Oral S01

Advanced Neural Network Processors and Systems

Date/Time 8/3 (三) 13:30-14:30

Chair(s) 李國君教授 / 成功大學電機工程學系

S01.1 13:30 — 13:45

An Ultra-Low-Power Neural Signal Processor for Seizure Prediction

Yi-Yen Hsieh, Yu-Cheng Lin, and Chia-Hsiang Yang
Graduate Institute of Electronics Engineering, National Taiwan University

This work proposes the world's first integrated neural signal processor for closed-loop neuromodulation. The area cost of the energy operator is reduced by 28% with an approximated energy operator (AEO). The proposed scaling-based Newton-Raphson divider achieves a 2.7x higher convergence speed. For the alternating direction method of multipliers (ADMM)-based SVM training, the proposed pointer-based matrix multiplication (PBMM) reduces 99.9% of operations. With LDL decomposition, the required number of multiplications is reduced by up to 82%. For seizure prediction, the chip achieves a sensitivity of 92.0% and a false alarm rate (FAR) of 0.57/h with a training latency of 8.44ms and a power dissipation of 2.31mW at 6.05MHz. The performance of seizure detection also surpasses the existing literature with the dedicated hardware implementation. Compared with a high-end CPU, this work achieves a 2.45x10^4x higher area efficiency and a 1.32x10^6x higher energy efficiency.

S01.2 13:45 — 14:00

A Lego-based Neural Network Design Methodology by using Flexible NoC

Kun-Chih Chen, Yi-Sheng Liao, and Cheng-Kang Tsai
Department of Computer Science and Engineering, National Sun Yat-sen University

Deep Neural Networks (DNNs) have shown superiority in solving the problems of classification and recognition in recent years. However, DNN hardware implementation is challenging due to the high computational complexity and diverse dataflow in different DNN models. A large body of research has focused on accelerating specific DNN models or layers and proposed dedicated designs to mitigate this design challenge. However, dedicated designs for specific DNN models or layers limit the design flexibility. This work takes advantage of the similarity among different DNN models and proposes a novel Lego-based Deep Neural Network on a Chip (DNNoC) design methodology. We work on common neural computing units (e.g., multiply-accumulation and pooling) and create some neuron computing units called NeuLego processing elements (NeuLegoPEs). These NeuLegoPEs are then interconnected using a flexible Network-on-Chip (NoC) to construct different DNN models. To support large-scale DNN models, we enhance the reusability of each NeuLegoPE by proposing a Lego placement method. The proposed design methodology allows leveraging different DNN model implementations, helping to reduce implementation cost and time-to-market. Compared with the conventional approaches, the

proposed approach can improve the average throughput by 2,802% for target DNN models. Besides, the corresponding hardware is implemented to validate the proposed design methodology, showing on average 12,523% hardware efficiency improvement by considering the throughput and area overhead simultaneously.

S01.3 14:00 — 14:15

Sparse Compressed Spiking Neural Network Accelerator for Object Detection

Hong-Han Lien[1] and Tian Sheuan Chang[2]
[1]Program of Artificial Intelligence, National Yang Ming Chiao Tung University
[2]Institute of Electronics, National Yang Ming Chiao Tung University

Spiking neural networks (SNNs), which are inspired by the human brain, have recently gained popularity due to their relatively simple and low-power hardware for transmitting binary spikes and highly sparse activation maps. However, because SNNs contain extra time dimension information, the SNN accelerator will require more buffers and take longer to infer, especially for the more difficult high-resolution object detection task. As a result, this paper proposes a sparse compressed spiking neural network accelerator that takes advantage of the high sparsity of activation maps and weights by utilizing the proposed gated one-to-all product for low power and highly parallel model execution. The experimental result of the neural network shows 71.5$\%$ mAP with mixed (1,3) time steps on the IVS 3cls dataset. The accelerator with the TSMC 28nm CMOS process can achieve 1024$\times$576@29 frames per second processing when running at 500MHz with 35.88TOPS/W energy efficiency and 1.05mJ energy consumption per frame.

S01.4 14:15 — 14:30

Real-Time Block-Based Embedded CNN for Gesture Classification on an FPGA

Ching-Chen Wang[1], Yu-Chun Ding[2], Ching-Te Chiu[3], Chao-Tsung Huang[2], Yen-Yu Cheng[3], Shih-Yi Sun[2], Chih-Han Cheng[3], and Hsueh-Kai Kuo[3]
[1]Ambarella
[2]Department of Electrical Engineering, National Tsing Hua University
[3]Department of Computer Science, National Tsing Hua University

This paper presents a block-based embedded convolutional neural network (CNN) for gesture classification on field-programmable gate array (FPGA) in real time. Gesture recognition is an important tool to spontaneous interact with human machine interface. Many CNN architectures using RGB images have been proposed for gesture classification. RGB based gesture classification may cause incorrect results under insufficient light or similar gestures. In addition, most of the CNN architectures cannot run in real time on edge devices due to their large number of parameters and DRAM data access. In this paper, a block-based CNN using RGB-D data is proposed for gesture classification. Adding depth images to RGB images boots the classification accuracy. A CNN architecture with block-based feature maps is built for embedded FPGA implementations. The total number of parameters of the proposed RGB-D embedded CNN (eCNN) model is only 0.17M and it achieves 99.96% and 99.88% accuracy with 32-bit floating point and 8-bit fixed point implementation for America Sign Language (ASL) data set. The RTL simulation of the proposed eCNN model has the average inference speed of 0.171 milliseconds at frequency of 250MHz for a single pair RGB-D image. Implemented on a FPGA integrated with Microsoft Kinect v2 achieve an inference time in 19.42 ms which achieves high accuracy and real-time performance.

Oral S02

Cutting Edge EDA Research

Date/Time 8/3 (三) 13:30-14:30

Chair(s) 麥偉基教授 / 國立清華大學資訊工程學系

S02.1 13:30 — 13:45

Compatible Equivalence Checking of X-Valued Circuits

Yu-Neng Wang[1], Yun-Rong Luo[1], Po-Chun Chien[2], Ping-Lun Wang[1], Hao-Ren Wang[2], Wan-Hsuan Lin[1], Jie-Hong Roland Jiang[1,2], and Chung-Yang Ric Huang[1,2]
[1]Department of Electrical Engineering, National Taiwan University
[2]Graduate Institute of Electronics Engineering, National Taiwan University

The X-value arises in various contexts of system design. It often represents an unknown value or a don't-care value depending on the application. Verification of X-valued circuits is a crucial task but relatively unaddressed. The challenge of equivalence checking for X-valued circuits, named compatible equivalence checking, is posed in the 2020 ICCAD CAD Contest. In this paper, we present our winning method based on X-value preserving dual-rail encoding and incremental identification of compatible equivalence relation. Experimental results demonstrate the effectiveness of the proposed techniques and the outperformance of our approach in solving more cases than the commercial tool and the other teams among the top 3 of the contest.

S02.2 13:45 — 14:00

Routability-driven Global Placer Target on Removing Global and Local Congestion for VLSI Designs

Jai-Ming Lin[1], Chung-Wei Huang[1], Liang-Chi Zane[1], Min-Chia Tsai[1], Che-Li Lin[2], and Chen-Fa Tsai[2]
[1]Department of Electrical Engineering, National Cheng Kung University
[2]Department of Physical Design Service, Global Unichip Corporation

Cell placement remains a big challenge in the modern VLSI design especially in routability. Routing overflow may come from global and local routing congestion in a placement. To target on resolving these problems, this paper proposes two techniques in a global placement algorithm based on an analytical placement formulation and the multilevel framework. To remove global routing congestion, we consider each net as a movable soft module and propose a novel congestion-aware net penalty model so that a net will receive a larger penalty if it covers more routing congested regions. Therefore, our placement formulation can be more easier to move nets away from routing congested regions than other approaches and has less impact on wirelength. In addition, to relieve local congestion, we propose an inflation technique to expand the area of a cluster according to its internal connectivity intensity and routing congestion occupied by the cluster. The experimental results demonstrate that our approaches can get better routability and wirelength compared to other approaches such as NTUplace4h, NTUplace4dr, and RePlAce.

S02.3 14:00 — 14:15

A Complete PCB Routing Methodology with Concurrent Hierarchical Routing

Shih-Ting Lin[1], Hung-Hsiao Wang[1], Chia-Yu Kuo[1], Yolo Chen[2], and Yih-Lang Li[1]
[1]National Yang Ming Chiao Tung University
[2]Wistron NeWeb Corporation

Trends in high pin density and an increasing number of routing layers complicate printed circuit board (PCB) routing, which is categorized as escape and area routing. Traditional escape routing research has focused on escape routing but has not considered the quality of area routing among chip components at the same time. In this work, we propose a complete PCB routing methodology, including simultaneous escape routing (SER), post-SER refinement, and gridless area routing. The SER completes the layer assignment of all nets and produces an escape order ensuring suitable escape and area routing on each layer. Length-matching constraints and differential pair routing are satisfied in each stage of the routing flow. The experiment results indicate that the proposed PCB routing method can complete routings for seven commercial PCB designs, whereas the commercial PCB tool cannot complete any of them.

S02.4 14:15 — 14:30

An Optimal Algorithm for Splitter and Buffer Insertion in Adiabatic Quantum-Flux-Parametron Circuits

Chao-Yuan Huang, Yi-Chen Chang, Ming-Jer Tsai, and Tsung-Yi Ho
Department of Computer Science, National Tsing Hua University

The Adiabatic Quantum-Flux-Parametron (AQFP), which benefits from low power consumption and rapid switching, is one of the rising superconducting logics. Due to the rapid switching, the delay of the inputs of an AQFP gate is strictly specified so that additional buffers are needed to synchronize the delay. Meanwhile, to maintain the symmetry layout of gates and reduce the undesired parasitic magnetic coupling, the AQFP cell library adopts the minimalist design method in which splitters are employed for the gates with multiple fan-outs. Thus, an AQFP circuit may demand numerous splitters and buffers, resulting in a considerable amount of power consumption and delay. This provides a motivation for proposing an effective splitter and buffer insertion algorithm for the AQFP circuits. In this paper, we propose a dynamic programming-based algorithm that provides an optimal splitter and buffer insertion for each wire of the input circuit. Experimental results show that our method is fast, and has a 10% reduction of additional Josephson Junctions (JJs) in the complicated circuits compared with the state-of-the-art method.

Oral S03

Advanced Analog Circuit Techniques

Date/Time 8/3 (三) 13:30-14:30

Chair(s) 蔡宗亨教授 / 國立中正大學電機工程學系

S03.1 13:30 — 13:45

A Self-powering Wireless Soil-pH and Electrical Conductance Monitoring IC with Hybrid Microbial Electrochemical and Photovoltaic Energy Harvesting

Yao-Wei Huang[1], Chun-Ting Chang[1], Chuan-Yi Wu[1], Chi-Wei Liu[1], Jing-Siang Chen[1], Cong-Sheng Huang[1], Ting-Heng Lu[1], Ling-Chia Chen[1], I-Che Ou[1], Sook-Kuan Lee[2], Yen-Chi Chen[2], Po-Hung Chen[1], Chi-Te Liu[2], Ying-Chih Liao[2], and Yu-Te Liao[1]
[1]National Yang Ming Chiao Tung University
[2]National Taiwan University

This paper presents an energy-autonomous wireless soil pH and electrical conductance measurement IC using soil-microbial and photovoltaic (PV) energy harvesting. The chip scavenges soil/PV energy while achieving a peak efficiency of 81.3% over a 0.05-14 mW power range. A readout circuit attains a sensitivity of -8.8 kHz/pH and 6 kHz·m/S. The 433 MHz chirp-modulation transmitter and self-frequency tracking receiver provide a data rate of 100 kb/s (uplink) and 10 kb/s at a sensitivity of -80 dBm (downlink).

S03.2 13:45 — 14:00

A 14-bit DACLS Ringamp-Based Pipelined-SAR ADC

Jia-Ching Wang and Tai-Haur Kuo
National Cheng Kung University

This work presents a 14-bit pipelined-SAR ADC using a DACLS ring amplifier which can reduce the residue amplification gain error and retain a wideband opamp design. In addition, a latency-reduced (LR) SAR logic is proposed to improve the SAR logic speed around 30%. Furthermore, a customized bypass windows design for the backend ADC of which reduces the power by 30%. Under 130MS/s, this ADC achieves 72.5 dB SNDR with 0.82 mW, resulting in Walden FoM and Schreier FoM of 1.8 fJ/conv. and 181.5 dB, respectively.

S03.3 14:00 — 14:15

An 8Mb DC-Current-Free Binary-to-8b Precision ReRAM Nonvolatile Computing-in-Memory Macro using Time-Space-Readout with 1286.4 - 21.6TOPS/W for Edge-AI Devices

Je-Min Hung[1], Yen-Hsiang Huang[1], Sheng-Po Huang[1], Fu-Chun Chang[1], Tai-Hao Wen[1], Chin-I Su[2], Win-San Khwa[2], Chung-Chuan Lo[1], Ren-Shuo Liu[1], Chih-Cheng Hsieh[1], Kea-Tiong Tang[1], Yu-Der Chih[2], Tsung-Yung Jonathan Chang[2], and Meng-Fan Chang[1]
[1]National Tsing Hua University
[2]TSMC

This study sought to achieve high capacity nonvolatile CIM (nvCIM) with high energy efficiency (EFMAC) and short computing latency (TAC) per output precision (TAC-OUT) for multiply-and-accumulate (MAC) operations with high input (IN) and weight (W) bit-precision. This proposed 22nm 8Mb ReRAM macro is the first uses a time-space readout scheme without DC-current to achieve the highest ever TAC-OUT (1.06-0.757 ns/b) and EFMAC (21.6 TOPS/W) from binary to 8bIN-8bW-19bOUT precision.

S03.4 14:15 — 14:30

A 0.8V Intelligent Vision Sensor with Tiny Convolutional Neural Network and Programmable Weights Using Mixed-Mode Processing-in-Sensor Technique for Image Classification

Tzu-Hsiang Hsu[1], Guan-Cheng Chen[2], Yi-Ren Chen[2], Chung-Chuan Lo[2], Ren-Shuo Liu[2], Meng-Fan Chang[2], Kea-Tiong Tang[2], and Chih-Cheng Hsieh[2]
[1]MediaTek
[2]National Tsing Hua University

A 0.8V intelligent vision sensor with a tiny convolutional neural network and programmable weights using the processing-in-sensor (PIS) technique is implemented for real-time inference applications of low-power edge devices. Using the proposed efficient mixed-mode PIS technique (3×3 convolution with ReLU, 2×2 maximum pooling, and 1×1 fully-connected), the prototype is configured to demonstrate a face detection task with an achieved accuracy of 93.6% and an ultra-low power consumption of 122.6μW at 250fps.

Oral S04

Artificial Intelligence Circuits and Systems (I)

Date/Time 8/3 (三) 14:30-15:30

Chair(s) 吳崇賓教授 / 國立中興大學電機工程學系
鄭湘筠教授 / 中央研究院資訊科技創新研究中心

S04.1 14:30 – 14:42

Hardware-Friendly Weight Pruning Algorithm for CNN Model Size Reduction using Universal Pattern Sets

Wei-Cheng Chou, Cheng-Wei Huang, Yung-Han Chen, and Juinn-Dar Huang
Institute of Electronics, National Yang Ming Chiao Tung University

Pattern-based weight pruning on CNNs has been proven an effective model reduction technique. In this paper, we first present how to select hardware friendly pruning pattern sets that are universal to various models. We then propose a progressive pruning framework, which produces more globally optimized outcomes. Moreover, to the best of our knowledge, this is the first paper dealing with the pruning issue of the first and also the most sensitive layer of a CNN model through a twostaged pruning strategy. Experiment results show that the proposed framework achieves 2.25x/2x computation/model reduction while minimizing the accuracy loss.

S04.2 14:42 – 14:54

A VLSI Implementation of a Local Binary Convolutional Neural Network

Yu-Tong Shen[1], Shan-Chi Yu[2], and Ing-Chao Lin[1]
[1]Department of Computer Science and Information Engineering, National Cheng Kung University
[2]Department of Engineering Science, National Cheng Kung University

In order to reduce the computational complexity of convolutional neural networks (CNNs), the local binary convolutional neural network (LBCNN) is proposed. In this work, we propose a platform that includes a weight preprocessor and layer accelerator for the LBCNN. Our weight preprocessor takes advantages of sparsity in the LBCNN and encodes the weight offline. The layer accelerator effectively uses the encoded data to reduce computational complexity and memory accesses for an inference. When compared to the state-of-the-art design, the experimental results show that the number of clock cycles is reduced 7.9 times, and memory usage is reduced 1.7 times. The synthesized results show that the clock period is reduced 5.7%; the cell area is reduced 19.7%, and the power consumption is reduced 15.2%. The total execution time is 8.38 times better, and the inference accuracy is not affected.

S04.3 14:54 — 15:06

A 62.45 TOPS/W Spike-Based Convolution Neural Network Accelerator with Spatiotemporal Parallel Data Flow and Sparsity Mechanism

Chen-Han Hsu, Yu-Hsiang Cheng, Zhao-fang Li, Ping-Li Huang, and Kea-Tiong Tang
National Tsing Hua University

Convolutional neural networks (CNNs) have been widely used for image recognition and classification in recent years. Low energy consumption is crucial in the circuit design of edge devices, and data reuse is one method of reducing energy consumption. In addition, spiking neural networks (SNNs) are receiving increasing attention due to their low power use. However, the temporal characteristic of SNNs causes repeated data access at different time steps, leading to high energy consumption. In this paper, a spiked-based CNN accelerator that can support various inference time steps and models is proposed. Spatiotemporal parallel data flows are employed to reuse data from different time steps and, convolution operations are used to reduce energy consumption. Furthermore, the accelerator is designed for high sparsity and event driven SNNs. The synthesis achieves power efficiency of 62.45 TOPS/W and area efficiency of 7.58 TOPS/kmm2

S04.4 15:06 — 15:18

Automated Quantization Range Mapping for DAC/ADC Non-linearity in Computing-In-Memory

Chi-Tse Huang, Yu-Chuan Chuang, Ming-Guang Lin, and An-Yeu (Andy) Wu
Graduate Institute of Electronics Engineering, National Taiwan University

Computing-in-memory (CIM) has demonstrated the great potential of analog computing in improving the energy efficiency of matrix-vector multiplications for deep learning applications. Albeit low-power feature of CIM, the non-linearity of digital-to-analog converters (DACs)/analog-to-digital converters (ADCs) causes deviation between the computed outputs and desired values, thus degrading classification accuracy. This paper proposes Automated Quantization Range Mapping (A-QRM) mechanism to mitigate the negative effect of non-linearity on model accuracy. Instead of fixing the quantization range for quantized deep learning models, the proposed A-QRM automatically finds a better quantization range that balances the model capability and quantization errors caused by the non-linearity. Experimental results show that our proposed A-QRM achieves 89.02% and 86.93% of top-1 accuracy in ResNet20 and VGG8 on Cifar-10, respectively, under the non-linearity of DACs/ADCs.

S04.5 🕓 **15:18 — 15:30**

Efficient Dilated Convolution Engine with Group-reordering and Normalization

Tong Wu, Kai-Ping Lin, Shih-Yi Sun, and Chao-Tsung Huang
Department of Electrical Engineering, National Tsing Hua University

Dilated convolution neural network models have demonstrated the ability to increase the receptive field without increasing the model size and complexity. However, direct implementation of dilated convolution is inefficient in terms of the input buffer size and hardware utilization with inserting zeros in the weight kernels. In this paper, we propose a power/area-efficient dilated convolution engine, which supports arbitrary dilation rates ranging from 1 to 16 between different layers. Compared to the baseline architecture, this design has 53.9% and 69.9% of area and power reduction, respectively.

Oral S05

Advanced Synthesis, Reliability and Security Techniques

Date/Time 8/3 (三) 14:30-15:30

Chair(s) 謝東佑教授 / 國立中山大學電機工程學系
温宏斌教授 / 國立陽明交通大學電機工程學系

S05.1 14:30 – 14:42

On Efficient Modeling and Test Generation of Hard-to-Detect Combinational Hardware Trojans

Tong-Yu Hsieh, Hsin-Hsien Lin, and Jun-Tsung Wu
Department of Electrical Engineering, National Sun Yat-sen University

In the context of hardware security, it is crucial to identify the most likely suspects of hardware Trojans in a target circuit. In the literature, this issue has also been widely investigated by targeting various types of hardware Trojans. In this paper, combinational hardware Trojans are considered and a much more efficient methodology is proposed to model Hard-To-Detect (HTD) hardware Trojans that are more hardware-security risky. The experimental results show that the proposed modeling methodology reduces 96.44% hardware Trojans, while that for the previous work is 70.84%. This allows us to put more focus on the riskiest lines where hardware Trojans may be implanted. The other major contribution of this work is that we also develop an efficient test generation methodology that can generate compact test patterns capable of detecting these identified HTD Trojans. In addition to aiming to detect as many HTD Trojans as possible, we also include a particular compression process in our pattern generation methodology in order to minimize the number of the generated test patterns. The experimental results show that the test pattern size is reduced by 84.48% when compared to the previous work. Also, the generated test patterns achieve 97.08% detection coverage of HTD Trojans.

S05.2 14:42 – 14:52

Enhancing ILP-based Identification of Rational-Weight Threshold Logic Gates

Ting-Yu Yeh[1], Yueh Cho[1], Yung-Chih Chen[1], and Wang-Dauh Tseng[2]
[1]Dept. of Electrical Engineering, National Taiwan University of Science and Technology
[2]Dept. of Computer Science and Engineering, Yuan Ze University

In CMOS-based current mode realization, the threshold logic gate (TLG) implementation with rational weights has been shown to be more cost-effective than the conventional TLG implementation without rational weights. The existing method for the rational-weight TLG identification is an integer linear programming (ILP)-based method, which could suffer from inefficiency for a Boolean function with a large number of inputs. This paper presents a two-stage method to enhance the ILP-based method. In the first stage, we propose some rules to directly transform the given conventional TLG into a rational-weight TLG. If the transformation does not reduce the implementation cost, we then conduct the second stage, in which we perform the ILP-based method with a simplified formulation to optimize the

rational-weight TLG. We conducted the experiments on a set of TLGs with 4 ~ 15 inputs. The results show that the proposed method has a competitive quality with an average ratio of 0.92, compared to the ILP-based method. Additionally, the proposed method spent only an average of approximately 21% of CPU time.

S05.3 14:52 — 15:02

A TMR Based Error-Tolerant Memory Protection Scheme without Additional Storage for Machine Learning Applications

Tong-Yu Hsieh, Yu-Ren Chiu, and Wei-Ji Chao
Department of Electrical Engineering, National Sun Yat-sen University

In this paper, a TMR (Triple Modular Redundancy)-based protection scheme for memories for machine learning applications is proposed. In particular, no additional storage is required to store the needed data copies, and thus the incurred cost can be greatly reduced. A pedestrian detection machine learning model is also employed as a case study. Our proposed scheme well exploits the inherent machine error-tolerability of AI machine learning models. A detailed error-tolerability analysis process is carried out. Accordingly, the non-critical memory bits that have negligible impacts on the machine detection results are identified and used to store the data copies of TMR for protecting the critical bits. Possible protecting models for critical bits based on the proposed scheme are then developed and discussed, together with their hardware implementation. The experimental results show that our scheme achieves effective memory protection with only negligible cost.

S05.4 15:02 — 15:12

Diagnosing Double Faulty Chains through Failing Bit Separation

Cheng-Sian Kuo[1], Bing-Han Hsieh[1], James Chien-Mo Li[1], Chris Nigh[2], Gaurav Bhargava[2], and Mason Chern[3]
[1]Department of Electrical Engineering, GIEE, National Taiwan University
[2]Qualcomm Technologies, Inc.
[3]Qualcomm Semiconductor Limited

Diagnosing scan chain faults plays a key role in ramping up production yield. High test compression ratios of modern designs increase the challenge of scan chain diagnosis. We propose a technique to address this problem through separating the superpositioned chain fault effects to diagnose chips with two faulty scan chains. Experiments are conducted on both simulated and silicon test data, and the proposed method showed improvements over commercial tools in resolution (2.38) and accuracy (92.0%). We did find failing chips that potentially have a systematic problem in the same double chains.

S05.5 15:12 — 15:22

Compiler of Reed-Solomon Error Correction Codec for IEEE Std 802.3bs Supporting a Very High Throughput of 400 Gbps

Lin Liu[1], Chi Lai[1], Shi-Yu Huang[1], and Ka-Yi Yeh[2]
[1]Department of Electrical Engineering, National Tsing Hua University
[2]Industrial Technology Research Institute

Error Correction is often indispensable in a modern digital communication system that transmits data at a very high speed. Recently published IEEE Std 802.3bs requires an astounding throughput of 200Gbps or even 400Gbps while using the Reed-Solomon Code (RS-Code) for Error Correction to protect the integrity of the transmitted data. An RS-Codec supporting such a high throughput demands sophisticated hardware implementation. In this paper, we present a Reed-Solomon Codec complier that makes contributions in two aspects. First, our parameterized Codec satisfying IEEE Std 802.3bs using RS(544, 514), in which the throughput can be boosted on demand by setting some "configuration". Second, our RS-Codec can easily produce an area and power efficient RS-Codec design satisfying a target throughput in just minutes while supporting easy process migration. Experimental results using 28nm and 90nm CMOS processes are presented to demonstrate its effectiveness.

S05.6 15:22 — 15:32

Multi-level Exploration for Single-Event Double-Node Upsets (SEDU) in Sub-65nm Radiation-Hardened Latches

Sam M.-H. Hsiao, Lowry P.-T. Wang, Ralf E.-H. Yee, and Charles H.-P. Wen
ECE Department, National Yang Ming Chiao Tung University

A single-event double-node upset (SEDU) may appear to result in an erroneous state of the storage element due to the scalability of transistor features. Therefore, SEDU must be well addressed from the perspective of circuit reliability, especially for safety-critical electronics. To better understand the technology scaling impact, we re-examine SEDU in different advanced technologies (including 7-nm finFET, 45-nm bulk CMOS, and 65-nm bulk CMOS). An integrated multi-level framework is developed with the current-source modeling derived from the device-level TCAD simulation, combined with voltage calculation derived from the circuit-level SPICE simulation.

To adequately capture the probability of errors occurring in the latch design under all possible scenarios, this paper also considers a variety of environmental factors, such as striking angles and technology nodes. Additionally, three classical latch designs (i.e., TMR, DICE, and HLR) have been implemented in different technologies and are well-calibrated for experiments. According to experiment results, it is evident that SEDU is highly dependent on both the physical layout of the design as well as its design style. DICE is found to be the most susceptible to SEDU in all three manufacturing technologies, whereas TMR and HLR can be immune to SEDU in the 45-nm and 65-nm technologies due to a lack of sufficient charge to upset more than two nodes. It is, therefore, essential to consider both the physical layout and the manufacturing technology employed for ensuring the robustness of a radiation-hardened design against particle strikes.

Oral S06

Sensor Interface Circuits

Date/Time 8/3 (三) 14:30-15:30

Chair(s) 洪振傑教授 / 逢甲大學電子工程學系
翁峻鴻教授 / 東海大學電機工程學系

S06.1 14:30 – 14:42

A Buffer Amplifier with Self-Adapted Current for Biomedical Applications

Zu-Jia Lo, Yuan-Chuan Wang, Yun-Jie Huang, Ren-Yong Hung, Yi-Heng Wu, and Sheng-Yu Peng
Department of Electrical Engineering, National Taiwan University of Science and Technology

A power-efficient amplifier with self-adaptive supply current is proposed for biomedical signal buffering application. Without extra sensing and control circuitries, the current consumption of the proposed amplifier increases spontaneously when the input signal is fast or large, achieving a high slew rate. The supply current dwindles back to the low quiescent level autonomously when the output voltage reaches equilibrium. Therefore, the proposed ACABA is power-efficient and suitable for processing physiological signals. A prototype version has been designed and fabricated in a 0:35 μm CMOS process occupying an area of 0:151 mm2. When loaded by a 10 pF capacitor, it consumes 3 μW to achieve a unity-gain bandwidth of 100 kHz, THD of - 60 dB under 1:8Vpp input amplitude, and a slew rate of 7:86V/μs.

S06.2 14:42 – 14:54

A Low-Power Sensing System of VEGF Concentration with Monolithic Electrodes and An All-Digital Sub-Sampling Delay-Locked Loop

Tsung-Wen Sun, Ren-Wei Cheng, and Tsung-Heng Tsai
Department of Electrical Engineering, National Chung Cheng University

A low-noise sensing system using sub-sampling technique to suppress the interface noise is presented. The CMOS MEMS capacitive transducer is integrated on the same chip for the point-of-care (PoC) applications. The biomarker is vascular endothelial growth factor (VEGF). The proposed sensing system converts the capacitance variations caused by the VEGF concentration at the transducer into digital codes with a low-power SAR-assisted time-to-digital converter (TDC). CMOS MEMS interdigitated electrodes are adopted as the transducer. Post-process etching and Au plating are implemented on the surface of the electrodes. Experimental results show that a capacitive resolution of 28.3 fF and a sensing range of VEGF concentration from 1 to 1000 pg/ml are achieved. The power consumption of the sensing system is merely 60.65 μW. The FOM of the proposed capacitance-to-digital converter (CDC) is 0.76 pJ/conv.-step.

S06.3 14:54 — 15:06

A 53mK-NETD Low Power ROIC for 80*60 Microbolometer CMOS Thermal Image Sensor with Ultra High Dynamic Range

Hsin Yu and Chih-Cheng Hsieh
Department of Electrical Engineering, National Tsing Hua University

This paper presents an 53-mK NETD 18-bits Microbolometer ROIC using current-mode folding-integration and hybrid ADC. By using the folding integration operation, wide dynamic range and the tolerance to pixel non-uniformity can be achieved without costing additional power and chip area. Moreover, The proposed coarse fine hybrid ADC achieves the integration capacitance (Cint) mismatch immunity. The prototyped ROIC was fabricated in 0.18 um CMOS process with chip area of 2.28x2.28 mm^2, and it consists of 80x60 pixel array with a pitch of 18 um. The final noise equivalent temperature difference (NETD) is 53-mK, and the power consumption is 12.8mW.

S06.4 15:06 — 15:18

A High-Sensitivity CMOS-MEMS IL-6 Sensing System with an All-Digital Phase-Lock-Loop-Based Capacitance-to-Digital Converter

Chun-Hung Tsai, Yi-Xian Chen, and Tsung-Heng Tsai
Department of Electrical Engineering and Advanced Institute of Manufacturing with High-tech Innovation, National Chung Cheng University

A high-sensitivity electrochemical sensing system integrating the CMOS MEMS electrodes and the readout circuits on a single chip is presented. In-house wet etching post process is used to release the MEMS structure. The Interleukin-6 (IL-6) concentration is detected through the capacitance variations using the proposed vertical interdigitated capacitive MEMS transducer and then digitized by an all-digital phase-locked-loop (ADPLL) based capacitance to digital converter (CDC). Parasitic capacitance and noise are greatly reduced owing to the whole system integration on a single chip. The capacitive MEMS transducer is inserted in the oscillator of the ADPLL. When the sensing capacitance changes, frequency drift will be produced. The ADPLL will force the oscillation frequency to track the reference frequency and produce the digital output codes. Moreover, the proposed CDC can cancel the errors result from post process variations, parasitic capacitance and circuit delays. The sensitivity and sensing range of the proposed sensing system are 0.8 nF/ppm and 0~1000 pg/ml, respectively. The prototype chip is fabricated in TSMC 0.18 μm 1P6M CMOS process with an area of 1.2*1.2 mm2, 1.5mW power consumption and a FoM of 31.65 pJ/step.

S06.5 🕓 **15:18 — 15:30**

A Switched-Capacitor Based Acoustic Feature Extraction for Voice Activity Detection

Meng-Ju Chiang and Soon-Jyh Chang
Department of Electrical Engineering, National Cheng Kung University

This paper presents a low-power acoustic feature extraction (AFE) for voice activity detection (VAD). It adopts analog signal processing to extract voice features. With the low feature rate, it is more power efficiency compared to the digital AFE. To overcome the issues of PVT variations, use the switched-capacitor (SC) bandpass filter (BPF) in this work with the op-amp sharing technique to reduce the power consumption. In addition, the chip is free from off-chip calibration. Measurement results of AFE chip, fabricated in a 0.18 μm CMOS process, has a core area of 0.65 mm2 and consumes 2 μW in 1.8 V supply voltage. It achieves an averaged 92%/97% speech/non-speech hit rate at a 3dB signal-to-noise ratio (SNR) with 32 ms latency, which corresponds to all the specifications of practical application.

Oral S07

Artificial Intelligence Circuits and Systems (II)

Date/Time 8/4 (四) 11:00-12:00

Chair(s) 黃柏蒼教授 / 國立陽明交通大學國際半導體產業學院
湯松年教授 / 中原大學資訊工程學系

S07.1 11:00 — 11:12

A Multi-Precision Neural Network Inference Acceleration System Based on FPGA

Yu-Tung Liu, Guo-Yang Zeng, and Tzi-Dar Chiueh
Graduate Institute of Electronics Engineering, National Taiwan University

Neural networks have achieved excellent results in many fields, such as image recognition, natural language processing, etc. With the increasing penetration of intelligent edge devices, fast and efficient neural network inference has become the focus of many research groups. This paper proposes a mixed-precision neural network architecture and low-power circuit implementation. The accelerator circuit supports weights in three precisions and achieves hardware usage flexibility through the im2col algorithm. Finally, we integrated the algorithm and hardware control flow with PyTorch to create a complete solution for users to perform quantized NN training and FPGA acceleration deployment under one framework. At a clock speed of 150MHz, our U250 FPGA accelerated system is 10.57 times more energy-efficient than the CPU. Another DLA solution using the SoC-based KV260 FPGA module is portable and consumes only a few Watts and achieves similar higher energy efficiency.

S07.2 11:12 — 11:24

An Integration-Friendly CNN Accelerator Architecture with High Utilization and Scalability

Chia-Heng Hu, I-Hao Tseng, Pei-Hsuan Kuo, Yu-Hsiang Huang, and Juinn-Dar Huang
Institute of Electronics, National Yang Ming Chiao Tung University

In this paper a highly scalable VLIW-driven CNN accelerator architecture is proposed. A new VLIW instruction, which specifies all settings of an entire convolution layer and natively supports layer concatenation, is defined. A multi-mode input aligner (MMIA) is developed to efficiently organize input data for various convolution modes. A zero-initial-latency (ZIL) buffer is created to further boost the performance. A strip-based dataflow is adopted to drastically minimize external DRAM accesses. The accelerator is also equipped with an AXI4 on-chip bus interface, an instruction queue, ping-pong DRAM I/O buffers, and is thus ready for efficient and easy SoC integration. An accelerator instance with 576 MACs has been implemented using TSMC 40nm process. The core logic only requires 490K gates and the total internal memory size is merely 286KB. The peak performance is 1440 GOPS @1.25GHz and the core power efficiency is 8.71 TOPS/W. Moreover, the proposed accelerator has also enabled a real-time image semantic segmentation system for autonomous driving on an FPGA system.

S07.3 11:24 — 11:36

A Linear Quantization Training Method for Hardware Constraints of In-Memory Computing Architecture

Hao-Wen Kuo, Zhaofang Li, Yu-Hsiang Cheng, Shih-Ting Lin, Meng-Fan Chang, and Kea-Tiong Tang
Department of Electrical Engineering, National Tsing Hua University

Analog computing is used in in-memory computing (IMC) to improve the energy efficiency of accelerators. However, because of hardware constraints, such as the interleaved structure of the components, and nonideal effects caused by the peripheral circuits, the performance of neural networks is hindered. Therefore, a five-dimensional searching direction method for backpropagation is proposed that can increase top-1 accuracy by 1.62% after quantization in ResNet-50 on ImageNet; in addition, a matrix-vector multiplication quantizer and deviation estimator are proposed that reduce the noise caused by nonideal effects and further improve the top-1 accuracy of ResNet-50 by 1.72% and of Mobile-Net v2 by 21.08%. The proposed methods improve the performance and show state-of-the-art results of neural networks in IMC environments.

S07.4 11:36 — 11:48

Tiling Strategy for CNN Zero-Skipping Accelerator

Yen-Xun Chen, Chih-Hung Kuo, Shi-An Zhan, and Kuan-Hung Chen
Department of Electrical Engineering National Cheng Kung University

For AI edge devices, input data should be partitioned into smaller tiles due to limited on-chip memory. Different tile sizes significantly affect data access times and on-chip memory size. In this paper, we explore different tile strategies to find the appropriate one with the best energy efficiency. We design a Zero-Skipping Accelerator (ZSA) that adopts row stationary data flow to reduce the data movement and achieve highly parallel computing. The Pre-processing module handles sparse features to save energy. Experiments show that a suitable tile size can achieve better performance and energy efficiency.

S07.5 11:48 — 12:00

Static Effective Weight Convolution for Deep learning Accelerators

Tz-Yuang Su, Chun-Yuan Chen, and Tian-Sheuan Chang
Dept. of Electric Engineering, National Yang Ming Chiao Tung University

Eliminating unnecessary operations with effective weight-based convolution (EWC) is a promising approach to accelerate convolutional neural network (CNN) processing. However, a previous approach with a dynamic search for effective weights results in complex control and a high area overhead. This paper presents a static effective weight-based convolution (SEWC) that uses four static effective weights instead of six dynamic effective weights to enable simple control and offline decomposition. Furthermore, the control overhead is reduced with simple weight information encoding. The multiplication after decomposition is transformed to simple addition and shift with a modified Booth's algorithm to completely remove multiplications. The experimental results show that this approach can reduce 32.3% of the area and 36.4% of the power consumption for the PE implementation compared to the previous design.

Oral S08

Advances in Modern EDA Methods and Technologies

Date/Time 8/4 (四) 11:00-12:00

Chair(s) 陳盈如教授 / 國立成功大學電機工程學系
鄭維凱教授 / 中原大學資訊工程學系

S08.1 11:00 — 11:12

Keeping Deep Lithography Simulators Updated: Graph-Based Smart-Sampling Schemes for Active Learning

Hao-Chiang Shao[1], Kuo-shiuan Chen[2], Hsing-Lei Ping[2], Weng-Tai Su[2], Chia-Wen Lin[2], Shao-Yun Fang[3], Pin-Yian Tsai[4], and Yan-Hsiu Liu[4]
[1]Dept. Statistics and Information Science, Fu Jen Catholic University
[2]Dept. Electrical Engineering, National Tsing Hua University
[3]Dept. Electrical Engineering, National Taiwan University of Science and Technology
[4]United Microelectronics Corporation

Learning-based pre-simulation (i.e., layout-to-fabrication) models have been proposed to predict the fabrication-induced shape deformation from an IC layout to its fabricated circuit. Such models are usually driven by pairwise learning, involving a training set of layout patterns and their reference shape images after fabrication. However, it is expensive and time-consuming to collect the reference shape images of all layout clips for model training and updating. To address the problem, we propose two active-learning strategies for sampling novel layouts beneficial to the model update process. Given a set of predefined newly-desgined novel layout patterns, the proposed methods aim to sample a reduced amount of representative layouts most worthy to be fabricated for acquiring their ground-truth circuit shapes. Experiments demonstrate our active-learning strategies' ability in selecting representative novel layouts for keeping a learning-based pre-simulation model updated.

S08.2 11:12 — 11:24

TCG-based Warpage-aware Floorplanning for Heterogeneous Integration

Yang Hsu[1], Min-Hsuan Chung[1], and Yao-Wen Chang[1, 2]
[1]Graduate Institute of Electronics Engineering, National Taiwan University
[2]Department of Electrical Engineering, National Taiwan University

In modern heterogeneous integration technologies, chips with different processes and functionality are integrated into a package with high interconnection density and large I/O counts. Integrating multiple chips into a package may suffer from severe warpage problems caused by the mismatch in coefficients of thermal expansion between different manufacturing materials, leading to deformation and malfunction in the manufactured package. To our best knowledge, the industry is eager to find a solution for the warpage optimization. In this paper, we propose the first warpage-aware floorplanning algorithm for heterogeneous integration. We first present an efficient qualitative warpage model for a multi-chip package structure based on Suhir's solution, more suitable for optimization than the time-consuming

finite element analysis. We then propose three perturbations for the transitive closure graph floorplan representation to optimize the warpage more directly and can thus speed up the simulated annealing process. Finally, we develop a force-directed detailed floorplanning algorithm to further refine the results by utilizing the dead spaces. Experimental results demonstrate the effectiveness of our warpage model and algorithm.

S08.3 11:24 — 11:36

Thermal-Aware Chiplet Placement on Interposer-Based 3D ICs via Thermal Weighted Nets and Refinement

Hong-Wen Chiou, Yu-Teng Chang, Ting-Yu Cheng, and Yu-Min Lee
Institute of Communications Engineering, National Yang Ming Chiao Tung University

Up to now, researches for the chiplet placement of interposer-based 3D ICs focus on the minimization of wire lengths. This work aims to simultaneously consider the issues of wirelength and thermal effect in chiplet placement. We introduce the concept of thermal weighted net that is applied to the advanced Branch-and-Bound method and develop a refinement method to improve the placement result for reducing the maximum temperature. Experimental results show that they can satisfy the temperature constraint at only sacrificing most of 5.88% total wirelength.

S08.4 11:36 — 11:48

Improving Pin Accessibility of Standard Cells under Power/Ground Stripes

Pei-Sheng Lu and Rung-Bin Lin
Yuan Ze University

This paper presents an approach to improving pin accessibility of standard cells under power/ground (P/G) stripes. Two strategies are used. First, hard-to-access cells located under M3 P/G stripes are swapped with easily accessible cells in close proximity. Second, filler cells are added at some places under P/G stripes to push hard-to-access cells away from P/G stripes. Experimental results show that cell swapping can reduce DRC violations by 86% while averagely maintaining timing performance. Filler insertion can reduce DRC violations by 80% but less effective in maintaining timing performance.

S08.5 🕓 **11:48 — 12:00**

Timing-Critical Path Analysis in Circuit Designs Considering Aging with Signal Probability

Jiun-Cheng Tsai, Aaron C.-W. Liang, Shang-Ming Liu, and Charles H.-P. Wen
ECE Department, National Yang Ming Chiao Tung University

Aging is an important determinant for the reliability of circuit designs. and has been addressed by a number of protection techniques based on static timing analysis (STA). The timing reported by STA, however, is often too optimistic without considering the functional behavior of the circuit. Furthermore, signal probability has also been found to be a significant factor in the aging effect. As such, we present in this paper a timing-critical path analysis that takes function and aging into account as well as signal probability. Functional timing analysis (FTA) eliminates the false paths and generates more accurate timing. Furthermore, machine learning can be used to build models for predicting the timing of each cell for various aging lifetimes and signal probabilities. Experimental results indicate that there can be a difference of up to 6% on path delay between STA and FTA. The path ranks also differ for most of the benchmark circuits after considering aging with signal probability, resulting in the delay differences of up to 7%. In conclusion, it is necessary to consider function, aging, and signal probability simultaneously when analyzing timing-critical paths in a circuit design.

Oral S09

Data Converters and Clock Generator

Date/Time 8/4 (四) 11:00-12:00

Chair(s) 彭朋瑞教授 / 國立清華大學電機工程學系
陳佳宏教授 / 國立陽明交通大學電機工程學系

S09.1 11:00 – 11:12

A 400KS/s 62.6pJ/c.-s 10-bits Column-Parallel Hybrid ADC Using Self-Adaptive Charge-Injection Cell and Single-Slope Conversion

Min Ruei Wu and Chih Cheng Hsieh
Department of Electrical Engineering, National Tsing Hua University

This work presents a column-parallel hybrid ADC using self-adaptive (SA) charge-injection (CI) cell and single-slope (SS) conversion for CMOS image sensor (CIS) applications. The proposed scheme solves the speed bottleneck in conventional SS ADC without the need of high frequency clock. To satisfy the column-to-column matching requirement in CIS readout, we proposed a SA ci-cell that achieves high linearity performance without calibration. The prototyped 10-bits column-parallel ADC is fabricated in 40nm 1P9M CMOS and achieves a sampling frequency of 400KS/s with only 40MHz counting clock with an energy efficiency FoMw of 62.6pJ/c.-s.

S09.2 11:12 – 11:24

A 77.16 dB SNDR 10MHz Bandwidth Continuous Time ΔΣ ADC with VCO-Based Quantizer

Guan-Yu Chen and Zhen-Jie Hong
Department of Electronic Engineering, Feng Chia University

This paper reports a continuous-time (CT) third-order delta-sigma (ΔΣ) analog-to-digital converter (ADC) in a 0.18 μm CMOS process for ultrasound testing over a 10 MHz signal bandwidth with a 1GHz sampling rate (Fs). There are two critical technologies are implemented to lower the power dissipation of this ADC. First, the conventional quantizer based on the flash ADC is replaced with a voltage-controlled oscillator (VCO)-based quantizer which works like an integrator providing a first-order noise shaping. This additional noise shaping help to save an integrator and lower the power. Another reason to choose a VCO-based quantizer is due to its implicit clocked averaging (CLA) can save additional dynamic element matching (DEM) circuits to reduce the mismatch of a digital-to-analog converter (DAC). Second, with such high Fs (1GHz), the OPA is adopted with active feedforward compensation to improve the unit gain bandwidth (UGB). It is an efficient way to enhance the UGB and does not cost too much power consumption simultaneously.

The prototype ADC achieves a 79.9 dB signal-to-noise ratio (SNR), 77.16 dB signal-to-noise and distortion ratio (SNDR), and 79.3 dB dynamic range (DR), respectively, over a 10 MHz signal bandwidth with a 1GHz sampling rate. The whole ADC total consumes 30.7 mW of power from a 1.8 V power supply.

S09.3 11:24 — 11:36

FPGA Multichannel Digital to Time Converter based on Multidimensional Delay

Tai-Sen Lin, Jane Chen, and Poki Chen
Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology

This paper presents a precise multichannel digital-to-time (DTC) converter, or digital pulse generator, based on multidimensional delay array. It is also proposed to further extend FPGA applications into the analog domain. Multidimensional delay array is composed of a delay matrix and a phase-locked loop (PLL) with phase shifting so that output phases are uniformly distributed within the reference clock period. For the proof of concept and performance evaluation, the proposed multichannel DTC based on multidimensional delay array is implemented and the achieved fine stage resolution is as high as 2 ps. The differential nonlinearity (DNL) and integral nonlinearity (INL) of this architecture is measured to be -4.7 ~ 6.7 LSB and -4.8 ~ 5.8 LSB respectively.

S09.4 11:36 — 11:48

A 12-bit 100KS/s SAR ADC with Split Capacitor Array DAC and Offset Calibration

Yun-Yen Huang and Chung-Chih Hung
Department of Electrical and Computer Engineering, National Yang Ming Chiao Tung University

This paper presents a 12-bit 100KS/s successive approximation register (SAR) analog-to-digital converter (ADC) in TSMC 0.18-μm process. To reduce the switching energy and save the total capacitance, a hybrid switching procedure with split capacitor array DAC is applied. The hybrid switching procedure combines two concepts of one side switching instead (OSSI) and higher bits switching instead (HBSI). Besides, a dynamic comparator with adjustable capacitor calibration to achieve low offset is used. With 1.8V supply voltage and 100K sampling rate, simulation results achieve 73.15dB signal-to-noise-and-distortion-ratios (SNDR) which lead to 11.86 effective number of bits (ENOB) at 49.97k input frequency. Its power consumption is 5.04-μW and figure-of-merit (FOM) is 13.5 fJ/conversion-step.

S09.5 11:48 — 12:00

A 197-ppm/°C 2.5-GHz Low-Voltage Oscillator Using Frequency-Locked Loop in 180-nm CMOS

Tsung-Ying Chen and Ching-Yuan Yang
National Chung Hsing University

In the paper, we present a low-voltage oscillator using the frequency-locked loop (FLL) circuit in 0.18μm CMOS process. With employing the FLL, The output frequency depends on passive components and divider ratio to reach low temperature variation and frequency programmable ability. Operation at 2.5 GHz, the temperature stability for the oscillator is about 197.76 ppm/°C from 0°C to 100°C and line sensitivity is 5.91 %/V from 0.9V to 1.1V.

Oral S10

Advanced Baseband Communication and Cryptosystem Design

Date/Time 8/4 (四) 13:30-14:30

Chair(s) 施信毓教授 / 國立中山大學電機工程學系
吳志峰教授 / 長庚大學電子工程學系

S10.1 13:30 — 13:42

Reconfigurable Filtered-OFDM Baseband Processor for Mobile Communications

Huan-Lun Tso[1], Chun-Yu Chen[1], Ching Hsu[2], Pei-Yun Tsai[3], and Yuan-Hao Huang[1, 2]
[1]Institute of Communications Engineering, National Tsing Hua University
[2]Department of Electrical Engineering, National Tsing Hua University
[3]Department of Electrical Engineering, National Central Univerisity[1, 2]

Orthogonal Frequency Division Multiplex-ing (OFDM) has been widely used in wireless communi-cation systems. Filtered-OFDM is one of the potential improved modulation waveform for the 5G/B5G communication systems. Filtered-OFDM is a per-subband-filtering multi-carrier modulation, in which whole bandwidth is divided into several subband each carrying individual data information. The per-subband-filtering approach can reduce out-of-band emission and then avoid adjacent-channel interference (ACI). This work designed and implemented a reconfigurable filtered-OFDM baseband processor that supports 14 different frame structures in terms of bandwidth, FFT size and subcarrier spacing. The proposed processor was designed and synthesized by using TSMC 40 nm CMOS process technology. The proposed filtered-OFDM processor chip achieves maximum throughput of 840 Mb/s at clock speed of 140.9 MHz with chip area of 12 mm^2.

S10.2 13:42 — 13:54

Beam Tracking Based on Deep Reinforcement Learning for Hybrid Beamforming MIMO Communications Systems

Po-Chuan Huang and Pei-Yun Tsai
Department of Electrical Engineering, National Central University

In this paper, beam tracking based on deep reinforcement learning (DRL) in hybrid beamforming multiple-input multiple-output (MIMO) systems is designed and verified. With channel state information (CSI), the deep deterministic policy gradient (DDPG) of DRL is employed to compute the analog precoder, called aBFNet, which achieves good channel capacity. In the last layers of the actor network of the aBFNet, output normalization is performed and is considered during back propagation so that the generated outputs satisfying the magnitude constraint can be applied to the phase shifters directly. In addition, time-varying channels are taken into consideration and can be tracked by the aBFNet. Simulation results demonstrate the feasibility of aBFNet and its good performance compared to the conventional hybrid beamforming algorithm.

S10.3 13:54 — 14:06

High-Performance Radix-4 Montgomery Modular Multiplier with Carry-Propagation Free Format Conversion

Yen-Jui Chen, Chun-Yi Wang, and Shiann-Rong Kuang
Department of Computer Science and Engineering, National Sun Yat-sen University

Carry-save arithmetic and higher radix are frequently adopted to enhance the performance of Montgomery modular multiplication. In this paper, a carry-save radix-4 Montgomery modular multiplier is proposed to speed up the modular multiplication while maintaining a short critical path and low hardware complexity. The proposed multiplier takes binary numbers as the input/output operands and adopts Booth encoding in which a triple of modulus is replaced with a negative of modulus to reduce the register usage. Moreover, the carry-save adder (CSA) is used to not only perform the intermediate addition operations but also convert the result of modular multiplication from carry-save representation to binary representation to avoid the long carry propagation. The critical path of proposed multiplier is further shortened by simplifying quotient calculation, rapidly generating two's complement form of negative multiples, and pipelining the execution of Booth encoding and selection. In addition, an efficient detect circuit is developed to suspend the unnecessary carry-save addition operations so that the extra clock cycles for format conversion can be significant reduced. Experimental results show that the proposed multiplier can achieve considerable area-time product improvement when compared with previous designs.

S10.4 14:06 — 14:18

A new natural plant self-awareness Phytosensing sensor based on Electrophysiology circuit design

Yu-Hsien Yao and Chi-Chia Sun,
Department of Electrical Engineering, National Formosa University

In recent years, plant electrophysiology has become a feasible invasive tool for describing the thinking of plant properties of many systems. In this paper, we proposed a new electrophysiological sensor presented as an autonomous device, which can transfer plant electrical potential data into plant self-awareness information, it will explore how plants can end up in plant factories or smart warming and artificial intelligence technology to achieve the concept of making plants "think themselves". Plant sensing can measure certain aspects of the internal state of natural plants. The proposed sensor circuit design is based on ECG and EEG measurement methods in electrophysiology respectively. The major contribution of this paper is to use electrophysiological measure method to collect data that the plants themselves behavior can be further analyzed through machine learning. The experimental result shows that with each sensing condition with the external stimuli through the proposed plant electrophysiology sensor module (light, heat, humidity, CO2, etc.) the potential changes can be observed.

S10.5 🕓 **14:18 — 14:30**

VLSI Implementation of RSA cryptography

Chia-Yu Liu[1] and Yuan-Ho Chen[1, 2]
[1]Dept. of Electronics Engineering, Chang Gung University
[2]Dept. of Radiation Oncology, Chang Gung Memorial Hospital

RSA cryptosystem is a common asymmetric cryptography. Due to the complexity of the operation, it is hoped that the hardware can effectively improve its operation speed.The architecture proposed in this paper includes interleaved modular multiplication, modular exponentiation, Baillie-PSW primality test, and extended Euclidean algorithm. According to the architecture proposed in this paper, it is finally implemented on the chip of the TSMC 90nm CMOS technology. The operating frequency was 28.57 MHz, and the average operation time is about 0.34 ms. Finally, the architecture can be scaled directly to 1, 024 bits without further modification.

Oral S11

Intelligent Digital System Designs and Biomedical Applications

Date/Time 8/4 (四) 13:30-14:30

Chair(s) 陳坤志教授 / 國立中山大學資訊工程學系
賴伯承教授 / 國立陽明交通大學電機工程學系

S11.1 13:30 — 13:42

A High-Speed Extreme Learning Machine Engine for Age-related Macular Degeneration Detection

Jie-Yi Ji[1], Cheng-Hung Lin[1,2], Cheng-Kai Lu[3], Jia-Kang Wang[1,4], and Tzu-Lun Huang[1,4]
[1]Department of Electrical Engineering, Yuan Ze University
[2]Biomedical Engineering Research Center, Yuan Ze University
[3] Department of Electrical Engineering, National Taiwan Normal University
[4]Department of Ophthalmology, Far Eastern Memorial Hospital

In this paper, a system on chip (SOC) design for the extreme learning machine (ELM) model is proposed. By training the model with optical coherence tomography (OCT) images, a classifier for detecting age-related macular degeneration (AMD) has been constructed. To simplify the hardware architecture, the general normalization is neglected, and the activation function is also replaced from sigmoid to ReLU. By reducing the calculation complexity, the area could achieve 0.36 mm2 and the accuracy reaches 89.04%. At the same time, the sensitivity and specificity could achieve 98.25% and 79.82%.

S11.2 13:42 — 13:54

Compression-Aware Projection with Greedy Dimension Reduction for Convolutional Neural Network Activations

Yu-Shan Tai, Chieh-Fang Teng, Cheng-Yang Chang, and An-Yeu (Andy) Wu
Graduate Institute of Electrical Engineering, National Taiwan University

Convolutional neural networks (CNNs) achieve remarkable performance in a wide range of fields. However, intensive memory access of activations introduces considerable energy consumption, impeding deployment of CNNs on resource-constrained edge devices. Existing works in activation compression propose to transform feature maps for higher compressibility, thus enabling dimension reduction. Nevertheless, in the case of aggressive dimension reduction, these methods lead to severe accuracy drop. To improve the trade-off between classification accuracy and compression ratio, we propose a compression-aware projection system, which employs a learnable projection to compensate for the reconstruction loss. In addition, a greedy selection metric is introduced to optimize the layer-wise compression ratio allocation by considering both accuracy and #bits reduction simultaneously. Our test results show that the proposed methods effectively reduce 2.91×~5.97× memory access with negligible accuracy drop on MobileNetV2/ResNet18/VGG16.

S11.3 13:54 — 14:06

A Low-Power Two-Lens Wireless Panoramic Micro-Endoscopy Implemented Using Voltage-Current Adjuster and 3D-PCB Stacking Technology

Sheng-Wei Hsu, Ching-Hwa Cheng, and Don-Gey Liu
Dept. of Electronic, Feng Chia University

The main purpose of this paper is to develop a low-power design technology for a wireless two-lense panoramic micro-endoscopy. The voltage current adjuster (VCA) replaces the voltage-adjust method and does not increase the additional silicon cost without using voltage converters. A built-in voltage measurement mechanism provides that the voltage level can be automatically adjusted. The system achieves a 32~68% power reduction for video decoders using VCAs. The proposed technique can reduce system power consumption without performance degradation. Scalable system functions are successfully validated by a 3D-PCB stacking technique. The 3D stacking system comparison with the conventional flatten design, with good performance, less power-consumption and small volume size.

S11.4 14:06 — 14:18

An Experience Sharing: A Panoramic-Vision Lesion-Finding Low-Power Wireless Endoscopic System Design and Implementation

FCU-BioElectronic research group, Ching-Hwa Cheng, and Don-Gey Liu
Department of Electronic, Feng Chia University

The limited field-of-image and lesion-position-loss of the endoscope are often the most problematic issues faced by junior surgeons. A battery-operation wireless four-lens panoramic-endoscopy is proposed. This design provides a panoramic vision and in-time lesion-guiding information during a surgical operation. This work contains a combination of several image processing techniques of image-stitching, viewing-synthesis, and lesion-guiding to a view. Lesion-guiding provides global positioning information and tracks the predefined lesion position during surgery. On the low-power hardware design system, the image processing encoder and decoder chips were designed by the voltage-domain, which were developed by the clustered voltage-domain technique. The technique of the designed chip is separated by two/four voltage-domain, and the voltage-scaling technique is used for each domain. The high-voltage domain maintains the chip performance and the low-voltage domain reduces the power consumption. This effective technique decreases power consumption without reducing the performance of the chip. The entire system is integrated by a personal computer, an embedded system, and image encoder, decoder chips. By applying the Multi-Vdd technique, the multiple-Vdd encoder and decoder chips can be quickly redesigned based on the power, delay-time, and gate-count optimization requirements. The power consumption of the encoder and decoder chip can be effectively reduced to 50% and 24%, respectively. The performance loss can be maintained within 5% of both designs. A wireless panoramic endoscopic system is successfully validated and demonstrated by integrated encoder and decoder chips. The whole system has been successfully validated by in vivo experiments with animals. The experimental results show that the proposed system can enhance the side-by-side image size to 155%. By our search, there is no similar work that can be as a comparison.

S11.5 🕓 **14:18 — 14:30**

Using Phase Portrait of Electrocardiography Signals to Analyze Atrial Fibrillation

Chin-Cheng Kuo[1], Shu-Yen Lin[1], and Yu-Wei Chiu[2, 3]
[1]Department of Electrical Engineering, Yuan Ze University
[2]Department of Computer Science and Engineering, Yuan Ze University
[3]Cardiology Department, Far Eastern Memorial Hospital

To analyze arrhythmia symptoms, such as Atrial fibrillation (AF), this study uses phase portrait as an analysis method for Electrocardiography (ECG) long-term lead. The role of AF detection method in Electrocardiography (ECG) long-term lead is also proposed. Our aim is to reconstruct the phase portrait for analyzing and extract the features of AF through the time delay of the ECG signal. This algorithm cooperates with the dispersion of the RR interval and the trajectory of the phase portrait for each P wave. The distance parameters from the trajectory are calculated, and the rule to identify the AF can be provided. The classification accuracy rate with dispersion (D) as the parameter is 70%. After adding the distance of parameters, the accuracy can be improved.

Oral S12

Power Management & Readout IC

Date/Time 8/4 (四) 13:30-14:30

Chair(s) 陳柏宏教授 / 國立陽明交通大學電機工程學系
陳景然教授 / 國立臺灣大學電機工程學系

S12.1 13:30 – 13:42

A V3-Controlled Buck Converter with Pseudo-Current Sensing Technology

Jing-Hao Yang[1], Yuh-Shyan Hwang[1],and Dong-Shiuh Wu[2]
[1]Department of Electronic Engineering, National Taipei University of Technology
[2]Department of Electronic Engineering, Lunghwa University of Science and Technology

This paper presents a V3-Controlled Buck Converter with Pseudo-Current Sensing Technology. The proposed converter achieves fast transient response and a wide output voltage range. The proposed buck converter is fabricated with TSMC 0.18μm 1P6M CMOS technology, and the chip area is 1.052 mm× 0.942 mm. The measured results show that the output voltage is 1.6V, the load current changes from 50mA to 500mA, and the transient response from 50mA to 500mA is 2μs and 2μs, respectively. When the load current is 400mA, the maximum power efficiency is 93.2%.

S12.2 13:42 – 13:54

An Adaptive 2nd-Order Delta-Sigma-Modulation Buck Converter with Transient-Accelerated-Circuits

Jian-Yuan Chen[1], Yuh-Shyan Hwang[1], and *Dong-Shiuh Wu[2]*
[1]Department of Electronic Engineering, National Taipei University of Technology
[2]Department of Electronic Engineering, Lunghwa University of Science and Technology

This paper presented an adaptive 2nd-order continuous-time delta-sigma-modulation (CT-DSM) current-mode controlled buck converter with transient-accelerated-circuits that features an integral loop filter with a superiority of oversampling and noise shaping for effective spurious-noise reduction. In addition, the Transient-Accelerated-Circuits (TAC) use dynamic-slope generator and current-sensor to speed up the transient response. Then, the proposed buck converter preserves a low spurious noise and fast transient response. The proposed buck converter is fabricated in TSMC 0.18um 1P6M CMOS processes with a chip area of 1.19mm × 1.19mm. The measurement results show that the transient recovery times are 2.7μs and 3.1μs, and the undershoot and overshoot voltages are 22mV and 24mV, when the load current changes from 50mA to 500mA and from 500mA to 50mA. The output spectrum with signal to noise ratio is 77dB was obtained across all sampling frequencies. The peak power efficiency is 94.88%, when the load current is 400mA and output voltage is 2V.

S12.3 13:54 — 14:06

An Adaptive On-Time Buck Converter with New Integral Current-Sensing and Dynamic Slope Compensation Techniques

Ming-Dao Luo and Jiann-Jong Chen
Department of Electronic Engineering, National Taipei University of Technology

An adaptive-on-time buck converter with new integral current sensing and dynamic-slope compensation techniques is presented in this paper. First, the adaptive on-time controlled scheme is adopted, which is based on the techniques of the on-time controlled by input voltage feedforward and output voltage feedback, by using this scheme, a relatively constant switching frequency can be obtained. Second, the proposed converter uses the integral current-sensing circuit with the linearized input-stage OTA to sense the inductor current, which increases the bandwidth of the converter's closed-loop gain and will speed up the converter's transient response. Third, the converter uses the dynamic-slope compensation techniques to eliminate subharmonic oscillation when the duty cycle is greater than 50%. The proposed converter is designed and fabricated in TSMC 0.18μm 1P6M CMOS process, and the chip area is 0.998mm × 0.949mm. When the load current changes from 50mA to 500mA and 500mA to 50mA, the transient responses are 2.6us and 2.5us, respectively, and the maximum efficiency is 93.2%.

Keywords: Buck converter, Adaptive-On-Time (AOT), Dynamic-Slope Compensation, Integral Current-Sensing

S12.4 14:06 — 14:18

A Cross-Correlation-Based Time-of-Flight Design for Chaos Lidar Systems

Yi-Cheng Lin[1], Ping-Hsuan Hsieh[1], Jian-Lun Hong[1], Yu-Hsiang Lai[2], Jun-Da Chen[2], Fan-Yi Lin[2], Yuan-Hao Huang[1], and Po-Chiun Huang[1]
[1]Department of Electrical Engineering, National Tsing Hua University
[2]Institute of Photonics Technologies, National Tsing Hua University

This work presents a time-of-flight circuit design for chaos lidar systems. It is implemented by a mixed-signal structure, with 1-bit digitization and a multiplier-and-integrator array to calculate the cross-correlation between the two signals to estimate the time difference and the corresponding distance. With an off-chip tunable low-pass filter, the ranging accuracy and detection distance can be adjusted more effectively and flexibly.

This work is implemented in CMOS 90-nm process. At the sampling rate of 5 GS/s, the measurement results demonstrate that the maximum ranging error is 3.38 cm and 1-sigma is 1.97 cm in the ranging range of 7.4m. With a pulse width of 100ns and a repetition interval time of 50 μs, this work consumes 113 mW under 1-V supply.

S12.5 🕓 **14:18 — 14:30**

A VRO-based TDC for a light sensor application

Jen-Chieh Liu[1], Jian-Sheng Li[1], Chi-Hua Chen[1], and Yu-Lung Lo[2]
[1]Department of Electrical Engineering, National United University
[2]Department of Electrical Engineering, National Kaohsiung Normal University

A vernier ring oscillator (VRO) based time to digital converter (VRO-based TDC) proposes for the timing resolutions of the coarse-tuning stage (CTS) and the fine-tuning stage (FTS) have a proportional relationship under the process, voltage and temperature (PVT) variations. In the input range, the CTS can be flexibly to extend the bit number for a wide input range. The timing resolutions of CTS and FTS are defined by the rise time and the fall time. Therefore, the timing ratio between CTS and FTS is a constant under the PVT variations. This 14-bit TDC was fabricated in a 0.18 μm standard CMOS process and the core area of 62×199 μm2. The measured timing resolution of the proposed VRO-based TDC was 125 ps and the input range was from 10 ns to 200 ns. The DNL and INL were less than ±1 LSB and ±1.64 LSB, respectively. The proposed VRO-based TDC is also integrated with a light sensor for internet of things (IoT) applications.

Oral S13

Digital Signal Processing ICs, Image/Visual & Multimedia Systems

Date/Time 8/4 (四) 14:30-15:30

Chair(s) 李佩君教授 / 國立臺灣科技大學電子工程學系
李宇軒教授 / 元智大學電機工程學系

S13.1 14:30 – 14:42

Modularized and Extensible Convolution Cell Design and Chip Implementation for AI Accelerator

Wei-Hsuan Chang and Chung-Bin Wu
National Chung-Hsing University

This paper proposes a modularized and extensible Convolution Cell for neural network accelerators. Each unit of convolution cell is composed of 72 PEs and can be arbitrarily expanded to above 72xN up to 2160 PEs. The Convolution Cell is designed to be quickly reconfigured into application requirements under certain memory size and hardware specifications. In addition, the Convolution Cell can support two convolution operation modes, 1x1 and 3x3, can be adapted to the network architecture of the Yolo series. In FPGA implementation, one Convolution Cell are configured to reach 7.69GOPS/W and two Convolution Cell to reach 15.22GOPS/W. In 40nm Chip implementation, according to analyzing area into million gate equivalent (MGE) and energy efficiency, one Convolution Cell can reach 66.5 GOPS/MGE and 316.4 GOPS/W.

S13.2 14:42 – 14:54

Architecture Implementation on FPGA for CNN with Gabor Feature Extraction

Yu-Wen Wang
Department of Electrical Engineering, National Cheng Kung University

This paper implements a hardware design that can calculate the 2-dimensional convolution with specific Gabor filters onto a Field Programmable Gate Array (FPGA). As a popular algorithm in the field of computer vision today, Convolutional Neural Network (CNN) requires a huge amount of computation and therefore requires hardware and special algorithms to accelerate. For reducing the number of additions and multiplications, we use the Eigen-transformation approach to transform the 16 Gabor filters into the 16 transformed filters with a high degree of symmetry and then pre-add the input pixels corresponding to the position of the repeated coefficients. We analyze the four models of processing units for the transformed filter bank proposed by the previous work in our lab and use the Xilinx XUPV5-LX110T Evaluation Platform for prototyping. Finally, we use the Xilinx Chipscope as an integrated logic analyzer for verification.

S13.3 🕓 **14:54 — 15:06**

A Parallel-Process Auto-Clear Time-to-Digital Converter Chip Integrated with SPAD Sensors for a LiDAR System

Yi-Chen Hsieh, Ching-Hwa Cheng, and Don-Gey Liu
Department of Electronic Engineering, Feng Chia University

The proposed design proposes to perform parallel computing processing on the two groups of TDCs, to improve the calculation speed of the overall TDC, reduce its death time, and maintain high resolution capability at the same time. And integrating SPAD and TDC on the same chip can speed up the TDC input and SPAD output transmission to reduce the delay time of signal transmission. The proposed design integrates the front-end SPAD optical sensor and TDC circuit, and adopts the TSMC-T18HVG2 process. The reference clock frequency is 100MHz. TDC shortest dead time is 20ns, resolution is 100ps, measurement range can reach 337.5 meters, and the chip area is 2.5 mm2. Two TDCs can operate at the same time; the speed of data transmission will be greatly improved to bringing a high-performance LiDAR system.

Oral S14

Emerging Algorithms and Future Computing Technologies

Date/Time 8/4 (四) 14:30-15:30

Chair(s) 林忠緯教授 / 國立臺灣大學資訊工程學系
陳元賀教授 / 長庚大學電子工程學系

S14.1 14:30 – 14:42

Partial Equivalence Checking of Quantum Circuits

Tian-Fu Chen[1], Jie-Hong R. Jiang[1], and Min-Hsiu Hsieh[2]
[1]Department of Electrical Engineering, National Taiwan University
[2]Quantum Computing Research Center, Hon Hai Research Institute

Equivalence checking of quantum circuits is an essential element in quantum program compilation, in which a quantum program can be synthesized into different quantum circuits that may vary in the number of qubits, initialization requirements, and output states. Verifying the equivalences among the implementation variants requires proper generality. Although different notions of quantum circuit equivalence have been defined, prior methods cannot check observational equivalence with respect to measurement between two quantum circuits whose qubits are partially initialized, which is referred to as partial equivalence. In this work, we devise algorithms for partial equivalence checking. Experiments demonstrate the generality of our method in checking quantum circuits whose partial equivalence cannot be verified by prior approaches and the effectiveness of our method in subsuming prior more restricted equivalence checking. Our result may unleash the optimization power of quantum program compilation to take more aggressive steps.

S14.2 14:42 – 14:54

Bearing Fault Diagnosis by Using Exponential Power Entropy and a Decision Threshold for Artificial Neural Network

Pavan Kumar MP and Kun-Chih (Jimmy) Chen
Department of Computer Science and Engineering, National Sun Yat-sen University

Rotating machinery is used in a variety of industries including petroleum, automotive and food processing, etc. which uses bearings to reduce the friction between the moving parts. Besides, bearings are prone to wear due to a variety of conditions such as speed variations, heavy loads, and lengthy periods of operation. Therefore, real-time monitoring and diagnosis of bearing faults will improve safety, avoiding unplanned downtime and lowering the cost. Condition-based maintenance (CbM) is usually employed to observe the faults that arises in a healthy condition of the machine, running by gathering sensing vibration signals. However, the massive sensing data increase the challenge to analyze the fault signal will result in high computing time. To solve this problem, we propose an Exponential Power Entropy (EPE)-based feature extraction and selection approach to minimize the computing data and used a feed-forward neural network for fault classification with a proposed decision threshold score to improve the prediction accuracy. Compared with the conventional approaches, the proposed method not only reduce the computing time by 64.8% but also increase the accuracy of fault diagnosis up to 99.2%.

S14.3 14:54 — 15:06

Edge AI Implementation for Heterogeneous Images Semantic Segmentation

Ming-Hwa Sheu, De-Yu Chen, S. M. Salahuddin Morsalin, Szu-Hong Wang, and Keng-Wei Lin
Department of Electronic Engineering, National Yunlin University of Science and Technology

The deep convolutional neural network-based semantic segmentation needs large-scale computations and annotations for data training to reach real-time inference speeds. The heterogeneous image semantic segmentation method extracts the features of visible and thermal images separately. We designed an efficient architecture with the multi-hybrid-autoencoder and decoder for Faster Heterogeneous Image (FHI) Semantic Segmentation. The proposed architecture has fewer layers resulting in lower parameters, higher inference speed. The specialty of this architecture is the discrete autonomous feature extraction framework for RGB image and Thermal (T) image inputs with individual convolutional layers. Later, we combined the 4-channels (RGBT) convolution features to reduce computational complexity and robust the model performances. The proposed FHI-Unet semantic segmentation model experimented on NVIDIA Xavier NX edge AI platforms with standard accuracy under the real-time inference requirement. The proposed FHI-Unet model has fastest real-time inference of 83.39 frames per second on edge AI implementation on the Multi-spectral Semantic Segmentation Dataset compared with the existing works.

S14.4 15:06 — 15:18

Polar Code Belief Propagation Decoder with Stage Stopping Scheme under Round Trip Scheduling

Pei-Hsuan Chen[1] and Cheng-Hung Lin[1, 2]
[1]Department of Electrical Engineering, Yuan Ze University
[2]Biomedical Engineering Research Center, Yuan Ze University

Capacity-achieving polar codes have grown attention in recent years, and the belief propagation (BP) algorithm is one way to decode them. With the help of Min-sum approximation and G-Matrix early termination schemes, it has been proven that complexity and computation can be further reduced without significant performance loss. Although G-Matrix early termination is a strong stopping criterion, there are still some redundant cycles that need to be reduced. In this paper, a stage-stopping BP algorithm with round-trip scheduling is applied to reduce the redundant computation. Different thresholds for different SNR regions are simulated to seek the proper combination. The results show that threshold=3 is suitable for the SNR region in{1,1.5,2,2.5,3} (dB), and threshold=3/3.5 can be selected for lower computation or better performance in {3.5} (dB). As for the high SNR region in {4,4.5,5} (dB), threshold=5.5 is suitable for better performance.

S14.5 🕓 **15:18 — 15:30**

Toward Optimal Topology Generation and Efficient Radius Selection for Parallelism-Aware Wavelength-Routed Optical Networks-on-Chip Design

Kuan-Cheng Chen[1], Yan-Lin Chen[1], Yu-Sheng Lu[1], and Yao-Wen Chang[1, 2]
[1]Graduate Institute of Electronics Engineering, National Taiwan University
[2]Department of Electrical Engineering, National Taiwan University

The wavelength-routed optical network-on-chip (WRONoC) emerges as a promising solution for multi-core system communication, providing high-bandwidth, high-speed, low-power, and low-latency transmission. However, as the number of cores in a WRONoC increases, some wavelength-routed optical network-on-chip (WRONoC) topologies could be infeasible with bandwidth and crosstalk constraints if bit-level parallelism is not considered during topology generation. In previous work, the parallelism was optimized only for the radius selection of microring resonators but not for topology generation. Further, existing parallelism-aware WRONoC designs are too time-consuming to handle the current design sizes efficiently. To remedy these drawbacks, we present a parallelism-aware WRONoC design flow to optimize parallelism in topology generation and radius selection. The proposed design flow guarantees to generate a parallelism-optimal topology for full connectivity; and parallelism-optimal topology for customized connectivity if the netlist meets certain conditions. Compared with the state-of-the-art methods, experimental results show a 31.6% improvement in parallelism. Besides, the proposed radius selection method can significantly reduce runtime without performance loss.

Oral S15

Advanced Circuit Techniques for RF/mm-Wave Applications

Date/Time 8/4 (四) 14:30-15:30

Chair(s) 簡俊超教授 / 國立臺灣大學電機工程學系
張裕鑫教授 / 國立虎尾科技大學電子工程學系

S15.1 14:30 – 14:42

A 187 μW Multi-Mode Wake-Up Receiver Achieving –97 dBm Via a Balun LNA

Pin-Chen Yeh, Shih-En Chen, and Kuang-Wei Cheng
National Cheng Kung University

This paper presents an energy efficiency wake-up receiver (WuRx) with a balanced-to-unbalanced low noise amplifier (balun LNA) to achieve high sensitivity. Based on an injection-locked oscillator (ILO) and an envelope detector, the prototype WuRx provides the capabilities of reliable demodulation for on-off keying (OOK), binary frequency shift keying (BFSK), and differential binary phase-shift keying (DBPSK) demodulation schemes. A 187 μW 433 MHz multi-mode WuRx is implemented in a 0.18-μm CMOS process. For a data rate of 200 kbps and packet error rate (PER) of 10-1 with false alarm rate (FAR) < 10-3/s, the receiver achieves sensitivities of–97/–96/–94 dBm under the OOK/BFSK/DBPSK demodulation schemes, respectively.

S15.2 14:42 – 14:54

A Type-3 FMCW Radar Synthesizer with Wide Frequency Modulation Bandwidth

Cheng-Tang Chen, Yu-Hong Yang, and Tai-Cheng Lee
Graduate Institute of Electronics Engineering and Department of Electrical Engineering, National Taiwan University

This work presents a 5-GHz frequency-modulated continuous wave signal generator (FMCW). The presented circuit aims to generate a wider frequency modulation bandwidth at a low center frequency that grants a significant advantage in saving power. Utilizing a fractional-N phase-locked loop (PLL)-based synthesizer with a 50-MHz reference frequency as the FMCW generator, the synthesizer modulates the carrier frequency across a range of 700 MHz. A Type-3 architecture with a switchable polarity is embedded to improve the linearity around the chirp turning-around points (TAPs) and switching-band points (SBPs). Switching Band Control (SBC) circuit and Multi-varactor LC Voltage Control Oscillator (MV-VCO) are proposed as an efficient method to increase the modulation range so that the FMCW signal generator can achieve 21.43-cm resolution in distance detecting. Fabricated in a TSMC 40-nm 1P10M CMOS technology, the proposed generator consumes 3.3 mW power and occupies 0.497 mm2 die area. The measured root-mean-square (rms) frequency error of the generated triangle chirp over 1.15 ms period is 625 kHz.

S15.3 14:54 — 15:06

A 17-21 GHz Current-Folding Frequency Tripler With >36-dBc Harmonic Rejection and 190.8-dB FoM in 90-nm CMOS

Chun-Hung Lin and Ching-Yuan Yang
Department of Electrical Engineering, National Chung Hsing University

A frequency tripler (FT) using a current-folding technique to achieve inherently nonlinear operation is presented. A built-in VCO generates the fundamental signal, and the proposed current-folding stage converts the fundamental input into the triple-frequency output, which is injected into a bandpass stage for harmonic suppression. Fabricated in 90-nm CMOS technology, the measured FT features 36 to 43-dBc harmonic rejection from 17.5 to 21 GHz (18.2% FTR), while consuming 3.5 mW only from 1.2-V supply. The measured phase noise (PN) of the VCO and the FT are -112.5 and -102.8 dBc/Hz at 1-MHz offset, respectively. Furthermore, the achieved figure-of-merit (FoM) of the FT are -180.52 and -190.87 dB at 1-MHz and 10-MHz offset, respectively. To the best of our knowledge, the proposed FT achieves outstanding performance on both harmonic rejection and FoM among the state-of-the-art multipliers.

S15.4 15:06 — 15:18

Simultaneous Noise and Input Matching for W-Band LNA Using Parallel Coupled Transmission Lines

Pin-Yi Huang and Yen-Chung Chiang
Department of Electrical Engineering, National Chung Hsing University

This proposed low-noise amplifier (LNA) for W-band applications adopts four stages of common source topology. The implement of the chip was fabricated by using TSMC 90-nm CMOS GUTM 1P9M process. With the parallel-coupled transmission lines between the voltage bias of gate terminal and the degenerative source terminal, this structure is applied to optimize among the specifications of noise, input matching and gain. From the measurement results, the peak gain is 13.15 dB at 73 GHz; the minimum noise figure is 4.83 dB at 77.5 GHz; input P1dB is –7.3 dBm at 78 GHz, and the corresponding IIP3 is –1.1 dBm. The chip area is 0.730×0.588 mm2, and the total power consumption is 28.14 mW under 1.2-V supply.

S15.5 15:18 — 15:30

A 94 GHz Low-Power Down-Conversion Mixer With Negative-R Gain Enhancement

Wei-Chien Wang, Sheng-Wei Hsu, and Chien-Nan Kuo
Institute of Electronics, National Yang Ming Chiao Tung University

A W-band single-balanced mixer is designed and fabricated in 40 nm digital CMOS technology. The mixer down-converts the 94 GHz rf input to the 10 GHz IF output. The entire circuit consumes dc power of 6.6 mW with the supply voltage of 1 V, while the core circuit consumes only 2.8 mW. The measured peak conversion gain achieves 4.3 dB at 94 GHz. The input 1-dB gain compression point (iP1dB) is -13 dBm. The core circuit occupies a small area of only 0.22 mm2

Oral S16

新創技術

Date/Time 8/4 (四) 16:00-17:00

Chair(s) 黃元豪教授 / 國立清華大學電機工程學系

S16.1 16:00 – 16:15

Intelligent Implantable Neuromodulation Medical Device

Chin-Fong Chiu
A-Neuron Electronic Corporation

In recent decades, mixed-signal IC technology has been widely used in implantable medical devices for intelligent diagnosis and treatment. Such medical ICs provide the functions of sensing and stimulation with integrated memory storage, wireless communications and power management to achieve implantable medical application with small form factor and ultra low power. A-Neuron was established in 2016 and realizes the electronic medical products to treat nerve damage, disorders and diseases that cannot be cured by current medicine. Current developed intelligent neuromodulation medical devices are used in neurological diseases, including refractory epilepsy and retinal degenerative diseases. It also cooperated with BETRC (Biomedical Electronics Translational Research Center, NYCU) to develop a closed-loop deep brain stimulation system to treat patients with Parkinson's disease.

S16.2 16:15 – 16:30

Hardware Accelerator Design Challenges for AI Computing in Data Centers

Juinn-Dar Huang
Neuchips Inc.

This short talk presents design challenges engineers are facing during the implementation of state-of-the-art accelerators optimized for various AI computing applications in data centers. Hardware engineers have to meet a set of tough timing and power constraints while integrating hundreds of thousands processing elements, memory blocks, and other functional units together on a die in 7nm and beyond. Software engineers are required to build a sophisticated compilation flow that translates various high-level AI models into low-level hardware instructions that fully exploit the hardware computing resources to maximize the performance. System engineers carefully deal with a bunch of critical system-level issues such as board-level signal integrity, voltage domain and power management, cooling and thermal control so that those add-on accelerator cards can operate properly as expected within a computing server in data centers.

S16.3 🕓 16:30 — 16:45

See the Wonders: Micro Eye Tracking Technology on Edge Devices for the Metaverse

Shao-Yi Chien
Ganzin Technology, Inc. (Graduate Institute of Electronics Engineering, National Taiwan University)

Eye tracking is viewed as one of the enabling technologies for AR/VR/smart-glasses, which are the edge devices for the Metaverse. Ganzin Technology, a start-up spun-off from National Taiwan University, focuses on creating the next generation eye tracking modules. Based on the in-house AI eye tracking algorithm, IC design, and hardware-software-integration capabilities, Ganzin's Aurora eye tracking technology is the most easy-to-install solution on the market. Our vision is to unlock the potential of the eye as a seamless interface into the extended reality world. In this talk, we will explain why eye tracking is crucial in the edge devices of the Metaverse and show the breakthrough that Ganzin has achieved.

S16.4 🕓 16:45 — 17:00

RF/ASIC ICs in B5G/6G SatCom Equipment and Phased Array Radars

Yu-Jiu Wang
Tron Future Tech

In this talk, we will first overview recent progress of active phased array antennas, and discuss their adoptions in B5G/6G SatCom, and radar market. Then, we will discuss RF/ ASIC IC requirements and usage in such systems. In the last part, we will introduce technological innovations by Tron Future Tech in both markets.