

Multimedia Data Mining and its Integration in Information Sector and Foundation: An Overview

Prantosh K Paul¹ and K.S.Shivraj²

¹Indian Institute of Engineering Science and Technology (IEST), Shibpur, West Bengal, India

²K L University, Vijayawada, Andhra Pradesh, India

E-mail: prantoshkpaul@gmail.com

(Received on 25 December 2013 and accepted on 20 February 2014)

Abstract - Information and Communication Technologies are one of the important component and toll. Virtually, the advent of Electronic resources and similar foundation use in Information Foundation and similar foundation has brought about significant changes in storage and communication of information. Data mining process consist of several process and stages, which are related to each other and interactive. This is the way of mining or extraction of data from the Database or Dataset. Extraction of Data with multimedia nature such as audio, video, images, text may be called as Multimedia Data Mining. In Information Foundation, Data Mining has wonderful role and importance. This paper is talks about Multimedia Information and Data Mining and its characteristics. Paper also talks about role and need of Multimedia Data Mining in Information and Similar Foundation.

Keywords: Information, Information Science, Data Mining, Multimedia Technology, Information Foundation, Multimedia Data Mining, MDM, Knowledge Management, Pattern Discovery, Multimedia Content

I. INTRODUCTION

Multimedia Data Mining is the broader aspect and periphery of information and knowledge or data mining. Multimedia Data Mining is a part of Multimedia Technology and needed for so many aspects such as media compression as well as storage, delivering streaming media, multimedia indexing, summarization, search and retrieval and so on. Multimedia Data Mining is combination of some step and procedure which includes problem definition, data preparation, implementation of the analysis, deployment of results. Multimedia Data Mining is applicable in

Information Foundation such as Data Centre, Information Centre, Information Networks, Knowledge Grid, and Information Systems and so on. Digital Repository, Digital Archive are the important sector in this field [04, 08]. Virtually, Multimedia Data Mining techniques are active and growing area of research in today Information Science and Technology circle. Actually, healthy and sophisticated Data Mining strategies, considering standards, features are available techniques needed for complete utilization of Multimedia Data Mining in Information Foundation and similar entities [10, 14].

II. OBJECTIVES OF THE STUDY

The main aim and objective of this study is includes but limited to as follows:

1. To know basic about Multimedia Data Mining; including its characteristics, features;
2. To know basic about Multimedia Data Mining and its need and value;
3. To know basic about the application and integration of Multimedia Data Mining in Information and Documentation Foundations;
4. To know about the challenges and opportunities of Multimedia Data Mining applications in Information Foundation and so on;
5. To learn about the latest architecture of Multimedia Data Mining briefly.

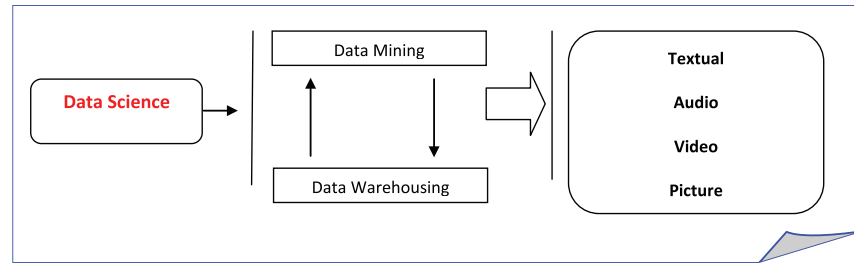


Fig. 1 Depicted the general activities of Data Mining and allied domain in terms Multimedia Services

III. MULTIMEDIA DATA MINING: BASICS

Data Mining is the important way and method of collection of Data from large database and helps in decision making and organizational activities. Data Mining is the automatic extraction of pattern of Data from Big Warehouse. As far as types of Data Mining are concerned, Textual Data Mining and Web Data Mining are most important[12, 18]. However, Multimedia Data Mining is new term than these two; and deals with collection of audio, video, picture, information and text from the huge database. Some expert explain that, the concept of mining is multimedia is also treated automatic annotation and some times as Annotation Mining. The Multimedia Data Mining is also known as main pattern discovery. Virtually, Multimedia Data Mining deals with some stages. Multimedia Data Mining is consisting with following segment or items:-

- a) Media Compression and Backup;
- b) Delivering Online and Streaming Media;
- c) Audio/ Video Indexing and Media Indexing;

- e) Designing and Development of Media and Multimodal User Interface.

As Data Mining is deals with so many steps and thus each and every step is need to fulfill for better and interaction Multimedia Data Mining. In first stage, i.e. Data selection storage which is requires the user to target or particular database or sub field or Data Records for possible Data Mining. However, in other sense, Domain understanding is also very much importance. Ultimately a proper understanding is essential. The next step is clearing and preprocessing step, however in Multimedia Data Mining, this step become not so much popular. As in Multimedia Data Mining, data kept in rational form and not at per with traditional data mining systems [20, 21].

As far as pre processing step is concerned, it is basically involves integration of data from different sources and making choice about representing or coding certain data fields that serve as input to the pattern discovery stages. As far as approach is concerned, it is also deals with so many approaches such as association, classification, clustering, regression, time series analysis and so on.

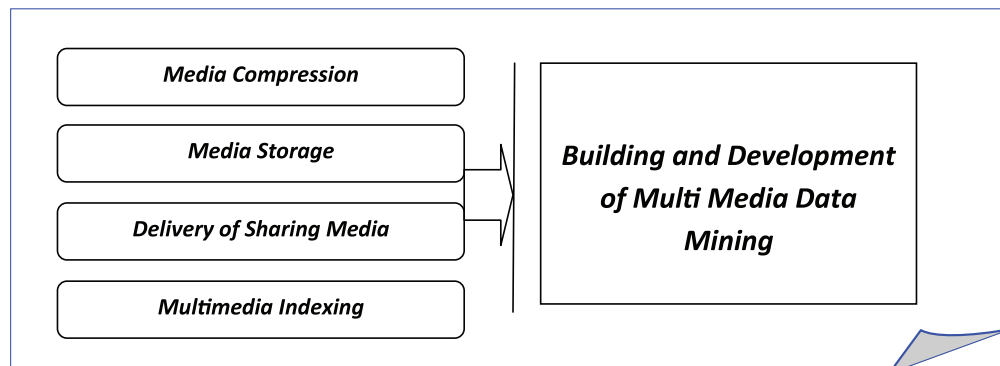


Fig. 2 Depicted Multimedia Data Mining gradients and aspects

The last stage in Data mining is interpretation and which not limited to evaluation of quality of discovery and its value, proper domain understanding and so on[18, 19]. However, Myat put importance in following for better data Ming projects.

- a) Problem Definition;
- b) Data preparation;
- c) Implementation of the Analysis;
- e) Deployment Results.

A. Multimedia Data Mining: Some standards

Multimedia Data Mining is deals with so many color, edges, shape and texture during Data Mining. As far as Multimedia Data Mining is concerned, following are important and useful:-

- 1. For spatial distribution of color in an image is concerned, color histograms may be used as main features;

- 2. For global shape property of the segmented region or a shape cloud be described by representing Fourier descriptors[23, 24];
- 3. As far as easiness is concerned, global descriptors are generally useful and able in compact representation and one learn prior to segmentation errors;
- 4. As far as MPEG-7 Visual Description standard is concerned, it provides several features such as
 - Color Layout Descriptor;
 - Color Structure Descriptor;
 - Dominant Color Descriptor;
 - Scalable Color Descriptor.

However, MPEG-7 is provides two main shape descriptors and some other user are based and additional semantic information;

- The MPEG 7 Audio standard defines two set of audio descriptors.

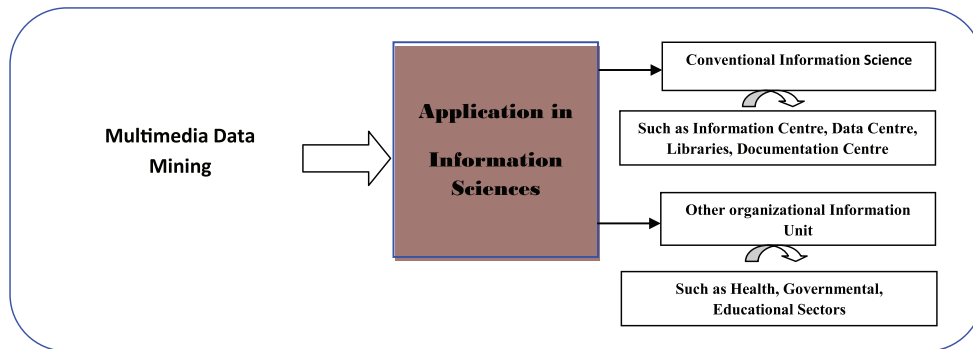


Fig. 3 Multimedia Data Mining and its conventional use place and emerging utilities

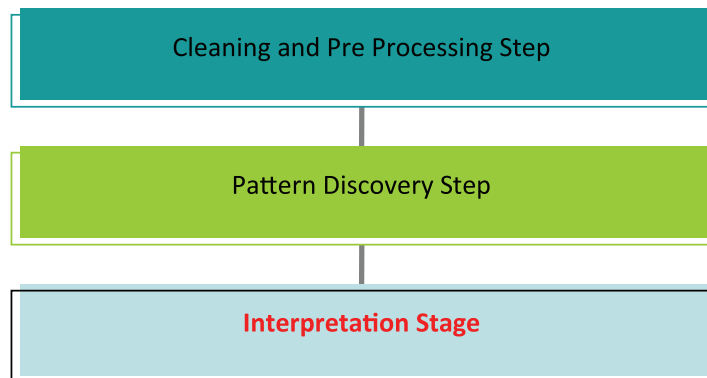


Fig. 4 Major stages of Multimedia Data Mining building

B. Multimedia Data Mining: Information Foundation

Multimedia Data Mining plays an important role in Information Science practice and mainly in Digital Repositories, Digital Archives, and Digital Libraries and so on. As far as Multimedia Data Mining is concerned, it has several standards such as

- Dublin Core;
- RDF;
- EAD;
- TEI
- SGML
- XML/HTML
- MARC and so on.

As far as our India is concerned, Multimedia Data Mining is applicable only in some of the organization and universities and still Digitalization process is in primary stages. Virtually, Multimedia Data Mining needs several things for its development. Like fundamentals conversation of content from physical to digital form, starting digital context and metadata in appropriate multimedia repository, content and information delivery via file transfer or streaming media. Today, information is disseminated in several formats and out of which multimedia based information system is important one. As far as its mining is concerned, it is treated as important; as user's demand need to fulfill properly.

IV. FEATURE POTENTIALS

Multimedia Data Mining is important for several reasons; virtually the concept of mining in multimedia is also referred to as automatic annotation or annotation mining. For better Multimedia Data Mining, it is essential to follow each and every step very carefully. Ultimately, Multimedia Data Mining is an Applied discipline, which is growing for several reasons. Depending upon nature, the Multimedia Data Mining may also change. Multimedia Data Mining has potential in Cloud Computing and virtualization and makes easy transformation in information from one place to another [25, 26].

V. CONCLUSION

Information is power and needed in most of the sector such as business, IT, health, and social sector and so on. And Multimedia Information System is contemporary word and valuable for multi format audio, video delivery to the large audience [27, 28]. The information is getting triple every year and thus, the large number of data becomes tough to manage. The size of Multimedia Information Systems is big than textual information system and thus it is tough to find out exact and needed information; hence Multimedia Data Mining is proposed by some of the renowned data and information scientist near about 30 years back.

REFERENCES

- [1] Cohen, E. B. (2004). Applying the Informing Science Framework to Higher Education: Knowledge Development, Management, and Dissemination. Konferencja Pozyskiwanie wiedzy i zarządzanie wiedzą (Proceedings of the Knowledge Acquisition and Management Conference) May 13-15, 2004 Kule, Poland.
- [2] Cohen, Eli B. and Nycz Malgorzata (2006). Learning Objects and E-Learning: an Informing Science Perspective. *Interdisciplinary Journal of Knowledge and Learning Objects* Volume 2, 2006.
- [3] Martin, S.B. (1998). Information technology, employment, and the information sector: Trends in information employment 1970–1995. *Journal of the American Society for Information Science*, 49(12), 1053–1069.
- [4] Michael Buckland and Ziming liu (1995). History of information science. *Annual Review of Information Science and Technology* vol. 30: 385-416.
- [5] Prantosh Kumar Paul, “ Information Scientist: Roles and Values with special Reference to their Appropriate Academic Programme and its availability in India:” *International Journal of Information Dissemination and Technology*, Vol. 2, No. 4, October-December-2012, Page-245-248, ISSN-2229-5984 [Indexed in DOAJ, EISRJC, J-GATE, Ulrich Directory, Google Scholar, Proquest, Index copernicus and other major databases]
- [6] Paul, Prantosh Kumar, D Chatterjee, R Bhatnagar, Uma Pricilda “Information Scientist: Contemporary innovative techno management roles with special reference to Information Scientist Vs Information Technologist: A Study”, *Indian Journal of Information Science and Applications [IJISA]*, Vol. 2. No. 1, Jan-Jun-2012, Academic Research Publication, New Delhi, Page-47-50, ISSN-2249-3689
- [7] Paul, Prantosh Kumar, D Chatterjee, M Ghosh “Neural Networks: Emphasizing its Application in the World of Health and Medical Sciences” *Journal of Advances in Medicine*, Vol. 1 No. 2, July-Dec, ISSN-2277-9744 Page-17-23, New Delhi Publisher, New Delhi.

- [8] Prantosh Kumar Paul, Ashok Kumar, Dipak Chatterjee “ Health Informatics and its Practice: Emerging Domain of Information Science-Indian Scenario” in Current Trends in Biotechnology and Chemical Research, Vol. 2 No. 2, July-Dec, 2012, Page- 83-87, ISSN-2249-4073 [Indexed in DOAJ, Index Copernicus, Google Scholar, CAS-USA]
- [9] Prantosh Kr. Paul, K L Dangwal, Asok Kumar Garg “ Education Technology and Sophisticated Knowledge Delivery” *Techno-Learn-International Journal of Education Technology*, ND Publisher, New Delhi, Vol. 2, No. 2, Page-169-175 ISSN-2231-4105
- [10] Prantosh Kr. Paul, K L Dangwal and Dipak Chatterjee, “Information Technology and Advance Computing and their interaction for healthy Education, Techning, and learning: The IKM Approach” *Asian Journal of Natural and Applied Sciences*, ISSN-2186-8468, Page-70-77 V-1, No. 4, December-2012, Leena and Luna International, Oyama, Japan
- [11] Paul,Prantosh Kumar, M K Ghose, “Cloud Computing: Possibilities, Challenges, and opportunities with special reference to its emerging need in the academic and working area of Information Science”, ICMOC, Procedia Engineering, 38 [2012], Page-2222-2227, DOI-10.1016/j.proeng.2012.6.267, 1877-7058 C- Published by-Elsevier,USA,
- [12] Prantosh Kr. Paul, K L Dangwal and Ramana Chettri, “ Quadrple Play Network: Emphasizing its possibilities for smarter University Educaation especially online knowledge delivery model” *Learning Community- International Journal*, ISSN-0976-3201 Vol. 4. No. 1, March, 2013 NewDelhi-Publishers,New-Delhi,[Indexed in EBSCO, Ulrich Directory, ICI, CAB, Proquest, Camell, ERIC, Index copernicus and other major databases]
- [13] Prantosh Kr. Paul, S Govindarajan, Dipak Chatterjee, “ Cloud Computing: Emphasizing Hybrid Cloud Computing on Android Computing Platform-An Overview” *International Journal of Applied Science and Engineering*, V.1, N1, ISSN-2321-0745, Page-21-28 New Delhi-Publishers,New-Delhi
- [14] Paul,Prantosh Kumar , R Rajesh, D Chatterjee, M K Ghose “Information Scientist: Technological and Managerial Skill requirement in 21st century” in *Information Studies* Vol. 19, No. 1, January, 2013, RCIS,Chennai, Page-29-36, ISSN-0971-6726
- [15] Paul, Prantosh Kumar, “MSc-Information Science [Geo Informatics]: Overview emphasizing twoproposed curriculum for sophisticated Geo Spatial development” *International Journal of Pharmaceutical and Biological Research (IJPBR)*”, Vol 4 Issue 5 Oct-Nov 2013, ISSN : 0976- 285X, Vol-218-227 [Indexed in DOAJ; 2. Ulrich’s Periodical Directory, USA; 3. EBSCO Publishing’s Electronic Databases, USA; 4. Indian Science Abstracts, India; 5.Index Copernicus, Poland. 6. NewJour,USA, 7. Google Scholar 8. Citeseerx]
- [16] Paul, Prantosh Kumar, “Environment and Sustainable Development with Cloud Based Green Computing:A Case Study” *Scholars Academic Journal of Biosciences (SAJB)*, 2013; 1(6):337-341 ISSN 2321-6883.
- [17] Paul, Prantosh Kumar, “Nutrition Information Networks: Possible domain and Future Potentials” *Scholars Academic Journal of Biosciences (SAJB)*, 2013; 1(6):342-345, ISSN 2321-6883
- [18] Prantosh Kr. Paul, K L Dangwal “Cloud Computing Based Educational Systems and iits challenges and opportunities and issues” *Turkish Online Journal of Distance Education-TOJDE* January 2014 ISSN 1302-6488 Volume: 15 Number: 1 Article 6, Page-89-98
- [19] Prantosh Kr. Paul, K L Dangwal, B Karn, “ Engineering Academics, Departments and Community : Emphasizing Some Educational Perspective of Information Science [IS], EDUCATIONAL QUEST: *An International Journal of Education and Applied Social Sciences*, 4(2): August, 2013: Page 141-146
- [20] Prantosh Kr. Paul, K L Dangwal, A Kumar “Information Infrastructure and Academic and Education World: The Role and Opportunities in Contemporary Perspective” *International Journal of Education for Peace and Development*: Vol. 1 No. 1 December 2013, Page-31-36
- [21] Reichman, F. (1961). Notched Cards. In R. Shaw (Ed.), *The state of the library art04(01)*, pp. 11–55). New Brunswick, NJ: Rutgers, The State University, Graduate School of Library Service.
- [22] Saracevic, T. (1996). Relevance reconsidered. *Information science: Integration in perspectives*. In *Proceedings of the Second Conference on Conceptions of Library and Information Science* (pp. 201–218), Copenhagen, Denmark: Royal School of Library and Information Science.
- [23] Saracevic, T. (1975). Relevance: A review of and a framework for the thinking on the notion in information science. *Journal of the American Society of Information Science*, 26(6), 321–343.
- [24] Saracevic, T. (1979a). An essay on the past and future of information science education. I. Historical overview. *Information Processing & Management*, 15(1), 1–15.
- [25] Saracevic, T. (1979b). An essay on the past and future of information science education. II. Unresolved problems of ‘extemalities’ of education *Information Processing & Management*, 15(4), 291–301.
- [26] Vakkari, S.P. (1996). Library and information science: Content and scope. In J. Olaisen, E. Munch-Petersen, & P. Wilson (Eds.), *Information science: From development of the discipline to social interaction*. Oslo, Norway: Scandinavian University Press.
- [27] Vickery, B.C., & Vickery, A. (1987). *Information science in theory and practice*. London: Butterworths.
- [28] Wersig, G., & Neveling, U. (1975). The phenomena of interest to information science. *Information Scientist*, 9, 127–140.
- [29] White, H.D., & McCain, K.W. (1997). Visualization of literatures. *Annual Review of Information Science and Technology*, 32, 99–168.
- [30] www.en.wikipedia.org
- [31] www.infosci.cornell.edu/
- [32] www.ischools.org
- [33] <http://www.libsci.sc.edu/bob/istchron/iscnet/ischron.html>