Chuan Zhang

Southeast University, China

chzhang@seu.edu.cn

Biography

Dr. Chuan Zhang received the B.E. degree in microelectronics and the M.E. degree in very-large scale integration (VLSI) design from Nanjing University, Nanjing, China, in 2006 and 2009, respectively, and the M.S.E.E. and Ph.D. degrees from the Department of Electrical and Computer Engineering, University of Minnesota, Twin Cities (UMN), USA, in 2012.

He is currently the Excellence Professor and the Purple Mountain Professor with Southeast University, Nanjing, China. His current research interests include low-power high-speed VLSI design for digital signal processing and digital communication.

Dr. Zhang now serves as an Associate Editor for the IEEE Transactions on Circuits and Systems-II. He served as an Associate Editor for the IEEE Transactions on Signal Processing and IEEE Open Journal of Circuits and Systems. He served as a Corresponding Guest Editor for the IEEE Journal on Emerging and Selected Topics in Circuits and Systems twice. He is also the Secretary of the Circuits and Systems for Communications TC of the IEEE Circuits and Systems Society. He is also a member of the Circuits and Systems for Communications TC, VLSI Systems and Applications TC, and Digital Signal Processing TC of the IEEE Circuits and Systems Society. He received the Best Contribution Award of the IEEE Asia Pacific Conference on Circuits and Systems (APCCAS) in 2018, the Best Paper Award in 2016, the Best (Student) Paper Award of the IEEE International Conference on DSP in 2016, the Best Paper Award Nomination of the IEEE Workshop on Signal Processing Systems in 2015, the Outstanding Achievement Award of the Intel Collaborative Research Institute in 2018, and the Merit (Student) Paper Award of the IEEE APCCAS in 2008. He also received the Three-Year University-Wide Graduate School Fellowship of UMN and the Doctoral Dissertation Fellowship of UMN.

Lecture 1: Advanced Baseband Processing Circuits and Systems for 5G/B5G Communications

This lecture focuses on Advanced Baseband Processing Circuits and Systems for 5G/B5G Communications: an emerging research field enabling 5G/B5G from theory to practice. By committing itself to the emerging techniques of baseband processing circuits and systems for 5G/B5G, this lecture means to bring a synthesized source and wide view of recent progress and existing challenges in this particular but very important research area, including: 1) channel decoding algorithms and implementations for 5G/B5G baseband processing; 2) massive MIMO detection algorithms and implementations for 5G/B5G baseband processing; 3) algorithm and implementation co-design for 5G/B5G baseband.

Lecture 2: Artificial Intelligence for 5G and Beyond 5G: Implementations, Algorithms, and Optimizations

Due to its undoubted significance, research combining “AI” and “5G and B5G” has drawn lots of attentions from both academia and industry. Although some initiatives related to “AI for 5G and B5G” have been named, their design, implementation, and optimization are unfortunately not complete and of course in infancy. Having lots of potential for AI’s new innovations, advances are required in network architectures, signal processing solutions, semiconductor technologies as well as in its optimization regarding the overall wireless system design. Much of the research has scattered on the design, implementation, and optimization of the corresponding circuits and systems.

This lecture would like to emphasize its uniqueness on “AI for 5G and B5G” related VLSI/IC designs and help readers to know the cutting-edge progresses from the perspective of circuits and systems. With a focus on bridging the gaps between theory and practical implementations, the goal of this lecture is to demonstrate the latest research progress on circuits and systems design for efficiently realizing machine learning in wireless communications. The lecture will bring together academic and industrial aspects to identify technical challenges and recent results related to this area.

Michael Peter Kennedy

University College Dublin, Ireland

peter.kennedy@ucd.ie

Biography

Michael Peter Kennedy received the B.E. degree in electronics from the National University of Ireland, Dublin, the M.S. and Ph.D. degrees from the University of California, Berkeley, and the D.Eng. degree from the National University of Ireland. He has published and lectured on a range of topics in the field of nonlinear circuits and systems including oscillators, chaos, neural networks, mixed-signal testing, phase-locked loops, delta-sigma modulation and frequency synthesis. He was made an IEEE Fellow in 1998 for his contributions to the study of Neural Networks and Nonlinear Dynamics. He was awarded the IEEE Third Millennium Medal, the IEEE Circuits and Systems Society Golden Jubilee Medal, and the RIA Parsons Medal. He has held faculty positions at University College Cork, where he also served as Vice-President for Research and Innovation, and University College Dublin, where he is currently Professor of Microelectronic Engineering. He has had visiting appointments at BME, EPFL, Imperial College London, and the University of Pavia. He has provided consulting services to a number of semiconductor companies and was founding Director of Ireland's Microelectronics Industry Design Association and the Microelectronic Circuits Centre Ireland. He served as President of the Royal Irish Academy from 2017 to 2020.

Lecture 1: Recent Advances in Frequency Synthesis

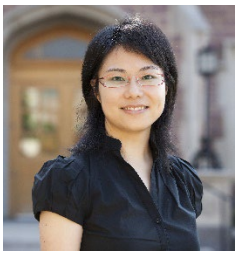
Frequency synthesizers are universally used in a wide range of applications including clocking, communications, instrumentation, and radar. The most common architecture is the fractional-N frequency synthesizer which uses a nonlinear finite state machine to produce the desired frequency. Both the finite state machine itself and interaction between its output and nonlinearities in the implementation can lead to unwanted spurious periodic output frequency components (spurs) and excess noise. Understanding of the origins of these effects has led to the design of novel mitigation strategies.

This talk will explain the underlying issues, explain recent innovations, and highlight open problems.

Lecture 2: The Good, the Bad, and the Ugly of Nonlinear Circuits and Systems

Most engineering education focusses on linear phenomena and linear methods of analysis. Many practical systems, such as power converters, inherently exploit nonlinearity; others, such as data converters, suffer degraded performance due to nonlinearities. As electronics technology is pushed to its physical limits, the need to appreciate how to use and how to mitigate nonlinearity is growing.

This talk will explain basic concepts in nonlinear circuits and systems and show how an appreciation of nonlinearity can lead to insights and ultimately innovative circuits and systems solutions in an increasing range of application domains.



Xuan (Silvia) Zhang

Washington University in St. Louis, USA

xuan.zhang@wustl.edu

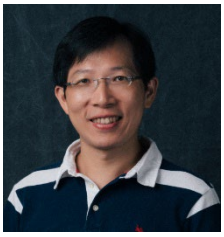
Biography

Dr. Xuan 'Silvia' Zhang is an Associate Professor in the Preston M. Green Department of Electrical and Systems Engineering at Washington University in St. Louis. Before joining Washington University, she was a Postdoctoral Fellow in Computer Science at Harvard University. She received her B. Eng. degree in Electrical Engineering from Tsinghua University in China, and her MS and Ph.D. degrees in Electrical and Computer Engineering from Cornell University. She works across the fields of VLSI design, computer architecture, and design automation and her research interests include hardware/software co-design for efficient machine learning and artificial intelligence, adaptive power and resource management for autonomous systems, and hardware security primitives in analog and mixed-signal domain. Dr. Zhang is the recipient of NSF CAREER Award in 2020, AsianHOST Best Paper Award in 2020, DATE Best Paper Award in 2019, and ISLPED Design Contest Award in 2013, and her work has also been nominated for Best Paper Awards at ASP-DAC 2021, DATE 2019 and DAC 2017.

Lecture 1: Exploring Autonomous Edge Intelligence in the Analog Domain

Recent rapid advancements of artificial intelligence (AI) and machine learning (ML) technologies have fueled many burgeoning smart applications from personalized recommendations to self-driving cars. Despite these unprecedented developments, the holy grail of enabling fully autonomous machine intelligence remains far from our grasp. One key challenge is the lack of performant and efficient hardware implementation to support these computationally demanding tasks, especially on edge devices with stringent resource constraints.

In this talk, I will present research from my lab to tackle the edge intelligence challenge from a unique angle that explores system/design abstractions beyond the conventional binary digital logic. Instead, we leverage the signal representation and information processing ability that are innate in the analog/mixed-signal (AMS) domain. Our AMS-domain method not only leads to novel resistive RAM (RRAM) based analog-to-digital converter designs that exceed the state-of-the-art performance limit, but it also transforms the peripheral circuits in in-memory computing (IMC) architectures which are crucial to delivering much-improved computing performance and efficiency. In the end, I will also reveal how an AMS approach can enrich our technological arsenals from several diverse domains such as learning-assisted design automation, cross-domain system security, and real-time and adversarial autonomy.



Wen-Hsiao Peng

National Chiao Tung University, Taiwan

wpeng@cs.nctu.edu.tw

Biography

Wen-Hsiao Peng received his Ph.D. degree from National Chiao Tung University (NCTU), Taiwan in 2005. He was with the Intel Microprocessor Research Laboratory, USA from 2000 to 2001, where he was involved in the development of ISO/IEC MPEG-4 fine granularity scalability. Since 2003, he has actively participated in the ISO/IEC and ITU-T video coding standardization process and contributed to the development of SVC, HEVC and SCC standards. He is currently a Professor with the Computer Science Department, National Yang Ming Chiao Tung University (NYCU), and was a Visiting Scholar with the IBM Thomas J. Watson Research Center, USA, from 2015 to 2016. He has authored over 70+ journal/conference papers and over 60+ ISO/IEC and ITU-T standards contributions. His research interests include learning-based video/image coding, deep/machine learning, multimedia analytics, and computer vision. Dr. Peng is Chair of the Visual Signal Processing and Communications (VSPC) Technical Committee of the IEEE Circuits and Systems (CAS) Society. He was Technical Program Co-chair for 2021 IEEE VCIP, 2011 IEEE VCIP, 2017 IEEE ISPACS, and 2018 APSIPA ASC; Publication Chair for 2019 IEEE ICIP; Area Chair for IEEE ICME and VCIP; and Review Committee Member for IEEE ISCAS. He served as Associate Editor-in-Chief/Lead Guest Editor/Guest Editor/SEB Member for IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Associate Editor for IEEE Transactions on Circuits and Systems for Video Technology, and Guest Editor for IEEE Transactions on Circuits and Systems II. He was Distinguished Lecturer of APSIPA for the 2017-2018 term.

Lecture 1: (3-hour tutorial) Advances in Design and Implementation of End-to-End Learned Image and Video Compression

The DCT-based transform coding technique was adopted by the international standards (ISO JPEG, ITU H.261/264/265/266, ISO MPEG-2/4/H, and many others) for nearly 30 years. Although researchers are still trying to improve its efficiency by fine-tuning its components and parameters, the basic structure has not changed in the past two decades. The arrival of deep learning recently spurred a new wave of developments in end-to-end learned image and video compression. The seminal work by Balle et al. connects the learning of an image compression system to learning a variational generative model, known as the variational autoencoder (VAE), opening up a new direction for constructing high-efficiency image/video coding systems based on advanced deep generative models. Recent results, particularly from the Challenge on Learned Image Compression (CLIC) at CVPR, indicate that this new type of compression technology achieves comparable or superior compression performance to VVC Intra (the state-of-the-art codec standardized and published in 2020) and has superior subjective quality, especially at the very low bit rates.

In this talk, I shall first (1) summarize briefly the progress of this topic in the past 3 or so years. In the second part, I shall (2) introduce the design concepts of VAE-based image compression. The third part switches gear to (3) explore the emerging area of end-to-end learned video

compression, which has attracted lots of research interest and publications since the advent of the first such system in 2019. Lastly, I shall (4) address the circuits and systems aspects of learned image/video codecs by exploring recent efforts in creating hardware-friendly, low-complexity models. The talk will be concluded with open issues and recent standardization initiatives by the IEEE and other communities.

Lecture 2: (1-hour seminar) Video/Image Coding for Machines

The traditional video/image coding is optimized primarily for preserving the waveforms of video/image signals in a form suitable for human perception. However, the emergence of artificial intelligence (AI)-enabled visual recognition systems often requires video/images (or more broadly, visual data) to be analyzed by machines for image processing and/or computer vision tasks. In many real-world applications, the video/image acquisition happens on resource-limited edge devices, with the compressed bitstream stored or transmitted to the cloud for analytics. This calls for a compact representation of video/images that is optimized not only for human perception but also for machine consumption. The demand has recently kicked off the development of a new type of coding technology in both academia and standards organizations (e.g. JPEG AI and MPEG). In this endeavor, deep learning is emerging as the enabling technology due to its great success in computer vision tasks and learned-based video/image compression. The wide variety of application scenarios makes this rapidly growing field a wide open space for research, inviting contributions from communities of various disciplines.

In this talk, I shall first (1) overview the recent advances in this area, particularly the standardization activities taking placing in JPEG and MPEG. In the second part, I shall (2) review some notable designs that adopt end-to-end learned systems (which replace the compression backbone with neural networks) as solutions. The third part (3) explores the learning-assisted approach (i.e. re-purposing or enhancing the traditional codecs by learning techniques without changing the codecs) to coding for machines. Lastly, I shall (4) discuss their circuits and systems implementations for resource-limited applications.

Yu (Kevin) Cao

Ryerson University, Canada

yu.cao@asu.edu

Biography

Yu Cao received the B.S. degree in physics from Peking University in 1996. He received the M.A. degree in biophysics and the Ph.D. degree in electrical engineering from University of California, Berkeley, in 1999 and 2002, respectively.

He is now a Professor of Electrical Engineering at Arizona State University, Tempe, Arizona. He has published numerous articles and two books on nano-CMOS modeling and physical design. His research interests include neural-inspired computing, hardware design for on-chip learning, and reliable integration of nanoelectronics.

Dr. Cao was a recipient of the 2020 Intel Outstanding Researcher Award, the 2009 ACM SIGDA Outstanding New Faculty Award, the 2006 NSF CAREER Award, the 2006 and 2007 IBM Faculty Award, and four Best Paper Awards. He is a Fellow of the IEEE.

Lecture 1: Reliable In-memory Computing with Unreliable Devices and Circuits

With the ever-increasing demand of AI algorithms and high-definition sensors, Contemporary microprocessor design is facing tremendous challenges in memory bandwidth (i.e., the von Neumann bottleneck), processing speed and power consumption. Leveraging the advances in device technology and design techniques, in-memory computing (IMC) embeds analog deep-learning operations in the memory array, achieving massively parallel computing with high storage density. On the other side, its performance is still limited by device non-idealities, circuit precision, on-chip interconnection, and algorithm properties.

In this talk, we will first review the state-of-the-art IMC design techniques, such as those based on resistive random-access memory (RRAM) and SRAM. Then based on statistical data from a fully integrated 65nm CMOS/RRAM test chip, we will illustrate the bottlenecks of current IMC system, including RRAM variations, the stability of machine learning models, peripheral circuits and interconnection. They interact with each other, limiting the inference accuracy and system energy-delay product (EDP). To efficiently explore design space, we will present a newly developed benchmark simulator, SIAM, which integrates device, circuit, architecture, network-on-chip (NoC), network-on-package (NoP) and DRAM access models to address the bottlenecks in data movement and robustness. Furthermore, we will demonstrate two methods to recover the accuracy loss: training for model stability before mapping to the hardware, and a hybrid SRAM/RRAM architecture for post-mapping recovery. These methods are applied to various datasets as well as a 65nm SRAM/RRAM test chip, helping shed light on future IMC research focus.

Gabriel A. Rincón-Mora

Georgia Institute of Technology, USA

rincon-mora@gatech.edu

Biography

Gabriel Alfonso Rincón-Mora is a Fellow of the National Academy of Inventors, Fellow of the IEEE, and Fellow of the Institution of Engineering and Technology. He was Design Team Leader at Texas Instruments in 1994–2003, Adjunct Professor at the Georgia Institute of Technology (Georgia Tech) in 1999–2001, Assistant/Associate/Full Professor at Georgia Tech since 2001, Director of the Georgia Tech Analog Consortium in 2001–2004, Director of the TI Analog Fellowship Program in 2001–2015, and Visiting Professor at National Cheng Kung University in Taiwan in 2011–2019. He was inducted into Georgia Tech's Council of Outstanding Young Engineering Alumni and named one of "The 100 Most Influential Hispanics" by Hispanic Business magazine. Other distinctions include the National Hispanic in Technology Award, Charles E. Perry Visionary Award, Three-Year Patent Award, Orgullo Hispano Award, Hispanic Heritage Award, State of California Commendation Certificate, IEEE Service Award, and included in "List of Notable Venezuelan Americans" in Science. His scholarly output includes 11 books, 8 handbooks, 4 book chapters, 42 patents issued, over 190 articles, 25 educational videos, over 26 commercial power-chip products released to production, and over 150 keynotes/speeches/seminars. URL: rincon-mora.gatech.edu.

Lecture 1: Unraveling Feedback Translations

Power supplies and other analog systems rely on feedback translations for control and stability. With sufficient loop gain, these translations are largely independent of loop gain and loop dynamics. Complex feedback systems, however, often lose the phase and gain margin they need for stable operation when loop gain is high. Unfortunately, understanding how these feedback systems behave and translate signals when loop gain is not high is largely algebraic and abstract. This talk introduces a new way of viewing and analyzing these feedback systems that is more intuitive and insightful. The presentation uses this method to show how looped amplifiers translate signals across frequency when forward gain and feedback translations alternate dominance. Understanding, innovating, and designing feedback and mixed translations this way is more straightforward.

Lecture 2: Switched-Inductor Power Supplies: Compact Control Loops

Switched-inductor power supplies are pervasive in electronics. This is because they deliver a large fraction of the power they draw from the input source with an output voltage or current that is largely independent of the load. Keeping the output voltage or current steady this way is ultimately the responsibility of the feedback controller. This talk uses insight and intuition to show how pulse-width-modulated (PWM) and hysteretic loops switch the inductor, offset the current or voltage they control, and respond to load dumps. The presentation also shows how summing comparators work and how they can contract, offset, and compensate (for reduced offset) these

control loops. With this background and understanding in hand, designing compact feedback controllers for switched-inductor power supplies is more straightforward.

Hongliang Li

University of Electronic Science and Technology of China, China

hlli@uestc.edu.cn

Biography

Hongliang Li is currently a full Professor in the School of Information and Communication Engineering, University of Electronic Science and Technology of China. His research interests include image and video processing, visual attention, object detection and segmentation, object recognition and parsing, and multimedia content analysis. He has authored or co-authored more than 150 technical papers in international conferences and journals, such as IEEE T-CSVT, IEEE TIP, IEEE TMM, etc. He is a co-editor of a Springer book titled “Video segmentation and its applications”. He was involved in many professional activities. He is an Associate Editor of IEEE Transactions on Circuits and Systems for Video Technology, IEEE Open Journal of Circuits and Systems, and Journal on Visual Communication and Image Representation, and an Area Editor of Signal Processing: Image Communication (Elsevier Science). He received the 2019 and 2020 Best Associate Editor Awards for IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), 2021 Best Editor Award and 2018 Outstanding Service Award to JVCI and its Community. He served as a General Chair of ISPACS 2017, the Technical Program Chairs of VCIP2016 and PCM 2017, and the Local Chair of IEEE ICME 2014. He served in program committees and session chairs for many IEEE conferences, e.g., AAAI2022, ISCAS2021, ISCAS2020, ISCAS2019, VCIP2021, and ICME2014. He is/was the member of IEEE CAS VSPC TC and IEEE CS MCTC. He is now a senior member of IEEE.

Lecture 1: Multi-Modal Scene Understanding Theory and Application

Multi-modal tasks aim to build models that can process and integrate the cross-modal information, including images, video, speech, text, etc., which have always been hot issues in computer vision, such as image caption, video caption, visual retrieval, and referring expression comprehension. This lecture first focuses on the video caption task that expects to generate descriptive sentences about the video. The prior information of word dataset in the captions is employed to predict POS tag, which can help the model to thoroughly understand video content and precisely map them into corresponding sentence components. Then this talk focuses on building more accurate vision-language mappings for the referring expression comprehension tasks. A key word extraction model is presented to extract key words from the language query to suppress the noise in the query and to highlight the desired object. In addition, a language-aware deformable convolution model adaptively samples a set of key points to capture fine-grained object information, the cross-modal bidirectional interactions are further introduced to learn precise object and language representations. Meanwhile, a query reconstruction network is built to confirm the vision-language consistency for bidirectional vision-language mapping. Finally, the challenges and applications of Multi-modal scene understanding in the future will be summarized.

Lecture 2: Visual Object Detection: Challenge and Approaches

Object detection is a fundamental research topic in computer vision and has made remarkable progress in recent years. However, natural images usually contain objects with various categories, sizes and semantic confusion, which makes it challenging to accurately locate these objects. It aims to simultaneously locate objects of interest and recognize their categories information. This lecture will discuss recent research directions and hot topics for the object detection. Firstly, a novel detection method is introduced by employing a set of growing cross lines as object represents. An object is flexibly represented as cross lines in different combinations, which will contribute to enhance discriminability of object features for classification and accurately find the object boundaries location. Then, by regarding the object regression as a classification problem, a new strategy is introduced to predict accurate object location, which provides different penalties for diffident samples and avoids the gradient explosion problem caused by samples with large errors. Furthermore, this talk will discuss how to attack the class label noise problem, which employs different losses to describe different roles of noisy class labels to enhance the learning. Finally, the challenges and opportunities of object detection in the future will be summarized.

Massimo Alioto

National University of Singapore, Singapore

malioto@ieee.org

Biography

Massimo Alioto is a Professor at the ECE Department of the National University of Singapore, where he leads the Green IC group, the Integrated Circuits and Embedded Systems area, and the FD-fAbrICS center on intelligent&connected systems. Previously, he held positions at the University of Siena, Intel Labs – CRL (2013), University of Michigan - Ann Arbor (2011-2012), University of California – Berkeley (2009-2011), EPFL - Lausanne.

He is (co)author of 330 publications on journals and conference proceedings, and four books with Springer. His primary research interests include ultra-low power and self-powered systems, green computing, circuits for machine intelligence, hardware security, and emerging technologies.

He is the Editor in Chief of the IEEE Transactions on VLSI Systems, Distinguished Lecturer for the IEEE Solid-State Circuits Society, and was Deputy Editor in Chief of the IEEE Journal on Emerging and Selected Topics in Circuits and Systems. Previously, Prof. Alioto was the Chair of the "VLSI Systems and Applications" Technical Committee of the IEEE Circuits and Systems Society (2010-2012), Distinguished Lecturer (2009-2010) and member of the Board of Governors (2015-2020). He served as Guest Editor of numerous journal special issues, Technical Program Chair of several IEEE conferences (ISCAS 2023, SOCC, PRIME, ICECS), and TPC member (ISSCC, ASSCC). Prof. Alioto is an IEEE Fellow.

Lecture 1: Green Technologies for Intelligent and Connected Circuits & Systems Powered by Renewable Energy Sources

Recent semiconductor scaling trends continue to support the evolution of silicon systems beyond the inevitable end of technology scaling, growing the deployment of intelligent and connect chips towards the trillion range by the end of the decade. Such evolution vastly outranges any application ever deployed by human beings, and its sustained growth is now fundamentally impeded by batteries as conventional source of energy. From a silicon chip viewpoint, batteries at the trillion scale severely limit advances in cost, form factor, system lifespan and chip availability over time. From a societal perspective, batteries in the trillions threaten economic and environmental sustainability of the underlying scaling trend, and hence its feasibility.

This keynote introduces key concepts and silicon demonstrations of a new breed of always-on silicon systems with ultra-wide power adaptation down to nWs, and no battery inside (or other energy storage). Adaptation to the highly-fluctuating power profile of energy harvesters is shown to enable next-generation pervasive integrated systems with cost well below 1$, size of few millimeters, long lifetime well beyond the traditional shelf life of batteries, yet at near-100% up-time. The principles are exemplified by numerous silicon demonstrations of sensor interfaces, processing, power management and wireless communications, as well as of full systems. Ultimately, the technological pathway discussed in this keynote supports the sustained growth of applications leveraging large-scale deployments of silicon systems, making our planet smarter. And greener too.

Lecture 2: Securing the Next Trillion of Chips via In-Memory and Immersed-in-Logic Design – Beyond Traditional Design Boundaries

Divide-and-conquer design methodologies facilitate building block design, but conflict with basic security requirements, while also precluding opportunities for efficient system integration and inexpensive embedment of security features. Indeed, conventional design partitioning vastly facilitates the identification of attack targets, and reduces the related effort by focusing on specific areas of the overall attack surface. At the same time, the insertion of security primitives as standalone blocks is inherently additive in terms of area, power, design effort and integration effort, limiting their embeddability in low-cost devices (i.e., the vast majority of the upcoming trillion chips for the Internet of Things).

In this keynote, the road towards ubiquitous hardware security is pursued from a primitive design perspective, designing PUFs and TRNGs that are inherently immersed in existing memory arrays and logic fabrics, and breaking the boundaries of traditional system partitioning. From a non-recurring engineering cost viewpoint, design and system integration entail lower effort and very low silicon area thanks to extensive circuit reuse, while also facilitating technology and design portability. At the same time, their immersed and distributed nature offers inherent physical-level obfuscation against several physical attacks targeting specific primitive instances with well-defined boundaries and ports, while also allowing full reuse of conventional techniques to protect memories and logic. Stricter data locality also facilitates architecture-level security, confining secure keys within the same logic module that they are used in (e.g., within the same cryptographic engine, or within the same memory encrypting its own data). Several silicon demonstrations are illustrated to quantify the benefits and the limits of existing techniques, and identify opportunities and challenges for the decade ahead. At the end of the keynote, fundamental directions on how to make hardware security more pervasive and unceasing are discussed.

Partha Pratim Pande

Washington State University, USA

pande@wsu.edu

Biography

is a professor and holder of the Boeing Centennial Chair in computer engineering at the school of Electrical Engineering and Computer Science, Washington State University, Pullman, USA. He is currently the director of the school. His current research interests are novel interconnect architectures for manycore chips, on-chip wireless communication networks, and heterogeneous architectures. Dr. Pande currently serves as the Editor-in-Chief (EIC) of IEEE Design and Test (D&T). He is on the editorial boards of IEEE Transactions on VLSI (TVLSI) and ACM Journal of Emerging Technologies in Computing Systems (JETC) and IEEE Embedded Systems letters. He was/is the technical program committee chair of IEEE/ACM Network-on-Chip Symposium 2015 and CASES (2019-2020). He also serves on the program committees of many reputed international conferences. He has won the NSF CAREER award in 2009. He is the winner of the Anjan Bose outstanding researcher award from the college of engineering, Washington State University in 2013. He is a fellow of IEEE.

Lecture 1: Processing-in-memory (PIM)-based Manycore Architecture for Training Graph Neural Networks

Graph Neural Networks (GNNs) enable comprehensive predictive analytics over graph structured data. They have become popular in diverse real-world applications. A key challenge in facilitating such analytics is to learn good representations over nodes, edges, and graphs. Unlike traditional Deep Neural Networks (DNNs), which work over regular structures (images or sequences), GNNs operate on graphs. The computations associated with GNN can be divided into two parts: 1) Vertex-centric computations involving trainable weights, like conventional DNNs, and 2) Edge-centric computations, which involve accumulating neighboring vertices information along the edges of the graphs. Hence, GNN training exhibits characteristics of both DNN training, which is compute-intensive, and graph computation that exhibits heavy data exchange. Conventional CPU- or GPU-based systems are not tailor-made for applications that exhibits such trait. This necessitates the development of new and efficient hardware architectures tailored for GNN training/inference. Both the vertex- and edge-centric computations in GNNs can be represented as multiply-and-accumulate (MAC) operations, which can be efficiently implemented using resistive random-access memory or ReRAM-based architectures. In addition, ReRAMs allow for processing in-memory, which helps reduce the amount of communication (data transfers) between computing cores and the main memory. This is particularly useful for GNN training as it involves repeated feature aggregation along the graph edges. The in-memory nature of ReRAM's computation significantly reduces the on-chip traffic leading to better performance. However, existing ReRAM-based architectures are designed to accelerate specifically either DNNs or graph computations. As GNN training exhibits characteristics of both DNNs and graph computations, these tailor-made architectures are not well suited for efficient GNN training. In this talk we will present design and performance evaluation of a novel ReRAM-based manycore architecture that caters to the specific characteristics exhibited by GNN training.

Lecture 2: Interconnect Meets Architecture: On-Chip Communication in the Age of Heterogeneity

Neural Networks, graph analytics, and other big-data applications have become vastly important for many domains. This has led to a search for proper computing systems that can efficiently utilize the tremendous amount of data parallelism that is associated with these applications. Generally, we depend on data centers and high-performance computing (HPC) clusters to run various big-data applications. However, the design of data centers is dominated by power, thermal, and physical constraints. On the contrary, emerging heterogeneous manycore processing platforms that consist of CPU and GPU cores along with memory controllers (MCs) and accelerators have small footprints. Moreover, they offer power and area-efficient tradeoffs for running big-data applications. Consequently, heterogeneous manycore computing platforms represent a powerful alternative to the data center-oriented type of computing. However, typical Network-On-Chip (NoC) infrastructures employed on conventional manycore platforms are highly sub-optimal to handle specific needs CPUs, GPUs, and accelerators. To address this challenge, we need to come up with a holistic approach to design an optimal network-on-chip (NoC) as the interconnection backbone for the heterogeneous manycore platforms that can handle CPU, GPU, and application-specific accelerator communication requirements efficiently. We will discuss design of a hybrid NoC architecture suitable for heterogeneous manycore platforms. We will also highlight effectiveness of machine learning-inspired multi-objective optimization (MOO) algorithms to quickly find a NoC that satisfies both CPU and GPU communication requirements. Widely used MOO techniques (e.g., NSGA-II or simulated annealing based AMOSA) can require significant amounts of time due to their exploratory nature. Therefore, more efficient, and scalable ML-based optimization techniques are required. We are going to discuss various features of a generalized application-agnostic heterogeneous NoC design that achieves similar levels of performance (latency, throughput, energy, and temperature) as application-specific designs.

Lecture 3: Bringing Cores Closer Together: The Wireless Revolution in On-Chip Communication

The continuing progress and integration levels in silicon technologies make complete end-user systems on a single chip possible. This massive level of integration makes modern manycore chips all pervasive in domains ranging from weather forecasting, astronomical data analysis, and biological applications to consumer electronics and smart phones. Network-on-Chips (NoCs) have emerged as communication backbones to enable a high degree of integration in manycore platforms. Despite their advantages, an important performance limitation in traditional NoCs arises from planar metal interconnect-based multi-hop communications, wherein the data transfer between far-apart blocks causes high latency and power consumption. The latency, power consumption, and interconnect routing problems of NoCs can be simultaneously addressed by replacing multi-hop wired paths with high-bandwidth single-hop long-range wireless links. In this talk, we will present design of the millimeter (mm)-wave wireless NoC architectures. We will present detailed performance evaluation and necessary design trade-offs for the small-world network-enabled wireless NoCs with respect to their conventional wireline counterparts in presence of both conventional CMP and emerging big data workloads. We will discuss how Machine Learning can be exploited to design energy efficient Wireless NoC architectures. We will finish this presentation by discussing how the wireless NoC paradigm can enable realization of datacenter-on-chip using heterogeneous processing cores.

Jennifer Blain Christen

Arizona State University, USA

Jennifer.Blainchristen@asu.edu

Biography

Jennifer Blain Christen received a bachelor's degree (1999), master's degree (2001) and doctorate (2006) in electrical and computer engineering from the Johns Hopkins University. She completed a postdoctoral fellowship at the Johns Hopkins School of Medicine in the Immunogentics Department. Blain Christen is currently leading the BioElectrical Systems and Technology group at Arizona State University where she is an associate professor in the School of Electrical, Computer, and Energy Engineering. Her research has recently focused on point of care/need diagnostics, wearable diagnostics, field deployable sensors, and flexible neural interfaces. She is an NSF CAREER awardee, NSF ADVANCE fellow, DII@FSE (diversity and inclusion) taskforce member, Fulton Entrepreneurial Fellow, and Flinn Foundation Scholar. She is the co-founder of FlexBioTech, a faculty-based startup with a mission centered on medical equity.

Lecture 1: Fluorescence-based Lateral Flow Immunoassays (LFIA) for Quantitative, Multiplexed Detection of HPV

Cervical cancer has a disproportionately high mortality rate in India and across many countries with high healthcare disparity. While cervical cancer is straightforward to treat, especially with early detection, lack of awareness and screening results in hundreds of thousand preventable deaths each year. We describe collaborative work with AIIMS (All India Institute of Medical Science) New Delhi to screen and treat in some of the most underserved communities. This talk will describe a low-cost, multiplexed fluorescence detection platform with high sensitivity and wide dynamic range. An overview of the design constraints for low resource settings will be given with respect to the design, manufacture, ruggedizing, and training. Our system features inexpensive 3×3mm interference filters with a high stopband rejection, sharp transition edges, and greater than 90% transmission in the passband. In addition to the filters, we improve signal-to-noise ratio by leveraging time for accuracy using a charge-integration-based readout using an integrate and fire inspired architecture. The fluorescence sensing platform provides a sensitivity to photon flux of approximately 1 × 104 photons/mm2 sec and has the potential for 2–3 orders of magnitude improvement in sensitivity over standard colorimetric detection that uses colored latex microspheres. We also detail the design, development, and characterization of our low-cost fluorescence detection platform and demonstrate 100% and 97.96% reduction in crosstalk probability and filter cost, respectively. This is achieved by reducing filter dimensions and ensuring appropriate channel isolation in a multiplexed configuration. Practical considerations with low-cost interference filter system design, analysis, and system performance are also discussed. We will present the results from the training set at AIIMS along with the clinical partnership and the role of transdisciplinary collaboration in a successful deployment of the study.

Lecture 2: Low-Cost COVID-19 Diagnostics for Low Resources Settings, Bringing Point of Need Testing to Low Resource Environments

Point of Need testing for infectious diseases has been of paramount importance in reducing transmission of COVID-19. Unfortunately, many of the most underserved populations lack access to high-sensitivity testing contributing to disproportionate mortality rates within these communities. An approach to nucleic acid-based testing using low-cost microfluidics and handheld fluorescence-based sensing is presented. The diagnostic system is designed for use by minimally trained individuals in a point of need setting (outside of a medical environment) with results available in under an hour. Unlike rapid antigen testing, the presented technology can detect RNA with high CT value (small number of RNA copies) using isothermal amplification of the nucleic acid. The system integration for the reader will be described covering the thermal, optical, and electrical integration. Challenges in maintaining closed-loop PID thermal control of reaction chambers while performing optical excitation and measurements in those chambers is detailed along with the electrical control and sampling of fluorescence over time which enables the low false positive/false negative rates. The importance of the multiplexed approach with the continual emergence of new COVID-19 variants will be discussed describing the approaches and benefits of mutation resilient design. Engineering tradeoffs in time-domain and charge-based detection will be described along with the challenges and lessons learned. In addition, the microfluidic approach that enables low-cost testing with simple mechanical actuation is described. This approach allows for a highly portable pneumatic- and pump-free instrument capable of implementing nucleic acid-based testing. The design approaches considered and implemented will be discussed along with the tradeoffs and limitations based on practical implementation. In addition, issues of supply chain and manufacture in the pandemic and their influence on the research and development will be discussed.

Lecture 3: Smart Nets – Machine Learning-based Approach to Protecting Marine Life

Fishing gear has proven to be an incredible threat to many of the most endangered marine megafauna species; it is in fact the biggest threat to many endangered sea turtle species. Bycatch (unintentional catch of species by fishers) reduction technology has proven effective in dramatically reducing the mortality of many marine species from sea turtles to marine birds. Unfortunately, current approaches to the design of bycatch reduction technology (BRT) involves a “guess and check” method with years long trials run to determine the efficacy of each design. Small variations in BRT require multi-million-dollar trials with labor intensive identification and counting of each organism in a fishing net. We will describe a machine learning-based approach to optimizing the BRT parameters using underwater imaging to observe the response of marine megafauna to emitted deterrent signals that are varied based to determine the tradeoffs between efficacy and energy efficiency. The data must be evaluated with respect to the modulation of the emitted signals by the highly dynamic marine environment. We present an approach to machine learning for behavioral identification along with the design of smart BRT capable of modulating the emitted cues. Furthermore, we present the design and testing of solar powered Smart Net buoys that emit programmable sensory cues that can be varied autonomously to enable remote testing and design space exploration in closed loop with the machine learning algorithms. LED-based optical emitters, solar panels and charging circuits, and microcontroller wake-up circuit will be described. Materials and methods for deploying up to 100 meters below sea level along optical challenges of imaging through the chambers in turbid environments. Finally, we will describe the importance of participatory design through workshops and collaboration with fishers emphasizing the importance of understanding the perspectives of all stakeholders.

Lecture 4: Fully Passive Wireless Sensors for Biosignal Acquisition

Wireless implanted sensors can record high-resolution biosignal inside body without the lengthy and cumbersome cables. Existing wireless systems often require intracranial wires to connect implanted electrodes to an external head-stage that is either battery or inductive powered, raising safety concerns such as infection, electronics failure, or heat-induced tissue damage. In this lecture, we introduce fully-passive wireless sensors aiming at continuously record biosignal from the cortex without battery, power consuming electronics, or the transcranial wiring with high risk of infection.

The proposed fully-passive wireless sensor uses radio frequency (RF) backscattering method is engineered to achieve wireless telemetry. In this method, RF energy is not consumed by the sensor but reflected to an external interrogator. This process is accomplished using only passive electronics: antenna, varactors, inductors and capacitors. As a result, the sensor exhibits several advantages over traditional wireless sensors, including its small-size, high flexibility and near-zero power consumption. These features make the device ideal for long term wireless implantable biosignal acquisition. This lecture will introduce how the sensors are designed, fabricated, and verified in vitro and in vivo. In addition to its use as an implantable device, the lecture will cover other application including wearable electronics, as I will demonstrate its capability to measure various biopotentials. In addition, the lecture will describe wireless, passive, flexible device for other biomedical field such as cardiac tissue stimulation. Finally, I will cover the integration of MEMS-based sensors and actuators to address multiple neurological disorders including hydrocephalus and epilepsy.