

3-10 모바일 프로세서에서의 뉴럴넷 SW 프레임워크*

기술 개요

Neural Software Framework for Mobile Applications

Neural network software framework adapted for mobile applications

| | | | |
|-----------------------|--|------------------------|---|
| Concept | <p>에너지 효율은 모바일 프로세서의 저전력화는 실용화의 가장 핵심 요소!</p> <p>지능정보 App. Operating System Energy Hypervisor</p> <p>Wake-up Load investigation ON-OFF Decision/Time Freq/Voltage Scaling</p> <p>뉴럴넷 타입에 따른 Hypervisor SW 최적화</p> | | |
| Service Offering | <ul style="list-style-type: none"> ■ Microminiature Nano-Core based Intelligence Information Many-Core Processor Architecture ■ Energy Hypervisor SW for Ultra Low-power Processing (5mW, in Active-mode) ■ High-performance & Light-weight System converging Intelligence Information Algorithms and Neural Net Parallelization | | |
| Comparative Advantage | <ul style="list-style-type: none"> ■ 256 Nano-Cores integrated Many-core Processor for Parallel Processing ■ Below 15Watt Power consumption at General-purpose Computing ■ Below 1Watt Power consumption at Low-power Hypervisor embedded Many-core Processor ■ Minimized implemented chip size of 144mm² degree and Optimized Power consumption of Many-core Processor ■ Energy Hypervisor execution time of 0.5ms degree | | |
| Patents(Domestics) | Application (√) Registration () | Patents(International) | Application (√) Registration () |

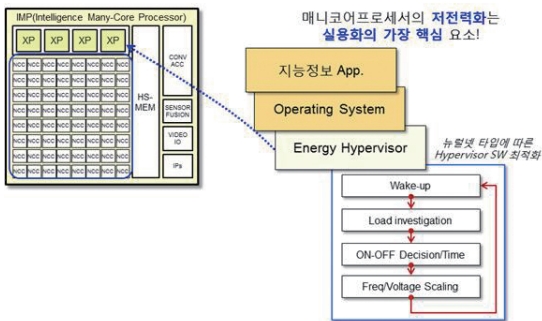
결과물 개요

| | | | | |
|----------|------------------------------|----------------|--------------------|-------|
| 개발목표시기 | 2018.06 | 기술성숙도 (TRL) | 개발 전 | 개발 후 |
| | | | TRL 3 | TRL 5 |
| 결과물 형태 | SW-Platform, Patent | 검증방법 | 자체검증, 3자검증(수요기업평가) | |
| Keywords | 뉴럴넷, 매니코어 병렬 API, 에너지 하이퍼바이저 | | | |
| 외부기술요소 | 100% 개발기술 | 권리성 | 특허, HW/SW-IP | |

기술의 개념 및 구성

기술의 개념

- 지능정보 에너지 하이퍼바이저 (Energy-Hypervisor) SW 개발
 - 뉴럴넷(Neural Net)의 초저전력 병렬화를 위한 하이퍼바이저 SW 기술
 - 병렬화에 따른 파라미터 및 메모리 사용 최적화 기술
 - 뉴럴넷 에너지 사용량 통합에 의한 코어별 Wide-Range V-F 조절 기술
 - 뉴럴넷 나노코어의 전압, 동작주파수 동적 조절 및 전력 차단 미세 조절 기술
 - 하이퍼바이저 기반의 매니코어를 시뮬레이션하여 전력소모량 측정
 - 뉴럴넷 매니코어의 활용을 위한 알고리즘 초병렬화 API 개발



기술개념도 : 저전력 매니코어를 위한 에너지 하이퍼바이저 SW

기술적 경쟁력

경쟁기술/대체기술 동향 및 현황

- 국내기술 동향
 - 4차 산업혁명의 도래에 대비하여 국내의 정책, 산업적 대응이 본격적으로 진행되어 인공지능 알고리즘 및 서비스의 개발이 이루어지고 있는 반면, 지능정보 처리를 위한 전용 지능형반도체 기술이 개발에 대한 관심이 고조되고 있음
 - 매니코어프로세서는 초병렬의 프로세싱 기술을 한 개의 반도체 칩에 구현하여 모든 디바이스에 고도의 컴퓨팅 능력을 구현하는 기술로서 혁신적인 사회서비스를 실현하는 핵심 반도체 기술

- 국내 지능정보 처리 저전력 매니코어프로세서 기술의 동향을 살펴보면, 프로세서 코어 설계 기술, 컴파일러, 지능정보처리 SW 기술 등의 핵심기술은 갖추고 있는 반면, 고성능의 지능정보 처리를 위한 매니코어프로세서 핵심 기술은 부족

- 지능정보 처리 기술의 다양한 서비스 적용을 위해서는 반도체 부품 기술의 확산과 함께 반도체에 적합한 실용적 알고리즘 기술의 개발이 필수적이나 현재의 알고리즘 연구는 대부분 방대한 연산 자원을 요구하는 이론적 인공지능 알고리즘 연구가 대부분임

■ 해외기술 동향

- 병렬컴퓨팅을 위한 반도체 기술은 반도체의 공정 스케일링의 한계에 따른 성능향상 한계를 극복하기 위한 방안으로서 2000년대 초부터 장기간 연구가 진행되었고, 최근 지능정보처리, 딥러닝(Deep Learning) 기술의 발전과 함께 주목을 받고 있음

- 지능정보 프로세싱을 위한 정부지원의 정책적인 지원을 바탕으로 전 세계 글로벌 기업의 매니코어프로세서 기술선점을 위한 연구가 활발히 진행되고 있으며, 특히 중국은 CAS(국가과학원, Chinese Academy of Science)를 중심으로 연구개발 진행

- 글로벌 기업은 정부의 지원과 시장지배력을 바탕으로 지능정보 처리를 위한 다양한 아키텍처의 매니코어프로세서를 개발하고 있음

- 전 세계적으로 4차 산업혁명의 흐름과 함께 인공지능 기술의 중요성이 높아지고 관련 기술이 발전함에 따라 요구되는 프로세서의 연산 성능이 크게 증가함에 따라 매니코어 및 뉴럴코어와 같은 새로운 반도체 설계 기술이 요구되고 있음

- 매니코어프로세서를 개발하는 대표적인 기관으로서 중국의 'National High Performance IC Design Center'와 인텔 등이 있으며 아래에 현황을 정리함

| 기업(국가) | 제품 | 전략 및 기술분석 |
|---|--|---|
| 국가고성능 반도체 설계센터 (중국: National High-Performance IC Design Center) |  〈 SW 26010 〉 | <ul style="list-style-type: none"> • (전략) 중앙 프로세서와 다수의 컴퓨팅 엔진 프로세서 코어를 한 개의 칩에 집적하여 어플리케이션 또는 알고리즘의 병렬화 특성에 따라서 서로 다른 코어를 응용하는 방식의 프로세서 • (분석) 일반적인 슈퍼컴퓨팅 연산을 목표로 하므로 중앙프로세서의 전력소모량과 컴퓨팅 엔진의 전력소모량이 수십와트로 매우 큰편 |
| Intel (미국) |  〈 Xeon Phi 〉 | <ul style="list-style-type: none"> • (전략) 자사가 강점을 가진 x86 프로세서 코어를 대형 병렬 프로세서로 집적하고 캐시 메모리 공유를 위한 아키텍처를 제공함으로써 고성능화 • (분석) x86 프로세서 코어 자체의 전력소모량이 매우 큰 편이므로 전력소모 측면에서 원천적으로 불리하여 소형 ICT 기기에 적용은 불가 |
| NVIDIA (미국) |  〈 Tegra-X2 〉 | <ul style="list-style-type: none"> • (전략) Tegra-X2는 지능정보 처리를 위하여A57 64비트 코어, 4개 Denver(Nvidia가 자체개발한 프로세서 코어), 512개의 CUDA 코어를 탑재 • (분석) NVIDIA의 셰이더 코어는 3차원 그래픽스 처리 전용으로 개발된 코어로서 전력소모량이 40W~100W에 이르기 소형의 ICT기기에 사용 불가 |
| Qualcomm (미국) |  〈 Zeroth; DSP 〉 | <ul style="list-style-type: none"> • (전략) Qualcomm은 스마트폰 AP (어플리케이션 프로세서)인 Snapdragon에 NPU (Neural Processing Unit)를 통합하여 고도의 지능형 컴퓨팅을 위한 개발 추진 • (분석) Qualcomm은 Zeroth 프로젝트를 통하여 뉴럴넷 프로세싱을 위한 특화형 IP를 AP에 통합하겠다는 계획은 발표하였으나, 개발 중단 |
| MIT (미국) |  〈 Eyeriss 〉 | <ul style="list-style-type: none"> • (전략) MIT의 Microsystems Tech. Lab.은 DNN(딥뉴럴넷) 중 AlexNet (토론토대의 Hinton 교수 그룹이 개발)을 가속하기 위한 뉴럴 네트워크 가속칩 개발 (출처: ISSCC 2016) • (분석) 특정 DNN에 한정적인 형태의 지능형반도체로서 다양한 딥뉴럴넷의 구현이 어려우며, 프로세서 코어가 내장되어 있지 않아 외부 AP를 사용함으로써 성능저하 |

우수성 및 차별성

| 경쟁기술 | 본 기술의 우수성 |
|-----------------|---|
| Xavier (NVIDIA) | <ul style="list-style-type: none"> • 경쟁기술은 그래픽스 전용의 Shader 코어 사용으로 전력소모량 20~30W 소요 • 본 기술은 뉴럴넷에 최적화된 초저전력 나노코어 설계 및 에너지 하이퍼바이저 기반 병렬화처리를 통해 전력소모 최적화한 1W 급 매패코어프로세서 구조 |
| MIC (Intel) | <ul style="list-style-type: none"> • 경쟁기술은 x86 자체의 전력, 625mm²의 면적에 의해 상당한 전력 소요 • 본 기술은 뉴럴넷에 최적화된 초저전력 나노코어 설계를 통해 144mm² 급 면적의 매패코어프로세서 구조 구현 |
| SW26010 (ICDC) | <ul style="list-style-type: none"> • 경쟁기술은 일반적인 슈퍼컴퓨팅 연산을 목표로 하므로 중앙프로세서의 전력소모량과 컴퓨팅 엔진의 전력소모량이 수십와트로 매우 큰편 • 본 기술은 뉴럴넷에 최적화된 초저전력 나노코어 설계 및 에너지 하이퍼바이저 기반 병렬화처리를 통해 전력소모 최적화한 1W 급 매패코어프로세서 구조 |

기술의 특성 및 성능

- 고객/시장의 니즈를 충족시키는 독특한 점
 - 국내최초, 세계수준의 지능정보 매패코어프로세서 설계 기술 독자 확보를 통한 인공지능반도체 시장진입 교두보 마련
 - 저전력 지능정보 매패코어프로세서 기술 개발을 통한 전력/요구성능최소화를 통한 상용화 핵심기술 마련
 - 실시간 지능정보 프로세싱 반도체 기술 확보를 통한 지능정보 처리 기술의 대중화
- 기술의 상세사항
 - 초저전력 5mW 급 나노코어(Ultra-Low-Energy Nano-Core) 설계
 - 부동소수점 연산기를 병렬로 탑재가능한 매패코어 아키텍처
 - 나노코어 eFlash 기반 에너지-하이퍼바이저 전용 모듈 구현
 - 지능정보 고도화를 위한 뉴럴넷 병렬화 및 스케줄링 하이퍼바이저 기술
 - 256-Core, 512-Thread, GHz급 나노코어 기반 매패코어 프로세서
 - 256 Core의 코어 동시작동시 1Watt급 저전력을 구현하는 에너지 하이퍼바이저 기술

표준화 및 특허

표준화 동향

- 주요 글로벌 기업을 중심으로 인공지능을 플랫폼화해 생태계로 만들어 가려는 추세로, 구글의 딥마이드랩, 페이스북의 FAIR, MS의 옥스퍼드 API, 오픈AI의 오픈 AI Gym, IBM의 머신러닝 플랫폼 시스템ML 등을 공개하고 있으나 상대적으로 기업보유 데이터의 공개는 미약한 수준이며 관련된 표준화 활동은 미비함

사업화 제약사항

시장 전망과 대응 방안

- 스마트폰이 촉발한 4차 산업혁명은 인공지능과 IoT 양대 축을 통해 바닥에서부터 사회적인 생활상을 바꾸고 있으며, 이를 통해 교통, 홈, 제조, 교육, 관광, 의료 등의 사회 전 분야에 걸쳐 고도화된 지능정보 서비스에 대한 수요가 급증함
- 국내외 프로세서 코어 설계기술, 컴파일러, 지능정보처리 SW 기술 등 핵심기술 고도화에 비해 고성능 지능정보 처리를 위한 매니코어프로세서 핵심 기술은 부족한 상황
- 지능정보 처리를 위해서는 초고성능, 초저전력의 프로세서 반도체 기술이 요구되어 투자비용 면에서의 Risk가 크기 때문에 타겟 응용 분야를 한정하여 기술수준 별로 최적화된 기술이전을 통해 시장경쟁력이 높은 제품 개발에 활용하는 전략이 필요

시장성

국내외 시장동향 및 전망

- 국내 및 해외 시장 동향 및 전망
 - 4차 산업혁명의 중추 기술, 인공지능 시장우위 창출을 위해 반도체 경쟁력은 필수
 - SK하이닉스는 스탠퍼드대와 뉴로모픽(Neuromorphic) 칩 연구개발 협약(16,10)을 맺고 인공지능 시대를 대비
 - 삼성전자는 D램과 낸드플래시를 조합해 일반 SSD보다 속도를 대폭 높인 Z-SSD를 일부 고객사에 공급(17.3)
 - 전통 반도체 업체인 인텔 · 엔비디아 등을 비롯해 구글 · 애플과 같은 인터넷 · 모바일 기업도 인공지능 시대가 보편화되고 있음을 인지하고 전용 칩, 프로세서 등 반도체 기술개발에 매진하고 있는 상황
 - 이에 반도체 강국인 우리나라에게도 새로운 기회가 될 것으로 예상되는 바, 높은 기술력을 바탕으로 인공지능 시대를 대비한 충실한 준비 필요

| (단위 : 억원) | | | |
|---|-----------------|----------------------|----------------------|
| 년도 | (2017년) 현재년도 | (2021년) 개발 종료후 1년 | (2023년) 개발 종료후 3년 |
| * 지능정보 ASIC/ASSP Data Processing, DSP, Microprocessor | | | |
| 세계 시장 규모 | 793,092 | 839,360 | 845,002 |
| 한국 시장 규모 | 23,761 | 25,147 | 25,316 |
| * 지능정보 Electronic Equipment Pivot Table, Data Processing-Compute-Server | | | |
| 세계 시장 규모 | 143,989 | 167,216 | 183,388 |
| 한국 시장 규모 | 4,550 | 5,284 | 5,795 |
| * 지능정보 Microprocessor-Computing&Industrial Components | | | |
| 세계 시장 규모 | 486,339 | 520,956 | 538,569 |
| 한국 시장 규모 | 14,969 | 16,035 | 16,577 |
| * 종합 | | | |
| 세계 시장 규모 | 1,423,420 | 1,527,532 | 1,566,959 |
| 한국 시장 규모 | 43,280 | 46,466 | 47,688 |
| ※ Gartner Worldwide Semiconductor Forecast 1Q2016 자료 인용 | | | |
| ※ 환율 : 1169.95 적용 (Gartner 자료 2017년 기준) | | | |
| ※ 한국 시장규모는 2017년의 시장규모를 바탕으로 2021/2023년은 동일 비율로 계산 | | | |

제품화 및 활용분야

| 활용분야(제품/서비스) | | 세부내용 |
|--------------|--------------|---|
| 1 | 자율주행이동체 | 지능정보를 실시간으로 처리하기 위한 지능정보 실시간 처리 매니코어프로세서 칩으로서 자율이동체의 영상정보 인식, 판단, 액추에이션 기능을 통합 처리 (인간형 로봇, 자율주행자동차, 자율이동체 등) |
| 2 | 소형 서버급 슈퍼컴퓨팅 | 자율주행자동차, 중형의 자율이동체 등에서는 고성능의 매니코어프로세서를 다수 상호연결하여 동작하는 고성능 소형 슈퍼컴퓨터급 컴퓨팅 성능을 요구하므로, 본 과제의 개발 결과물인 매니코어프로세서를 응용하여 시스템을 구현 (자율주행차, 중형 자율이동체 등) |

시장규모

| 활용분야 (제품/서비스) | 관련 시장 규모(5년), 단위: 억달러 | | | | |
|---|-----------------------|----------------|----------------|----------------|----------------|
| | 2018 (선택기능) | 2019 (선택기능) | 2020 (선택기능) | 2021 (선택기능) | 2022 (선택기능) |
| 1 자율주행이동체 | 458.88 | 527.72 | 680.27 | 1069.52 | 1470.24 |
| 2 소형서버급 슈퍼컴퓨팅 | 3.14 | 4.52 | 6.51 | 9.38 | 13.5 |
| ※ 자율주행자동차 시장 및 정책동향(월간SW중심사회, 2017.7), 드론 시장 및 산업동향(융합 Weekly TIP, 2017.1), 관계부처합동 지능형 로봇 2014 실행계획(한국정보화진흥원, 2014) 자료를 토대로 ETRI에서 가공 | | | | | |
| ※ Micro Server IC Market Global Forecast to 2022 (RESEARCH AND MARKETS, 2017.1) | | | | | |

기대효과

기술도입으로 인한 경제적 효과

- 초저전력 지능정보 처리 스케줄링과 멀티코어 프로세서 설계 기술, 하 이퍼바이저 기반 SW 등 기존 해외 IP 및 지식재산 수입에 의존해온 매 니코어프로세서의 설계 핵심기술의 국산화를 통해 국내 중소기업의 해 외 IP 및 지식재산 수입에 의한 외화 유출을 축소
 - ※ 2010년부터 2014년까지 국내 지식재산권 국제수지적자 29조원 ('국내 지식재산권 국제수지적자 29조원, "국내지식재산과 시장 활성 화 필요", 에너지경제, 2015년 9월 15일')
- 모바일 디바이스 시장의 레드오션화 및 성장 둔화로 인한 대한민국 IT 경쟁력 공백을 매니코어프로세서 설계기술 및 초저능 컴퓨팅 서버 기술 기반의 초저능 지능정보처리 기술 서비스로 대체하여, 급격히 성장하는 해외 인공지능 시장 대비 및 국내 기술 경쟁력 확보
 - ※ 인공지능 시장의 성장 규모는 연평균 82.9%로 예상, 기술 분야별로 는 딥러닝(Deep Learning) 관련 기술이 104.5%로 최대 ('인공지능 시장 전망' , 미래에셋대우, 2016년 11월)
 - ※ 국내 인공지능 시장, 즉, 인공지능 서비스와 인공지능을 위한 프로세 서 반도체의 통합 생산 규모는 2030년 27.5조원까지 성장 예상 ('인 공지능 시장 전망' , 미래에셋대우, 2016년 11월)

기술사업화로 인한 파급효과

- 매니코어프로세서는 모든 디바이스에서 실시간 저전력 지능정보 처리 를 가능하게 하는 기술로서 뛰어난 성능의 지능을 반도체와 SW 융합 기술로 구현하는 4차 산업혁명 선도 기술
- 지능정보 고성능, 저전력 매니코어프로세서 기술확보를 통해 4차 산업 혁명 시대의 반도체 핵심기술 확보
- 인공지능반도체 기술확보를 통해 자율주행자동차, 자율이동비행체 등 을 위한 새로운 반도체 시장 진입 교두보 확보