IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 3, Issue 5, Oct -Nov, 2015
**ISSN: 2320 – 8791 (Impact Factor: 2.317)**
**www.ijreat.org**

# Video Structure Retrieval Using Video Extraction Analyzer

**Noraida Haji Ali[1] , Fadilah Harun[2], Noor Maizura Mohamad Nor[3]**

**[1]School of Informatics and Applied Mathematics,**
**Universiti Malaysia Terengganu, Terengganu, Malaysia**

**[2]School of Informatics and Applied Mathematics,**
**Universiti Malaysia Terengganu, Terengganu, Malaysia**

**[3]School of Informatics and Applied Mathematics,**
**Universiti Malaysia Terengganu, Terengganu, Malaysia**

**Abstract**

Video Extraction System Analyzer (*VidExA*) has been developed in order to solve one of the problems encountered in the difficulty in determining the contents of the existing data in a video structure. The video structure is poor and inconsistent, but a video still have own basic structure. The basic structure of the video is a hierarchical structure consisting of a video program, scene, shot and key frames. *VidExA* is aimed to help the user understand the basic structure of video data that is complicated and complex. In general, *VidExA* is developed to extract the data of the structure contained in a video that was input by the user. Normally, the user is quite difficult to understand and define the data that exists in a video structure. There are two (2) modules of the system include the following: (1) Extraction: A video is extracted to get a basic structure and their relation between scene, shot and keyframe, (2) Analysis: This system will generate an analysis of relation structure had in a video. Based on *VidExA* operations, it can help the user know the originality of the video. Besides that, we can determine each relation between the structures in the video.

***Keywords: Multimedia Database, Video Database, Video, Video Basic Structure.***

## 1. Introduction

A multimedia database system it is an efficient way to handle the huge of existing data like, texts, image, audio and video [1]. Multimedia database system is a system structure and constructs information for retrieval of multimedia content [2]. A maintaining and managing large collections of media objects it's used multimedia database requirements as a tool. It supports the various multimedia data. Multimedia databases contain different data types such as text, images, graphics, animation, audio and video. Special methods are required for huge storage, access, indexing and retrieval of different data types in multimedia database. Video is one of the challenging and have a lot of issues associated with it [3]. This is because the video contains information on all other types of media into a single data stream [4]. This is an issue occurs when develops a video increase when one changes from images to image sequences, or video clips. Also, each video program has own rules and formats. The development of *VidExA* is to solve the issues to retrieve the structure of video.

*VidExA* developed to identify a framework that has been developed through the application. *VidExA* purpose of this module is to solve one of the problems encountered in the difficulty in determining the contents of the existing structure in a video. *VidExA* is expected to help the user to understand the basic structure of video content that is complicated and complex. In general, the *VidExA* was developed to extract the contents of the structure contained in a video that was input by the user. Normally, the user is difficult to understand and define the basic structure of the content appearing in a video.

This paper is organized as follows: The following section discusses related works had done before. In Section III, explained about the framework design of *VidExA*, input, process and output. In Section IV, elaborate the experiments and result. Finally, in Section V our conclusions are stated.

## 2. Related Works

Video database has been widely used in various fields. However, video database still has some issues that need to be resolved. Among the main problems faced by the current database systems is the shortage of a common way to represent complex queries [5]. This is due to the gap between the way consumers think and query language that is used in most systems. Multimedia data manipulation is not as easy as using a conventional database. The use of database can display video clearly but the main problem is

to achieve the video data structure itself. This problem occurs when we want to ensure that content data in the database must be commensurate with the content specified in the query. Each query responses mentioned in media data, it must have a technique that attempts to analyze the data content to get different semantics associated with media data [5]. Sometimes, the structure of video mutually compatible and linked between them cannot be achieved easily. Furthermore, there is no approach that supports video and display the schema object.

An important issue in the development of a video is formal modeling techniques for achieving video data. These models have a high capacity for abstract multimedia information and define the semantics [6]. By using spatial modeling constrained by the size of each frame and the frame layout. Constraints must be taken into account by giving the value of each coordinate x and y for each corner of the frame by using constraints or state differences in the relative position of the frame in each performance group [7]. Structuring of video data using several different techniques, such as temporal and spatial relations for the design of a database [8]. All these techniques are focused on design a good database. However, the structuring of video data have some limitations, such as the size and model complex objects in a variety of conditions for conducting the process of indexing, search and retrieval [6]. Therefore, the investigation conducted by researcher [9] attempted to resolve issues related to multimedia data using temporal specifications. Temporal relationship is to determine the length of the relationship between multimedia objects. It uses the power of temporal Petri net to model the temporal relationship and interactive. However, the temporal and spatial approach is still difficult to achieve a video data and determine the relationship of each data structure. Therefore, a mechanism should be established to determine the relationship between the structure of video data and content structure is reached.

## 3. Framework Design of *VidExA*

*VidExA* purpose is to extract and analyzes the basic structure of the video data. It was developed to overcome the problem of access video data and the difficulty in determining the relationship. *VidExA* framework consists of several components, Input, Process and Output. Each component in this framework based on fundamental components of the framework. Input is an important component in the framework as the main source for processing. In this framework, it involves the input video to be analyzed in terms of its structure. Process is used to describe the process of the implementation of this module. This process was developed based on the needs identified. Processes involved in the development *VidExA* extraction and analysis is the process of video data structure. Output also is the end result after the information is processed in engine components. Output of this framework is the list of scenes, shots, and keyframe within the structure of the video. Moreover, from the output can determine the relationships found in the structure of the video data. Authenticity of a video can also be proved by the content of the resulting structure.

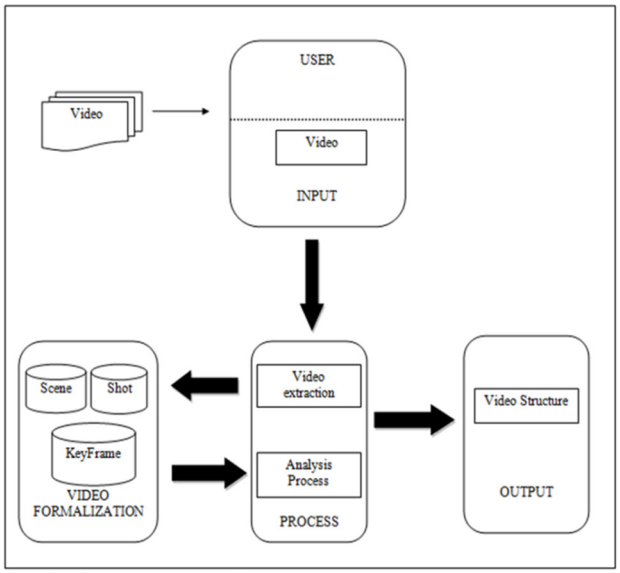


Fig. 1 Basic Framework of *VidExA*

Based on Figure 1, there are three main sections described in detail. Briefly, it consists of input, process and output. Input consists of video input by the user. Meanwhile, the process consists of two processes, namely, the extraction process video and analysis process. Part of structure video output is also generated.

- Input

In *VidExA*, video clip used as input files. Video input may consist of various original formats such as AVI, MP4, MP3, 3GP and so on. The structural characteristics of video that has been extracts from the basic structure, namely, scenes, shots and keyframes.

- Process

The second component is the process. This component consists of two processes involved in the development of video tools, namely the extraction and analysis process. Each component plays a different role. The extraction process will be generated to make the search hierarchical structure for each scene, shot and keyframes found in both video inputs involved. In the extraction process, there are three sub processes to be implemented, namely, the comparison between the framework for determining the number of shots you have, check the percentage of similarity

framework for producing scene and extraction of contents of the video data structure. Meanwhile, the analysis process is a process of analyzing the structure that has been extracted. The analysis process will make access to the structures that have been extracted for analysis every existing relationships in a video input. Any results that have been obtained will be compared between the original video with a video that has been modified to ensure that the basic structure of the actual content of a video clip.

- Output

The main output is required in this study is a list of structures in the extract when video is included. List structure can describe a relationship in the video. Output is generated; it can determine the authenticity of a video clip to be added depends on the structure of the list of video data that has been extracted. Each structure can be determined through the process of extraction and analysis process can be carried out to determine the authenticity of a video.

### 3.1 The Extraction Process

Video extraction process is carried out to find a list of hierarchical structures. Extraction algorithm using this structure must comply with certain rules in order to achieve the desired structure list. Inputted video clips by using search methods when input is entered manually into the system. Inputted video consists of a variety of formats and sizes, suitable for a video. Each video will be inserted through the conversion process in the form of a frame to block sequentially using *FrameGrabber* functions to determine the number of frames available. Each frame will be converted into the form of a block of image and stored in the format Portable Network Graphic (PNG). Converted resulted will be used to determine the percentage difference for each existing frame. Each frame will be tested sequentially. The first frame will be tested with the percentage of the second frame. If the percentage of the frame is> 40 then, the keyframe can be identified based on the differences that occur. Number of percentage differences that occur is to determine the number of scene contained in video data structure.

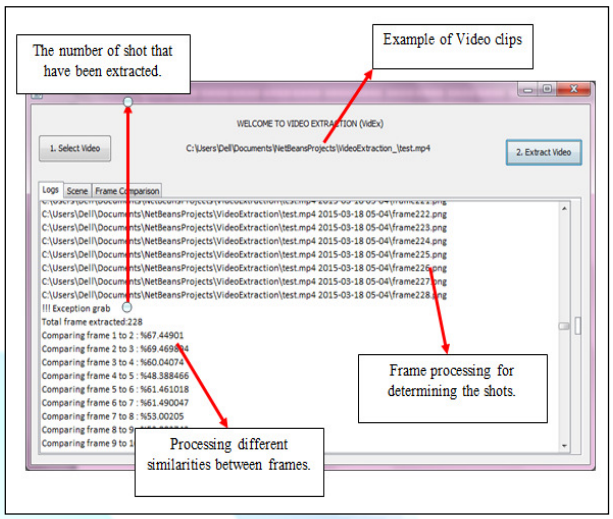


Fig. 2 Extraction process

Figure 2 showed the display for the performance of the extraction process. In this section, the video clip through frame processing for determining the number of shots you have. Then, the number of frames that have been extracted display. The next process of determines the percentage difference between the processing of each frame to determine the number of scenes that are inherent in a video clip.

### 3.2 The Analysis Process

*VidExA* analysis process is a process of retrieval structures that have been extracted for analysis every relationship that exists between the structures of video data. This process involves several important steps. The steps in this process must be followed to ensure that the process works well analysis.

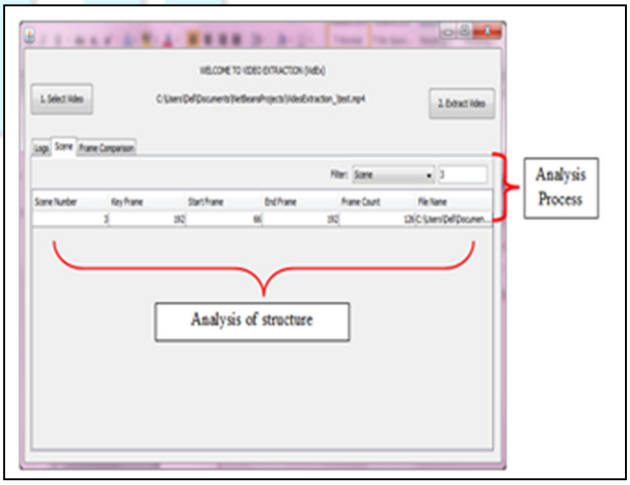

Fig. 3 Analysis Process

This section makes a search based on the structure of the selected option. Users can select a search made it, scene, shots or keyframe. The results of this search will be displayed in the text area provided as in Figure 3 above. Based on Figure 3, the search field shows an example of a search for scene. Scene '5' fed as a search. List of associations related to the 5th scene are listed in the textbox.

## 4. Experiment and Results

Testing and evaluation process is one of the most important phases in software development. The original objective was to ensure that software development can be achieved. In this section, each step described for a test run. The experimental design was developed for this purpose, taking into account elements in it. Next, this section also describes in detail the results obtained from tests which have been carried out. The main objective of this test is performed to test the effectiveness of *VidExA* which acts as an analyzer for modeling object-oriented structure. This equipment is expected to perform the extraction of the basic structure of the video. These tests involve multiple sets of data were included as input to match.

### 4.1 Testing of *VidExA*

Testing the basic structure of the extraction process of a video clip is done and the resulted list on Table 1. Methods of extraction process used to develop *VidExA*. Some video clips tested to get all the basic hierarchical structure of the video, scenes, shots and keyframes. The original and trimming video tested to get the basic data video structure.

Table 1. Comparison Of The Structure After Trimming Process

| Video | Scene | Shot | KeyFrame |
|---|---|---|---|
| Video A | 258 | 3007 | 258 |
| Video A edit | 217 | 3565 | 217 |
| Video B | 232 | 3115 | 232 |
| Video B edit | 160 | 3745 | 160 |

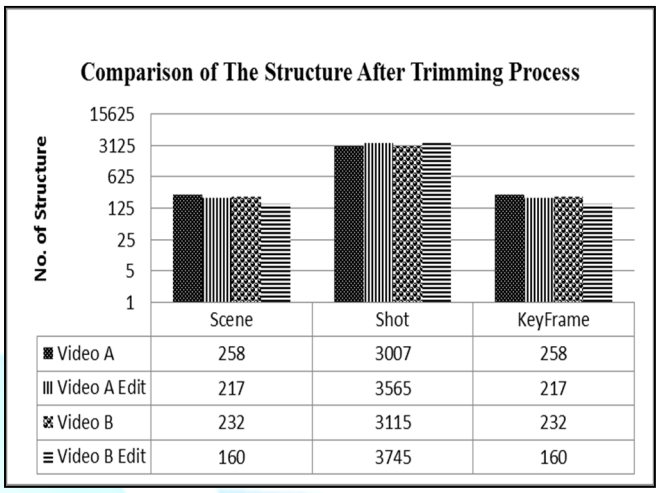


Fig. 4 Graph comparison

Figure 4 showed a comparison of the number of structures contained in the video A and B. Input for this search is a video clip of the same. The first video A, is an original video and was extracted to determine the list of the existing structure. Next, the video has been edited and re-extract structure for the result sought. After the two video is extracted, found the number for the three different structures. Thus, the authenticity of a video can be sure with content structure which results from the extraction.

As in Table 1, a video clip has been tested to prove the extraction process to extract the video in the hierarchical structure. The used of this application is essential to detect any changes made illegally on an original video clip. This can be evidenced by the number of scene, shot and keyframes. It can be compared with the number of the existing framework. For example, in original video had 3007 of shot but in video that had trims (modified version of the original video) contains 3565 the number of shot. Authenticity of a video clip can be determined by making a comparison based on the number of frames and structures that have been extracted from the original video clip. If there are any changes in it will affect the original structure featured in the video. Therefore, it can identify the authenticity of the video.

### 4.2 Evaluation of *VidExA*

*VidExA* performance compared with other video editing tools such as the Table 2.

Table 2. Comparison Between System Performances

| Tools Name / Features | *VidExA* | Virtual Dub | AV Cutty | HandyS aw DS |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| **Video file format** | **All format** | AVI | AVI | AVI |
| **Image format (sequences of frames)** | **JPG** | BMP | BMP | BMP |
| **Scene extraction** | **Yes** | No | Yes (scene determined by the user) | Yes (scene determined by optical and processing date) |
| **Shot extraction** | **Yes** | Yes | Yes | Yes |
| **Keyframe extraction** | **Yes** | No | No | No |
| **Determine originality of video** | **Yes** | No | No | No |

According to Table 2, it compares the performance of *VidExA* with others tools like AVCutty, VirtualDub and HandySaw DS. Some features have been compared in this study. The first features are a video file format used in this study. *VidExA* capable of supporting a wide range of video file formats such as MP3, MP4, 3GP and AVI. Other tools only support AVI video format only. The second features, image file format that was extracted. *VidExA* save shaped JPG image file while the three others store image files in bitmap format (BMP). Images are stored using JPG file format has a much larger scale and more clearly than the pixel BMP file format. VirtualDub is only able to extract a shot only. Whereas because AVCutty able to extract two types of structures only, scene and shot only. It's because, of the scene generated by AVCutty specified by the user. Users must first decide which scenes you want extracted and this was done manually. HandySaw DS is also capable extracting scene and shot only. Each scene must selected manually by the user based on two criteria; sight for optical (real view) or the date an action was performed. Next, *VidExA* able to extract the scene, shot and keyframes in one process and it can save you time in the extraction process. In addition, *VidExA* also not complicate the user with a choice of manually because it involves one process. With the use of *VidExA*, it can produce an analysis of the authenticity of the video are tested. It can be determined based on the list of successful structure extracted using this *VidExA*.

## 4. Conclusion

The main objective of this study is the emphasis on the extraction of the hierarchical structure of a video clip. The framework is a basic conceptual structure that aims to solve or identify any problems or errors and usually consists of a set of tools, materials or components. Similarly, the development *VidExA* framework developed for the purpose of resolving the difficulties of access for the data structure of a video clip. In a video clip, it contains a basic structure, scene, shot and keyframes. *VidExA* consist two main processes, namely the process of extraction and analysis process. Through the extraction process, the list data structure to be extracted in detail and will list structure. While the analysis process aims to achieve re-list structure generated by the search made. The results of this analysis process can determine the existence of any structure made by the search and the relationship of each structure can be determined.

**Acknowledgments**

This research was supported by a Fundamental Research Grant Scheme (FRGS), Vot: 59298.

## References

[1] D. Ainscough. "*An Introduction to Multimedia Database System*", 2010 [cited 2014 14 June]; Available from: http://blueadmiral.com/notes/computing/cdd/cdd_introbd.html.

[2] S.K. Jalal, "Multimedia Database: Content and Structure", Workshop on Multimedia and Internet Technologies, 2001.

[3] A. Lilly Suriani, M. Ali, I. Hamidah, and A. Fatimah, "Video Data Modelling To Support Hybrid Query", International Journal of Computer Science and Network Security, 2007. 7(9): pp. 53.

[4] F.C. Jian, L. Hui, and W.J. Tao. "Video Hierarchical Structure Mining", in International Conference on Communications, Circuits and Systems Proceedings. 2006.

[5] D.A. Keim, and V. Lum. "Visual query specification in a multimedia database system", in Proceedings of the 3rd Conference on Visualization'92. 1992: IEEE Computer Society Press.

[6] A. Ghafoor. "Multimedia Support Infrastructure (MSI)", National Science Foundation 2006 [cited 2013 20 November]; Available from: http://www.cs.purdue.edu/.

[7] M. Koprulu, N.K. Cicekli, and A. Yazici, "Spatio-temporal querying in video databases", Information Sciences, 2004. 160(1–4): pp. 131-152.

[8] A. Duda, and C. Keramane. "Structured temporal Composition of Multimedia Data", in iw-mmdbms. 1995.
[9] C. Djeraba, and H. Briand, "Temporal and interactive relations in a multimedia database system", in Multimedia Applications, Services and Techniques (ECMAST'97). 1997, Springer. pp. 457-473.