VIDEO-SYNOPSIS: AN AUTOMATIC SURVEILLANCE VIDEO ABSTRACTING SYSTEM

PROJECT REFERENCE NO.: 39S_BE_0655

COLLEGE: VIVEKANANDA COLLEGE OF ENGINEERING AND TECHNOLOGY, PUTTUR
BRANCH: DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
GUIDE: MR. HARIVINOD N
STUDENTS: MS. CHAITRA P V
           MS. HARSHITHA R
           MS. NIKITA P V
           MS. PRAJNA S K

INTRODUCTION:

With the rapid development of the digital-media industry, the huge video datasets captured by various resources such as webcams, surveillance cameras, are growing at an explosive speed. The amount of captured video is growing with the increased numbers of video cameras, especially the increase of millions of surveillance cameras that operate 24 hours/day.

In most site-seeing surveillance videos, we often observe large amounts of free space in the scene backgrounds, where no active objects move at all. It is time consuming to review entire, lengthy videos, captured by surveillance cameras, to find interesting objects, since most surveillance videos contain only a limited number of important events. Hence, end users often prefer briefer, condensed representations of long video sequences in surveillance videos. This is generally referred as video synopsis or abstraction. It is memory intensive to store and transfer the entire captured videos while retaining less important objects; thus, video synopsis can effectively reduce memory storage for large video datasets.

OBJECTIVE:

The main objective of this project is to develop a software system to condense the given surveillance video by preserving the essential activities. The activity in the video is condensed into a shorter period by simultaneously showing multiple activities, even when they originally occurred at different times. The summary will be in the video format with reduced size containing the actions present in original video.

METHODOLOGY:

The proposed system analyzes the video for interesting events and records an object-based description of the video. This description lists, for each camera, the interesting objects, their duration, their location, and their appearance. The reduction of the original video to an object-based representation enables a fast response to queries. After a user query about a second (shorter) time period, the requested abstract video is generated having only objects from the desired time period. A two-phase process is used for synopsis of video, is shown in Figure 1.
The steps can be summarized as follows.

**Phase 1: Online Phase.**
This phase can be done in real time.

1.1. *Creating a background video by temporal median:* The background for each time can be computed using a temporal median over a few minutes before and after each frame. For example, median over 3 minutes.

1.2. *Object detection and segmentation:* Background subtraction is combined together with min-cut to get smooth segmentation of foreground objects.

1.3. *Inserting detected objects into the object queue:* All detected objects, represented as tubes in the space time volume, are stored in a queue awaiting user queries.

**Step 2: Response Phase.**
This phase constructs a synopsis according to a user query. This phase may take a few minutes, depending on the amount of activity in the time period of interest. This phase includes the following:

2 1. *Constructing a time-lapse video of the changing background.* Background changes are usually caused by day-night differences, but can also be a result of an object that starts (stops) moving.

**CONCLUSION:**

We have captured 3 videos and used it as our dataset. Hall.avi, Mall.avi and College.avi. The results are summarized in Table 1. Fig 2 illustrates the indoor scenario captured in a hall. This is taken in a controlled environment. The empty frames in the scenarios are removed in the final output. The length of the video is 3:10 min. Snapshots of public area, a shopping mall surveillance video, are shown in Fig 3. This is an example for uncontrolled environment. As the crowd cannot be predicted. Here a video of length 0:16min was taken for processing. In Fig 4, the snapshots of the surveillance video captured at the college entrance are depicted. This video recording illustrates an outdoor scenario, which can be crowded at peak hours. This video is of length 2:29 min.
Video synopsis is a useful tool for summarizing long surveillance videos captured by digital cameras. It has been proposed as an approach for condensing the activity in a video into a very short time period. This condensed representation can enable efficient access to activities in video sequences and can enable effective indexing into the video. This enables fast browsing of surveillance video by bringing dynamic events widely separated in time closer together.

Video synopsis can make surveillance cameras more useful by giving the viewer the ability to view summaries of endless video stream. Video Synopses summarizes hours of surveillance video recordings into a short "brief" that takes only minutes to review. It filters out the empty frames present in the input surveillance video and provides a summary which is also a video that consists of only the activities that were present in the given video. We consider large changes as essential activities. Very minute changes are ignored. Video synopsis may also be applied for long movies consisting of many shots. In this case the method should be used for shot boundary detection and create video synopsis on each shot separately.

Table 1: Comparison Between Input And Output Video

<table>
<thead>
<tr>
<th>Name of the video</th>
<th>Original video duration (in minutes)</th>
<th>Summary video duration (in minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hall.avi</td>
<td>3:10</td>
<td>1:09</td>
</tr>
<tr>
<td>Mall.avi</td>
<td>0:16</td>
<td>0:08</td>
</tr>
<tr>
<td>Vpt.avi</td>
<td>2:29</td>
<td>0:18</td>
</tr>
</tbody>
</table>
**FUTURE WORK:**

Video synopsis is less applicable for a video with already dense activity where all locations are active all the time. An example is a camera in a busy train station. Future work can involve simultaneous display of detected activities. In some cases, the video so obtained is very condensed with objects and events, making it difficult for a user to search for any particular object.

Currently, proposed method works on videos with static backgrounds. It can be extended to handle videos captured by moving cameras. This is a challenging aspect, due to the dynamic backgrounds, and the different types of camera motion, such as panning, zooming, and jittering, etc. It can also be then extended to use video search tools for presenting relevant events first.