JONG HWAN KO

jhko@skku.edu ⋄ (+82)31-290-7695
Assistant professor, School of Electronic and Electrical Engineering, SungKyunKwan university
2066, Seobu-ro, Jangan-gu, Suwon-si, Gyeonggi-do, Korea

RESEARCH INTERESTS

My research aims to enable intelligent multimedia processing capabilities at the Internet of Things (IoT) edge devices that have limited resources and dynamic conditions. To achieve this goal, I explore design optimization of multimedia processing at the intersection of model-driven (signal processing) and data-driven (deep learning) approaches, incorporating cross-layer interactions between algorithm, architecture, and digital circuit design.

EDUCATION

Ph.D. Electrical and Computer Engineering Aug. 2013 - Aug. 2018
Georgia Institute of Technology, Atlanta, GA
- Thesis topic: Energy-efficient image processing for intelligent sensor systems
- Thesis advisor: Prof. Saibal Mukhopadhyay

M.S. Electrical Engineering and Computer Science Mar. 2004 - Feb. 2006
Seoul National University, Seoul, Korea
- Thesis topic: Materialization models in flash memory storage systems
- Thesis advisor: Prof. Sang Lyul Min

B.S. Mechanical and Aerospace Engineering Mar. 2000 - Feb. 2004
Double Major in **Computer Science and Engineering**
Seoul National University, Seoul, Korea

RESEARCH EXPERIENCE

Assistant Professor Mar. 2019 - Present
SungKyunKwan University (SKKU), Suwon, Korea
- Machine-learning based image/audio computing for IoT systems

Research Intern May 2017 - Aug. 2017
Microsoft Research, Redmond, WA
- Efficient deep neural networks for speech processing
  - Studied the effect of precision scaling on the performance and processing time of deep neural networks
  - Designed a DNN-based real-time voice activity detection and speech enhancement engine for MS Cortana

Graduate Research Assistant Aug. 2013 - Aug. 2018
Georgia Institute of Technology, Atlanta, GA
- OROEB: On-line Real-time Optimal Energy Balancing for Self-Powered Environment Adaptive Sensor
  - Managed project objectives and led team members as a project leader
  - Designed scalable and adaptive image processing techniques based on region-of-interest-based coding
  - Implemented and tested an ASIC of an autonomous self-powered wireless sensor node (IBM 130nm)
- Energy-efficient and robust deep learning for the IoT edge devices
  - Designed a weight compression method based on adaptive image encoding for lower memory demand
  - Proposed an entire frequency-domain convolution approach for energy-efficient CNN training
  - Designed deep learning algorithms robust to adversarial attacks and image perturbations

Senior Research Engineer Jan. 2006 - Feb. 2019
Agency for Defense Development, Daejeon, Korea

- Tactical Information and Communication Network (TICN)
  - Conducted quality-of-service analysis and optimization of multimedia data in distributed networks
  - Proposed enhanced signaling/routing protocols for tactical voice over IP systems
- A Tactical Embedded Platform for Communication and Control of Unmanned Vehicles
  - Performed initial idea development, proposal writing, and project management as a project lead
  - Designed an energy-efficient RF module and signaling protocol for real-time communication of UAVs

Visiting Researcher July 2004 - Aug. 2004
Saarland University, Germany

- Worst Case Execution Time (WCET) Analysis for Real-Time Systems
  - Participated in a project for WCET analysis tool development

Graduate Research Assistant Mar. 2004 - Feb. 2006
Seoul National University, Seoul, Korea

- Low-Power/High-Performance Flash Memory Controller
  - Designed a low-power filesystem (flash memory translation layer) and firmware for flash memory controllers
  - Studied the effect of different write-back schemes on the performance of flash memory systems
- Real-time Scheduler and Execution Time Analysis Tool
  - Developed a software tool for task scheduling of an engine control unit

HONORS AND AWARDS

1st Place Winner of the Cybersecurity Demo Day Apr. 2018
Institute for Information Security and Privacy, Georgia Institute of Technology

Travel Grant for DAC PhD Forum Finalists Aug. 2017
Association for Computing Machinery (ACM)

Best Paper Award Aug. 2016
International Symposium on Low Power Electronics and Design (ISLPED 2016)

Best Presentation Award July 2016
As a mentor for Summer Undergraduate Research in Engineering/Science Program, Georgia Tech

Best Poster Award July 2014
As a mentor for Summer Undergraduate Research in Engineering/Science Program, Georgia Tech

Ph.D. Study Fellowship 2013 - 2017
Agency for Defense Development

Outstanding Achievement Award Aug. 2012
Agency for Defense Development

Grand Prize in the Innovative Research Project Proposal Contest Dec. 2010
Agency for Defense Development

Outstanding Achievement Award Aug. 2012
Agency for Defense Development

Graduate Study Scholarship 2004 - 2005
Agency for Defense Development

Best Presentation Award for the Bachelor's Thesis Jan. 2004
Department of Mechanical and Aerospace Engineering, Seoul National University

Academic Excellence Scholarships 2000 - 2003
Department of Mechanical and Aerospace Engineering, Seoul National University

PUBLICATIONS

International Journal

[J7] Arvind Singh, Nikhil Chawla, **Jong Hwan Ko**, Monodeep Kar, and Saibal Mukhopadhyay, "Energy Efficient and Side-Channel Secure Cryptographic Hardware for IoT-edge Nodes," *IEEE Internet-of-Things Journal (***IoT-J***), Vol. 6, no. 1, pp. 421-434, Feb. 2019 (impact factor:7.6)*

[J6] **Jong Hwan Ko**, Duckhwan Kim, Taesik Na, and Saibal Mukhopadhyay, "Design and Analysis of a Neural Network Inference Engine based on Adaptive Weight Compression," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (***TCAD***), Vol. 38, no. 1, pp. 109-121, Jan. 2019 (impact factor: 1.94).*

[J5] Mohammad F. Amir, **Jong Hwan Ko**, Taesik Na, Duckhwan Kim, and Saibal Mukhopadhyay, "3-D Stacked Image Sensor With Deep Neural Network Computation," *IEEE Sensors Journal (***Sensors-J***), Vol. 18, no. 10, pp. 4187-4199, May. 2018 (impact factor:2.51).*

[J4] **Jong Hwan Ko**, Taesik Na, and Saibal Mukhopadhyay, "An Energy-Quality Scalable Wireless Image Sensor Node for Object-Based Video Surveillance," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems (***JETCAS***), Vol. 8, no. 3, pp. 591-602, Sept. 2018 (impact factor:2.54).*

[J3] Taesik Na, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Clock Data Compensation Aware Digital Circuits Design for Voltage Margin Reduction," *IEEE Transactions on Circuits and Systems I (***TCAS-I***), Vol. 64, no. 9, pp. 2295-2307, June 2017 (impact factor:2.41).*

[J2] **Jong Hwan Ko**, Mohammad F. Amir, Khondker Zakir Ahmed, Taesik Na, and Saibal Mukhopadhyay, "A Single-Chip Image Sensor Node with Energy Harvesting from a CMOS Pixel Array," *IEEE Transactions on Circuits and Systems I (***TCAS-I***), Vol. 64, no. 9, pp. 2295-2307, June 2017 (impact factor:2.41).*

[J1] **Jong Hwan Ko**, Burhan A. Mudassar, and Saibal Mukhopadhyay, "An Energy-Efficient Wireless Video Sensor Node for Moving Object Surveillance," *IEEE Transactions on Multi-Scale Computing Systems (***TMSCS***), vol. 1, no. 1, pp. 7–18, Sep. 2015.*

International Conference

[C23] Taesik Na, Burhan A. Mudassar, Priyabrata Saha, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Mixture of Pre-processing Experts Model for Noise Robust Deep Learning on Resource Constrained Platforms," *International Joint Conference on Neural Network (***IJCNN 2019***), July 2019.*

[C22] **Jong Hwan Ko**, Taesik Na, Mohammad Faisal Amir, and Saibal Mukhopadhyay, "Edge-Host Partitioning of Deep Neural Networks with Feature Space Encoding for Resource-Constrained Internet-of-Things Platforms," *15th IEEE International Conference on Advanced Video and Signal-based Surveillance (***AVSS 2018***), Nov. 2018.*

[C21] Mohammed F. Amir, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "An Image Sensor SOC with Energy Harvesting Mixed-Vth Pixel Generating 5.8uW/mm2 Power Density and 0.77 Frames/second Self-Powered Frame Rate," *Design Automation Conference (***S3S 2018***), Oct. 2018.*

[C20] Burhan Mudassar, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Edge-Cloud Collaborative Processing for Intelligent Internet of Things: A Case Study on Smart Surveillance," *Design Automation Conference (***DAC 2018***), June 2018.*

[C19] Burhan Mudassar, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "An Unsupervised Anomalous Event Detection Framework with Class Aware Source Separation," *International Conference on Acoustic, Speech, and Signal Processing (***ICASSP 2018***), Apr. 2018.*

[C18] Taesik Na, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Noise-Robust and Resolution-Invariant Image Classification with Pixel-Level Regularization," *International Conference on Acoustic, Speech, and Signal Processing (***ICASSP 2018***), Apr. 2018.*

[C17] **Jong Hwan Ko**, Josh Fromm, Shuayb Zarar, Matthai Philipose, and Ivan Tashev, "Limiting Numerical Precision of Neural Networks to Achieve Real-Time Voice Activity Detection," *International Conference on Acoustic, Speech, and Signal Processing (***ICASSP 2018***), Apr. 2018.*

[C16] Taesik Na, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Cascade Adversarial Machine Learning Regularized with a Unified Embedding," *Design, Automation, and Test in Europe (***ICLR 2018***), Apr. 2018.*

[C15] Saibal Mukhopadhyay, Marilyn Wolf, Mohammed F. Amir, Evan Gebahrdt, **Jong Hwan Ko**, Jae Ha Kung, and Burhan A. Mudassar, "The CAMEL Approach to Stacked Sensor Smart Cameras," *Design, Automation, and Test in Europe (***DATE 2018***), Mar. 2018.*

[C14] Taesik Na, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Adversarial Machine Learning Regularized with a Unified Embedding," *Annual Conference on Neural Information Processing Systems, Machine Learning for Audio Signal Processing Workshop (***NIPS Workshop 2017***), Dec. 2017.*

[C13] **Jong Hwan Ko**, Josh Fromm, Shuayb Zarar, Matthai Philipose, Ivan Tashev, "Bit Precision Control of Deep Neural Networks for Efficient Speech Processing," *Annual Conference on Neural Information Processing Systems, Machine Learning for Audio Signal Processing Workshop (***NIPS Workshop 2017***), Dec. 2017.*

[C12] (Invited) **Jong Hwan Ko**, Yun Long, Mohammad Faisal Amir, Duckhwan Kim, Taesik Na, Amit Trivedi, and Saibal Mukhopadhyay, "Energy-Efficient Neural Image Processing for Internet-of-Things Edge Devices," *IEEE International Midwest Symposium on Circuits and Systems (***MWSCAS 2017***), Aug. 2017.*

[C11] Taesik Na, **Jong Hwan Ko**, Jaeha Kung, and Saibal Mukhopadhyay, "On-Chip Training of Recurrent Neural Networks with Limited Numerical Precision," *International Joint Conference on Neural Networks (***IJCNN 2017***), June. 2017.*

[C10] **Jong Hwan Ko**, Burhan Mudassar, Taesik Na, and Saibal Mukhopadhyay, "Design of an Energy-Efficient Accelerator for Training of Convolutional Neural Networks using Frequency-Domain Computation," *Design Automation Conference (***DAC 2017***), June. 2017 (Oral, acceptance rate: 23.8%).*

[C9] **Jong Hwan Ko**, Duckhwan Kim, Taesik Na, Jaeha Kung, and Saibal Mukhopadhyay, "Adaptive Weight Compression for Memory-Efficient Neural Networks," *Design, Automation, and Test in Europe (***DATE 2017***), Mar. 2017 (Oral, acceptance rate: 24%).*

[C8] Taesik Na, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Clock Data Compensation Aware Clock Tree Synthesis in Digital Circuits with Adaptive Clock Generation," *Design, Automation, and Test in Europe (***DATE 2017***), Mar. 2017 (Oral, acceptance rate: 24%).*

[C7] **Jong Hwan Ko**, Mohammad Faisal Amir, Taesik Na, and Saibal Mukhopadhyay, "A Low-Power Wireless Image Sensor Node with Noise-Robust Moving Object Detection and a Region-of-Interest Based Rate Controller," *42th Annual Government Microcircuit Applications and Critical Technology Conference (***GOMACTech 2017***), Mar. 2017.*

[C6] Khondker Zakir Ahmed, Mohammad Faisal Amir, **Jong Hwan Ko** and Saibal Mukhopadhyay, "Reconfigurable 96x128 Active Pixel Sensor with 2.1W/mm2 Power Generation and Regulated Multi-Domain Power Delivery for Self-Powered Imaging," *42th European Solid-State Circuit Conference (***ESSCIRC 2016***), Sep. 2016.*

[C5] **Jong Hwan Ko**, Taesik Na,and Saibal Mukhopadhyay, "An Energy-Efficient Wireless Video Sensor Node with a Region-of-Interest Based Multi-Parameter Rate Controller for Moving Object Surveil-

lance," *IEEE Advanced Video and Signal-based Surveillance (***AVSS 2016***), Aug. 2016. (acceptance rate: 28.4%)*

[C4] **Jong Hwan Ko** and Saibal Mukhopadhyay, "An Energy-Aware Approach to Noise-Robust Moving Object Detection for Low-Power Wireless Image Sensor Platforms," *International Symposium on Low Power Electronics and Design (***ISLPED 2016***), Aug. 2016 (acceptance rate: 23.2%)* – **Best paper award**

[C3] **Jong Hwan Ko**, Khondker Zakir Ahmed, Mohammad Faisal Amir, and Saibal Mukhopadhyay, "A SelfPowered Wireless Video Sensor Node for Moving Object Surveillance," *41th Annual Government Microcircuit Applications and Critical Technology Conference (***GOMACTech 2016***), Mar. 2016*

[C2] Arvind Singh, Monodeep Kar, **Jong Hwan Ko**, and Saibal Mukhopadhyay, "Exploring power attack protection of resource constrained encryption engines using integrated low-drop-out regulators," *International Symposium on Low Power Electronics and Design (***ISLPED 2015***), Aug. 2015.*

[C1] **Jong Hwan Ko**, Burhan A. Mudassar, and Saibal Mukhopadhyay, "Adaptive Wireless Video Sensor Node Using Content-Aware Pre-Processing for Moving Target Identification," *Annual Government Microcircuit Applications and Critical Technology Conference (***GOMACTech 2015***), Mar. 2015*

PATENTS

[P10] Ivan Tashev, Shuayb Zarar, Matthai Philipose, and **Jong Hwan Ko**, US Patent Application, An approach to efficiently process audio signals by limited numerical precision of neural networks, Mar. 2018.

[P9] **Jong Hwan Ko**, Hyun-sung Lee, Jae Hyun Ham, Jeung-Won Choi, Jeong-ki Pack, and Mingyeong Seo, Korean Patent No.10-1488793, Simulating body for evaluating communication channels near a human body and Method for manufacturing the same, Jan. 2015.

[P8] **Jong Hwan Ko**, Juman Park, Jeung-Won Choi, and Jeong-ki Pack, Korean Patent No.10-1486847, Numerical phantom model manufacturing method for analyzing on human effects of Radio Frequency, Jan. 2015.

[P7] Jae Hyun Ham, **Jong Hwan Ko**, Hyun Sung Lee, and Jeung-Won Choi, Korean Patent No. 10-1471440, MAC frame structure for dynamic Ad Hoc network, and operating method, Dec. 2014.

[P6] Hyun-sung Lee, **Jong Hwan Ko**, Juman Park, and Jeung-Won Choi, Korean Patent No.10-1464667, Real-time Ad Hoc wireless transmission system based on Chirp Spread Spectrum ranging and Method thereof, Nov. 2014.

[P5] Hyungjoo Lee, Mijeong Hoh, **Jong Hwan Ko**, Jeung-Won Choi, Jaehyun Kim, Sung Hyung Lee, Sang Hun Lee, and Kihun Kim, Korean Patent No. 10-1447666, Adaptive Buffer Management Method based on Priority of Voice over IP Traffic, Sep. 2012.

[P4] **Jong Hwan Ko** and Jeung-Won Choi, Korean Patent No.10-1357688, Method and System for SIP Call Process to Ensure QoS using VoIP Exchange, Apr. 2012.

[P3] Bong-soo Roh and **Jong Hwan Ko**, Korean Patent No.10-1084629, A buffer Space Securing Method and Apparatus for Real Time Data Transmission According to Data Increase and Decrease Rate, Nov. 2011.

[P2] **Jong Hwan Ko**, Mijeong Hoh, and Jeung-Won Choi, Korean Patent No.10-1199448, Apparatus, Method and Recording Device for Prediction of VoIP based Speech Transmission Quality using Extended E-model, Oct. 2011.

[P1] **Jong Hwan Ko**, Bong-soo Roh, and Jung-hoon Kim, Korean Patent No.10-1046597, Method for Processing Data in Network, Device for Transceiving Data and Computer Recordable Medium, June 2011.

TEACHING EXPERIENCE

Computer Programming for Engineers Spring 2019
SungKyunKwan Univ., Suwon, Korea

Advanced VLSI (Teaching Assistant) Fall 2015
Georgia Institute of Technology, Atlanta, GA

Computer Architecture (Teaching Assistant) Spring 2004
Seoul National University, Seoul, Korea

INDIVIDUAL STUDENT GUIDANCE

Ph.D. Thesis Advisor
SungKyunKwan Univ., Suwon, Korea

- Hyochan Kim Spring 2019 -

Graduate Advisor for MS Thesis
Georgia Institute of Technology, Atlanta, GA

- Burhan A. Mudassar Spring 2014 - Spring 2015
  Design and implementation of a content aware image processing module on FPGA

Mentor for SURE (Summer Undergrad Research in Engineering/Science) Program
Georgia Institute of Technology, Atlanta, GA

- Jennifer Camacho Summer 2016
  Sign language translator using an artificial neural network - **Best presentation award**
- Maurisa Orona Summer 2015
  Visual image display using an FPGA compatible camera and MATLAB
- Luis D. Nieves Summer 2014
  Power consumption and image quality tradeoff analysis - **Best poster award**

Graduate Advisor for Special Problem Courses
Georgia Institute of Technology, Atlanta, GA

- Abhijit Gadad Spring 2017
  FPGA implementation of a neural network weights decoder
- Abhinav Roy Burman Spring 2017
  FPGA implementation of a neural network inference engine
- Pallavi Kinnera Spring 2016
  Variable quality factor JPEG design
- Ramnath Ramakrishnan Spring 2016
  Hardware design of a variable window median filter for image blurring

TECHNICAL REVIEWER ACTIVITIES

- *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*
- *IEEE Internet of Things Journal (IoT-J)*
- *IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS)*
- *IEEE Transactions on Multi-Scale Computing Systems (TMSCS)*
- *IEEE Signal Processing Letters (SPL)*
- *IEEE Access*
- *Design Automation Conference (DAC)*
- *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*
- *International Symposium on Circuits and Systems (ISCAS)*
- *Design, Automation, and Test in Europe (DATE)*
- *ACM/IEEE Supercomputing Conference (SC)*

TALKS

- **"Resource-Efficient and Robust Multimedia Processing for IoT Edge Devices"**
  KyungHee University, Yongin, Korea, Mar. 2019
- **"Smart Image Sensor Systems: Incorporating Deep Neural Networks into Image Sensors"**
  24th Asia and South Pacific Design Automation Conference (ASP-DAC) Tutorial, Tokyo, Japan, Jan. 2019
- **"Efficient Encoding and Transmission of Video Data over Ultra-low Bandwidth Tactical Wireless Link"**
  DTC Technology, Troll Systems Headquarters, USA, Nov. 2018
- **"Limiting Numerical Precision of Neural Networks to Achieve Real-Time Voice Activity Detection"**
  International Conference on Acoustic, Speech, and Signal Processing (ICASSP), Canada, Apr. 2018
- **"Resource-Aware and Robust Multimedia Processing for IoT Edge Devices"**
  Rice University, Houston, TX, USA, Mar. 2017
- **"Efficient Neural Network Design for Real-Time Speech Enhancement"**
  Microsoft Research, Redmond, WA, USA, Aug. 2017
- **"Energy-Efficient Neural Image Processing for Internet-of-Things Edge Devices"**
  IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), Aug. 2017
- **"Design of an Energy-Efficient Accelerator for Training of Convolutional Neural Networks using Frequency-Domain Computation"**
  Design Automation Conference (DAC), June. 2017
- **"Adaptive Weight Compression for Memory-Efficient Neural Networks"**
  Design, Automation, and Test in Europe (DATE), Mar. 2017
- **"Clock Data Compensation Aware Clock Tree Synthesis in Digital Circuits with Adaptive Clock Generation"**
  Design, Automation, and Test in Europe (DATE), Mar. 2017
- **"An Energy-Efficient Wireless Video Sensor Node with a Region-of-Interest Based Multi-Parameter Rate Controller for Moving Object Surveillance"**
  IEEE Advanced Video and Signal-based Surveillance (AVSS), Aug. 2016
- **"An Energy-Aware Approach to Noise-Robust Moving Object Detection for Low-Power Wireless Image Sensor Platforms"**
  International Symposium on Low Power Electronics and Design (ISLPED), Aug. 2016
- **"A Tactical Smart Communication Module for Real-Time Control of Unmanned Vehicles"**
  Nordic Semiconductor Headquarters, Oslo, Norway, June 2012
- **"Tactical IP Switching Systems for Military Information Networks"**
  Alcatel Lucent Headquarters, Paris, France, June 2011
- **"A Low-Power and High-Performance Flash Memory Controller for Embedded Systems"**
  Saarland University, Germany, July 2004

REFERENCES

- Dr. Saibal Mukhopadhyay
  Professor, School of ECE, Georgia Institute of Technology
  saibal.mukhopadhyay@ece.gatech.edu
- Dr. Sang Lyul Min
  Professor, School of CSE, Seoul National University
  symin@snu.ac.kr
- Dr. Matthai Philipose
  Researcher, Microsoft AI and Research
  matthaip@microsoft.com
- Dr. Sudhakar Yalamanchili
  Professor, School of ECE, Georgia Institute of Technology
  sudha@gatech.edu

SUMMARY OF COMPLETED RESEARCH

The rapid growth of Internet of Things (IoT) is likely to create billions of connected edge devices for multimedia sensing/processing, each operating under tight resource constraints and dynamic conditions. My dissertation research aims to answer the following question: "How can we enable intelligent multimedia processing capabilities at the edge devices with limited resources and dynamic environment?" To address this question, I pursue interdisciplinary research that optimizes multimedia processing at the intersection of model-driven (signal processing) and data-driven (deep learning) approaches, incorporating cross-layer interactions between algorithm, architecture, and circuit design [Figure 1].

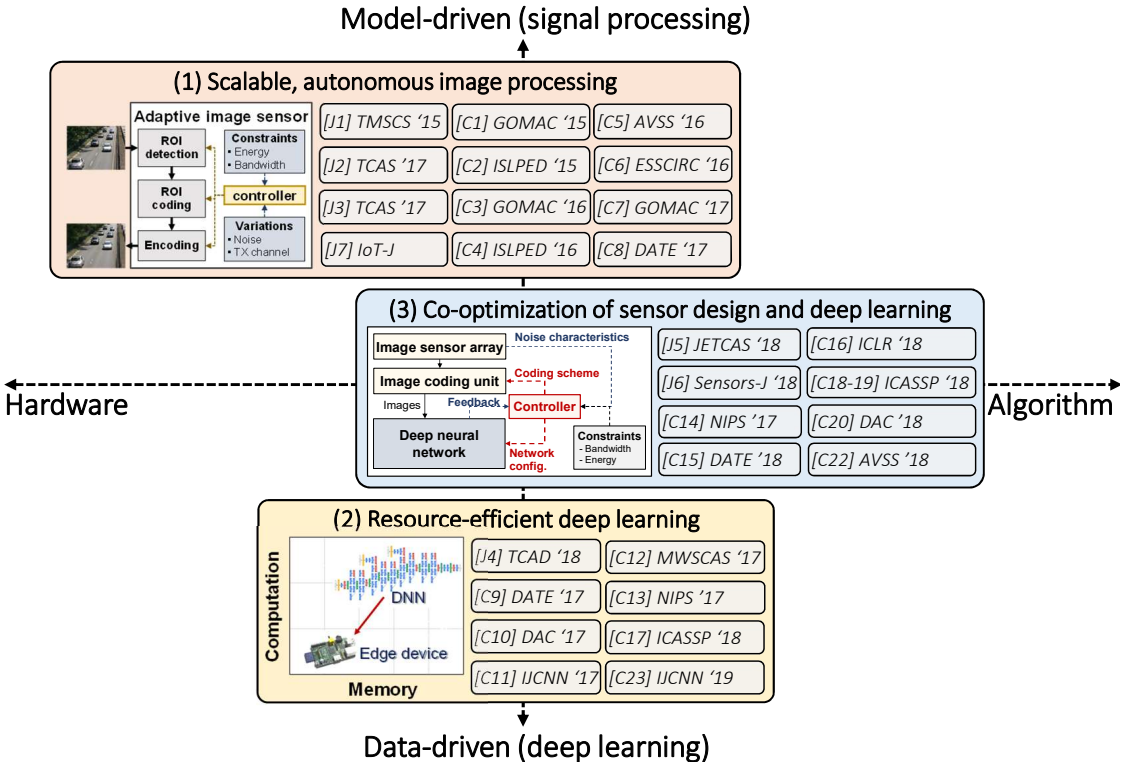
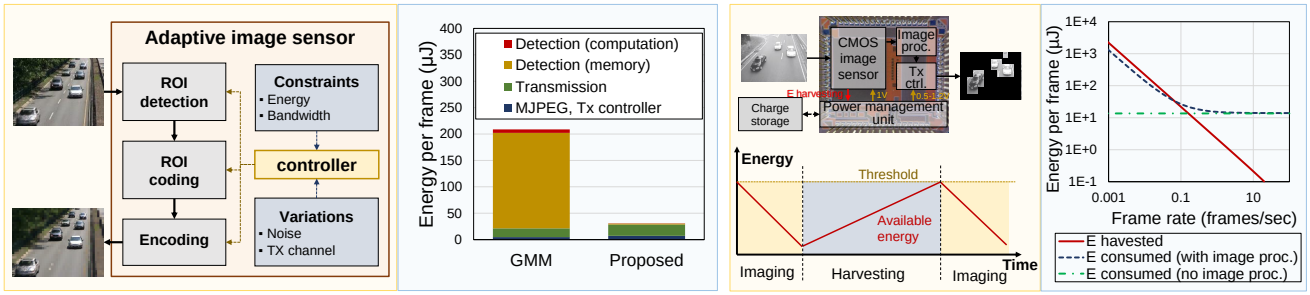


Figure 1: Projection of research topics in a design space from hardware to algorithm, and from model-driven to data-driven approaches.

I first tackled the challenges of existing model-driven algorithms for resource-constrained and dynamic IoT environment by designing **(1) scalable and autonomous image processing algorithms and hardware**. I designed image processing techniques that are scalable to available resources while adaptive to variations in the input noise and transmission channel conditions. The adaptive algorithms together with circuit optimization methods were utilized to demonstrate an autonomous self-powered image sensor node that can sustain its operations without external control or energy supply. To make resource-demanding data-driven multimedia processing approaches applicable to edge devices, I explored **(2) resource-efficient deep learning techniques** that reduce the computation/memory demands of deep neural networks (DNNs) with minimum performance loss. Utilizing my hybrid expertise in sensor platform design and DNN optimization, I am currently exploring **(3) co-optimization of image sensor design and deep learning algorithms**. As a countermeasure to the vulnerability of DNNs to the dynamic environment such as structural/inherent noise and input scene variations, I am developing noise-robust and variation-adaptive deep learning algorithms. I am also investigating design optimization that leverages the interactions between a sensor platform and a DNN to maximize the performance and energy-efficiency of deep learning based sensor platforms. The significance of my work has been recognized by 7 IEEE journal publications and 20+ conference presentations, including the Best Paper Award at ISLPED 2016.

(1) Scalable and autonomous image processing algorithms and hardware

A critical goal in the image sensor node design is to deliver high-quality visual information under stringent energy and bandwidth constraints. This goal becomes more challenging under dynamic conditions such as environmental noise and variations in a wireless channel condition. I tackled this challenge by designing energy-quality scalable region-of-interest (ROI) based image processing techniques with adaptiveness to the dynamic conditions [Fig. 2(a)]. I enhanced the noise robustness of a simple moving object detection method with very low computation/memory overheads, achieving 8.8 × lower energy than Gaussian mixture model at the same ROI quality [Fig. 2(b)] [C1, C4].



(a) (b) (c) (d)

Figure 2: (a) Diagram of adaptive image processing, (b) Energy comparison of the proposed ROI detection with state-of-the-art approach, (c) Block diagram of the self-powered image sensor node (top) and autonomous frame rate control (bottom), (d) Self-power performance depending on the frame rate.

To minimize the buffer overhead, a low-cost controller dynamically controls the ROI coding parameters according to varying transmission bandwidth, while consuming 61% less energy than the H.264 rate controller at the target ROI quality [J1, C5, C7]. As wireless image sensors for remote sensing are usually deployed in the areas where human intervention is prohibitive, one of the most critical capabilities of sensor nodes is an autonomous operation that does not require external control input or energy supply. In response, I implemented the first complete single-chip design of a self-powered image sensor node that includes an energy-harvesting pixel array with a power management unit, an image processing engine with an on-chip SRAM for moving object detection, and a transmission controller [Fig. 2(c)][J2, C3, C6]. The frame rate is autonomously controlled depending on the energy level to maximally utilize the available energy. For further gain in the energy efficiency, I explored various circuit optimization techniques including block-level pipelining and power gating, a low-voltage memory operation, low-area hardware implementation [C2], and voltage margin reduction using clock-data compensation [J3, C8]. The sensor node test results demonstrate 2.1uW of the maximum harvested power, which can self-power the system to process a frame at every 7 seconds [Fig. 2(d)].

(2) Resource-efficient deep learning techniques

While recent advance in deep neural networks has facilitated intelligent multimedia processing, their large computation and memory demands have been a critical challenge for the IoT edge devices. To reduce their memory demand with high flexibility, I proposed a weight compression method based on image encoding that does not require re-training of the network [J4, C9]. By adaptively applying the encoding level of each weight block depending on its error sensitivity, it achieves 15.3× compression of AlexNet for Imagenet with 1% accuracy loss [Fig. 3(a)].

For lower computation demand of convolutional neural networks, I designed the entire frequency-domain convolution approach that simplifies convolution operations without the Fourier transform overhead at every layer, reducing the computation demand of AlexNet by 8× [Fig. 3(b)][C10]. Both computation and memory demands of training process are reduced by dynamically controlling the neuron/weight precision, achieving 4.6× energy reduction of recurrent neural network training [C11]. I also applied the precision control approach to speech applications, achieving significant lower processing time (up to 30×) while providing comparable performance to the state-of-the-art approaches [C13, C17].

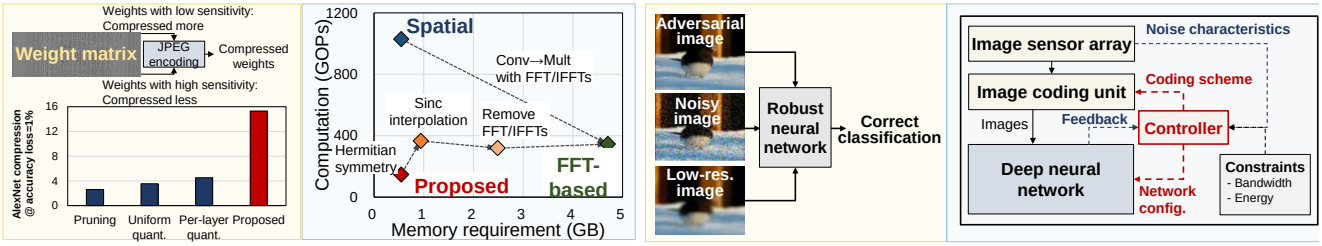


Figure 3: (a) adaptive weight compression approach (top) and performance comparison (bottom), (b) Computation/memory reduction from proposed frequency-domain convolution approach, (c) Concept of robust deep learning techniques and (d) diagram of sensor platform optimization exploiting interaction with a deep neural network.

(3) Co-optimization of image sensor design and deep learning algorithms

My current research focuses on bridging the sensor platform design and deep learning algorithms by jointly optimizing them for higher performance and energy-efficiency. One of the challenges of deploying deep neural networks in sensor platforms is the variations in the input images; structural noise (adversarial images), inherent random noise (image perturbation), and input scene variability [Fig. 3(c)]. To enhance the robustness to adversarial attacks, I co-worked on designing a DNN training algorithm that can ignore unknown pixel level perturbations using an embedding space for both classification and pixel-level similarity learning [C14, C16]. This algorithm is modified to enable noise-robust and resolution-invariant classification, by minimizing the distance between a pair of embeddings from clean and perturbed images [C18]. For detection and localization of anomalous events in videos with variations in the input scenes, I designed a novel framework based on unsupervised learning with reconfigurable target class-based separability [C19].

To maximize the performance and efficiency of deep learning based image processing, I am currently exploring sensor platform design optimized for deep learning by leveraging the interactions between a sensor platform and a DNN [Fig. 3(d)]. I first studied how conventional quality-centric image processing at the image sensor node affects image classification of deep neural networks [J5]. Inspired by this study, I have investigated the use of feedback control from a DNN to adapt image coding schemes for higher classification accuracy [C15, C20, C21]. By designing a 3D-stacked image sensor integrated with DNNs, I studied the effect of transistor- and temperature-induced noise from an image sensor on the neural network accuracy [J6]. Based on the trade-off study between energy, accuracy, and throughput of an image-sensing edge platform with an embedded DNN, I am exploring partitioning of DNN inference between the edge and the host to optimally utilize the available bandwidth and energy [C22].