

최신 ICT 이슈

II. 퀄컴 IoT 센서 전용의 신형 칩 발표, AI 보안 카메라 등 개발 촉진

- 퀄컴은 IoT 엣지 디바이스의 비전 처리 전용 브랜드로 'Qualcomm Vision Intelligence Platform'을 만들고, 그 첫 번째 행보로 SoC(시스템온칩) 제품 2개를 발표하였음

- ▶ 퀄컴은 칩뿐만 아니라 관련 소프트웨어도 제공하는데, 이 중에는 학습된 신경망을 SoC에 구현하여 AI 추론을 실행하는 소프트웨어도 포함되어 있기 때문에, 칩과 소프트웨어를 조합하면 가령 AI 보안 감시 카메라의 개발을 보다 용이하게 할 수 있음

- ▶ 퀄컴은 Vision Intelligence Platform(비전 인텔리전스 플랫폼)의 응용 분야로 산업 및 생활용 감시 카메라와 스포츠 카메라, 웨어러블 카메라, 360도·180도 가상현실(VR) 카메라, 로봇틱스, 스마트 디스플레이 등을 꼽고 있음

- ▶ 퀄컴의 보도자료에 따르면 카메라와 스마트 도어락 제품을 판매하는 리코(RICOH)와 중국 KEDACOM(Keda Communications) 등이 비전 인텔리전스 플랫폼을 이용하여 다양한 카메라 기기를 개발할 계획임

- ▶ 이번에 발표된 2 종류의 SoC인 'QCS605'와 'QCS603'은 모두 삼성전자의 1세대 10nm FinFET 공정(10LPE)에서 제조하게 되며 현재 샘플 제품을 출하 중에 있음

- ▶ 퀄컴과 대만의 알텍(Altek)은 현재 QCS605를 기반으로 한 360도 VR 카메라의 레퍼런스 디자인을 제공 중이며, QCS603을 기반으로 한 산업용 감시 카메라의 레퍼런스 디자인은 2018년 하반기에 제공할 예정임



<자료> Future Source

[그림 1] QCS605칩과 360도 VR 카메라

- 비전 인텔리전스 플랫폼에서 주목 받고 있는 AI 추론 처리는 “퀄컴시 엔진”을 통해 제공되

* 본 내용과 관련된 사항은 산업분석팀(☎ 042-612-8296)과 최신ICT동향 컬럼리스트 박종훈 집필위원(soma0722@naver.com ☎ 02-576-2600)에게 문의하시기 바랍니다.

** 본 내용은 필자의 주관적인 의견이며 IITP의 공식적인 입장이 아님을 밝힙니다.

며, 퀄컴은 학습된 신경망을 AI 엔진에 맞게 최적화하는 소프트웨어도 제공하고 있음

- ▶ ‘QCS605’와 ‘QCS603’은 모두 퀄컴의 스마트폰용 SoC인 ‘스냅드래곤(Snapdragon)’ 프로세서에서 이동통신(LTE)용 모뎀 회로를 제외했을 뿐 동일한 칩 내부 구성을 채택하고 있으며, QCS605가 상위 제품으로 QCS603보다 동작 주파수가 높고 기능이 많음
- ▶ CPU 코어를 예로 들면, 두 칩 모두 ARMv8-A 아키텍처에 퀄컴 독자적인 마이크로 아키텍처인 ‘카이로(Kyro) 300’을 통합했으나, 상위 제품인 QCS605는 동작 주파수 2.5GHz인 ‘카이로 300 골드’ 2개와 동작 주파수 1.7GHz인 ‘카이로 300 실버’ 6개 등 총 8코어로 구성
- ▶ 반면, 하위 제품인 QCS603은 동작 주파수 1.6GHz인 ‘카이로 300 골드’ 2개와 동작 주파수 1.7GHz인 ‘카이로 300 실버’ 2개 등 총 4코어 구성을 채택하였음
- ▶ CPU 코어를 제외하면, 두 제품 모두 GPU 코어로는 퀄컴 ‘아드리노(Adreno) 615’, DSP(디지털신호처리) 코어로는 퀄컴 “헥사곤(Hexagon) 685 벡터 프로세서”, ISP(이미지 신호처리) 프로세서로는 퀄컴 ‘스펙트라(Spectra) 270’을 탑재하고 있음
- ▶ 이 밖에 IEEE 802.11ac 지원 와이파이 회로(QCS605는 2×2 MIMO 지원), 블루투스 5.1 회로, GNSS(글로벌 내비게이션 위성시스템) 수신 회로, 동영상 압축 신장 회로, 오디오 처리 회로, 디스플레이 출력 회로 등을 탑재하고 있어 기능은 매우 풍부한 편
- ▶ 비전 인텔리전스 플랫폼의 AI 추론 처리가 이루어지는 “퀄컴 AI 엔진” 시스템에서는 신경망 처리 전용 회로가 아닌 앞서 기술한 CPU 코어, GPU 코어, DSP 코어 등 3종의 프로세서 코어를 이용함
- ▶ 구체적으로 살펴보면 “스냅드래곤 신경 처리 엔진(Neural Processing Engine: NPE)”이라 불리는 소프트웨어가 학습된 신경망을 3종의 프로세서 코어에 분산하는 형태로 구현함
- ▶ 이를 위해 퀄컴은 텐서플로우(Tensorflow)나 카페(Caffe) 등 업계 표준 프레임워크를 통해 학습된 신경망을 AI 엔진에 최적화하는 소프트웨어도 함께 제공하고 있음

■ **한편, 반도체 소프트웨어 프로그래밍 업체인 카덴스 디자인 시스템(Cadence Design Systems)도 비전 처리 및 AI 추론 처리를 위한 DSP 코어 신제품을 발표하였음**

- ▶ 신제품 “텐실리카 비전(Tensilica Vision) Q6 DSP”(이하 비전 Q6)는 DSP 연산을 수행하는 회로로, 비전 처리 및 AI 추론 처리용 IC 설계 시에 이용됨
- ▶ 이전 제품인 ‘비전 P6’ DSP에 비해 연산 성능이 1.5배, 전력 효율이 1.25배 증가했으며, 16nm FinFET 프로세스에서 비전 Q6를 구현하는 경우 최대 동작 주파수는 1.5GHz임

- ▶ 카덴스 측에 따르면 비전 Q6는 비전 P6와 마찬가지로 단독으로 비전 처리나 AI 추론 처리도 가능하지만, 규모가 큰 CNN(Convolutional Neural Network)을 처리하는 경우 상위 제품인 ‘비전 C5 DSP’와 조합하여 사용하는 것이 효율적이라고 함



<자료> Cadence IP

[그림 2] 카덴스 디자인 시스템의 DSP 제품별 주 사용 용도

- ▶ 비전 Q6나 P6는 256 병렬처리가 가능한 반면 비전 C5는 1024 병렬처리가 가능하기 때문인데, 이러한 이유로 비전 C5는 소형화 면에서 칩 면적이 커지는 단점이 있음
- DSP 코어 신제품은 AI 추론 처리를 위한 지원 기능이 풍부해졌으며, 카덴스는 비전 처리나 광범위한 데이터 처리를 위한 다양한 개발 플로우도 준비하고 있음
 - ▶ 추가된 지원 기능 중 대표적인 것은 “안드로이드 신경망(Android Neural Network: ANN)” API 지원으로, 안드로이드용 앱에서 비전 DSP의 호출이 용이하게 되었음
 - ▶ 또한, 대응 가능한 학습 프레임워크도 늘었는데, 지금까지 지원한 카페 외에도 텐서플로우와 ‘텐서플로우 라이트’도 지원함
 - ▶ CNN 중에 사용자 고유의 커스텀 계층을 넣는 작업도 쉽게 할 수 있게 했는데, MobileNet, SegNet, YOLO 등 업계에서 잘 알려진 학습된 CNN의 동작 검증을 시행할 수 있음
 - ▶ 카덴스 측은 학습된 신경망에서의 개발 흐름뿐만 아니라 비전 처리나 광범위한 데이터 처리를 위한 개발 플로우도 준비하고 있으며, OpenCL 또는 C/C++, OpenVX 등으로 작성한 디자인도 DSP에 구현할 수 있게 하려는 것임

[참고문헌]

- [1] Qualcomm, 4. 11, <https://bit.ly/2GVAKVK>
 [2] Business Wire, 4. 12, <https://bit.ly/2ES2Hfs>