THINK INDIA JOURNAL

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

An Extensive Review on Different Strategies of Multimedia Data Mining

Pradeep Kumar

Assistant Professor

Department of Computer Science and Engineering
Lovely Professional University
Jalandhar, Punjab
pradeep.16473@lpu.co.in

***Abstract*— Multimedia data mining is the process of extracting the media data from the data store. In today's time, there is a huge amount of multimedia data generated from different data sources every single second. The multimedia data is generally used for sharing the information or conveying the messages through different media such as text, audio, images, video, animation and graphics. As nowadays, we are using huge amount of digital sources for communication so it is also important to manage such a huge data in a proper order. Managing the data in proper order also help us to filter or mine the data from the data stores. The data should be organized in an appropriate manner and should be categorized properly so that the required data should be retrieved from the data store within less amount of time. Multimedia data mining is used for the purpose of extracting the data from the data store.**

***Keywords*— Multimedia Mining, Static data, Dynamic Data, Clustering**

I. INTRODUCTION

Information mining and information disclosure gives us another approach to determine the issue that we can't make the best of expanding gigantic information [1]. Information mining is the key innovation in computerized library. Individuals need not exclusively to recover the data of content yet additionally to oversee and recover video data. It makes the interactive media information mining the hotspot of research. Mixed media Data Mining is the procedure to be utilized to find the certain, powerful, important and understandable information from a lot of sight and sound information. This can be done b by breaking down the element of seeing and hearing. Moreover, it can likewise give us the capacity of choice to determine the issue [1]. This paper examines the fundamental speculations of Data Mining and Knowledge Discovery and presents the procedure, techniques and information portrayal of sight and sound information mining. At that point, it dissects the component of data recovery and interactive media mining engine (MME) [1].

In the course of recent decades, information quarrying or mining has been a viable methodology for separating covered information from colossal assortments of managed advanced information that put away in databases. Interactive media information alludes to the examination of a lot of mixed media data so as to discover designs or factual relationships [2]. Rapid changes in data innovation have radically changed the capacities and exercises of sight and sound. It incorporates sound, video, discourse, content, picture and web, which are getting progressively accessible and are practically unstructured or semi organized information [2].

II. RELATED WORK

Sight and sound information digging is utilized for extricating intriguing data for interactive media for example, sound, video, pictures, illustrations, discourse, content and mix of a few kinds of informational collection which are altogether changed over from various organizations into advanced media [2]. Sound

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

mining is a subfield of information mining which is utilized to discover intriguing data of known information from mixed media databases [3]. Mining of interactive media information requires at least two information types for example, content and video or content video and sound. Mining in interactive media is alluded to programmed comment or explanation mining [3]. Multimedia data can be of different types such as audio, video, images, graphics and animation.

The different types of multimedia data will be combined to do the research. Sometimes, the text data mining could be an area of research whereas, sometimes the video and animation mining could be the research fields. It has been analyzed that the difficulty occurs for extracting the media data from huge amount of data stores. The research for extraction of the data came into existence in the early 90's [4].

There are many algorithms have been developed so far for the extraction of media data. The media data stored in multimedia database management systems. These data bases are used for retrieving the data by applying a technique for data mining. The technique could be the implementation of an algorithm of data retrieval. Multimedia data stores are used for storing, updating and fetching the necessary information, which is required by user [5]. Fig. 1 below shows the different types of mining that could be implemented [6].Multimedia Data Mining is widely applied for those kinds of data stores where we store large amount of structured, semi-structured and unstructured data. The media elements can be stored by using those data stores which are implemented on the basis of object oriented approach [3].

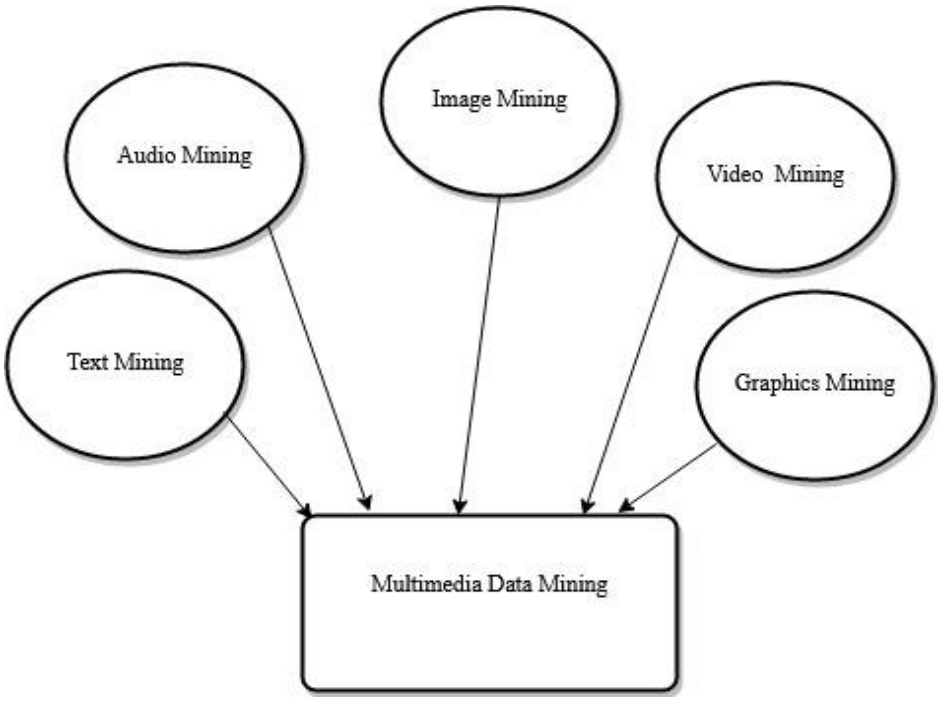

Fig. 1 Multimedia Data Mining [6]

Structured data is the data which is stored by using the relational approach means using rows and columns whereas unstructured data is the data which needs special mechanism to store it. Video, animation and images come under unstructured data. As compared to structured data, unstructured data is quite difficult to store and manage [3].

III. MULTIMEDIA DATA ARCHITECTURE

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

Multimedia data architecture is used for storing the different types of media data elements. The data can be either structured data or unstructured data. First of all, data need to be categorized and then we have to store it in the data ware houses. The data mining tool is used to extract the data elements from the data bases. Before applying the data mining technique, we have to convert the unstructured data to structured data [6].

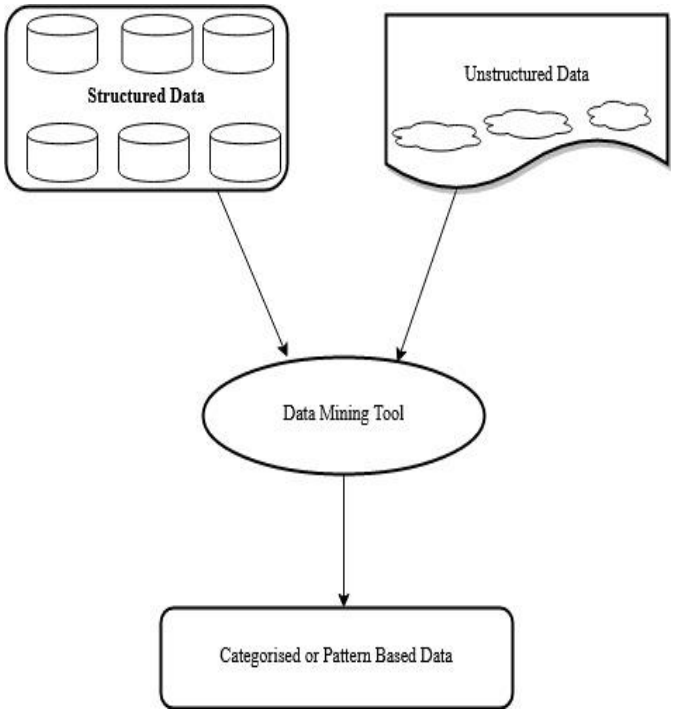

Fig. 2 Conversion of Structured and Unstructured data [3]

The figure above shows the process of applying the data mining mechanism for the conversion of unstructured media data into structured data. The conversion can be done by applying the data mining mechanism. The data which we collected must be stored in an organized manner so that we can retrieve the required data within fraction of seconds from the large databases. Basically, the initial data considered as a raw data because it has to be arrange in an organized form in the data ware houses. The raw data is then arranged in an organized form by using various data arrangement tools. If any user want to fetch the data then data mining algorithms are required to retrieve the necessary data from the data ware house.

Data mining is the process of extracting the media data from the data stores. Multimedia data mining process consists of different sub processes. First of all, we have to collect the data. Then the features are extracted from the collected media data. Features are actually the different parameters of the media data e.g. for image data the parameters could be its luminance, chrominance, saturation etc. After extracting the features, we need to classify these features into a specific category.

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

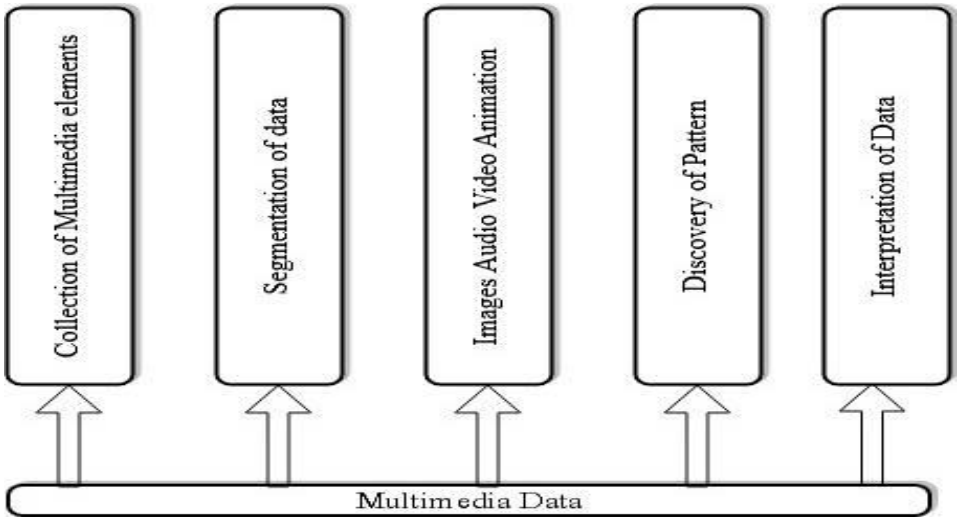

Fig. 3 Architecture of Multimedia Data Mining [3]

Fig. 3 shows the various levels that are the integral part of the architecture of multimedia data mining. The gathered media data will need to be represented in different patterns. Patterns are framed on the basis of media data whether it is image, video, animation or textual data. The media data fetched from the media data store on the basis of pattern recognition. The hidden patterns are collected from the stored media data. Different types of pattern identification schemes such as integration, classification and clustering etc. can be applied for generating the patterns of the media data elements [3]. The ending phase of the architecture of data mining process is interpretation phase. This phase is used for the verification of data which has been generated. Moreover, this phase is also responsible for knowing that whether we have to go back to the previous phases for generating the valuable and quality oriented data, or not [3].

IV. MULTIMEDIA DATA MINING STRATEGIES

There are many processes, techniques and strategies are available for the successful implementation of data mining process. During the survey, we found that we need to learn the multimedia storage and searching techniques for implementing any of the strategy for media data mining process. Most of the processes include feature extraction, searching and pattern recognition phases.

A. Clustering Based Strategy for Data Mining

Generally, clustering is the process where we made the groups of such data items which are very much similar in nature. Similarity can be based upon different patterns of the data. After the media data collection, the next step would always be arranging the data into different groups. This seems to be very easy but technically it would be very much a tedious task. Clustering is implemented on the basis of neural network and different mapping based procedures [7]. One can also follow the incremental clustering based strategy at different aspect ratios using wavelet transform mechanism that is based upon k-means algorithm.

B. Classification Rules Based Strategy for Data Mining

Classification rules based strategy can also be used for the efficient and effective data mining process. This strategy also works on to manage the media data elements efficiently. Generally, we work on finding the semantic meaning of the data for the storing purpose. Hidden-Markov model can also be used for the classification of the data such as video and images [7]. There is another method which is based upon

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

combined multi-model feature and classification rules [8] which can be employed for the arrangement of media data in an appropriate way.

C. Association Rules Based Strategy for Data Mining

Besides classification rules based strategy, there is another strategy for data mining process that is based on the concept of association rules. This strategy is used to find the relationship between the different variables which have been considered for storing the data in the media data stores. Different types of association can be identified and implemented on the basis of the logic of association rules. Image data can be segmented to find the similarity in different regions of images [9]. Multilevel based rules can also be imposed on different image objects [10]. FP- Tree based algorithm can also be used for the implementation of association rules [11]. With the help of association rules, the multiple descriptors can be identified for the same image data.

D. Statistical modeling Based Strategy for Data Mining

In this strategy, different types of models need to design to find the relationship amongst different type of media data element. These models help us to find the relation between various image features and their keywords [7]. Another method based upon co-occurrence model can also be used to determine the link and relation between different features of an image data [15].

V. CHALLENGES IN MULTIMEDIA MINING

The Major challenges that occurred in the field of multimedia data mining are content based storing and retrieval of media data, searching of media data, classification and prediction analysis of media data etc. [16].

A. Text Based Storage Retrieval of Data

One of the challenges is to retrieve the data from the storage hub on the basis of content. In early days, textual description was used for storing the media data in the data store. For storing the data there always would be a need of textual descriptor. Textual descriptor is an expert person who writes the description about the media data. Such description is also called a meta-data. But this was very tedious and time consuming process because textual descriptor needs to first watch the media data and then analyze the data to write its meta-data. Moreover, problem occurs when one has to retrieve the data because there could be two words with the same name in the meta-data. During the retrieval process, it might create confusion so the chance of getting the right data from the data store deteriorates.

B. Content Based Storage Retrieval of Data

To resolve the issue of text based retrieval of data, an advanced mechanism has been developed. This advance mechanism is called as content based storage retrieval of data [17]. In this approach, the logic of query by example has been implemented. In case of query by example, the user first uploads the demo media data say demo image. Then system software will extract the features from this demo image. These extracted features will then compare with already stored features. If the percentage of matching features reaches the threshold limit then all such images for which feature matches will be extracted from the media data store and will be presented to the user for further operations. This process is very much efficient but challenge is, the complex and complicated algorithms need to be implemented which requires really great expertise in the field of software development.

C. Classification and Forecast Analysis

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

Multimedia data is usually stored in media data ware house. This data is requiring the classification. Classification is the process of categorizing the media data on the basis of its requirement in different domains. Media data can also be demanded by any user for doing any kind of forecasting analysis. If a user is working in the field of entertainment like producing and directing movies then sometimes he/she might require the data for doing the forecast analysis. In such cases, the media database is very much helpful. Because, one can store, categorize and retrieve the data at any point of time.

Image and video data came at a very high rate and most frequently. For retrieving this type of data, system requires a very high processing power and speed. So high performance based processors need to employed. Hence, it increases the cost of whole process implementation. The challenge in this process is how to reduce the cost expenditure. If we could able to reduce the cost then the data retrieval system becomes very much cost effective.

VI. CONCLUSION

Multimedia data mining is one of the grooming research areas. In today's time, most of the smart phone users are using the social media for sharing their messages, images and videos. Problem is how to store this huge amount of media data in data ware houses. More over, it is also a very challenging task to categoriste and filter out such a huge amount of data. Many algorithms and strategies need to be designed for the efficient retrieving of the media data. In this paper, we reviewed the various strategies which can be used in data mining process. We found that still we need lots of up gradation in the existing strategies so that we can fetch the media data efficiently from the media data stores. We have also reviewed about the multimedia data architecture which helps in understanding the whole idea of media data representation in data store. We have also reviewed the various challenges that can be encountered for the smooth storage of media data in the data ware houses.

REFERENCES

[1] https://ieeexplore.ieee.org/stamp/stamp.jsp?tp= &arnumber=5690870&isnumber=5688492

[2] S. Kotsiantis, D. Kanellopoulos, P. Pintelas, Multimedia Mining, "Multimedia mining', WSEAS Transactions on Systems, Vol. 3, No. 10, pp.3263–3268.

[3] Mylavarapu Kalyan Ram, Dr. M. Venkateswara Rao, Challapalli Sujana, "An Overview on Multimedia Data Mining and Its Relevance Today," International Journal of Computer Science Trends and Technology (IJCST), Volume 5, Issue 3, pp 118-113, Jun 2017.

[4] Sarla More, Durgesh Kumar Mishra, "Multimedia Data Mining: A Survey" Pratibha International Journal of science, spirituality, business and technology, vol. 1, no.1, March 2012 ISSN, pp 2277-7261.

[5] Bhavani Thuraisingham, "Managing and mining multimedia Databases" International Journal on Artificial Intelligence Tools, Vol. 13, No.3 (2004), pp 739-759.

[6] Bhavani Thuraisingham, Managing and Mining Multimedia Databases, Published by CRC Press, 2001.

[7] Shu-Ching Chen, Mei-Ling Shyu, Chengcui Zhang, and Jeff Strickrott, "Multimedia Data Mining for Traffic Video Sequences," Proceedings of the Second International Workshop on Multimedia data

ISSN: 0971-1260
Vol-22-Issue-3-July-2019

Mining MDM/KDD'2001), in conjunction with the Seventh ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 78-85, August 26, 2001, San Francisco, CA, USA.

[8] Shu-Ching Chen, Mei-Ling Shyu, Chengcui Zhang, MinChen, A multimodal data mining framework for soccer goal detection based on decision tree logic, International Journal of Computer Applications in Technology, Volume 27 , Issue 4, 2006, Pages 312-323.

[9] Ordenoz C, Omiecinski E. Discovering association rules based on image content. In: ADL 99: Proceedings of the IEEE Forum on Research and Technology Advances in Digital libraries Washington, DC: IEEE Computer Society, 1999, pp. 38.

[10] Tseng, V.S.; Ming-Hsiang Wang; Ja-Hwung Su, A New Method for Image Classification by Using Multilevel Association Rules, Data Engineering Workshops, 05-08, April 2005, pp. 1180 – 1180.

[14] Ankur M. Teredesai , Muhammad A. Ahmad, Juveria Kanodia and Roger S. Gaborski, "Knowledge and Information Systems", Volume 10, August, 2006, Springer London.

[15] Mori Y, Takahashi H, Oka R. Image-to-word transformation based on dividing and vector quantizing images with words in MISRM'99 First International Workshop on Multimedia Intelligent Storage and Retrievel management, 1999.

[16] Mittal, Ankush, "An overview of multimedia content-based retrieval strategies", Publication: Informatica, October 1, 2006.

[17] Yu H, Wolf, "Scenic classification methods for image and video databases", in SPIE International Conference on Digital Image Storage and Archiving Systems, Vol. 2606, 1995, pp. 363-371.

[18] S. Kotsiantis, D. Kanellopoulos, P. Pintelas, Multimedia Mining, (2004), ''Multimedia mining'', WSEAS Transactions on Systems, Vol. 3, No. 10, pp.3263–3268.

[19] V. Vijayakumar, R. Nedunchezhian, "A study on video data mining", International Journal of Multimedia Information Retrieval, October 2012, Volume 1, Issue 3, pp 153-172.