ISSN:2321-1156

International Journal of Innovative Research in Technology & Science (IJIRTS)

PROFICIENT RETRIEVAL OF SIMILAR VIDEOS

M.Narmada Devi[1],V.Keerthana[2],V.Poojalakshmi[3],A.Packialatha[4]
[123]Students,[4]Assistant Professor,
Department of Information Technology,
Jeppiaar Engineering College,
Chennai-600119.

Abstract ***—In recent times, videos form a large part of data in the internet. Owing to its ability to attract audience from all age groups, a large number of websites dedicated to browsing and viewing videos have mushroomed. But, unfortunately, search engines do not give adequate importance to retrieve videos. In all the video-oriented websites, the search is based on keywords typed in, which produces undesirable output in terms of relevancy. To achieve content-based video retrieval, we propose Proficient Identification of Similar Video System. In the proposed system, the background and object key frame of each and every shot of the video is extracted. Then the morphological shape of the background and objects are derived and represented as a matrix of 0 and 1 based on grid encoding. The Eigen values of these grid encoded matrices are used as metadata for indexing and retrieving videos. An index table is maintained based on the Eigen values of the background key frame matrices. The index table is structured in such a way that the background key frame matrices of videos with similar Eigen values shares the adjacent field in the table. The Eigen values of objects in the video are associated with the background key frame in the index table by means of linear chaining. Thus, we have effectively utilized the low level features of video to achieve ontological retrieval of videos.***

Key terms – Grid Based Encoding, Eigen values, linear chaining.

1. INTRODUCTION

Videos, of late, account for a major chunk of data in the online server. Videos have become one of the key means of dissemination of information and its use in the academic stream is assuming huge proportions. Besides, they satiate the needs of those who look for online entertainment, and this explains the existence of a large number of websites dedicated to browsing and viewing videos. Keywords are typed in to retrieve videos, but not much importance is given to criteria 'relevancy'. Since videos possess great entertainment value, they reach out to people of almost all age groups with internet access. When such is the case, relevancy and instant retrieval become a major need, which has been ignored till now.

In all the video-oriented websites, mining is based on keywords the user types in. Using the keywords, the search engine will look for the matching tags available in the videos. Each video will have many tags associated to it. *[Here tags refer to the concept on which the video is based on]. Most of the websites allow the user who uploads the video to specify their own tags. So, the tags are completely independent of the website's vision. In other websites, the words in the name of the video specified by the user will be used as tag words. Neither of the methods deals with the actual content of the video, but just takes words as filtering criteria for a video-based search.

The existing system exhibits the following flaws: 1.Browsing time is very high, since the results produced are vast. 2. Results are not relevant as the tag words may be generic. 3. There is no filtering process for redundant videos.

Hence, we propose a novel idea of filtration with the help of content-based video retrieval by taking into account the actual content of the video and not any tag/word provided by the user.

2. RELATED WORK

Significant work has been done in relation to our proposal with a similar objective, but from a different perspective. The initiation started way back in 2002, when video started garnering more attention from the online user. But, it was only able to propose a theoretical procedure for structure analysis of the images of the video for better filtration and retrieval [4], and failed to explain its practical implementation.

An enhanced work [13] to retrieve videos focused on multimodal visual features (colour and shapes) as well, it employed k-means algorithm for classifying the videos which suffers serious drawbacks like the prediction of k-value, dissimilarity in clusters, etc.

A further developed works [14] concentrated on motion estimation and edge detection and represented them in the form of histograms. The histograms employed in visual processing are much prone to noise interference and quantization errors.

Later, to overcome the difficulty of variation in the dimension between videos, a proposal came up to match low with high dimensional videos this contributed to video comparison factor [7]. With all the technological advancements came the new idea of feature extraction for comparison of videos in content matter with the help of video signature [6]. Even though this idea gave good similarity results, it is not practical to implement it in a busy network like Internet because of its high complexity and time consuming factor. Since time mattered a lot, indexing was simplified with the help of vector-based mapping, which slices the videos [8] and uses pointers, which performed great solely. Later, dynamic binary tree generation [9] came into being to save storage space, but it consumed time.

A very similar proposal to ours but complicated in its implementation came up, which uses threshold and colour histogram [10] to do content-based analysis. Next came a completely dedicated searching and retrieval method for MPEG-7 [5], which is not much in use in recent days. Personalized video searching with re-usability depending on the user came up with high caches [3]. This can come in handy for private use, but not for public consumption. When queries become difficult to express, a proposal came up to implement an application-based technology, combined with multi-touch exploitation which would result in compelling the user to give entry to an external application inside their browser [2].Finally, the basis of our proposal was from a content-based retrieval idea [1] which uses a complex grid method to retrieve videos using symbolization,[11] which is feasible but complex. We have attempted to minimize the complexity level with high responsive and relevant videos with limited time consumption.

3. SYSTEM DESCRIPTION

We propose an efficient content-based video retrieal system for identifying and retrieving similar videos from a very large video database. Here, searching is based on the input given as a video clip rather than the caption.

We store the video in an efficient manner so that retrieving is easier and more relevant. This is mapped by two-level indexing - first segregated on the basis of Eigen values of background encoded matrix, followed by object key frame encoded matrix.

The proposed system involves two different tasks 1) to store videos in an efficient way and 2) to retrieve video accurately using indexing techniques. The process of storing videos in a database involves five complex tasks as shown in figure (1) namely Video preprocessing, key frame generation, Pencil Shaded Pattern Extraction, Grid based Encoding and Indexing. (The user interface provides the input video to be stored.)

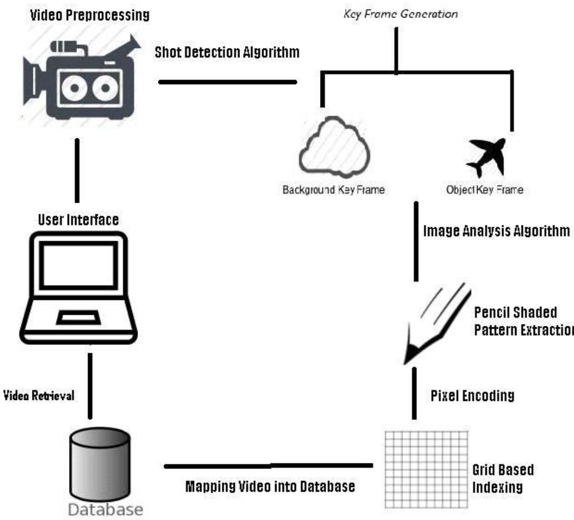

Fig. 1 System architecture diagram

4. IMPLEMENTATION

The proposed idea could be implemented in four phases.

4.1 VIDEO PRE-PROCESSING

Since the video is the most unstructured form of data, knowledge cannot be directly obtained from it. Hence, applying video pre-processing techniques becomes inevitable. As a shot acts as the basic entity of a video, shot detection algorithm has been applied to partition the video into shots. A shot is again broken down into a set of key frames/images.

4.1.1 BACKGROUND KEYFRAME GENERAION

This phase involves the generation of complete background key frame. The background generation is based on the principle of temporality. According to the principle of temporality, the moving object acts as the foreground of the key frame and the object that remains fixed acts as the background of the key frame.
Initially, the video shot is converted into multiple frames of pictures and the frames are numbered. Each frame $N_i[k]$ is compared with its consecutive frame $N_{i+1}[k]$ and if the pixel value remains the same for both the frames, the corresponding value is updated in the

background key frame. If the pixel value goes unmatched, the comparison is made between the next consecutive key frames. The value of unfilled pixels can be computed from the surrounding pixels.

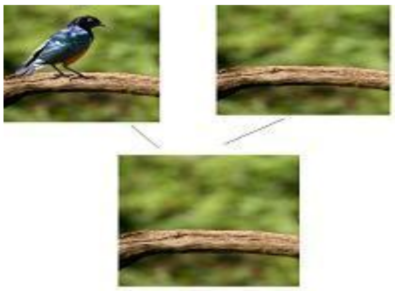

Fig.2 Background Key frame Generation

4.1.2OBJECT POSITION IDENTIFICATION

This phase involves the extraction of objects in the key frames of video shots. The simple subtraction of the constructed background key frame and any of the key frame maps the object in the video shot.

Fig.3 Identification of Object Position by filtering the pixel that do not match

4.1.3 PENCIL SHADED PATTERN EXTRACTOR

Once the background and object key frames are built, those key frames are led to the pencil shaded pattern extractor works based on simple image analysis algorithm.

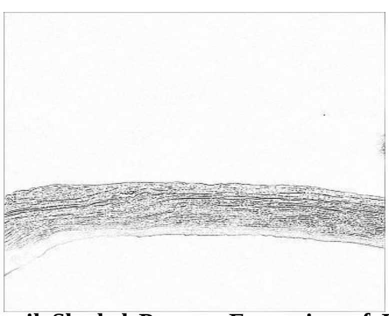

Fig.4 Pencil Shaded Pattern Extraction of Background Key Frame

Image analysis algorithm computes a value for pixel based on its four adjacent pixels. If all the four pixels have the same quantized colour, it is marked as an interior pixel and it remains not shaded. If any one of the four pixels have different quantized colour, it is marked as a border pixel and shaded.

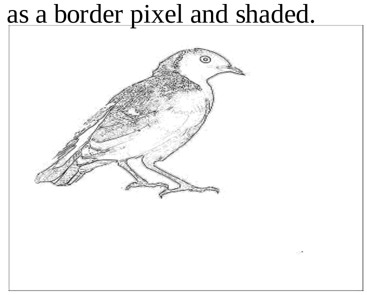

Fig.5 Pencil Shaded Pattern Extraction of Object Key Frame.

4.1.4 PIXEL ENCODED BY GRID BASED IDEXING

The morphological shape of background and object is fixed on to a grid of fixed cell size in a manner that the shape is justified to the top left corner. The scanning of grid is performed from left to right and top to bottom.

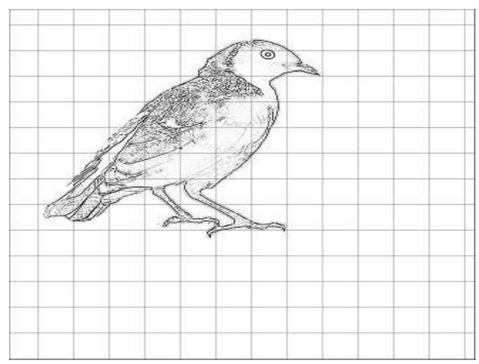

Fig.6 Pixel Based Encoding of Object Key Frame

one is assigned to the cells of the grid partially or wholly covered by the shape and 0 to the cells outside of the shape boundary which gives us a sequence of numbers which can be used for shape representation. The sequence of binary digits represents the background and the object key frame [12]. But, it must be noted that the binary number obtained for the same shape with a different orientation in space or with a different scale will be different.

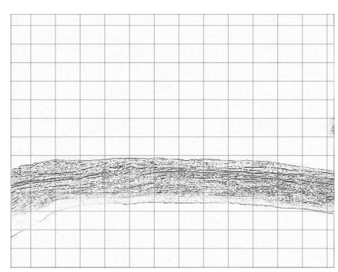

Fig.7.Pixel Based Encoding of Background Key Frame

Let us consider a grid matrix G of order 4*4 and the matrices of background key frame B_k and object key frame (of order 4*4) where pixel contains shape are represented as 1 else 0.Pixel encoding involves the product of grid matrix G with B_k or O_k leaving the resultant matrices PB_k or PO_k.Further,the characteristic equation of PB_k and PO_k are defined and Eigen values of PB_k and PO_k are computed which are used for indexing.

$$G=\begin{pmatrix} 1\,1\,1\,1 \\ 1\,1\,1\,1 \\ 1\,1\,1\,1 \\ 1\,1\,1\,1 \end{pmatrix} \qquad Bk=\begin{pmatrix} 0110 \\ 1001 \\ 0111 \\ 0101 \end{pmatrix}$$

$$PB_k = \begin{pmatrix} G11*Bk11 & G12*Bk12 & G13*Bk13. \\ G21*Bk21 & G22*Bk22 & G23*Bk23. \\ G31*Bk31 & G32*Bk32 & G33*Bk33. \\ G41*Bk41 & G42*Bk42 & G43*Bk43. \end{pmatrix}$$

$$PO_k = \begin{pmatrix} G11*Ok11 & G12*Ok12 & G13*Ok13. \\ G21*Ok21 & G22*Ok22 & G23*Ok23. \\ G31*Ok31 & G32*Ok32 & G33*Ok33. \\ G41*Ok41 & G42*Ok42 & G43*Ok43. \end{pmatrix}$$

$Det(PB_k - \lambda I)=0$ [15] computes the Eigen value for the background key frame matrix and $Det(PO_k - \lambda I)=0$ computes the Eigen value of the object key frame matrix .

4.2 MAPPING INTO DATABASE

An index table is maintained based on the Eigen value of the background key frame. The index table

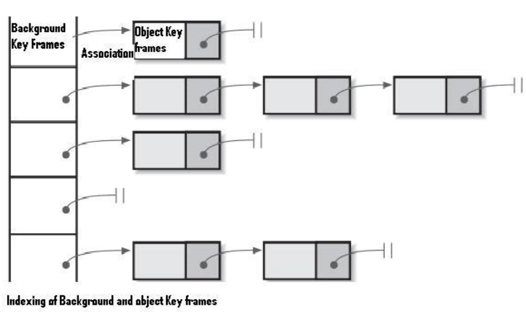

Fig.8 Associating Background and Object Key Frame in Indexing Table

is structured in such a way that the background key frame matrices of videos with similar Eigen values shares the adjacent field in the table. The objects in the video are associated with the background key frame in the index table by means of linear chaining.

4.3 RETRIEVING RELEVANT VIDEOS

To retrieve similar videos more efficiently and effectively, several key issues need to be noted. A video has to be represented compactly and informatively. The second issue is to measure the similarity between videos. Thus, searching in large databases over a raw video data is computationally expensive. To overcome this, we extracted the background and object key frame of each shot of the input video, then grid encoded matrices are computed and their Eigen values are compared with the Eigen values of videos in the indexing table to retrieve the most appropriate video.

Since Two levels searching is performed based on the Eigen value of background as well as the Eigen value of the object, the precision rate reaches a new peak value. As it is well proved that the similar matrices share same Eigen vector [15].

Therefore, the user can select any retrieved video and playback the video clip. Figure 9 shows one of the sample examples of the retrieval result. The retrieval result will be even better when the backgrounds are masked out. On the other hand, if the background becomes much clumsy or its area increases, the results will degrade gradually.

But the current video search engines are based on lexicons of semantic concepts and perform tag-based queries. These systems are generally desktop applications or have simple web interfaces that show the results of the query as a ranked list of key frames. For each query result, the first or similar frames of video clip is shown. These frames are obtained from the video streaming database, and are shown within a small video player. Users can then play the video sequence and, if interested, zoom in each result, displaying it in a larger player that shows more details and allows better video detection. The extended video player also allows searching for visually similar video clips.

Therefore, at the bottom of the result lists, there are the concepts which are related to the video results. By selecting one or more of these concepts, the video clips returned are filtered in order to improve the information retrieval process. The user can select any video element from the results list and play it as they need. This action can be repeated for other videos, returned by the same or other queries. Videos out of the list can be moved along the screen, resized or played. Therefore, the overall retrieval process is simple, effective and quick.

Fig.9 Sample Example of Retrieval result

5. EXPERIMENTAL SETUP

We implemented the shot detection algorithm by using SID video cutter. The naïve pixel algorithm used for generation of background and object key frame is implemented using mat lab codes. The pencil shaded patterns of the background and object key frame are obtained using Foto Sketcher tool. The grid based pixel encoding matrices are obtained using Mat lab codes. The Eigen values are computed and stored in Oracle Multimedia using ODBC connection of mat lab.

'Fig.10. Proficient Retrieval System

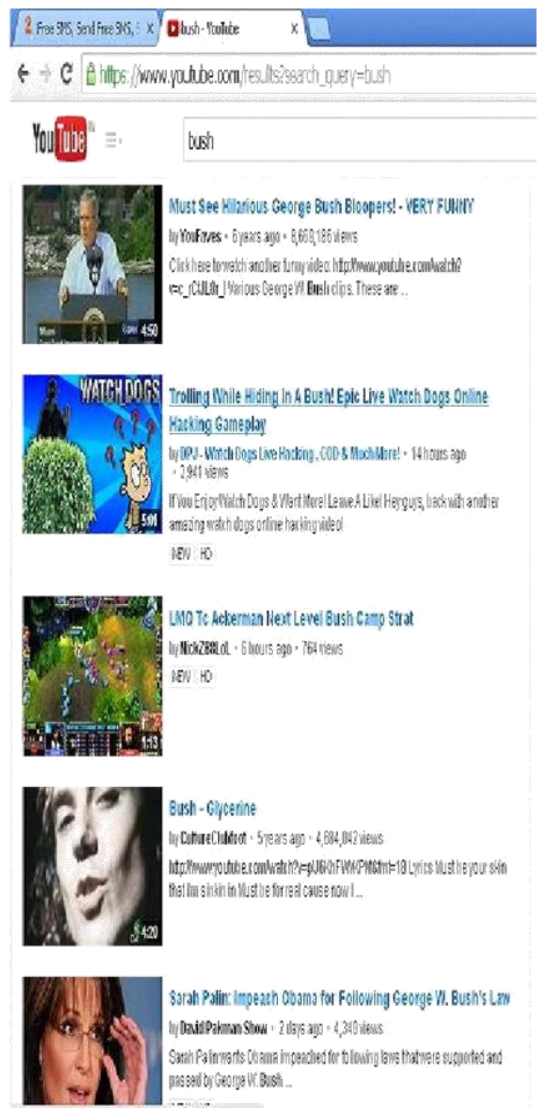

Fig.11 Existing Video Retrieval System

6. CONCLUSION

In this paper, we have presented an effective CBVR (Proficient Identification of Similar Videos). We obtain the morphological shape of the background and the object in the video shot and they are encoded based on the grid to obtain pixel encoded matrices, which acts as the metadata for the indices to retrieve videos. In further works, we will try to develop a much more effective algorithm to obtain the background and object key frame, since the naïve pixellize algorithm used is slightly prone to noise disturbance in the pixels.

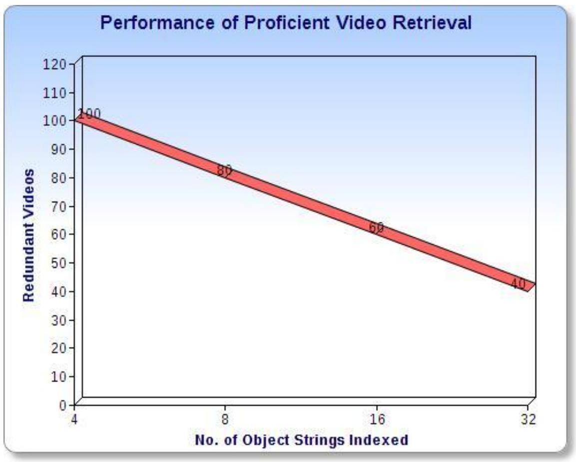

Fig.12 Evaluation of Proposed System based on Object Key Frames

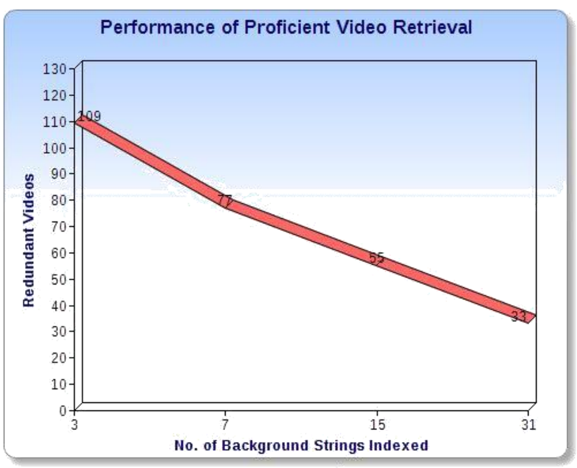

Fig.12 Evaluation of Proposed System based on Background Key Frame

7.REFERENCES

[1] Xiangmin Zhou,Xiaofang Zhou, Lei Chen,Yanfeng Shu,Bouguettaya and A Taylor, "Adaptive subspace symbolization for content based video Detection ",IEEE on vol. 22,no.10, 2010.

[2] Marco Bertini, Alberto Del Bimbo and Andrea Feracani, "Interactive Video Search and Browsing system", 9th International workshop on CMBI, 2011.

[3] Victor Valdes and Jose M.Martinez, "Efficient Video Summarization and Retrieval",9th International workshop on CMBI, 2011.

[4] Ng Chung Wing," ADVISE: Advanced Digital Video Information Segmentation Engine", IEEE, 2002.

[5] Quan Zheng and Zhiwei Zhou,"An MPEG-7 Compatible Video Retrieval System with Support for Semantic Queries", International Conference on CECNet,2011.

[6] Sen-Ching, S.Cheung, and Avideh Zakhor, "Fast Similarity Search and Clustering of Video Sequences on Wold Wide Web", IEEE on vol.7,no.3, 2005.

[7] B. Cui, B.C.Ooi, J.Su and K.L.Tan, "Indexing High Dimensional Data for Efficient similarity search", IEEE transactions on Knowledge discovery and Data Engineering, vol.17,no.3, 2005.

[8] H.Lu, B.C.Ooi, H.T.Shen and X.Xue,"Hierarchical indexing structure for efficient similarity search in video retrieval", IEEE transactions on Knowledge discovery and Data Engineering,vol.18,no.11, 2006.

[9] V.Valdes,J.M.Martinez,"Binary tree based on line Video Summarization",Proceedings of the 2nd ACM TRECVid Video Summarization Workshop, pages 134-138,2008.

[10] Sathish Kumar,L Varma and Sanjay N Talbar, "Dynamic threshold in Clip Analysis and Retrieval" ,IEEE ,2011.

[11] PackiaLatha A, ChandraSekar A," Adept Identification of Similar Videos", IJERSS, vol.1,no.4,2014.

[12] Ranjen Praker,"Principles of Multimedia", Tata McGraw hill,2006.

[13] Madhav Gitte, Harshal Bawaskar, Sourabh Sethi, Ajinkya Shinde," Content based video retrieval system", IJERT, vol.3, no.6, 2014.

[14] T.N.Shanmugam and Priya Rajendran,"An enhanced content based video retrieval system based on query clip", IJRRAS, vol.1, no.3, 2009.

[15] T.Veerajan,"Engineering Mathematics", T -Hill, 2005.