

A Survey on Various Techniques used for Video Retrieval

Kusuma Bai R

Dept. of Computer Science & Engineering, New Horizon
College of Engineering, Bangalore, India,
kushmarajruk77@gmail.com

Priti Badar

Dept. of Computer Science & Engineering, New Horizon
College of Engineering, Bangalore, India,
priti_badar@yahoo.co.in

Abstract: *In today's modern world internet is widely being used by everyone. Many people access the internet to get their desired information. This information can be of text, image, audio or even the combination of all the form. In brief this method of retrieving information is known as the multimedia information retrieval. In this paper we discuss about the multimedia information retrieval. Video retrieval is one form of the information retrieval. YouTube is the most popular video application. CCTV is used almost everywhere these days. This paper discussed about how the information is retrieved from the CCTV footage. Here we have discussed about the video retrieval and its importance in various applications. We have also summarized about various techniques that is being used for the video retrieval. In future a model can be designed in such a way that the video retrieval can be done just by analysing the users mood.*

Keywords: *Multimedia information retrieval; Video retrieval; YouTube; CCTV.*

I. INTRODUCTION

A. Multimedia information retrieval

Multimedia information retrieval can be defined as the method that is used for searching and finding for multimedia documents. Multimedia literally means more than one media. It can be audio, video, text, image or a combination of everything. Search engines play a vital role for retrieving any kind of information from the internet.

Multimedia information can be retrieved in many ways. For example, if you come across a monument and if any details about it is required then a picture of it can be captured and sent to the internet and later information will be retrieved in the form of text. A medical image can be given as input query and can be used to get the details about the status of the patient's health [1].

Information can be retrieved in many ways. Text based retrieval is the basic way for retrieving any information. By giving a query in the form of text information in various methods can be retrieved. The retrieved information can be in the form of an image or a video or document that is in the form of text. Therefore

there are various types of retrieval such as image retrieval, audio retrieval, video retrieval etc.

B. Video retrieval

Video retrieval is a form of extracting the required video shot from a large video database. Video retrieval can be done in many ways. It can be retrieved by a simple text/phrase, it can also be retrieved by giving a piece of video itself as a query, image can also be given as the input to retrieve the video, it can also be retrieved by audio query.

YouTube is an example of retrieving the video by giving the text/phrase. Based on the input query given this application will play the appropriate video. It will also work in the same fashion when the voice note is used instead of the text. There are many applications of video retrieval. Like seen above, it is useful in YouTube which one of the example for multimedia. CCTVs are used in many places these days. It can be public places like malls, airports etc. It is also used in the main roads to monitor the vehicle movement. There are cases where it is required to access these recording videos to get some information. In such cases video retrieval plays an important role.

There are many techniques that are being used for the video retrieval. The following section discusses the various such methods that are used to implement the video retrieval.

II. TECHNIQUES USED FOR VIDEO RETRIEVAL

Content based video retrieval is the most common and simple method to retrieve the video from the database videos. In this technique we make use of SVM classification also. So by combining both CBVR and SVM it is easier to classify the video and also to retrieve the desired video.

Video is nothing but a group of frames. It is required to convert the image frame into a suitable colour model. In this technique grey-scale colour model is chosen. Once this was done a csv file is used for classification which splits the entire data set into both training image and testing image.[2]

There are cases where it is required to extract some information from the video if any unusual event occurs. It can be either from the CCTV video or from the airborne

video. It is easier to retrieve the video from CCTV rather than airborne because of its low contrast. Since the size of surveillance videos are large it is required to compress it so it becomes easier to query it.

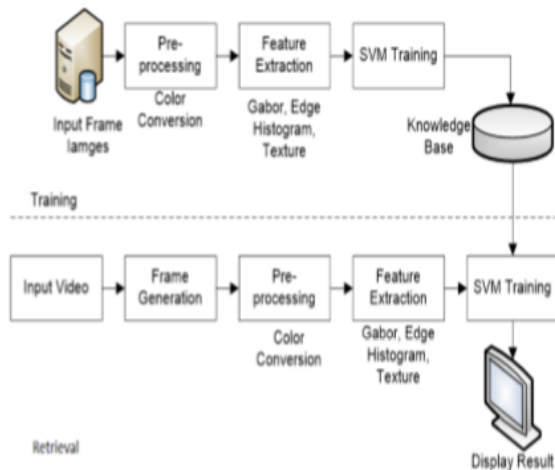


Fig 1. Video retrieval using CBVR and SVM

Video is nothing but group of frames. It is required to convert the image frame into suitable colour model. In this technique grey-scale colour model is chosen. Once this was done csv file are used for classification which splits the entire data set into both training image and testing image [2].

There are cases where it is required to extract some information from the video if any unusual event occurs. It can be either from the CCTV video or from the airborne video. It is easier to retrieve the video from CCTV rather than airborne because of its low contrast. Since the size of surveillance videos are large it is required to compress it so it becomes easier to query it.

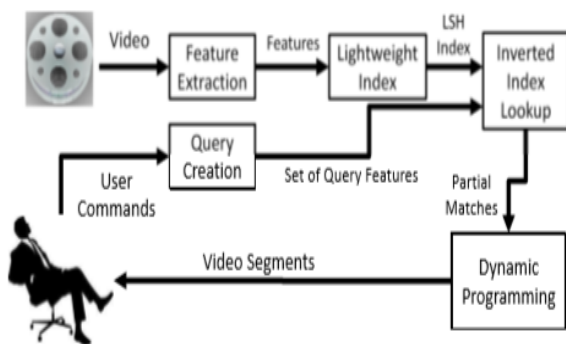


Fig 2. Video retrieval using dynamic programming

In CCTV video, we use HORN and SHUNCK method for the estimation of footage motion and for airborne tracklets. The required feature is extracted from the video and is put into a lightweight index. Dynamic programming is utilized to combine the partial matches along with full matches. Once the user gives the required query matching output is given [3].

E-lecturing videos play a major role in various organizations, research departments etc. This e-lecturing video involves both audio tracks and visual screen forms. The retrieval of lecture video is analyzed with the help of automated indexing of video and the search of video. The result is more efficient when the audio and text in the video are combined for the video retrieval more than the individual techniques [4].

As the internet is being used by almost everyone these days there are more than 100 million of hours of video that is uploaded and used in the internet. To retrieve video from such a big database is difficult by using the context based match method. This is because of the huge number of frames present in the video. User can give an image query of his/her choice and can retrieve the desired video from the video database. To overcome this problem, adaptive shot detection and a method called feature extraction was used for the faster retrieval of the video [5].

Video abstraction can be defined as the process of making a video short and precise in such a way that, if this is given a user query it will retrieve the original full video. This video abstraction will eliminate all the unwanted and redundant details. This can be either static or dynamic. This concept of video abstraction is integrated in surveillance video and for the retrieval of the video.

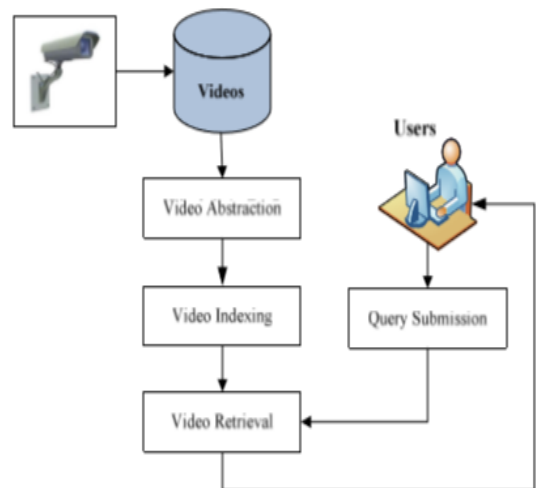


Fig 3. Video abstraction

The video from the CCTV are stored in the video database. After this process the video abstraction is applied on it so that all the unwanted and repeated information are removed from it. In the second step video indexing is done. This is mainly done for analyzing the video to detect, track and classify the objects. Finally, video retrieval is done based on the query submitted by the user [6].

It is possible to retrieve the video just by giving the image as a query. We try to implement the fisher vector and video descriptors. The results showed that this technique is more efficient in terms of memory and time

when compared to frame based method. In this process first the image query given by the user is shown as descriptors. To retrieve the video clips in a short list format the image query is queried against the index [7].

There are many challenges in the process of video retrieval. Problems can be because of the semantic gap i.e. the desired retrieval results expected from the user will not be up to the mark and also there are no enough matching techniques for the video retrieval. To overcome these two major challenges a new system was proposed and it works as follows:

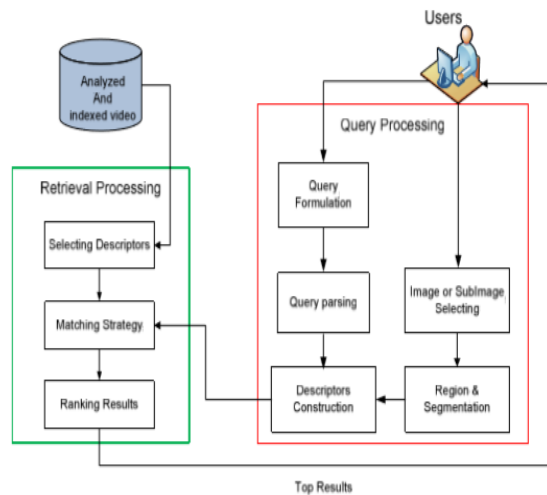


Fig 4. Process to reduce semantic gap

It is divided as two parts, query part and retrieval part. This approach supports different types of query and also region based and EMD blob based matching techniques were used for the matching purpose [8].

Text based video retrieval is commonly used technique for the video retrieval. However, this method is not efficient because it may ignore visual context. To overcome this problem a technique called video search reranking(VSR) was proposed. When the user gives text query to retrieve the video the result will be used to select few videos according to the user query.

The resulted video query will be measured according to its importance and also corresponding weight will be calculated, it represents the amount of likeness for the video query. These weights are then used for the reranking purpose. Along with the visual feature even the tag feature will be combined. This will increase the performance of the VSR [9].

Augmented reality combines both the real object and virtual object. It is nothing but creating a virtual object in the real life environment. It is possible to retrieve the video from such augmented reality with help of image queries. This process requires four different step to retrieve the video from the augmented reality. First the key frames are extracted from the video and then a feature extraction called THOG is applied. To find the nearest neighbour hashing technique is used. And finally

reranking of the video list is done by measuring the similarity between them.

III. CONCLUSION AND FUTURE WORK

This paper gives a brief introduction about the multimedia information retrieval and types of available retrieval methods. It also discussed about the video retrieval and its importance in various applications. In the second section of the paper the major techniques used for the video retrieval is explained in detail along with their downside.

In future a model can be designed in such a way that users can retrieve the desired video from YouTube just by analyzing the user's mood.

REFERENCES

- [1] <https://medlineplus.gov/paralyses>
- [2] Twinkle Sharma, Sachin Kumar, Naveen Yadav, Kritika Sharma, Piyush Bhardwaj, Air-Swipe Gesture Recognition Using OpenCV in Android Devices
- [3] BoYu Gao, HyungSeok Kim, Hasup Lee, Jooyoung Lee, Jee-In Kim, Effects of Continuous Auditory Feedback on Drawing Trajectory-Based Finger Gestures , Year: 2018.
- [4] D. H. De La Iglesia, G. Villarrubia, J.F. De Paz and J. Bajo, Design And Implementation Of A Low-Cost Universal Control For Intelligent Electric, , IEEE Latin America transactions, vol. 16, no. 5, Pages: 1328-1336, Year: May 2018
- [5] Tsung-Hsien Tsai, Jun-Wei Hsieh, Hung Chun Chen, and Shih-Chin Huang, Reverse Time Ordered Stroke Context for AirWriting Recognition, 10th International Conference on Ubimedia Computing and Workshops, Year: 2017
- [6] Basheer Alwaely , Charith Abhayaratne, Graph Spectral Domain Feature Representation for in-Air Drawn Number Recognition, , 25th European Signal Processing Conference (EUSIPCO), Pages: 370-374, Year: 2017
- [7] Zhangjie Fu, Jiashuang Xu, Zhuangdi Zhu, Alex X. Liu and Xingming Sun, Writing in the Air with WiFi Signals for Virtual Reality Devices, Year: 2018
- [8] Duo Lu, Kai Xu, Dijiang Huang,, A Data Driven In-Air-Handwriting Biometric Authentication System, , IEEE International Joint Conference on Biometrics, Pages: 531-537 ,Year: 2017
- [9] Chenyu Huang, Zhice Yang, Huangxun Chen, Qian Zhang, Signing in the Air w/o Constraints: Robust Gesture-based Authentication for Wrist Wearables, Year: 2017