A Novel Mining System for Criminal Issues from a Video File Within Cloud Computing Environment

Emad Mohammed Ibbini, Mohammed Salameh Ibbini, Kweh Yeah Lun

***Abstract*– This paper presents a description of a novel mining system which mines the different occurrences (instances) of the same object from a video file.**

The framework of the system consists of four steps: segmenting the video file into stable tracks, extracting objects and their features from the tracks, grouping these tracks into clusters based on their residing objects, and finally mining the instances of each object in the shared pool of configurable computing resources within cloud environment for more security.

The paper also presents a critique and feedback for the system and proposes an idea to improve its performance.

***Index Term*-- segmenting ,mining system, cloud computing, Instances ,Objects.**

1. INTRODUCTION

Images and video captures wide area in our life (digital images, news, movies, and surveillance systems). The increased connectivity to the web provides more facilities to exchange image and video files around the world.

Video mining is the process of discovering knowledge in video files. Some times we need to get the different instances of some object in a video file. This process could be important in movies and surveillance systems. This paper is a good contribution in this field within cloud environment .

The paper describes a system for mining the different instances of the same object from video file. System starts with video segmentation which is presented in section (2), section (3) presents the second step of the system which talks about obtaining the object from the video segments, in section (4) the mining process is described, section (5) shows the implementation and results.

Both sections (6, 7) respectively show a critique and proposed improvement for the system, and finally section (8) concludes the paper and talks about future work.

2. CLOUD COMPUTING - OVERVIEW

Cloud computing model divided into public, private or hybrid. Cloud computing is where the organization outsources data processing to computers owned by the vendor.

Primarily the vendor hosts the equipment while the audited entities still has control over the application and the data. Outsourcing may also include utilizing the vendor's computers to store, backup, and provide online access to the organization data.

Management Components of Cloud Computing [3]

Cloud management system is separated into four layers: the Resources & Network Layer, Services Layer, Access Layer, and User Layer.

Each layer contains a set of operations:

The Network Layer manages the physical resources.

The Services Layer contains the types of cloud services, namely, NaaS, IaaS, PaaS, the service orchestration function and the cloud operational function.

- The Access Layer includes API termination function, and Inter-Cloud peering and federation function.
- The User Layer includes End-user function, Partner function and Administration function.

Segmentation video within cloud environment

This object extraction will be more effective in a cloud computing environment when the load will be on remote machines rather than clients with minimum processing power and better performance for video segmentation as services are delivered on demand to outside organizations over the internet.

3. VIDEO SEGMENTATION

In order to control video file and to efficiently mine its contents we need to segment it into smaller pieces. In general video is segmented hierarchically into smaller pieces which are in turn segmented into smaller and smaller pieces until we have small frames of video as the lowest level of segmentation.

The segmentation process starts with dividing the video into shots which is the basic element of video. Shot can be defined as an uninterrupted sequence of frames recorded from a single camera operation. Each shot in turn is divided into tracks where each track contains more than one frame. Figure (1) shows the followed segmentation.

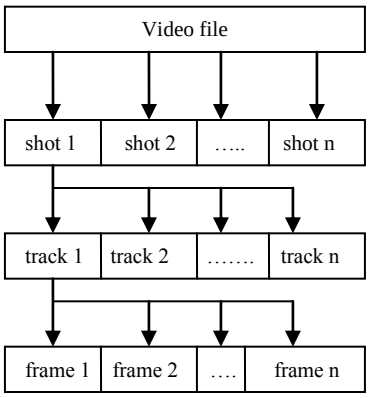

Fig. 1. video segmentation

4. OBTAINING OBJECTS FROM THE SHOTS

To obtain the representation and description of objects from the shots of the video we need two types of algorithms, *region extraction* algorithm, and *feature description* algorithm.

Region extraction algorithm is used to detect the presence of some object in the video shot and feature description algorithm is used to detect and describe the local features of the existed object in the video.

The system uses a combination of MSER and SIFT algorithms as a region extractor and feature descriptor respectively. Any combination of another extractor and descriptor algorithms works fine with this system.

Both of the extracted regions and described features are used to group the tracks of the video into clusters. For each shot, system assigns the tracks of the same object to a single cluster.

All the instances of the same object are grouped into the same cluster. The following algorithm is used to group the tracks of the shot into clusters depending on the object which occurs in the tracks.

Algorithm (1): clustering tracks of the shot
Initialization phase:

Assign the longest track of the shot to the first cluster.

Iteration phase:

Take the next track and compute the distances between it and the current clusters.
The closest cluster (C) from the track is identified.*
The distance between the track and the closest cluster is compared to some predefined threshold value.
If the distance is less than the threshold value then assign the track to the closest cluster(C), else create a new cluster and assign the track to it.*

Termination phase:

The algorithm terminates when no changes occurs to the tracks (tracks are stable).

End algorithm

Figure (2) shows the results of applying this algorithm on some shot.

5. MINING SIMILAR OBJECTS

Obtaining the objects from the shots is not enough to mine the instances of the target object from the whole video. We have more than one shot in the video file and hence we have more than one cluster for the same object generated by the different shots. This is demonstrated in figure (3).

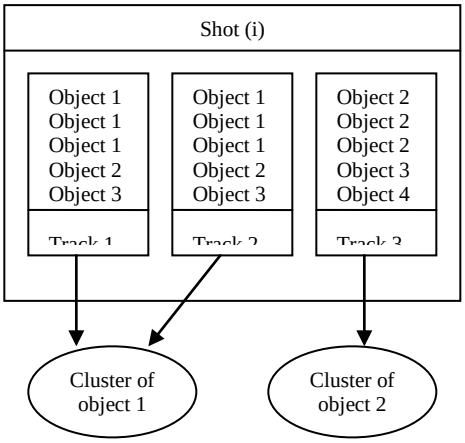

Fig. 2. clustering the tracks of the shot

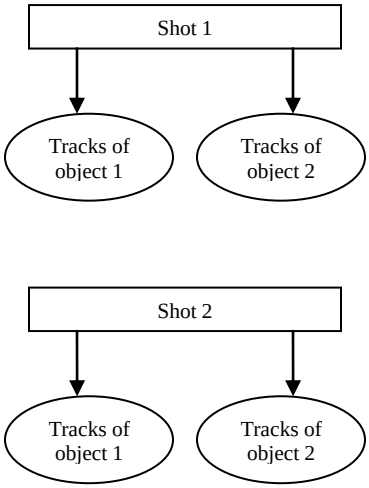

Fig. 3. different shots generating clusters of the same object

The instances of the same object may be dispersed in different shots of the video file. To mine the instances of the same object, system merges the different clusters of the same object into a single large cluster. Figure (4) shows the merge process.

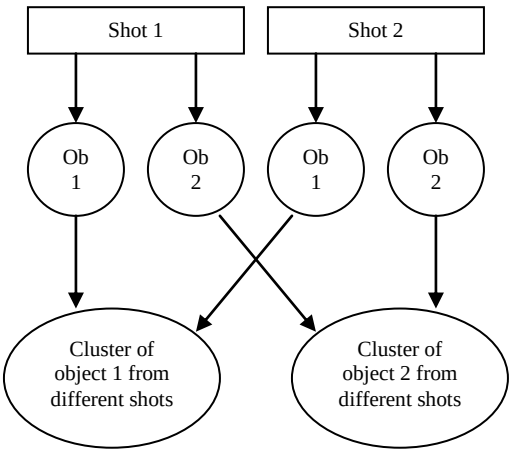

Fig. 4. the instances of the same object in different shots are combined to form a single cluster

Now the extracted instances of the same object are in the same cluster. System can mine the required object easily.

6. Implementation and Results

This system serves for the purpose of full movie files, so it was applied to a full movie *"groundhog day"*.

The movie was segmented into stable tracks. The combination of MSER and SIFT algorithms was used to extract and describe the objects of the movie. Algorithm (1) was used to group tracks of each shot into clusters depending on their objects.

The system mined the instances of the objects from different clusters and classified them into relevant groups.

The system considers the mined object as a valid if it appears at least in five different shots of the movie. Under this condition 147 objects were mined from the whole movie.

Table (1) shows a sample of the mined objects, number of correctly mined clusters, number of missed clusters, and number of wrongly mined clusters. Recall and precision is computed for each mined object and promising results is gotten.

Table I
results of some of the mined objects

Mined Object	Correct	Miss	False	Precision	Recall
	81	292	1	0.988	0.217
	70	131	1	0.986	0.348
	6	4	0	1.000	0.600
	5	2	0	1.000	0.714
	13	1	0	1.000	0.929
	5	1	0	1.000	0.833
	5	0	0	1.000	1.000
	5	0	0	1.000	1.000
	5	18	0	1.000	0.217
	6	6	0	1.000	0.500
	23	3	0	1.000	0.885
	8	4	1	0.889	0.500

7. Where is the Gap

Examining the sample of the results in table (1), we can conclude the following advantages of the proposed system.

1. System mines the objects at a high precision ratio (89-100%).
2. System mines the object at an accepted Recall ratio (70-100%)
3. System mines more instances of the same object compared to other systems.

Some results in table (1) have a large number of missed clusters which is greater than the correctly mined clusters. Other results have a low Recall ratio (21%) how to justify that. Examining figure (2), it could be seen that each track may contain more than one object. For example system assigns both track 1 and track 2 to the same cluster, cluster 1, because both of them contain more occurrences of object 1. System follows this strategy for all the tracks in all the shots.

That justifies the large number of missed clusters and the low ratio of Recall in some of the results.

8. Proposed Improvement

The proposed improvement is derived from the idea of *"image background subtraction"*, which is used by *M.Manoria* and *S.Dave* in their paper *"object extraction in data mining framework for video sequence"*. The authors propose a system to mine the moving objects from a video file.

8.1 Background subtraction

Video or image may contain several objects some of them could be in the area of our interest and some other are out of our interest.

Background subtraction is one of the mostly used ways to extract the objects in which we are interested from a given image or video.

Since this paper talks about video rather than image, we talk about background subtraction in video mining. To mine interesting objects from video, video is segmented into frames in order to enable object recognition and feature extraction.

System recognizes the interesting object of the frame and considers the remaining objects as a background. System uses the features of the background to construct a background model which is subtracted from the original frame.

8.1 Deploying background subtraction in the system

This subsection exploits the idea of background subtraction to enhance the performance of the system.

As mentioned in section (4), system has a single cluster for each object in the video. The cluster of each object contains the tracks of that object. However the track may contain occurrences of other objects as figure (2) shows.

Now we add a new step to the system framework. This step extracts the occurrences of the other object from each object cluster. The new step is an algorithm whose target is to get pure clusters of objects. That is each cluster of object contains only the occurrences of this object, no more no less.

Algorithm (2): background subtraction

- *Start with the first cluster*
- *Parse the first track and examine if it contains objects of other clusters*
- *Apply background subtraction for each of these objects one by one and move the resulted track to its appropriate cluster*
- *Iterate for all tracks in the current cluster*
- *Apply the algorithm for the entire clusters*

End Algorithm

CONCLUSION AND FUTURE WORK

Mining video depends mainly on the segmentation approach and the algorithms of object recognition and feature description.

The proposed system mines the different instances of the same object from video file through three steps: segmenting the video, extracting the objects and their features, and finally grouping the similar objects into clusters.

System is implemented on a complete movie and the results show that there is a gap. This gap comes from the non purity of the clusters. An improvement algorithm is proposed. It receives the non pure clusters as an input and return pure clusters as an output and because lose of trust we use this proposed mining system in Cloud environment for more security.

Future work is to realize and implement the algorithm and to apply the system with its improvement algorithm. The time complexity of the algorithm is $O(n^3)$ it needs more studying and improvement.

REFERENCES

[1] Arasanathan Anjulan and Nishan Canagarajah. "A novel video mining system". IEEE, 2007
[2] Prof. Manish Manoria and Ms. Stuti Dave. "Object Extraction in Data Mining Framework for Video Sequence". 2003
[3] Kangchan Lee , “Security Threats in Cloud Computing Environments”, International Journal of Security and Its Applications,Vol. 6, No. 4, October, 2012.
[4] Maximally Stable External Regions (MSER) algorithms, www.wikipedia.com.
[5] Scale-Invariant Feature Transform (SIFT). www.wikipedia.com