Pervasive Video and Audio

Fahim Kawsar, *Nokia Bell Labs Cambridge, Cambridge CB3 0, U.K. and also TU Delft, 2628 CD Delft, the Netherlands*

Romit Roy Choudhury, *University of Illinois at Urbana Champaign, Champaign, IL, 61820, USA*

Ganesh Ananthanarayanan, *Microsoft Azure for Operators*

For a long time, cameras and microphones have been studied to give machines the ability to see and hear the world around them. The remarkable progress in artificial intelligence (AI) in recent years has pushed this perception ability to a new high. AI-enabled cameras and microphones are now able to replicate human-like reasoning and cognition, be it detecting cancers better than a human doctor, identifying people's sexual orientation, or understanding human speech. Naturally, these interminable possibilities across a wide range of verticals, ranging from public safety, infrastructure efficiency, entertainment, education, and health, have led to the deployment of billions of cameras and microphones around us—in our homes, workplaces, and cities.

As this pervasiveness keeps disrupting a multitude of business dynamics, it also uncovers myriad challenges—some are technical, some are political, and some are ethical. For instance, how should these systems be architected? What are the new applications enabled? How can we design a low-latency network for these systems? What are the required training and signal-processing algorithms? What processor architecture and OS support are required for a QoS guarantee? What are the challenges in deploying this technology? And, how can they be secured against cyber-attacks?

We reflect on some of these challenges and catalyze new research in pervasive video and audio in this Special Issue.

Our first article, "A Case for Camera-as-a-Service," puts forward a compelling vision toward "software-defined camera" equipped with adequate storage and compute capabilities to perform machine-learning operations natively. Drawing on the technical advancement in edge-scale ML accelerators and lightweight vision models, Xu *et al.* discuss the technical challenges and early results in transforming dumb cameras into autonomous data processors and uncovering a new paradigm of "Camera-as-a-Service" to redefine today's video analytics services.

In our second article, "Adaptive Resolution-Based Tradeoffs for Energy-Efficient Visual Computing Systems," LiKamWa *et al.* reflect on the optimization techniques for device driver and media frameworks to radically improve the energy efficiency of vision-based applications. They discuss sensor management techniques for energy-reactive image capture and media frameworks for dynamic resolution scaling toward designing highly adaptive visual computing systems.

Our third article, "Low Quality for High Quality: Exploiting 2K Frames for Supporting Efficient 4K-Quality Pervasive Video Streaming Applications," addresses the challenge of efficiency for cloud-based video streaming applications. Choi and Ko present a new codec scheme that exploits lower resolution frames to reconstruct higher resolution frames while reducing overhead and computational latency for a seamless video streaming experience.

In our final article, "Cognitive Audio Interfaces: Mediating Sonic Information with an Understanding of How We Hear," we shift our focus on audio, and in particular, how we perceive and interact with the sounds around us. Borrowing principles from causal reasoning and audiority cognition literature, Ananthabhotla *et al.* present an "audio-first" approach to cognitive interface design grounded on multiple case studies that deconstruct sensor data into human-meaningful sound objects.

These four articles shed some light in that direction and coalesce a set agenda for pervasive video and audio research. We hope that future research on this space will provide solutions for many, if not all, of the unresolved concerns.

1536-1268 © 2021 IEEE
Digital Object Identifier 10.1109/MPRV.2021.3076500
Date of current version 25 May 2021.

FAHIM KAWSAR currently leads the pervasive systems research at Nokia Bell Labs Cambridge and holds a Design United Professorship at TU Delft Netherlands. Contact him at fahim.kawsar@nokia-bell-labs.com.

ROMIT ROY CHOUDHURY is currently a professor and Jerry Sanders scholar with the Department of Department of Computer Science and Department of Computer Science, University of Illinois at Urbana Champaign. Contact him at croy@illinois.edu.

GANESH ANANTHANARAYANAN is currently a principal researcher at Microsoft Azure for Operators. Contact him at ga@microsoft.com.