**JOURNAL OF CRITICAL REVIEWS**

ISSN- 2394-5125 VOL 09, ISSUE 01, 2022

# Advanced data extraction and classification using image frame assessment technique

**R.Manthru Naik**,
Asst. Prof., Assoc.Prof. dept of CSE, Malla Reddy Institute of Technology, Kompally, Maisammaguda, Dulapally, Secunderabad, 500100
**Dr.D Narahari**,
Assoc.Prof., Assoc.Prof. dept of CSE, Malla Reddy Institute of Technology, Kompally, Maisammaguda, Dulapally, Secunderabad, 500100
**Dr.Lalitha kumari pilli**,
Assoc.Prof., Assoc.Prof. dept of CSE, Malla Reddy Institute of Technology, Kompally, Maisammaguda, Dulapally, Secunderabad, 500100

**Abstract:**

Database technology developments that include multimedia data are needed to handle the ever-increasing amount of multimedia data. The development of digital media has resulted in a major shift in extraordinarily large multimedia databases (digital camera, digital video and digital television, e-book and mobile phone, etc). It is the major objective of this study to examine and enhance the present photo retrieval system.

**Keywords:** Information Retrieval, Image pixel comparison, Pixel threshold value, Image frame value, Image mining, Image frame comparison.

## 1. INTRODUCTION

There are many different types of retrieval techniques used in the field of information retrieval (IR), but they all revolve on searching for information in and within documents, including metadata about those documents. Even though the phrases "data retrieval" and "document search" are sometimes used interchangeably, each term has its own body of literature, theory, practice, and technology. Data retrieval is one of them. Computing, mathematics, library science, information science, and architecture all play a role in information retrieval (IR), which is an interdisciplinary field that draws on a variety of other fields as well. In the last 10 years, worldwide research has focused on content-based video analysis and retrieval as an essential technology. In order to effectively retrieve specific information based on video content and to provide better ways for entertainment and multimedia applications it is urgently needed the advanced technologies for organizing, analyzing, representing, indexing, filtering, retrieving and mining the vast amount of videos.

### 1.1 Image Retrieval

Image mining necessitates the retrieval of pictures based on predetermined criteria. There are three stages of increasing complexity to the requirements specifications [1]:

a. Color, texture, form, and placement of picture parts are some of the fundamental attributes that may be used to retrieve images at Level 1.

b. Image retrieval based on derived or logical properties, such as a specific kind of item or particular object or person, is at the second level.

c. Level 3 image retrieval entails a large degree of high-level thinking about the meaning or purpose of objects and situations portrayed in the images.

## 2. THE FOUNDATIONS OF DATA MINING

There has been a protracted process of study and development that has led to data mining methods. First, computers were used to store data, then access to that data was improved, and finally technologies were developed that enable users to interact with their data in real time. Data mining moves this evolutionary

ISSN- 2394-5125 VOL 09, ISSUE 01, 2022

process from a focus on accessing and navigating past data to a focus on anticipating and delivering future information. There are three technologies that are developed enough to facilitate data mining in the business community:

- A massive amount of data gathering
- Multiprocessor machines that are very powerful.
- Algorithms for data mining


**Figure 1. Data Mining Preprocessing Steps**

### 2.1 Research issues in image mining

Image mining, by definition, is the process of identifying patterns in a large number of photographs. Because image mining focuses on identifying patterns in a vast collection of photos, it is distinct from low-level computer vision and image processing approaches, which concentrate on interpreting and extracting characteristics from a single image. When it comes to working with a huge collection of photos, image mining and content-based retrieval seem to have certain similarities, but image mining goes far beyond. The purpose of image mining is to find interesting patterns in a collection of photographs. Perhaps the most widespread misperception about picture mining is that it is just a matter of reusing current data mining methods to analyze photographs.This is not true, since relational databases and image databases have significant differences.

#### 2.1.1 Absolute versus relative values.

These semantically relevant data values may be found in relational database management systems. For instance, it's well-known that 35 is a mature age. Although the data values itself may not be relevant in an image database, the context could. It is possible for a grey scale value of 46 to look darker than an equal value of 87 in the context of a brightly lit scene.

#### 2.1.2 **Spatial information**:

Unlike relational databases, image databases require implicit spatial information for interpretation, whereas relational databases don't. This is another important difference between relational databases and image databases. So, image miners aim to circumvent this challenge by first extracting position-independent characteristics from pictures before trying to mine valuable patterns from the images themselves.

#### 2.1.3 **Unique versus multiple interpretation,**

ISSN- 2394-5125 VOL 09, ISSUE 01, 2022

Image features that allow for a variety of interpretations of the same visual patterns are a crucial third difference. A pattern-to-class association (interpretation) technique used in conventional data mining will not function effectively here. To meet the unique requirements of mining valuable patterns from photos, a new class of discovery algorithms is required. For image pattern mining to be successful, new discovery algorithms for mining patterns from image data must be developed.

## 3. CONTANT BASED IMAGE RETRIEVAL SYSTEM

The use of computer vision to the image retrieval issue is known as content-based image retrieval (CBIR), also known as query by image content (QBIC), and content-based visual information retrieval (CBVIR). Higher- and lower-level vision concepts are combined in this approach. Semantic information about an image's structural content may be gleaned by higher-level analysis techniques such as perceptual organization, inference, and grouping. Image texture, shape and color histograms are used in lower-level analysis. There are two sorts of picture searches: target and category. Searching for specific images, such a registered brand, a historical snapshot or an individual artwork is the purpose of target search. A category search is used to locate photographs that are relevant to the user, such as images of landscape or skyscrapers that they may not have known about before.


**Figure 2. Context based Retrieval**

### 3.1 Image-based applications

The development of image-based computer applications followed the development of text-based computer applications because of computer capacity constraints. Weather forecasting was one of the first uses of image-based computer applications, but other early uses included:

- Creating and modifying maps
- Chemistry, physics and statistical analysis may all benefit from visual examination of experiment results.
- Mathematical functions may be shown
- Architecture, landscaping, interior design, fashion, and other creative fields were among the first to use CAD/CAM (computer-aided design and manufacture) in shipbuilding.
- Image-based applications have developed in scope in tandem with computer power, and now include:
- Weather, temperature, agriculture, human activities... can all be tracked via satellite imagery analysis.
- Medical image analysis
- Submarine, ship, and aircraft navigation

ISSN- 2394-5125 VOL 09, ISSUE 01, 2022

- Fingerprints and a picture ID are required for security reasons

Art, cultural and natural science institutions have discovered new consumers for their picture data as broadband Internet capability has become more widely accessible.

### 3.2 Image Mining Algorithm Steps:

Mining associations in the context of a picture requires algorithms. The following are the four main stages in the image mining process:

1. A picture is segmented into areas that are identified by region descriptors, or "blobs," each of which is meant to represent an individual item.

2. Comparing one picture to another for item identification and record-keeping: Assign an identifier to each item. The preprocessing algorithm is the name given to this procedure.

3. Images with recognized items may be generated in this step to help you understand the association rules.

4. Create an item using a data mining technique.

### 3.3 Image Indexes

At least the first three of the four primary index kinds stated above, listed below in decreasing frequency:

1. Atomic indexes, used for basic RDB characteristics providing context (photographer/artist, owner) or structural information (encoding type, size).

2. Semantic metadata elements such as title, caption, topic, and description may be used to construct term indexes based on semantic metadata attributes such as these:

3. Low-level properties such as colour, texture, and form may be automatically retrieved and used to build structural feature indexes.

4. "Exam" and "prayer" are examples of high-level semantic features that may be used to create semantic feature indexes.

## 4. VIDEO DATA MINING

### 1. Video shot detection.

For video feature extraction, shot boundaries are the fundamental units of video parsing, which is why shot detection is the initial step. A multi-filtering architecture, which included pixel-level comparison, histogram comparison, and segmentation maps, was used in this research to improve upon prior work.

### 2. Indexing Of Frame

Indexing is a critical component of any data retrieval effort. As previously said, there are a variety of systems that use various methods to retrieve work and the recovered picture may be the same query image, a portion of the query image or even individual objects in the image. If you want to get the 10 most similar pictures to a specific image, you may include the term "next-neighbour" in your query. The performance of current systems is inversely proportional to the size of the database. If the system is to be efficient, it should be based on how many photos are similar rather than the overall number of images.

### 3. Color histogram

In order to create an image's colour histogram, one must first quantize the image's colours and then count the number of pixels in each colour. An picture may be projected into three colour channels using a colour space

ISSN- 2394-5125 VOL 09, ISSUE 01, 2022

(such as YUV) (Y, U, and V). Consequently, each picture is separated into three colour components, each of which may be viewed as an image with grey levels in accordance with a certain colour channel. The histograms of the image's colour components may be used to generate the feature vector. To create a histogram for a grayscale picture, the bin number, N, must be specified in advance; each pixel in the image is then placed into the bin whose centre is closest to the value of that pixel. The histogram's bar value is determined by counting the pixels in the bin. After everything is said and done, a feature vector of an image may be derived from the histogram's values, which correspond to the feature vector's coefficients.

## 4. Experimental Results:

**Table 1 Different Frame splitting time**

| Number Of Frames Spitted | Search Time |
|---|---|
| 100 | 2 Seconds |
| 1000 | 5 Seconds |
| 10000 | 14 Seconds |


**Figure 3. The model input frames of different types of test videos**

**Figure 4. Example Content Based Image Retrieval**

ISSN- 2394-5125 VOL 09, ISSUE 01, 2022


**Figure 5. Cartoon Input Image Vs Ouput image extrcted.**


**Figure 6. Sports Input Image Vs Output image extracted.**

**Table 2. Input frame Vs Time taken**

| frmcnt | milliseconds | categor |
|---|---|---|
| 25 | 789 | Cricket |
| 50 | 785 | Cricket |
| 75 | 780 | Cricket |
| 100 | 756 | Cricket |
| 152 | 705 | Cricket |
| 125 | 758 | Cricket |

ISSN- 2394-5125 VOL 09, ISSUE 01, 2022

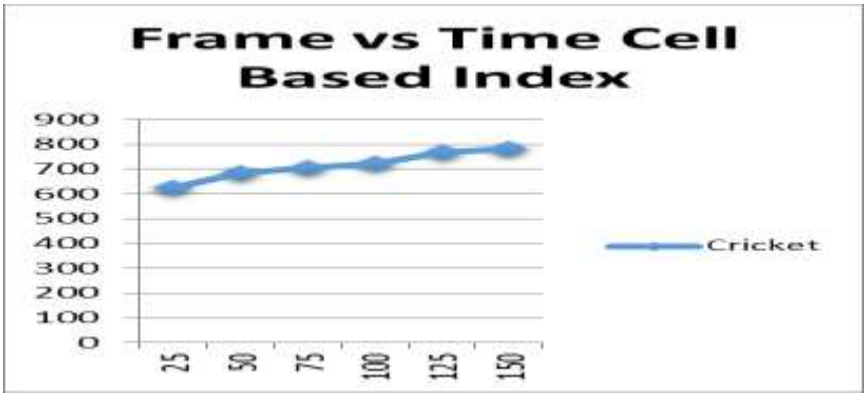


**Figure 7. Performance graph of Sports video file**

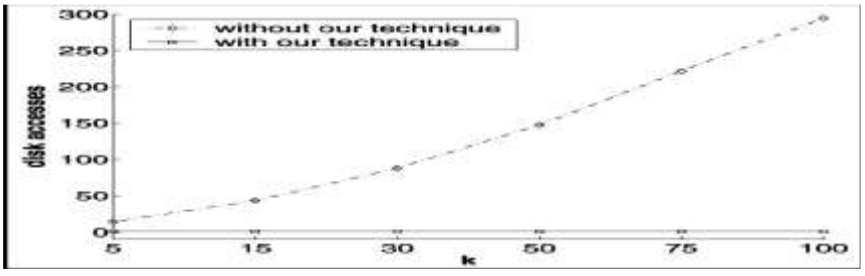


**Figure 8. Sampling Query Comparison**

## 5. CONCLUSSION

Large data sets are subjected to data mining as a means of uncovering patterns and information that may otherwise go unnoticed. Alpha-numeric databases are the primary focus of data mining research, whereas multimedia data mining has received less attention. An early stage of multimedia data mining, video data mining is examined for its present state and problems in this study. In order to mine the large volumes of video data for useful information, it is imperative that the difficulties outlined above be addressed.

## REFERENCES:

[1] D.Saravanan, Dr.S.Srininvasan " Query Point Techniqeus for CBIR Systems ", In the Proc. Of Advances in Image Processing, Mining Computing Technology", Pages 425-432,2010.

[2] 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

[3] Z. Barceló's, E. L. Flores; C. A. Z. Barceló's, S. F. Silva, M. A. Batista," A Multi-Dimensional Similarity Modeling and Relevance Feedback Approach for Content-Based Image Retrieval", Systems, Signals and Image Processing, 2009,Page 1-5.

[4] Sun Yen , Wang Zheng-xuan , Wang Dong-mei "An Image Retrieval Method Based on Relevance Feedback and Collaborative Filtering", 2010.

[5] D.Saravanan " Effective text data extraction using Hierarchical clustering technique" Int. Journal of Engg., and Advanced Technology(IJEAT),Feb 2019, Vol. 8, Issue 3-S, Pages 741-743(ISSN 2249 8658)June 2019

[6] Yihun Alemu, Jong-bin Koh, Muhammed Ikram, Dong-Kyoo Kim "Image Retrieval in Multimedia Databases: A survey", Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing,2009

[7] D. Hilbert, “Uber die stetige Abbildung einer Linie auf ein Flachenstuck,” Math. Annalen, 38-40. [10]I. Bartolini, P. Ciacci, and F. Waas, “Feedbackbypass: A New Approach to Interactive Similarity Query Processing,” Proc. 27th Int’l Conf. Very Large Data Base (VLDB ’01), pp. 201-210, 2001.

[8] R. Brunelli and O. Mich, “Image Retrieval by Examples,” IEEE Trans. Multimedia, vol. 2, no. 3, pp. 164-171, 2000.

[9] D.Saravanan, ”Information retrieval using image attribute possessions” Soft computing and signal processing , Advances in Intelligence systems and computing 898, Springer. DOI:10.1007/978-981-13-3393-4_77, Pages 759-767. March 2019.

[10] Fan, J.. Luo, H. "'Emantic Video Classification by Integrating Flexible MixtLre Model witlh Adaptive EM Algorithm",ACMSIGMM, 2003, pp.9-16.

[11] T. Gevers and A. Smeulders, “Content-Based Image Retrieval: An Overview,” G. Medioni and S.B. Kang, eds., Prentice Hall, 2004.

[12] MAGE MINING: ISSUES, FRAMEWORKS AND TECHNIQUES i Zhang Wynne Hsu Mong Li Lee Department of Computer Science, School of Computing National University of Singapore Singapore, 117543

[13] J. Zhang, W. Hsu and M. L. Lee. An Informationdriven Framework for Image Mining, in Proceedings of 12th International Conference on Database and Expert Systems Applications (DEXA), Munich, Germany, September 2001.

[14] Zhang Qi-dong, Wu Jian-hua, Gao Li-qun. Medical image retrieval based on non subsampled Contourlet transform and Zernike moments [J]. Chinese Journal of Scientific Instrument, 2009, 30(6): 1275-1280.

[15] Y Lu , C Hu, X Zhu et al1 A unified semantics and feature based image retrieval technique using relevance feedback1 In : Proc of the 8th ACM Int’l. Multimedia Conf. Los Angeles: ACM Press, 2001.

[16] D.Saravanan,” Efficient Video indexing and retrieval using hierarchical clustering techniques”, Advances in Intelligence systems and computing, Volume 712, Pages 1-8, ISBN:978-981-10-8227-6, Nov-2018

[17] A. Vailaya, M. Figueiredo, A. K. Jain, and H.J. Zhang, Image Classification for Content-Based Indexing , IEEE Transactions on Image Processing, vol. 10, no. 1, pp 117-130, January, 2001.

[18] 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.