A Systematic Review of Existing Data Mining Approaches Envisioned for Knowledge Discovery from Multimedia

Benaka Santhosha S1, Chitra Kiran N2

1Department of Electronics & Communication Engineering, Coorg Institute of Technology, Kodagu, Karnataka, India
2Department of Electronics & Communication Engineering, Sai Vidiya Institute of Technology, Bengaluru, Karnataka, India

ABSTRACT

The extensive use of multimedia technologies extended the applicability of information technology to a large extent which results in enormous generation of complex multimedia contents over the internet. Therefore the number of multimedia contents available to the user is also exponentially increasing. In this digital era of the cloud-enabled Internet of Things (IoT), analysis of complex video and image data plays a crucial role. It aims to extract meaningful information as the distributed storages and processing elements within a bandwidth constraint network seek optimal solutions to increase the throughput along with an optimal trade-off between computational complexity and power consumption. However, due to complex characteristics of visual patterns and variations in video frames, it is not a trivial task to discover meaningful information and correlation. Hence, data mining has emerged as a field which has diverse aspects presently in extracting meaningful hidden patterns from the complex image and video data considering different pattern classification approach. The study mostly investigates the existing data-mining tools and their performance metric for the purpose of reviewing this research track. It also highlights the relationship between frequent patterns and discriminative features associated with a video object. Finally, the study addresses the existing research issues to strengthen up the future direction of research towards video analytics and pattern recognition.

Keyword:
Data mining
Internet of things (IoT)
Multimedia

Corresponding Author:
Benaka Santhosha S,
Department of Electronics & Communication Engineering,
Coorg Institute of Technology,
Kodagu, Karnataka, India.
Email: benaka.research@gmail.com

1. INTRODUCTION

Data mining is a well-known process of knowledge discovery and also exploring significant patterns from a massive set of data. The extensive use of data mining in the field of information technology made it an active research area thus several commercial products and research prototypes are witnessed. The current research trends reveal a matter of fact that more emphasize has been put on corporate data—typically in alphanumeric data-base where very less focus has been laid towards mining of multimedia data [1] Zaiane, Han, & Zhu, 2000). Multimedia data mining till date has been conceptualized for different types of files such as audio, image, and video. In the recent times, accessibility to a huge amount of video contents in both internet and television require implicit knowledge extraction, and it has become a crucial task owing to its non-structured nature. Video analysis or mining is even more complicated task than analyzing still images [2], [3]. However, a video object consists of a collection of a time-ordered sequence of images where the subject in each image is statistically correlated with other. The video content consists of both temporal and
textual information which creates complexity during computational mining operations. A video content can be classified into three different categories which are namely (a) Low-level feature information, (b) Syntactic Information and (c) Semantic Information. The low-level feature information is subjected to concern about the features of a video object such as color, texture, shape and so on. Syntactic information refers to different salient objects their spatial-temporal coordinates and correlations. Semantic information conveys about the sequence of the video from every aspect and also describes what is visually perceived by the viewer [4], [5]. The following Figure 1 exhibits a basic structure of video hierarchy from a segmentation viewpoint.

Video segmentation is a part of video processing which invariably decomposes video track into smaller units. However, the visually based segmentation identifies shot boundaries where the motion based segmentation tracks down the pans and zooms. However, this manuscript intends to contribute towards investing the existing research trends about the efficient knowledge discovery from multimedia contents to maximize the accuracy of analyzed contents. The paper is organized with a pattern where Section 1.1 talks about the background of Video Data Mining concept where Section 1.2 discusses the key research problems identified in this field. Section 2 highlights the existing video mining approaches from theoretical perspectives. Section 3 talks about the existing research contributions using their addressed problems, applied techniques and the performance parameters considered. Finally, the study extract research gap in Section 4 after reviewing the conventional approaches and their contribution concerned followed by conclusion in Section 5.

1.1. Background of Video Data Mining

Video data mining deals with extracting meaningful information from a video data object sequence considering an implicit knowledge discovery process. Visual interpretation of meaningful patterns in a video frame sequence is quite a challenging task as video object comprises complex different patterns of semi-structured and unstructured data. It also includes pattern discovery process while patterns are identifiable in video databases [6]. However, pattern discovery in video databases performed considering an extension of still image mining followed by mining of temporal image sequences [7]. The process also not only meant to extract content, structure, the spatial or temporal correlation between moving objects of video content rather it emphasizes more on extracting patterns concerning object activities and events from a vast amount of video data. There exist certain dissimilarities which makes video data mining a unique from other related areas.

a. Video Data Mining Vs. Video Processing: The relationship between video processing and Video data analytics is quite subjective from different contexts. Video data mining refers to the process of extracting meaningful patterns from a video sequence while video processing focuses on mostly feature extraction.

b. Video Data Mining Vs Pattern Recognition: Both the areas are inclined into feature extraction steps but the video data mining differs in terms of pattern specificity recognition, and Pattern recognition deals with classifying special samples with the help of existing model while video mining indulging into a study which performs detecting of rules and patterns irrespective of any video processing operations.

c. Video Data Mining Vs. Video Information Retrieval: The difference in this context is very much similar to the difference that exists between the traditional data base management systems and the data mining [8]. The prime objective of video mining is to find out correlation and patterns which are yet to understand from a set of video data bases. Video mining performs information retrieval from the video databases and further performs mining operations to recognize the
patterns and trends where video scripting plays a significant role. The following Figure 2 shows a general framework intended to represent overall video data mining operations.

![Figure 2. An Overview of Video Data Mining](image)

The above-stated figure exhibit that how multimedia data mining reinforces feature extraction and clustering process to find a significant pattern from multimedia data. Several studies are found to talk about the architectural process for multimedia data mining which involves three different prime tasks. Firstly it focuses on pre-processing of video data objects which comprises pixels, key frames, segments, scene, etc. secondly it involves extraction of different types of features such as physical, motion, relation features from a video object which helps further in knowledge discovery and pattern recognition [9], [10].

1.2. Key Research Problems in Video Mining

This section briefly talks about the existing research problems associated with the operational design aspects of video mining algorithms and the problem encountered in current research track as well. Video mining often referred as an emerging field of video analytics which influences the futuristic data science from different aspects. The unsupervised learning of different audio-visual patterns makes it an operationally challenging process. Video data mining and data management open up a new era of smart applications which includes intelligent content filters, surveillance, personal video recommendation, or content-based advertisement. The core challenges are to predict semantic features from primitive features. There should be a generalized framework which can have high efficiency on detecting semantic features from general video contents and further apply that to any types of videos. The prime challenge is to build up a framework which ensures efficient extraction of multiple semantics from the videos with the use of primitive features [11].

1.3. Conventional Video Data Mining Approaches

The current research trends are being witnessed by employing various video data mining techniques to a large extent. The prime goal of every mining technique is to extract significant knowledge from the video databases very efficiently and within a short period. However, the extracted data should provide maximum accuracy in analyzed contents. Since many years various video data mining approaches are being proposed which can be crudely classified into the major five categories such as 1) Video pattern mining [12], 2) Video clustering and classification, 3) Video association mining, 4) Video content structure mining and finally 5) Video motion mining. A brief discussion of these video data mining approaches is given below:

1.4. Video Pattern Mining

This process aims to detect various spatial patterns modeled in advance within a video object. What a set of sequential characterized events such as dialogue or presentation image belongs to a medical video is referred in this context. The existing video pattern mining techniques are clustered into two prime categories such as (a) mining similar motion patterns and secondly (b) mining similar objects [13].

1.5. Video Clustering and Classification

It is a process of clustering and classifying the video units concerning different categories. However, clustering meant for performing an unsupervised learning to discover certain significant knowledge from a
dataset. The extensive clustering process during video mining includes performs a sequential operations of (a) shot detection, (b) Key frame extraction, (c) feature extraction and similarity measurement, (d) clustering of key frames followed by (e) semantic interpretation and query retrieval. The clustering has a potential of extracting meaningful patterns from a video sequence which have a significant impact on different aspects such as indexing, surveillance, activity discovery and event recognition [14]. A process of video clustering is depicted below:

- **Figure 3. Functional blocks of Video Clustering Processes**

Clustering and classification analysis in association with above highlighted (Figure 3) functional blocks exclusively tries to find the unique patterns of moving object in a video sequence. Clustering of similar shots in a moving sequence of frames produces more precise video object with less redundancy and noisy environment. There are several clustering algorithms, which are categorized into partitioning methods, hierarchical methods, density-based methods, and grid-based methods and model-based methods.

### 1.6. Video Association Mining

It refers to the operations which involve discovering associations that exist among video frames. This technique extracts the knowledge from a video sequence by incorporating two different functional stages where in Stage-I the video content is segmented into certain units where an analysis further carry out to extract significant features or data patterns. In Stage-II the operational units perform consecutive association mining to extract meaningful knowledge from feature descriptors [15].

### 1.7. Video Content Structure Mining

As video contents comprising of complex unstructured data patterns thus it makes access to video content in a database computationally challenging. To make this operation efficient the prime objective is to convert unstructured data patterns into structured patterns (Figure 1) using video structure mining. The synthetic level composition which detects structured patterns from a video data makes random access to the contents faster. It also defines the fundamental logic structure within a video object constituting a relationship between semantic multi-modality concepts and computer low-level features. The following Figure 4 shows how semantic concepts are detected utilizing a multimodal content analysis and A-priori algorithm.

### Figure 4. A process of semantic concept detection [16]
### 1.8. Video Motion Mining

It is one of the most prominent mining techniques, which characterizes the motion of objects from a video sequence. The motion contains temporal information, and it is quite challenging to retrieve visual information from the motion object regions. The camera motion also poses some key issues in video motion detections such as camera located in a static position while objects are moving, the camera is moving while objects are also moving and a set of cameras recording the same objects [17]. The next section further highlights a summary of existing research contributions using their addressed problems, applied techniques and the performance parameters considered.

### 2. REVIEW OF LITERATURES

Apart from the below mentioned studies, the following are few of the existing literature found relevant to the video mining and knowledge extraction. The study of Wlodarczak [36] reviews few of the significant deep learning methods, recently adopted into the scope of multimedia and video data mining. In the study of Kumar [37], a method that efficiently classifies animal objects from images has been introduced. The study also emphasized on eliminating the background objects from a given image by using a graph cut based technique. Further, the segmented animal images are divided into blocks and passed through a processing where color texture moments are further extracted by referring each block segment. Further, the concepts of probabilistic neural networks and K-nearest neighbors are taken into consideration to perform the classification. Shahbaz et al. [38] conceptualized a novel approach well capable of classifying satellite images. It also objectifies different unique patterns from satellite images very efficiently. An experimental testbed has been created to validate the performance efficiency of the proposed method where the method has been tested with a set of 42 satellite images. The performance analysis further shows that the proposed technique outperforms the conventional video mining approaches with an accuracy of 80%. The study of Benoit [39] mostly emphasized on improving the performance concept detectors (CD) and further conceptualized an analytical model by improvising the quality aspects of semi-supervised learning enabling refinement framework. The study employed a self-training paradigm capable of expanding the training dataset with automatically labeled data. The numerical analysis and simulation also exhibited the prime role of extracted visual features and text metadata to enhance the performance of the concept classifiers concerning different unlabeled video objects. The performance of the proposed method has been validated considering a dataset comprising 21,000 entities which shows expanding the training set with labeled shots and CD significantly excels better outcomes. Research in similar direction performed by Cao and Wang [40] where the authors presented a novel concept which is subjected to enhance the performance of image mining modeling. The outcomes obtained after simulating the proposed technique conveyed its superiority as compared to other image mining schemes. In the study of Devasena and Hemalatha [41], a novel approach for video analytics has been introduced. The technique integrated with a unique LIM based clustering paradigm which uses self-organizing maps to discover unique patterns from video frames and also distinguish novelty in the frames belongs to a video sequence. The design and implementation of the proposed framework have been carried out in a numerical computational environment where a set of sample videos are considered to test the performance of the proposed model. The experimental outcomes further exhibited that it provides promising results and ensure it's adaptability to object detection systems including remote video surveillance in defense for national and international border tracking.

Saravanan and Srinivasan [42] have identified the need for an efficient Video Frame Based retrieval system. Vaduva [43] presented the image information mining based on a communication channel concept. The authors considered this approach for a real meaning based semantic annotation of very high-resolution remote sensing images. The scene content is described using a multi-level hierarchical information representation. Feature hierarchies are discovered considering that higher levels are formed by combining features from the lower level. Such a level to level mapping defines our methodology as a deep learning process. Wang [44] have the framework of MapReduce is explored for large-scale multimedia data mining. Yang et al. [45] design follows a model-view-controller (MVC) pattern for applying semantics. Table 1 show the summery of Existing Studies towards data mining key activities.
Table 1. Summary of Existing Studies towards data mining key activities

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Concepts or problem Focused</th>
<th>Application Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tien et al. [18], Huang et al. [19]</td>
<td>2008</td>
<td>Symbolic streams mining, visual feature mining, support vector machine</td>
<td>Sports, movie, surveillance</td>
</tr>
<tr>
<td>Chaudhury et al. [20]</td>
<td>2009</td>
<td>Object Mining, Support Vector Machine</td>
<td>Movie/Tv shows, sports, video</td>
</tr>
<tr>
<td>Anjulan and Canagarajah [21], Gadon et al. [22]</td>
<td>2011</td>
<td>Sequential pattern mining</td>
<td>Movie/Tv shows, sports, video, surveillance</td>
</tr>
<tr>
<td>Gilbert et al. [23], Harikrishna et al. [24], Jiang et al. [25]</td>
<td>2011</td>
<td>Hierarchical visual event pattern mining</td>
<td>Surveillance</td>
</tr>
<tr>
<td>Cui et al. [26]</td>
<td>2012</td>
<td>Temporal association rule mining</td>
<td>Security and surveillance</td>
</tr>
<tr>
<td>Anjulan and Canagarajah [21], Gaidon et al. [22]</td>
<td>2012</td>
<td>Video data model, Video data mining</td>
<td>Movie/Tv shows, sports, video, surveillance</td>
</tr>
<tr>
<td>Xiaolin et al. [28]</td>
<td>2013</td>
<td>Data mining based Radar Track simulation</td>
<td>Radar data mining</td>
</tr>
<tr>
<td>Gowsikhaa et al. [27]</td>
<td>2014</td>
<td>Slack based measure (SBM) model and data mining</td>
<td>Education Efficiency</td>
</tr>
<tr>
<td>Vijaykumar and Neduzechuan [28]</td>
<td>2014</td>
<td>Commercial decision making and Data mining</td>
<td>Data Analysis</td>
</tr>
<tr>
<td>Xiong et al. [29]</td>
<td>2016</td>
<td>An optimization problem formulated to cluster behavior patterns (bPs)</td>
<td>Video segmentation in the field of Security and surveillance</td>
</tr>
<tr>
<td>Zhou et al. [30]</td>
<td>2016</td>
<td>Video object segmentation</td>
<td>Video game development, surveillance</td>
</tr>
<tr>
<td>Park et al. [31]</td>
<td>2017</td>
<td>Discovering knowledge in TV rating data</td>
<td>Tv shows, water and sport reports</td>
</tr>
<tr>
<td>Palazzo et al. [32]</td>
<td>2017</td>
<td>Video anomaly detection</td>
<td>Security and video surveillance systems</td>
</tr>
</tbody>
</table>

2.1. Statistics of Existing Research Trends on Video Mining

This section exclusively highlights the current research statistics about the video data mining and analytics. There are not much-published research articles found in IEEE Xplore digital library within a timeline of 2008-2017. The statistics of the published contents till date is represented in Figure 5.

The study also emphasized on exploring existing studies focused on strengthening different video data mining applications. The data collected exclusively referring IEEE Xplore digital library which shows very fewer works have been carried out on video data analytics considering different application aspects such as security and video surveillance, video game applications, etc. The following Figure 6 shows an overview of the current statistics subjected to the number of research articles published on video data analytics for different video applications.

![Figure 5. Research Statistics on Video Data Mining](IEEE Xplore)

![Figure 6. Statistics of research on Video data mining applications](IEEE Xplore)

The above figure shows that video data mining has been most extensively and widely adopted into security and surveillance applications since 2008. The statistics also conveys that very less contribution to movie/TV show analytics. The research gap extracted briefly illustrated below.
3. RESEARCH GAP

After reviewing the above-stated existing literature, concerned different aspects of video data mining, the study outlines most significant issues needed to be addressed to reinforce the existing video mining techniques.

a. Mining of Semantic Concepts: Very few studies are found to emphasize on mining of semantic concepts from different intelligent video applications. Most of the works found to carry out theoretical discussion on predictive semantic feature problems whereas no extensive simulation to determine the effectiveness of video analysis based on primitive features has been witnessed.

b. Less Focus towards Non-Deterministic Approaches: Most of the existing studies focuses on deterministic approaches where very few found to apply non-deterministic approaches during knowledge extraction from a video object.

c. Few Benchmarking: Very less effective studies are found till data where computational complexity and benchmarking of the proposed solution towards video mining highly ignored.

d. No-Optimization: Very fewer studies considered algorithm optimization to maximize the system throughput from an operational viewpoint.

4. CONCLUSION

The proposed study intends to perform and in-depth analysis of the conventional video mining techniques and their performance efficiency. The study also highlights a comprehensive overview of different video mining techniques and their adaptability into different systems for efficient knowledge discovery process. The investigational study depicted the fact that the existing solution approaches lacks computational efficiency and doesn’t achieve an optimal trade-off between maximum accuracy in analyzed contents and operational constraints. It also outlines the existing research issues which are needed to be minimized to make this research track more effective and operative.

REFERENCES


BIOGRAPHIES OF AUTHORS

Benaka Santhosha S, Assistant Professor, Department of Electronics and Communication, Coorg Institute of Technology, Ponnampet. I have done BE Degree in Electronics and Communication from VTU in 2010. I have M.Tech Degree in Digital Electronics and Communication Systems from VTU in 2012. I have 5 years of experience in teaching. I am perusing PhD in VTU. My Area of Interest is Image Processing, Signal Processing, and Multimedia.

N. Chitra Kiran, Professor and Head, Department of Electronics & Communication Engineering, Sai Vidya Institute of Technology, Bengaluru. She has done PhD in Electronics & Communication Engineering.