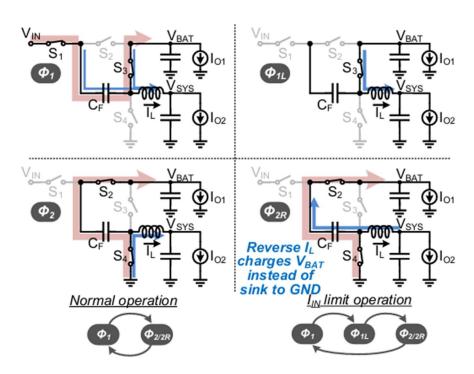
KAIST 전기및전자공학부 박사과정 박수언

Session 2 Hybrid Converters

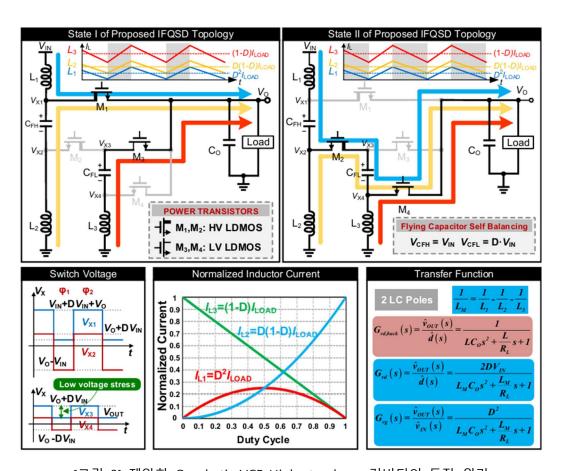
이번 ASSCC 2024의 Session 2는 hybrid converter를 주 내용으로 총 4편이 제출되었다. 이 중 2편의 논문 (2-3, 2-4)에서 유망한 산업 분야인 배터리 충전과 데이터센터용 전력 공급 IC를 다루어 이를 살펴보고자 한다.

#2-3 논문은 경북대학교에서 발표한 논문으로, 종래의 V_{BAT} , V_{SYS} 용 primary-secondary charger를 one-stage로 regulation하는 IC를 제안한다. 제안한 IC는 one-stage regulation으로 V_{SYS} regulation 효율이 증가하며, V_{BAT} CC mode charger 용 2:1 SC charger와 inductor를 결합한 하이브리드 토폴로지에 의해 면적, 즉 IC 제작 비용이 감소하는 장점이 있다. V_{SYS} 는 PWM으로 제어되고, V_{IN} 이 unplug된 상황에서는 V_{BAT} 으로부터 전력을 공급받는 시나리오를 구현하였다. 또한, Charger의 I_{IN} limit 기능은 컨버터의 duty를 감소시킴으로써 2:1 SC의 Rout을 증가하는 방식으로 지원하는 방법을 제안하였다. 그 결과, 제안한 Charger용 IC는 I_{IN} limit 기능과 함께 0.77A/mm²의 on-die current density와 CaS(Charge and Supply) 모드, peak P_{OUT} point에서 92.5%의 높은 효율을 달성하였다.



[그림 1] 제안한 one-stage V_{BAT} , V_{SYS} regulation operation phase

#2-4 논문은 Northwestern Polytech 대학의 논문으로, 데이터센터용 Quadratic-VCR High-stepdown 컨버터를 제안한다. 제안한 컨버터는 Inductor-First 구조를 가져 EMI가 작으며, Triple-path로 부하를 공급하여 인덕터 DCR과 스위치의 온저항에서 발생하는 conduction loss가 감소하는 장점이 있다. 또한, conventional buck과 같이 RHP zero가 없는 장점이 있다. 출력 전압 제어 방법으로서 Cout의 ESR을 이용한 ripple-based control의 단점인 큰 ESR의 사용은 컨버터의 과도응답특성과 효율을 감소시키기 때문에, 작은 ESR로 ripple-based control이 가능한 In-chip Ripple Compensation을 제안하였다. 그 결과, 제안한 Ripple Compensation을 사용한 COT(Constant On-Time) 방식으로 출력 전압을 제어하며, 12V-to-1V에서 90.4%의 효율을 달성하였다. 하지만, 본 구조는 3개의 인덕터를 사용하여 power-density가 감소하는 구조적 단점을 가지며, 이 부분은 데이터센터용 IC로서 다소 아쉽다.



[그림 2] 제안한 Quadratic-VCR High-stepdown 컨버터의 동작 원리

저자정보



명예기자 박수언

● 소속 : KAIST 전기 및 전자공학부 박사과정

● 연구분야 : Power Management IC 설계

● 이메일 : tndjs12221@kaist.ac.kr

● 홈페이지 : https://idec.or.kr

KAIST 전기및전자공학부 박사과정 박수언

Session 18 Temperature Resilient Analog Circuits

이번 ASSCC 2024의 Session 18은 온도 변화에 Immune한 회로 설계 기법을 주 내용으로 총 4편의 논문이 제출되었다. 이 중 MOS-based Energy-efficient Temperature sensor와 Polynomial Temperature Compensated TCXO에 대한 2편의 논문을 살펴보고자 한다.

#18-2 논문은 KAIST에서 발표한 논문으로, MOS-based temperature sensor(TS)의 temperature conversion을 에너지 소모 관점에서 최적화한 TS를 제안한다. 제안한 TS는 시간에 대해 변화가 작은 온도의 본질적인 특성에 기인하여, LSB부터 MSB로 최종 코드를 탐색하는 LSB-First SAR Logic을 적용한 CDAC을 사용, 탐색 과정에서의 energy 효율을 높인 장점이 있다. 또한, Temperature-sensitivity를 높인 Asymmetric Floating Inverter Amplifier(FIA) Based의 preamp와 strong-arm comparator를 결합한 comparator는 에너지 효율을 2차적으로 증가시키며, Asynchronous CLK의 사용은 conversion에 필요한 energy를 크게 감소시켰다. 위 기법들에 의해 제안한 TS는 2 point calibration 이후 3-sigma 기준 -1.93/+1.44 °C의 낮은 inaccuracy를 가지며 1.2V supply에서 conversion 당 41.9pJ의에너지 소모, 71.8mK의 우수한 온도 resolution을 갖는다.

Presented Energy-efficient MOS-based Temperature Sensor

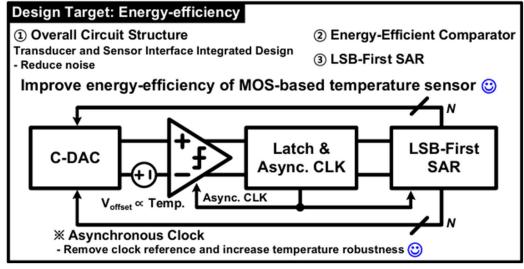
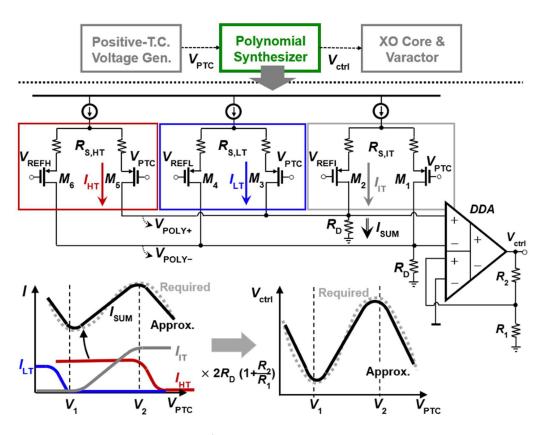


그림 3 제안한 Energy-efficient MOS-based 온도 센서의 디자인 기법

#18-3 논문은 Taiwan University에서 발표한 논문으로, Pierce oscillator 기반의 Crystal oscillator(XO)의 Polynomial한 온도에 대한 주파수 변화 특성을 보상하는 기법을 제안한다. 제안한 보상 방법은 XO의 cubic한 온도에 따른 주파수 변화 특성을 differential pair의 large signal 특성을 활용해 인풋 TR에 흐르는 전류의 off point를 설정하여 XO의 polynomial을 inversion한 근사 파형을 만들었다. 또한, startup 속도를 개선시키기 위해Gm-boosting technique을 추가하였다. 그 결과 테스트한 5개의 샘플에서 -40~80°C의 온도 range에서 ±2.5ppm의 주파수 variation과 1kHz에서 -98.4dBc/Hz의 Phase noise를 달성하였다.



[그림 4] 차동증폭기의 대신호 특성을 활용한 XO의 Polynomial한 온도에 따른 주파수 특성 모사

저자정보



명예기자 박수언

● 소속 : KAIST 전기 및 전자공학부 박사과정

● 연구분야 : Power Management IC 설계

● 이메일 : tndjs12221@kaist.ac.kr

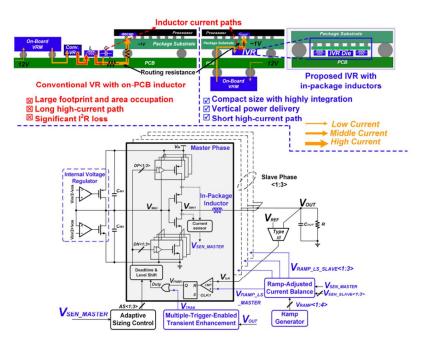
● 홈페이지 : https://idec.or.kr

고려대학교 전기전자공학과 박사과정 이윤호

Session 5 DC-DC Converter

이번 A-SSCC 2024의 Session 5는 DC-DC Converter와 관련된 총 4편의 논문이 발표되었으며, 이 들 논문에서는 LC oscillator 기반과 switched capacitor 기반의 converter 등 다양한 topology의 converter 기술이 소개되었다. 그 중에서도 고구동 application을 위한 4-phase integrated voltage regulator와 portable device를 위한 buck-boost battery charger 기술에 관한 두 가지 논문에 대해 알아보고자 한다.

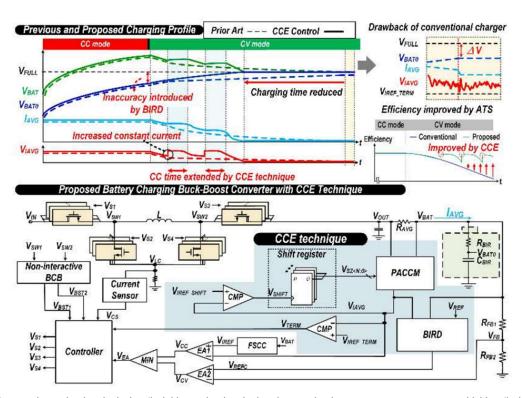
#5-3 본 논문은 AI와 고성능 GPU의 발전으로 인해 프로세서의 전력 소모가 급격히 증가함에 따라 발생되는 전력 손실과 발열 이슈를 해결하고자 하였다. 이 문제점을 해결하기 위해서는 voltage regulator(VR)가 processor와 최대한 가까이 배치되면서 높은 current density와 고효율을 제공해야 한다. VR 설계에서 가장 큰 병목 현상은 인덕터의 크기로, 이를 극복하기 위해 본 논문에서는 그림 1과 같이 in-package inductors를 사용하여 효율적이고 집적도가 높은 솔루션을 제안하였다.



[그림 1] Vertical power delivery utilizing the proposed IVR

작은 크기의 SMD 인덕터(In-package inductor)를 활용한 high-frequency, 4-phase integrated voltage regulator (그림1)를 설계하여 비용 효율적이면서도 고효율을 구현하였으며, 다중 트리거 기반 과도 응답 기술을 도입해 fast load transient 상황에서도 낮은 undershoot를 유지하였다. 또한, ramp-adjusted current balance technique을 통해 multiphase current를 조정함으로써 안정성과 신뢰성을 개선하였고, four 7.2nH 0402 inductors integrated in the flip-chip chip-scale package를 통해 전압 조정기를 프로세서 근처에 배치함으로써 전력 손실과 발열을 최소화하였다. 그 결과, 본 설계는 최대 7A의 출력 전류와 7A/mm²의 전류 밀도를 달성하였으며, 전압 강하를 254mV(PWM alone)에서 88mV로 65% 개선하였다. 또한, 89.0%의 최대 효율을 보이고 있다.

#5-4 본 논문은 portable devices에서 전력 수요와 배터리 용량 증가에 대응하여 더 빠르고 효율적인 충전 솔루션을 제안한다. 기존 charging profile에서 built-in resistance detection(BIRD) 기술은 온도 및 상태 변화로 인해 부정확한 결과가 발생하여 충전 시간이 길어지는 문제가 있어, 이를 해결하기 위해 그림 2와 같이 constant current extension 기술을 도입하여 충전 효율성을 높이고 종료 시점을 정확히 판단하는 방식을 제안하였다.



[그림 2] 기존 충전 방식과 제안하는 충전 방식 비교 (위)와 CCE technique를 포함한 제안하는 buck-boost battery charger (아래)

이에 더해서, precision adjustable charging current monitor를 통해 충전 전류를 모니터링하여 노이즈 영향을 최소화하고 정확도를 향상시켰으며, fast start constant current 기술을 적용해 constant current mode 동작을 앞당겨 충전 시간을 단축하였다. 또한, non-Interactive BCB을 도입하여 충전 회로의 안정성을 강화하고 작은 충전 전류 센싱 정확도를 개선하였다. 그 결과, 충전 시간이 기존 대비 50.5% 단축되어 2배 빠른 충전 속도를 달성하였고, 출력 전류 및 전압 리플이 84% 감소했으며, 최고 97.1%의 효율을 달성하였다. 칩 면적은 3.1mm²이다.

저자정보



이윤호 박사과정 대학원생

● 소속 : 고려대학교

● 연구분야 : Power management ICs

• 이메일 : uknow@korea.ac.kr

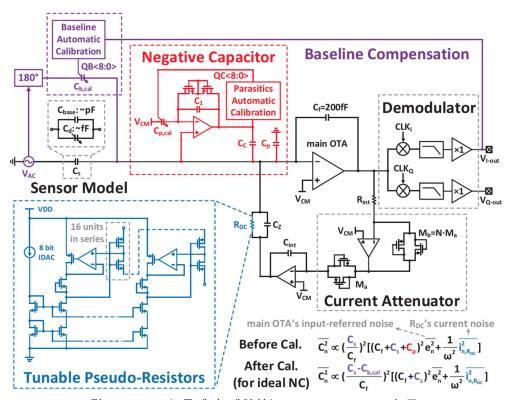
• 홈페이지: https://sites.google.com/site/kubasiclab/home

고려대학교 전기및전자공학부 박사과정 안재웅

Session 14 Amplifiers and Power Management

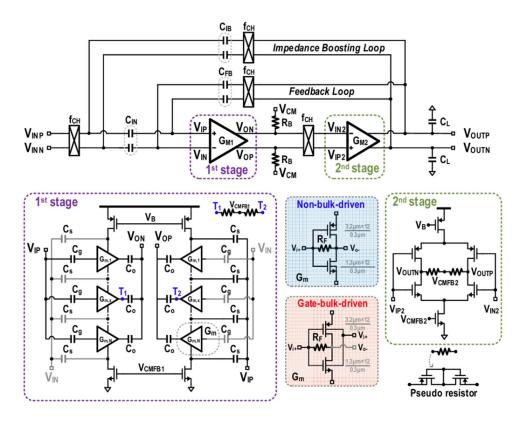
이번 2024 IEEE A-SSCC의 Session 14에서는 Amplifiers and Power Management라는 주제로 총 4편의 논문이 발표되었다. 이 중 #14-1은 증폭기 설계의 정밀도를 높이기 위한 새로운 보정 기술을, #14-2는 에너지 효율을 극대화하는 저전력 설계를 다루었다.

#14-1 논문에서는 capacitive sensing을 위한 precision lock-in amplifier를 제안하며, auto-calibrated negative capacitor를 통해 기생 커패시턴스를 제거하여 bandwidth과 resolution을 개선하였다. 기생 커패시턴스는 해상도를 저하의 주요 요인이므로, 이를 제거하는 기술이 필수적이다. 또한, automatic baseline capacitance calibration을 도입하여 SNR을 보장하고, tunable pseudo-resistor network를 통해 122nA까지의 DC 전류를 허용하였다. 이를 통해 2.06zFrms의 용량 해상도를 구현하였으며, 62MHz의 대역폭과 5.3mW의 낮은 전력 소모를 달성하였다.



[그림 1] #14-1 논문에서 제안하는 lock-in amplifier의 구조

#14-2 논문에서는 gate-and-bulk-driven stacking amplifier를 사용하여 0.59의 소음 효율계수(NEF)와 0.42의 전력 효율계수(PEF)를 구현하였다. 초저전력 환경에서 신호 품질을유지하려면 NEF와 PEF를 동시에 극대화하는 설계가 요구된다. 이 논문에서는 3-stacking topology를 통해 6배의 current reuse를 달성하고, capacitive negative feedback network를통해 high gain을 유지하고 offset과 noise를 억제하였다. 또한, temperature 및 supply variation에 견고한 설계를 통해 신뢰성을 높였으며, 33.6nW의 전력 소모로 2.16μVrms의입력 잡음 성능(IRN)을 유지하였다.



[그림 2] #14-2 논문에서 제안하는 stacking amplifier의 구조

저자정보



안재웅 박사과정 대학원생

● 소속 : 고려대학교

● 연구분야 : 디스플레이 드라이버 / 픽셀 보상 / 터치 센서

• 이메일 : ajw1104@korea.ac.kr

• 홈페이지: https://sites.google.com/site/kubasiclab/home

고려대학교 전기및전자공학부 박사과정 안재웅

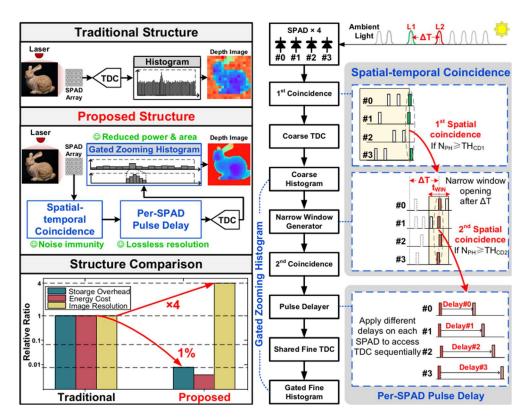
Session 24 Advanced Sensor and Imaging Technologies

이번 2024 IEEE ASSCC의 Session 24는 Advanced Sensor and Imagin Technologies 라는 주제로 총 4편의 논문이 발표되었다. SPAD LiDAR Sensor, Smart Ring에 적용되는 PPG Acquisition Circuit, CMOS Imager의 주제로 발표가 되었으며, 이번 리뷰에서는 그 중 두 가지 주제를 다룰 예정이다.

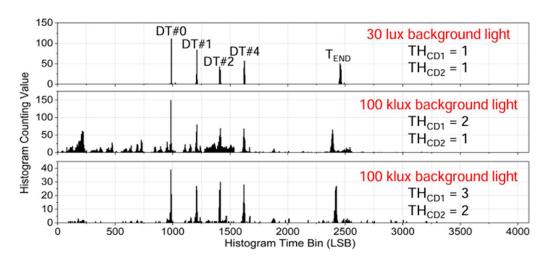
#24-1 Single-photon Avalanche Diode (SPAD)-based LiDAR 기술은 고급 운전자 보조 시스템(ADAS)에서 주로 활용되며, 레이저로부터 얻은 거리 정보를 추출하기 위해 background light (BGL) 데이터를 효과적으로 필터링해야 한다. 그러나 SPAD 픽셀 배열의 증가로 인해 대량의 데이터 처리, 저장 및 전송 과정에서 전력 소모와 레이아웃 크기 증가 문제가 발생한다. 본 논문에서는 이를 해결하기 위해 [그림 1]과 같이 새로운 detection scheme을 제안하였다.

제안된 방식은 ΔT 간격으로 두 번의 레이저 펄스를 방출하고, 각각 coarse TDC와 fine TDC를 활용해 데이터를 처리한다. 마이크로 픽셀은 4개의 SPAD로 구성되며, 첫 번째 레이저 펄스를 통해 coarse time 데이터를 추출하고, 이를 저장 및 peak value와 비교한다. Peak value를 초과한다면, 두 번째 레이저 펄스는 좁은 time window (t_{win}) 을 통해 수신된다. 이 과정에서 각 SPAD의 데이터는 다양한 delay를 적용받아 단일 fine TDC에 순차적으로 입력된다. 이를 통해 BGL을 효과적으로 제거하면서 storage와 cost를 줄일 수 있고 또한 기존의 coincidence 기법의 해상도 문제를 극복할 수 있었다.

실험 결과, 본 논문은 100 klux 환경에서 98.3%의 BGL 데이터 압축률을 달성하고, 픽셀당 14.6 μ W의 저전력 소모를 확인하였다. 또한, [그림 2]와 같이 pixel histogram을 통해 100klux의 상황에서도 효과적으로 background light가 제거됨을 증명하였다.



[그림 1] 제안하는 detection scheme



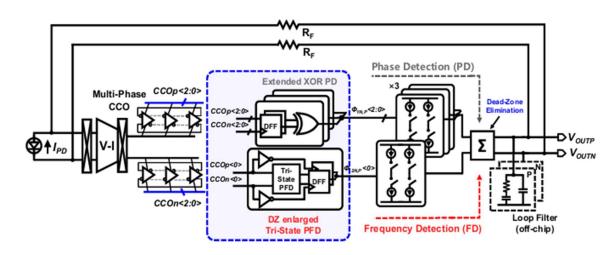
[그림 2] 4개의 SPAD를 이용하여 측정한 pixel histogram

#24-3 Photoplethysmography (PPG)는 심박수와 혈중 산소 포화도를 판단하기 위하여 사용되는 non-invasive biological signal detection 기술이다. 그러나 기존의 기술들은 사용자의 불편함, 비용, 유연하지 않은 구조 등으로 인해 한계가 있었다. 이에 본 논문에서는 TFT 회로를 이용하여 패키징 문제와 유연성 문제를 해결한 fully flexible ring을 제안하였다. 또한 TFT의 성능 저하 문제점을 해결하기 위하여 fully VCO 기반 구조와 dead-zone free TIA 구조를 사용하였다.

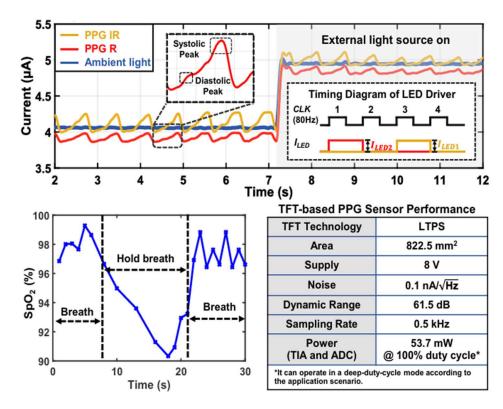
제안된 VCO-TIA는 [그림 3]과 같이 VCO, PFD, charge pump (CP), 그리고 loop filter (LP)로

구성되어 있다. VCO를 time-domain integrator로 사용하여 Infinite DC gain을 얻을 수 있으며, PFD에서는 제안된 기술을 통해 dead zone을 제거하여 nonlinearity와 noise를 제거할 수 있었다. 또한 제안한 0-1 MASH ADC를 통해 input nonlinear error를 제거하고 noise level을 줄일 수 있었다.

실험 결과, PPG 신호는 주변 광 조건에 따라 차이를 보였으며, SpO_2 값이 호흡에 따라 변화하는 것이 [그림 4]를 통해 확인되었다.



[그림 3] 제안된 VCO-TIA 회로 다이어그램



[그림 4] PPG, SpO₂ 그리고 TFT-based PPG sensor 성능 측정 결과

저자정보



안재웅 박사과정 대학원생

● 소속 : 고려대학교

● 연구분야 : 디스플레이 드라이버 / 픽셀 보상 / 터치 센서

• 이메일 : ajw1104@korea.ac.kr

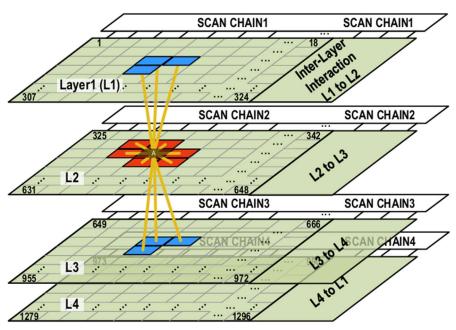
● 홈페이지: https://sites.google.com/site/kubasiclab/home

포항공과대학교 반도체대학원 박사과정 박은빈

Session 20 Innovative Computing and Circuit Design

이번 2024 IEEE A-SSCC 학회의 Session 20은 Innovative Computing and Circuit Design라는 주제로 총 5편의 논문이 발표되었다. 이 세션에서는 고성능 컴퓨팅 및 회로 설계에서 혁신적인 접근법을 통해 다양한 응용 시나리오에서의 효율성과 정확도를 향상시키는 데 초점을 맞추었다.

#20-2 논문에서는 3D 아날로그 Ising 머신(CAIM)을 설계하여 조합 최적화 문제를 해결하기 위한 새로운 접근 방식을 제안하였다. 기존 2D 또는 2.5D Ising 머신이 제한된 상호 작용 자유도와 크기 확장성 문제를 가지고 있었던 반면, 본 논문은 Hexagonal Close-Packed(HCP) 구조를 채택하여 스핀 당 12개의 상호작용 자유도를 제공하며, 기존 구조대비 두 배 이상의 연결성을 확보하였다. 또한, HCP 구조는 상호작용 간 배선을 최적화하여 층 간 복잡성을 크게 줄였으며, 4층 HCP 구조의 1296 스핀 레이아웃을 구현함으로써 확장 가능성과 설계 효율성을 입증하였다.



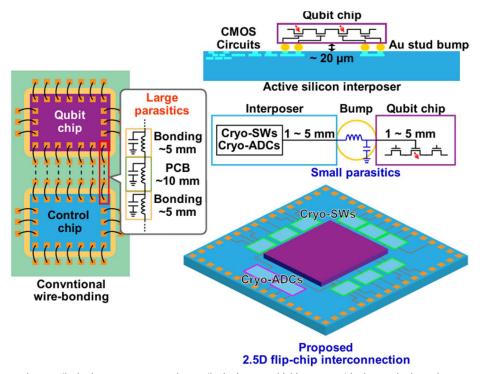
[그림 3] 4층 구성의 3차원 HCP 구조로 구성된 제안된 CAIM

본 논문은 아날로그 스핀 코어의 안정성을 확보하기 위해 Supply-And-Bulk Injection Locking(SABIL) 기술을 도입하였다. 기존 주입 잠금 방식(TIL)이 상호작용에 따른 삼중 안

정성 문제를 유발하여 수렴 속도가 느려지는 반면, SABIL은 공급 전압 및 벌크 주입을 활용하여 스핀 코어가 빠르고 안정적으로 수렴할 수 있도록 지원한다. 이를 통해 수렴 주기를 5 사이클로 단축하였으며, 기존 TIL 대비 58.3%의 개선된 수렴 속도를 달성하였다. 또한, SABIL은 기존 아날로그 Ising 머신 대비 8.5배 넓은 풀인 범위를 제공하여 주파수 변동에 강건한 특성을 나타낸다.

측정 결과, 본 3D Ising 머신은 최대 18×18×4 스핀 문제를 효과적으로 해결하며, 조합 최적화 문제인 max-cut 문제에서 높은 정확도와 에너지 효율을 입증하였다. 특히, SABIL을 활용한 경우 3D 문제를 빠르고 정확하게 해결할 수 있었으며, 실험적으로 평균 에너지 최소화 정확도가 98%에 도달함을 확인하였다. 본 논문은 3D 구조와 안정적인 아날로 그 스핀 설계를 결합하여 기존 Ising 머신의 한계를 극복하였으며, 조합 최적화 및 기타 NP-난해 문제 해결을 위한 하드웨어 설계에서 중요한 진전을 보여준다.

#20-3 논문에서는 극저온 환경에서 스핀 큐비트 제어를 위한 초고속 펄스 셰이퍼를 제안하였다. 본 연구는 스핀 큐비트 제어를 위한 1ns 시간 해상도의 Cryo-ADC와 Cryo-SW를 포함하는 2.5D 실리콘 인터포저 구조를 설계하였다. 기존의 외부 제어 시스템이 긴배선으로 인해 신호 무결성과 열 유입 문제를 겪었던 것과 달리, 본 논문의 설계는 큐비트와의 거리를 1mm 이내로 줄여 신호 정확도와 열 관리 성능을 크게 향상시켰다. 이를통해 고속 게이트 펄스 제어가 가능해졌으며, 양자 컴퓨팅의 확장성 문제를 해결하는 데기여하였다.



[그림 4] 제안된 cryogenic 펄스 셰이퍼를 포함한 2.5D 실리콘 인터포저 구조

Cryo-ADC는 극저온 환경에서 낮은 전력 소비와 고해상도를 동시에 달성하도록 설계되었다. 3.3V 전압에서 동작하는 완전 동적 SAR 구조를 채택하여 평균 35µW의 전력으로 11비트 해상도를 제공하며, 50MHz의 유효 대역폭을 지원한다. 또한, 동적 샘플링 및 부트스트랩 스위치를 활용한 설계를 통해 기존 구조 대비 입력 정밀도를 83% 개선하였다. Cryo-SW는 큐비트와의 신호 전송 거리를 최소화하여 신호 감쇠를 줄였으며, 이를 통해 2-큐비트 연산에서 요구되는 10mV 이하의 제어 정확도를 만족시켰다.

측정 결과, 본 펄스 셰이퍼는 100mK 이하의 온도에서 안정적으로 작동하며, 외부 펄스 제너레이터 대비 신호 감쇠를 0.3dB로 줄이고, 전압 변동성을 6.5mVpp(1.8mVrms)로 최소화하였다. 이를 통해 2-큐비트 연산에서 필요한 광대역 신호와 높은 정확도를 구현하였으며, 양자 컴퓨팅 환경에서 큐비트 제어의 효율성과 신뢰성을 높였다. 본 논문은 극저온 회로 설계와 양자 컴퓨팅의 융합을 통해 스핀 큐비트 제어 기술의 새로운 기준을 제시하며, 향후 양자 프로세서의 대규모 스케일링 가능성을 확대하는 데 기여할 것으로 기대된다.

저자정보



박은빈 박사과정 대학원생

● 소속 : 포항공과대학교

● 연구분야 : HW설계 및 딥러닝 최적화

• 이메일 : eunbin@postech.ac.kr

● 홈페이지:

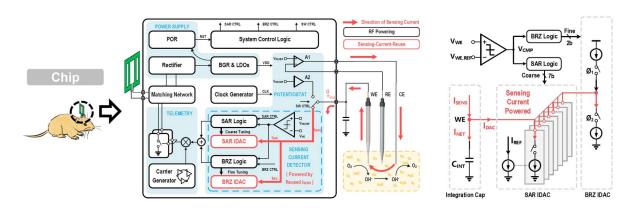
https://sites.google.com/view/epiclab/member/ebpark

KAIST 전기및전자공학부 박사과정 윤기찬

Session 4 Emerging Wireless and Battery-Free Systems

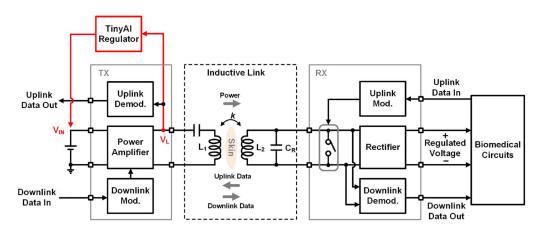
이번 2024 IEEE A-SSCC의 Session 4는 Emerging Wireless and Battery-Free Systems이라는 주제로 총 4편의 논문이 발표되었다. 이 세션에서는 무선 전력 전송 및 데이터 전송 회로가 포함된 센서 인터페이스 시스템에 관련된 논문이 2편, 새로운 통신 방식 제안을 주제로 2편의 논문이 연구 결과를 제시하였다. 본 review에서는 그 중 4-1, 4-2, 4-3을 리뷰하고자 한다.

#4-1 본 논문은 Zhejiang University에서 발표한 연구로, 외상성 뇌손상 등과 같은 뇌 질환 치료에 도움을 주는 뇌 산소 모니터링 (Cerebral oxygen monitoring)인 PO_2 센싱 시스템을 소개한다. 무선동작 및 장기간 동안 안정적인 사용을 위해, 전력 소모를 크게 줄여야 하고, 이를 위한 저전력 센싱 시스템 및 무선 전력 전송 및 데이터 전송 시스템을 제안한다. 기존의 potentiostat을 활용한 센싱 시스템은 TIA나 전류 미러, ADC 등 추가적인 전력 소모가 필요한 회로를 필요로 한다. 본 논문에서는 센싱 전류를 추가적인 전력 소모 없이, sinking 전류원을 사용하는 IDAC으로 디지털화 하는 기술을 제안하였다. 증폭단이 없어짐으로써 발생하는 해상도 문제는 bipolar residue zoom-in IDAC을 통해 해결하였다. 결과적으로 PO_2 센싱 회로가 포함된 시스템 중 가장 낮은 전력 소모인 25μ W를 달성하였으며, 해상도도 0.028%로 2배 이상 좋아졌다. 또한, backscattering 방식을 활용한 434MHz의 주파수를 가지는 passive TX를 통해 무선 전력 전송 및 데이터 전송 회로를 구현하여, 저전력으로 2m 거리에서 통신이 가능한 시스템을 구성하였다.



[그림 1] 무선 뇌 산소 모니터링 시스템 및 제안된 potentiostat

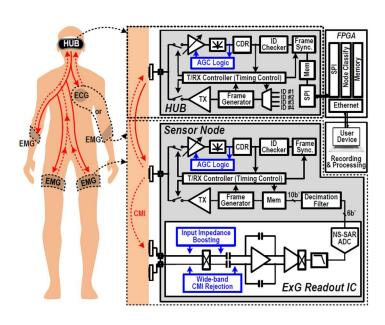
#4-2 본 논문은 Shanghai Jiao Tong University에서 발표한 연구로, WPDT 시스템에서 RX regulation을 위한 새로운 기술 및 개선된 데이터 전송 방식을 제안하였다. 기존에는 RX 에 local loop을 구성하거나, TX로 uplink를 통해 regulation을 진행하였다. 하지만, 전력 효율을 감소시키거나, regulation 속도가 제한된다는 단점을 각각 가지고 있었다. 이를 해 결하기 위해, 본 논문에서는 TX에 local loop을 구성하였다. TX의 coil에 걸리는 전압을 센 싱 및 neural network를 training하여, RX의 load 상태를 예측하고 TX의 V⋈을 조절하여 전력 효율을 개선하였다. 또한, 데이터 전송을 개선하기 위하여 Adaptive-BPSK (A-BPSK) 와 Auto Time Tracking Cyclic OOK (ATT-COOK)를 각각 uplink와 downlink에 적용 및 제 안하였다. 기존 BPSK는 각 bitstream을 전송하기 위해 6.5T의 PLL locking 시간이 필요한 반면, A-BPSK는 변화하지 않는 bitstream (11 혹은 00)을 전송할 때에는 1T의 시간으로도 인코딩 할 수 있도록 하여, 평균적으로 1.73배 빠른 데이터 전송 속도를 확보할 수 있었 다. COOK은 한 carrier cycle에서 스위치를 1번 닫는 위치를 통해 데이터를 전송하기 위 해 에너지 손실을 최소화할 수 있다. 하지만, 스위치를 닫는 타이밍에 따라 전력 손실이 커질 수 있고, RX 코일에 걸리는 전압이 최소가 되었을 때를 변조 타이밍으로 선택하여 전력 손실을 최소화하고, 변조 depth를 13.5% 향상시켰다. 결과적으로, 기존 대비 3.7배 높은 전력 전송, 1.4배 빠른 downlink 및 가장 빠른 uplink 속도를 보여주었다.



[그림 2] 제안하는 WPDT 시스템

#4-3 본 논문은 KAIST에서 발표한 연구로, 생체신호 (ECG, EMG)를 4채널의 센서 노드로 부터 측정하고, 그 데이터를 body channel communication (BCC)를 통해 무선으로 허브로 전달하는 시스템을 구현하였다. BCC는 몸을 매개체로 저전력 무선 통신을 가능하게 한다. 하지만, 사람의 움직임에 따라 body channel 특성이 달라지게 되고, 그에 따라 신호의 크기가 time-varying하게 달라지고, real-time communication을 어렵게 한다. 본 논문에서는 automatic gain control (AGC) 기술을 활용하여 채널 특성 변화를 극복하고, OOK 변조를 통해 데이터 전송 효율을 높이고자 하였다. 기존 AGC 기술이 포함된 demodulator 같은

경우에는 일반적으로 여러 stage의 증폭기와 정류기를 필요로 하여, 높은 전력 소모가 요구된다. 제안하는 AGC 기술은 각 증폭기의 전압이득을 feedback factor 조정을 통해 programmable하게 만들어, 필요로 하는 증폭기와 rectifier의 수를 줄임으로써 전력소모를 크게 줄일 수 있었다. 또한, BCC 동작은 body channel을 통해 데이터를 전달하기 때문에, ExG recording 회로에게 큰 common-mode interference (CMI)를 인가하게 된다. 본논문에서는 CMI 제거 회로 및 입력 임피던스 향상 회로를 통해 BCC와 ExG 신호 간의 간섭을 해결하여 concurrent한 동작이 가능하도록 하였다.



[그림 3] 생체신호 측정 인터페이스 및 BCC 통신 시스템

저자정보



윤기찬 박사과정 대학원생

● 소속 : KAIST

• 연구분야 : Mixed-signal IC & Sensor interface IC

• 이메일 : ygc980215@kaist.ac.kr

• 홈페이지 : https://impact.kaist.ac.kr

한국과학기술원 바이오및뇌공학과 박사과정 석동열

Session 17 Biosignal and Biomedical Applications

올해 2024 아시아고체회로학회(ASSCC)의 바이오-신호 및 바이오-의료분야 응용회로 부문(Session 17)에 선정된 논문은 5개로 주로 바이오-의료목적에 특화된 센서 및 계측 (instrumentation) 회로 연구를 다루고 있다. 본 부문에 게재된 논문의 제목과 연구 목적과 방법론을 요약해보면(표1), 웨어러블(wearable)이라는 키워드가 주요하게 다가온다. 보다 세부적으로 보면, 1) 웨어러블 바이오 센서 소형화 및 저전력화, 2) 웨어러블 환경의한계를 극복을 위한 인터페이스 회로의 성능 개선, 3) 웨어러블 바이오 회로의 응용분야 멀티모드 센서 및 인터페이스 회로 개발로 정리해 볼 수 있다. 이 리뷰에서는 게재된 다섯 건의 연구 중 #17-1, #17-4 두 건을 살펴보고자 한다.

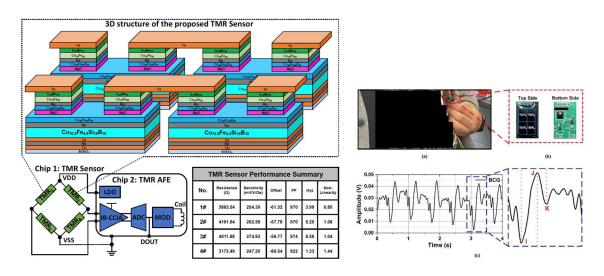
[표 1] 바이오-신호 및 바이오-의료분야 응용 부문 (Session 17) 선정 논문 5개

번호	제목	요약
#17-1	A Non-contact TMR-Based Active Electrode for BCG	MEMS 기술 기반으로 소형화 된 TMR(Tunneling
	Recording	Magnetoresistance) 심탄도(BCG, Ballistocardiogram)
		센서 및 인터페이스 회로 개발
#17-2	A 1.96Vpp Input-Range 3rd-order Noise-Shaping	웨어러블 환경(큰 입력 범위, 높은 입력임피던스
	Pipelined-SAR ADC with Self-Boosted Impedance and	요구, 저전력 소모, 저잡음 등) 극복을 위한 고성능
	Active Charge Sharing for Biosignal Acquisition	ADC 개발(방법은 논문 제목 참조)
#17-3	RMMIC: A 40µW Reconfigurable Multimodal Multi-	다양한 생체신호(ECG, EOG, PPG, EIT) 수집을 위한
	channel Interface Circuit System for Bio-signals	인터페이스 회로 소형화 및 저전력화, 단일 증폭기
	Monitoring	및 디지털 컨버터 활용을 위한 컨트롤러 활용
#17-4	A 1.27-mW Battery-Powered Multimodal Biosensor	웨어러블 로봇 내 움직임을 예측 및 감지하기 위한
	Interface IC for Wireless Motion Intention Recognition	무선 저전력 생체신호 및 압력센서(ExG, CDC,
	in Wearable Robots	pMMG, EIT 등) 인터페이스 회로 소형화 개발
#17-5	A 1.92nJ/Conv Pulse-Width Locked-Loop Time Domain	저항변화를 이용한 변형(strain) 센서를 PWLL(Pulse-
	Readout IC with VCO-Integrator and Pipeline TDC for	width Locked Loop)을 이용하여 구현, 센서 소형화
	Wheatstone Bridge Wearable Strain Sensing System	및 저전력화 달성

#17-1 심탄도(BCG, Ballistocardiography)는 심장박동에서 발생하는 물리적 진동을 전기신호화 한 것으로 심전도(ECG)와 함께 활용하여 혈류 속도, 혈압 등의 모니터링에 활용되는 기술이다. 본 연구는 항저우의 저장연구실, 베이징 중국과학원의 공동 연구로써 MEMS 기술을 활용하여 고감도의 소형화 TMR(Tunneling Magnetoresistance) 센서를 구현하여 이를 실시간 심탄도 측정에 활용할 수 있음을 실험을 통해 보여주고 있다. 또한, 최종으로 구현된 실험장치는 그 크기로 볼 때 웨어러블 등 소형화 장비에서의 활용 가능

성을 보여주고 있다.

논문은 기존 NiFe 또는 CoFe 층을 대체하여 비정형 결정구조를 가지며 높은 자성저항 (magnetoresistance) 값을 갖는 연자성체 CoFeSiB 층을 도입한 것이 TMR 센서 성능 향상에 기여하는 주요한 요소로 언급하고 있다. $0.18\mu m$ 1P6M CMOS 공정을 활용하여 TMR Sensor에 해당하는 $3.08~cm^2$ 크기의 Chip 하나와 계측회로에 해당하는 $1.98~mm^2$ 크기의 별 개의 Chip을 하나를 제작했으며, 각 Chip을 인쇄회로기판의 앞 뒤에 배치하여 2.8cm~x 1.7cm 크기의 능동전극(active electrode)을 제작하여 실험했다. TMR센서는 $\pm 100\mu T$ 수준의 입력범위, 10pT 수준의 민감도(sensitivity)를 주요 성능 지표로 보여주고 있으며, 계측회로를 IC기반으로 소형화 제작하여 결합함으로써 제안된 센서의 활용성을 높여 보여주고 있다.

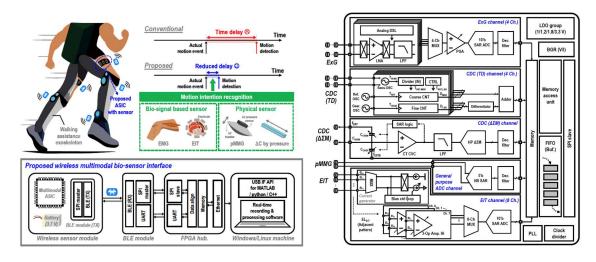


[그림 1] 연구 #17-1이 제시하고 있는 TMR Sensor와 계측회로의 개요도와 In-vivo BCG 측정결과

#17-4 본 논문은 KAIST, 삼성전자, SK Hynix, NYU Abu Dhabi의 공동연구로 사람의 동작에 관한 다양한 정보를 동시에 수집하여 웨어러블 로봇과의 상호작용에 활용하기 위한 기반연구이다. 연구에서 수집하고자 목적하는 신호로는 ExG(EEG, ECG, EMG 등), CDC(capacitance-to-digital converter), pMMG(pressure mechanomyography), EIT(electrical impedance tomography)가 있으며, 각각의 신호를 수집하기 위한 인터페이스 회로를 하나의 칩 안에 모두 구현하였다. 제작된 칩에는 1) ExG 수집을 위한 CCIA(chopper-stabilized capacitive instrumentation amplifier)를 비롯하여 2) 시간-도메인 4채널 CDC, 3) DSM 기반의 CDC, 4) pMMG 센서 정보를 읽어 내기 위한 5bit의 NS SAR, 5) EIT 수집 목적의 전류생성기 및 프론트엔드를 포함한 임피던스 측정기가 포함되어 있다.

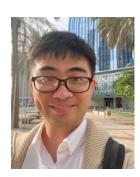
180nm CMOS 공정을 활용하여 14mm² 크기의 칩을 제작하였으며, 각 인터페이스 모듈에서 주어진 입력에 대해 예상되는 출력신호를 얻을 수 있음을 실험적으로 보여주고 있다. 각 모듈의 성능지표로는 1) 4채널 ExG 센서에서 최대 60dB 이득, $0.8\mu V_{rms}$ (0.5 ~

500Hz) 수준의 입력 변환 노이즈, 2) 시간-도메인 CDC에서 $0\sim300pF$ 의 입력 범위(해상도 최대 $1.8fF_{rms}$), 3) EIT 센서에서 최대 $0.33\Omega_{rms}$ 수준의 해상도를 보여주고 있으며, 전체 회로의 전력소모는 1.27mW 수준이다.



[그림 2] #17-4 웨어러블 로봇 사용자 동작 감지를 위한 인터페이스 회로 연구개요 및 구조도

저자정보



석동열 박사과정 대학원생

● 소속 : 한국과학기술원

● 연구분야 : 바이오메디컬 응용회로 설계(센서 및 신호처리)

• 이메일 : sukd10@kaist.ac.kr

한국과학기술원 바이오및뇌공학과 박사과정 석동열

Session 21 Application-Specific Processors

올해 2024 아시아고체회로학회(ASSCC)의 특수목적 프로세서(Application-Specific Processors) 부문(Session 21)에 선정된 논문은 5개로 주로 일반적인 연산 프로세서를 활용할 때 시간, 메모리, 전력 등의 측면에서 비효율적으로 계산이 이루어 질 수밖에 없는 다양한 응용 분야에서 특정 형태의 연산을 효과적으로 효율적으로 수행하기 위한 하드웨어 연구를 다루고 있다. 본 부문에 게재된 논문의 제목과 연구 목적과 방법론을 요약해보면(표2), 생체신호나 시각적 정보처리를 웨어러블 장치 또는 인공지능 로봇 등에서 빠르고 전력 효율적으로 수행하기 위한 특수목적 프로세서, 데이터 압축해제, 조합론적-최적화문제(COP) 등 기존의 프로세서가 수행 할 경우 시간이 오래 걸리는 연산을 가속하는 하드웨어 등이 소개되어 있다. 이 리뷰에서는 게재된 다섯 건의 연구 중 #21-3 한 건을 살펴보고자 한다.

[표 2] 특수목적 프로세서 부문 (Session 21) 선정 논문 5개

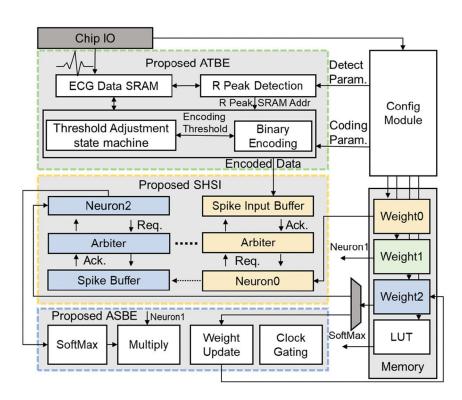
번호	제목	요약
#21-1	ROC-Spin: A 28nm 2,000 Ring-Oscillator-Collapse	조합론적-최적화문제(COP) 해결을 위한 이징 모델의
	Spins for Solving Combinatorial Optimization	스핀을 링-오실레이터의 듀티 사이클로 구현, 복잡도
	Problems	높은 COP 문제를 빠르게 해결할 수 있음을 보임
#21-2	A Unified Microrobotic Visual-Perception Processor	마이크로로봇의 시각인지프로세서에서 이루어지는
	with 62.2-FPS/mm2 and 103-uJ/frame Navigation in	Cholesky 분해를 가속화 하는 연산모듈 ReMMDA를
	28nm	고안하여 시간적, 하드웨어 측면에서 연산성능 향상
#21-3	A 0.04 μJ/Classification High-Accuracy Energy-	SNN 층 간 불필요한 연산을 줄이고, 심전도파형
	Efficient ECG Processor with SNN On-Chip	인코딩 방식 제안, 마지막 층 가중치를 역전파보정
	Backpropagation and Adaptive Threshold Encoding	하는 방법으로 지연을 줄이고 정확도를 높인 연구
#21-4	A 1.12nJ/Pixel High-Accuracy and Memory-Efficient	하이브리드 다운 샘플링, 배경제거, 관심영역 설정
	Real-Time Object Detection Processor for	등을 하드웨어로 구현하여 1.12nJ/pixel 수준의
	Neuromorphic Vision Sensors	고메모리 효율의 실시간 사물 감지 프로세서 개발
#21-5	A 43.3 bit/cycle Inflate Accelerator Featuring Static-	데이터 압축 해제를 위한 inflate 알고리즘을 하드웨어
	Dynamic Huffman Decoder with Multiple	가속기 형태로 개발한 연구로 해당 과정에 사용되는
	Checkpoints and Optimized End-Of-Block Control	Huffman Decoder, LZ77 Decoder가 압축 데이터를
	for Hyperscale data	병렬로 처리하도록 하여 처리 속도를 향상함

#21-3 SNN(spike neural network)은 신경세포의 작동방식을 모방한 인공 신경망 네트워크로 이벤트 발생시 반응하는 특성을 가지고 있어서 저전력 방식의 생체신호 분류 목적의 프로세서에 자주 사용된다. 본 연구는 중국 전자과기대학, Hangzhou Shiguang Xight Inc.

Nanjing Houmo Al Inc.의 공동연구로 진행되었으며, SNN 방식의 심전도(ECG) 분류 하드웨어에서 적응형 임계 값을 활용한 심전도 파형 디코딩 방식(ATBE, Adaptive Threshold Binary Encoding) 과 시간 효율적인 각 층(layer)간 Handshaking 방식의 연산 모듈(SHSI, SNN Hand Shaking Inference), 역전파를 통한 마지막 층 가중치 보정(ASBE, Adaptive SNN Backpropagation Engine)을 도입하여 지연시간을 줄이고, 정확도를 높일 수 있음을 보여주었다.

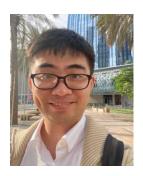
ATBE는 심전도 전위 값의 변화를 추적하여 정상적인 심전도의 파형을 구성하는 P, Q, R, S, T파의 위치 정보를 96 비트의 벡터에 담아내는 알고리즘으로 전위 값의 변화를 검출하는 한계를 적응조절(adaptive control) 하여, 인코딩 된 데이터의 유효성을 높이는 모듈이며, SHSI는 앞 층에서 전달되는 스파이크 정보를 입력 버퍼(spike input buffer)에 저장한 후 전달 된 경우에 대해서만 연산하고, 연산 중에 앞뒤 레이어에서의 스파이크를 방지하는 방식으로 SNN 연산을 단순화하여 지연 시간을 줄였다. 또한, 소프트맥스(softmax) 모듈을 이용하여, 입출력 값 비교를 통한 가중치 변화를 계산하여 마지막 층가중치를 보정하는 ASBE를 통해 전체 시스템의 정확도를 향상시켰다.

55nm CMOS 공정을 활용하여 0.31mm² 의 칩을 제작하였으며, 분류 건 당 0.04 μ J 수준의 에너지 소모와 98.6%의 분류 정확도 0.3ms 수준의 지연을 성능지표로 보고하였다.



[그림 3] #21-3 ECG 분류 목적의 SNN에서 지연시간 감소, 저전력화 및 정확도 향상 연구 개요도

저자정보



석동열 박사과정 대학원생

● 소속 : 한국과학기술원

● 연구분야 : 바이오메디컬 응용회로 설계(센서 및 신호처리)

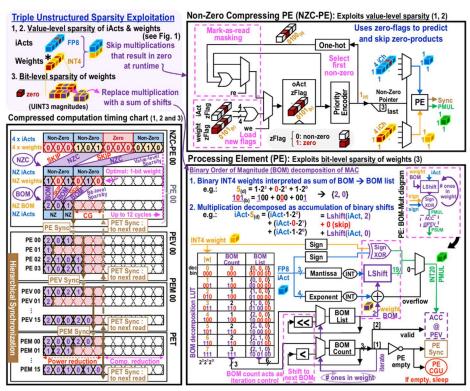
• 이메일 : sukd10@kaist.ac.kr

포항공과대학교 반도체대학원 박사과정 박은빈

Session 3 Efficient AI and DSP Processors

이번 2024 IEEE A-SSCC 학회의 Session 3에서는 Efficient AI and DSP Processor라는 주제로 총 5편의 논문이 발표되었다. 이 세션은 주로 고효율 AI 및 DSP 프로세서 설계를 통해 다양한 응용 시나리오에서 전력 소비를 줄이고 성능을 극대화하는 데 중점을 두었다.

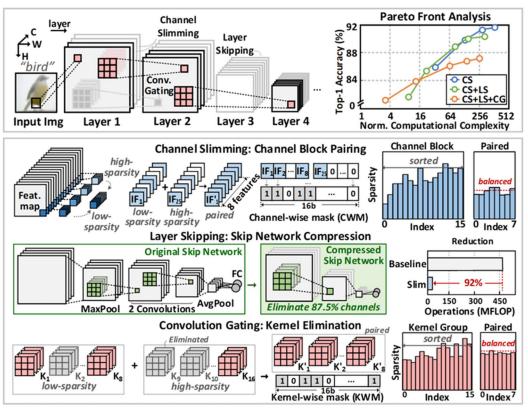
#3-2 논문에서는 WhiteDwarf라는 고효율 신경망 추론 엔진을 발표하였다. 이 엔진은 ** 트리플 비정형 희소성(triple unstructured sparsity)**과 모델 압축 기술을 결합하여 CNN 및 MLP 모델의 연산 효율성을 극대화한다. 특히, 활성화, 가중치, 비트 수준의 희소성을 활용해 데이터 처리 효율성을 높이고, INT2-4와 FP8 같은 적응형 정밀도를 통해 메모리접근과 전력 소비를 줄였다. 이를 통해 ResNet-50 모델은 원래 크기의 5.7%로 압축되었고, 정확도 손실 없이 ImageNet 데이터셋에서 74.7%의 성능을 유지하며, 12.24 TFLOPS/W의 에너지 효율성을 달성하였다.



[그림 1] triple unstructured sparsity에 관한 설명

WhiteDwarf는 RISC-V 기반의 커스텀 컨트롤러와 Huffman 코딩 기술을 사용하여 다양한 신경망 모델을 지원하고 메모리 접근을 최소화한다. 또한, 클럭 게이팅(clock gating)과 같 은 전력 최적화 기술을 통해 유휴 상태의 전력 소비를 줄였다. 본 논문은 트리플 희소성 활용과 적응형 정밀도를 통해 엣지 컴퓨팅과 같은 자원 제약 환경에서의 딥러닝 적용 가 능성을 크게 확대하며, 차세대 신경망 가속기의 설계 방향성을 제시한다는 점에서 큰 의 의를 가진다.

#3-5 논문에서는 동적 신경망(dynamic neural network)의 하드웨어 지원을 통해 에너지 효율을 극대화하는 딥러닝 가속기를 제안하였다. 이 가속기는 채널 슬리밍(channel slimming), 레이어 스키핑(layer skipping), 컨볼루션 게이팅(convolution gating) 같은 최신 최적화 기법을 결합하여 다양한 운영 시나리오에 적응할 수 있는 설계를 구현하였다. 특히, 채널 및 커널 데이터를 페어링하여 연산 효율을 극대화하고, 불필요한 연산을 줄여 메모리 접근과 전력 소비를 효과적으로 감소시켰다. 이러한 기술을 통해 8.1-to-353 TOPS/W의 에너지 효율성과 높은 처리량을 실현하였다.



[그림 2] 채널 슬리밍, 레이어 스키핑, 컨볼루션 게이팅에 관한 설명

이 가속기는 CIFAR-10 데이터셋을 기반으로 실험한 결과, 83~93%의 정확도를 유지하면 서도 최대 3.94배 높은 에너지 효율과 55.8배 높은 면적 효율을 기록하였다. 또한, 동적 데이터 흐름(dynamic dataflow)을 활용해 프로세싱 유닛(PE)의 활용도를 극대화하고, 레이어 스키핑을 통해 비핵심 연산을 효율적으로 제거하였다. 이러한 설계는 엣지 AI와 같은 자원 제약 환경에서 딥러닝의 효율적 운영을 가능하게 하며, 하드웨어 차원에서 동적 신경망의 잠재력을 효과적으로 실현한 점에서 큰 의의를 가진다. 본 논문은 다양한 에너지 -성능 요구를 충족시키는 차세대 딥러닝 가속기의 설계 방향성을 제시하며, AI 응용 분야에서의 실용적 기여 가능성을 입증하였다.

저자정보



박은빈 박사과정 대학원생

● 소속 : 포항공과대학교

● 연구분야 : HW설계 및 딥러닝 최적화

• 이메일 : eunbin@postech.ac.kr

● 홈페이지:

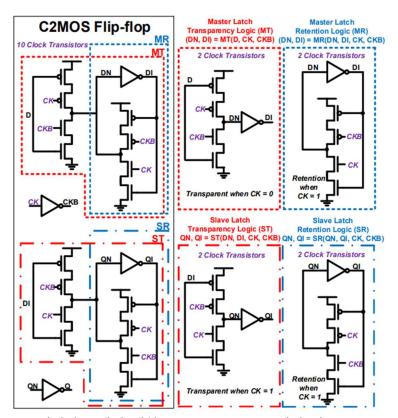
https://sites.google.com/view/epiclab/member/ebpark

KAIST 전기및전자공학부 석사과정 박민하

Session 15 Energy-Efficient Circuit Level Techniques

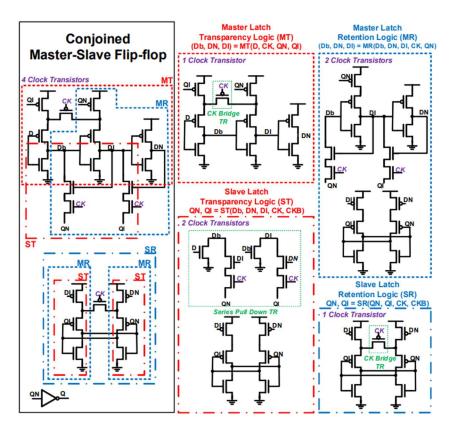
Session 15 Energy-Efficient Circuit Level Techniques에서는 Asynchronous Spiking Neural Network, Time-multiplexed Random-Access Processing-in-Memory, Static Contention-free Conjoined Master-Slave Flip-flop, Asynchronous Non-Volatile-Memory-based Computing-In-Memory Neuromorphic Processor 등 총 5편의 논문이 발표되었다. 이 중 Static Contention-free Conjoined Master-Slave Flip-flop, Asynchronous Non-Volatile-Memory-based Computing-In-Memory Neuromorphic Processor에 대해 살펴보고자 한다.

#15-1 본 논문은 Low-Voltage, Low-Power, 그리고 Low-Area를 갖는 Contention-free Master-Slave Flip-Flop을 제안한다. 기존 C2MOS 또는 flip-flop의 경우 static contention-free와 낮은 전력 소비를 동시에 달성하지 못하는 문제점을 갖고 있거나, 이를 달성하더라도 큰 면적을 차지하는 문제점이 있다.



[그림 1] 독립적인 두개의 래치(Master/Slave)로 구성된 기존 C2MOS 구조

위와 같이 기존의 C2MOS 구조는 총 24개의 트렌지스터와 10개의 클록 트렌지스터를 요구한다. 그러나 제안된 CMSFF의 구조는 다음과 같다.



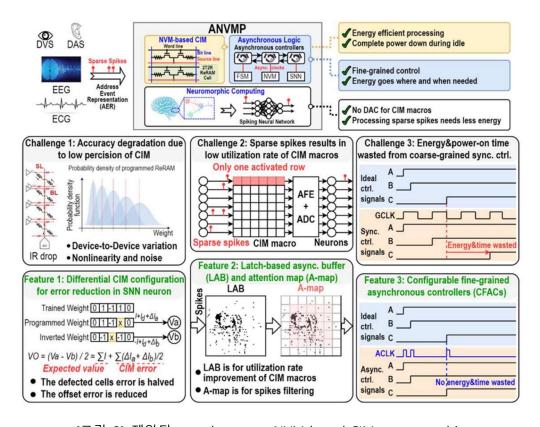
[그림 2] 제안된 CMSFF 구조

CMSFF는 Master와 Slave 래치를 유기적으로 결합하여 트렌지스터와 클록 제어 회로를 줄였다. Master latch의 transparency와 데이터 유지 기능을 slave latch와 공유하며 클록 트렌지스터를 기준으로 2개에서 1개로 줄여, Transparent 상태와 retention 상태 간의 전환을 효율적으로 수행하도록 하였다. 결과적으로, 트렌지스터 수를 22개로 줄였으며, 면적 및 전력을 절감할 수 있었다.

28nm 공정을 통해 test chip을 제작하였고, CMPSFF를 포함한 7개의 플립플롭과 비교실험을 수행한 결과, CMSFF는 클록 전력을 1.0V에서 60%, 0.4V에서 63%까지 절감하였다. 면적 측면에서는 REFF 대비 20%, TGFF 대비 6%의 면적 감소를 달성하였다. 또한 몬테카를로 시뮬레이션에서 CMSFF는 static contention-free를 입증하였고, 저전압 환경에서도 안정적으로 동작함을 보여주었다.

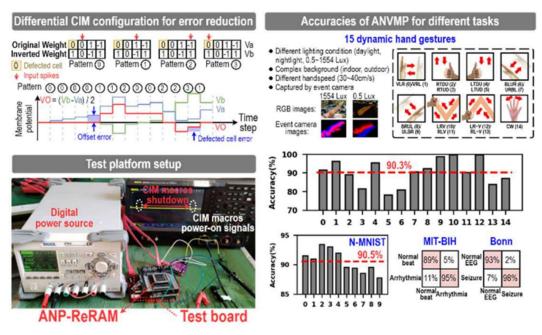
#15-2 본 논문은 저전력 엣지 AI 애플리케이션을 위해 설계된 비휘발성 메모리 기반 컴퓨팅-인-메모리(CIM) 뉴로모픽 프로세서(ANVMP)에 관한 연구이다. ANVMP는 28nm 공정으로 제작되었으며, 다양한 작업에서 높은 에너지 효율성과 낮은 전력 소비를 달성합니

다. 엣지 AI 장치에서 저전력과 다기능성에 대한 요구가 증가하고 있고, 기존 연구는 mW급 전력 소모 또는 단일 작업에 국한된 문제를 가진다. 다음 그림 3은 ANVMP가 해결하고자 하는 주요 과제와 이를 해결하기 위한 핵심 기술을 시각적으로 나타내고 있다. 제안된 ANVMP가 기존 뉴로모픽 설계의 한계를 어떻게 극복하였는지를 한 눈에 보여주는 중요한 도식으로, 본 연구의 주요 기요를 나타낸다.



[그림 3] 제안된 asynchoronous NVM-based CIM neuromorphic

ANVMP 설계의 도전 과제는 NVM 기반 CIM의 오차와 SNN 뉴런의 누적 오차, 희소 스파이크로 인한 CIM 매크로 활용률 저하, 그리고 글로벌 클럭 사용으로 인한 에너지 낭비였다.



[그림 4] ANVMP의 오차 보정 능력과 성능 평가

따라서, 제안된 설계에서는, Differential CIM configuration으로 CIM 오차를 줄이고 평균 정확도를 2% 향상시켰다. 그리고 비동기 버퍼(LAB)와 A-map으로 불필요한 전력 소비를 줄였다. 이를 통해 power-on 시간을 최대 45.1%까지 절약할 수 있었다. 또한 세분화된 비동기 컨트롤러CFAC)를 통해 동적 주파수 제어로 CIM 매크로의 power-on 시간을 41% 단축하였다.

제안된 ANVMP는 비휘발성 메모리(NVM) 기반 CIM, 스파이킹 신경망(SNN), 비동기 logic를 통합하였고, 결과적으로 보면, 이미지 분류에서 90.5%, 손 제스처 인식에서 90.3%, 심전도(ECG) 부정맥 감지에서 94%, 뇌파(EEG) 발작 감지에서 95.6%의 정확도를 보였다. 또한 전력 소비는 평균적으로 52.6µW였으며, 디지털 코어가 전체 전력의 63.5% 차지하였다.

저자정보



박민하 석사과정 대학원생

● 소속 : KAIST

● 연구분야 : 디지털 회로 설계● 이메일 : mhpark@ics.kaist.ac.kr

● 홈페이지 : https://idec.or.k

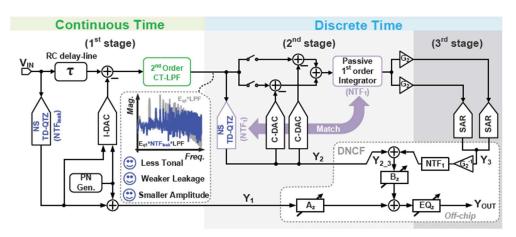
KAIST 전기및전자공학부 박사과정 윤기찬

Session 19 Continuous-time ADCs

이번 2024 IEEE A-SSCC의 Session 19는 Continuous-time ADCs라는 주제로 총 4편의 논문이 발표되었다. 4편의 논문 모두 continuous-time 영역에서 전력 효율적으로 높은 해상도를 얻기 위해 noise-shaping technique을 적용한 논문으로, 2개의 논문은 최근 많은 연구가 진행되고 있는 TDC를 활용한 MASH 형태의 pipelined ADC였다.

#19-1 본 논문은 University of Macau에서 발표한 연구로, 320MHz의 대역폭을 가지는 CT/DT hybrid pipelined ADC를 제안한다. 기존에는 고속 동작 CT domain의 고해상도 ADC를 만들기 위해서 OSR이나 NTF의 차수를 높이거나, 혹은 quantizer의 해상도를 높이는 방법을 택해왔다. 하지만, 이러한 방식은 power-hungry한 loop filter나 quantizer를 필요로 함에 따라 전력 효율성이 좋지 않았다. 반면 DT domain으로 구성하면, 저전력 달성이 가능하지만 sampling 동작에 따른 aliasing 효과가 나타나게 된다. 본 논문은 이러한 두 가지 특성을 합치기 위해, 첫번째 stage는 CT로 구성하여 anti-aliasing 효과를 얻고, 뒤의 stage들을 DT domain로 구성하여 전력 효율성을 높이고자 하였다.

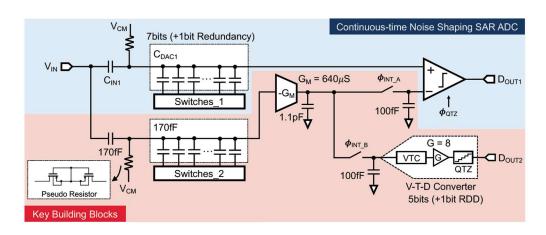
첫번째 stage는 NS TD-QTZ의 결과를 RC delay-line과 I-DAC을 통해 residue voltage를 생성 및, SAB를 사용하여 증폭 및 DAC image를 필터링하여 전달한다. 2번째 stage 또한 NS TD-QTZ를 사용하여 첫 stage의 residue voltage 및 NTF leakage들을 shaping해주며, 마지막 stage에선 고속동작을 위해 6b TI SAR ADC를 구현하였다. 결과적으로, 320MHz의 넓은 BW에서 65dB가 넘는 SNDR과 57.7fJ/conv.-step으로 우수한 FoMw을 달성하였다.



[그림 1] 제안하는 CT/DT hybrid pipelined ADC

#19-2 본 논문은 KAIST에서 발표한 연구로, CT NS-SAR ADC와 TDC를 결합한 1-0 MASH 형태의 ADC를 제안한다. 일반적으로 NS-SAR ADC는 DT domain에서 구현이 되는데, 최근에 CT domain에서의 구현이 제안된 바 있다. CT DSM과 유사하게 anti-aliasing 특성을 가지고 있지만, time-varying하는 입력 신호를 sampling 없이 quantization하면서 error가 발생하고, quantization 과정 중 residue voltage가 존재하지 않아 sampling speed가 제한된다는 단점이 있다. 본 논문에서는 입력 stage를 하나 추가하여, sampling speed를 최대화할 수 있게 하였으며, 동작 속도가 빨라지며 발생하는 input dependent error를 1-0 MASH 구조를 통해 해결하고자 하였다.

먼저, capacitively coupled replica input stage (CCIS)의 추가를 통해, SAR quantization을 진행하는 동안에도 CT 적분기가 계속해서 적분을 할 수 있게 하였다. 또한, 그 적분기의 출력을 TDC를 통해 quantization함으로써, quantization noise를 더 낮추고, 첫번째 stage 의 input dependent error를 shaping할 수 있었다. 그 과정에서 발생하는 적분이 출력 CM 전압 변화 문제를 새로운 VTC의 ramp generator 방식을 제안함으로써 해결할 수 있었다. 결과적으로, 기존에 수 MHz로 제한되던 sampling speed를 400MHz까지 크게 늘릴 수 있었으며, inherent anti-aliasing 특성으로 인해 DT NS-SAR ADC에 비해 더 작은 input capacitor를 사용할 수 있었다. 또한, 23.2fJ/conv.-step으로 우수한 FoMw을 달성하였다.

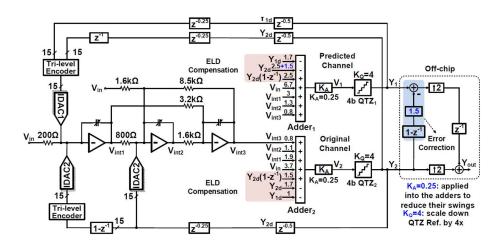


[그림 2] 제안하는 CT NS-SAR ADC와 TDC를 결합한 구조

#19-3 본 논문은 Shanghai Jiao Tong University에서 발표한 연구로, 180MHz의 대역폭을 제공하는 2채널 time-interleaving 3차 CT-DSM을 제안하였다. CT-DSM의 대역폭을 넓히기 위해서는 일반적으로 MASH 구조 (더 높은 차수의 NTF) 혹은 quantizer의 해상도를 높게 가져가게 되는데, 이는 전력 효율성이 떨어진다. 다른 방식으로는, CT-DSM에 time-interleaving 기술을 구현하는 것이다. 단순히 integrator를 2개로 구현하는 것이 아닌 각채널이 다른 채널의 정보를 예측하면서 recursive하게 동작하는 extrapolated TI-DSM이제안된 바 있으며, 이를 구현함에 있어 feedforward path와 ELDC를 위한 전력이 크게 증가하고, DAC error compensation이 어렵다는 문제가 있었다. 본 논문에서는 기존 전압 기

반 feedforward를 G_m -cell 및 IDAC을 통해 구현함으로써, 기존의 저항들이 가지는 loading effect를 줄였다. 또한, source degeneration technique을 활용할 때, capacitor를 parallel로 추가하여 LHP zero를 추가함에 따라 1.64배 넓은 대역폭을 가질 수 있었다. 또한 TI CT DSM에는 다양한 feedback 신호가 필요하게 되고, 이를tri-level IDAC을 사용하여 IDAC의 unit cell 개수를 줄였다. 또한, 상보적인 신호를 생성하는 것을 reverse encoding을 통해 구현함으로써 DAC에서 발생하는 mismatch를 감소시켰다.

결과적으로, 해당 논문은 5.76GHz의 빠른 동작이 가능하였고, 180MHz 대역폭에서 68.3dB의 SNDR을 얻을 수 있었다. 그 결과, 제안한 ADC는 비슷한 동작 주파수인 수 GHz sampling speed CT DSM들 중 가장 높은 164.5dB의 FoM_{SDR}을 달성하였다.



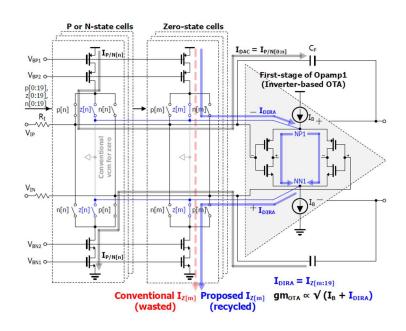
[그림 3] 제안하는 time-interleaving CT DSM 구조

#19-4 본 논문은 Samsung에서 발표한 연구로, 앞선 3개의 논문과는 다르게, 24kHz의 대역폭을 가지는 상대적으로 느린 주파수 대역을 사용하는 audio application을 위한 CT DSM을 제안하였다. Audio용 ADC는 wearable device 등에 많이 사용되면서, 저전력 및 높은 DR을 요구하고 있다. Low noise 특성을 달성하기 위해 kT/C noise가 있는 DT ADC 들 보다는 CT DSM이 주로 사용되고 있다. CT DSM에서 가장 주된 noise는 첫번째 integrator의 input-referred noise와 DAC cell의 thermal noise이며, negative-R 혹은 OTA staking 등을 통해 integrator의 noise를 줄일 수 있으나, 추가적인 noise를 발생시키거나 swing range 등이 제한될 수 있다.

본 논문에서는 tri-level current steering DAC을 사용하되, zero 상태에 있는 DAC의 전류를 낭비하지 않고 integrator의 OTA의 bias current로 재사용함으로써, gm을 키우고 noise를 낮추는 DIRA 방식을 제안하였다. 하지만, 이러한 동작은 signal dependent한 error를 AC ground에 발생시키고, 하모닉 성분을 발생시킨다. 이를 해결하기 위하여, quantizer의 결과와 FIR 필터를 통해서 신호의 크기를 tracking하고, 그에 따라 bias current를

adaptive하게 변화시키는 ADB 기술 또한 제안하였다. 추가적인 전력 소모가 필요함에도 불구하고, 일반적으로 오디오 신호의 크기가 작기 때문에, ADB회로가 동작하는 시간이 길지 않아 전력 소모 등에는 큰 영향을 주지 않는다고 한다.

결과적으로, 해당 회로는 작은 신호가 들어왔을 때 DR을 4.2dB 향상시켰으며, 큰 신호가 들어왔을 때 하모닉 성분을 -95.1dB까지 줄일 수 있었다. 412mW를 소모하면서 24kHz 대역폭에서 105.8dB의 높은 DR을 달성하여 180.2dB의 우수한 FoM_{S,DR}을 달성하였다.



[그림 4] 제안하는 DAC current recycling amplifier 구조

저자정보



윤기찬 박사과정 대학원생

● 소속: KAIST

• 연구분야: Mixed-signal IC & Sensor interface IC

• 이메일 : ygc980215@kaist.ac.kr

• 홈페이지 : https://impact.kaist.ac.kr

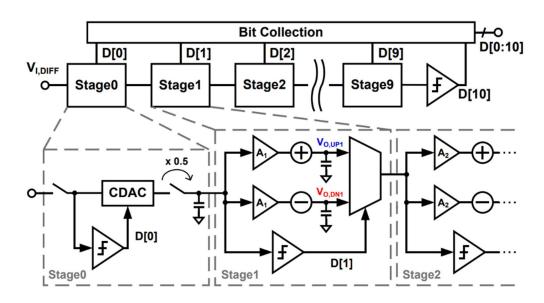
DGIST 전기전자컴퓨터공학과 박사과정 위정윤

Session 9 High-Speed Nyquist ADCs

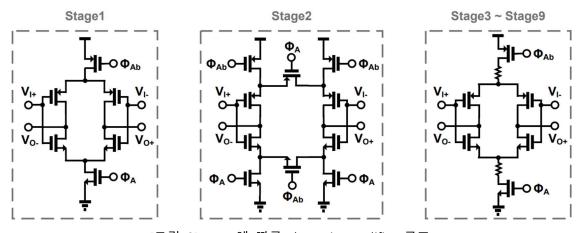
이번 2024 IEEE A-SSCC의 Session 9은 High-Speed Nyquist ADC라는 주제로 총 5편의 논문이 발표되었다. 최근 고속 어플리케이션에 적합한 ADC로써 time-interleaved ADC (TI-ADC), pipelined ADC, 그리고 두 구조의 장점을 결합한 time-interleaved pipelined ADC 기반의 ADC 구현 기술이 활발하게 연구되고 있다. 고속 ADC 구현의 핵심은 고속 입력을 빠르고 정확하게 처리하는 것에 있다. #9-2 논문에서 제시된 바와 같이, TI-ADC는 입력 신호를 여러 개의 sub-ADC로 나누어 시간적으로 분할하여 처리함으로써, 저속의 sub-ADC를 사용하여 넓은 대역폭을 구현할 수 있다는 장점이 있다. 반면, #9-1, 9-3 논문과 같이 pipelined ADC는 여러 단계의 sub-ADC를 직렬로 연결하여, 낮은 정밀도의 sub-ADC를 사용하더라도 높은 정밀도를 달성할 수 있다. 이러한 이유로, 고속과 높은 정밀도를 동시에 요구하는 어플리케이션에서는 #9-5 논문에서 제안된 time-interleaved pipelined ADC 구조가 자주 활용되고 있으며, 에너지 효율을 위해 sub-ADC로 SAR ADC 구조가 자주 채택되고 있다. 향후 연구에서도 높은 에너지 효율을 유지하면서 환경 변화에 강인한 고속 ADC를 개발하는 것이 중요한 과제가 될 것으로 보인다.

#9-3 본 논문에서는 amplifying-and-select stage를 제안하여 기존 pipelined ADC의 속도한계를 개선했으며, replica-biased supply-regulation 기술을 결합한 dynamic amplifier을 사용하여 에너지 효율을 높였다. 기존 pipelined ADC는 sampling, conversion, 그리고 residue amplifying 순차대로 conversion을 수행하여 각 pipeline 단계의 긴 처리 시간이전체 변환 속도를 제한하는 한계를 지니고 있다. 본 논문에서 제안된 amplifying-and-select stage는 가능한 residue 값을 미리 생성 및 증폭시킴으로써 pipeline ADC의 단계별처리에 따른 병목 현상을 해소했다. 본 논문에서는 1b SAR ADC를 sub-ADC로 사용하여두 가지 가능한 residue voltage 값을 미리 생성하고, amplifying함으로써 다음 stage의 conversion 준비 시간을 단축시켰다. 또한, 에너지 효율적인 dynamic amplifier (DA)를 사용하여 회로 복잡도 증가에 따른 전력 소모 문제를 최소화했지만, 기존 DA의 좁은 input common-mode 범위와 온도 민감성은 pipelined ADC의 안정적인 동작을 저해하는 요소로 작용할 수 있다. 본 논문에서는 replica-biased supply-regulation 기술과 더불어, 후속 stage 이후의 DA tail node에 저항 추가, 입력 sampling 동안 charge injection, clock

feedthrough에 의해 발생하는 common-mode 전압의 변화를 상쇄하기 위한 bootstrap sampling switch 등 여러가지 기술들을 사용하여 DA의 단점을 극복했다. 이런 기술들로 구현된 pipelined ADC는 3.1GS/s에서 동작 가능하며, 넓은 온도에서도 7.28b의 분해능 및 153.7dB-FoMS를 달성했다.

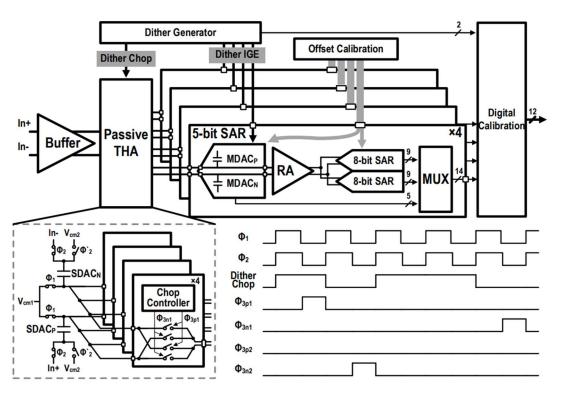


[그림 1] amplifying-and-select 기술이 적용된 제안된 pipelined ADC



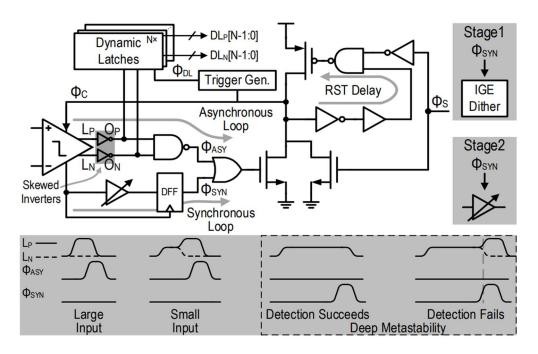
[그림 2] stage에 따른 dynamic amplifier 구조

#9-5 본 논문에서는 passive track-and-hold (T/H) 회로와 input-independent calibration을 사용하여 에너지 효율적으로 skew와 각 채널 간의 대역폭 mismatch 문제를 해결하고, metastability reduction 기술을 제안하여 에너지 효율적인 SAR ADC를 sub-ADC로 사용하



[그림 3] 제안된 4-way Time-interleaved-pipelined SAR ADC

는 과정에서 발생하는 comparator metastability 문제를 해결함으로써 고속 ADC의 성능 저하를 야기하는 주요 요인들을 개선하고 제안한 시스템의 성능을 향상시켰다. 그림3는 제안된 4-way TI-pipelined SAR ADC를 보여준다. 차동 입력 신호는 입력 버퍼와 passive T/H 회로를 거쳐 샘플링 되며, 1.5GHz의 고주파 입력 신호를 처리하기 위해 3GS/s의 샘 플링 속도가 필요하다. 이를 위해 4개의 채널에서 각각 750MS/s의 sampling 속도로 입 력 신호를 분담하여 처리한다. 각 sub 채널은 two-stage pipelined SAR ADC로 구현되었 으며, 특히 2단에서는 에너지 효율을 극대화하기 위해 2개의 sub-ADC로 추가 분할되었 다. 제안된 T/H 회로는 bottom plate 샘플링 방식을 통해 signal DAC (SDAC)에 입력 신호 를 샘플링한 후, dither chop 스위치를 거쳐 각 서브 채널의 multiplying DAC (MDAC)에 passive 하게 전달된다. Dither chop 스위치를 거치면서 발생하는 오프셋과 SDAC와 MDAC의 비율에 의해 결정되는 T/H의 이득 오차는 입력에 독립적인 형태로 calibration 될 수 있다. 제안된 passive T/H는 입력 신호와 무관하게 offset과 이득 오차를 보정하여 스위치의 임계 전압 변화에 둔감한 특성을 가지므로, 외부 환경 변화에 강인한 회로 구 현이 가능하다. 다음으로, 그림 4은 본 논문에서 제안된 metastability reduction 기술을 보여준다. 제안된 metastability reduction 기술은 빠른 conversion을 위해 asynchronous loop을 사용하고, 동시에 metastability 발생을 검출을 위해 synchronous loop을 병렬로 구현되어 있다. comparator에서 metastability가 발생하면 병렬로 구성된 루프는 최종적으 로 동기 루프에 의해 동작하게 된다. 동기 루프 종료 시점에서 LP와 LN의 값이 여전히 불안정한 상태로 유지된다면, 이를 통해 metastability 발생을 정확하게 감지할 수 있다. 본 시스템은 28nm CMOS 공정을 사용하여 0.095mm2의 면적으로 구현되었으며, Nyquist 입력에서 56.7dB의 SNDR과 75.2dB의 SFDR을 달성하였으며, 15.5fJ/conversion-step FoMW를 달성했다.



[그림 4] 제안된 metastability reduction 기술

저자정보



위정윤 박사과정 대학원생

● 소속 : DGIST 전기전자컴퓨터공학과

• 연구분야 : Analog ICs for biomedical application

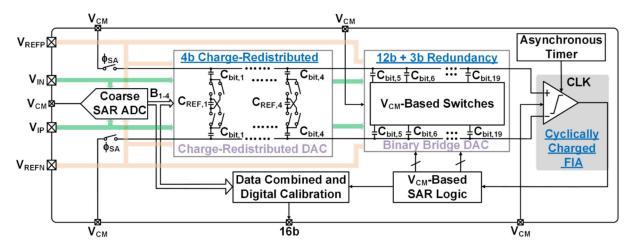
이메일 : wiejung@dgist.ac.kr홈페이지 : http://ins.dgist.ac.kr

DGIST 전기전자컴퓨터공학과 박사과정 위정윤

Session 23 Energy-Efficient and High-Resolution ADCs

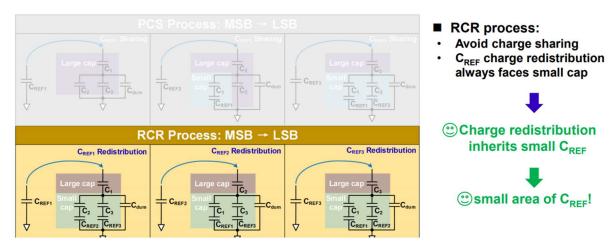
이번 2024 IEEE A-SSCC의 Session 23은 Energy-Efficient and High-Resolution ADC라는 주제로 총 4편의 논문이 발표되었다. 높은 에너지 효율을 자랑하는 SAR ADC는 최근 고분해능이 요구되는 다양한 분야에서 널리 활용되고 있다. 본 세션에서는 SAR ADC의 분해능을 높이면서 동시에 높은 에너지 효율을 유지하기 위한 다양한 방법들이 제시되었다. 첫째, 논문 #23-1에서는 SAR DAC의 비트 수를 16비트까지 높여 직접적으로 분해능을향상시키는 방법을 제안했다. 둘째, 논문 #23-2에서는 고차 loop filter를 사용하여quantization 노이즈를 감소시키고, oversampling ratio (OSR)을 낮춰 전력 소모를 줄이는동시에 분해능을 높이는 방법을 제시했다. 마지막으로 논문 #23-4에서는 저전력 residue amplifier을 제안하여 고차 loop filter을 구현함으로써 분해능을 향상시키는 방법을 제안했다. 이번 세션을 통해 SAR ADC의 분해능을 향상시키기 위한 다양한 접근 방식이 제시되었으며, 이는 고성능 저전력 ADC 개발에 기여할 것으로 기대된다.

#23-1 본 논문에서는 reservoir-redistributed DAC (RCR DAC)를 제안하여 회로 면적을 줄이고 전력 소모를 최소화하여 고효율 ADC를 구현했다. Charge redistribution DAC은 기존 VCM-based switch DAC와는 달리, CDAC 비트와 함께 CREF의 비트를 조절하여 reference 전압을 동적으로 변화시킴으로써, 더욱 넓은 출력 범위와 높은 분해능을 달성하는 기술이다. 본 논문에서 제안된 RCR DAC는 기존 passive charge redistributed DAC(PCR DAC)의 단점인 CREF 사이즈 문제를 해결하여 회로 면적을 효과적으로 줄였다. 기존 PCR DAC는 정확한 charge redistribution을 위해 MSB 비트 capacitor보다 CREF가 항상 커야 하는 제약이 있다. 본 논문에서는 RCR DAC는 새로운 charge redistribution 방식을 제안하여 작은 CREF 사용이 가능하게 했다. 또한, 본 논문에서는 floating-inverter-based amplifier를 comparator의 pre-amplifier로 사용하여 전체 시스템의 에너지 효율을 극대화했다.



[그림 5] 제안된 초소형 저전력 16비트 SAR ADC

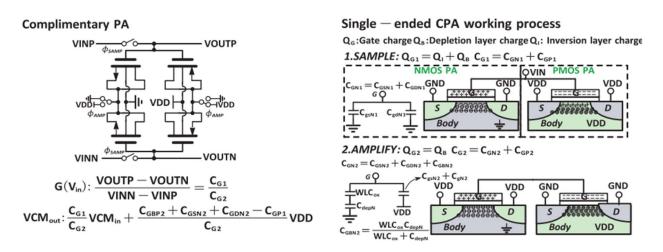
본 기술들을 통해 구현된 ADC는 그림 6과 같이 SAR-SAR sub-ranging 구조를 사용하여 두 단계의 conversion을 수행한다. 첫 번째 단계에서는 coarse SAR ADC를 통해 coarse conversion을 수행하고, 두 번째 단계에서는 4비트 RCR DAC와 12비트 DAC를 이용하여 fine conversion을 수행한다. 본 기술들을 통해 구현된 ADC는 저주파 입력에서 83.3dB SNDR과 103.3dB SFDR을 달성했고, Nyquist 입력에서 80.9dB SNDR과 100.3dB SFDR을 달성했다. 또한, 추가적인 calibration 없이 ±5% 전원전압 변화에서 2.5dB 이내의 SNDR 변화와, 3.5dB 이내 SFDR 변화를 달성했으며, -40~100°C의 온도 변화에서도 worst SNDR과 SFDR은 각각 79dB와 97dB 이상을 유지함으로써 전압, 온도변화에 안정적인 시스템을 구현했다.



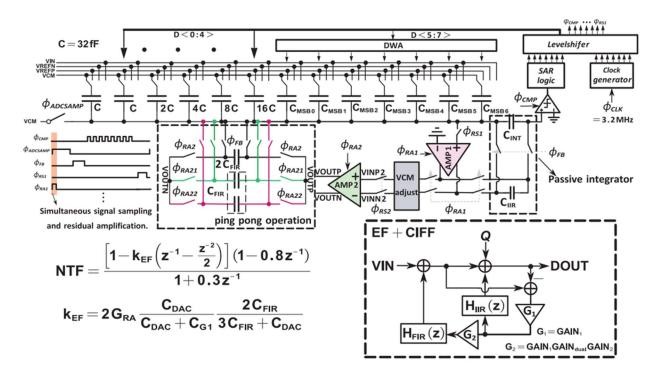
[그림 6] 제안된 reservoir-redistributed DAC (RCR DAC)

#23-4 본 논문에서는 complementary parametric amplifier을 제안하고, 이를 활용 및 검증하기 위해 8b EF+CIFF 토폴로지 기반의 3차 noise-shaping SAR (NS-SAR) ADC를 구현했다. 높은 에너지 효율과 정밀도를 요구하는 분야에 빈번히 사용되는 noise-shaping SAR ADC는 quantization 이후 발생하는 quantization noise를 고주파 영역으로 shaping

하는 기능을 탑재한 loop filter를 설계했다. 이를 통해 저주파 대역에서의 SNDR을 향상 시켜 저전력, 고정밀도를 요구하는 어플리케이션에 적합한 ADC를 구현했다. 제안된 complementary parametric amplifier (CPA)는 gate capacitance를 사용하여 passive amplification을 수행한 다는 점에서 기존 parametric amplifier (PA)와 동일하지만, 제안된 CPA는 P, N-type 입출력 MOSFET을 동시에 사용하여 P 또는 N-type만 단일 사용한 PA보 다 두배 넓은 입력범위를 갖는다는 장점이 있다. 더불어 gate capacitance에 sampling된 전하가 그대로 amplification에 사용되기 때문에, kT/C noise를 제외한 channel thermal noise current나 flicker noise과 같은 다른 noise source의 영향을 받지 않고, 전력 소모는 오로지 capacitor의 charging에 의해서만 발생하여, single-ended CPA기준 E=2·C_{G1}· Δ V_{IN}·V_{DD} 식으로 표현되는 매우 적은 양의 에너지만 소모한다는 장점이 있다. 그림 8은 CPA 기술이 내장된 NS-SAR ADC를 보여준다. 제안된 NS SAR ADC는 EF+CIFF loop filter 를 사용하여 효율적인 구조를 구현했다. 내장된 CPA는 첫 번째 residue amplifier, common-mode 보정, 두 번째 residue amplifier로 구현되어 있고, EF loop과 CIFF loop은 CPA의 첫 번째 residue amplifier을 공유함으로써 하드웨어 효율성을 높였다. CPA의 discrete한 amplifying 특성으로 인해 DAC는 두 번째 amplifier와 common-mode 조정이 작동하는 동안 첫 번째 샘플링을 동시에 수행하도록 구현되었다. 180nm CMOS 공정을 사용해 0.241mm²의 면적으로 제작된 본 논문은 10kHz 대역폭에서 250의 OSR로 동작할 때, 1.8V 공급전압에서 24.9µW, 아날로그 부분에서는 8.9µW의 전력을 소모했다.



[그림 7] 제안된 complementary parametric amplifier



[그림 8] complementary parametric amplifier 기반 3차 noise-shaping SAR ADC

Chip variation 확인을 위해 측정된 4개의 칩에서 모두 88dB 이상의 SNDR을 달성하였고, 이를 에너지 효율로 환산했을 때, 176.2dB FOMs를 달성하였다. 제안된 CPA의 효용성을 검증하기 위해 제안된 본 시스템은 높은 에너지 효율을 갖는 3차 noise-shaping SAR ADC를 구현함으로써 CPA의 안정성 및 효용성을 성공적으로 검증했다.

저자정보



위정윤 박사과정 대학원생

● 소속 : DGIST 전기전자컴퓨터공학과

• 연구분야 : Analog ICs for biomedical application

● 이메일 : wiejung@dgist.ac.kr

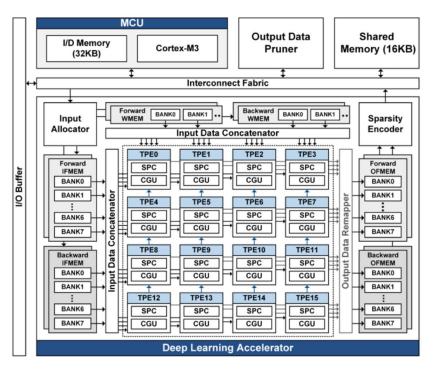
● 홈페이지: http://ins.dgist.ac.kr

경북대학교 전자전기공학부 박사과정 박승현

Session 11 AI Accelerators

이번 A-SSCC 2024의 11번째 세션, AI Accelerators에서는 대규모 신경망 처리 및 효율적인 학습 가속화를 위한 다양한 하드웨어 아키텍처와 최적화 기법들이 중점적으로 논의되었다. 세션의 논문들은 공통적으로 데이터 처리 효율성, 전력 및 면적 효율성, 확장성을 높이는 데 초점을 맞췄다. 또한 다양한 AI 모델과 애플리케이션을 지원하기 위한 하드웨어 최적화, 데이터 재사용 전략, 그리고 고효율 연산 구조 설계가 중요하게 다루어졌다.

#11-1



[그림 1] 제안하는 딥러닝 SoC의 시스템 아키텍처

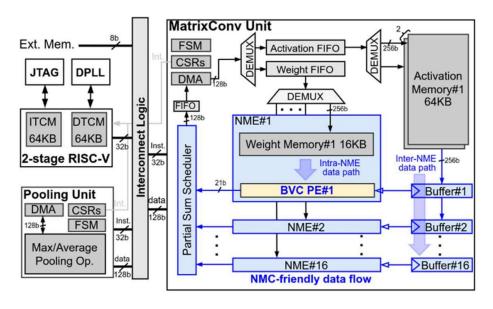
이 논문은 대규모 신경망의 높은 계산 요구 사항을 효과적으로 처리하기 위해 확장 가능하고 재구성 가능한 딥러닝 SoC를 제안한다. 멀티칩 솔루션의 한계로 지적되는 낮은 하드웨어 활용도와 데이터 이동의 비효율성을 해결하기 위해 채널 및 평면 기반 데이터 분할, 양방향 데이터 흐름, 그리고 희소성 데이터를 활용하는 최적화된 아키텍처를 설계하였다. 이를 통해 SoC는 칩 간 데이터 이동량을 최소화하고, 계산 효율성을 극대화하며,

대규모 신경망의 다양한 구조를 지원할 수 있게 되었다. 특히, 제안된 아키텍처는 다차원 하드웨어 병렬성을 적용하여 신경망의 깊은 층에서도 높은 칩 활용도를 유지한다.

SoC의 구조는 딥러닝 가속기, 출력 데이터 필터링 장치, Coretex-M3 MCU로 구성되며, 딥러닝 가속기는 4x4 텐서 프로세싱 요소(TPE) 배열로 구성된다. TPE는 희소성 데이터를 효율적으로 처리하기 위한 희소성 처리 코어(SPC)와 채널 그룹 유닛(CGU)을 포함한다. 또한, TPE 배열은 다차원 병렬성을 고려한 설계로 CNN, GCN, Transformer 모델에서 높은 성능을 발휘한다. 예를 들어, 희소성 데이터의 비율이 75% 이상인 경우 CNN과 Transformer 모델에서 최대 75%의 지연 시간을 줄였으며, GCN에서는 93%의 지연 시간 감소를 달성하였다. 추가적으로, 데이터 분할 및 양방향 데이터 흐름은 칩 간 데이터 충돌을 방지하고, 칩 활용도를 83%까지 향상시켜 16칩 시스템에서 12.6배의 처리량 향상을 달성하였다.

이 SoC는 28nm CMOS 공정으로 제작되었으며, 칩 면적은 7.8mm², 소비 전력은 10.4-273.8mW로, 최고 성능 9.83TOPS를 기록하였다. 면적 효율은 1.26TOPS/mm², 에너지 효율은 141.4TOPS/W로, 기존 멀티칩 시스템 대비 면적에서 5.7배에서 최대 21배, 에너지 효율에서 1.7배에서 45.6배 우수한 성능을 보였다. 이 연구는 고성능 및 고효율의 딥러닝 SoC 구현 가능성을 입증한다.

#11-3



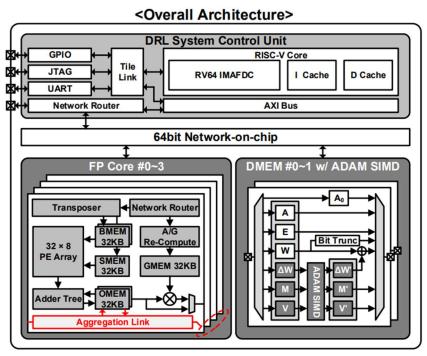
[그림 2] near-memory-computing 데이터플로우와 Booth-value-confined 다해상도 처리 아키텍처

이 논문은 엣지 AI 애플리케이션을 위해 높은 전력 효율과 처리 성능을 제공하는 Booth-value-confined(BVC) 가속기와 근접 메모리 컴퓨팅(NMC) 기반 프로세서를 제안한다. 기존의 von-Neumann 아키텍처와 비교해 NMC 기반 구조는 전력 효율에서 우위를 보였지만, 메모리 불연속성과 비효율적인 데이터 매핑 전략으로 인해 전체적인 지연과

소비 전력이 증가하는 문제가 있었다. 이를 해결하기 위해 본 연구는 BVC 근사 연산기, NMC 친화적인 데이터 흐름, 메모리 공간 연속성을 고려한 신경망 매핑 전략을 도입한다. 프로세서는 RISC-V 서브시스템, 디지털 PLL, 프로그래머블 풀링 및 매트릭스 연산 유닛으로 구성되며, BVC 근사 연산기가 포함된 16개의 NME가 주요 연산을 담당한다.

제안된 BVC 연산기는 Booth 알고리즘을 기반으로 3가지 정밀도(BVC3/6/8)를 지원하며, 전력과 면적을 기존 멀티프리시전 연산기 대비 각각 82%, 70% 절감한다. 이를 통해 각가중치의 재사용을 극대화하고 데이터 전송량을 줄여 시스템 레벨에서 효율성을 극대화한다. 또한, 가중치와 활성화 데이터를 재사용하며 메모리 공간을 연속적으로 유지하여메모리 액세스 지연을 최소화한다. VGG16 및 ViT-Tiny 모델에 대한 테스트 결과, 평균 전력 효율은 10.14-15.11 TOPS/W, 피크 효율은 12.92-29.11 TOPS/W를 기록하였으며, 메모리 지연을 최대 31.4% 감소시켰다. 22nm 공정으로 제작된 이 프로세서는 13.61-210.67mW의 전력에서 동작하며, Activation(INT4)과 Weight(BVC3)를 조합한 경우 최고 33.98 TOPS/W의 전력 효율을 달성하였다. 본 연구는 높은 전력 효율과 데이터 재사용성을 통해 엣지 AI 프로세서 설계에 있어 새로운 최적화 방향을 제시하였다.

#11-4



[그림 3] EnTADRL의 아키텍처

이 논문은 연속적이고 효율적인 심층 강화 학습(DRL) 가속화를 위한 SoC(System-on-Chip)인 EnTADRL을 제안한다. DRL은 게임 AI, 자율주행 등과 같은 시퀀셜 의사결정 문제

에서 인간 수준의 성능을 달성하지만, 복잡한 연산 구조와 높은 메모리 대역폭 요구 사항으로 인해 기존 AI 가속기에 직접 매핑하기 어렵다. 이를 해결하기 위해 EnTADRL은 세 가지 주요 최적화를 제안한다: 사용자 정의 명령어 세트 아키텍처(ISA), 네트워크-온-칩(NoC) 및 스트리밍 데이터플로우를 활용한 SoC 수준 최적화, 활성화 기울기 재계산 및 Tanh 양자화를 포함하는 부동소수점(FP) 코어 최적화, 그리고 파이프라인화된 ADAM SIMD를 활용한 효율적인 가중치 업데이트 지원이다.

제안된 EnTADRL은 RISC-V 기반의 DRL 시스템 제어 유닛, 4개의 FP 코어, 그리고 2개의 데이터 메모리(DMEM)로 구성된다. FP 코어는 활성화 기울기 재계산 유닛과 32×8 2D PE 어레이를 통합하여 메모리 발자국을 14.3% 감소시키고, 계산 효율성을 높인다. 또한, ADAM SIMD는 전용 메모리 구조와 파이프라인화된 연산 흐름을 통해 높은 메모리 대역 폭을 지원하며, 지연을 최소화한다. 스트리밍 데이터플로우와 이중 버퍼링 기법은 메모리 전송과 계산 사이의 병목현상을 완화하며, 전체 연산 주기를 23.2% 줄인다.

EnTADRL은 28nm CMOS 공정으로 제작되어 12.96mm²의 다이 면적과 2.8MB의 온칩 SRAM을 포함한다. 10MHz에서 200MHz의 동작 주파수와 0.65V에서 0.9V의 전압에서 작동하며, 최고 409.6 GFLOPS의 성능과 2.4 TFLOPS/W의 에너지 효율을 달성한다. 특히, DRL 가속기를 위한 기존 솔루션과 비교하여, EnTADRL은 손실 계산, 노이즈 삽입, ADAM 연산을 포함한 복잡한 DRL 연산을 온칩에서 완벽히 처리하여 실질적인 종단 간 가속화를 지원한다. 이 연구는 DRL을 위한 새로운 하드웨어 아키텍처의 가능성을 제시하며, 실제 환경에서의 DRL 성능 최적화를 크게 향상시킨다.

저자정보



박승현 박사과정 대학원생

● 소속 : 경북대학교

● 연구분야 : 딥러닝 가속기 설계● 이메일 : ijjh0435@gmail.com

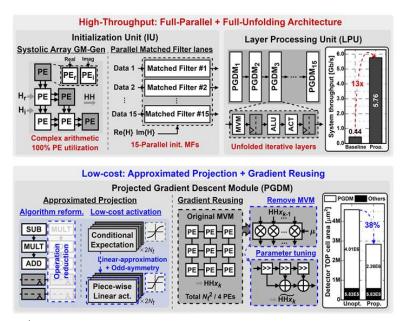
• 홈페이지 : https://ai-soc.github.io/

한국과학기술원 전기및전자공학부 석사과정 권재훈

Session 27 Security/Signal Processing Systems

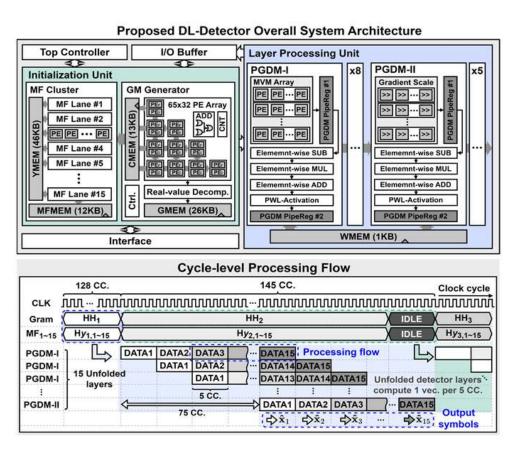
이번 2024 IEEE A-SSCC의 Session 27 Security/Signal Processing Systems에서는 SRAM 기반 Physically Unclonable Function (PUF) 설계, SOIPUF(Symmetrical Obfuscated Interconnection PUF) 기반 Hybrid Strong PUF 설계 및 Resistance 향상, DL 기반 efficient MIMO Detector 설계, HBM3와 NPU 기반의 Chiplet architecture를 위한 효율적인 Digital PHY 설계라는 주제로 총 4편의 논문이 발표되었다. 이 중 DL 기반 efficient MIMO Detector 설계, HBM3와 NPU 기반의 Chiplet architecture를 위한 효율적인 Digital PHY 설계에 대한 2개의 논문을 살펴보고자 한다.

#27-3 본 논문은 POSTECH에서 발표한 논문으로, Deel Learning (DL) 기반의 Massive multiple-input multiple-output (MIMO) uplink detector를 최초로 silicon으로 구현하였으며, 5.76 Gb/s의 throughput과 79.7 pJ/b의 energy efficiency을 달성했다. 128×32 MIMO 환경에서 4~256QAM modulation을 지원하며, 기존의 minimum mean square error (MMSE) 방식에 비해 4.3 dB의 성능 향상을 이뤄냈다. [그림 1]은 제안된 DL-detector의 high-throughput, low-cost를 위한 최적화 방법을 나타낸 그림이다.



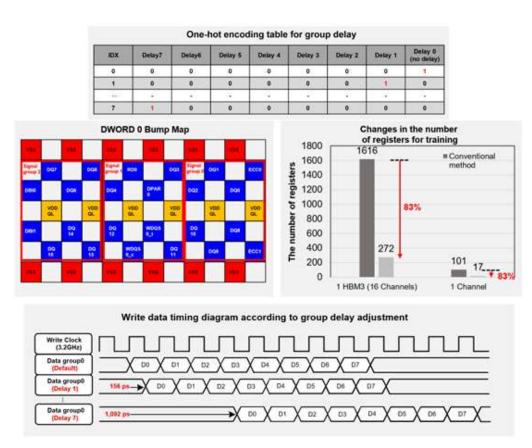
[그림 1] Overview of the proposed high-throughput and low-cost optimization strategies of DL-detector.

Contribution은 1번째로 Fully Unfolded DL-based MIMO Detector를 설계하여 병렬성을 높인 것이다. Unfolding technique은 loop를 제거하고 병렬화하는 설계 기술이고 이를 통 해 throughput을 증가시킬 수 있다. 이에 대한 자세한 내용은 textbook [1]에서 참고할 수 있다. 기존 직렬 방식의 DL 연산과 달리, 모든 DL layer의 연산을 병렬로 실행하여 13 배의 throughput 향상을 이끌어냈다. Deep Learning에 대한 자세한 설명은 textbook [2] 에서 참고할 수 있다. 2번째로 Initialization을 빠르게 할 수 있도록 한 것이다. 본 논문의 초기화 단계에서는 fully parallel MF와 gram-matrix generator를 사용해 128 cycles 내에 초기화를 완료한다. 이때 gram-matrix는 Symmetric Property를 활용하여 half-sized systolic array와 메모리만으로 계산할 수 있다. 또한 15 parallel MF lane을 통해 연속적인 입력 데이터를 병렬로 처리한다. 마지막으로 Approximate Projection과 Gradient Reusing 이다. Approximate Projection은 기존의 projection 연산을 최적화하여, linear connection 연산의 50% 감소를 달성했다. 또한 ReLU 기반 nonlinear unit을 단순화하여 divisor와 exponent 연산자를 제거했다. Gradient Reusing은 마지막 6개의 DL layer에서는 Gradient Reusing을 적용하여 새로운 gradient를 생성하는 것이다. 이를 통해 MVM 연산을 완전히 제거함으로써, 전체 연산 cost의 38% reduction을 달성했다. [그림 2]는 전체적인 chip architecture와 cycle-level로 processing flow를 나타낸 그림이다.



[그림 2] Overall chip architecture and cycle-level processing flow.

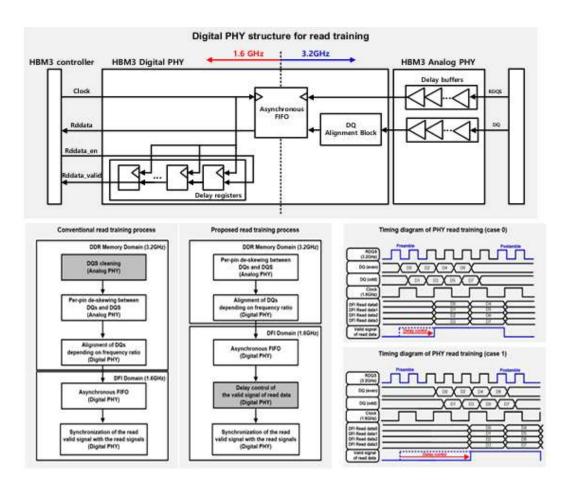
#27-4 본 논문은 AI SoC Research Division, Electronics and Telecommunication Research Institute (ETRI)에서 발표한 논문으로, HBM3 Digital PHY를 chiplet 기반 AI processor에 최적화한 설계에 대한 연구이다. 12nm CMOS 공정을 사용하였고 6.4Gb/s/pin의 데이터 전송 속도를 달성했다. 기존의 HBM2E와 비교하면, HBM3는 clock frequency가 1.2GHz에서 3.2GHz로 3배 향상되었는데, 이러한 고속 동작을 지원하기 위해 고속, 저전력의 Digital PHY 설계가 필요하여 본 논문에서 최적화를 제시한 것이다. chiplet architecture에 대한 내용은 textbook [3]에서 참고할 수 있다. 또한 textbook [4]에서 PHY 설계의 개념과, timing 문제 및 설계 기법 등을 참고할 수 있다.



[그림 3] Grouping signal delays based on HBM3 bump map.

Contribution은 1번째로 group delay 방식을 개선한 것이다. 기존에는 data, command, address 신호에 대해 개별적으로 delay를 조정했지만, 본 논문에서는 HBM3 bump map을 기반으로 grouped signal 단위로 delay를 조절하는 방식을 제안했다. 이를 통해 delay register 수를 83% 줄였고, PHY training time도 단축시킬 수 있었다. 이에 대한 figure는 [그림 3]을 통해 제시되었다. 2번째로 Async FIFO 구조를 단순화한 것이다. 기존의 async FIFO는 full flag로 상태를 모니터링해야 했지만 본 논문에서는 full flag 없이도 동작 가능한 async FIFO 구조를 제안했다. FIFO의 data depth를 8로 설정하여, full state activation을

방지했다. 또한 루프백 카운터를 latency가 적은 lookup table(LUT)로 대체해 고속 신호 동작에 적합한 구조를 설계했다. 3번째로 새로운 Read Training 알고리즘을 제안한 것이다. 기존의 RDQS cleaning을 없애고, FIFO의 clock 신호로 RDQS를 직접 연결했다. 또한 high-frequency domain(3.2GHz) 대신 low-frequency domain (1.6GHz)에서 delay 제어를 수행해 전력 소모를 줄이고 알고리즘 복잡도를 낮췄다. 이에 대한 figure는 [그림 4]에서볼 수 있다.



[그림 4] Proposed HBM3 PHY structure for read training.

참고문헌

- [1] Parhi, Keshab K. VLSI digital signal processing systems: design and implementation. John Wiley & Sons, 2007.
- [2] Goodfellow, Ian. "Deep learning-ian goodfellow, yoshua bengio, aaron courville." Adapt. Comput. Mach. Learn (2016).
- [3] Manna, Kanchan, and Jimson Mathew. Design and Test Strategies for 2D/3D Integration for NoC-

based Multicore Architectures. Springer Nature, 2019.

[4] Khan, Shoab Ahmed. Digital design of signal processing systems: a practical approach. John Wiley & Sons, 2011.

저자정보



권재훈 석사과정 대학원생

● 소속 : 한국과학기술원 전기및전자공학부

• 연구분야 : Digital Circuit Design, ECC Hardware Design

● 이메일 : jhkwon@ics.kaist.ac.kr ● 홈페이지 : https://ics.kaist.ac.kr/

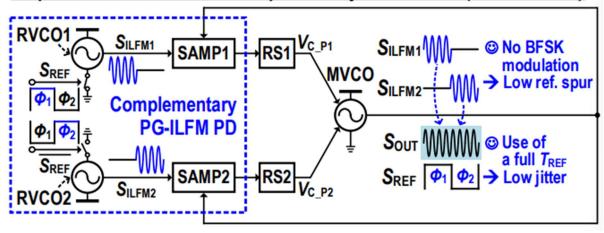
단국대학교 파운드리공학부 석사과정 조혁준

Session 13 Low-Jitter PLLs and Low-Phase-Noise Oscillators

Session 13에서는 Low-Jitter PLL과 Low-Phase-Noise Oscillator를 주제로 한 총 4편의 논문이 발표되었다. 이 중, 특히 Low-Jitter을 강조한 PLL과 관련된 두 편의 논문을 살펴보고자 한다. 하나는 W 및 D 대역의 Sub-THz 주파수를 다룬 논문이며, 다른 하나는 십여 GHz 범위에서 동작하는 Low-Jitter PLL 설계에 관한 연구이다.

#13-1 본 논문은 서울대학교, KAIST에서 발표한 논문으로, 두개의 상보적으로 작동하는 RVCO를 활용하여 BFSK Modulation과 Reference Spur를 억제하고, 낮은 Jitter 및 높은 KPD를 달성하는 PLL 구조를 제안한다. 서로 반대 위상의 Reference Clock을 통해 RVCO를 위상 정렬하며 RVCO의 Phase Noise를 제거합니다. 또한 두개의 RVCO를 상보적으로 동작 시켜 TREF 전체에서 Phase Error를 감지하며 이를 통해 KPD를 향상시켰다. 추가적으로, Frequency-offset canceller와 Reference Clock의 Duty Cycle Corrector 같은 회로를 포함하여 성능을 더욱 개선하였다. 그 결과 119.5GHz 출력 신호에서 65.6fs의 RMS Jitter, -52dBc의 Reference Spur를 달성하였다.

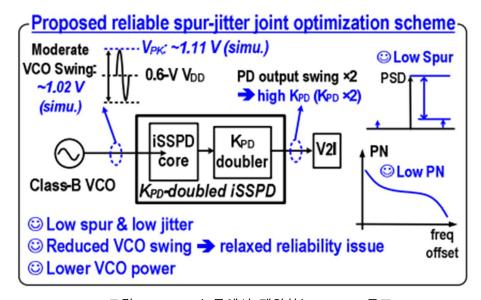
Proposed D-Band PLL with Complementary PG-ILFM PD (CPG-ILFM PD)



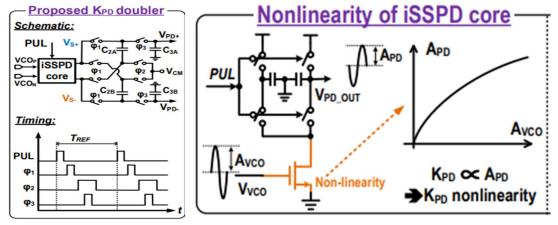
[그림 1] #13-1 논문에서 제안하는 Complementary PG-ILFM PD 구조.

#13-2 본 논문은 중국 과학원 반도체 연구소에서 발표한 논문으로, 낮은 Jitter와 Spur를 동시에 달성하기 위해 적절한 VCO Swing과 KPD doubler를 사용하는 SS-PLL을 제안한다.

제안된 구조는 기존 방식이 높은 VCO Swing을 요구했던 것과 달리, KPD doubler를 사용하여 ISSPD Core의 출력 Swing을 두 배로 증폭함으로써 높은 KPD를 확보하면서도 낮은 VCO 전력으로 동작할 수 있도록 설계되었다. KPD Doubler는 커패시터와 스위치를 활용한 Passive 방식으로 설계되어 Phase Noise와 Jitter 성능에 부정적인 영향을 미치지 않는다. 동시에 낮아진 VCO Swing 덕분에 ISSPD Core에서 MOSFET 크기를 줄여 VCO로 전달되는 간섭을 최소화하고, Spur 성능을 개선하였다. 출력 Swing 감소로 인한 KPD 저하는 KPD doubler를 통해 보상하여 비선형성을 억제하고 신호 품질을 유지한다. 이 결과, 제안된 PLL은 -100.63 dBc의 Spur, 32.1 fs의 Integrated jitter (1 kHz to 100 MHz), 그리고 -261.4 dB의 FOM(성능 지수)를 달성하였다.



[그림 2] #13-2 논문에서 제안하는 SS-PLL 구조



[그림 3] #13-2 논문에서 제안하는 K_{PD} doubler 와 ISSPD MOSFET의 Nonlinearity 특성

저자정보



조혁준 석사과정 대학원생

● 소속 : 단국대학교 파운드리공학부

● 연구분야: Biomedical Circuits ● 이메일: sdi5236@dankook.ac.kr

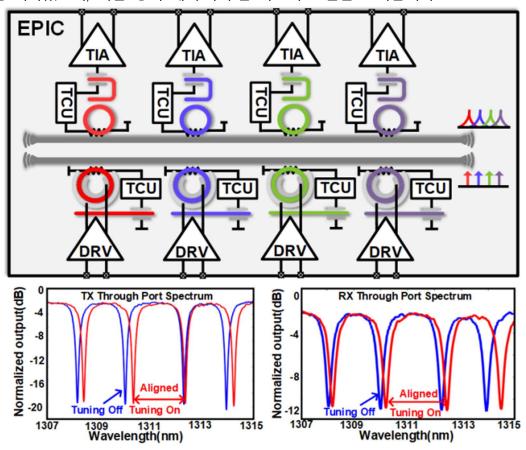
● 홈페이지: https://sites.google.com/dankook.ac.kr/acs-lab

서울대학교 전기정보공학부 박사과정 박현준

Session 22 Advanced Wireline Transceivers

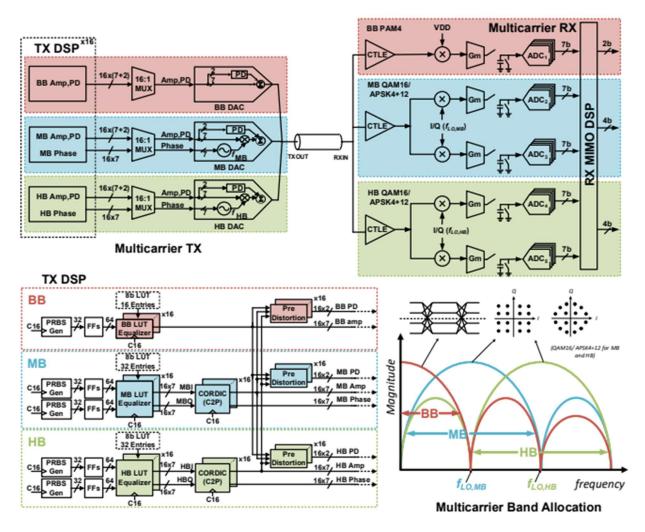
이번 ASSCC 2024의 Session 22에서는 High-Speed Wireline Transceivers에 관련된 4편의 논문이 발표되었다. 이 중 2편의 논문(#22-1, #22-2)은 고속 데이터 전송에서의 에너지효율과 신뢰성을 극대화하기 위한 설계 기술에 대해 다루었다.

#22-1 이 논문은 중국 반도체 연구소에서 발표한 연구로, 45nm SOI-CMOS 공정을 이용해 200Gb/s 전송 속도와 3.5pJ/bit의 에너지 효율을 달성한 WDM 실리콘 포토닉 트랜시 버를 제안했다. 하이브리드 방식의 기생 효과를 줄이기 위해 광학 및 전자 회로를 단일칩에 통합했으며, 비대칭 구동단과 TIA 설계를 통해 높은 삽입 손실에서의 비선형 반응을 완화하고 대역폭 손실을 보완하여, 50Gb/s에서 2pJ/bit, 64Gb/s에서 1.6pJ/bit의 에너지효율을 기록했으며, 기존 방식 대비 우수한 속도와 효율을 보여줍니다.



[그림 1] OTRX 시스템 Diagram 및 TRX TP 스펙트럼 측정결과

#22-2 은 텍사스 A&M 대학교와 광주과학기술원(GIST)에서 발표한 논문으로, 56Gbps 속도와 7.82pJ/bit의 에너지 효율을 달성한 ADC/DAC 기반 다중반송파(Multicarrier) 트랜시 버를 제안하였다. 40.8dB의 고손실 채널에서도 동작 가능한 RX MIMO-DSP를 설계하여 기존 방식 대비 우수한 지터 내성과 스펙트럼 효율을 제공했습니다. 트랜시버는 다중 반송파를 이용하여 진폭 및 위상 변조를 수행하며, 5.6GS/s의 낮은 대역폭에서 각 반송파를 처리함으로써 심볼 시간을 5배 늘리고 지터 내성을 향상시켰습니다. 송신단은 Polar DAC 기반으로 구성되어 PAM4와 QAM16, APSK4+12와 같은 변조를 지원하며, 수신단은 CTLE, 혼합기, 적분기 및 5.6GS/s ADC로 구성되어 채널 손실을 보정한다. RX MIMO-DSP는 ISI와 ICI를 FIR 필터로 제거하며, 필터 계수는 sign-sign LMS 알고리즘으로 학습한다. 측정결과, 트랜시버는 1.21ps RMS 지터 조건에서도 <10^-4의 BER을 유지했으며, baseband와 mid/high-band에서 각각 14mV의 전압 마진과 -18.28dB EVM을 달성했다. 56Gbps 전체시스템은 438mW를 소모하며, 유사한 데이터 속도를 가진 기존 PAM4/DMT 기반 트랜시버보다 우수한 지터 내성과 에너지 효율을 보여준다.



[그림 2] 다중 캐리어 트랜시버 구조도

저자정보



박현준 박사과정 대학원생

● 소속 : 서울대학교

• 연구분야 : HBM, Chord Signaling, Information Theory

• 이메일 : spp098@snu.ac.kr

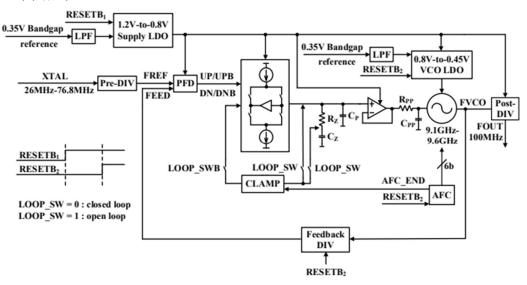
● 홈페이지: https://sites.google.com/view/wschoi?pli=1

서울대학교 전기정보공학부 박사과정 박현준

Session 26 Clocking and Memory Interfaces

이번 ASSCC 2024의 Session 26에서는 Advanced Clocking 및 Memory Interface 기술에 관한 4편의 논문이 발표되었다. 이 중 2편의 논문(#26-1, #26-2)은 고속 데이터 전송과 클럭 안정성을 위한 PLL 및 DDR5 RCD IC 설계 기술을 다루었다.

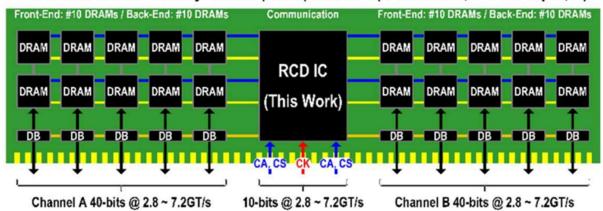
#26-1 삼성반도체 인도연구소와 삼성전자 화성에서 발표한 논문으로, PCIe Gen6 레퍼런스 클럭 생성을 위해 설계된 저지터 PLL을 제안하였다. 2nm MBCFET 공정을 사용하여 - 45℃에서 125℃까지의 온도 변화를 지원하며, 45-fsrms 누적 지터와 6μs의 빠른 락 타임을 달성했다. 기존 PCIe 시스템은 저지터 TXCO를 사용하지만, 높은 비용과 면적 요구 사항이 단점인데, 이를 해결하기 위해 저비용 온칩 PLL 설계를 제안하며, 1.2V 공급 전압에서 동작하는 아날로그 블록을 안정적으로 설계하기 위해 새로운 신뢰성 중심 아날로그설계 기법과 고급 레이아웃 기술을 적용했다. 제안된 PLL은 저대역폭 설계를 통해 PCIe 규격 내에서 지터 피킹을 방지하며, LC-VCO를 사용해 플리커 노이즈를 줄이고 지터 성능을 개선했다. 또한, 공급 노이즈로 인한 deterministic 지터를 완화하기 위해 2단 Cascaded 레귤레이터 구조를 도입했다. 빠른 락 타임을 위해 24 FREF 사이클 만에 0.17%의 해상도를 달성하는 bang-bang AFC 기법을 적용하여 기존 방식보다 25배 빠른 성능을 구현했습니다. 측정 결과, 제안된 PLL은 100MHz에서 ±4%의 세밀한 튜닝 범위와 ±2%의 온도에 따른 주파수 변화를 지원하며, 기존 연구 대비 뛰어난 전력, 면적, 및 성능 지수를 보여주었다.



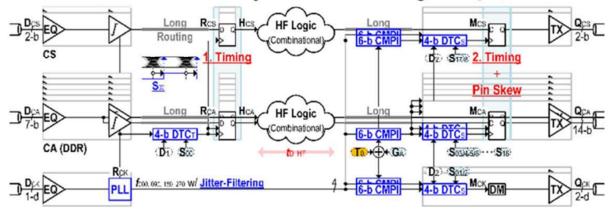
[그림 2] 제안된 Low jitter PLL

#26-2 삼성전자와 MIT에서 발표한 논문으로, DDR5 DIMM에서 사용되는 7.2GT/s 레지스터링 클럭 드라이버(RCD) IC 설계에서 고속 데이터 전송과 안정적인 동작을 위해 핀 간스큐(<±1.6ps) 및 병렬 데이터 타이밍을 보정하는 전략을 제안했다. 제안된 RCD IC는 28nm CMOS 공정을 기반으로 제작되었으며, PI-내장 PLL과 PVT 변화에 강한 DTC를 통해 DCK 입력의 지터(4.72psrms)를 0.20psrms로 줄여 시스템 안정성을 크게 향상시켰다. 또한, A-DAC 기반 CMPI는 위상 및 진폭 왜곡을 보정해 INLp-p를 1.06LSB로 개선하고, ±1.6ps의 스큐 보정을 달성했다. 고속 논리와 TX를 위한 PI-오프셋 코드 및 DTC 기반의 2단계 스큐 보정은 DRAM의 요구사항을 만족시키는 정확한 데이터 전달을 보장한다. 측정 결과, 제안된 RCD IC는 3.6GHz에서 24배의 지터 감소를 달성했으며, DTC는 0.24mW/GHz의 우수한 에너지 효율과 높은 선형성을 보여줬다. 이를 통해, 차세대 DDR5 메모리 모듈의 신뢰성과 성능을 크게 향상시킨 인터페이스로 제안되었다.

DDR5 Dual In-Line Memory Module (DIMM) Interface (2-Channel A, B / 2-Group A, B)



Architecture of RCD IC with Timing and Skew Calibration @ 1-Group A, 1-Channel A



[그림 2] 7.2GT/s RCD IC for a DDR5 DIMM interface

저자정보



박현준 박사과정 대학원생

● 소속 : 서울대학교

• 연구분야 : HBM, Chord Signaling, Information Theory

• 이메일 : spp098@snu.ac.kr

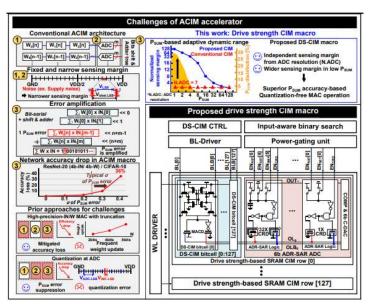
● 홈페이지: https://sites.google.com/view/wschoi?pli=1

KAIST 전기 및 전자공학부 석사과정 윤지원

Session 8 SRAM-Based Computing-in-Memory

이번 2024 IEEE ASSCC의 Session 3은 SRAM-Based Computation-in-Memory라는 주제로 총 5편의 논문이 발표되었다. 이 세션에서는 메모리 내에서 데이터를 직접 처리함으로써 데이터 이동과 연산에 소모되는 에너지와 시간을 크게 줄일 수 있도록 하는 혁신적인 아키텍처를 제안하는 것에 중점을 두었다.

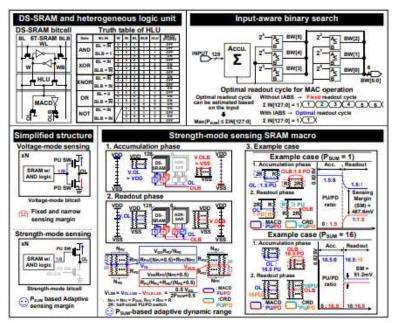
#8-3 본 논문은 DGIST, 고려대학교에서 발표한 논문으로, Computation In Memory (CIM)에서 발생하는 Partial Sum(PSUM) 오류를 해결하기 위해, 고정밀 입력/가중치 사용 및 MAC 출력을 절단하거나, 저해상도 ADC를 활용하여 Partial Sum을 양자화하는 기존의 해결 방법대신 Drive Strength 기반 SRAM CIM 접근법을 제안함으로써 정확도 손실 문제점을 해결하고자 하였다. 제안된 DS-SRAM CIM macro는, 128x128b DS-CIM bitcell array, Input-Aware Binary Search(IABS), Power Gating Unit (PGU), 128 row-wise ADR-SAR ADCs, 그리고 DS-CIM controller로 이루어져 있다.



[그림 1] 아날로그 CIM macro 설계의 난점 및 제안된 DS-SRAM CIM

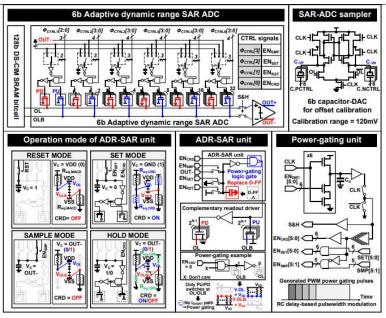
비트셀은 6T SRAM, Heterogeneous Logic Unit (HLU), MAC 드라이버(MACD)로 구성되며, PSUM 값은 연결된 PU 스위치 수로 표현된다. Strength 기반 MAC 연산은 저 PSUM 영

역에서 높은 감지 여유(SM)를 제공하며, 연산은 축적 단계와 읽기 단계로 수행된다. 축적 단계에서 비트별 AND 연산이 수행되고, 읽기 단계에서는 6비트 ADR-SAR ADC가 CRD와 COMP를 통해 이진 탐색으로 PSUM 출력을 읽는다. IABS는 입력 합계를 기반으로 ADC 변환 주기를 결정하고, 추가적인 스위치를 통해 충분한 SM을 확보한다. PSUM 값이 입력 합계를 초과할 경우 IABS를 비활성화해 읽기를 수행한다



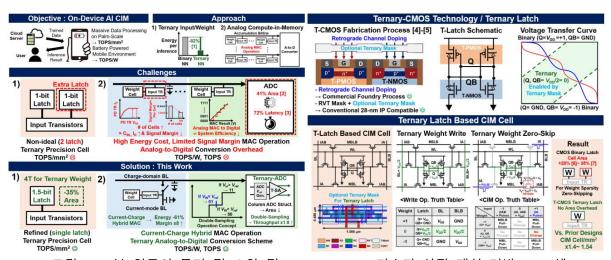
[그림 2] DS-SRAM bitcell의 Truth table과 IABS, 그리고 각 mode에서의 Sensing margin 분석

ADR-SAR ADC는 Partial Sum 출력을 읽기 위해 설계된 6비트 ADC로, 샘플링 커패시터와 NAND, 제어 스위치로 구성되며 Reset, Set, Sample, Hold 네 가지 모드로 작동한다. Reset 모드에서는 샘플링 커패시터를 초기화하고, Set 모드에서 CRD가 OL/OLB에 연결되며, Sample 모드에서는 COMP 출력을 커패시터에 저장하고 Hold 모드에서 데이터를 유지한다. IABS는 최적의 ADC 변환 주기를 예측하고 PGU는 이에 맞춰 전력 게이팅 펄스를생성해 전력 소모를 절감한다. 이를 통해 기존 디지털 SAR ADC 대비 면적을 35% 줄이고, 전력 소모를 최대 40% 절감한다. 정밀도는 샘플링 커패시터의 불균형 보정과 하이브리드 SAR 구조로 달성되며, 변환 정확도는 98.5%에 이른다. 또한 노이즈 저감 메커니즘과 고정밀 COMP를 통해 SNR이 기존 대비 25% 이상 향상된다. 결과적으로 ADR-SAR ADC는 낮은 전력 소모와 높은 정확도를 유지해 메모리 기반 연산에서 PSUM의 오류 없는 읽기와 에너지 효율적인 성능을 제공한다.



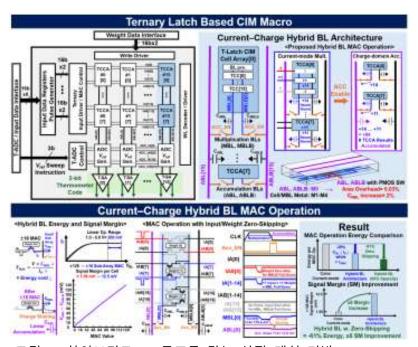
[그림 3] 제안된 ADR-SAR ADC, ADR-SAR 유닛의 동작 모드 및 Power Gating 유닛

#8-5 본 논문은 UNIST 와 삼성전자가 공동 발표한 연구로, 28nm T-CMOS 기술을 기반으로 한 단일 latch 삼진 셀과 에너지 효율 및 신호 마진을 개선한 Double-Sampling Ternary ADC 를 제안한다. 리소스가 제한된 온디바이스 AI 환경에서는 면적효율과 전력 효율이 핵심 요소로, 네트워크 양자화와 삼진 입력/가중치 (-1, 0, +1)를 사용하면 에너지를 82% 절감할 수 있다. 그러나 삼진 가중치를 지원하기 위한 추가적인 Latch 와 아날로그 CIM 에서 MAC 연산이 증가함에 따라 효율성 저하 문제가 발생한다. 이를 해결하기 위해 T-CMOS 기술을 활용한 단일 latch 삼진 셀 (T-latch)이 제안된다. 이설계는 29%~35%의 면적 오버헤드를 피하고, CIM 셀 밀도를 1.4~1.54 배 향상시켜 면적효율성과 전력 효율을 크게 개선함으로써, 제한된 리소스에서 높은 성능을 유지하는 데중요한 역할을 한다.



[그림 1,2] 본 연구의 동기 및 요약 및 Tenary-CMOS 기술과 삼진 래치 기반 CIM 셀

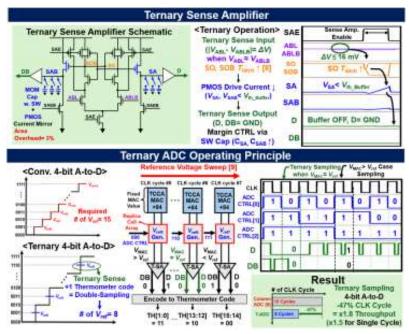
제안된 삼진 래치 기반 CIM 매크로는 128 개의 삼진 CIM 셀 배열(TCCA)로 구성되며, 각 열은 8 개의 TCCA 와 이중 샘플링 삼진 ADC 를 포함한다. 입력 데이터는 레지스터에 저장되며 값이 "+1" 또는 "-1"일 때 펄스를 생성한다. 상단과 우측에는 가중치 업데이트를 위한 쓰기 주변 회로(드라이버 및 디코더)가 위치하고, 좌측 하단에는 삼진 ADC 제어부와 3 비트 제어 코드 입력부가 배치된다. 128 개의 TCCA 는 전류-전하하이브리드 비트라인(BL) 구조를 통해 MAC 연산을 수행한다. 기존 단일 전류 모드 비트라인 대신, 8 개의 전류 모드 곱셈 비트라인(MBL)과 1 개의 전하 도메인 누적 비트라인(ABL)으로 분리해 에너지 효율과 신호 마진 문제를 해결했다. 각 8 개의 TCCA 는 전류 모드 곱셈을 수행한 후 ±16 의 MAC 결과를 8 개의 MBL 에 저장하고, 이후 PMOS 스위치를 통해 ABL 에 ±128 의 MAC 결과를 누적한다. 이 구조는 면적 오버헤드 없이 신호 마진을 8 배 향상시키며, 전력 소모를 줄인다. 또한, 입력과 가중치가 0 일 때 동적 전력을 소모하지 않는 zero-skipping 기능을 통해 에너지 절감을 극대화한다.



[그림 3] 하이브리드 BL 구조를 갖는 삼진 래치 기반 CIM macro

삼진 센스 증폭기(T-SA)는 기존 센스 증폭기에 삼진 감지($V_{MAC} \approx V_{REF}$) 기능을 추가해 이중 샘플링을 수행한다. 이를 통해 기준 전압(V_{REF}) 단계를 기존 15 에서 8 로 줄여 처리량을 기존 방식 대비 1.5 배, 1.8 배 증가시켰다. T-SA 는 컬럼 피치에 맞게 설계되어 면적 오버헤드 없이 각 열에 통합되며, 효율적인 ADC 변환을 지원한다. 기준 전압은 TCCA 하단에 위치한 Replica T-Latch CIM 셀 배열을 통해 생성되고, 8 개의 연속적인 클럭 사이클 동안 이중 샘플링으로 2 개의 Thermometer 코드를 생성한다. 측정 결과, 제안된 CIM 매크로는 단일 클럭 사이클에서 하이브리드 MAC 연산과 삼진 A-to-D 변환을 성공적으로 수행했으며, 167 MHz 클럭 주파수에서 높은 성능을 보였다. 면적당

143 TOPS/mm²와 전력당 892 TOPS/W 를 달성했고, 비트당 정규화 시 323 TOPS/mm²/b 와 2,007 TOPS/W/b 성능을 기록했다. 이는 삼진 래치 기반 CIM 셀, 이중 샘플링 삼진 ADC, 하이브리드 비트라인 구조, 그리고 zero-skipping MAC 연산 덕분이다. 결론적으로, 제안된 CIM 매크로는 온 디바이스 AI 환경에서 높은 에너지 효율과 신호 마진을 달성하며, 면적과 전력 소모 측면에서도 뛰어난 성능을 보여준다.



[그림 4] 삼진 센스 증폭기 기반 이중 샘플링 삼진 ADC

저자정보



윤지원 석사과정 대학원생

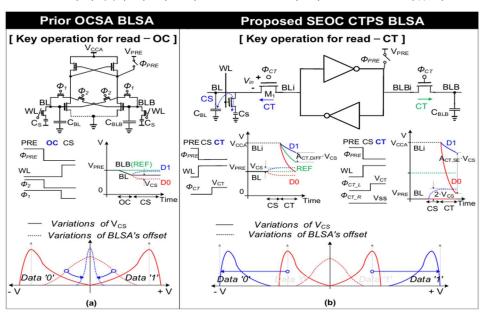
소속: 한국과학기술원 (KAIST)
연구분야: 디지털 회로 설계
이메일: jwyoon@kaist.ac.kr
홈페이지: https://idec.or.k

고려대학교 전기전자공학과 박사과정 한창우

Session 28: Advanced Memory Technology

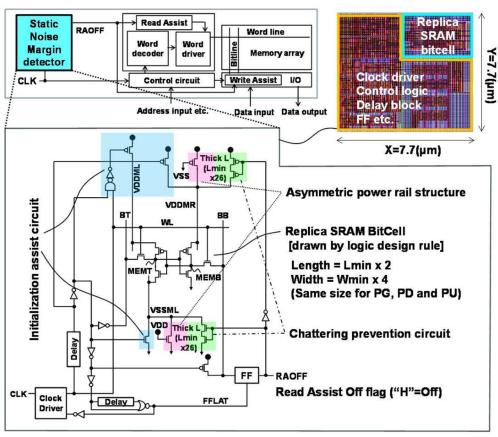
이번 A-SSCC 2024의 Session 28에서는 차세대 메모리 기술을 중심으로 총 5편의 논문이 발표되었다. 저전력, 고속 동작, 안정성을 목표로 한 다양한 메모리 기술이 소개되었으며, 특히 DRAM, SRAM과 같은 메모리 솔루션의 성능 개선과 전력 최적화 방안이 집중적으로 논의되었다. Session 28 안에서도, DRAM과 SRAM의 주요 기술을 다루는 3편의 논문을 살펴보고자 한다.

#28-1 본 논문은 삼성전자에서 발표한 논문으로, Sub-1V DRAM에서 bit-line sense amplifier (BLSA)의 성능을 개선하기 위한 single-ended offset compensation BLSA (SEOC BLSA)를 제안한다. DRAM의 에너지 효율을 위한 공급 전압이 낮아지면서 기존 cross-couple latch 기반의 BLSA는 저장 용량 감소와 트랜지스터의 임계 전압 변동성 증가로 인한 신뢰성 문제를 겪었지만, SEOC BLSA의 도입으로 앞선 문제를 해결하였다. SEOC BLSA는 ground precharge (GND PRE)와 charge-transfer (CT) topology를 도입하여 데이터 sensitivity를 대폭 향상시켰으며, PVT (프로세스, 전압, 온도) 변화에도 안정적인 동작을 보였다. 특히 데이터가 모두 0일 때 전류 소비를 최대 63% 절감했고, 기존 방식보다 최소 공급 전압을 0.85V까지 낮추어 저전력 DRAM 운영에 적합함을 입증하였다.



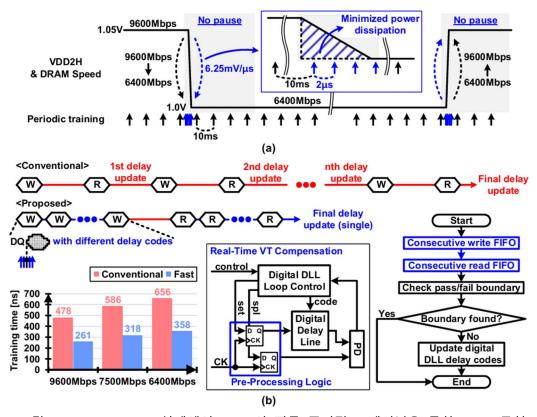
[그림 1] 기존 BLSA 구조 및 제안된 SEOC CTPS BLSA 구조 비교

#28-3 본 논문은 르네사스 일렉트로닉스에서 발표한 논문으로, High Density (HD) SRAM 의 static noise margin (SNM)을 감지하는 센서를 활용한 영역 효율적인 어시스트 기술을 제안한다. SRAM은 PVT (프로세스, 전압, 온도) 조건에 아주 민감하며, 특히 HD 셀에서는 SRAM의 SNM이 감소되는 등 안정성 저하 문제가 심각한 상황에 직면해왔다. 이를 보완하기 위해 word line under drive (WLUD)와 같은 어시스트 기술을 사용했지만, 이는 전력소비가 크고 액세스 시간을 느리게 만드는 단점이 존재했다. 또한, 기존 PVT 센서를 활용한 방식은 큰 면적을 차지하고 초기화 시간이 길어 실용성이 떨어지는 단점까지 존재했다. 본 논문에서는 앞선 문제를 모두 해결할 수 있는, 최소 안정성 셀을 복제한 구조의 SNM Detector를 제안하였고, 이를 통해 불필요한 어시스트를 제거해 SRAM의 액세스 시간을 25% 개선하고, 전력 소모를 읽기 측면에서 32.8%, 쓰기 측면에서 24.4% 감소시키는 데 성공했다. 또한, 0.015%의 무시할만한 작은 면적 오버헤드를 가졌으며, 온도 변화에 따른 어시스트 활성화/비활성화를 조절하는 효율성 또한 제공하였다. 결과적으로, 제안된 SRAM의 SNM Detector는 HD SRAM에서 공정 및 온도 변화에 따른 불안정성을 해결하고, 전력 효율과 성능을 동시에 개선할 수 있는 실용적인 솔루션을 제시할 수 있다.



[그림 2] 제안된 SNM Detector 회로도 및 레이아웃

#28-4 본 논문은 삼성전자에서 발표한 논문으로, LPDDR5X PHY 설계에서 adaptive driver strength control (ADSC)와 빠른 주기적 트레이닝(fast periodic training)을 통해 완전한 dynamic voltage and frequency scaling (DVFS)을 구현하여 메모리 성능을 유지하면서도 전력 소비를 줄이는 방법을 제안한다. 제안된 설계는 DRAM 전원 전압 (VDD2H)을 1.05V에서 1.0V로 낮추는 등 전력 절감을 달성하면서도, 전압 변동으로 인해 발생할 수 있는 쓰기 valid window margin (VWM) 저하를 극복할 수 있다. 이를 위해 빠른 주기적 트레이닝 방식을 적용하여 DVFS 전환 중에도 메모리 접근 차단(blackout) 없이 안정적인 동작을 보장한다. ADSC는 데이터 속도에 따라 드라이버 강도를 조정하여 전력 효율을 최적화하며, 고속 데이터 전송에서도 안정적인 신호 무결성을 제공한다. 실험 결과, DRAM 전력 소비는 기존 DVFSC 방식 대비 14% 이상 감소했고, 6400Mbps에서 쓰기 VWM이 26.4%, 2700Mbps에서 18.9% 개선되었다. 또한 PHY와 I/O 전력 소비가 각각 15.18%, 23.31% 감소하며, 읽기/쓰기 성능도 향상되었다. 이 기술은 고속 데이터 처리와 저전력을 모두 만족시키는 차세대 LPDDR5X DRAM 설계의 중요한 요소로 평가된다.



[그림 3] LPDDR5X PHY 설계에서 ADSC와 빠른 주기적 트레이닝을 통한 DVFS 구현

저자정보



한창우 박사과정 대학원생

● 소속 : 고려대학교 전기전자공학과

● 연구분야 : 차세대 반도체 소자 및 회로

• 이메일 : cwoo0105@naver.com

● 홈페이지: https://sites.google.com/view/kudclab

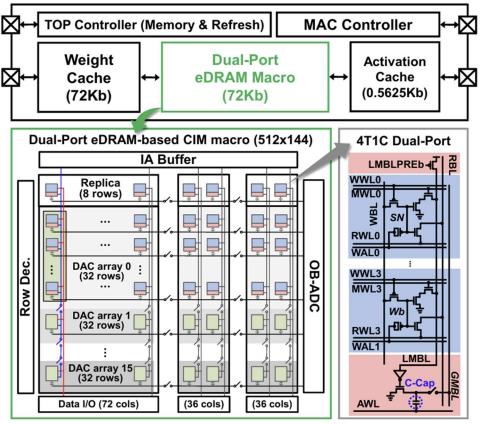
2024 A-SSCC Review

KAIST 전기및전자공학부 박사과정 엄소연

Session 25 High Density Computing-In-Memory

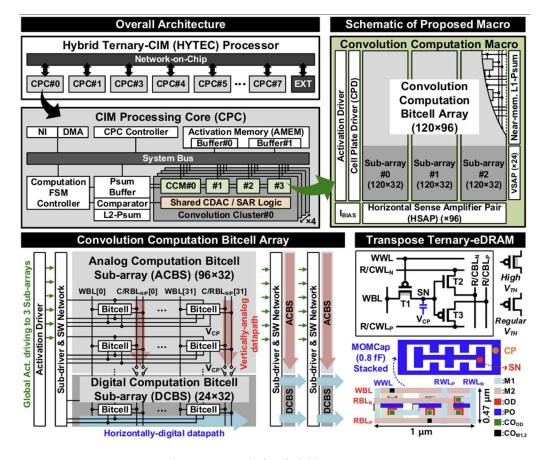
이번 2024 IEEE ASSCC의 Session 25는 High Density Computing-In-Memory 라는 주제로 총 4편의 논문이 발표되었다. 이 세션에서는 eDRAM, ROM, Flash 기반 Compute-in-Memory (CIM) 가속기와 새로운 데이터 변환 기법이 소개되었으며, FPGA를 활용한 다양한 AI 응용 사례를 중심으로 성능과 효율성을 높이기 위한 기술적 접근이 다뤄졌다. 본후기를 통해 3개의 논문에 대해 상세히 살펴보고자 한다.

#25-1은 연세대학교에서 발표된 Dual-Port eDRAM Compute-in-Memory (CIM) 가속기로, eDRAM 기반 메모리와 연산을 통합하여 성능과 전력 효율을 극대화한 새로운 아키텍처를 제안한다. 4T1C 듀얼 포트 구조를 채택하여 데이터 리프레시와 CIM 연산을 병렬로수행함으로써 기존의 eDRAM CIM 구조에서 발생하는 성능 병목과 리프레시 오버헤드 문제를 효과적으로 해결했다. Adaptive Refresh Tracking (ART)는 PVT(공정, 전압, 온도) 조건에 따라 최적의 리프레시 주기를 동적으로 조정하여 불필요한 리프레시를 줄이는 데 중점을 두었다. 이를 통해 리프레시 전력을 최대 60% 절감하고 시스템 전반의 에너지 효율을 높이는 데 성공했다. 또한, Data Conversion Reduction Scheme (DCRS)을 도입하여 DAC와 ADC의 면적 및 전력 소모를 최소화하였다. DAC와 ADC는 기존 CIM 시스템에서높은 에너지 소비와 신호 마진 저하를 초래했으나, DCRS는 데이터 변환 프로세스를 간소화하고 하드웨어 자원을 줄이는 방식으로 이러한 문제를 해결하였다. 신호 마진은 8.9배향상되었고, 메모리 효율은 이전 설계 대비 90.8% 증가했다. 이 가속기의 성능은 CIFAR-10 데이터셋과 VGG-16 모델로 평가되었다. Dual-Port eDRAM CIM은 평균적으로 기존 설계 대비 27.5% 더 높은 처리량을 제공하며, 전체 전력 오버헤드는 3.16%, 면적 오버헤드는 4.2% 증가에 불과했다.



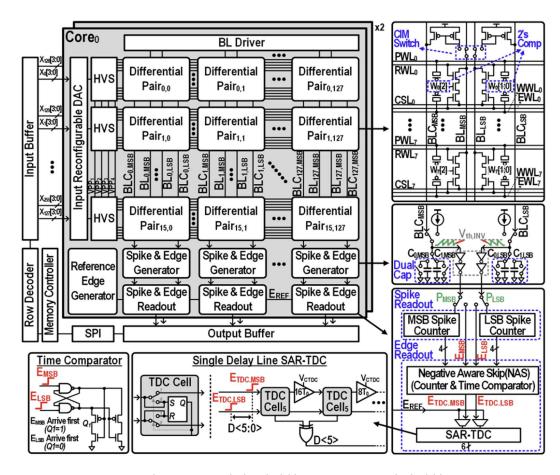
[그림 1] #25.1에서 제안한 DPe-CIM 가속기 구조

#25-2는 UNIST에서 발표한 논문으로, Hybrid-domain Ternary CIM 프로세서를 소개한다. 본 논문은 ternary neural network(TNN) 기반의 AI 연산에서 에너지 효율성을 극대화한 설계로 주목받았다. 본 연구는 eDRAM 기반 고밀도 셀과 수직 아날로그 및 수평 디지털 연산을 결합한 하이브리드 연산 구조를 통해 기존 CIM 구조의 한계를 극복하고 성능을 대폭 향상시켰다. 수직 아날로그 MAC 연산의 선형성을 높이기 위해 게이트 전압 바이어 싱 기법을 적용하여 연산 정확도를 86.8% 향상시켰으며, 1.5-bit/cycle SAR ADC를 사용하여 ADC 면적을 50% 줄였다. 수평 디지털 Psum 연산은 메모리에서 데이터를 효율적으로 분배하고 병렬 연산을 수행하여 데이터 병목 현상을 방지했다. 이러한 설계는 하드웨어 자원의 활용도를 극대화하면서도 높은 처리량을 유지하도록 도왔다. TNN 모델로 CIFAR-10 데이터셋을 평가한 결과, Hybrid-domain CIM은 89.2%의 높은 정확도를 유지하면서도 기존 eDRAM CIM 대비 4.88배 높은 시스템 에너지 효율성과 1.63배 높은 매크로셀 밀도를 달성하였다. 특히, 수직 아날로그 연산과 수평 디지털 연산을 결합한 설계는 MAC 연산의 병렬성과 속도를 극대화하여 연산 병목을 최소화하였다.



[그림 2] #25.2에서 제안한 HYTEC 구조

#25-4는 KAIST에서 발표된 Embedded Flash 기반 Compute-in-Memory (CIM) 아키텍처로, 플래시 메모리의 비휘발성 특성을 활용하여 엣지 디바이스에서 데이터 이동 지연과전력 소비를 줄이는 동시에, 높은 에너지 효율과 처리량을 제공하도록 설계되었다. 본 연구는 Negative-Aware-Skip (NAS) 알고리즘을 통해 불필요한 연산을 줄이고, 전력 소모를 최소화하였다. NAS는 음수 출력을 제거하여 CIM 연산의 효율성을 높였으며, 이 과정에서데이터 변환 속도를 크게 향상시켰다. SAR-TDC (Time-to-Digital Converter)는 변환 과정을 디지털화하는 데 있어 파이프라인 구조를 적용하여 변환 지연을 최소화하였다. 플래시 메모리 매크로는 65nm CMOS 공정을 기반으로 설계되었으며, 1.86mm²의 코어 면적을 차지한다. 이 구조는 플래시 메모리의 고속 데이터 읽기 및 쓰기 기능을 활용하여 연산 병목을 방지하고, 고밀도 데이터 저장과 연산을 통합하여 639.38GOPS의 처리량과 1517.16TOPS/W의 에너지 효율을 달성하였다. NAS 알고리즘은 CIM 연산에서 음수 데이터의 비효율성을 제거함으로써 전력 소비를 추가로 줄였고, MSB와 LSB 간의 차동 전류구성은 신호의 선형성을 극대화하였다. 이 설계를 통해 기존 NVM 기반 CIM 구조 대비전력 소비를 최대 42.5% 줄이는 동시에 높은 처리 성능을 유지하였다.



[그림 3] #25.4에서 제시한 eflash-CIM 아키텍처.

저자정보



엄소연 박사과정 대학원생

● 소속: KAIST 전기및전자공학부

● 연구분야: Computing-In-Memory Processor

● 이메일 : soyeon.um@kaist.ac.kr ● 홈페이지 : https://ssl.kaist.ac.kr/

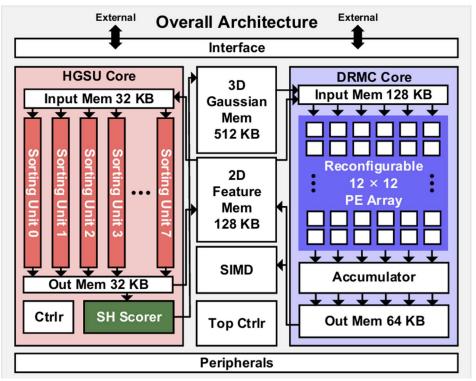
2024 A-SSCC Review

KAIST 전기및전자공학부 박사과정 엄소연

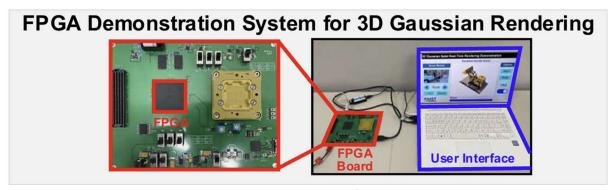
Session 7 FPGA Implementation for Emerging Applications

이번 2024 IEEE ASSCC 의 Session 7 에서는 AI on FPGA 라는 주제로 총 4 편의 논문이 발표되었다. 이 세션에서는 이벤트 비전 센서를 활용한 객체 검출, 고속 3D 렌더링, 실시간 SLAM 가속기, 그리고 포스트 양자 암호를 위한 FPGA 가속기가 소개되었다. 본후기를 통해 2 개의 논문에 대해 상세히 살펴보고자 한다.

#7.2 는 KAIST에서 발표된 논문으로, 3D Gaussian Splatting (3DGS)을 활용한 고품질 3D 렌더링 FPGA 가속기를 소개한다. 3DGS는 이전의 Neural Radiance Fields (NeRF) 방식보다투명 및 반사 장면 렌더링에서 더 우수한 품질을 제공하며, 적은 학습 데이터로도효율적으로 장면을 표현한다. 본 연구는 Hierarchical Group Sorting Unit (HGSU)를설계하여 불필요한 가우시안을 사전에 제거함으로써 VR (Volume Rendering) 단계의계산량을 줄이고, 처리량을 1.2 배향상시켰다. Dimension Reconfigurable Matmul Core (DRMC)는 다양한 매트릭스 연산을 유연하게 처리할 수 있도록 설계되어, 자원 활용도를극대화하며 1.4 배높은 처리량을 달성하였다. 또한, Importance-Based SH Evaluation를통해높은 수준의 디테일이 요구되지 않는 가우시안에서는 기본 밴드만 계산하여 메모리사용량을 67.2% 줄이고 처리량을 9.3 배향상시켰다. Cyclone V FPGA 에서 구현된 이가속기는 66.6 FPS 의실시간 렌더링 속도를 제공하며, 기존 대비에너지효율과 정확도모두에서 우수한 성능을 입증하였다. 특히 Synthetic NeRF Lego 데이터셋과 T&T Truck데이터셋에서 테스트 결과, PSNR 손실을 0.4dB로 유지하면서도 높은 품질의 렌더링결과를 보여주었다.



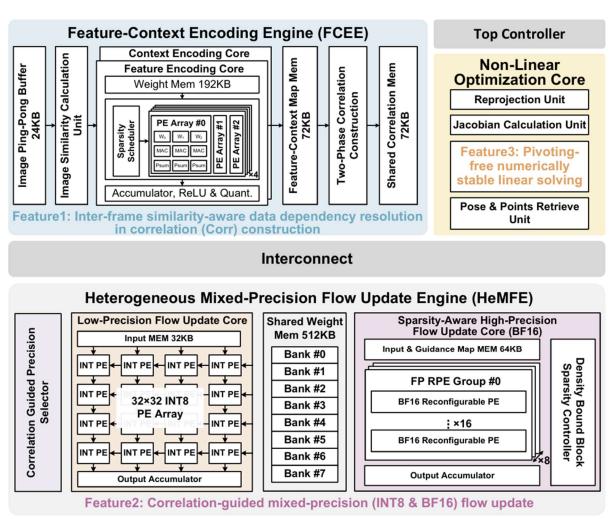
[그림 4] #7.2에서 제시한 3D 렌더링 FPGA 가속기 아키텍처.



[그림 5] #7.2에서 제시한 FPGA를 활용한 데모 시스템.

#7.3 은 Fudan University 에서 발표된 논문으로, Optical-Flow 기반의 Simultaneous Localization and Mapping (SLAM) FPGA 가속기를 소개한다. 본 연구는 SLAM 의 data dependancy 을 완화하고 연산 효율성을 높이기 위해 Inter-Frame Similarity 를 활용한 Correlation Construction 기법과 Correlation-Guided Mixed-Precision Flow Update 방식을 도입하였다. Inter-Frame Similarity 는 연속 프레임 간의 유사성을 기반으로 data dependancy를 줄이며, 연산 과정을 두 단계로 분리하여 연산 속도를 4.4 배 향상시켰다. 혼합 정밀도 연산(Mixed-Precision)은 상관도(Correlation)를 기반으로 정밀도를 조정하여 대부분의 데이터는 INT 로 처리하고 일부 중요 데이터만 FP 로 처리함으로써 연산량을 줄였다. 이를 통해 3.7 배 높은 속도를 달성하면서도 정확도 손실을 0.5% 이하로

유지하였다. 또한, 비선형 최적화 코어(NLOC)는 Householder 변환을 활용해 수치적으로 안정적인 방식으로 Jacobi 행렬을 계산하고 해를 구했다. 본 가속기는 Xilinx XCZU7EV FPGA에서 구현되어 65 FPS의 처리량과 프레임당 110mJ의 에너지 소비를 달성하였으며, 기존 대비 500 배 이상의 에너지 효율성을 입증하였다. 제안된 아키텍처는 모바일 로봇 및 SLAM 응용에서 실시간 고해상도 맵 생성을 위한 새로운 가능성을 제시한다.



[그림 6] #7.3에서 제시한 SLAM FPGA 가속기 아키텍처.

저자정보



엄소연 박사과정 대학원생

● 소속: KAIST 전기및전자공학부

● 연구분야 : Computing-In-Memory Processor

● 이메일 : soyeon.um@kaist.ac.kr

● 홈페이지: https://ssl.kaist.ac.kr/