Introduction to the Issue on Distributed Processing in Vision Networks

DISTRIBUTED vision networks is a multidisciplinary topic that interfaces the fields of computer vision, signal processing, pervasive computing, embedded programming, wireless sensor networks, and ambient intelligence. It creates opportunities for design paradigm shifts in these fields through emphasizing distributed vision processing and collaborative fusion of information over the network. In this premise, novel smart environment applications can be envisioned that are scalable, real-time, adaptive, interpretive, context-aware, and user-centric in nature.

Technological advances in the design of image sensors and embedded processors have facilitated the development of efficient embedded vision-based applications. Design of scalable, network-based applications based on high-bandwidth data such as video requires a change of paradigm in the processing methodologies. Instead of streaming raw images to a powerful, central processing unit, each network node can employ local processing to translate the observed data into features and attributes, which are shared with other network nodes to produce a collaborative deduction about the event of interest. The major advantages of distributed processing are hence improved scalability and reliability of the network and better bandwidth utilization.

The multidisciplinary research opportunities arising from the domain of networked cameras can influence the direction of the various research disciplines involved. For example, in signal processing, information sharing among the nodes of the network provides research opportunities in the design of data exchange and fusion methods and distributed algorithms that utilize spatial and temporal redundancies in the data. Many applications of distributed camera systems are based on detection and tracking of humans, objects, or their various attributes. These applications are often developed based on signal processing tools from estimation theory, filtering, optimization, or statistical reasoning areas, which can find new research directions under a distributed processing assumption. As an example, distributed estimation, belief propagation, and decision making techniques need to be examined taking into account the processing capabilities of the nodes as well as the network bandwidth limits and application latency requirements. In addition, spatio–temporal data fusion algorithms employing information obtained by the network across the dimensions of space, time, and feature levels need to be developed and analyzed.

The goal of this special issue is to provide insight into the design methodologies for vision networks from a distributed signal processing perspective. Papers presented in this issue offer perspectives onto the different aspects of deployment, operation, and application development for camera networks.

Digital Object Identifier 10.1109/JSTSP.2008.2001726

Network deployment and operation: Topics in network deployment, operation, and clustering of camera networks are covered in four papers:

Information about camera locations and the field of view of neighboring cameras often allows hand-over of visual object tracking by information passing between dynamically forming camera clusters. In "Distributed Object Tracking using a Cluster-Based Kalman Filter in Wireless Camera Networks," by Medeiros, Park, and Kak, cooperation among the cameras in the form of a decentralized Kalman filter is examined when the cameras in a neighborhood form a cluster in a dynamic fashion which moves along the path of the tracked object and carries the object position estimates. In "Optimal Camera Network Configurations for Visual Tagging," Zhao, Cheung, and Nguyen describe a model for placing cameras on the basis of the object's visibility in the field of view of different cameras. The paper focuses on triangulating visual objects for identifying different individuals, with an emphasis on privacy protection in video surveillance networks. Self organization is an enabling attribute for deployment of scalable networks. In "Spatial Partitioning in Self-organizing Smart Camera Systems," Hoffmann, Wittke, Hahner, and Muller-Schloer propose a decentralized, self-organizing system architecture for wireless networks of smart cameras with PTZ capabilities. The paper "Lightweight People Counting and Localizing for Easily-Deployable Indoors WSNs" by Teixeira and Savvides describes person counting and tracking utilizing motion histogram generation from address-event image sensors. These imagers offer processing advantages through efficient computing of low-level features such as motion histograms.

Sensor fusion and collaborative processing: Four papers describe different sensor fusion, data exchange, and collaborative processing techniques. Systems based on different multiresolution, multispectral, multimodal fusion, as well as signal-level and decision-level fusion are presented and their applications are examined:

In "Target Detection and Tracking with Heterogeneous Sensors," Zhou, Taj, and Cavallaro present a multimodal detection and tracking algorithm for sensors composed of a camera mounted between two microphones. Target localization is performed based on color-based change detection in the video modality, and based on time-difference-of-arrival estimation between the two microphones in the audio modality. A particle filter-based tracking algorithm integrates measurements from the heterogeneous sensors, and produces a meta data report. The paper "A Visible/Infrared Fusion Algorithm for Distributed Smart Cameras" by Chen, Lin, and Wolf describes an algorithm for visible/infrared fusion for video surveillance. Their technique combines signal-level and decision-level multispectral information fusion and uses a hierarchical probabilistic model to fuse and update the image information from different spectra and processing levels. The paper discusses how this kind of fusion addresses low-contrast imaging conditions, which often cause a challenge in real-world surveillance applications. High-resolution image sensing may be necessary in certain applications. However, the amount of acquired data introduces challenges to the operation of camera networks by demanding excessive processing and communication loads. The paper "Energy-Aware High Resolution Image Acquisition via Heterogeneous Image Sensors," by Kim, Rahimi, Lee, Estrin, and Villasenor, describes a heterogeneous camera network that uses simple stereo image sensors to compute pan/tilt/zoom information to enable efficient operation of a separate set of high-capability cameras. In "Distributed Multilevel Data Fusion for Networked Embedded Systems," Klausner, Tengg, and Rinner consider multisensor fusion. Their model considers fusion at multiple levels of abstraction as well as data flow and resource constraints in the sensor network. This leads to a generic framework for data fusion that can support different types of sensing and processing modules. The paper examines application to vehicle classification using video and audio sensing to demonstrate the benefits of multilevel information fusion.

Application development: Four papers describe algorithms addressing different applications based on camera networks:

In the paper "Exploitation of Inter-Frame Redundancy for Real-Time Volumetric Reconstruction of Arbitrary Shapes," Ruiz and Macq present a system for real-time 3-D shape reconstruction based on voxel (or volumetric) representation. Their method exploits inter-frame redundancy for efficient reconstruction. The paper proposes an approach for distributing the computation between the observing cameras and a master reconstructing node to achieve a scalable and real-time operation. The paper "Detecting Dominant Motions in Crowds" by Cheriyadat and Radke describes an algorithm to automatically identify dominant motions in crowded scenes. Tracking individual objects in such scenes is difficult due to inter- and intra-object occlusions. Instead of tracking individual objects, the paper proposes to extract low-level feature points, which are tracked by optical flow. The feature-point trajectories are then clustered into dominant motions. The paper "Robust Tracking in A Camera Network: A Multi-Objective Optimization Framework" by Song and Roy-Chowdhury addresses the problem of tracking multiple targets in a sparse network of non-overlapping cameras. This is performed through the use of both short-term and long-term characteristics. In "Active Multicamera Networks: From Rendering to Surveillance," Stancil, Zhang, and Chen introduce the challenges in the design and operation of multicamera networks by focusing on two classes of applications, namely collaborative scene reconstruction and rendering, and distributed surveillance. Surveys of existing methods as well as a comparison of the requirements for these applications are provided.

The special issue aims to offer a view to the potentials of the field of vision networks. Given the multidisciplinary nature of this field and its relation with several other research domains, to fully cover the research potentials in this area there is obviously a multitude of aspects that need to be further examined. Incorporation of concepts from the different fields allows for the development of efficient methods of processing information from pixel-level data to high-level decision and interpretation. In addition, this also enables extending the application scope of camera networks to novel domains of smart environments that fuse vision-based observations with other modalities as well as contextual information and behavior models to serve the users according to their states and preferences. This provides a lively field for research, pushing the technology and relevant conceptual developments in smart homes, entertainment, gaming, robotics, smart buildings, assisted living, smart classrooms, and many other application domains.

HAMID AGHAJAN, *Lead Guest Editor*
Stanford University
Department of Electrical Engineering
Stanford, CA 94305 USA
hamid@wsnl.stanford.edu

RICHARD KLEIHORST, *Guest Editor*
NXP Semiconductor Research
Eindhoven, The Netherlands
richard.kleihorst@nxp.com

BERNHARD RINNER, *Guest Editor*
Klagenfurt University
Klagenfurt, Austria
Bernhard.Rinner@uni-klu.ac.at

WAYNE WOLF, *Guest Editor*
Georgia Institute of Technology
Atlanta, GA 30332 USA
wolf@ece.gatech.edu

Hamid Aghajan (M'06) received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1991 and 1995, respectively.

From 1994 to 2002 he held positions in the silicon valley industry and co-founded a startup company in telecommunications. In 2003 he established the Wireless Sensor Networks Lab at Stanford University, where he is currently a faculty member in the Electrical Engineering Department. Recent work in his research group consists of vision-based algorithms for assisted living, occupancy sensing for ambience control in smart buildings, human gesture analysis for gaming, virtual reality, and smart displays, and interactivity-based environment discovery—all based on distributed vision-based reasoning in multicamera networks.

Dr. Aghajan is co-editor-in-chief of the *Journal of Ambient Intelligence and Smart Environments*.

Richard Kleihorst received the M.Sc. and Ph.D. degrees in electrical engineering from Delft University of Technology, The Netherlands, in 1989 and 1994, respectively.

In 1994, he joined the VLSI design group of Philips Research Laboratories, Eindhoven, The Netherlands, where he worked on single chip MPEG2 encoding, embedded compression techniques, parallel image processing and fault-tolerance computation techniques. Since 2006, he has been with NXP Semiconductor Research, where he is heading the Xetal team active on parallel image processing architectures and applications. he has over 90 international patent applications including 25 in the US. He is (co-)author of over 100 scientific publications in the field of video processing and architectures.

Bernhard Rinner (M'95–SM'04) received the M.Sc. and Ph.D. degrees in telematics from Graz University of Technology, Austria, in 1993 and 1996, respectively.

He is Full Professor and Chair for Pervasive Computing at Klagenfurt University, Austria, where he is currently serving as vice dean of the Faculty of Technical Sciences. He held research positions at Graz University of Technology and at the Department of Computer Sciences, University of Texas at Austin. His research interests include parallel and distributed processing, embedded systems as well as mobile and pervasive computing. He has authored and co-authored about 100 papers for journals, conferences and workshops, lead several research projects and served as reviewer, program committee member, program chair and editor-in-chief.

Wayne Wolf (M'83-F'98) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1980, 1981, and 1984, respectively.

He is the Farmer Distinguished Chair and Georgia Research Alliance Eminent Scholar at the Georgia Institute of Technology, Atlanta. He was with AT&T Bell Laboratories from 1984 to 1989. He was on the faculty of Princeton University from 1989 to 2007. His research interests include embedded computing, embedded video and computer vision, and VLSI systems.

Dr. Wolf has received the ASEE Terman Award and IEEE Circuits and Systems Society Education Award. He is a Fellow of the ACM.