

# Multimedia Retrieval Using Web Mining

Manda Jaya Sindhu, Y. Madhavi Latha, V. Samson Deva Kumar, Suresh Angadi

**Abstract:** Multimedia is a media that uses a combination of different forms which includes audio, text, video, images, etc...So, for the purpose of better understanding of user it's better to retrieve Multimedia data rather than text. Web mining is retrieving the content using data mining techniques from World Wide Web. Multimedia retrieval using web mining deals with the retrieval of useful patterns based on the user's search requirements. The images and video clips were retrieved by crawling the World Wide Web. The main theme of this paper is to know the mining details of various types of data in form of multimedia retrieval where all kinds of data or information is retrieved based on the search pattern of the user.

**Index terms:** Multimedia, Web Mining, World Wide Web, Data Mining.

## I. INTRODUCTION

These days World Wide Web is the major source for retrieving the information. The entire process of applying the techniques for extracting the data from World Wide Web is Web mining. A Multimedia database system includes a multimedia database management system (MMDBMS), which manages and also provides support for storing, manipulating and retrieving multimedia data from a multimedia database, a large collection of multimedia objects, such as image, video, audio and hypertext data, is discussed in [1]. Multimedia gives a lot of information on each entity but not the same information for each entity. The difference between Multimedia Mining and Structured data mining is the sequence or time element. Multimedia often captures an entity changing over time. Video and audio are clearly ordered, and even text has little meaning without sequence. Time series mining analyzes the change to one or more values over time. Multimedia is more complex-as the sequence progresses, the concept being may change as well. Understanding and representing changes in the mining process is necessary to mine multimedia data [2].

## II. TYPES OF WEB MINING

Web mining is categorized into three types

### A. Web Content Mining:

Web Content Mining is the way of retrieving useful information from Web. The retrieved information may contain text, images, audio and video.

### B. Web Structure Mining:

Web Structure Mining is the way of discovering the structure information from the Web, which is further divided into two types based on the structure information.

**1. Hyperlinks:** Hyperlink connects a Web page either in the same Webpage or on different Web page.

**2. Document Structure:** Web page can also be arranged in a tree-structured format, based on HTML and XML tags within the page.

### C. Web Usage mining:

Web Usage Mining is the application of Data mining techniques in order to obtain the useful patterns from the Web, which is a huge repository of different patterns.

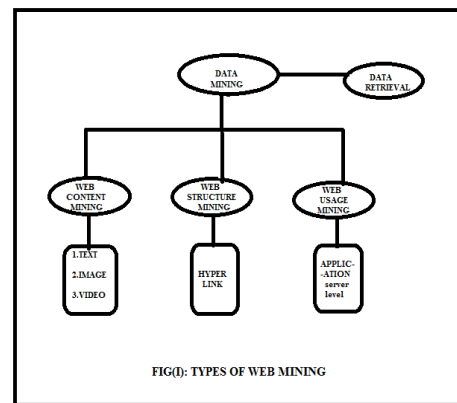


Fig 1: Types of Web Mining

## III. MULTIMEDIA DATABASE MINING:

Multimedia is the major focus for many researchers around the world. While retrieving Multimedia data from the database the core process followed is Knowledge Discovery Process. A Multimedia data mining system prototype, Multimedia Miner-includes the construction of a multimedia data cube which facilitates multiple dimensional analyses of multimedia data, primarily based on visual content, and the mining of multiple kinds of knowledge, including summarization, comparison, classification, association and clustering [3].

In Multimedia documents, Knowledge discovery deals with non-structured information. In general, the multimedia files from a database must be first pre-processed to improve their quality followed by feature extraction. With the help of generated features, information models can be devised using data mining techniques to discover significant patterns as shown in Fig 2 Multimedia data mining refers to pattern discovery, rule extraction and knowledge acquisition from multimedia database [4].

Revised Manuscript Received on 30 March 2013.

\* Correspondence Author

Manda Jaya Sindhu, B. Tech Student, Department of Electronics and Computers, K L University, Guntur (Andhra Pradesh), India.

Y. Madhavi Latha, Asst. professor, Department of Electronics and Computers, K L University, Guntur (Andhra Pradesh), India

V. Samson Deva Kumar, Project Manager, SOUTH CENTRAL RAILWAY WWO, S/w Training and Development Center, Vijayawada (Andhra Pradesh), India.

Suresh Angadi, Asst. Professor, Department of Electronics and Communications, K L University, Guntur (Andhra Pradesh) India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

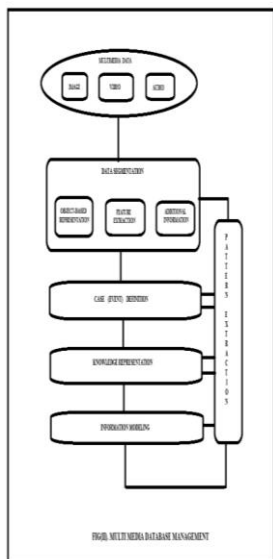


Fig 2: Multimedia Database

IV. TYPES OF MULTIMEDIA MINING:

Different types of Multimedia Mining are discussed below

A. Text Mining:

Text Mining is a type of mining where data is extracted only in text format from the data bases. Text mining can be said as an extension of Data mining. Here the data can be retrieved by specifying the attributes or key words. Text mining or text data mining, the process of finding useful or interesting patterns, models, directions, trends, or role from unstructured text, is used to describe the application of data mining techniques to automated discovery of knowledge from text [5].

B. Image Mining:

Image mining is the concept used to detect unusual patterns and extract implicit and useful data from images stored in the large data bases. Therefore, we can say that image mining deals with making associations between different images from large image databases shown in Fig 3. Image mining is used in variety of fields like medical diagnosis, space research, remote sensing, agriculture, industries and also handling hyper spectral images. Images include maps, geological structures, and biological structures and even in the educational field, explained in [6].

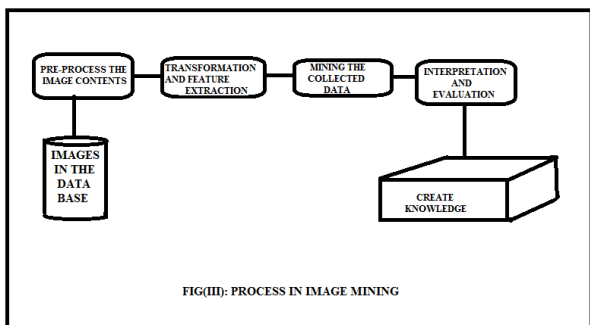


Fig 3: Process In Image Mining

C. Video Mining:

Mining of video data is complicated than mining an image data. Video is a collection of moving images like animation. There are three types of videos in video mining.

1. The produced (includes movies, news videos and dramas).

2. The Raw (includes traffic videos, surveillance videos etc...).

3. The Medical Video (includes ultra sound videos including echocardiogram etc...)

The Fig 4 given below clearly represents how the video mining takes place in a multimedia retrieval using web mining concept.

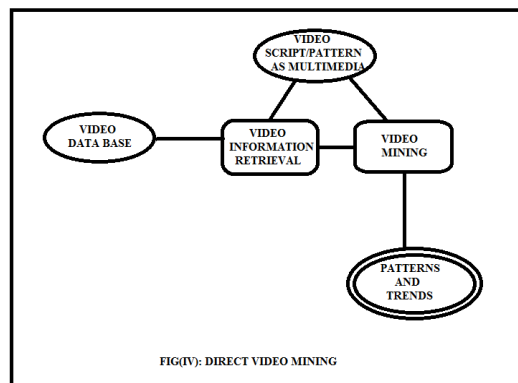


Fig 4: Direct Video Mining

D. Audio Mining:

Audio is a continuous media type as video the techniques used in audio are similar to the video data extraction. Audio can be in the form of radio, speech, etc... To mine audio data, first it has to be converted into text using speech transcription techniques. Audio data can also be mined directly by using audio information techniques and then mining selects the audio data. Audio mining is very simple in designing when compared to video mining.

V. PROCESS OF MULTIMEDIA MINING

The process of applying multimedia mining in order to retrieve different types of data is represented in Fig 5. Data collection is the first and foremost point of a learning system, as the quality of raw data is the factor which determines the overall achievable performance. The main aim of data pre-processing is to discover the important patterns from the raw data, which includes the concepts of data cleaning, normalization, transformation, feature selection etc...

Learning can be of straightforward, if informative features can be identified at pre-processing stage. Detailed procedure depends highly on the nature of raw data and problem's domain. The product of data pre-processing is the training set. Given a training set, a learning model has to be chosen to learn from it and make multimedia mining model more iterative. Higher complexity found on compared data mining with multimedia mining: a) the huge volume of data, b) the variability and heterogeneity of the multimedia data (e.g. diversity of sensors, time or conditions of acquisition etc) and c) the multimedia content's meaning is subjective. Application and system of multimedia data mining based on process discussed is surveyed in [8].

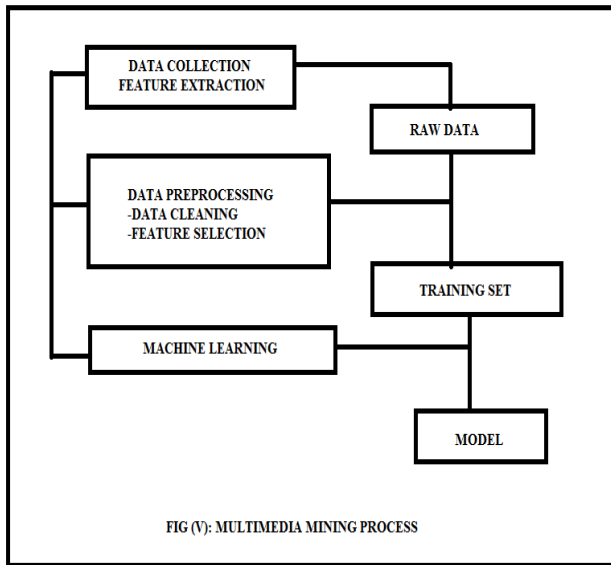


Fig 5: Multimedia Mining Process

**VI. APPLICATIONS OF MULTIMEDIA MINING:**

The video and audio data mining can be found in the Mining Cinematic Knowledge project [9], which created a movie mining system by examining the suitability of existing concepts in data mining to multimedia.

Moreover, the analysis and mining of traffic video sequences in order to discover information (Such as vehicle identification, traffic flow, and the spatial-temporal relations of the vehicles at intersections) provide an economic approach for daily traffic operations. There are some multimedia data mining frameworks [10] for traffic monitoring systems. Furthermore, various methods for the detection of faces in images and image sequences are reported in [11]. Detection of generic sport video documents seems almost impossible due to the large variety in sports.

**VII. CONCLUSION**

The Multimedia mining, knowledge extraction plays crucial role in multimedia knowledge discovery. The paper was discussed in the use of Multimedia database management systems and also mining of different kinds of Multimedia data. Process of Multimedia mining is also discussed. Some of the issues in Multimedia mining are too much of data is lost when the sequence of multimedia is ignored. But in audio and video mining a basic problem rises which is a combination of information across multiple media. Finally web mining is an integration of heterogeneous information sources.

**ACKNOWLEDGEMENTS**

I take this opportunity to remember and acknowledge the co-operation, good will and support both moral and technical extended by several individuals out of which the thought of making this paper had evolved. So, I am greatly elated and thankful to our guides who had supported.

**REFERENCES**

1. Yoshitaka A. and Ichikawa T., "A Survey on Content-based Retrieval for Multimedia Databases". IEEE Transaction on Knowledge & Data Engineering, Vol 11, 1999, pp. 81-93.
2. Bhavani Thuraisingham, "Managing and Mining Multimedia Databases" at International Journal on Artificial Intelligence Tools Vol. 13, No. 3 (2004) 739-759

3. Osmar R. Za\_ane Jiawei Han Ze-Nian Li Sonny H. Chee Jenny Y. Chiang, "MultiMediaMiner: A System Prototype for MultiMedia Data Mining," Intelligent Database Systems Research Laboratory and Vision and Media Laboratory report, 2009.
4. Dianhui Wang, Yong-Soo Kim, Seok Cheon Park, Chul Soo Lee and Yoon Kyung Han, "Learning Based Neural Similarity Metrics for Multimedia Data Mining" Soft Computing, Volume 11, Number 4, February 2007, pp. 335- 340.
5. Chakrabarti S. " Mining the web: Analysis of Hypertext and Semi Structured Data", morgan Kaufmann, San Francisco, CA.
6. C. Lakshmi Devasena et al. / International Journal on Computer Science and Engineering (IJCSE), ISSN : 0975-3397 Vol. 3 No. 3 Mar 2011 1155-1167.
7. Rosenfeld A., D. Doermann, D. DeMenthon, Eds., Video Mining, Kluwer, 2003.
8. Pravin M. kamde, Dr. siddu, P. Algur, "A Survey on web mining Multimedia"/The International Journal Of Multimedia & Its Applications (IJMA).
9. Wijesekera D. and D. Barbara, "Mining cinematic knowledge: Work in progress", in Proc. Of International Workshop on Multimedia Data Mining (MDM/KDD'2000), Boston, pp. 98-103.
10. Dailey D., F. Cathey, and S. Pumrin, "An algorithm to estimate mean traffic speed using uncalibrated cameras". IEEE Transactions on Intelligent Transportation Systems, Vol. 1, 2000, pp. 98-107.
11. Snoek C.G.M. and M. Worring, Multimodal Video Indexing: "A Review of the State-of-the-art Multimedia Tools and Applications", Vol. 25, No. 1, 2005, pp. 5-35.

**AUTHOR PROFILE**



**Manda Jaya Sindhu** pursuing B.Tech (IV/IV) in Electronics and Computers branch at K L University.

**Y. Madhavi Latha** is presently working as Asst.professor in department of electronics and computers at K L University



**V.Samson Deva Kumar** is the ProjectManager in SOUTHCENTRALRAILWAY WWO at S/w training and development center in vijayawada region.



**Suresh Angadi** is presently working as Asst. professor in department of electronics and communications at KL University. He received his B. Tech degree in electronics and communication in G.V.P College of Engineering, vizag. 2007 And completed M.Tech in Maulana Azad National Institute of Technology (MACT) in 2009, Bhopal. He has published SEVEN international journals of repute.

