TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 14574 - 14579

Efficient image Data Extraction using Image Clustering Technique

D.Saravanan[1], Dr. S. Nirmala Sugirtha Rajini[2], Dr. K. Dharmarajan[3]
[1]Faculty of Operations & IT, ICFAI Business School (IBS), Hyderabad, The ICFAI Foundation for University, Higher Education (IFHE), (Deemed to be university u/s 3 of the UGC Act 1956), Hyderabad-India.
[2]Professor, Department of Computer Applications Dr.MGR.Educational and Research Institute, Maduravoyal, Chennai,
[3]Associate Professor, Department of Information Technology, VISTAS, Vels University, Chennai.
sugirtharaja77@yahoo.com[2], dharma.scs@velsuniv.ac.in[3]

Article Info
Volume 82
Page Number: 14574 - 14579
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 28 February 2020*

Abstract:
Extracting the domain knowledge from huge data sets are more complex today because of growing technology. This technology brings huge sets of data from every users. This gives the complex data sets, from this content extraction of needed knowledge are challenging task for many users. The proposed frameworks achieves full automation via a knowledge-based video indexing and retrieve an appropriate result, and replace a presented object with the retrieval result in real time. Along with this indexing mechanism a histogram-based color descriptors also introduced to reliably capture and represent the color properties of multiple images. Including of this a classification approach is also carried out by the classified associations and identified by proper tickets. Proposed techniques works well in all type of data sets. Experimental outputs verified this results.

Keywords: *Image extraction, Pixel comparison, threshold values, Fame grouping, Frame classification, Data extraction, Frame sequences.*

I. INTRODUCTION

Extracting the need domain knowledge from the complex data sets are more thirsted area today for most of the researchers. This files gives different meaning retrieval of document, image, video or audio file. This retrieval process now it is most challenge and complex operation for any researchers because of increasing the data size. Increasing the demand of multimedia data sets allows every users to create and store the data effortlessly. This causes increasing the unstructured data sets. From this data sets extracting the needed are challenging task for many researchers. Data's like multimedia is the combination of all media, from this type of data file it is really a challenge task to the user to bring the needed information. Today technology allows the users to extract the needed content by using any image attributes such as image, pixel, text, motion, time, frame and any image features. Users gets the flexible to use any one of the image attribute as a key elements to search the needed content from the data repository. Extracting the right knowledge users use different methodology and different procedures. Extracting the information first user need to arrange the, data's in a proper form it easy to the user bring the needed information in single search. For this data are cluster and store. From this complex data sets user need to create a grouping, it helps to extract the needed information from the specified group sets. Creation of grouping is done by one of the image attributes help of this user need to create a smaller groups it reduce the searching time and also bring the accurate data retrieval. In each group closeness of items which present in the group are higher compare with the other groups item sets. Creating of grouping various techniques are available. Each technique works well in any one type of image

January - February 2020
ISSN: 0193 - 4120 Page No. 14574 - 14579

features. None of the techniques not works for more than one image attributes. Example some grouping done by image pixel values are similarities between two pixels. Here extraction or grouping based on the image pixel only works well. In this group user want to extract any text or motion based information user won't get proper knowledge. This grouping works only for pixel based operations. Like based on the group creation information's are extracted. From this user can extract only based on the group and the procedure works well based on the grouping property. For that reason today most of the image mining technique required a special attention to retrieve the need knowledge.

A. Disadvantage of surviving structure

Existing technique unable to find the exact cluster from the data points.
Cluster formation take longer time duration.
Showing poor performance identify the outlier.
Statistical not improves the quality of clustering.
Well situated for specific video files, not suitable for all set of video files.
Well formation of cluster in some set of video files only.

II. PROPOSED METHODOLOGY

The proposed clustering called CURE that eliminate more outlier and produce good cluster results. CURE performs its operations differ from other clustering technique it first scan the image and identified specified data points in the image and picking the selected points done through scanning. After collecting those points it check the similarities between the points generated, points are closer or similar points are collected and grouped separately. Points which are not similar are far away from the generated points are grouped separately. Like the entire image frames are scanned and every time of iteration like various groups are formed. After performed the groups again the big scanning among the groups are done. This algorithm works well in large data base that is image size is more and complex. The process are performed in two step process in step one random grouping are done in the second scanning it generate the desired cluster based on the image quality.

A. Advantages of projected arrangement

Creating more cluster during the small amount of time eliminating the outlier

The algorithm is robust to the presence of outliers.

Forming the cluster more highly compare the existing techniques.

Works well in all type of video files.

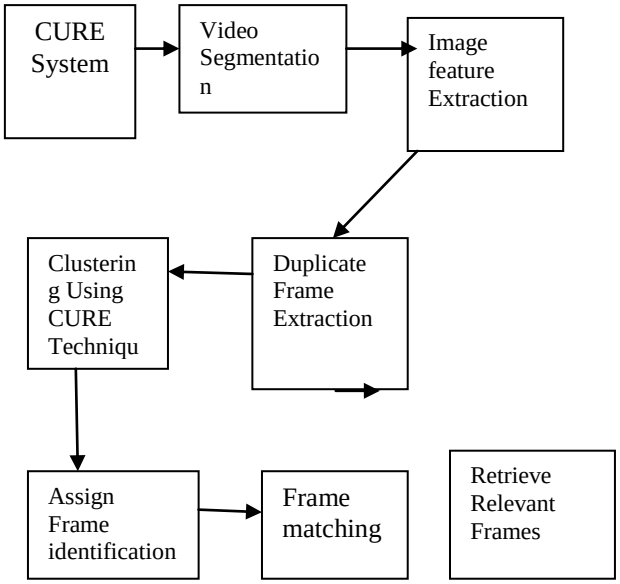

Fig. 1. Block Diagram for proposed technique

III. EXPERIMENTAL SETUP

A. Image mining

Image extraction or image knowledge extraction are complex operation in data mining. Where never user want to extract the needed information from the complex data sets user need to consternate two important information in his process. First after converted into static images every image frames are identified by a common identifies. Naming this identifier one of the most important task if you assign same identifier for two UN similar frames it gives confusion during the extraction process. For that naming the object one of the most important

TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 14574 - 14579

task in image mining. [7] .Second which domain this image extraction are done it also consider in detail. Because information always extracted from complex data sets for that user need to identify from which domain or which part this application are used. Sometime naming the identifier based on the image attribute sets. For that reason the domain also consider very seriously. In image mining user need to specify the query using image attribute values. This value are check with the stored data base values. Based on the similarity measure information are extracted and returned to the user.

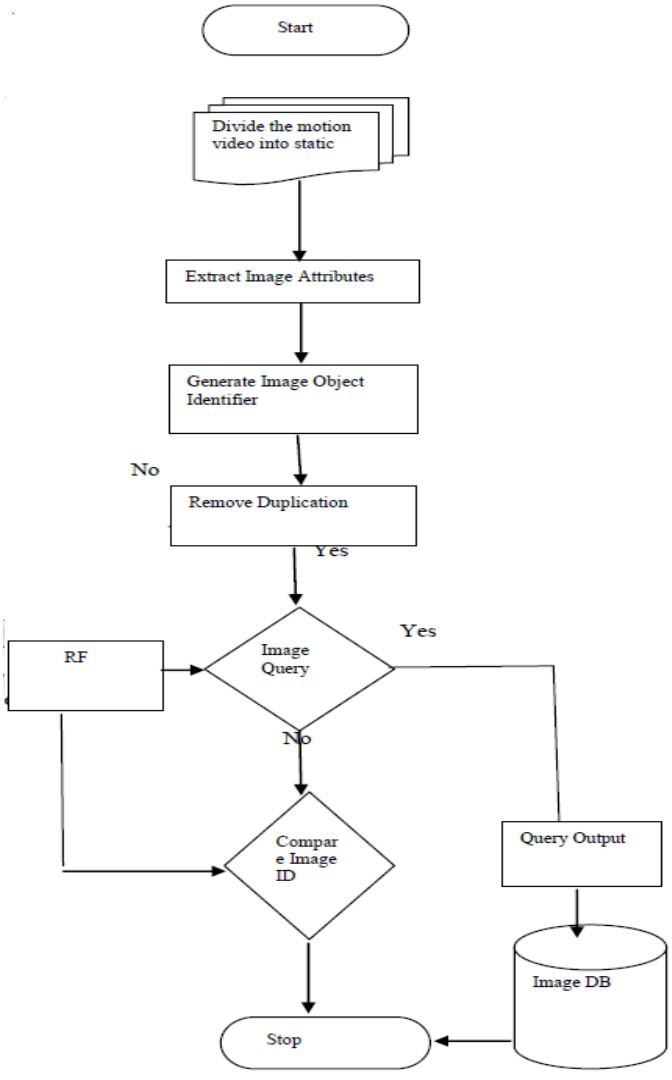

Fig. 2. Flow chart for the overall process

B. Information extraction procedure:

Extracting the needed information from the stored data sets based on the following steps:

1. Feature Extracting: Video files are divided in to frames each frame image feature are extracted.
2. Object identification and record creation: Compare one frames with other frame each frame features are recorded with separate frame identification.
3. Apply image mining steps to produce object

C. Video Classification

Ordering the image data is one of the predominant steps in image mining. Here information are extracted based on image attributes or image properties. Information's are collected from various data sources each sources of information some common attributes are identified and arranged in the common name. This name helps the user to extract the information it also helps to reduce the searching time. For that the collected information's are arranged in the unique name based on the image attributes. Initially this process started with dynamic image is converted into static image then unwanted images are removed using image attribute values. This process show in the fig 2 and fig 3.

D. Video Clustering

Image grouping can be done with help of image any of the image attributes. Like image pixel values. Image pixel average value, difference between pixel, pixel density in the frame, time duration between the frames, text position in the frame like number of properties are available. Using any of this property user normally create clustering. Data points with the cluster are very closer and the data points of one cluster to other cluster having more difference. Data points are very similar or very closer points are collected and grouped under common name or common identifier. User can use this name or identifier to retrieve the data sets. This grouping helps the user to get some common or get some knowledge based on the users input query either in the form of text based or image based input query. Normally this grouping can done based on

TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 14574 - 14579

the similar image properties. Whenever we make similar cluster the cluster are grouped under common values or common nature. It reduce the searching time and also brings the effective result while user searching any specific information form this domain. Here we constructed image cluster this cluster differ from the tradition text or other property based clustering. Because based on the various image properties which mentioned the above the cluster get formed here. Before we can make image cluster first the dynamic videos sets are converted into static images. After the image preprocessing steps are applied to remove the unwanted and noisy data object. After purifying the objects any of the image properties are consider to form the clusters. Normally this process takes more time than the traditional clustering operations. This operations are listed in the fig 4 & Table 1, 2 and 3.

Pseudo code for removing outlier

```
loop i = 0 To i1 - 2
   loop  j = 0 To j1 - 1
      Edg1 = DisplayBM.GetPixel(i, j)
      Edg2 = DisplayBM.GetPixel(i, j + 1)
      If Averageimagevalue < Edg1.G Or
Averageimagevalue < Edg2.G Then
         If Edg1.G <> Edg2.G Then
            DisplayBM.SetPixel(x,y,
ColorTranslator.FromWin32(RGB(255, 0, 255)))
         End If
```

IV. RESULT ANALYSIS

Today the data sets are complex in nature data's consisting of combination of everything from this complex data sets extracting of needed content are challenging task for many users. In the proposed method we are using video data sets. In any type of video file the proposed method forming more number of cluster compare with the existing procedure. In this all video file, which data files are giving more clustering with minimum time? This work can extended in future as how the cure clustering gives Different solution to all video data file. And performance of all video data file in graphic representation.

Fig. 3. Selecting the user query input

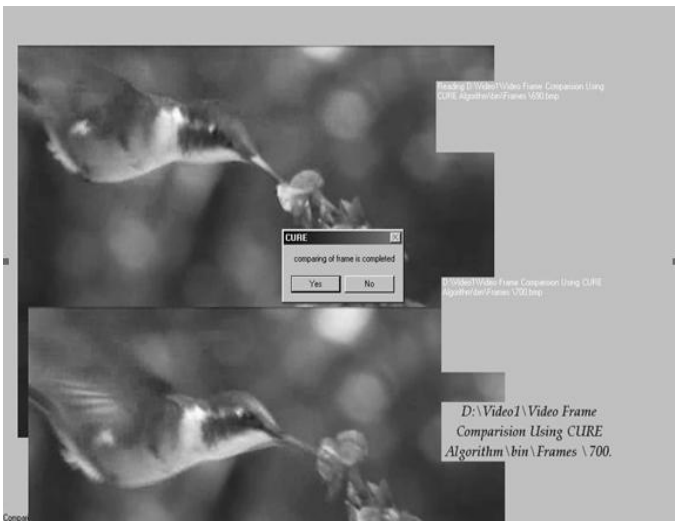

Fig. 4. Assessment done through deposited values set process

January - February 2020
ISSN: 0193 - 4120 Page No. 14574 - 14579

Cluster	Mini	Sec
7881	0	8
8015	0	6
8015	0	5
8221	0	7
8965	1	8
7895	0	6
7889	1	8
8558	1	4

Fig.5. Result cluster and time taken

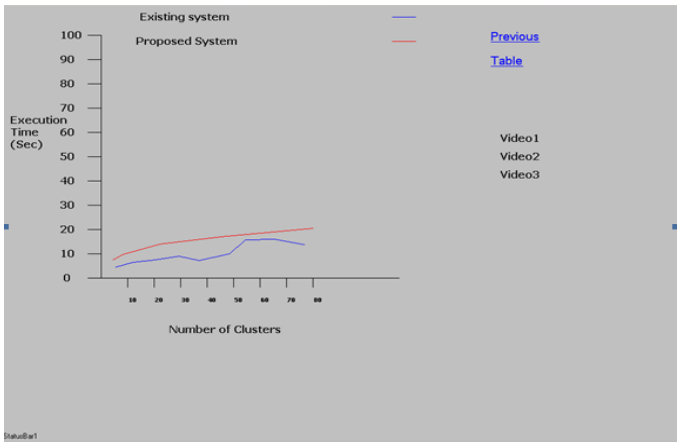

Fig. 6. Performance Graph Cure

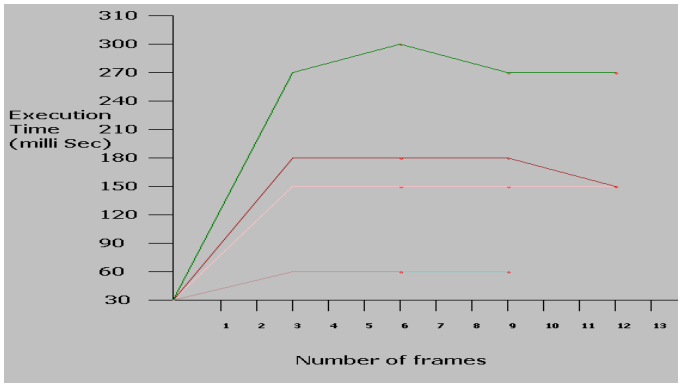

Fig. 7. Ouput for differnet video files.

Cyan- Forest file , Brown- Animated file , Rosy brown-Debat file , Red- Motion Green- Birds groups file.

TABLE 1
Grouping Vs Duration

Video frame type	Frames	Cluster	Time(msec)
1	1	08	120
1	2	12	140
2	1	07	134
2	2	14	150
3	1	06	124
3	2	10	132
4	1	09	120
4	2	14	140

TABLE 2
Frames Vs Duration

ID	Frames	Number Cluster	Time(msec)
1	0	7589	4:16:02
2	1	7548	4:16:10
3	2	7452	4:16:07
4	3	7524	4:16:11
5	4	7852	4:16:06
6	5	7125	4:16:05
7	6	7458	4:16:12
8	7	7536	4:16:08
9	8	7892	4:16:01
10	9	7412	4:16:44
11	10	7521	4:16:03

V. REFERENCES

[1]. D.Saravanan", Information extraction using user opinion procedure", International journal of recent technology and engineering (IJRTE), Pages 10403-10407, Volume 8, Issues-4, Nov 2019

[2]. D.Saravanan", Image substance extraction using data mining clustering method" International journal of innovative

January - February 2020
ISSN: 0193 - 4120 Page No. 14574 - 14579

technology and exploring engineering (IJITEE), Pages 2735-2739, Volume 9, Issues 2, Dec-2019

[3]. A. Imran, A. Moreno, and F. Cheikh, "Exploiting visual cues in non- scripted lecture videos for multi-modal action recognition," in Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on, 2012, pp. 8–14.

[4]. D.Saravanan, Dr. Vaithyasubramaniyan, Dr. Dennis joseph "Effective Utilization of image information using Data mining technique" , Intelligent systems reference library 172, Volume 172, Pages 207-215, ISBN 978-3-030-32643-2, Nov 2019

[5]. Vaithyasubramanian, S., & Christy, A. (2015). A scheme to create secured random password using Markov chain. In Artificial Intelligence and Evolutionary Algorithms in Engineering Systems (pp. 809-814). Springer, New Delhi.

[6]. D.Saravanan,"Issues and considerations for effective text data retrieval", International journal of innovative technology and exploring engineering(IJITEE), Pages 1442-1445, Volume 9, Issues 1, Nov 2019.

[7]. T.-J. K. P. Monserrat, S. Zhao, K. McGee, and A. V. Pandey, "Notevideo: Facilitating navigation of blackboard-style lecture videos," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2013, pp. 1139–1148.

[8]. W. Hu, N. Xie, L. Li, X. Zeng, and S. Maybank, "A survey on visual content-based video indexing and retrieval," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 41, no. 6, pp. 797–819, 2011.

[9]. D.Saravanan, Multimedia data Retrieval Data mining image pixel comparison technique", Lecture notes on Data Engineering and communications Technology 31,Aug 2019, Pages 483-489, ISBN 978-3-030-24642-6, Chapter 57

[10]. Vaithyasubramanian, S., Christy, A., & Saravanan, D. (2016). Access to network login by three-factor authentication for effective information security. The Scientific World Journal, 2016.

[11]. Pathirage Kamal Perera. "Traditional medicine-based therapies for cancer management." Systematic Reviews in Pharmacy 10.1 (2019), 90-92. Print. doi:10.5530/srp.2019.1.15

[12]. Vishnu, V., Shefrin, S., Sreelaxmi, C.S., Nair, S.C.Aquasomes: An excellent and promising system for novel drug delivery(2018) International Journal of Pharmaceutical Research, 10 (2), pp. 207-215. https://www.scopus.com/inward/record.uri?eid=2s2.085048042893&partnerID=40&md5=12f4c117a60fd6ed1d9d6effc07692f9

[13]. Joseph bell (2018) anti-oxidant nutrition: an adjuvant with a balance diet. Journal of Critical Reviews, 5 (1), 4-. doi:10.22159/jcr.2018v5i1.23343

[14]. Monterosso, D.M., Kumar, V. and Zala, K., 2019. Spiritual Practices in The Era of Smartphones & Social Networking: A Comparative Study. International Journal of Psychosocial Rehabilitation. Vol 22 (2) 45, 57.

[15]. Shafti, S.S. and Ahmadie, M., 2018. Improvement of Psychiatric Symptoms by Cardiac Rehabilitation in Coronary Heart Disease Vol 22 (2) 80, 89.

[16]. Bonsaksen, T., Opseth, T.M., Misund, A.R., Geirdal, A.Ø., Fekete, O.R. and Nordli, H., 2019. The de Jong Gierveld Loneliness Scale used with Norwegian clubhouse members: Psychometric properties and associated factors Vol 22 (2) 88, 100.