Effective Video Data Retrieval Using Image Key Frame Selection

D. Saravanan

Abstract Owing to the rapid growth of multimedia technology, multimedia information is easily accessed by any user and the same information construction and distribution are also very easy. Due to technology development, the multimedia information increases due to variety of factors: it can be uploaded by unprofessional users nowadays. Due to the low quality and the large number of duplicated video files available, this leads video extraction more and more complex. The general method of representing each video segment is shot that consists of series of frames. Among this series, the input frame based shot method is specifically assisted for searching the video content as clients provided image query/search where an image will be matched with the indexed key frames with assist of resemblance distance. As a result, the key frames selection is most significant, and several methods are used to automate the process. This paper proposes a new technique for key frame selection. The proposed method shows significantly good and the experiments prove the above statement.

Keywords Video data mining · Data mining · Key frame extraction · Knowledge extraction · Multimedia data · Data extraction · Video dates · Frames

1 Introduction

Data mining is a process of detecting knowledge from a given huge set of data. Of the available huge data set, multimedia is the one which contains diverse data such as audio, video, image, text and motion, and video data play a vital role in the field of video data mining. In short, the application of video data is called video data mining. Data mining helps the users to retrieve the efficient content using data mining preprocessing operations. Increasing the quantity of video content reduces

D. Saravanan (✉)
Faculty of Operations & IT, IFHE University, Hyderabad 501 203, Telangana, India
e-mail: Sa_roin@yahoo.com

S.C. Satapathy et al. (eds.), *Proceedings of the First International Conference on Computational Intelligence and Informatics*, Advances in Intelligent Systems and Computing 507, DOI 10.1007/978-981-10-2471-9_15

the quality of the content; flexibility of the network produces duplicated information. Technology brings the huge amount of multimedia data sets. From this data set it is very difficult to the user, extract the needed informaton. For reduce the burden (Searching) data mining helps remove the unwanted infromation from the given data set. Nowadays, information is shared in the form of images instead of text information. Technology allows the users to share and upload this information in an easy manner. For retrieving these multimedia content such as image and motion, pixel values are quite difficult. This brings today most challenging operations for both users and researchers. While some basic forms of multimedia retrieval are available on the Internet, these tend to be inflexible and have significant limitations. The key frame selection is having two main issues: 1. The number of key frame(s) utilized, (The first issue is tackled by where the amount of key frames for every shot will be decided arbitrarily using the shot length). 2. The significant representative frame(s) selection in a shot. (The second issue is generally complicated for choosing the frames automatically with maximum semantic value. This issue is handled through minimizing the redundant frames with the help of the methods, for example, relevance ranking). Existing methods for searching video to identify co-derivatives have substantial limitations: they are sensitive to degradation of the video; they are expensive to compute; and checking the whole video files is quite complex, and also comparing the entire video content is not possible. Existing techniques perform direct comparison of video features between the query clip and the data being searched, which is computationally expensive.

2 Existing System

- There is no proper indexing and retrieval process available. Existing indexing techniques are suitable for only few sets of video files.
- Vast amount of video files are currently available on the web. There is not proper mechanism to arrange these contents. Proper arrangement reduces the searching time of the user.
- Current technique focuses on text-to-image retrieval; it never produces good result. Every search engine never returns the same type of information even if the users' query may be the same.

2.1 Drawback of Existing System

- This approach does not consider about the false positives and false negatives in the given video.
- The fault tolerance value is not reduced.

- Incorporating genetic variations into the design will affect the accuracy of this work.
- It considers only the action-based video mining.
- Human motion detection specifies only a particular region.

2.2 Advantage of Proposed System

- Compared to other search method, this method can also reduce the detection time.
- It first pre-processes the query image and extracts the features of that image.
- Trained videos are stored in the database, and the features of the trained videos are clustered using the extracted features of the query input image.
- Finally, features matching procedure is implemented to identify the similar features and to retrieve the relevant video.
- This method provides an efficient video retrieval using an image as input.
- Efficient clustering process is implemented.
- Features matching provides an efficient and accurate similar video retrieval

3 Literature Survey

Effective Multimedia Content Retrieval [1]. This paper brings the effective multimedia content retrieval using hierarchical clustering algorithm. Clustering provides grouping the data set effectively; it also reduces the searching time. Real-Time Human Pose Recognition in Parts from Single Depth Images [2]. The problem of predicting the human pose recognition in parts in a single depth image is discussed here. A new method should be proposed to quickly and accurately predict the position of the body joints from a single depth image. In Video Mining with Frequent Item set Configurations, [3] a new method for mining frequently occurring objects and scenes from videos is proposed. Object candidates are detected by finding recurring spatial arrangements of affine covariant regions. Content Based Image Retrieval using Color Histogram; [4] this paper brings the information retrieved based on the content; image can be extracted using image features such as text, pixel value, motion, frame value. Here, they proposed image feature vector technique, construct color histogram. An enhanced technique based on content-based image retrieval and video indexing [5]; this paper brings the new indexing technique for video data file using one of the image features like color. With the help of global color histogram (GCH) and histogram analysis, they create an indexing operation.

4 Experimental Setup

Experimental setup consists of a four-step operation. First, creation of an admin database; here, video frames are converted as shots, and those values are stored in the database [6]. Followed by users' image query, here, user input frames are taken, whose values are compared with existing stored value, and then, creation of key frame indexing; it reduces the searching time. The last phase called video retrieval is the process of user input frame matched with stored value, and similar value returned to the user. This experiment is conducted for various video files such as song, game, debate, news, and animated sets of video files. The entire process is constructed and tested using JAVA coding. Experiment is conducted as: for each video file, one input frame is selected, and then, based on the input, output extractions are done.

4.1 Creation of Admin Database

Key frame selection process is done with a two-step process. First phase called training phase or user side phase. The second phase is input query phase or server side phase. In the first phase, video files are converted as frames; after successful elimination of duplication frames, frames histogram values are calculated and stored in the admin database. Using file handling method, the duplicate files are eliminated. Key frame is determined by the frame that is the most same to the average frame. In the second phase, user input query; here, user input frame is compared with our stored frame value. Frames those are matching with input frame are extracted and returned to the user (Fig. 1).

4.1.1 Video Training Set Algorithm

Step1: Select input video file.
Step2: Segment the input file extract frames.
Step3: Eliminate duplication using image histogram technique.
Step4: Select key frames.
Step5: Per from Client and Servicer side operations.
Step 6: Based on users input, do video frame matching.
Step7: Retrieve and send matched result.

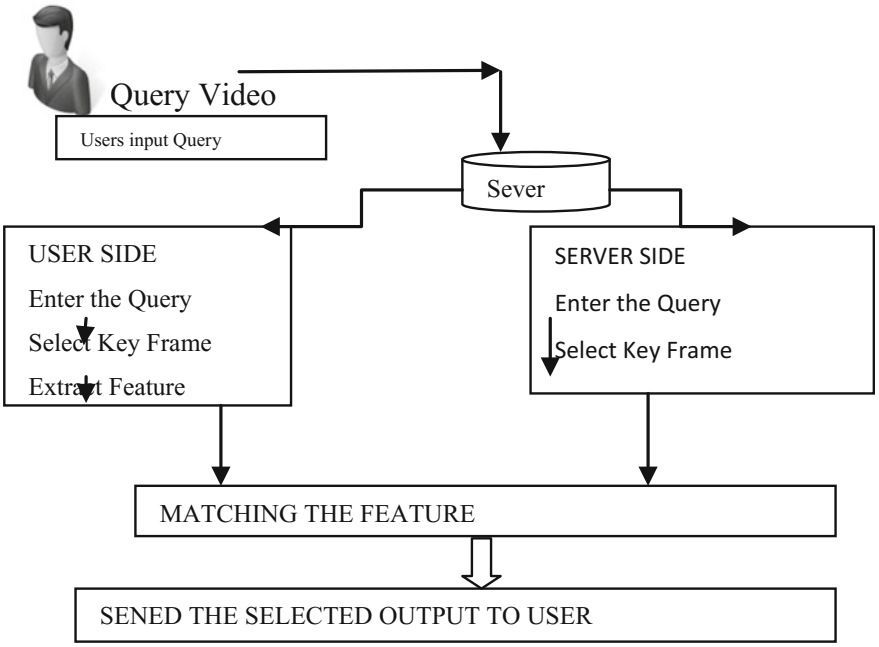


Fig. 1 Proposed architecture

4.2 User's Image Query

Fine matching stage is specially designed to choose which image in the image database is the most relevant one with that of the query image. Also, the measure to signify the degree of similarity has to be described. In this phase, input query image features are extracted using image feature extraction [7]. Extracted value is compared with admin-stored database values. The matching frames are returned to the user (Fig. 2).

Pseudo code for duplication removal:

```
Grey value = lngGrayScaleValue = (0.299 * clrPixel.R) +
(0.587 * clrPixel.G) + (0.1114 * clrPixel.B)
Grey = Σ Grey value
'Grey' gave the value of the grey value of the whole image.
Grey value = Image1.Greyvalue – Image2. Grey Value
If Grey value < threshold then Duplicate Image
```

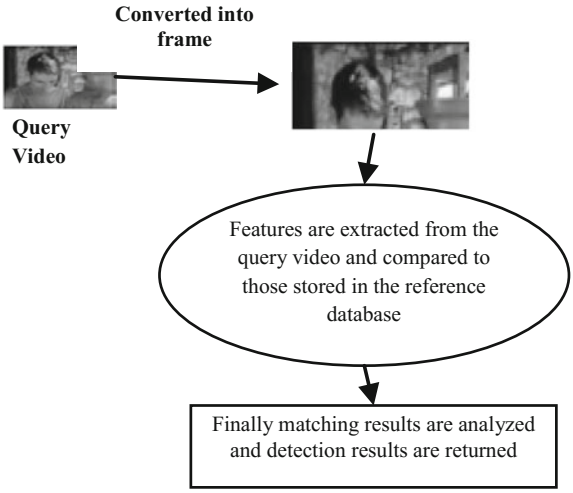


Fig. 2 User's video query

4.3 Video Frame Retrieval

Image retrieval is done with various techniques like histogram, image indexing, and algorithmic techniques. Indexing is the process to improve the performance and to reduce the searching time. Frame retrieval consists of the following steps:

Step1: Select the input frame.
Step2: Calculate the pixel value of input frame.
Step3: Convert the value obtained in Step2 into histogram value.
Step4: Compare this value with the existing value stored in the database (Fig. 3).

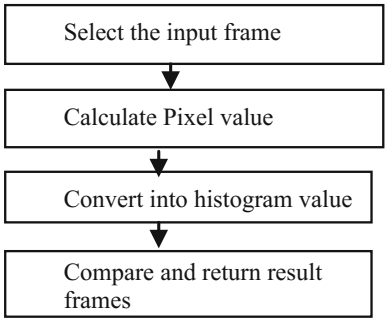


Fig. 3 Steps of video frame retrieval process

4.4 Creation of Key Frame Indexing

Due to the availability of video content in the web, it is very difficult to organize and retrieve the content. Increasing the usage of video data file, today, it is necessary for video data indexing. Many number of indexing techniques are available today, but the experimental result shows that each technique supports a particular type of video files only. For this reason, improvement is needed in each methodology. Here, we proposed key frame indexing technique for image retrieval, and the experimental results prove that the proposed technique provided better results (Fig. 4).

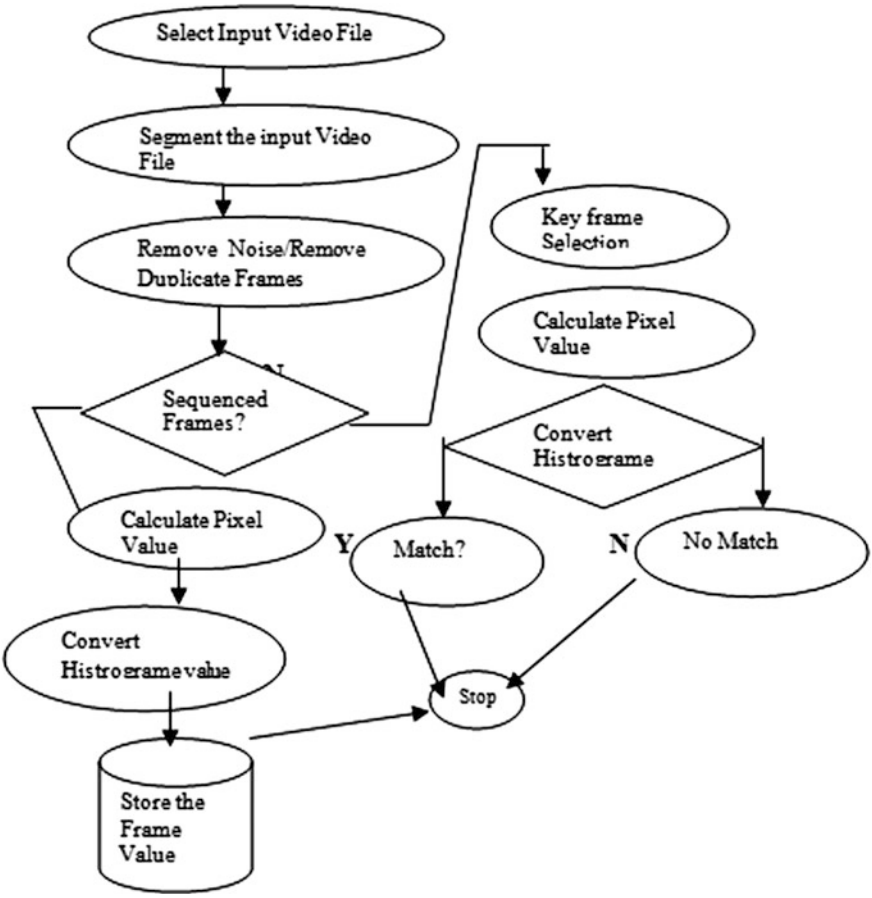


Fig. 4 Flowchart of overall video information retrieval process

5 Conclusion and Future Enhancement

This work brings out an efficient technique for video information retrieval. Because of the network development and its flexibility, the usage of image information has increased in recent years. This is because of the increasing demand for image information, and also, it provides very easy access to the video contents. Every day, the amount of these image files is increasing due to various factors like usage of mobile phones, twitter, facebook, YouTube and more. Creation and distribution of these files are very easy. Still, we suffer to manage and retrieve these contents. There are no efficient techniques or tools presently available. This paper brings the key frame image indexing technique for retrieving and efficiently indexing the video files. Results proved that the proposed method is more efficient.

5.1 Future Enhancement

The future enhancement of this technique by adding additional features will produce more accurate results.

6 Experimental Output

See Figs. 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14.

Fig. 5 Video segmentation

Fig. 6 Video shots (Frames)

Fig. 7 Duplication removal operations

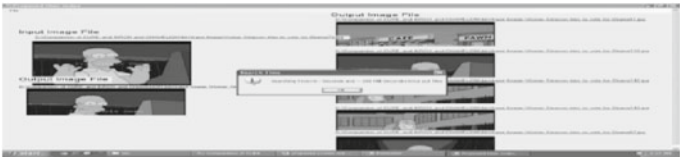
Fig. 8 Training the input frames

Fig. 9 Clustering the frames

Fig. 10 Key frame selection

Fig. 11 Output for selecting input key frame

Case 1: 1 input and 8 output files



Case 2: 1 input and 15 output files

Fig. 12 Image comparison based on user input image Case 1: 1 input and 8 output files. Case 2: 1 input and 15 output files

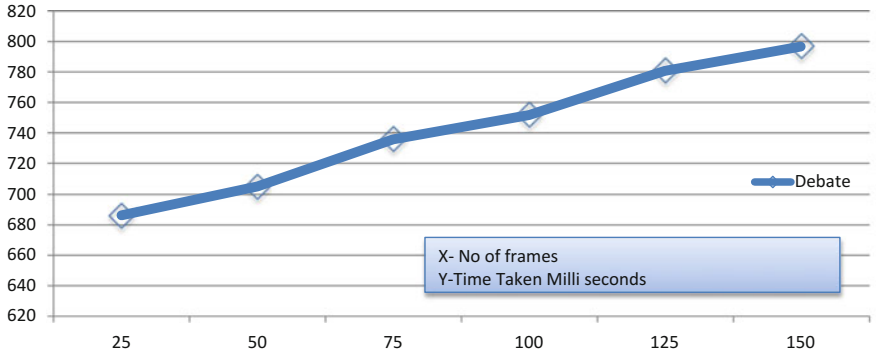


Fig. 13 Performance graph for news video file

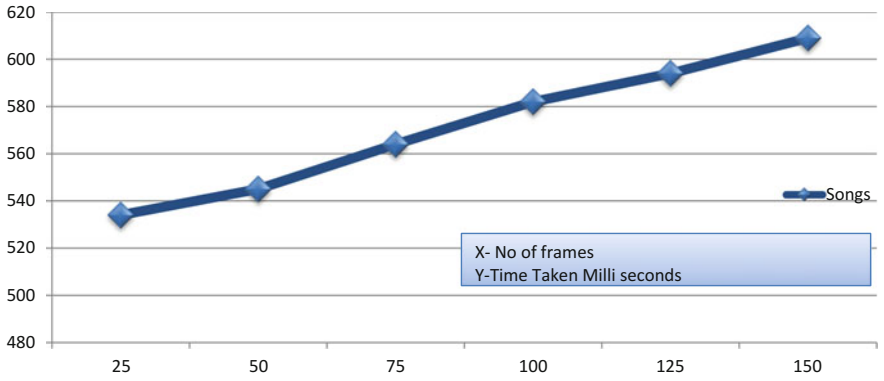


Fig. 14 Performance graph for song video file

References

1. D. Saravanan.: Effective Multimedia Content Retrieval. International Journal of Applied Environmental Sciences, Volume 10, Number 5, pp. 1771–17783, (2015).
2. Jamie Shotton, Alex Kipman, Andrew Fitzgibbon, Mark Finocchio,: Time Human Pose Recognition in Parts from Single Depth Images, Communications of the ACM, Volume 56 Issue 1, Pages 116–124, (January 2013).
3. H. Sundaram et. al., "Video mining with frequent item set configuaration", CIVR 2006, LNCS4071, Pages 360-369, 2006.
4. D. Saravanan, A. Ramesh Kumar, 2013 "Content Based Image Retrieval using Color Histogram", International journal of computer science and information technology (IJCSIT), Volume 4(2), 2013, Pages 242–245.
5. Ghorpade Krishna Bhimraj, Swapnalini Pattanaik, "An enhanced Technique Based on Content Based Image Retrieval and video Indexing" Proceedings of 19th IRF International Conference, 1st February 2015, Pune, India, ISBN: 978-93-84209-85-8, Pages 45–50.
6. D. Saravanan, V. Somasundaram 2014. Matrix Based Sequential Indexing Technique for Video Data Mining, Journal of Theoretical and Applied Information Technology 30th September 2014. Vol. 67 No. 3.
7. D. Saravanan, Performance Anlaysis 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