

Effective Video content Retrieval using Image attribute standards

Saravanan Devaraj¹

¹The ICFAI Foundation for Higher Education(Deemed –to-be-University), IBS Hyderabad, Telangana, India

Abstract:

Technology brings the important of image retrieval for various operations like education, defense, medical and more. Creation of video files are very easy to any user, retrieve the same is one of the major challenge for user community. Retrieve the particular image files really a complex process for user, for this user need a specific domain knowledge are required. This leads to an urgent focus on imager retrieval process, there are various existing algorithm are available, all of these works for particular video set only, none of this never work for all set of video files. In the proposed technique new video data retrieval technique proposed with help of image pixel values. First images are trained for extraction, during this UN wanted images are removed using pixel difference. After removing noisy data image frames are grouped and stored separately for further operations. The proposed technique works well for all type of video files, experiments verify this.

Keywords: Data mining, Key frame image, Image pixel, Color values, Histogram, Clustering, Hierarchical Clustering Image extraction.

Received on 18 April 2018, accepted on 07 May 2018, published on 12 June 2018

Copyright © 2018 Saravanan Devaraj, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.12-6-2018.154818

1. Introduction

Multimedia is an irreversible trend in computing today. Because multimedia not only provides the most expressive way of representing and browsing, but also enhances the overall quality and quantity of information, most areas of information technology are embracing multimedia. As the result of these developments, multimedia, and especially video, will bring dramatic improvements to human-computer interaction [1]. Video is by far the most powerful and expressive multimedia because it is a streaming media with high resolution and multiple channels with real multimedia including audio, visual and text information. These properties of video make it a popular medium for capturing and presenting information. At the same time as video has multiple channels to convey different media, and it requires massive storage space, it also presents technical challenges for the improvement of data management. Many computer vision problems require the efficient and effective organization of huge-dimensional data for information retrieval purposes [2]. Unsupervised learning, mostly clustering, provides a way to handle these challenges. Clustering is a technique to split a set of objects (frames) in groups (video) such that similar objects are grouped together, while objects that are not similar fall in different clusters [3, 4]. The choice of the notion of similarity (or distance) among objects that are needed to be clustered is of crucial importance for the final result. Clustering algorithms have no a-prior knowledge about the data domain, its hidden structure and also the number of hidden classes in

which data are divided is unknown. For this characteristic, clustering is often referred as un-supervised learning in contrast to classification (or supervised learning) in which the number of classes is known and for each class a certain number of examples are given. The independence of clustering algorithms from the data domain is at the same time the secret of its success and its main drawback. In fact since clustering does not need any a-prior knowledge of the data domain; it can be applied to a wide range of problems in different application areas [5]. In contrast, general purpose procedures do not allow applying (even trivial) problem dependent optimizations and consequently they typically perform worse than ad-hoc solutions.

1.1 Existing system

- In existing system, multimodal video indexing approach was used for detecting various semantic concepts consisting of mainly objects and scenes.
- This research work focused on efficient and scalable methods for annotating Web videos at various levels including objects, scenes, actions, and high-level events.

1.1.1 Issues in Existing system

- Many existing works have only employed the visual modality for event recognition, it is important to emphasize that video analysis is intrinsically multimodal, demanding multidisciplinary knowledge and tools from many fields

^{*}Corresponding author. Email: sa_roin@yahoo.com

- Scalable indexing methods and parallel computational platforms are also becoming an important part of modern video analysis systems.
- One well-known challenge is the long-standing semantic gap between computable low level features (e.g., visual, audio, and textual features) and semantic information that they encode.
- Image duplication is not done.
- Image retrieval process takes long time.
- Key input selection are done[6]
- Method works well in particular type of video files only.

2. Proposed System

2.1 Proposed Architecture

The input video are converted into frames, are grouped in to set of basic units this is the first step for video data mining. Each frame the image features like object, color, text, motion is extracted in order to differentiate the one frame with other. After this duplicate frames are eliminated using image RGB threshold value, this process is applied to all existing frames [7, 8]. After determining a threshold according to the RGB values, duplicate frames are removed, frames which never match with adaptive threshold valued. After successful removal of duplication rest frames are grouped with help of clustering technique and image key frame selected for image retrieval through this image are extracted.

- RGB values are used to remove redundant frames in a video.
- Model based Clustering algorithm is used to cluster the frames.
- SIFT feature is used to retrieve images in an unconstrained video.

2.1.1 Advantage of proposed system

- Reduced Complexity
- Increased Accuracy
- Increased Efficiency
- Key frame identification is done.
- Duplication frame are identified and eliminated.

3. Experimental Setup

- 3.1 Error Removal
- 3.2 Video Segmentation
- 3.3 Clustering Algorithm
- 3.4 Image Retrieval and feature extraction

3.1. Error Removal

Error removal is the preliminary process of our proposed framework. Video input files are converted in segments. Each segments works as a shot of the given input video file. Today technology brings image as important effective tool for exchange information instead of textual information exchange. Creation of any video files is easy task for many users. Quality of this input files are affected by many factor poor image quality, images are not captured by trained person, poor lighting, poor focal and more. For this video are converted into frames, not of unwanted information's are available after it spitted into frames. It is necessary any video data mining, first remove this unwanted frames. Unwanted frames i.e. repeated frames, frames as poor quality, bull red images, pixel quality and more. Those unwanted information are removed from the group of frames. In video files images are ordered based on the time manner this process gives the meaning of the input frames. Error identifications are done with various images input functions such as

- Color
- Shape
- Characters Position available in the frame.
- Time
- RGB Value and more

The above any of the functions used for differentiate one frame with others. Proposed work image Pixel values are consider. It helps to find the difference between one image frames to other image. Input of this module is user's sample video. This received video has converted into number of frames. This process is explained in terms of flow graph below (Fig 2)

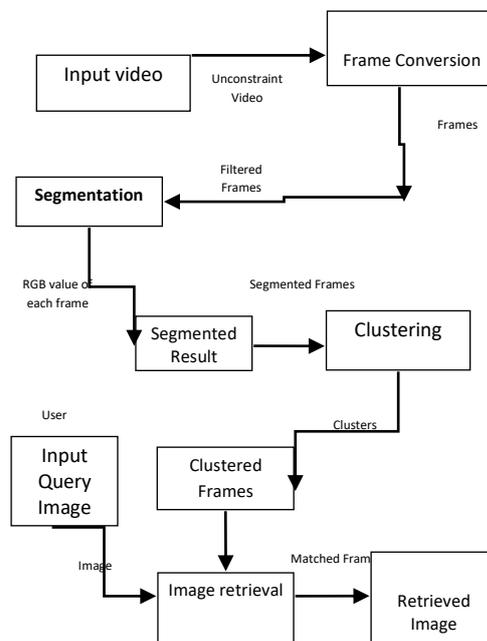


Figure.1 Proposed system flow diagram

3.2. Video Segmentation

Here segmented video frames Calculate RGB value for removing redundant frames. In brief, initially calculate red, green and blue values for each frame. After calculating RGB values, set threshold for removing redundant frames. If RGB values is greater than threshold hence the frame is considered as valid else neglect the frame. The process explained in the below flow graph (Fig 3)

3.3. Clustering Algorithm

The segmented images are taken as an input to the clustering process. In clustering is the best step for identifying the relationship between the data set. Video clustering differs from the traditional clustering process. Video file consist of audio, motion, image, text and graphic among this complex information grouping of information is a challenge process. Most video data's are unstructured for this any video data mining needed pre-processing steps. Video files are converted to image frame most of the frames are redundant and some frames are treated as noisy i.e. unwanted frames. For this video data mining process perform pre-processing step, it remove unwanted image frames i.e. noisy data, repeated frames are removed from the data base. After removing error, information are integrated and stored separately for further operations.

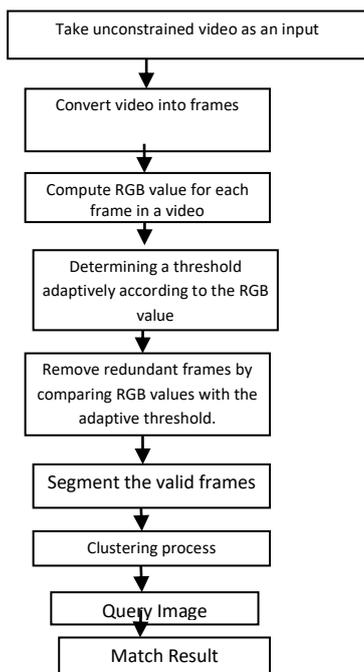


Figure.2 Frame Error removal process flow diagram

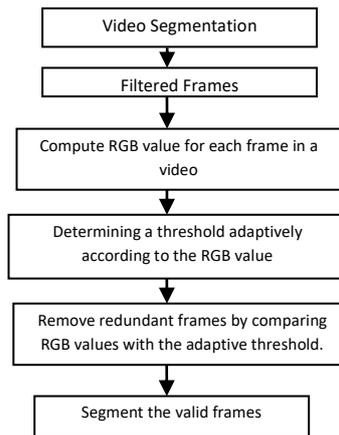


Figure.3 Frame Segmentation process

3.3.1 Pseudo code for cluster formation

```

    {
        return ((Cluster)obj).members.size() - members.size();
    }
    public double getDistance(Feature feature) {
        double d = 0.0D;
        for(int i = 0; i < mean.length; i++) {
            double d1 = mean[i] - feature.descriptor[i];
            d += d1 * d1;
        }
        return Math.sqrt(d);
    }
    float mean[];
    HashSet members;
}
    
```

3.4. Image Retrieval and feature extraction

After eliminating the duplicate frame form the give image data set, images are grouped using clustering technique . Clustering process is explained above. After clustering images are trained for the recovery process. This process normally done in two step process one from the client side other from server side. In server side video frames are trained and stored in the server data base. It helps reduce the searching time. In client side every time user gives image as an input , feature of those input frame are taken that values are compared with existing data base values. Values which matches are consider as similar frames, it retrieved and given output as users input query. This module describes the details of image retrieval and feature extraction process. Initially user gives query image to the server. The server takes a query image. Server compares the query image with clustered images. This feature eventually checks the clustered image. Finally, the matched values are displayed to the user.

3.4.1 Algorithm for Image retrieval process

- Step1: Take different videos set i.e. cartoon, debugging, song etc.
- Step2: Perform segmentation process and extract the frames of that video set.
- Step3: Apply image preprocessing step and extracted frames.
- Step4: Segment the frames using the image low level features
- Step5: Group the frames using cluster concept.
- Step6: Store the clustered frames in the database

Step7: Select key frame and give an image input query
 Step8: Based on input the server extracts the features.
 Step9. Retrieves the related video to the requested user.

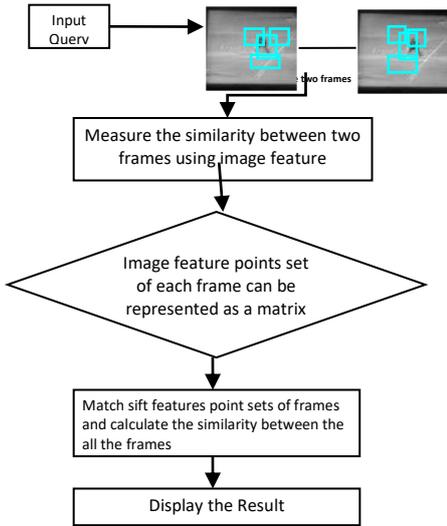


Figure.4 Image Retrieval process

4. Experimental outcomes



Figure.5 Frame conversion for Sports video file.



Figure.6 Frame conversion for Cartoon type video file

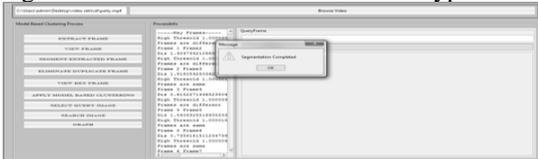


Figure.7 Image Segmentation

Table.1 Video Frames Vs Duplication Removal process

| frmcnt | milliseconds | category |
|--------|--------------|---------------|
| 25 | 4153 | Entertainment |
| 50 | 4189 | Entertainment |
| 75 | 4219 | Entertainment |
| 100 | 4255 | Entertainment |
| 125 | 4359 | Entertainment |
| 150 | 4375 | Entertainment |

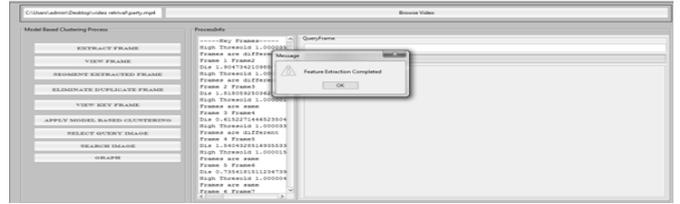


Figure.8 Eliminating Duplicate frames using image histogram technique



Figure.9 After duplicate eliminating frames grouping using clustering

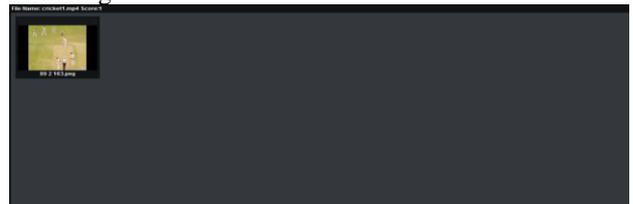


Figure. 10 Selecting Key Frame

Table 2 Videos Vs Cluster formation.

| S.No | Video | Frame | Cluster | Time |
|------|-------|-------|---------|------|
| 1 | 4 | 1 | 2 | 760 |
| 2 | 4 | 2 | 2 | 789 |
| 3 | 4 | 3 | 2 | 789 |
| 4 | 4 | 5 | 2 | 778 |
| 5 | 4 | 4 | 2 | 660 |
| 6 | 4 | 8 | 1 | 669 |
| 7 | 4 | 9 | 1 | 987 |
| 8 | 4 | 10 | 1 | 885 |
| 9 | 4 | 11 | 3 | 552 |
| 10 | 4 | 12 | 3 | 778 |
| 11 | 4 | 12 | 3 | 887 |
| 12 | 4 | 14 | 2 | 858 |
| 13 | 4 | 15 | 2 | 788 |
| 14 | 4 | 1 | 2 | 555 |
| 15 | 4 | 2 | 3 | 788 |
| 16 | 4 | 3 | 5 | 980 |
| 17 | 4 | 4 | 5 | 555 |
| 18 | 4 | 6 | 5 | 787 |
| 19 | 4 | 8 | 5 | 888 |
| 20 | 4 | 9 | 5 | 558 |
| 21 | 5 | 9 | 5 | 887 |
| 22 | 5 | 10 | 6 | 888 |
| 23 | 5 | 10 | 6 | 887 |
| 24 | 5 | 10 | 6 | 888 |
| 25 | 5 | 10 | 6 | 852 |
| 26 | 5 | 11 | 6 | 859 |
| 27 | 5 | 11 | 6 | 889 |
| 28 | 5 | 11 | 6 | 887 |
| 29 | 5 | 12 | 6 | 887 |
| 30 | 5 | 12 | 7 | 888 |

Figure.11 Select Frame based on Key frame selection

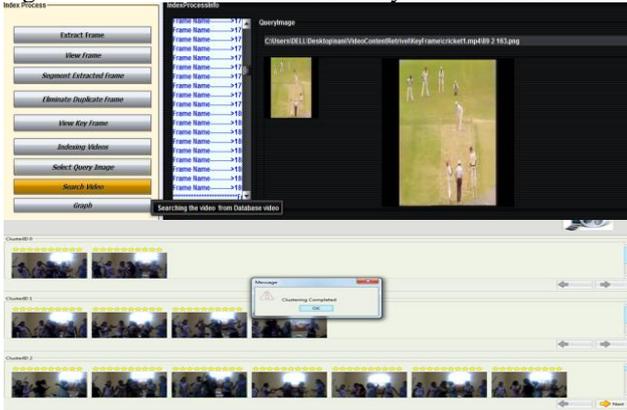


Figure.12 Cluster Formation 1 input 2, 3 7 group of cluster.

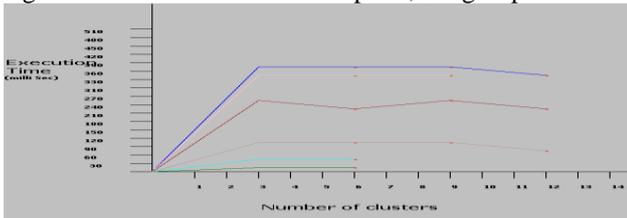


Figure.13 performance graph of Cluster formation.

Table .3 Video files Vs Duplication Removal

| Video name | Number of | Duplicate frames |
|------------|-----------|------------------|
| Cartoon | 10 | 3 |
| Graphics | 26 | 16 |
| Meeting | 15 | 2 |
| Globe | 27 | 12 |
| Song | 15 | 2 |

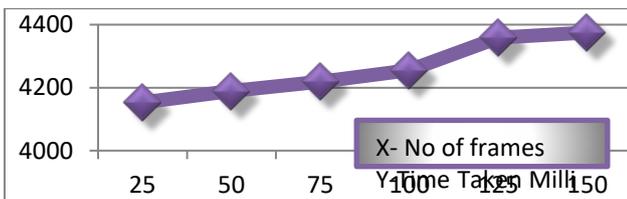


Figure. 14Performance graph of frame count Vs time

Table 4 Frame Count Vs Cartoon File

| Frame count | Time(Mil.Sec) | Category |
|-------------|---------------|----------|
| 25 | 4153 | Cartoon |
| 50 | 4189 | Cartoon |
| 75 | 4219 | Cartoon |
| 100 | 4255 | Cartoon |
| 12 | 4359 | Cartoon |
| 150 | 4375 | Cartoon |

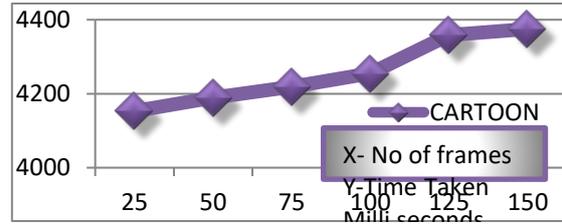


Figure.15. Performance graph of frame count Vs Time

Conclusion and future enhancement

In this paper we have presented a new framework for image retrieval in unconstrained video. Our system utilizes four major techniques for efficient image retrieval. Initially user gives sample video as an input. The received video was converted into frames. Then, image histogram is applied to remove the errors in an image. After error removal process repeated and unwanted frames are removed using the RGB value. After completing frame extraction process, clustering is utilized for clustering the images. Moreover, Image feature is utilized for predicting the image based on the user query. Our proposed system efficiently retrieves the images results are verified this.

References

- [1] .D.Saravanan, "Effective Multimedia Content Retrieval", *International Journal of Applied Environmental Sciences* ,Volume 10, Number 5 (2015), pp. 1771-17783.
- [2] Barbu, A., Bridge, A., Coroian, D., Dickinson, S., Mussman,S., Narayanaswamy, S., Salvi, D., Schmidt, L., Shangguan, J.,Siskind, J.M.,Waggoner, J.,Wang, S.,Wei, J., Yin, Y., Zhang, Z. "Large-scale automatic labelling of video events with verbs based on event-participant interaction", In: *Proc. Of International Conf. on Computer vision and pattern recognition*Xiv:1204.3616v1,pp.01-13,2012.
- [3] D.Saravanan, "Performance Anlalysis of video data image using Clustering Technique", *Indian journal of science and technology*,Vol 9(10), DOI: 10.17485/ijst/2016/v9i10/79731, March 2016, ISSN (Print) : 0974-6846. pp.01-06.(IF 1.05).
- [4] D. Saravanan and S. Srinivasan, "Data mining framework for video data",In: *Recent Advances in Space Technology Services and Climate Change (RSTSCC)*, 2010, Chennai, 2010, pp. 167-170
- [5] Jiang YG SUPER: "Towards real-time event recognition in internet videos" In: *Proc. of ACMinternational conference on multimedia retrieval,(ICMR)*, pp. 849-852. (2012).
- [6] D.Saravanan "Text information Reterival using Data mining Clustering Technique" *International Journal of Applied Engg.Research*" Vol.10,No 3(2015) , pp.7865-7873.
- [7] Saravanan,"Segment Based Indexing Technique for Data file". *Procedia of computer Science*,87(2016), pp.12-17,2016.
- [8] Saravanan,"Video Substance Extraction Using image Future Population based Techniques", *ARPN Journal of Engg., and applied science*, Vol.11, No(11), pp.41-7045