Video analytics for security

Surveillance cameras have emerged as very effective and important aspect of security and monitoring. Unfortunately however, they suffer from two serious challenges. First, their effectiveness in preventing a mishap is limited by the alertness levels of humans who are expected to monitor a grid of live feeds from many cameras 24X7. Since humans are not known for large attention spans, more often than not the mishap misses the eyes of the on-duty guards and the purpose is defeated. The recorded CCTV footages then at best serve to understand what happened, as a post-mortem analysis. We have developed software prototypes which would take-in live video streams coming directly from the cameras, analyse them for events like asset tampering or intrusion or loitering, etc. and generate real-time alerts in form of emails or mobile notifications or visual flash on screen or audio. We are now working on improving the user experience and packaging it as an easy-to-use product.

The second challenge is that even for post-mortem analysis, stored video footages don't lend themselves well to analysis. They consume a lot of space, take up a lot of time and are amongst the hardest to analyse. A 24 hour CCTV footage is typically about 20 GB of data! It is almost a torture asking a human to go through the entire footage to identify events of interest. There is a need for a summarisation tool which can enable watching hours of videos in minutes. This is precisely what we have developed (in collaboration with researchers at NCETIS) and we are currently working on various other advanced aspects of video summarisation. For example, consider locating the instant of time in video when a person wearing dark colored shirt enters a demarcated room, from a CCTV footage of about 24 hours! Or consider a tool which would process hours of video and emit statistics in terms of number of people seen or number of cars seen, etc. In this project, we are working on building tools which will help in such analysis of the stored video content.

Real-time Analytics on Live View

- Motion / Inactivity
- Intrusion
- Loitering
- Asset Tampering
- Face, Age, Gender, Ethnicity
- Apparel
- Object left behind

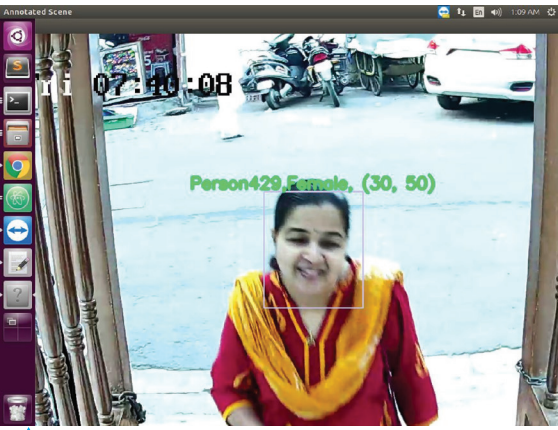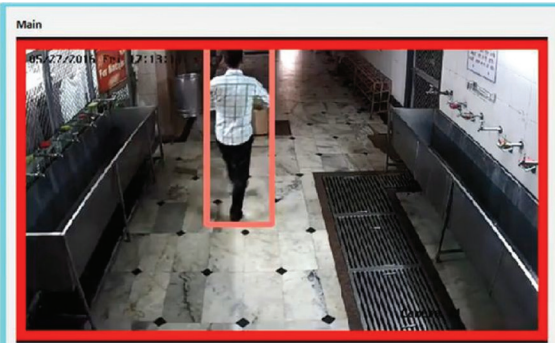Email / Mobile / Audio / Visual Alerts

Smart Analysis of Videos

- Face, text, age, gender, color, objects
- Search
- Summarize
- Statistics

Video Player

Real-time alerts

Face, age, gender recognition

Video summarisation
Video summarisation is an interesting research area in itself and includes identifying different activities seen in the video with respective duration, extracting key events from the video and producing trailer or highlights.

Several algorithms produce video summaries based on motion. This is not enough because there may be several portions of videos with lot of motion and yet not valuable enough to appear in the summary. More advanced algorithms look at minimising the redundancy across frames and choosing as diverse and as representative a subset of frames as possible. These suffer from not being able to adapt to different notion of significance across domains. We are developing a novel video summarisation technique based on learning a submodular mixture which would learn to adapt to different domains and produce summaries very close to those produced by humans as ground truth.

Smart search and statistics on videos

We are using state-of-the art deep models for an improved understanding of videos thereby enabling searching for useful content in videos and also to produce interesting statistics from videos. Given a search query (example, 'person wearing yellow' or 'shark images'), a retrieval problem involves finding objects and images which match the given query. Similarly, one could provide an example image, and retrieve similar frames from the video. Statistics include metrics like number of people seen in the video, number of objects in the video (could be cars or computers or guns or anything else depending on the domain).

Applications of submodular subset selection in computer vision

Submodular functions are a special type of set functions with a unique 'diminishing returns' property. This makes them naturally model the notions of diversity, coverage and representativeness. Some of our research projects have explored the use of submodular subset selection in data summarisation, active learning and video summarisation.

Smart search

Prof. Ganesh Ramakrishnan, Department of Computer Science and Engineering, ganesh@cse.iitb.ac.in