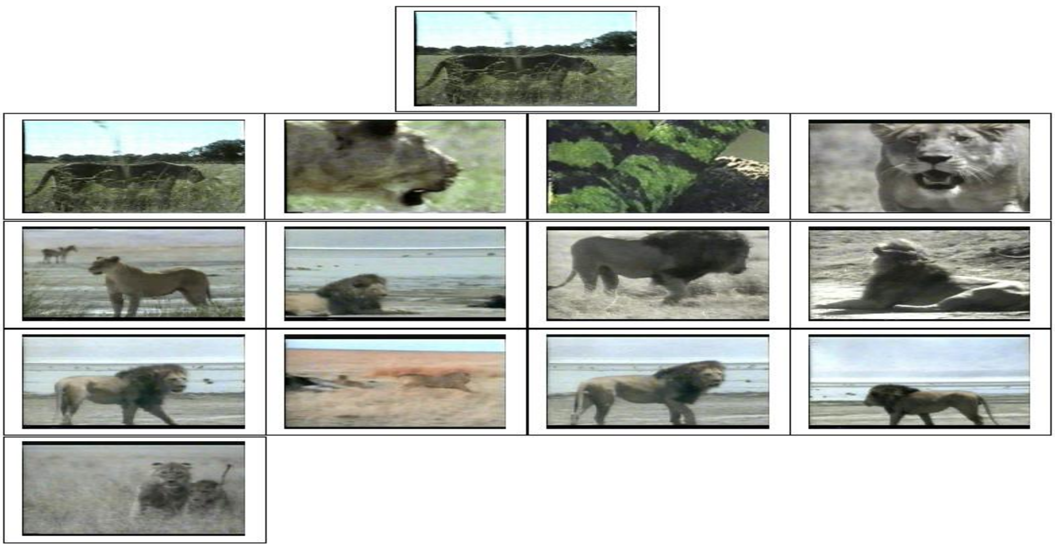

Figure 5: Lion keyframe, 12 matches from the Top 13 Using Combined Method with M2k-means

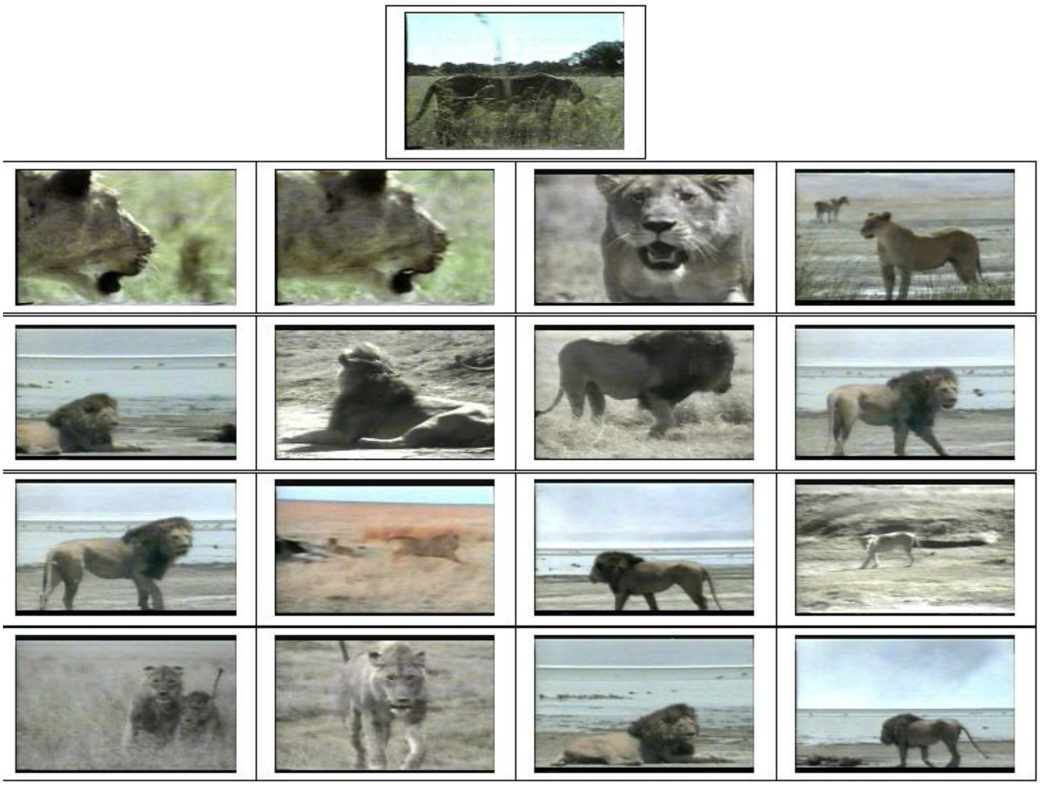

Figure 6: Lion keyframe, 16 matches from the Top 16 Using BPNN with Bias Node on Batch Training with M2K-means

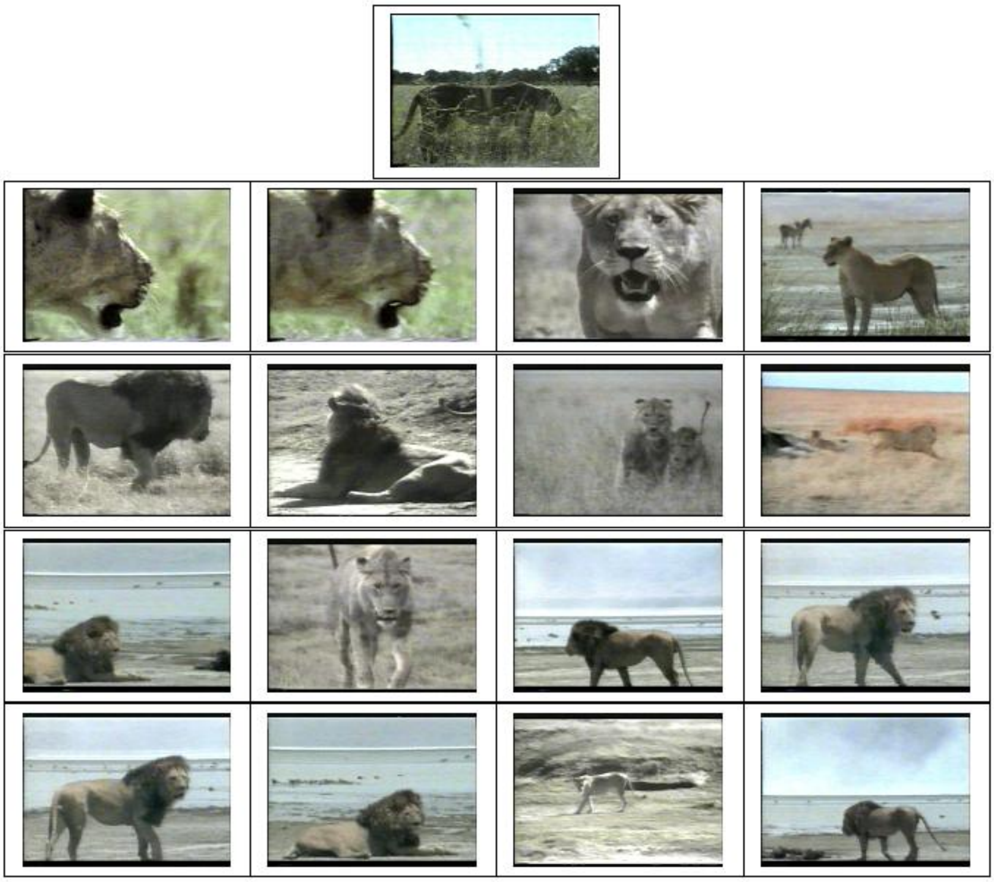

Figure 7: Lion keyframe, 16 matches from the Top 16 Using RBFN with Bias Node on Batch Training with M2K-means

CONCLUSIONS

We have designed and implemented a content based video retrieval system that evaluates the similarity of each key frame in its data store to a query key frame in terms of color characteristics, and returns the videos shot within a desired range of similarity. The main contribution of this approach is the consideration of both feature extraction and distance computation as a whole process. A video shot is represented by key frame, and a new metric is defined to measure the distance between a query image and a shot. For the color content extraction, a well-known and powerful techniques, Color Histogram, Color Moment and Color Correlogram are used in one method by combining all these features to produce efficient method called Combined Method. At search time, the query image is not compared with all the images in the database, but only with a small subset. Here we used M2k-means clustering algorithm, this algorithm is an efficient algorithm and produce efficient retrieval. Beside the above, a nonlinear combination of features of the key frame is used and is based on BPNN and RBFN. It is very efficient and produces 100% retrieval of key frame. We conclude that color methods are useful in content based video retrieval. Combining color information usually improves the performance of the method; and it further improves the performance if the M2k-means clustering algorithm and nonlinear combiner (BPNN & RBFN) are used. These methods proved an efficient performance when used in content based video retrieval.

REFERENCES

[1] Deng Y., Manjunath B., (1997), Content-Based Search of Video Using Color, Texture And Motion, IEEE International Conference on Image Processing, Vol 2, pp. 13-16.
[2] Dr. Shahbaa I. Khaleel, Dr. Nidhal Hussein, “Color Image Retrieval Based on Content by Using Linear and Nonlinear Combiner”, International Journal of Enhanced Research in Science, Technology & Engineering, ISSN: 2319-7463, Vol. 8 Issue 8, August-2019.
[3] Dr. Shahbaa I. Khaleel, Dr. Nidhal Hussein, “Colored Image Retrieval Based on Content by Using Integration of Features for Many Color Space Models”, International Journal of Enhanced Research in Science, Technology & Engineering, ISSN: 2319-7463, Vol. 8 Issue 8, August-2019.
[4] Dr. Shahbaa Ibrahim Khaleel, “Color Image Compression Based on Fuzzy Clustering”, International Journal of Enhanced Research in Science, Technology & Engineering, ISSN: 2319-7463, Vol. 8 Issue 7, July-2019.
[5] Dr. Shahbaa I. Khaleel, Dr. Nidhal Hussein, “Image Retrieval Based on Content by using Clustering Algorithms”, International Journal of Enhanced Research in Science, Technology & Engineering, ISSN: 2319-7463, Vol. 8 Issue 7, July-2019.
[6] Lew M. S.,(2001),Principles of Visual Information Retrieval, Springer-Verlag London Berlin Heidelberg, London.
[7] Huang J., (1998), Color-Spatial Image Indexing And Applications, PhD Thesis, Cornell University.
[8] Aner A., (2002), Video Summaries And Cross-Referencing, PhD Thesis, Columbia University.
[9] Rosenfeld A., Doermann D., Dementhon D., (2003), Video Mining, Kluwer Academic Publishers, London.
[10] Chen A. L., LIU C., and KUO T. C., (1999), Content-Based Video Data Retrieval, Vol 23, No 4, pp. 449-465.
[11] Karmouch A., Abu-Hakima S., Ahmed M., (2004), Keyframe Extraction and Indexing For Multimedia Databases, Vision Interface, pp. 506-511.
[12] Fauvet B., Bouthemy P., Gros P., Spindler F., (2004), A Geometrical Key-Frame Selection Method Exploiting Dominant Motion Estimation in Video, Springer-Verlag Berlin Heidelberg, pp. 419-427.
[13] Girgensohn A., Boreczky J., (2000), Time-Constrained Keyframe Selection Technique, Multimedia Tools and Applications, Vol 11, pp. 347-358.
[14] Shahbaa Ibraheem Khaleel Al-Taey, (2006), Efficient Content-Based Multimedia Retrieval With Neural Network, PhD Thesis, College of Computer Science and Mathematics, Mosul University, Iraq.