

This commentary is part of a series in TODAY's Science section, in collaboration with the National University of Singapore School of Computing. This series explores a selection of computer science research projects carried out here and how they can improve people's lives.



MOHAN KANKANHALLI

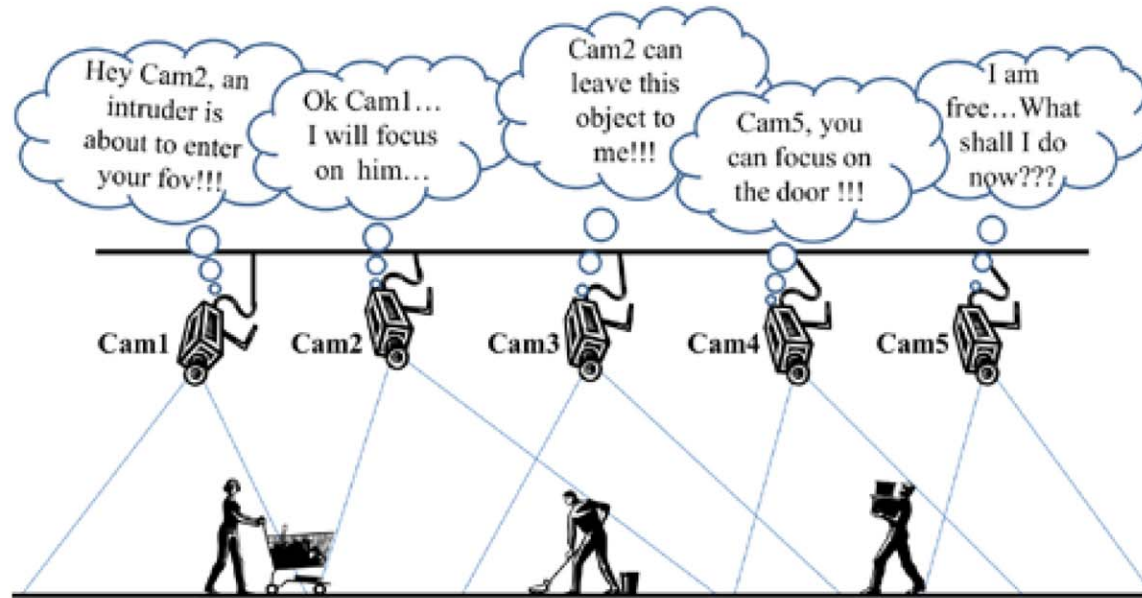
To most people, video surveillance means having a security officer monitor multiple screens of video footage taken by cameras located at different parts of a shopping mall in a bid to deter or detect crime. That is going to change in the future. With cheap cameras and ubiquitous Wi-Fi, it is now feasible to organise networks of cameras to work together to significantly boost security.

Imagine if a camera could detect a suspicious person, and as the person moves around, one camera alerts the next to keep watch on the person. Or, instead of having security personnel monitor hours of surveillance video footage — a tedious and expensive approach — let the computer monitor all the video feeds, and send alerts to security personnel when it detects suspicious events. These scenarios are still some time away, but our research continues to work towards creating such a future.

There are currently several automated video surveillance systems for different applications; for instance, for doorway surveillance (detect-

## NETWORK OF CAMERAS CAN BE USED TO EXTEND SURVEILLANCE

# When CCTV cameras work together as one



ing if anyone is entering or exiting a monitored doorway), people counting at virtual checkpoints, number plate detection and recognition, parked vehicle detection and sterile zone monitoring. These are dedicated systems where a single camera is installed and calibrated at a certain location.

Despite great efforts in research, existing systems still face a huge challenge in extending surveillance across a network of cameras. How can multi-

ple cameras be coordinated to monitor different people and spaces? How can security efficiently search for a person or vehicle from hours and hours of video footage?

As we work to find solutions to these problems at the National University of Singapore (NUS), we have developed two novel technologies: Active camera sensing and video synopsis.

### ACTIVE CAMERA SENSING

Currently, most Pan-Tilt-Zoom (PTZ) cameras are manually controlled by the security officer, who could use the same camera to observe various regions at high image resolution. However, manual control of cameras is practically impossible when the number of cameras increases.

With the help of a mathematically sophisticated algorithm, we can automate PTZ cameras, typically used in areas such as building lobbies or campuses, to observe targets such as people or vehicles. For instance, a group of cameras can be told to observe all the moving people in a lobby and the cameras will make sure that each moving person is always tracked by at least one camera.

In fact, different cameras can even be coordinated to simultaneously do different tasks.

This method, called "active camera sensing", can therefore be used in a number of applications. For example, it can be used to capture detailed face images of individuals, observe a highly crowded area for potential problems, or for taking photos of abandoned ob-

Active sensing by multiple cameras at work. GRAPHIC: MOHAN KANKANHALLI

● Dr Mohan Kankanhalli is the vice-provost of Graduate Education at the National University of Singapore and a professor at its School of Computing.

jects. It can even be used for analytics such as people counting.

With active camera sensing, cameras and social media may one day work together, where the former is programmed to take note if there is a sudden cluster of tweets from one location. The cameras could then turn their attention to the location to see if anything, such as a riot or accident, is in progress.

### VIDEO SYNOPSIS

With more cameras being used, the amount of surveillance footage is also increased. Currently, surveillance video streams from hundreds of cameras are routinely stored. If a terrorist incident, such as the one during the 2013 Boston Marathon, occurs, human operators are deployed to view these archived videos.

However, browsing through multiple videos takes a long time and may be unacceptable when the perpetrator is on the run and can potentially cause great damage. We therefore need a method to watch these videos faster than real-time.

One straightforward method would be to play the footage at a fast-forwarded rate to reduce the viewing time. But this is dangerous since a critical video segment may be inadvertently missed.

To overcome this problem, we have developed a method of video synopsis that creates a short video summary that preserves people and activities. Multiple activities can be shown simultaneously, even when each activity originally occurred at a different time, thus condensing the footage.

With our flexible technique, our system can focus only on certain aspects such as people or movement, and the footage automatically skips the irrelevant parts.

Another option we can offer is to create a single frame of, say, all the people who passed through a location during a specified time period.

With video synopsis, the human operator thus is able to view multiple activities in each camera in parallel at one glance. Doing so also enables all sorts of useful analytics data to be gathered, allowing security personnel to browse and analyse large amounts of videos from multiple cameras very efficiently.

We are currently in talks to see how video synopsis can be deployed in a setting, such as in a mall.

The combination of active camera sensing and video synopsis opens up a vast array of applications in video surveillance. We can expect our cameras in future to work together cooperatively to successfully safeguard our homes, shops, roads, buildings and nation. With its analytics capabilities, it can even help design them better.