

Insights to Video Analytic Modelling Approach with Future Line of Research

Madhu Chandra G.
Research Scholar
Department of Electronic & Communication
Engineering
M S Engineering College,
Bangalore
Visvesvaraya Technological University,
Belagavi, India

Sreerama Reddy G. M., PhD
Professor & HOD
Department of Electronics &
Communication Engg.
C Byregowda Institute of Technology,
Kolar, India

ABSTRACT

The area of Video analytics has made some significant improvement due to advancement in image processing and datamining techniques. However, the inclination is still towards image contents and less to mined contents owing to many unsolved issues. Although concept of mining is more than 2 decade old, but mining approaches are yet to be standardized in the area of video surveillance system. With evolution of newer set of challenges in video capturing, existing mining models finds itself less applicable due to unstructured format of dynamic frames. Hence, this paper discusses about video analytics and presents a brief discussion of frequently used mining approaches in video as well as discussed some recent studies in this direction in order to scale the degree of effectiveness in existing system. The paper also presents research gap and provided solution as future line of research as a possible way to overcome the research gap.

Keywords

Analytics, Data Mining, Knowledge Discovery, Semantics, Video.

1. INTRODUCTION

With the increasing trends of automation and data analysis, video analytics has completely taken a new shape. The existing video capturing devices like CCTV (Closed Circuit Television) is able to continuously capture live feeds but cannot perform identification of specific object based on user's requirement. A normal human can only keep the concentration on multiple surveillance screens not more than 20 minutes. Moreover manual surveillance requires more involvement of human intervention which is near to impossible if the quantity of the surveilled area is very large. Hence, a smart technique is required that can perform monitoring of specific event or objects within the video. It can be defined as a computational technique to process and analyze a stream of video. At present various forms of visual captured feeds are either stored in DVR (Digital Video Recording) system or are either compressed or uploaded over cloud. Hence, such a massive data goes completely unused until and unless a datamining approach is applied on the top of it [1]. However, there is a large range of complexities associated with datamining as compared to data management techniques [2]. Usually, the conventional datamining approach can be easily applied over structured data whereas in case of video data captured from multiple cameras are highly unstructured in nature. This is the main reason which results in less applicability of conventional data model representation in video mining approach [3]. Therefore, in order to develop a video analytical system, it is required to convert the unstructured data to structured data and then only

mining could be applicable. Otherwise, it will result in slower response time rendering it less applicable in real-time applications of video-analytics. The existing system are more on offline analytics and very less on online analysis as it can't handle the dynamics of the video frames and its associated uncertain parameters. The present meaning of video analysis is either an experimental prototype on controlled research environment or performs analysis of the offline video dataset. The application of video analytics ranges from analysis of real-time scenario as well as to analyze offline video file in order to extract the relevant information of interest [4]. A typical video has three types of contents e.g. i) audio, ii) video, and iii) text contents. Hence, in order to perform datamining over a video, it is required that a standard framework or a model to be chosen first in order to close the bridge of semantic gap between low and high level of features and concepts. Not only this, it is quite a difficult task to select mining approaches over a video. One of the biggest challenges is to extract a specific pattern with robust modelling. An efficient video analytics must have the capability of interpreting and applying the extracted knowledge for dynamic ranges of the video frames. There are multiple beneficial factors for the video analytics e.g. i) it assists in minimizing range dimensional space of storage and hence perform better storage optimization, ii) it allows sophisticated training and learning approaches to precisely detect the target semantics, iii) it also has capability to explore contents of media for ensure proper delivery of media contents as well as to allow all the advanced application to be driven by media type, and iv) it must have the customizable modelling for carrying out video indexing.

Therefore, this paper discusses about the core insights of the video analytics and measure the level of effectiveness of the existing techniques of video analytics. The paper also present possible solution to overcome issues in the form of future work. Section 2 discusses about an essential concept of video analytics where the discussion is made with respect to video processing, information retrieval, and datamining. Section 3 discusses about the frequently used techniques of analytics followed by discrete discussion of recent research contribution in this regards in Section 4. Section 5 highlights the points for research gap and Section 6 discusses about future Line of research meant to overcome research gap. Finally, Section 7 makes concluding remarks.

2. ESSENTIALS OF VIDEO ANALYTICS

Basically, video analytics are not a mere operation on mining approaches towards video, but also perform value added task in the knowledge discovery [5]. A typical video consists of multiple information which is the source of interest for any

data mining operation. There are three types of information existing within a video e.g. i) low-level information (e.g. texture, color, shape, etc.), ii) semantic information (e.g. spatial factors-characters, location, objects etc. and temporal factor-video sequences and movement of objects), and iii) Syntactic information (salient objects, respective position and timing attributes etc.) [6]. Another significant factors that plays a crucial role in designing video analytics are modalities. There are three types of video modalities viz. i) auditory (music, speech, surrounding sounds), ii) visual (seen of video), and iii) textual modalities (text contents of video) [7]. In order to understand the core process of video analytics, it is essential to understand the basic process of video processing and video retrieval system followed by data mining operation on video.

2.1 Video Processing

Video processing can be defined as an operation that takes the input of video and suitable performs various mathematical or signaling operations in order to make the video suitable for subjecting it to specific algorithm. The prime step in video processing is to perform transformation that converts non-structured video to structured video data. Depending on the types of contents, the process of structuring the video data is of two types i.e. scripted and unscripted video data. The scripted data are generated by controlled environment (movie file) while unscripted one is directly generated from real scene (e.g. CCTV footage, traffic scene etc.). Normally, the video clip is bigger in size and hence it is necessary to suitably partition it avoid long sequences of video. Fig.1 shows that for a given video, initially scenes are identified followed by identification of video group, shots, and key frames. All these respective information are seeked during the mining operation over video data.

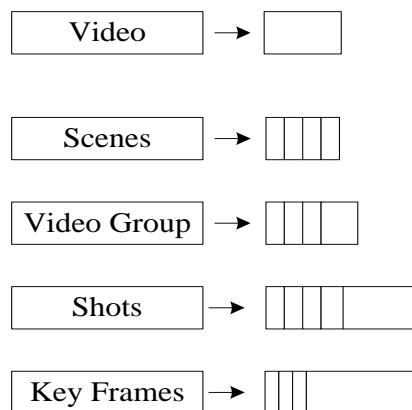


Fig 1: Hierarchy of video frame

Fig.1 shows some of the essential elements of a video which are required to be extracted in order to apply video mining operations. A *video* can be defined a frame sequences with both images and sound. A *scene* can be defined as a aggregation of adjacent groups that are temporally and semantically associated. A *video group* can be defined as an transitional component between the semantic scene and physical shots. A *shot* can be defined as frame sequences obtained from one camera. A *key frame* can be defined as visual contents with saliency features. Following are the important operations involved in video processing [8].

2.2 Video Data Model

It is highly essential to have a standard data model of a video as conventional model or even an object-oriented data framework doesn't support extraction of specific information

from the video *contents*. The prime reasons behind this are absence of any knowledge-based frameworks that can extract semantic contents from raw video data. Moreover there is no standard representation of query for complex video structures. The video data model is highly important as in order to perform video analytical operation a robust data model is required.

2.3 Video Segmentation

Segmentation has always been the first step of video processing as it assists in segmenting the track of video into a very smaller units with a supportability of subsequent important task e.g. semantic representation, video indexing, tracking etc.

2.4 Feature Extraction

There can be multiple forms of features that can be extracted from a single video frame e.g. color, histogram, edge, texture, etc. However, with respect to the term *video*, the features can have only two classification i.e. descriptive feature (caption, keyword, time, size etc.) and content-based feature (whole image, selected image patch). There can be also three types of features (e.g. spatial features, audio feature, and temporal features) which play a crucial role in performing data mining operation over a video.

2.5 Video Information Retrieval

Video information retrieval system can be defined as a specific form of multimedia database management that responds to user's query for a particular video by returning matching video. The manipulation of the video data is taken care of in order to perform and efficient video information retrieval system. The most significant step involved in this process are to identify boundary of shot, generation of features (low-level) from key frames. Apart from this, the most significant method of video information retrieval system is as follows:

2.6 Video Abstraction

Also known as video summarization, video abstraction is a process of consolidating the essential frames of the video. The process retains only the frames with non-repetitive information in sequence and considers it as a summary of the entire video. There is another similar process called as video skimming. This process is also used in building video analytics [9].

2.7 Video Annotation

The process of video annotation involves appending a suitable keyword which is displayed exactly at the same time of commentaries in video. The annotation is done using text which acts for a purpose of video retrieval system. There are three types of video annotation process e.g. rule-based, statistics based, and hybrid annotation based. Video annotation plays a significant role in video analytical operation as text-based annotation assists in fulfilling the semantic needs for solving query problems in mining operation [10].

2.8 Data Mining on Video

The process of data mining when applied to a video data performs implicit discovery of knowledge and develops relationship among the video data. The datamining techniques are responsible for involuntary extraction of structures as well as contents of the video, correlational factors of spatial and temporal features, discovery of hidden patterns in a video, etc. However, the process of video information retrieval system

shouldn't be confused with video mining techniques. Fig.2 shows the relationship between video information retrieval systems with mining system.

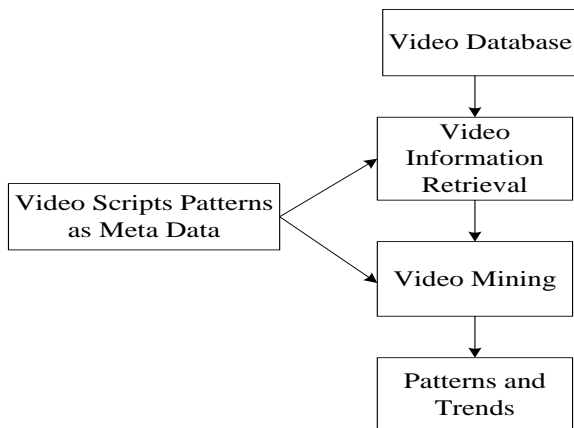


Fig 2: Relation between video information retrieval and mining

Video information retrieval system is just a part of mining operation on video. However, there are multiple related area e.g. pattern recognition, video processing, and video information retrieval system, which are somewhat thought of similar technique with video analytics. Both video mining process and pattern recognition deals with extraction of features but with different patterns. The motive of video mining approach is to generate all possible forms of patterns whereas motive of pattern recognition is only to recognize certain standards of patterns. Both video mining and video processing are quite subjective in nature. The motive of video mining is to acquire pattern from sequences of frame whereas video processing is more involved in feature extraction process. The difference between video information retrieval system and video mining process is that former is more over data management and latter is more involved in data mining operation. For clear understanding, the standard operation of video mining is shown in Fig.3.

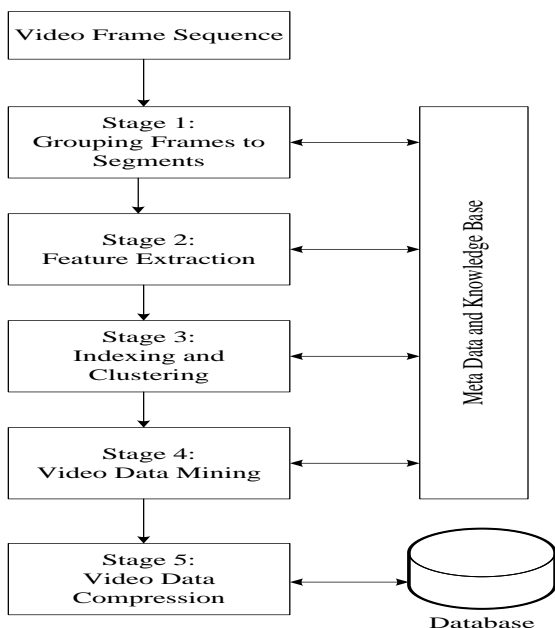


Fig 3: Standard Video Mining Process

Video mining technique can be represented as form of automated pattern discovery process where a mining

operation is performed over spatial and temporal features of a video dataset. The spatial feature will consists of texture, text, color, edge, shape, etc. whereas the temporal features will be the motion of the object. At present, majority of the video comes with text-related features that can be deployed for highlighting significant events within the video dataset. The source of an origination of the text from the video are i) superimposed text (e.g. subtitles), ii) scene text, iii) and inbuilt recognition of speech within the video. The next section gives elaborated insight of the techniques of the video analytics introduced by the research community till date.

3. TECHNIQUES ON VIDEO ANALYTICS

This section discusses about the technique of video analytics introduced by various researchers till date. The various standard and frequently used techniques are as follows:

3.1 Video Clustering

The mechanism of video clustering technique involves performing clustering to the video data into multiple categories. It can also be represented as a learning technique (unsupervised) for knowledge discovery from specific dataset. Fig.4 highlights an important steps involved in clustering process. One of the challenging problem in this operation is to perform clustering on video sequences for extracting information about specific activity for a given stream of video.

The clustering mechanism significantly assists in discovery of specific activity, surveillance, indexing video etc. [11]. There are various types of clustering algorithms e.g. hierarchical [12], grid-based [13], partitioning [14], density-based [15], etc. Clustering phenomenon directly assists in solving classification problems in video mining operation.

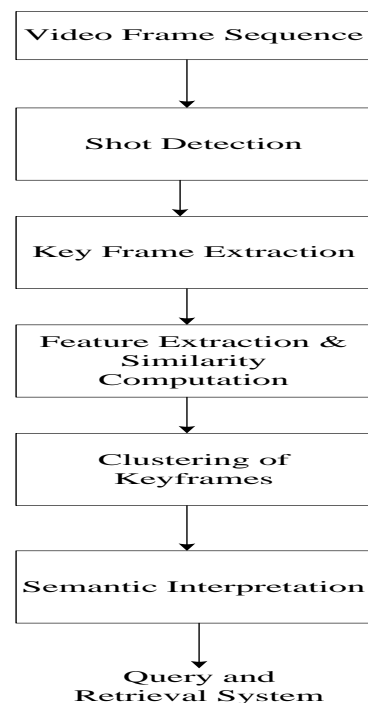


Fig 4: Mechanism of Video Clustering

3.2 Semantic Operation

There are various techniques which majorly uses association rule as a part of video mining technique. Such technique

basically integrates association rule with video processing in order to extract knowledge. Fig.5 shows a unique technique where video streams are subjected to semantic based operation in order to generate association rules. Usage of such concept can be seen in the work done in [16] [17].

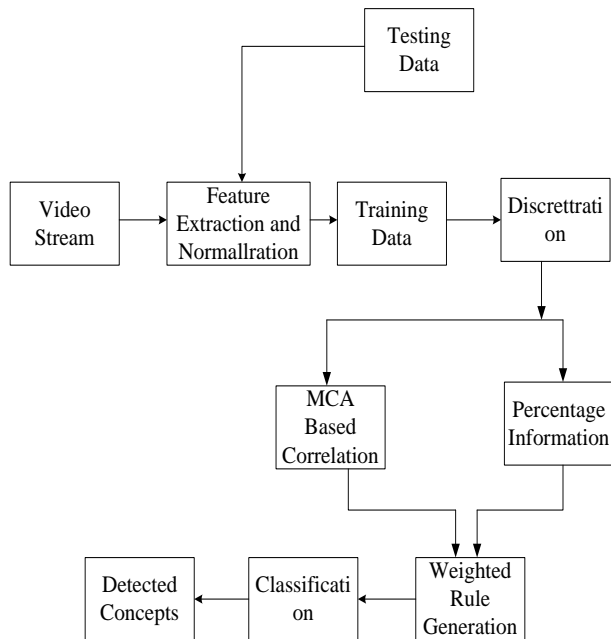


Fig 5: Semantic Operation in Video Mining

3.3 Pattern Mining

Video pattern mining is a technique of identifying specific pattern (predefined) that corresponds with specific video events. Usage of pattern mining concept can be seen in the work done in [18] [19] [20].

3.4 Structure Mining

The prime task of video structure mining is to perform detection of potential content structure for the purpose of performing arbitrary access to video. It can be also defined as identification of logical structure of a video using multiple techniques. Usage of structure mining concept can be seen in the work done in [21] [22].

3.5 Motion Mining

Motion can be stated a prime temporal feature present in one video and they are more consistent/objective in nature compared to other video features. There are various approaches for detection of moving objects considering camera motion as it plays a significant role in video motion detection over analytical operation. Usage of motion mining concept can be seen in the work done in [23] [24].

The next section discusses about the applications of video analytics.

4. APPLICATIONS OF VIDEO ANALYTICS

As video analytics takes the input video and convert it to a specific form of structured data hence there are higher chances of extraction of more features which can be further subjected to sophisticated mining operation. Various applications of video analytics are as follows:

4.1 Monitoring Human Behaviour

It is possible perform identification of the human behaviour using video analytics. The process (Fig.6) takes the video feed and applies segmentation followed by background/foreground extraction. Usually, in this segmentation, existing techniques uses head motion detection, motion tracking, and contact detection in order to understand the activity of the perform. The system can than perform classification of activity based on need. Such forms of application can be seen in the work carried out by [25], [26], [27].

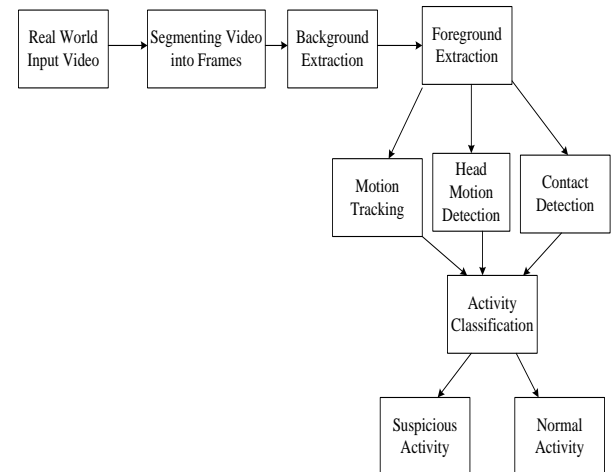


Fig 6: Scheme to Monitor Human Behaviour

4.2 Movie Classification

Classification of a movie file may be one of the most demanded applications in the big world of entertainment. In such mechanism, video frame sequences are considered as an input followed by detection of shot boundaries. This step is then followed by extraction of various features (could be both temporal and spatial) e.g. shot length, motion, color etc. Various forms of optimization techniques could be used to perform classification process. Such forms of application can be seen in the work carried out by [28], [29], [30].

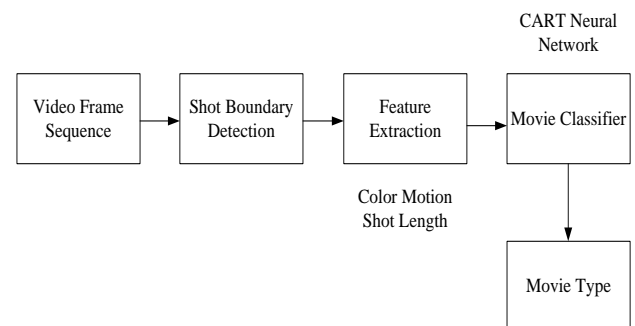


Fig 7: Scheme to Classify Movie

4.3 Commercial Video Mining

Various videos like commercials, news video, sports video, documentaries etc. can be used for performing video mining operation. Such videos are taken as an input from where the video shots are detected. Feature extraction is followed by pre filtering of the data and classification rules. Such forms of application can be seen in the work carried out by [31], [32], [33].

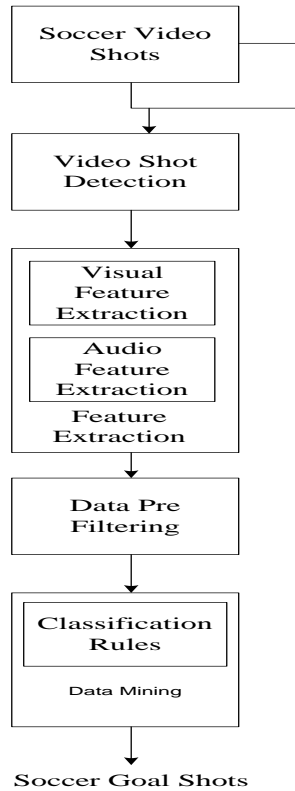


Fig 8: Scheme to perform commercial video mining

4.4 Identification of Abnormal objects

Video analytics can also be used for identifying abnormal objects from the given scene capture of a video. The technique allows taking the raw video from which the audio and visual features are extracted followed by video segmentation. A binarized approach for classification of data is obtained in order to perform online classification techniques, which finally results in identification of abnormal segments. Such forms of application can be seen in the work carried out by [34], [35].

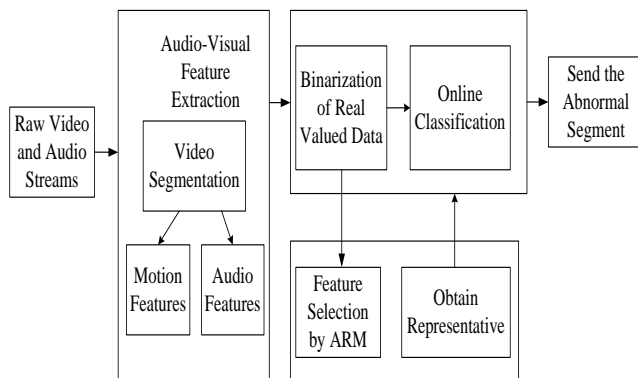


Fig 9: Identification Processes of Abnormal Objects

Apart from this, there are multiple utility of video analytics e.g. i) identification of suspicious tampering of camera, ii) identification of vehicles parked beyond time limit, iii) capturing customer traffic at sales point to improve customer management, iv) identification of region of more interest and less interest in a commercial location, v) tracking minor movement of patient under coma stage, vi) improves transportation decision system by leveraging navigational and

traffic-based information to customers. Hence, there are multiple positive advantages where video analytics could be useful. The next section discusses about existing techniques.

5. EXISTING TECHNIQUES OF VA

This section discusses about the existing techniques of the Video Analytics (VA).

The most recent study towards video analytics was carried out by Angelov et al. [36] intended to carry out real-time analysis of video. The technique uses iterative estimation of density and local average of clustering. The technique applies Lucas Kanade algorithm as well as RANSAC algorithm to perform better detection of object from air borne vehicle. Aradhya and Pavithra [37] have presented a study that performs identification of text from the videos with multiple language. The technique uses Gabor filter, wavelet transform, and k-means clustering in order to perform video indexing. Although, the study is not directly an application of video analytics but its mechanism is highly assistive in video analytics. Aryanfar et al. [38] have used a conventional classifier (support vector machine and naïve Bayes) in order to recognize the human action. The outcome of the study was compared with existing mechanism to find approximately 90% of recognition rate. Most recent, the actual study towards video analytics was seen in the work of Ayed et al. [39]. The author have used MapReduce model in order to perform mining of video dataset. The study outcome was testified with visual outcomes, classic model time, and processing time taken by MapReduce. Cai et al. [40] have discussed about the video analytics in order to gauge the engagement of customers. Author has used statistical approach to perform prediction of the video. Chen et al. [41] have presented a model called as PeakVizor which intends to extract the patterns of sophisticated learning mechanism of online courses. The study uses Glyph visualization technique in order to perform data abstraction as well as peak detection from MOOC (Massive Open Online Courses) videos. The study also assists in formulating correlation among different online users. Kim et al. [42] have developed a technique that can perform identification of outlying behaviour of user. The analytical module design by the authors considers the inputs as movement of users from webcam and that apply multimodal feature for collecting data followed by data refinement. The data for facial expression, facial movement, log data, audio data to perform classification of the outliers. Study on video surveillance was also carried out by Mao et al. [43]. The author have presented a unique traceability mechanism for system over video surveillance system. The technique also introduces a background modelling using trajectory generation for the purpose of food traceability system.

Shao and Fu [44] have discussed about video analytics which uses multiple view and self-learning algorithm. The study uses three different models e.g. a learning model, feature extraction model, and self-learning model. Xu et al. [45] have presented a technique that enables performing mining operation on three dimensional data of video residing over cloud environment. The technique also performs cryptographic process over the data. The outcome of the study was evaluated with respect to the time required for original and encrypted video coding. Riahi et al. [46] have studied outlier detection using Bayes network over relational dataset. The technique presents a simple framework for performing anomaly detection from a given relational data using Bayes network based learning. The system also has a rating mechanism. Descriptive Statistical parameters (e.g. minimum,

maximum, standard deviation, mean) were used for evaluating the success metric of multiple dataset

Potapov et al. [47] have presented a technique that performs summarization of a particular video. The technique allows performing a segmentation process to generate a segments with semantic factor. The technique uses kernel temporal segmentation process which allows rejecting the frames with nearly corresponding values which after applying to support vector machines performs summarization of video. Thilagavathy et al. [48] have presented a technique that uses neural network for performing extraction of text from multimedia file system. The technique initially obtain key frame from the video where the text region indicator is used along with minimum spanning tree. Cong et al. [49] have presented a study that perform searching anomalies in video considering the cluttered scene. The technique perform segmentation based on distance and time and then uses descriptor. In order to perform detection of anomaly, the technique uses abnormal occasion matching technique. Study towards anomaly detection was carried out by Bertini et al. [50]. Similar work has also been carried out by Jiang et al. [51]. However, the technique differs from others by incorporating context-awareness. The technique uses spatio-temporal approach to identify the point anomaly, sequential anomaly, and co-occurrence anomaly for video objects. The study uses a hierarchical data mining approach on both single and double objects video.

Deldjoo et al. [52] have presented a technique that can perform extraction of visual features in order to perform video recommendation based on its contents. Nomiya et al. [53] have presented a mechanism that can extract emotion-based information from the video. The technique uses an unsupervised learning-based approach using multiple facial expression factors. Devasena and Hemalatha [54] have presented a unique mining technique for videos which is based on clustering process. The technique also uses self organizing map for further training process. The study outcome was tested using 5 different video dataset. Kumar et al. [55] have developed a technique which perform faster extraction of text from the video, which is potentially important to develop a design of video analytics. Saligrama et al. [56] have presented a statistical approach for identification of video anomaly. The authors have also discussed various other strategies for activity detection mechanism in video. Surekha et al. [57] have presented a unique mining approach that integrates conventional mining with social network analysis. The author have also discussed about different feature and characteristics of data over internet to perform mining operation. The study outcome was tested with respect to centrality concept of the social network analysis. However, the study is more inclined on sentiment analysis from the video with precision of 88% found in the outcome study. Table 1 summarizes the existing techniques.

Table 1: Summary of Existing Techniques

Authors	Techniques	Advantage	Limitation
Angelov et al. [36]	Lucas-Kanade Algorithm, RANSAC	Faster processing	Not meant of sophisticated analytical operation
Aradhya [37]	Video indexing using wavelets, Gabor Filter, k-means clustering	-supports video indexing for English & Kannada language	Not tested for analytics
Aryanfar et al. [38]	support vector machine and naïve Bayes	-Good Recognition rate	Not tested for analytics
Ayed et al. [39]	Mining with MapReduce, Wavelets	-Reduce processing time	-No comparative analysis -No complexity Analysis
Cai et al. [40]	Statistical Analysis	Good precision	-No comparative analysis -No complexity Analysis
Chen et al. [41]	PeakVizor, Glyph visualization	-Interactive detection of user's online behaviour	-Doesn't work online -cannot perform classification -No complexity Analysis
Kim et al. [42]	Morphological study of facial expression	-ideal for online games	-No comparative analysis -No complexity Analysis
Mao et al. [43]	Video traceability	Accurate target extraction	-Doesn't include high level of data extraction or mining. -No benchmarking
Shao and Fu [44]	Self-learning algorithm, deep learning	-Can read variable illuminative frame condition	-No Numerical analysis -Algorithm Effectiveness not discussed
Xu et al. [45]	Encoding 3D video, Cryptography	-Secured data encoding	-No comparative analysis -No complexity Analysis -use of Encryption leads to complexity. -Not applicable for online video.
Riahi et al. [46]	Statistical approach, Bayes network based learning	-Effective correlation-based rating system	-No comparative analysis -No complexity Analysis
Potapov et al. [47]	Kernel-based temporal segmentation, support vector machine	-better performance of temporal segmentation	-No numerical analysis -No benchmarked studies -No complexity analysis
Thilagavathy [48]	Neural network, minimum spanning tree, Optical character recognition	Simple technique to extract text	-No numerical analysis -No benchmarked studies -No complexity analysis
Cong et al. [49],	Spatio-temporal segmentation	-Good performance of true	- No complexity analysis

Bertini et al. [50]		positive rate.	
Jiang et al. [51]	Spatio-temporal context	Good realization of detection rate	-No benchmarked studies -No complexity analysis
Deldjoo et al. [52]	Feature extraction	-better entropy performance	-Doesn't support mining - No complexity analysis
Nomiya et al. [53]	Emotion-based facial expression recognition	-better detection accuracy	-Not applicable for video analytics with multiple and big number of faces -Outcomes not benchmarked
Devasena [54]	Clustering using Lorenz Information Measure, Self-Organizing Map	Supports simple video mining approach	-No benchmarking -No discussion of complexity handling.
Kumar et al. [55]	Text extraction from video	Simple & cost efficient approach, better recall rate	No discussion of complexity handling.
Surekha et al. [56]	Social network analysis	88% accuracy	-No applicable for online videos -No Scope of optimization -No discussion of time complexity.

6. RESEARCH GAP

This section discusses about the research gap after reviewing the existing techniques of video analytics:

6.1 Less Supportability of Existing Techniques

All the mining techniques discussed in this paper as well as presented till date just consider video as a structured data which is far from reality. Majority of the upcoming future technology will use visual sensors in order to capture the visuals which are interconnected with each other. This will mean that overall database may be quite unstructured one. None of the work till date has explained how the unstructured data is converted to structured data. Conventional datamining algorithms can be never used over unstructured and bigger streams of data as it doesn't have any standard data model representation. At present, we have MapReduce to perform this operation, but again MapReduce has its own limitations towards video processing.

6.2 Less Work towards Outlier Detection

At present, there are certain degree of work that has considered a smaller clip of video to detect abnormal event. However, outlier detection is quite a difficult task in video analytics. Extraction of outlier as patterns using unsupervised learning technique is again a big research gap owing to more number of uncertain variances and differences among the feature. Also less focus is laid on considering contextual factors.

6.3 Less Standard Studies

Majority of the research work implemented till date are not benchmarked and doesn't speak about algorithm computational complexity. Because of this, it is quite a hard task to find the best algorithm performance as the better performance in papers are claimed on the basis on controlled environment of standard dataset, which doesn't ensures its performance with real-time environment. Majority of the retrieval results is designed on the basis of low-level features, which yield poor outcomes in terms of sensitivity, specificity, f1-score, etc.

6.4 Less Research on Event Mining

There are massive number of focused on object detection, tracking, and recognition, however, there is no single work focused on event mining. However, there are 25 journals and

106 conference papers published between the year 2010 to 2016 found to discuss to some extent about event mining from video. The existing system has used textual information majorly in doing this task. But, neither textual information nor visual information separately cannot perform event information with satisfactory outcomes. We don't find much mathematical modelling in this regard either.

Hence, there is a serious need to perform in-depth investigation on video analytics as we feel that this is still in developing stage and needs more attention. Cost effective datamining with faster response time is yet to be seen in this regards

7. UPCOMING LINE OF RESEARCH

This section presents our future work which is motivated by the findings of our survey to existing approaches of video analytics.

7.1 Framework for Outlier Identification in Video

The prime aim of this framework is to perform identification of outlier in the video. The schematic diagram of the proposed stages is shown in Fig.10. There are two types of video event outlier i.e. i) primary outlier and ii) secondary outlier. The primary outlier can be said to be a significant event in a video that is very much different from its temporal and spatial adjacent events. The secondary outlier can be said to be occurrences of multiple number of abnormal events. A closer look into the existing research towards video analytics will show that majority of its research are more focused on addressing the problems of primary outlier but quite less on secondary outlier. Example of primary outlier is an object in a video with abnormal structure or velocity, whereas example of secondary outlier can be traffic scenarios. Hence, a tertiary outlier would be much worst i.e. exploring events that have minimal feasibility of occurrence in a video.

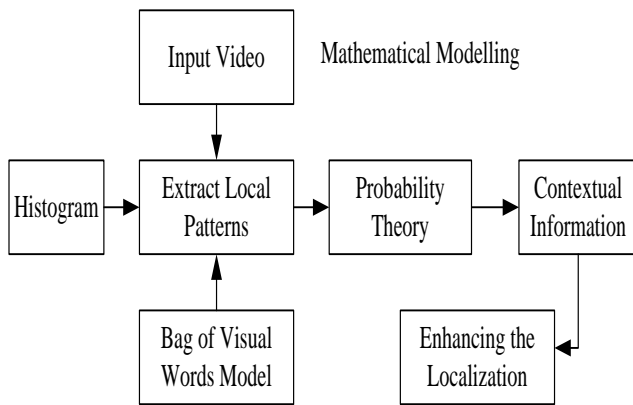


Fig 10: Schema to be adopted in Outlier Detection

This section therefore illustrates the problems being identified followed by methodology to be adopted and its associated benefits.

7.2 Problem Identification

Study towards automated video surveillance will call for identification of all possible outliers, which can be represented as problem of behaviour classification. Normally, there are various forms of visual sensors that capture the event representing the behaviour of monitored targets with presence of certain behaviour to be considered as outlier. The raw signal is normally subjected to the process of feature extraction that is again subjected to learning approaches for identification of outliers. The applications pertaining to automated identifications of outliers are quite useful for minimizing the need of unnecessary data quantity to be processed. Unfortunately, till date there were no research works focusing on in-depth investigation on behaviour classification by identifying the outliers from video.

7.3 Methodology To be Adopted

The core technique adopted for identification of outlier is by investigating the availability of local patterns from the input video. At the initial level, a histogram is constructed for the purpose of extracting local features with more complex patterns. The next step is to apply the probability theory for evaluating the amount of outliers available in the data. For precise in identification of outliers, the system integrates context information of both temporal and spatial factors. As an outcome, a precise identification of outliers can be expected. Following are the indicative steps to be followed for developing this framework:

- Develop a novel mathematical modelling that can extract the local patterns of the video.
- The system will also use bag of visual words model for extracting the local patterns in the video
- To develop an empirical modelling for minimizing the errors during the reconstruction of an image using sparsity of a vector.
- To develop a learning approach along with construction of novel mechanism of histogram calculations for further automating the feature extraction process.
- To develop a novel identification model for enhancing the localization characteristics of outliers in the video.

- To consider multiple challenging scenarios considering both real-time and synthetic video (and video sequences) dataset for testing the success rate of this technique.
- To measure the effectiveness of this model with respect to true / false positive.

7.4 Benefits

This model can offer a true estimation of outliers for a given video. As the proposed study considers adopting probability theory, hence, its response can be expected to be instantaneous as well highly accurate. It is also anticipated that proposed system will offer 95% precision for identification of outliers from multiple forms of video inputs that will positively affect the sophisticated analysis of the video.

8. CONCLUSION

This paper has discussed about the significance of video analytics as an emerging research area of unsupervised extraction of significant patterns from a given contents of video. The paper also emphasize that conventional mining techniques are not so easy to be applied over video as it is quite a challenging task to establish a hidden relationship about th segments or objects present within the components of the video. This makes the area of video analytics more interesting and worthy to carry out further investigation. The paper has also reviewed all the significant papers published during the year 2010 to 2016 and explored significant research gap. After reviewing the limitation and research gap, we discuss about our plan for future line-of research in order to address some of the significant problems in video analytics. We believe that our upcoming video analytical framework will be able to overcome some of the existing research gap to a larger degree.

9. REFERENCES

- [1] D.T. Larose, C.D. Larose, "Data Mining and Predictive Analytics", John Wiley & Sons, pp. 824, 2015
- [2] M. Hofmann, R. Klinkenberg, "RapidMiner: Data Mining Use Cases and Business Analytics Applications", CRC Press, pp. 525, 2013
- [3] Karâa, Wahiba Ben Abdessalem, "Biomedical Image Analysis and Mining Techniques for Improved Health Outcomes", IGI Global, pp. 335, 2015
- [4] O. Vermesan, P. Friess, "Internet of Things: Converging Technologies for Smart Environments and Integrated Ecosystems", River Publishers, Technology & Engineering, pp. 364, 2013
- [5] Marbán, Óscar, Gonzalo Mariscal, and Javier Segovia. "A data mining & knowledge discovery process model." Data Mining and Knowledge Discovery in Real Life Applications 2009 (2009): 8.
- [6] Hakeem, Asaad, Himaanshu Gupta, Atul Kanaujia, Tae Eun Choe, Kiran Gunda, Andrew Scanlon, Li Yu et al. "Video analytics for business intelligence." In Video Analytics for Business Intelligence, pp. 309-354. Springer Berlin Heidelberg, 2012.
- [7] G. Khanvilkar, D.B. Meshram, "Video Data Mining: Event Detection from the Association Perspective using FP-growth Tree", International Journal of Engineering Research and Applications (IJERA) Volume, 1, 2011

- [8] Kohun, Frederick G., and Gary J. DeLorenzo. "Simplified Procedures in Digital Video Editing: Concepts and Technological Alternatives", Retrieved, 22nd June, 2016
- [9] N. Dimitrova, H-J. Zhang, B. Shahraray, I. Sezan, T. Huang, and A. Zakhor, "Applications of video-content analysis and retrieval." *IEEE multimedia* 9, no. 3, 42-55, 2002.
- [10] R. Agrawal, T. Imielinski, and A. Swami, "Database mining: A performance perspective", *IEEE transactions on knowledge and data engineering*, Vol./ 5, no. 6 , pp. 914-925, 1993.
- [11] B.V. Patel and B. B. Meshram. "Content based video retrieval systems." *arXiv preprint arXiv:1205.1641*, 2012
- [12] G. Karypis, E-H. Han, and V. Kumar. "Chameleon: Hierarchical clustering using dynamic modeling." *Computer* 32, no. 8, 68-75, 1999.
- [13] D. Saravanan, and S. Srinivasan. "Video image retrieval using data mining techniques." *Journal of computer applications (JCA)* 1 39-42, 2012
- [14] S-C. Chen, M-L. Shyu, C. Zhang, and J. Strickrott, "Multimedia Data Mining for Traffic Video Sequences." In *MDM/KDD*, pp. 78-86. 2001
- [15] S.P. Algur, P. Bhat, "Web Video Object Mining: Expectation Maximization and Density Based Clustering of Web Video Metadata Objects", *IJ. Information Engineering and Electronic Business*, pp. 69-77, 2016
- [16] V.A. Paliwal, N. R. Adam, H. Xiong, and C. Bornhovd. "Web service discovery via semantic association ranking and hyperclique pattern discovery." In *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 649-652. *IEEE Computer Society*, 2006.
- [17] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. Luisa Damiani et al. "Semantic trajectories modeling and analysis." *ACM Computing Surveys (CSUR)* 45, no. 4, pp. 42, 2013.
- [18] J. Han, J. Pei, and X. Yan, "Sequential pattern mining by pattern-growth: Principles and extensions." In *Foundations and Advances in Data Mining*, pp. 183-220. *Springer Berlin Heidelberg*, 2005.
- [19] K. Saxena, and D. S. Rajpoot. "A Way to Understand Various Patterns of Data Mining Techniques for Selected Domains." *arXiv preprint arXiv:0911.0781*, 2009
- [20] R. Srikant, and R. Agrawal, "Mining sequential patterns: Generalizations and performance improvements", In *International Conference on Extending Database Technology*, pp. 1-17. *Springer Berlin Heidelberg*, 1996.
- [21] T. Srivastava, P. Desikan, and V. Kumar, "Web mining-concepts, applications and research directions", In *Foundations and advances in data mining*.. *Springer Berlin Heidelberg*, pp. 275-307, 2005.
- [22] R. Jain, and G.N Purohit, "Page ranking algorithms for web mining." *International journal of computer applications* 13, no. 5, pp.22-25, 2011.
- [23] J.H. Oh, J.K. Lee, and S.k. Kote. "Real time video data mining for surveillance video streams." In *Pacific-Asia conference on knowledge discovery and data mining*, pp. 222-233. *Springer Berlin Heidelberg*, 2003.
- [24] J. Meng, J. Yuan, M. Hans, and Y. Wu, "Mining motifs from human motion." In *Proc. of EUROGRAPHICS*, vol. 8. 2008
- [25] G.M. Weiss, "Data mining in telecommunications", In *Data Mining and Knowledge Discovery Handbook* (pp. 1189-1201). *Springer US*, 2005
- [26] D. Pallez, L. Brisson, and T. Baccino. "Towards a human eye behavior model by applying Data Mining Techniques on Gaze Information from IEC." *arXiv preprint arXiv:0803.3186*, 2008
- [27] H. Banaee, M. U. Ahmed, and A. Loutfi, "Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges." *Sensors* 13, no. 12, 17472-17500, 2013
- [28] K.I. Asad, T. Ahmed, and Md S. Rahman. "Movie popularity classification based on inherent movie attributes using C4. 5, PART and correlation coefficient." In *Informatics, Electronics & Vision (ICIEV)*, 2012 *International Conference*, pp. 747-752. *IEEE*, 2012.
- [29] S. Kabinsingha, S. Chindasorn, and C. Chantrapornchai. "Movie Rating Approach and Application Based on Data Mining." *International Journal of Engineering and Innovative Technology (IJEIT)* Volume 2, 2012
- [30] M. Saraee, S. White, and J. Eccleston. "A data mining approach to analysis and prediction of movie ratings." *Transactions of the Wessex Institute*, pp.343-352, 2004
- [31] C.K.A. Bhatt, and M. S. Kankanhalli. "Multimedia data mining: state of the art and challenges." *Multimedia Tools and Applications* 51, no. 1, pp.35-76, 2011.
- [32] P. Thirumurugan, and S. Hasan Hussain, "Event detection in videos using data mining techniques." *International Journal of Computer Science and Information Technologies* 3, no. 2, pp.3473-3475, 2012
- [33] A. Divakaran, K. Miyahara, K. A. Peker, R. Radhakrishnan, and Z. Xiong. "Video mining using combinations of unsupervised and supervised learning techniques." In *Electronic Imaging 2004*, pp. 235-243. *International Society for Optics and Photonics*, 2003.
- [34] S. Shekhar, R. Michael Evans, M. J. Kang, and P. Mohan. "Identifying patterns in spatial information: A survey of methods." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1, no. 3, 193-214, 2011
- [35] J.Y. Lee, and W. Hoff, "Activity identification utilizing data mining techniques." In *Motion and Video Computing, WMVC'07. IEEE Workshop*, pp. 12-12, 2007
- [36] P. Angelov, P. S-Tehran, C. Clarke, "AURORA: autonomous real-time on-board video analytics", *Neural Comput & Applica*, 2016
- [37] V.N.M. Aradhya and M. S. Pavithra, "A comprehensive of transforms, Gabor filter and k-means clustering for

- text detection in images and video", *Applied Computing and Informatics*, 2014
- [38] A. Aryanfar, R. Yaakob, A. A. Halin, N. Sulaiman, K. A. Kasmiran, and L. Mohammadpour, "Multi-View Human Action Recognition Using Wavelet Data Reduction and Multi-Class Classification", *Procedia Computer Science*, Vol. 62, pp.585-592, 2015.
- [39] A.B. Ayed, M. B. Halima, and A. M. Alimi, "MapReduce Based Text Detection in Big Data Natural Scene Videos", *Procedia Computer Science*, Vol. 53, pp.216-223, 2015.
- [40] P. Vollucci, B. Le, J. Lai, W. Cai, Y. Ye, G. Nacula, and D. Wroblewski, "Online Video Data Analytics", 2015
- [41] C.H.E. N. Qing, Y. Chen, D. Liu, C. Shi, Y. Wu, and H. Qu, "PeakVizor: Visual Analytics of Peaks in Video Clickstreams from Massive Open Online Courses", 2015.
- [42] Y.B. Kim, S. J. Kang, Sang Hyeok Lee, Jang Young Jung, Hyeong Ryeol Kam, Jung Lee, Young Sun Kim, Joonsoo Lee, and Chang Hun Kim. "Efficiently detecting outlying behavior in video-game players." *PeerJ* 3 (2015): e1502.
- [43] B. Mao, J.He, J. Cao, S.W. Bigger, and T. Vasiljevic, "A framework for food traceability information extraction based on a video surveillance system", *Procedia Computer Science*, Vol. 55, pp.1285-1292, 2015.
- [44] M. Shao, and Y.Fu, "Deeply Self-Taught Multi-View Video Analytics Machine for Situation Awareness", *AFA Cyber Workshop*, White Paper, 2015
- [45] T. Xu, W. Xiang, Q. Guo, and L. Mo, "Mining cloud 3D video data for interactive video services", *Mobile Networks and Applications*, Vol. 20, No. 3, pp. 320-327, 2015.
- [46] F. Riahi, O. Schulte, and Q. Li, "A Proposal for Statistical Outlier Detection in Relational Structures", In *AAAI Workshop: Statistical Relational Artificial Intelligence*, 2014.
- [47] D. Potapov, M. Douze, Z. Harchaoui, and C. Schmid, "Category-specific video summarization", In *European conference on computer vision*, pp. 540-555, 2014.
- [48] A.K. Thilagavathy, Aarthi, and A. Chilambuchelvan, "Scene text extraction from videos using hybrid approach", In *Advances in Computing and Information Technology*, Springer Berlin Heidelberg, pp. 739-748 2013.
- [49] Y. Cong, J. Yuan, and Y. Tang, "Video anomaly search in crowded scenes via spatio-temporal motion context", *IEEE transactions on information forensics and security*, Vol. 8, No. 10, pp.1590-1599, 2013.
- [50] M. Bertini, A.D. Bimbo, and L. Seidenari, "Multi-scale and real-time non-parametric approach for anomaly detection and localization", *Computer Vision and Image Understanding*, Vol. 116, No. 3, pp. 320-329, 2012.
- [51] F. Jiang, J. Yuan, S.A. Tsafaris, and A. K. Katsaggelos, "Anomalous video event detection using spatiotemporal context", *Computer Vision and Image Understanding*, Vol. 115, No. 3, pp. 323-333, 2011.
- [52] Y. Deldjoo, M. Elahi, P. Cremonesi, F. Garzotto, P. Piazzolla, and M. Quadroni, "Content-Based Video Recommendation System Based on Stylistic Visual Features", *Journal on Data Semantics*, pp.1-15, 2016.
- [53] H. Nomiya, A. Morikuni, and T. Hochin, "Unsupervised Emotional Scene Detection for Lifelog Video Retrieval based on Gaussian Mixture Model", *Procedia Computer Science*, Vol. 22, pp. 375-384, 2013.
- [54] C. Devasena, Lakshmi, and M. Hemalatha, "Video Mining using LIM Based Clustering and Self Organizing Maps", *Procedia Engineering*, Vol. 30, (2012): 913-921.
- [55] A. Kumar, A. K. Kaushik, and R. L. Yadav, "A robust and fast text extraction in images and video frames", In *Advances in Computing, Communication and Control*, Springer Berlin Heidelberg, pp. 342-348, 2011.
- [56] V. Saligrama, J. Konrad, and P-M. Jodoin, "Video anomaly identification", *IEEE Signal Processing Magazine*, Vol. 27, No. 5, pp. 18-33, 2010
- [57] A. Sureka, P. Kumaraguru, A. Goyal, and S. Chhabra, "Mining youtube to discover extremist videos, users and hidden communities", In *Asia Information Retrieval Symposium Springer Berlin Heidelberg*, pp. 13-24, 2010.