

Survey on Analysis of Various Techniques for Multimedia Data Mining

Priyanka A. Wankhade¹, Prof. Avinash P. Wadhe²

¹Research Scholar (M.E), CSE, G. H Raisoni college of engineering and Management, Amravati

²Faculty, CSE, G. H Raisoni college of engineering and Management, Amravati, Email- Wankhade_priyanka.ghrcemamecse@raisoni.net

ABSTRACT – Data mining is an art which is used and applied discipline for grown out the various statistical pattern recognition, learning machine, and artificial intelligence which is combined with business decision making to optimize and enhance it. Initially the techniques of Data mining have been applied to the already structured data from the database. The highly usage of computers makes data mining affordable for small companies but on the other hand the invention of cheap massive memory and digital recording electronic devices allowed the misuse of the private sector such as corporate, governmental and private documents. For ex-e-mail message from customer, recording from telephonic conversation between customers and operators. To handle such condition multimedia data mining is available. The aim of multimedia data mining is to process media data alone or combination with other data for finding patterns useful for business.

Keyword-data mining, multimedia, text mining, image mining, audio mining, video mining.

INTRODUCTION-Multimedia data mining is used to exploration of audio, video, image and text data together by automatic or semi-automatic means in order to discover meaningful patterns and rules. When all the needed data get collected, computer program analyse data and look for meaningful connection. This information is used by government sector, marketing sector etc... There are lot of use of Multimedia data mining in Today's society. (e.g. The use of traffic camera footage, to show the traffic flow). Whenever the planning of new street will going on, in that location this information can be used. Basically there are four types of Multimedia data mining that are text, image, audio, video. All these four types of multimedia data mining use techniques for the further process. In the following section the description of the process and techniques for the multimedia data mining is given. Multimedia Data mining is the way of extracting useful data from the huge data. The unstructured or semi-structured data is sorted by the multimedia data mining. Pravin M. Kamde, Dr. Siddu. P. Algur[5] describes that World wide web is an important and popular medium for knowing all types of information which are related to sports, news, education, booking, business, science, engineering etc.. In today's competitive world the ability to extracting hidden knowledge from such type of information is very important. The entire process of applying computer methodology on such types of big information and extracting the useful knowledge from that is successfully done by multimedia data mining. Xingquan Zhu, Member, Xindong Wu, Ahmed K. Elmagarmid, Zhe Feng, and Lide Wu[12] explain that organization which deals with the big digital assets they have a need of that type of tool which deals with the retrieving information extraction from such collection. In

this situation the use of multimedia data mining is get processed. In fig 1 the basic process of multimedia data mining is shown.

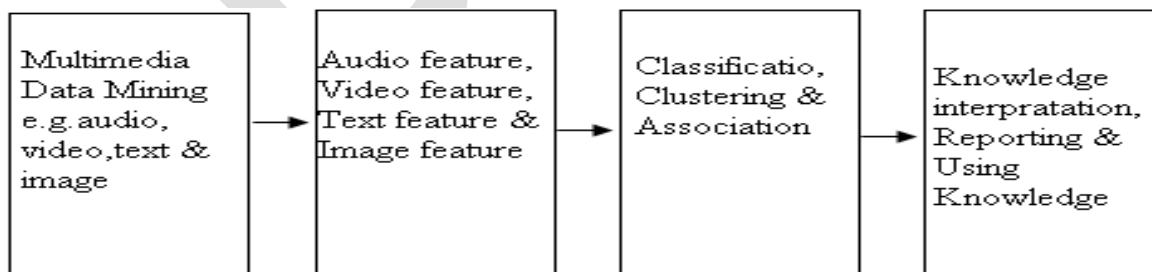


Fig.1 Multimedia Data mining Process

LITERATURE SURVEY-Bhavani Thuraisingham [10] explains that Multimedia data mining process is done by using some important techniques. In figure 1 the basic process of Multimedia data mining with their techniques is given. In the figure text, audio, video, and image are combine explained here the common process for mining all types of multimedia are shown. Starting point is the selection of multimedia type .i.e. audio, video, Images, text; it can also be called as raw data. Then the goal of text feature, audio feature, video feature, image feature is to discover important features from raw data. At this stage the data pre-processing is done. Data pre-processing includes feature extraction and transformation. If the informative features can be identified at the feature extraction stage. Detailed procedure depends highly on raw data. Finally the result of all these stages gets in the final stage. Knowledge Interpretation and reporting and using knowledge. It is the post processing and result evaluation stage. S. Kotsiantis, D. Kanellopoulos, P.Pintelas [11] describes that, as compare to data mining, multimedia data mining covers higher complexity resulting from: i) The huge volume of data, ii) The variability and heterogeneity of the multimedia data and iii) The multimedia content. A. Hema,E.Annasaro[1] explain All the views and ideas of all authors in field of multimedia data mining. The need of image mining is mainly focussed. Image mining have the great importance in the geological field, biological field, and pharmacy field. Pattern matching technique plays a vital role in Image mining. The process of extraction of useful Information hidden inside the image can be retrieve by pattern matching technique also. Xincheng ,Mihaela Vorvoreanu, and Krishna Madhvan[2] give knowledgeable information for those students or people who spend their more time on social media sites such as twitter, Facebook and you tube .And their elder ones worry about them, but by mining video also student can be focus on their study also. The focus of the paper is highly on the engineering students. NingZhong, Yuefeng Li, and Sheng-Tang Wu[3] describe the effective discovery of pattern which is used in text mining. Digital data on the internet is growing day by day, for turning such type of data in useful form, the need of text mining is occur. patterns can be discovered with Pattern Taxonomy model, Pattern deploying Method, inner pattern evolution. K. A. Senthildevi, Dr.E.Chandra[4] deals with the all technique's used in audio mining. In the areas of speech, audio processing and dialog the need of data mining is emerge. Speech processing process on the speech data mining voice data mining, audio data mining ,video data mining and conversation data mining . Speech data mining is useful for improving system operation and extracting business intelligence. And voice data mining (VDM) find and retrieve group of spoken documents such as TV or FM. audio of birds and pet animals recorded. Video data mining is use for the surveillance video. Conversation data mining is used in call centre. All the problems. Issue of caller gets understood. Pravin M.Kamde, Dr.Siddu.P.Algur [5] the diagrammatical representation of web mining taxonomy, mining multimedia database, text mining, image mining ,video mining, multimedia mining process are explain. Classification model, Clustering model and Association rule are some technique use for multimedia mining.Cory Mc Kay.David Bainbridge[6] describe the greenstone digital library software for extraction of musical web mining audio. feature extraction extension. JMIR is software tool is use for other resources. JMIR includes the components that are jAudio, jSymbolic , jWebminer2.0, jLyrics, ACE 2.0, jMusic Meta Manager, lyric feature, jMIR utilities, ACE XML.

A.Text mining with Information extraction

Ning Zhong, Yue feng Li, and Sheng-Tang Wu [3] say that there is lot of information is in the textual form. This could be library data or electronic books or web data. The one problem face by text data is, it is not well structured as relational data. In many cases it can be unstructured or it may be semi-structured. So the "Text Mining" is useful for describing the application of data mining techniques to automated discovery of useful and interesting knowledge from unstructured or semi-structured text.)Raymond J. Mooney and Un Yong Nahm [9] describes that, there are several techniques are proposed for text mining. That are conceptual structure, association rule mining, episode rule mining, decision trees and rule induction method. with attachment to this Information Retrieval technique is widely use for performing task such as document matching, ranking and clustering. From large text database, extraction of patterns and association is done by text mining. For text document, identifying the keywords that summarizes the content is needed. Words can occur frequently, Such as "the", "is", "in", "of" are no help at all, since they are avoided in every document. During the pre-processing stage these common English words can be removed using "stop-list" Bhavani Thuraisingham [10] describe that One can form association from the keywords. In one article the keyword can be "Belgium, nuclear weapons" and keyword in another article can be "Spain, nuclear weapons". The data mining could make the association that author from Belgium and Spain write articles on nuclear weapon. Fig 2 shows the process of Text mining. Xincheng ,Mihaela Vorvoreanu, and Krishna Madhvan[2] give knowledgeable information for those students or people who spend their more time on social media sites such as twitter, Facebook and you tube .And their elder ones worry about them, but by mining video also student can be focus on their study also. The focus of the paper is highly on the engineering students. Fig 2 shows the process of text mining.

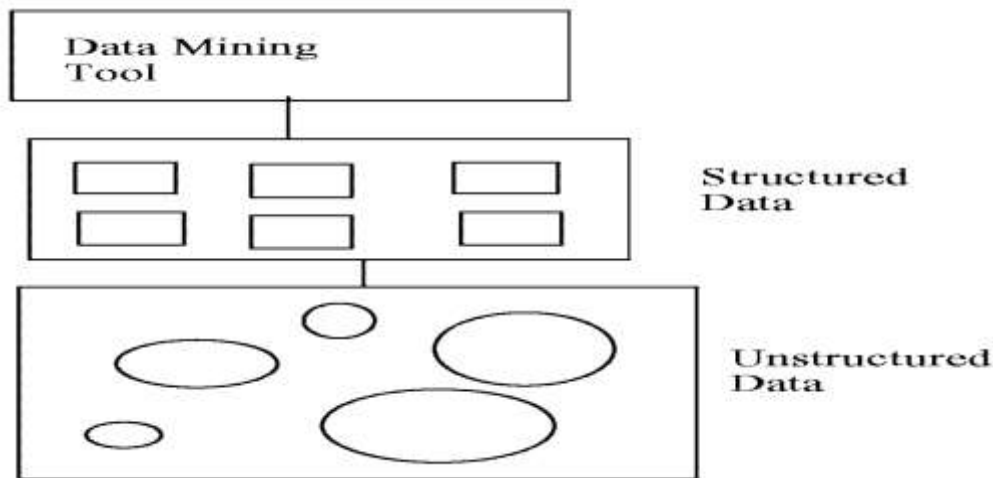


Fig.2 converting unstructured data to structured data for mining. Bhavani Thuraisingham [10]

B. Image mining with Information extraction- Nearly the techniques use for all types of multimedia data mining are identical, But the structure of various multimedia types are different, so according to that, the process of the mining of various multimedia type is different. Sometimes question get arises, if there is an availability of image processing so exactly what is the use of image mining? Image processing applications are in various domains, such as medical imaging for detection of cancer, Satellite images processing for space and intelligence application. Images include the geographical area, structure of biology. Tao Jiang ,Ah-Hwee Tan[7] explains that, important application of Image mining is, image mining not only detect the outcome from unusual pattern in image but also identify recurring themes in image, both these thing are done at the level of raw images and with higher-level concept. to find existence of pattern within a given description, the Matching technique is used. A.Hema, E. Annasaro[1] says that In the field of image mining, image matching is the vital application. There are so many techniques have been developed till today and still research for developing the optimized matching technique is going on. Nearest neighbourhood technique, least square method technique, coefficient of co-relation technique, relational graph isomorphism technique all these are matching techniques. Nearest neighbourhood technique is an important technique used in applications where objects to be matched are represented as n-dimensional vector. Fig 3 shows the process of image mining

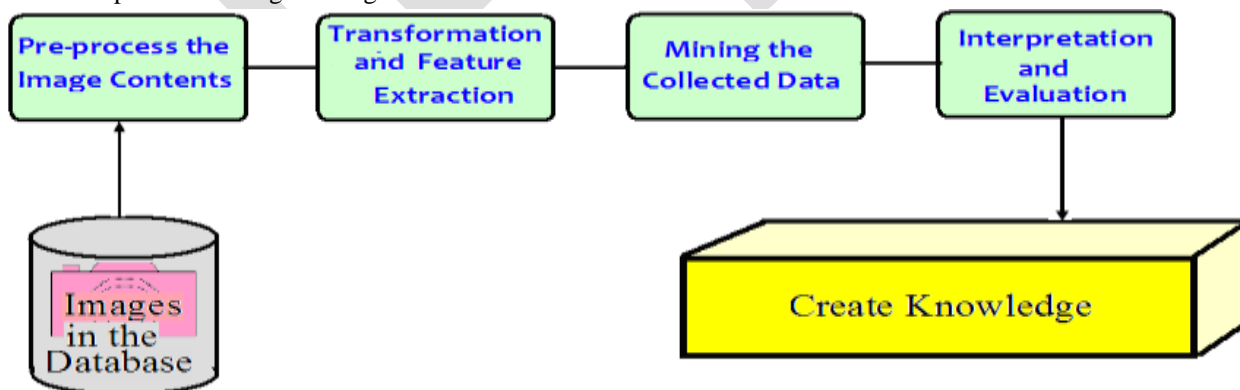


Fig.3. Image Mining Process.) Pravin M .Kamde, Dr. Siddu. P. Algur [5]

Video mining with feature extraction- Video mining is the third type of multimedia data mining. Video is the combination of images so the first step for successful video mining is to have a good handle on image mining. Ajay Divakaran, Kadir A. Peker, Shih-Fu Chang, Regunathan Radhakrishnan, Lexing Xie[11] says that, In terms of feature extraction, video feature extracted for each shot

based on detected shot boundaries. There are totally five video feature extracted for each shot, named as, pixel_change, histo_change, background_mean, background_var, and dominant_color_ratio. when the raw data is taken for the information extraction in video mining these five features are help for mining the video.) Mei-Ling Shyu, Zongxing Xie, Min Chen and Shu-Ching Chen[8] describes The basic techniques for video data mining, that are pre-processing of raw data, classification and association. In pre-processing of raw data technique, the important terms are get considered, that are, video shot detection and classification, video text detection and recognition, camera motion characterization, and salient audio event detection. Now in association Mining technique there are three terms are get considered that are video data transformation, definition and terminology, and video association mining. Video mining is day after day improving their techniques in various ways. Fig 4 shows the direct video mining process.

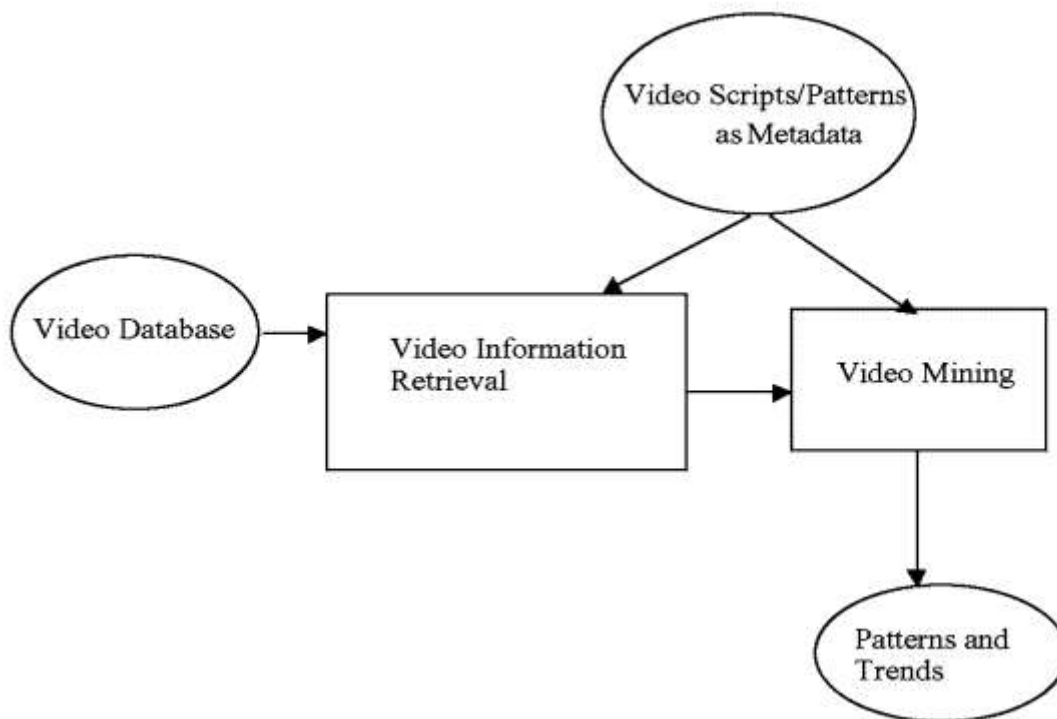


Fig.4.Direct video mining Bhavani Thuraisingham [10]

Audio mining with feature extraction-In multimedia application, Audio data plays vital role. Cory McKay. David Bainbridge[6] describes that Music information basically have two categories. a) Symbolic and b) Audio information. Audio is now became the continuous media type like videos. The techniques used in audio mining is similar to techniques used in video mining. audio data can be available in any form such as speech, music, radio, spoken language etc. The primary need for mining the audio data is the conversion of audio into text, using speech transcription technique this process can be done. other techniques are also available for this such as keyword extraction and then mining the text. Audio mining is that type of technique which is used to search audio files. K.A.Senthildevi, Dr.E.Chandra[4] explains that there are two main approaches of audio mining. 1) Text based indexing and 2) Phoneme based indexing. Text based indexing deals with the conversion process of speech to text. And Phoneme based indexing doesn't deals with conversion from speech to text, but instead works only with sound. Fig 5 shows the process of Audio mining.

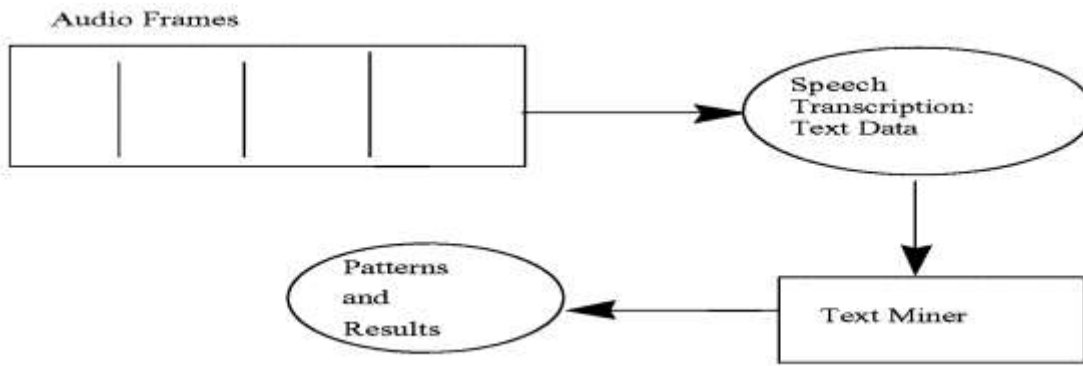


Fig.5.Mining text extracted from audio. [10]

APPLICATION of MULTIMEDIA DATA MINING

Multimedia data mining is the big application for all types of sectors and field. In today's society multimedia is the essential part of all kinds of work. Some application of multimedia data mining is as follows.

- A. To Know the geographical condition, agriculture, forestry, crops measurement, monitoring of urban growth, mapping of pollution, mapping of ice for shipping. Identification of sky object the satellite data is used.[5]Pravin M. Kamde, Dr.Siddu.P.Algur
- B. The use of Audio and video mining can be seen in movie mining system.
- C. For vehicle identification, traffic flow, and the spatio-temporal relations of the vehicle at intersection, the mining of traffic video sequences is used.[8]Mei-Ling Shyu,ZongxingXie, Min Chenand Shu-Ching Chen.
- D. For detecting sports video, or in big shops, the video mining is used.

CONCLUSION and FUTURE SCOPE

In this paper the description of techniques which is needed by the multimedia data mining is given. In text mining two approaches are used for information extraction. In first approach the general knowledge can extract from direct text. And in second approach, one can extract structured data from text documents. In image mining matching technique is used for finding the existence pattern of an image. To handle video mining, one should know all about the image.

REFERENCES:

- [1]A.Hema, E.Annasaro. "A SURVEY IN NEED OF IMAGE MINING TECHNIQUES", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE) ISSN (Print) : 2319-5940 ISSN (Online) : 2278-1021, Vol. 2, Issue 2, February 2013.
- [2]Xinchen,MihaelaVorvoreanu,andKrishnaMadhvan,"Mining Social Media Data for Understanding Students' Learning Experiences", IEEE computer society.1939-1382(c)2013 IEEE.
- [3]NingZhong, Yuefeng Li, and Sheng-Tang Wu, "Effective Pattern Discovery for Text Mining", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 1, JANUARY 2012.
- [4]K.A.Senthildevi, Dr.E.Chandra, "Data Mining Techniques and Applications in Speech Processing - A Review", International Journal of Applied Research & Studies(IJARS) ISSN 2278 – 9480Vol.I / Issue II /Sept-Nov, 2012/191.
- [5]PravinM. Kamde, Dr.Siddu.P.Algur. "A SURVEY ON WEB MULTIMEDIA MINING", the International Journal of Multimedia & Its Applications (IJMA) Vol.3, No.3, August 2011.

[6] Cory McKay, David Bainbridge, "A MUSICAL WEB MINING AND AUDIO FEATURE EXTRACTION EXTENSION TO THE GREENSTONE DIGITAL LIBRARY SOFTWARE", 12th International Society for Music Information Retrieval Conference (ISMIR 2011).

[7] Tao Jiang, Ah-Hwee Tan, "Learning Image-Text Associations", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 21, NO. 2, FEBRUARY 2009.

[8] Mei-Ling Shyu, Zongxing Xie, Min Chen and Shu-Ching Chen, "Video Semantic Event/Concept Detection Using a

Subspace-Based Multimedia Data Mining Framework", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 10, NO. 2, FEBRUARY 2008.

[9] Raymond J. Mooney and Un Yong Nahm, "Text Mining with Information Extraction", Multilingualism and Electronic Language Management: Proceedings of the 4th International MIDP Colloquium, September 2003, Bloemfontein, South Africa, Daelemans, W., du Plessis, T., Snyman, C. and Teck, L. (Eds.) pp.141-160, Van Schaik Pub., South Africa, 2005.

[10] Bhavani Thuraisingham, "Managing and Mining multimedia Database", International Journals on Artificial Intelligence Tools, Vol. 13, No. 3, 739-759, 20 March 2004.

[11] Raymond J. Mooney and Razvan Bunescu. "Mining Knowledge from Text Using Information Extraction", SIGKDD Explorations. Volume 7, Issue 1.

[12] Ajay Divakaran, Kadir A. Peker, Shih-Fu Chang, Regunathan Radhakrishnan, Lexing Xie, "Video Mining: Pattern Discovery Versus Pattern Recognition", IEEE International Conference on Image Processing (ICIP) TR2004-127 December 2004.