

FPGA 기반 연산가속장치를 활용한 로봇 모빌리티 자율주행을 위한 통합제어장치 구현

이용규, 이범진
현대케피코 로봇항공개발팀

YongGyu.Lee@hyundai-kefico.com, BeomJin2@hyundai-kefico.com

Implementation of Integrated Control Device for Autonomous Driving of Robot Mobility Using FPGA-based Computational Accelerator

Lee Yong Gyu, Lee Beon Jin
Robotics & UAM Development Team, Hyundai Kefico

요 약

본 논문에서는 로봇 자율주행을 위해 필요하는 고부하 연산을 가진 SW 알고리즘들을 빠르게 처리하기 위한 연산처리 장치에 대한 연구로 주 CPU 연산기와 FPGA 기반 연산가속기의 구조를 갖는 이기종 통합제어장치에 대한 연구하였다. 이 통합제어장치를 기반으로 영상데이터들을 통해서 객체검출 AI 알고리즘인 Yolo-v4 에 대한 연산가속 성능을 실험하기 위해 연산가속기인 FPGA 에 AI 가속 로직블록인 DPU 를 설계하였고 주 연산기에서 응용 SW 를 실행하고 연산가속기에서 가속화된 AI 결과들을 확인하여 표시하였다. AI 연산 가속성능 비교를 위해 단독 CPU 와 GPU 플랫폼과 알고리즘 성능 비교를 실시하였고 본 연구가 우수한 가속 성능을 갖는 점을 확인하였다.

I. 서 론

최근에 엣지컴퓨팅 개념이 사회적으로 대두됨에 따라 로컬환경에서 수집된 데이터들을 즉각적인 처리에 대한 필요가 점진적으로 증가하고 있다.[1] 이를 통해 실시간 처리, 저부하 네트워크 트래픽, 데이터 보안과 같은 이점을 취할 수 있다. 그리고 로봇, 드론, 차량, 물류 등의 모빌리티의 자율주행을 위해 인공지능(AI) 기술은 필수적인 구성요소가 되었고 고부하 및 대용량 연산을 갖는 AI 기술은 실시간 처리가 가능한 연산장치가 필요로 하다. 널리 알려진 바와 같이 전통적인 연산장치인 중앙처리장치(CPU)로 실시간성 보장하기 어렵고 그래픽처리장치(GPU)로는 실시간성을 보장 가능하지만 GPU 장치가 소비하는 전력량이나 방열을 위한 장치들과 같은 부가적인 비용이 매우 크다. 그래서 [2-3] 같은 많은 연구에서 FPGA 연산장치를 대안으로 채택하였고 이에 대한 결과로 실시간성과 전력소비량에 대한 효과적인 결과를 얻었다. 하지만 모빌리티 자율주행은 AI 기술뿐만 아니라 주행환경 인지, 판단, 경로계획, 통신, 보안 및 인터페이스까지 다양한 기술들이 구성요소로 필요하고 있다.[4] 그래서 AI 기술에 대한 연산처리 가속뿐만 아니라 다른 요소기술에 대한 가속처리도 함께 요구되고 있다. 게다가 사용자 어플리케이션을 위해서도 연산 능력을 배분하여야 한다. 그래서 단일 연산기를 갖는 제어장치에서는 막대한 비용을 소모하여 단일 연산기의 성능을 최대화하거나 연산최적화 방안들이 존재한다.

본 논문에서는 일반적으로 사용되는 단일 연산장치 구조의 제어장치를 벗어나 사용자 인터페이스를 담당하는 기능과 연산가속을 담당하는 기능을 갖는 두개의 이기종 연산 장치를 포함하는 통합제어장치를 구현하여 자율주행에 필요한 방대한 연산에 대해 효과적으로 대응하고자 한다.

II. 본 론

자율주행기술의 기술적 구성요소로 인지, 판단, 제어로 크게 3 가지 요소로 이뤄진다. 그림 1 과 같이 자율주행 시스템 환경에서 인지영역은 시스템의 주변 환경을 정확하게 인식하는 목표로 카메라, 초음파센서, 라이다, GPS 와 같은 센서들로 환경정보를 취득하는 기능을 수행한다. 판단영역에서는 인식영역에서 수집된 데이터들을 기반으로 정확한 판단을 목표로 딥러닝, 패턴분석, 빅데이터 분석과 같은 알고리즘을 이용하여 환경 분석·판단 기능을 수행한다. 이 기능을 위해 제어장치에서 연산처리장치가 필요로 한다. 제어영역에서는 취득된 결과를 바탕으로 시스템에서 필요로 하는 제어기능들을 수행한다.

자율주행 제어장치에 적용되는 연산처리장치는 사용자 인터페이스를 위한 주 연산처리기와 연산가속기로 나뉘며 그림 2 와 같은 제어장치 내부에 연산처리구조를 갖도록 설계하였다. 주 연산처리기는 서비스 공급자와 사용자 측면에서 사용하기 친숙한 CPU 로 설계하였다.

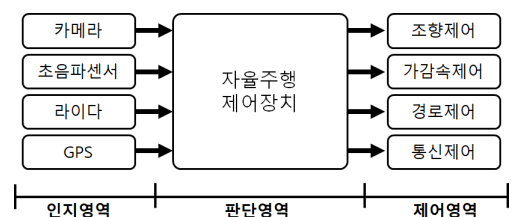


그림 1. 자율주행 시스템 환경

연산가속기는 다양한 고부하 알고리즘을 연산가속하기 위해 프로그래밍이 가능한 점과 연산가속기 구성을 위해 소비되는 자원을 최소화하며 연산가속 성능을 향상할 수

있는 효과적인 장치가 필요하다. 이를 위해 FPGA 를 설계하였고 적은 소비전력과 로직블럭에 의해 연산가속 성능이 뛰어나고 프로그래밍 가능하여 다양한 알고리즘 가속에 적합하다. 본 연구에서는 주 연산처리기로 인텔 코어 CPU 를 설계했고 연산가속기로 AMD XILINX 사의 ZYNQ MPSoC [5]를 설계하였고 주 연산처리기와 연산가속기간 데이터 교환을 위해 이더넷 통신과 짧은 길이 커맨드 전달을 위해 직렬통신을 반영하였다. 부가적으로 연산처리장치에 동반되는 DRAM 과 데이터 저장장치인 eMMC 메모리들을 부착하였다.

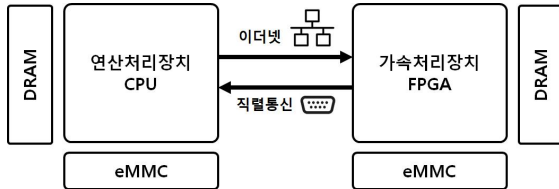


그림 2. 이기종 연산처리장치 구조

연산가속기로 적용한 ZYNQ MPSoC 는 하드코어 ARM Cortex-A53 CPU 를 내장한 연산장치로 그림 3 과 같이 CPU 영역과 로직블럭영역을 갖고 있다. CPU 영역에서는 주 연산 처리장치와의 데이터통신, FPGA 운영 및 시퀀스, 로직블럭에 대한 바이너리 파일들을 저장·관리한다. 로직블럭영역에서는 연산가속에 필요한 실제적인 로직블럭을 설계하여 연산가속을 수행한다.

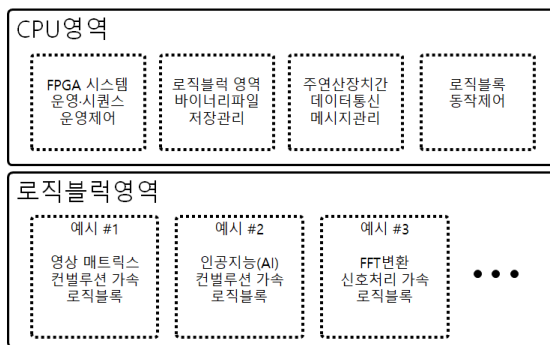


그림 3. FPGA 내부구조와 기능설계

III. 구현

본 논문에서는 자율주행에 필요한 알고리즘에서 연산가속이 필요한 기술들 중에 AI 알고리즘을 연산가속을 구현하였다. AI 알고리즘은 크게 컨벌루션, 뉴럴네트워크, 활성화 함수 연산들로 이뤄졌고 알고리즘 수행동안 반복적으로 수행된다. 그래서 이 연산들을 대상으로 연산가속 로직블럭을 구성하였고 이 블록은 AMD Xilinx, Vitis AI DPU 로 구성하였다. DPU 설계에 대한 FPGA 자원은 표 1 와 같이 사용되었다. 그리고 연산가속 실험을 위해 외부 카메라로 취득된 영상을 연산 가속기에 전달하여 객체감지 알고리즘인 Yolo-v4 AI 알고리즘을 수행하였다. AI 모델로는 Yolo-v4 Leaky[6]를 사용하였고 모델의 입력이미지 크기는 416x416 이다.

FPGA 자원	사용량
LUT	54,854
Flip Flop	83,977
RAM	177
DSP	356

표 1. DPU 설계에 대한 FPGA 자원 사용량

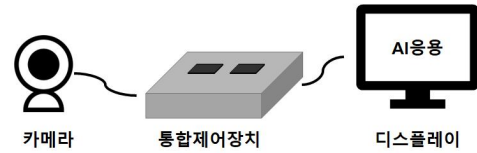


그림 4. AI 연산가속을 위한 실험환경 구성

총 300 장의 이미지를 대상으로 단독 CPU 또는 GPU 같은 다른 연산처리장치들과의 연산가속 성능비교 실험을 실시하였다. 단독 CPU 의 경우, 인텔 i5-10210U 를 사용하는 인텔 NUC 플랫폼[7] 환경에서 실시하였고 GPU 경우, Nvidia Jetson Nano Kit[8] 환경에서 실시하였다. AI 추론성능은 아래 표 2 와 같이 나타냈고 DPU 최적화 설계면에서 성능향상의 가능성이 존재하였다.

FPGA 자원	FPS (평균치)
[7]	0.5 FPS
[8]	1.4 FPS
Ours	4.4 FPS

표 2. FPGA 자원 사용량

IV. 결론

본 연구에서 많은 부하를 갖는 자율주행 SW 알고리즘들을 효과적으로 처리할 수 있는 제어장치에 대한 연구를 하였다. 연산처리 방안의 고려로 연산장치 구조에 대해 주 연산과 연산가속 기능으로 분할한 시스템을 제안하였고 연산가속 기능을 FPGA 를 통해 설계하였다. 이 구조상에서 AI 알고리즘 가속을 위해 설계 및 실험한 결과 임베디드 시스템 상에서 효과적인 성능을 확인할 수 있었다.

참고 문헌

- [1] 정현석, "4 차 산업혁명에서의 클라우드컴퓨팅 활용에 관한 연구", 한국통신학회논문지, Vol 44, No.6, pp.1213-1222.
- [2] 오은영외 3 명, "엣지 컴퓨팅 응용 서비스 가속을 위한 FPGA 기술", 한국통신학회지(정보와통신) Vol 35, No.11, 2018/10, pp.32-39.
- [3] L. Lu, Y. Liang 외 2 명, "Evaluating fast algorithms for convolutional neural networks on FPGAs", IEEE 25th Annu. Int. Symp., 2017/04, pp. 101- 108.
- [4] 박진선, "상용 수준의 자율주행 및 자동화를 위한 인공지능 기술 소개", 정보과학회지, Vol 40, No 6, pp.41-45.
- [5] <https://www.xilinx.com/products/silicon-devices/soc/zynq-ultrascale-mpsoc.html>
- [6] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020.
- [7] <https://www.intel.com/content/www/us/en/products/sku/189239/intel-nuc-10-performance-kit-nuc10i5fnh/specifications.html>.
- [8] <https://developer.nvidia.com/embedded/jetson-nano-developer-kit>.