Multimedia Data Mining

Dr. Prerna Mahajan[1], Vartika Bisht[2]
[1] IITM, GGSIPU/MCA, New Delhi, India
Email: hod.csdept@iitmipu.ac.in
[2] IITM, GGSIPU/MCA, New Delhi, India
Email: vartikabisht55@gmail.com

Abstract — **It has been seen that necessity is the mother of invention and human being fulfill all their requirements by developing unbelievable things. Recently, people are taking advantage of detectors for extracting useful information. In this paper we discuss Multimedia Data Mining to perform above task and multimedia data categorization.**
With advancement in technology an outsized quantity of multimedia system information has been taken to analysis centers for learning various factors altogether. Associate approach was planned to use data mining for multimedia data is called Multimedia Data Mining. Multimedia Data Mining is pattern discovery, rule extraction and data acquisition from database. To extract data from multimedia database, some techniques are used in this research we are using Multimedia Data Mining to extract the patterns for above to problems. For this, two common terms exist. The first one is, the multimedia system information data should be segregated and emerge as objects and second is, the behavior of these objects gets monitored for immediate higher cognitive purposes. Model and index the information for on-line analysis, storage or later pattern mining.
After studying categories of Multimedia Data Mining video surveillance system has been taken into consideration and some techniques are observed to detect motion for instance, motion segmentation, simultaneous partition and class parameter estimation (SPEPE) and background subtraction. This paper also helps to compare the clustering techniques that can be used for video surveillance that is used in traffic control systems.

Index Terms — **Multimedia data mining, Background Subtraction, Motion Segmentation, Video Surveillance**

I. Introduction

In today's time, with the advancement in technologies, we have seen that many times, road accidents, train accidents occur despite of sophisticated control mechanisms. Various kinds of problems may occur in traffic control, for instance, We all have seen that in railway system generally the temporal order of two trains collapse and one has to wait after leaving the metro station to get into the upcoming metro station in between somewhere on the track.

The other complications emanate when passenger board a train and unanticipatedly the doors get closed because the drivers simply uses mirror to check whether or not all have boarded the train or not.

In this paper, we also propose to identify various techniques which are used to find out patterns using MDM.

- To control the overlapping lines traveling in the same direction that can cause metro congestion i.e. to provide train stopping accuracy
- Recognition of visual events from video sequences i.e. Object and Event detection

These issues may be resolved to some extent using MDM techniques. MDM focuses to handle these botherations. The standard data mining process consists of many stages and therefore the overall process is inherently interactive and unvarying [1]. The data mining process contains following stages -

(1) **Domain** understanding- needs learning however the results of data-mining are going to use all relevant previous information.

(2) **Knowledge selection**- needs the user to focus on information or select a set of fields or knowledge records to be used for data processing.

(3) **Knowledge preprocessing, improvement and transformation**- Its goal is to find necessary choice from data, take away the noise and normalize it.

(4) **Discovering patterns** - It involves different methods like time-series analysis, and visualization at intersection of artificial intelligence, association, classification, clustering, regression, machine learning, statistics and database systems.

(5) **Interpretation** - It figure out the standard of discovery and its value to work out whether or not the previous stages ought to be revisited or not.

(6) **Reporting and using discovered knowledge** - it's the ultimate stage of the information mining method that consists of reporting and to use the discovered knowledge. This stage is application dependent [2].

Thus, MDM State of the Art review is organized in given figure (fig-1)-

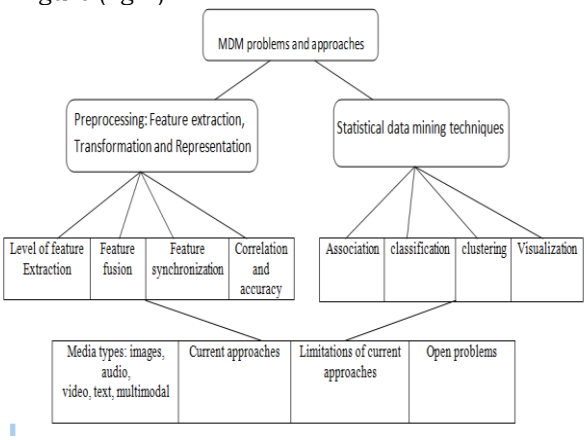

Figure 1: MDM Techniques: State of art

There are three types of multimedia data- static, dynamic and dimensional

- Static media- time-independent: image and graphic object.
- Dynamic media- time-dependent audio, video and animation.
- Dimensional media- 3D game and computer aided drafting programs.

Here we use the combination of static and dynamic data to find out different patterns.

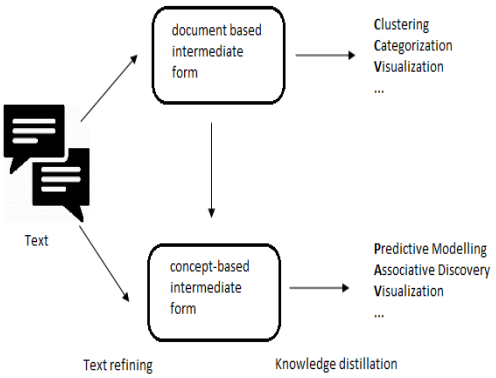

Figure 2- Refining text framework

A. Text mining

In this, the text data get processed and provide useful information known as patterns. This uses a technique of finding nuggets, finding patterns, text visualization. It leads to find and rank the text data which is useful. This refining leads to provide valuable information. Example- digital library, creating SMS, MMS.

It converts unstructured text into an intermediate form. Intermediate form can be document-based. Knowledge distillation from a document-based intermediate form deduces patterns or knowledge across documents. A document-based intermediate form is often projected onto a concept-based intermediate form by extracting object information relevant to that concern. Knowledge distillation from a concept-based intermediate form deduces patterns or knowledge across objects or concepts. Given below is the framework of refining the text [3].

B. Image mining-

In this information has been congregated from large collection of images. It helps in categorizing and annotating information It has two variants-

a) Mining the large stack of image

b) Combine data mining of large collection of image and associated alphanumeric data

This can be explained by considering the following case - a collection of weather satellite imagery of various cities in the India that has been recorded over an extended period of time. Here image data mining objective might be to find if there is some pattern that exists for an individual city (over time) or if there is some pattern that exists between different cities. This is the example of the first type (includes photos).

An example of the second type (like medical images) might involve medical imagery and patient (alphanumeric data) records. To develop an accurate diagnosis or prognosis both image data (such as X-rays, SPECT, etc.) and patient data (such as weight, prior health conditions, family history, etc.) can be examined together to find interesting clustering [4].

The fundamental challenge in image mining is to find out how low-level pixel representation confined in a raw image or image sequence can be processed to observe high-level image objects and relationships. Fig. given below illustrates the typical Image Mining Process

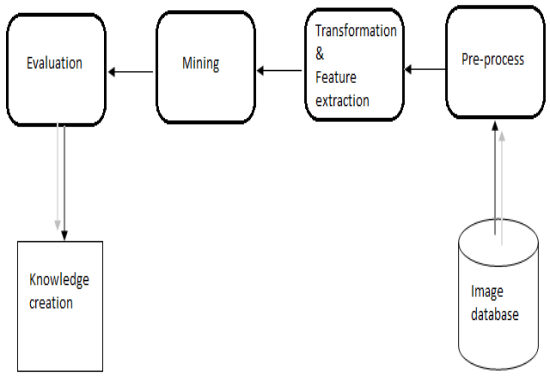

Figure 3- Evaluation of image refining

C. Audio mining

It is a stratagem by which constituent of audio signal can be explored and evaluated with a process called wavelet transformation. As we know that web, database and other digitized information storehouse contain growing volume of audio content, example sporting events, telephone, etc. audio mining uses audio indexing and further audio indexing uses speech recognition to enquire the whole file and turn out searchable index of content bearing words and their location. Indexing audio content thus enables searching. Audio mining has 2 approaches-

i) Text based indexing -

It helps to convert speech into text and then identifies word in dictionary that can contain thousands of entries. If a name is not in given dictionary, the system will automatically choose the most similar word it can find.

ii) Phoneme based indexing

It works only with sounds. In this, the system first analyses and identifies sound in a part of audio collection to create Phoneme based index. Then it uses dictionary of many phonemes to convert a user's search term. Here, phoneme is the smallest unit of speech in the language used for speech [5].

D. Video mining-

Video mining used to find out similar patterns. Such patterns can help to explain, classify, and search video contents. There are many related applications, such as commercial detection where investigation has been done to find speech patterns. One of the major areas of interest of video mining is video surveillance system.

Video Surveillance System

We know that intelligent video surveillance system deals with the monitoring of real time environment. It monitors the transient and specific object within an environment. It not only provides security but also helps in external environmental video surveillance [6]. Since in metro we have to put our

focus on object detection and motion of train/object/person, this can be done by various motion detection methods like motion segmentation and background detection.
Motion segmentation-it is basically a threshold taken by difference between the present image and sequence images and assuming that the background does not change over entire frames [8]. For this, some of the techniques are used in which SPCPE is of them. The Simultaneous Partition and Class Parameter Estimation (SPEPE) algorithm is an unsupervised video distribution approach to partition video frames. This technique is simple and quick in several applications; however some issues arise when tracking multiple objects or when an object stops [7]. Hence, we tend to communicate the opposite technique, background subtraction, at the expense of changing the background.

Background subtraction

Video surveillance system has common case in which static camera is used to detect a scene. Detecting intruding objects is an essential step in analyzing the scene. A usually applicable assumption is that the images of the scene without the intruding objects exhibit some regular behavior. If scene has been revealed, the disturbed object can be detected by spotting the parts of the image that don't fit to expectation. This process is usually known as "background subtraction" [9]. The blueprint of background subtraction is to deduct the image from mentioned image that models the background scene. The key steps involved for background subtraction are:

- Background modeling constructs a reference image representing the background.
- Threshold choice determines applicable threshold values utilized in the subtraction operation to get a desired detection rate.

We are performing background subtraction using different clustering algorithms
This research paper not only helps to understand how to provide train stopping accuracy and object and event detection but also presents a comparison of sophisticated clustering methods.

MDM Operation Using Clustering based Techniques:

Clustering is used to organize objects into groups having similar members. It is one of the unsupervised learning data mining technique. In unsupervised classification, issue arises when we want to group a collection of unlabeled multimedia files into meaningful clusters.
Most of the researchers worked on finding patterns (abnormal stream) from videos. Initially the focus was on color feature in order to perform clustering and determine different patterns but the drawback was that in the early day's single stationary camera were used which had many more limitations and that was used only to focus on single activity at stationary places. Later on moving camera at the stationary places has been used which has cover large geographical area but still having some lacunas. Previously the focus was on the fundamental classification technique which was unable to produce the accurate results. Each technique which was developed has it merits and demerits. But monitoring of the videos at real time and find out the abnormal activities is really a challenging work for the techno savvy researcher. In the recent years there were a many more clustering algorithms are used such as k- means algorithm, Dynamic Oriented graph, Markov Model but still need a new technique in order to produce the richer output [10].

II. Comparison

Here a table is used to compare the existing techniques on the basis of their working, advantages and disadvantages. The algorithm uses both Background Subtraction and Symmetrical Differencing methods to obtain the moving object or targets.

Techniques	Principle	Advantages	Disadvantages	Application
Clustering Segments by K-means Algorithm [11]	The algorithm uses both Background Subtraction and Symmetrical Differencing methods to obtain the moving object or targets. It has the ability to segment the moving object integrity, and achieve real-time motion detection after building the background.	• K-means clustering provides pixel-wise object tracking. In order to achieve the robust object tracking under complex condition (such as wired objects, cluttered background), a new reliability-based K-means clustering algorithm is applied to remove the noise background pixel (which is neither similar to the target nor the background samples) from the target object.	• It is only suitable to work indoors, such as the subway station, airport, etc, since it is able to track the target object by checking the differences between the observed image and the predefined background model. It also suffers from the random noise (such as rains or snows) and is not suitable to work in the dynamic scene. • Although K-means clustering is one of the most popular pixel-wise clustering algorithms and can be directly applied into object tracking, two problems degrade its performance: • When the noise data not belonging to any cluster exist, K-means clustering will wrongly classify them into some	• It is used to analyze student record. • It is used to control traffic • It is used in metro stations

			predefined clusters; • The wrongly assumed number of clusters sometimes leads to the wrong clustering result. Some researchers brought out some improvements on K-means clustering, while object tracking such problems still affect the performance of K-means clustering	
Probabilistic Latent Semantic Analysis (PLSA) [12]	It is a statistical methodology serving the purpose of dissection or close examination of two-mode and co-occurrence data. The constituents are: • Low-level HDP (Hierarchical Dirichlet Process) • Higher-level temporal motif	• Ambiguous patters which seem to be having multiple inferences are captured. • Distribution pattern over space of topics in images is very well captured. • HDP • It is efficient and scalable	• Needs other models like Hidden Markov Model to capture temporal patterns • Over fitting, can't jointly process • Not supervised • Can't handle multiple cameras	The practical application is that, a low-dimensional miniature model can be reached, out of observed patterns based on their resemblance to certain hidden patterns, similar to as seen in latent semantic analysis technology/procedure, which is the building foundation for PLSA. • Measure for Abnormality Rating Model Behavior on Synthetic Activities, Abnormality Rating in a Metro Station. • PLSA is adapted to capture frequently co-occurring pixels in images
Dynamic Oriented Graph (DOG) [13]	In this, video cameras situated at stationary places obtain images and art algorithm is applied to those images so that the image can be tracked or classified.	• It is one of the new approaches used to detect and predict abnormal behavior of object. • DOG classifier is extremely fast, leaning, classifying and predicting on line	• DOG Classifier doesn't detect danger condition when interaction of multiple objects occurs. • For example, if there is rush in platform and suddenly metro arrives, people started running and interaction between them occurs lead to stampede. In such case DOG will not able to recognize that pattern	• It is used in video cameras to obtain color images. • It is used to find out stationary object. • It observes actions and then characterizes them by means of node's structure. • It is one of the new approaches used to detect and predict abnormal behavior of object.
Markov model [14]	It is commonly known method in which future state depends only on present state not on sequence of events. For finding the patterns a video sequence has been taken into consideration. Each time next image is taken and then the difference between the image and the image taken from reference is measured. Find the threshold of taken output to find the connected pattern. If the requirement has not been fulfilled proceed with new track. The tracked file will provide the information about the location of object. "Bounding box " is used to track the information of that object and then "shape template" is calculated.	• It is one of the most original method to find different activities of trajectory object	• It is not applicable for more different type of activities. • For example, the object with totally different speed then a normal man in platform. This is not necessary to be performed in metro platform only. It can be used in tracks also	• It is used to store the information about the location of moving object • The information is tracked form "track file"
Bipartite graph co	Simple features of an object are taken to	• It automatically select important video signal. • It is used for co-embedding	• Since, it is an unsupervised technique; it does not select supervised features.	It is used to track moving objects.

clustering [15]	find out the unusual activities in video. It is an unsupervised technique. In this, video is divided into equal length segments and prototypes are made and then prototype segment co-occurrence matrix is computed.	problems. • It automatically select the important feature signal to utilize extremely simple features	• Complex activities are not recognized by it.	It can be used for security purpose at home
Total Motion (TM), Object Motion (OM), Camera motion (CM) [16]	It uses accumulation of quantized Pixel differences among all frames in a shot. It also examines each frame in a shot to check whether there are any camera motion changes and if so it computes the amount of motions and their directions. It computes Object Motion (OM) by compensating camera motion changes. OM and TM both are represented as two dimensional matrixes. Finally, Camera motion (CM) can be computed subtracting OM from TM such that CM =TM - OM. The matrices representing TM and OM are showing not only the amounts of motions of the object but also show the exact locations of motion of the object.	• It measures overall motion in a single shot. • It is more effective and gives specific information of motion in a shot.	• It seems that sometimes experimental data set has limited number of shots	• it is used in wildlife photography • it is used at areas with less traffic
Two-level hierarchical clustering [17]	It uses 2-D tensor histograms to compute motion features and color features are represented by 3-D color histograms. Cluster validity analysis is further applied to automatically determine the number of clusters at each level.	• No apriori information about the number of clusters required. • Easy to implement and gives best result in some cases.	• Algorithm can never undo what was done previously. • this is not a suitable choice for most applications, because it can lead to clusters that are quite heterogeneous internally, and the usual object of clustering is to obtain homogeneous clusters	• it is used to find out the distance between I and j records • it is used to predict the cost impact

CONCLUSIONS

This paper provides a quick review of MDM. It explains both static and dynamic object's mining. The comparison among the techniques has been shown on the basis of their basic method, merits, demerits and applications. Some techniques only provide information segmentation and some provide both information segmentation and object tracking. The comparison table shows that each technique has different principle in contrast of solving traffic control problem

REFERENCES

[1] Dr. S.Vijayarani1 and Ms. A.Sakila2," MULTIMEDIA MINING RESEARCH – AN OVERVIEW", International Journal of Computer Graphics & Animation (IJCGA) Vol.5, No.1, January 2015

[2] Manjunath T.N, Ravindra S Hegadi, Ravikumar GK, "A Survey on MDM and Its Relevance Today" IJCSNS International Journal of Computer Science and Network Security,VOL.10 No.11, November 2010

[3] https://en.wikipedia.org/wiki/Text_mining

[4] A.Kannan, Dr.V.Mohan, Dr.N.Anbazhagan,"An Effective Method of Image Retrieval using Image Mining Techniques", The International journal of Multimedia & Its Applications (IJMA) Vol.2, No.4, November 2010

[5] https://en.wikipedia.org/wiki/Audio_mining

[6] Wei Niu, Long Jiao, Dan Han, and Yuan-Fang Wang,"Real-Time Multi-person Tracking in Video Surveillance", Department of Computer Science University of California Santa Barbara, CA 93106

[7] Shu-Ching Chen , Mei-Ling Shyu , Chengcui Zhang , Jeff Strickrott ,"MDM for Traffic Video Sequences" , School of Computer Science, Florida International University, Miami, FL 33199

[8] Luca ZAPPELLA aXavier LLADÓ a and Joaquim SALVI a a, "Motion Segmentation: a Review", Institute of Informatics and Applications, University of Girona, Girona (Spain)

[9] Ashish Kumar Sahu1 Abha Choubey, "Motion Detection Surveillance System Using Background Subtraction Algorithm", International Journal of Advance Research in Computer Science and Management Studies

[10] Mr.Avinash P Ingle, Ms. Snehlata Dongre, "A Survey on Data Mining Techniques for Surveillance of Real Time Video Streams", International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-2 Number-4 Issue-6 December-2012

[11] Jinghua Wang and Guoyan Zhang, "Video Dat a Mining based on K-means Algorithm for Surveillance Video", 2011 IEEE

[12] Jagannadan Varadarajan, Jean-Marc Odobez, R´emi Emonet "Mult i-camera Open Space Human Act ivity Discovery for Anomaly Det ect ion", 8th IEEE Internat ional Conference on Advanced Video and Signal-Based Surveillance, 2011.

[13] Jae Young Lee, William Hoff," Activity Identification Utilizing Dat a T echniques", IEEE Workshop on Mot ion and Video Comput ing (WMVC'07).

[14] Duarte Duque, Henrique Santos and Paulo Cortez," Prediction of Abnormal Behaviors for Intelligent Video Surveillance Systems", Proceedings of the 2007 IEEE Symposium on Computat ional Intelligence and Data Mining (CIDM 2007).

[15] Hua Zhong, Jianbo Shi, Mirk´o Visontai, " Detecting Unusual Activity in Video", Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04).

[16] JungHwan Oh, P raveen Sankuratri, "Automatic distinction of camera and object mot ions In video sequences", C0 2002 IEEE

[17] Ngo. C Pong T. Zhang H," On Clustering and Ret rieval of Video Shot s", In: P roc of ACM Multimedia 2001,Ot tawa,Canada(2001),51-60.