Academic City University College, Accra, Ghana
Society for Multidisciplinary & Advanced Research Techniques (SMART)
Trinity University, Lagos, Nigeria
SMART Scientific Projects & Research Consortium (SMART SPaRC)
Harmarth Global Educational Services
ICT University Foundations USA
IEEE Computer Society Nigeria Chapter

33rd ECOWAS iSTEAMS ETech Multidisciplinary Conference (ECOWAS-ETech)

Segment Based Indexing Technique for Video Data File

John-Bunya Nyuiemedi Kluste
School of Technology
Ghana Institute of Management & Public Administration
GreenHill, Accra Ghana
E-mail: jbklutse@gmail.com

ABSTRACT

Due to the advanced technology, a video is an efficient means of exchanging information in the format of displaying a short text message. Video capture is a simple process, but related video retrieval is challenging, necessitating the indexing of the videos. Retrieval is the process that used a user query to find a video. Depending on how the query result output system handled the query, the query may return an image or video in particular. In this project, we develop a segment-based indexing technique for video files. Here, video will be broken down into a hierarchy similar to movie storyboards. For instance, a multi-stage abstraction is used in a hierarchical based video search to help users logically locate particular video segments or frames. The reduced bandwidth and decreased delays of the video through the network of searching and reviewing are highlighted in this paper. Results from experiments support this.

Keywords: Segment, Indexing, Video, Data, Message, Security, Outputs

Proceedings Citation Format
John-Bunya Nyuiemedi Kluste (2022): Segment Based Indexing Technique for Video Data File. Proceedings of the 33rd ECOWAS iSTEAMS Emerging Technologies, Scientific, Business, Social Innovations & Cyber Space Ecosystem Multidisciplinary Conference. University of Ghana/Academic City University College, Ghana. 29th Sept – 1st Oct, 2022. Pp 81-88. www.isteams.net/ecowasetech2022. dx.doi.org/10.22624/AIMS-/ECOWASETECH2022P14

INTRODUCTION

The process of extracting knowledge from a large set of data is known as data mining. The development of Information Technology has generated large amount of databases and huge data in various areas. The research in databases and information technology has given rise to an approach to store and manipulate this precious data for further decision making. Data mining is a process of extraction of useful information and patterns from huge data (Ramageri, 2008). It is also called as knowledge discovery process, knowledge mining from data, knowledge extraction or data /pattern analysis. Data mining is a logical process that is used to search through large amount of data in order to find useful data.

The goal of this technique is to find patterns that were previously unknown. Once these patterns are found they can further be used to make certain decisions for development of their businesses. Data mining is a process of extracting previously unknown knowledge and detecting the interesting patterns from a massive set of data.

Thanks to the extensive use of information technology and the recent developments in multimedia systems, the amount of multimedia data available to users has increased exponentially. Video is an example of multimedia data as it contains several kinds of data such as text, image, meta-data, visual and audio. It is widely used in many major potential applications like security and surveillance, entertainment, medicine, education programs and sports (Velusamy, 2012). The objective of video data mining is to discover and describe interesting patterns from the huge amount of video data as it is one of the core problem areas of the data-mining research community (Saravanan, 2015).

Compared to the mining of other types of data, video data mining is still in its infancy. Video data is crucial to the field of video data mining because it is one of the many types of data that multimedia contains, including audio, video, image, text, and motion. Video data mining is the short name for the use of video data (Saravanan, 2016). Various documents can use data mining techniques. Video data can be easily collected and stored, but it can be difficult to extract information from video data. According to Saravanan, choosing the right key frames is crucial. Several techniques are used to automate the process, but they have two main drawbacks;

- 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).
- 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).

Saravanan states that the objectives of the study were;

- RGB value of frame is used to eliminate the redundant frame.
- Segment based indexing technique is used to indexing the key frames.
- Hierarchical clustering algorithm is used to clustering the frames

Existing System

The approaches now in use for segmenting, annotating, and categorizing video lectures largely concentrate on either low-level audio features or low-level visual information. Some of the works use advanced semantic techniques like writing, erasing, speaking, and instructor gestures as well as sketching, scrolling, and explaining. The current method used color moments and Decision Trees to categorize video frames as narrative or text-based (slide, online, or whiteboard).

Issues in Existing System

- Less prediction accuracy
- Increased time complexity
- Key frames are often not enough to represent information in a shot.

Proposed System

- RGB feature is used to remove redundant frames in the query video.
- Segment based Indexing technique is used to segment and indexing the frames.
- Hierarchical clustering mechanism is used to cluster the frames.

Advantage of Proposed System

- Can be easily automated
- Segments reveal temporal structure well (eg. In hierarchy)
- Supports queries by image similarity

Functional requirement

- Input: The video inquiry serves as this work's input.
- Behavior: This project behaves by converting the incoming video into a number of frames. Duplicate frames are gotten rid of using the frame extraction procedure. Segment the frames using a segment-based approach. Finally, the user finds the pertinent frame for the specified query image.
- Output: This project's output is a retrieved, pertinent image.

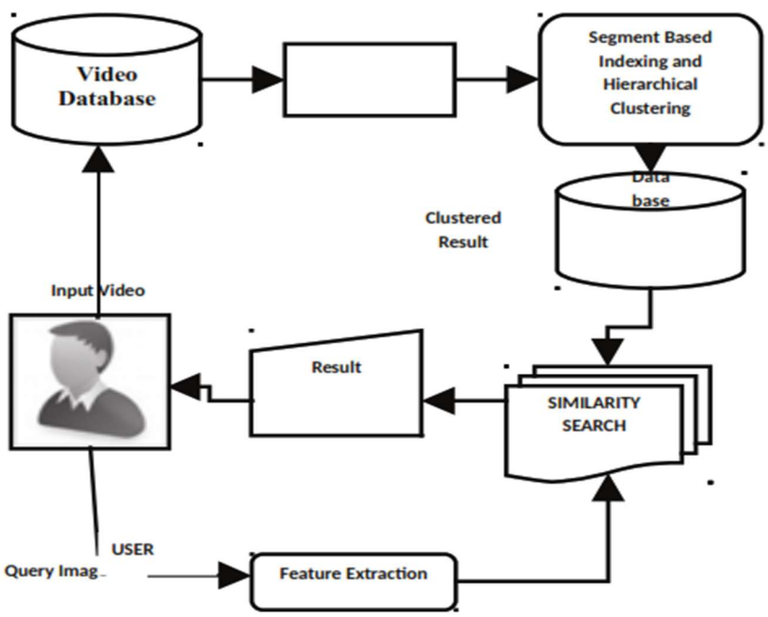

Fig 1 Proposed System Architecture Diagram

2. RELATED LITERATURE

According to Sariq et al., the increasing popularity of flexible and blended learning has led to non-scripted lecture videos being used as learning resources on a regular basis in most higher education institutions. The majority of the time, these videos are provided as Learning Objects. Non-scripted lecture videos can be automatically segmented and have their structure extracted by taking advantage of visual cues that are present in the video.

The identification of potential key frames, index points, key events, and pertinent meta-data using such visual cues is helpful for e-learning systems, video surrogates, and video skims (Sariq et al.). MUST-VIS provides a multi-modal algorithm for lecture segmentation based on audio/text and video, and annotates segments using keyword clouds, which provide immediate access to the information content, while considering the main speaker actions.

The MUST-VIS system for the Media-Mixer/VideoLectures.NET is shown in this study. Grand Challenge for Temporal Segmentation and Annotation (Chidansh et al.). Once more, Chidansh details the creation and assessment of the new NoteVideo and NoteVideo+ systems for identifying the conceptual "objects" of a blackboard-based video and using that information to create a summarized image of the video and use it as an in-scene navigation interface that enables users to jump directly to the video frame where that object first appeared rather than navigating it linearly through time (Toni-Jan et al.).

Experimental Setup

The subsequent procedures make up the experimental setup.

- Video Preprocessing: In the image pre-processing, the camera's input image, which can be either a video file or a digital image, is used. In video files, the images are divided into frames that often have noise and other impurities but may be cleaned up during preprocessing. Finally, the user receives the required image based on their request.
- Training of Images: Following the extraction of image features like texture and matrix conversion, the database's pixel values are trained by labeling the image features. A matrix with M rows and N columns will be created by converting the intensity at each point to its x, y, and RGB value. After that, the database is labeled with the images. As a result, retrieving it from the database is simple. The attributes of the image serve as the labeling. The picture is now kept in the database.
- Frame Extraction: The extracted frames are processed using the converted frames as an input. Each and every frame in the video's frame extraction process had its RGB value calculated. Algorithm for frame Extraction:
  Step 1: First, the image is given as the input to from the camera.
  Step 2: Initially, this image is a raw image where it contains noise.
  Step 3: Then, the Features like Texture, Color and Shape are extracted by the RGB values.
  Step 4: Features values and database values are matched. The content of the image is also retrieved

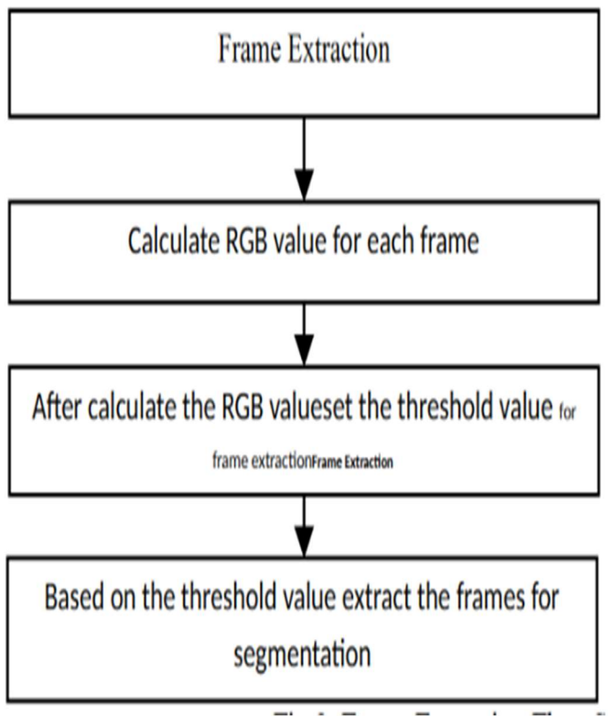

Fig 2 Frame Extraction Flow Steps

- Segment Based Indexing: The specifics of the segment-based indexing process are described in this module. A video may contain stories when indexing is done using segments. Each of the stories is made up of a number of scenes, each of which is divided into shots, each of which is made up of a number of individual frames.

- As a result, we have defined a frame in this indexing framework as a single image or picture, a shot as a succession of frames with similar characteristics, a scene as a succession of shots that correspond to a semantic content, and a story as a succession of scenes that reveals a single semantic story. Using a different version of this hierarchy, it is shown that videos typically serve one or more purposes, such as entertainment and information. With the aid of hierarchical clustering, the indexed frames are further grouped. For various video files, experiments are carried out using different video frames, such as 75,150.

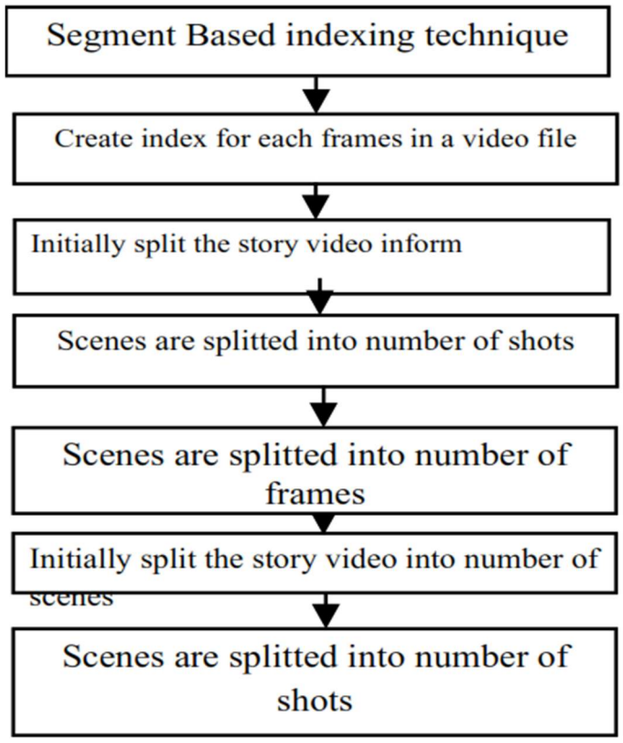

Fig 3. Segment Based Indexing Technique process flow diagram

- Query Retrieval process: The user provides the storage server with the query image in the query retrieval module. Query image features should be extracted. The clustered result is then compared to the feature extracted frame. Finally, the user sees the matched frames.

3. FINDINGS

In the article Saravanan created aa indexing for video file by using segment based indexing technique. Saravanan divided into a hierarchy which is in storyboards of film making. The study brings out the reduced bandwidth and reduced delays the video through the network of searching and reviewing.

4. RECOMMENDATIONS FOR PRACTICES AND DESIGN

The author suggested a concept for automated video search based on frames or images using segment-based indexing. It is organized hierarchically and clearly. In order to retrieve pertinent images from the database, the user can also use the image as a query. By avoiding repeated frames overall, the use of the RGB feature speeds up search times. Indexing, which uses segment frames, also speeds up searches.

5. EXPERIMENTAL OUTCOMES

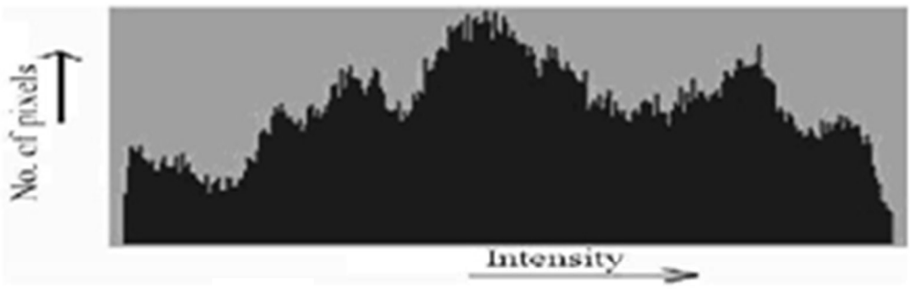

Fig 4: ig 5.Histogram Generation

Fig 5: 6. Select Video

Fig 6: . Frame Conversion process

6. CONCLUSION

An effective framework for image retrieval from the provided video was proposed by Saravanan. The RGB feature and segment based indexing mechanisms are the two key strategies for effective retrieval. The given video has redundant frames removed using the RGB feature. The method for indexing the individual frames is segment-based. The video frames are finally clustered by the hierarchical clustering mechanism. The query image is processed, compared to the corresponding image in the database, and the corresponding result is displayed when the user inputs the query image. We would like to emphasize that the framework we have described is very effective and shows the outcome quickly.

REFERENCES

8. Chidansh Bhatt, Andrei Popescu-Belis, Maryam Habibi, Sandy Ingram, Stefano Masneri, Fergus McInnes, Nikolaos Pappas, Oliver Schreer" Multi-factor Segmentation for Topic Visualization and Recommendation: the MUST-VIS System" Proceedings of the 21st ACM international conference on Multimedia, Pages 365-368.
9. Ramageri B. M. (2008). Data Mining Techniques and Applications. Indian Journal of Computer Science and Engineering Vol. 1 No. 4 301-305
10. Saravanan (2015). Text information Reterival using Data mining Clustering Technique" International Journal of Applied Engg. Research,10(3), 7865-7873.
11. Saravanan D. (2016). Segment Based Indexing Technique For Video Data File. Procedia Computer Science, 87, 12 – 17
12. Shariq Imran, Alejandro Moreno, Faouzi Alaya Cheikh. Exploiting Visual Cues in Non-Scripted Lecture Videos for Multi modal Action Recognition. Eighth International Conference on Signal Image Technology and Internet Based Systems, Pages 8-14
13. Toni-Jan Keith Monserrat, Shengdong Zhao, Kevin McGee, AnshulVikram Pandey" NoteVideo: Facilitating Navigation of Blackboardstyle Lecture Videos" Proceedings of the SIGCHI Conference on Human Factors in Computing Systems
14. Velusamy V. (2012). A Study on Video Data Mining. International Journal of Multimedia Information Retrieval 1(3)