K.N. Plataniotis
and C.S. Regazzoni

Visual-Centric Surveillance Networks and Services

Research in surveillance networks and services is an active discipline that is expanding in breadth. Due to the current state of world affairs, commercial products that capitalize on the security potential of digital surveillance systems dominate the engineering services market. The events of recent years have increased the demand for security. It is therefore not surprising that the development of intelligent systems for real-time acquisition, transmission, processing, and understanding of surveillance-related information is on the forefront of research and development efforts. More than ever before, it is important to maintain the safety and security of citizens, buildings, public infrastructures, and other public assets.

The new demands and the increasing maturity of algorithms and techniques have created new services in diverse application areas. The problem of remote surveillance has received growing attention in recent years, especially in the context of public infrastructure monitoring for transport applications (i.e., railway, airports, and motorways), safety or quality control in industrial applications, and improved public security (monitoring of indoor and outdoor environments), just to list a few examples.

The development of a surveillance system requires multidisciplinary expertise, including knowledge of signal and image processing, computer vision, communications and networking, pattern recognition, and sensor development and fusion. Although video surveillance research is an interdisciplinary area drawing upon many traditional disciplines, there is no doubt that signal and image processing methodologies form the core of any surveillance system. Signal processing methods are the enabling technologies that allow end-users to utilize multiple signals, to extract temporally and spatially extended reasoning about events in their surroundings, and to interact with the system and other users.

DEVELOPMENT OF VIDEO SURVEILLANCE

During the maturation process that followed the early years of video surveillance research, the most promising functionalities proved to be digital video processing, recognition, and communication. The same enabling technologies are poised to take advantage of the inexpensive, smart-sensing devices that are now available to practicing engineers. Despite an impressive growth over the last 20 years, the full potential of the digital approach to surveillance networks is yet to be achieved. Whereas current applications view surveillance sensors as standalone intelligent information sources, the modularity of the digital video surveillance framework and the availability of inexpensive devices with embedded processing capabilities will enable the design of distributed architectures in which networked homogeneous and heterogeneous sensors coexist and cooperate, allowing users to perceive and interpret monitored events in real time.

The stimulus for the development of modern surveillance systems is the obvious limitation of the human perception system with regard to surveillance and monitoring. Although humans can focus (for a short period of time) on events they consider important, the amount of information that they process and interpret is usually limited. Since their inception, surveillance systems have been viewed as information processing tools that extend human perception and reasoning. In the surveillance paradigm, extended perception was originally associated with the ability to communicate and process signals in remote control rooms. In this way, the surveillance agent was able to monitor a number of physical locales from a distance. Digital video coding and communications methods were the enabling technologies for this task.

More recently, research efforts in extended perception have shifted towards the development of solutions to digitally enhance the attention-focusing performance of the human operator. Although it is difficult for a human operator to concentrate on event detection for long periods of time, reasonable motion detection methods can be designed for static cameras that utilize background subtraction and change detection methods to assist with this task. When multiple cameras are used, the potential benefit is even greater, provided viewing priorities have been established.

MULTISENSOR SYSTEMS

Most currently available surveillance systems are not only capable of enhancing the performance of the operator by presenting the feeds from monitoring cameras but can also interact at a higher level of reasoning. Surveillance systems available on the market today are capable of automatically detecting the semantic nature of monitored events and providing synthetic coded descriptions of the events along with compressed video sequences of the monitored surroundings. In this way, the system not only augments the real-time event detection performance of the human operator, but it can also retrieve and query events of interest using metadata information associated with archived events. Despite the limitations of such systems in regard to the number and type of detected events as well as in detection performance, the perception extension they provide to humans is a breakthrough in the everyday use of video object segmentation and classification tools designed by the image processing community. Until now, the available solutions have utilized a single remote sensor with long time periods to augment human perception. The use of multiple, possibly heterogeneous, networks of sensors distributed over large areas presents an interesting challenge for researchers specializing in this field. The question is the extent to which one can design multisensor processing tools that allow us to monitor larger areas for long time intervals, utilizing signals that can provide us with information at multiple resolution levels depending on the importance of events as defined by current assignments of mined historical data.

This question is the subject of this special section on video and signal processing for surveillance networks and services, which highlights solutions based on techniques for multisensor data acquisition and processing. Surveillance systems are based on the strong interaction of these techniques, and their development requires multidisciplinary expertise (signal and image processing, computer vision, communications and networking, pattern recognition, and sensor development).

A MULTIDISCIPLINARY APPROACH

To keep pace with the expanding needs of the ever-increasing surveillance markets, surveillance engineers and end-users must take advantage of the many innovative solutions and methods proposed by the research community. This special section aims to provide critical insight into surveillance systems and services by offering a variety of articles, covering various aspects of the surveillance paradigm, written by contributors from academia and industry. While several modalities and sensing devices can be used in surveillance networks, this special section focuses on visual-centric surveillance systems and applications.

Visual data acquisition, processing, communication, and understanding are considered to be indispensable aspects of surveillance. This is due to several factors:

- Temporally organized visual information is the main source of information used by humans regarding their surroundings.
- The cost of visual (mostly video) sensors is considerably lower compared to the cost of other surveillance sensors.
- The large body of knowledge in the areas of visual data processing and communications allows for the development of effective and robust real-time surveillance systems.

However, prior to the development and deployment of a successful visual-based surveillance system, a number of key issues should be addressed. For example, processing and understanding video in the context of a surveillance system is much more demanding than processing visual data in traditional computer vision applications due to the high variability and irregularity of the monitored scenes. The high variability necessitates the use of sophisticated image processing algorithms to process and analyze the visual information collected. It also requires the deployment of robust scene description and pattern recognition methods to analyze the information. The ability to adapt to changing scene conditions and automatically acquire models of the event patterns are emerging issues that will dominate research for the development of the new generation of surveillance systems.

On the other hand, the availability of inexpensive and faster computing hardware, coupled with the efficient and versatile video sensors (i.e., multicameras and active cameras) available today, enables the creation of complex surveillance system architectures that can drive development and introduce new applications. For example, surveillance architectures employing multicamera systems can provide coverage across a wide area, ensuring object visibility over a vast depth range, and can be employed to disambiguate occlusions. This is of paramount importance in such application areas as public monitoring in indoor and outdoor environments (i.e., office towers, civic centers, retail banking offices, shopping malls, sport arenas, transportation hubs, and parking lots).

Finally, the multisensor aspect of a visual-centric surveillance system constitutes an important direction for future development. Multisensor surveillance systems have a competitive advantage in processing the pertinent information as it is captured by different types of sensors (e.g., video cameras, infrared cameras) over the same monitored area. The development of new processing techniques and the deployment of new sensors capable of providing real-time information as it relates to different scene characteristics can help both to enlarge the size of the environment under surveillance and to improve performance.

IN THIS ISSUE

By nature, this special section cannot cover all these important developments and certainly cannot provide a complete overview of the state-of-the-art in this vast area. However, the four articles provide a good starting point for the researchers and practitioners willing to tackle the obstacles and take advantage of the opportunities in the area of surveillance system development.

In the first article titled "High-Resolution, Slow-Motion Sequencing," S. Chaudhouri and D.R. Taur discuss high-resolution methodologies that capitalize on the relative motion information between successive frames in surveillance video to generate still surveillance images of increased spatial resolution. In this way, an area of interest is slowed down and super resolved so that activity that is not clearly visible in the original surveillance video can be identified and later subjected to standard surveillance operations such as tracking, recognition, and identity verification. In the same article, C.P. Mammen offers an industrial perspective on the slow-motion, high-resolution application for both real-time and post-facto surveillance.

The second article, "Active Video-Based Surveillance Systems," written by

G.L. Foresti, C. Micheloni, L. Snidaro, P. Remagnino, and T. Ellis, describes in detail the low-level image and video processing techniques needed for the implementation of an advanced visual-based surveillance system. Key technical problems are identified, and solutions are outlined. Issues pertinent to surveillance systems deploying multicamera clusters that operate in real time are highlighted. The article concludes with the authors' view on future developments in the area.

In the article titled "Smart Video Surveillance," the research team from IBM's T.J. Watson Research Center offers an industrial perspective on the important problem of situation awareness in the context of surveillance and security. A comprehensive, nonintrusive solution is proposed based on the concepts of multiscale, spatial-temporal tracking through the use of real-time video analysis, active cameras, object classification, and movement pattern analysis. Performance evaluation issues are discussed in a systematic manner, implementation challenges are highlighted, and future directions are provided.

Finally, in the article "Imaging for Concealed Weapon Detection," H.-M. Chen, S. Lee, R.M. Rao, M.A. Slamani, and P.K. Varshney present solutions to the problem of detecting concealed weapons. To this end, signal and image processing techniques are utilized to denoise, enhance, fuse, and extract shape information. Fusion of data from heterogeneous sensors is used to provide an extended perception of a monitored event. Particular emphasis is given to the problem of assembling the information gathered by the sensors in a coherent reference system. The article provides a tutorial overview on the taxonomy of sensors and signal processing support required for this task.

The special section covers only a limited number of works and initiatives in the area of surveillance systems and applications. Developments are introduced by researchers and practitioners at a very fast pace, and it is impossible to survey and present all of the information in a single issue. Nevertheless, the guest editors hope that the included articles will be helpful to those who are interested in contributing to this ever-expanding area in which academic research and industrial efforts come together to better serve the end-user.

ACKNOWLEDGMENTS

We would like to thank all those who have helped to make this special issue possible, especially the authors and the reviewers of the articles, as well as Marco Guainazzo, who provided coordination and logistical support. We'd also like to gratefully acknowledge Editor-in-Chief Ray Liu and Area Editor Petar Djuric for offering us the opportunity to edit this special issue. [SP]