

CLUSTERING OF VIDEO INFORMATIONS USING BRICH METHODOLOGY

D. Saravanan

Faculty of Operations & IT, IFHE University, IBS Hyderabad

ABSTRACT

Information extraction one of the emerging field today because of the huge data sets. Creation and uploading information's are easy to any user, retrieve the relevant content is most challenge task. Due to that number of researchers and research work progress in this area. This extraction process required basic domain knowledge, extraction of image data sets this knowledge are highly essential. Extraction of video data are most challenge task to the user because of the complex nature of the video. In video data mining, video data are grouped into frames. The vast amount of video frames, fast retrieval of needed information is important one. In this paper, a Birch-based clustering method for content based image retrieval is proposed.

Key words---Data Mining, Clustering, Hierarchical Clustering, Video Data, Clustering Algorithm.

I. INTRODUCTION

Technology brings the data sets hugely from this data sets retrieve the relevant information's are challenge task for any user group. Extraction relevant information it is necessary to know the nature of the data sets, for this domain of image mining one of the most thrust area for many researchers. There are number of different technique and model available in this area, but all works well for certain type of image files only. Mostly used techniques are classified as with help of algorithm information are extracted or apply the input based on the content information's are extracted. For complex data type content based mining is one of the best technique than model based technique because it extracts information based on relevance of data available in the database. This process, reduce the searching time and brings the effective output.

Unsupervised methods are applied when evolving data is more difficult to detect is the main problem in data stream mining. We are going to clarify some points:

1. Data stream clustering problem definitions.
2. Research having specific difficulties.
3. Different approaches like assumptions,

approaches and heuristics and some solutions to solve many problems are the basis.

Clustering in data stream process are more complex than normal clustering technique because this data sets are nor recorded easily most of this type of data sets are dynamic in nature. Example temperature data sets, information's are recorded in one particular place not having consent value it gets changed over period of time. Size of data get increased due to variety of reasons, handling such large data we need a special mechanism. Storage is another biggest problem for such a complex data types. For this reason, the type of data sets is analyzed in dynamic way. Data are observed due to flow of information main drawback for this data sets user unable to visit again. Data pass through the analysis areas once if we need the same data point this technique is not applicable. This kind of problems are solved by data streaming technique, information's are monitored and recorded once this information's are flow form one side to other side. Due to this technique, facing some problem, and this technique never works like our traditional data mining technique. Data mining techni-

ques work based on stored data sets. Many research works are carried in the particular domain, number of methods and algorithms are developed to improve the quality of the research work. All this force to the researches to develop a new algorithms and technique in this domain. This stream-ing applications are used many of our real word life they are:

Communication and association field: Streaming techniques are very useful for monitoring the customer dialing information's, communication interchange, safety monitoring are effectively done by this technology.

Commerce operations: It industrial and commerce applications it helps Varsity of way in banking it helps to identify the customer card transaction are monitored, in financial operation like stock markets it helps to record the online stock exchange transaction, in production sectors good flow details are recorded effectively.

Healthcare sector: In most important in health care sector it helps to grouping the information based on the symptom, it helps to groups the information based on the observations.

Monitoring operations in Business: Many business transactions are monitored and recorded properly, this operation is done effectively by using this technique. It helps to find the maximum work load of the system, work task, load everything is recorded effectively. Discovering the evolution of workload in an e-commerce' server, which can help in dynamically fine tune the server to obtain better performance.

II. EXISTING SYSTEM

Technology brings the data sets more and more complex and high volume of data. Due to increasing in complex data sets there is no proper storage medium for analyzing this complex data sets. Existing technique information are flow from one side of the information point to moves other side so that data set which we need to compute again not possible. Because of dynamic nature it is very difficult. This brings the huge burden to the user some of the drawbacks are listed below:

2.1 Issues in Existing System

Existing technique brings various trouble they are categorized into two main drawbacks:

1) Data's are normally flow form one side of the input point to other side so that only once we can compute the data points.

2)Data are dynamic nature so that data evaluation change over period.

III. PROPOSED SYSTEM

Information are very similar or closer they are all grouped in one this is the basic concept user for clustering because of this property it is used in industrial and academic areas and it user for foundation steps for many procedures and methods. This technique is effectively used for may application such as image mining, video data mining, text mining, financial analytics, marking segmentation, health care industry like most of the places this technique works very nicely. Variety of clustering technique are available based on the usage this algorithm classified either bottom up approach or top down approach. It means small similar data points are grouped together and make first level groups, form this which groups are similar they form second level cluster like all data points merged into one big groups. This technique called bottom up approach, same technique in reverse process one big data points are divided and form single data points. It is called top down approach.

Clustering using bottom up approach not suitable for image type of information. Every clustering algorithms are suitable for certain applications only. Based on the information either text or information either bottom up approach or top down approaches are used. No single method best suited for all the process. Based on the information user must choose best clustering algorithm, Choosing the best algorithm itself challenge task to the user. Best algorithms produce good clustering else number of outlier produced by the technique so choosing the best one is more important task here. Bottom up approach is not suited for dynamic operation such as video type of dataset this technique never produced good results. Experimental results also verified this concept clearly. There are number of reasons are there for they never produce good result for video file every frame need to analyze first based on image analyze only information is stored, no single analyse produce good results. For video file every frame information ie pixel value, color value, motion time, character available in the frame all need to consider so that scanning of single frame take long time, due to that analyze process take more time. This is one of the major drawback for bottom up approach, frame separation alone takes long time, so time durations are high due to that this technique are not suited for video data files. Image analyze required backtracking operation ie some of the operation are repeated based on the finite solution most of the clustering process ie dived techniques never allows backtracking it leads difficulty to the user to obtain best results. Because of the above reasons image type of input files, it is required technique may combined with other clustering technique. Proposed technique works well for major type of video files,

experiments also verified proposed technique produce good result.

Proposed technique works well for even small amount of cluster also, suppose your data set size is increased even for that it produces good results. Techniques works well in online data sets or stored data sets. During the clustering formation this technique forms a triangular representation for the given data sets. Experimental results prove that proposed technique produce good competence and precision. Steps for forming clustering

Step 1: Form a shot from given video data file.

Step 2: Remove noise from the data sets.

Step 3: Apply Proposed technique.

Step 4: Technique analyze the input file and formed tree structure.

Step 5: Second time analyze remove outlier from the tree structure.

Step 6: process repeated for rest of the data set.

IV IMAGE MINING ALGORITHM STEPS

Proposed technique works well in image type of input data set. First the input video files are converted into shots, shots are meaningful information for video data files. Splitted video files consisting of lot of repeated and noisy data's, it is removed from the data set else due to noise processing time of clustering get increased. Detailed steps are shown below.

1.Characteristic Extracting: Every multimedia frames combination of various type of information's such as audio, text and motion. It is necessary to first extract the basic information from the input file. Dynamic video files are divided into meaningful segments. Each segment is treated as one object or one frame.

2. Frame recognition: Every frame gives useful information to the user community, it is necessary to differentiate the one frame with other. Frames are properly analyzed and find the difference between the frames. Finally, each frame is assigned with unique identifier. It helps to recall the frames easily. In second step need to create the identifier to the input frames.

3.Forming relationship: Find the relationship between the frames and assign a common group.

4. Use the proposed technique find the best output ie clustering group.

IV. EXPERIMENTAL OUTCOMES



Fig 1. Open Image



Fig 2. Clustering Duplicate Frame Elimination
Table1: Duplication Elimination

Video name	Number of Input	Number of output	Duplicate frames
cartoon	7	6	1
Graphics	16	10	6
Meeting	15	14	1
Globe	15	13	2
Song	15	14	1

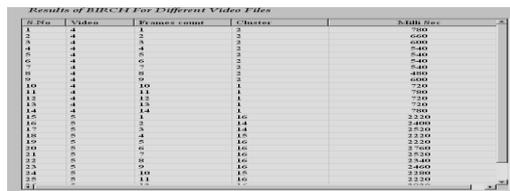


Fig 3. Cluster Formation

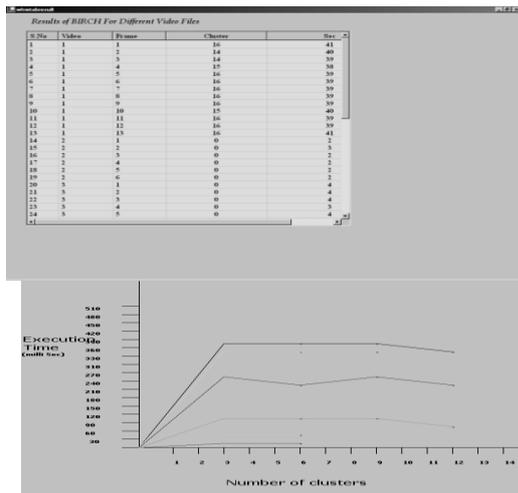


Fig 4. Output Comparison of Five Video file



Fig 5. Brich Cluster formation Time(millisecond)
Fig 6 Compare the frame

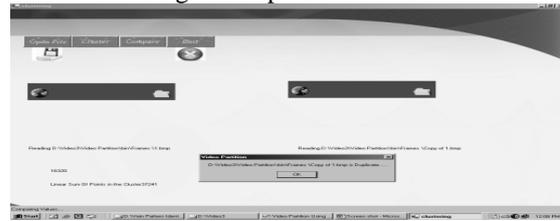


Fig 7. Comparison Result of frame

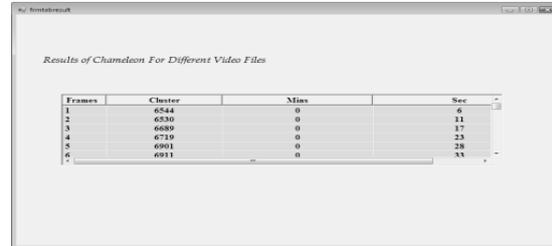


Fig 8. Brich cluster Result

V. CONCLUSION

This paper brings the difficulties for clustering formation in video data files. Most of the existing techniques works well for specific video files only it gives hug attention to the research community to bring new techniques in this specific area. Algorithm efficiency are measured based on the input data sets. Data set consisting of huge noise and repeated data points clustering formations are good. It is necessary to clan the data set before it processed, based on cleanness outputs get vary. Proposed technique overcome the drawback of the existing clustering techniques. Proposed technique works on two phased approach. Experimental results proved proposed technique bring good result.

VI. FUTURE ENHANCEMENTS

Proposed technique works based on some of the image property ie image pixel value, time and content availability Technique may further extend with extraction of images based on other image properties.

REFERENCES

- [1] D. Saravanan, S.Srinivasan, Video Data Mining: Imaged Retrieval Using Histogram Difference, *In the Proc. Of the National Conference on Recent Trends in Computer Science and Engineering NCRTCSE-10*, Pp. 147-149 (2010).
- [2] D. Saravanan, S. Srinivasan, A Study of Hierarchical Clustering Algorithms suitable for video data mining, *ICOM* (2011).
- [3] D. Saravanan, S. Srinivasan, Video Data Information Retrieval Using Data Mining *Concepts in the Proc. Of the National Conference on Computer and Communication Technology NCCCT-11*, PP. 96-98 (2011).
- [4] D. Saravanan, Pervise initiated peer distribution multiplexing. *Pak. J. Biotechnol. Vol 13(Special Issue)* 347-350 (2016)
- [5] Martin Ester, Hans-Peter Kriegel, Jorg Sander and Xiaowei Xu, A density-based algorithm for discovering clusters in large spatial database with noise *In I&I Conference on Knowledge Discovery in*

Databases and Data Mining (KDD-96), Portland, Oregon (1996).

- [6] D. Saravanan, Information Retrieval using: Hierarchical clustering algorithm. *International journal of Pharmacy & Technology* 8(4): 22793-22803 (2016)
- [7] D. Saravanan, Effective Video Data Retrieval Using Image Key frame selection. *Advances in Intelligent Systems and computing* Pages 145-155 (2016).
- [8] Martin Ester, Hans-Peter Kriegel, and Xiaowei Xu, A database interface for clustering in large spatial databases, *Int'l Conference on Knowledge Discovery in Databases and Data Mining (KDD-95)*, Montreal, Canada, August (1995).
- [9] Tian Zhang, Raghu Ramakrishnan, and Miron Livny, Birch: An efficient data clustering method for very large databases, *Proceedings of the ACM SIGMOD Conference on Management of Data*, Montreal, Canada, Pp. 103-114 (1996).
- [10] D. Saravanan, Organized and protected communication in mobile group. *International Journal of Control Theory and Applications* 9(24):9-16 (2016).
- [11] D. Saravanan, Design and implementation of feature matching procedure for video frame retrieval. *International Journal of Control Theory and Applications* 9(7): 3283-3293 (2016).
- [12] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, The R*-tree: an efficient and robust access method for points and rectangle, *Proc. Of ACM SIGMOD*, Atlantic City, NJ, Pp. 322-331(1990)