

## 최신 ICT 이슈

### I. 차세대 동작 인식 기술 발표 봇물, 소형화·저전력화가 특징

마이크로소프트가 키넥트를 발표한 이래 동작 인식 기술은 게임뿐만 아니라 의료, 제조업 등 다양한 분야에서 새로운 인터페이스 방식으로 점차 확산되어 가고 있음. 최근 반도체 및 센서 기술의 발전에 힘입어 새롭게 업그레이드되고 있는 차세대 동작 인식 기술은 소형화, 경량화, 저전력화가 특징이어서 스마트폰을 비롯해 보다 다양한 기기에 탑재될 가능성이 많아 동작 인식의 활용이 보다 광범위하게 이루어질 수 있는 모멘텀을 확보하고 있음

#### ◎ 집적회로 기술 관련 국제 컨퍼런스인 ‘International Solid-State Circuits Conference(ISSCC) 2018’에서 차세대 동작 인식 기술이 잇따라 발표되었음

- ▶ 손이나 몸의 움직임을 인식하여 기기 조작의 인터페이스로 사용하는 동작(gesture) 인식 기술은 2010년 출시된 마이크로소프트의 동작 입력 컨트롤러 ‘키넥트(Kinect)’를 계기로 대중에게 알려지기 시작했음
- ▶ 키넥트는 MS의 게임 콘솔인 X 박스와 짝을 이루며 주로 게임 분야에서 활용되었으나, 그 밖에도 직접 기기 조작이 어려운 수술실에서 이용하는 등 점차 응용처가 확대되어 왔음
- ▶ 최근에는 VR(가상현실)/AR(증강현실) 애플리케이션에서 활용이 적극 모색되고 있으며, 스마트폰이나 헤드마운트디스플레이(HMD)에 동작 인식 기술을 탑재하려는 움직임도 시작되었음
- ▶ 2월 중순에 샌프란시스코에서 열린 ISSCC 2018 컨퍼런스에서는 이처럼 새로운 분야에 적용 가능한 차세대 동작 인식 기술들이 다수 선보였으며, 신기술들은 이전의 동작 인식 기



<자료> Marrent

[그림 1] 수술실의 동작 인식 인터페이스

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

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

술에 비해 소비 전력과 비용은 절감하고, 계측 성능은 대폭 강화한 것이 특징임

◎ 차세대 동작 인식 기술들은 우선 적외선을 사용하지 않아 전력 및 비용을 크게 절감하였고, 우리나라의 KAIST가 선보인 손 동작 인식 IC 칩이 대표적

- ▶ 동작 인식 입력 컨트롤러에서 사용하는 심도 이미지 센서의 경우, 빛을 쬐 후 되돌아 올 때까지의 비행시간으로 거리를 추정하는 ToF(Time of Flight) 방식이 일반적임
- ▶ 혹은 특정한 2 차원 패턴(도트)광을 대상물에 쬐이고, 촬상 소자가 획득하는 패턴광의 왜곡을 바탕으로 심도를 측정하는 ‘SL(Structured Light)’ 방식을 사용하기도 함
- ▶ ToF 나 SL 방식 모두 비교적 정밀하게 동작을 인식할 수 있지만 적외선을 이용하므로 발광 소자나 수광 소자 등의 광학 소자가 필요하기 때문에, 그 만큼 소비 전력의 증가나 비용 상승으로 이어짐
- ▶ 이런 문제를 해결하고자 2018 년 ISSCC 에서는 적외선을 이용하지 않는 방법의 제안이 잇따랐으며, 대표적으로 KAIST 는 스테레오 카메라를 이용하여 전력 소모가 작은 손짓 인식 IC 의 프로토타입을 제작하여 선보였음
- ▶ KAIST 는 좌우 한 쌍의 카메라 이미지에 대해 대응점을 탐색하고 시차 정보로부터 이미지의 심도 정보를 얻는 스테레오 매칭 방식을 채택하였음
- ▶ 스테레오 매칭 방식은 ToF 방식 등에 비해 전력 소모가 작은 장점이 있지만, 손짓의 오인식이 발생하기 쉽다는 문제가 있었는데, KAIST 는 손의 식별과 심도 추정에 인공지능 알고리즘의 하나인 CNN(Convolutional Neural Network)을 이용하여 인식의 정확도를 높였음

		<table border="1"> <tr><td>Tehnology</td><td>65nm 1P8M CMOS</td></tr> <tr><td>Die Area</td><td>4mm x 4mm (16mm<sup>2</sup>)</td></tr> <tr><td>On-chip Memory</td><td>CNN-Stereo Engine 490.5 KB ICP-PSO Engine 291 KB</td></tr> <tr><td>Supply Voltage</td><td>0.7V ~ 1.2V</td></tr> <tr><td>Clock Frequency</td><td>10MHz ~ 100MHz</td></tr> <tr><td colspan="2">CNN Performance</td></tr> <tr><td>Energy Efficiency</td><td>0.9V/60MHz 1.80TOPS/W</td></tr> <tr><td colspan="2">1-Hand Real-time Scenario (0.85V/50MHz)</td></tr> <tr><td>Power</td><td>9.02mW</td></tr> <tr><td colspan="2">2-Hands Real-time Scenario (1.1V/90MHz)</td></tr> <tr><td>Power</td><td>31.20mW</td></tr> <tr><td>Latency</td><td>30ms (33.3fps)</td></tr> <tr><td>HGR Range</td><td>20cm ~ 40cm</td></tr> <tr><td>Hand Tracking Error</td><td>Max 10.6mm Average 4.3mm</td></tr> </table>	Tehnology	65nm 1P8M CMOS	Die Area	4mm x 4mm (16mm <sup>2</sup> )	On-chip Memory	CNN-Stereo Engine 490.5 KB ICP-PSO Engine 291 KB	Supply Voltage	0.7V ~ 1.2V	Clock Frequency	10MHz ~ 100MHz	CNN Performance		Energy Efficiency	0.9V/60MHz 1.80TOPS/W	1-Hand Real-time Scenario (0.85V/50MHz)		Power	9.02mW	2-Hands Real-time Scenario (1.1V/90MHz)		Power	31.20mW	Latency	30ms (33.3fps)	HGR Range	20cm ~ 40cm	Hand Tracking Error	Max 10.6mm Average 4.3mm
Tehnology	65nm 1P8M CMOS																													
Die Area	4mm x 4mm (16mm <sup>2</sup> )																													
On-chip Memory	CNN-Stereo Engine 490.5 KB ICP-PSO Engine 291 KB																													
Supply Voltage	0.7V ~ 1.2V																													
Clock Frequency	10MHz ~ 100MHz																													
CNN Performance																														
Energy Efficiency	0.9V/60MHz 1.80TOPS/W																													
1-Hand Real-time Scenario (0.85V/50MHz)																														
Power	9.02mW																													
2-Hands Real-time Scenario (1.1V/90MHz)																														
Power	31.20mW																													
Latency	30ms (33.3fps)																													
HGR Range	20cm ~ 40cm																													
Hand Tracking Error	Max 10.6mm Average 4.3mm																													



(a) 손 동작 인식용 IC의 주요 사양

(b) 손 동작 인식용 IC의 데모 장면

<자료> ISSCC

[그림 2] KAIST가 시험 제작한 손 동작 인식용 IC

- ▶ 구체적으로는 말하면, KAIST의 방식은 20~40cm의 손 동작 인식 범위에서 추적 오차(hand tracking error)가 평균 4.3mm 정도에 불과하고, 프로토타입 IC의 소비 전력은 한손 인식 시에 9.02 mW, 양손 인식 시에 31.2mW로 매우 낮음
- ▶ KAIST 팀은 프로토타입 IC를 구현한 USB 동글을 스테레오 카메라 장착 노트북에 끼워 손짓을 인식하는 데모를 선보였으며, 가상 공간에 나타난 입방체를 손으로 잡거나 늘리는 모습을 보여주었음

◎ **밀리미터파를 이용하여 저전력으로 동작을 인식하려는 기술도 나왔는데, 이런 방식의 데모는 인피니온(Infineon)과 텍사스 인스트루먼트(Texas Instruments, TI)가 선보였음**

- ▶ 인피니온은 ISSCC에 논문을 투고·발표한 것이 아니라, ISSCC의 담당 위원으로 추천 및 승인된 기업이 자사 제품과 기술을 홍보할 수 있는 ‘인더스트리 쇼케이스(Industry Showcase)’를 통해 기술을 공개하였음
- ▶ 인피니온은 소비 전력 100mW 미만의 소형(12.5mm×9mm×0.8mm) 동작 인식 밀리미터파 레이더(트랜시버) IC를 개발했으며, 측정 거리가 10m로 길다는 점도 특징
- ▶ TI는 시험 제작한 밀리미터파 레이더용 트랜시버 IC를 사용하여 동작 인식 데모를 보여주었으며, 이 IC는 이번 ISSCC에서 구두 발표된 것으로 주로 자율운전에서 차량 밖의 상황을 감지하기 위해 사용될 것이라고 함
- ▶ 그러나 TI는 원거리뿐만 아니라 근거리에서도 사용할 수 있음을 어필하기 위해 손가락의 움직임을 인식할 수 있다는 것을 데모 세션에서 선보였으며, 데모는 운전석 주변의 조작 패널에 적용하는 것을 가정한 것이었음

◎ **이 밖에도 노르웨이의 스타트업인 ‘엘립틱 랩(Elliptic Labs)’이 인더스트리 쇼케이스에서 초음파를 이용한 동작 인식 기술을 선보였음**

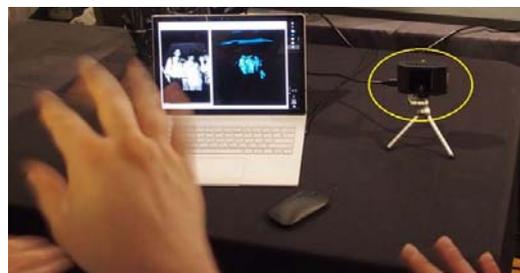
- ▶ 엘립틱 랩은 초음파의 ‘가상’ 센서로 손짓을 인식하며, 자신들의 인식 알고리즘을 채택한 소프트웨어를 스마트폰에 탑재하면, 전용 센서를 추가하지 않고도 동작 인식을 가능하게 해주기 때문에 ‘가상’ 센서라 부르고 있음
- ▶ 엘립틱 랩은 이미 스마트폰에 탑재되어 있는 스피커를 초음파 송신기로, 마이크를 수신기로 이용하기 때문에 만약 스마트폰에 여러 개의 스피커가 내장되어 있다면 동작 인식의 정확도는 더욱 높아진다고 함

◎ 동작 인식에 사용하는 ToF 방식의 심도 이미지 센서에서도 큰 진전이 있었으며, 마이크로소프트는 이번 ISSCC 에서 100 만 화소가 넘는 ToF 센서용 수광 소자를 발표하였음

- MS가 발표한 이미지 센서의 유효 화소 수는 1024×1024 픽셀, 화소 사이즈는 3.5 $\mu$ m이며, 이 화소에 초당 30 프레임의 프레임 속도로 작동시킬 경우 소비 전력은 650mW 로 낮은 편이고, 칩 크기는 9.8mm×5.4mm 로 65nm 제조 공정에서 만들어진다고 함
- 또한, ‘글로벌 셔터(global shutter)’를 지원하기 때문에 빠른 움직임을 보다 정확하게 인식할 수 있게 된 것도 특징임
- 전자제어 셔터에는 롤링 셔터(Rolling Shutter)와 글로벌 셔터 방식이 있으며, 롤링 셔터는 포컬 플레인 에 있는 두 개의 차광막을 순차적으로 열고 닫는 구조여서 이미지 센서가 한 라인씩 신호를 읽게 되므로, 뒤로 갈수록 증가된 신호량의 적체 문제가 발생
- 반면, 글로벌 셔터는 전체를 동시에 노광시킨 후 한번에 받아버리는 구조여서 한 프레임의 촬영 시점이 동일하기 때문에 시차에 의한 왜곡이 없는 장점이 있고 동영상 촬영에 보다 적합한 것으로 알려져 있음
- 통상 ToF 센서는 손을 빠르게 휘두르는 상태와 같이 빠르게 이동하는 객체가 있는 상황에서 배경광을 노이즈로 인식하기 때문에 거리 측정의 정밀도가 떨어지기 쉬움
- ToF 센서용 수광 소자는 통상 두 프레임 간의 차이를 취하여 배경광을 제거하므로, 다음 프레임이 되기 전에 객체가 움직이면 이전 프레임의 객체가 잔상으로 남아 노이즈가 되어 거리 측정 정밀도가 저하되는데, 글로벌 셔터 방식에서는 이런 문제가 거의 없어짐

◎ MS는 시연을 통해 프로토타입 센서와 레이저 광원을 조합한 ToF 센서에 의한 측정 결과 등을 보여 주었는데, 2013년에 나온 2세대 키넥트에 비해 성능이 크게 향상되었음

- 고무적인 것은 광원을 포함해도 이 ToF 센서의 소비 전력이 950mW 로 1W 를 밑돈다는 것인데, 이는 ‘모바일 기기’에도 탑재될 수 있음을 의미하기 때문
- 심도의 불확실성(Depth Uncertainty)은 3000 lux 미만의 밝기 환경에서 측거 범위의 0.2% 미만이었고, 측거 범위는 1024 픽셀 ×1024 픽셀로 구동할 경우 0.4~4.2m 였



<자료> ISSCC

[그림 3] MS의 새로운 ToF 센서 시연 장면

으며, FOV(시야각)은 수직과 수평 모두 120 도였음

- ▶ 시연 결과 MS 의 새로운 ToF 센서는 2 세대 키넥트에 비해 성능이 크게 향상된 것으로 보이는데, 가령 화소 수는 약 4 배 이상이며 소비 전력은 크게 낮아졌음
- ▶ 2 세대 키넥트는 ToF 센서의 소비 전력이 컸기 때문에 냉각 기기가 컸고 냉각 팬도 붙어 있었으며, 따라서 키넥트 본체의 크기로는 도저히 모바일 기기에 탑재할 수 없었음
- ▶ 이에 비해 MS 가 데모 세션에서 보여준 프로토타입의 ToF 센서 모듈은 마우스 정도의 크기였기 때문에 기술 발전 여하에 따라서는 스마트폰의 주변기기 나아가 스마트폰에 탑재 될 수 있는 가능성을 보여주었음

◎ 소형화-경량화-저전력화가 특징인 차세대 동작 인식 기술이 가장 먼저 활용될 분야로는 VR 과 AR, 자동차 분야 등이 꼽히고 있으며, 특히 VR 시장의 촉매제가 될 것으로 기대됨

- ▶ 2016 년에 고사양의 HMD 들이 속속 선보이며 VR 시장이 본격화될 것이란 전망이 계속해서 제기되고 있으나 아직은 현실화되지 못하고 있음
- ▶ 여기에는 여러 원인이 있겠으나 센서를 외부에 놓고 HMD 가 그 신호를 받아들여 위치와 동작을 인식하는 “아웃사이드-인 방식”으로 인한 번거로움도 한 요인으로 꼽힘
- ▶ 이에 비해 2 세대 HMD 는 센서 카메라를 HMD 의 전면에 배치하여 사용자의 위치와 움직임을 인식하는 ‘인사이드-아웃’ 방식을 채택함으로써 PC 가 필요 없는 스탠드얼론 형태를 취하고 있어, 차세대 동작 인식 기술은 더 가볍고 편리한 HMD 출현을 촉진할 것으로 보임
- ▶ 자동차 분야에서는 운전자의 몸짓을 감지해 전방 주시 태만, 졸음 운전, 운전 중 전화 사용 등에 대한 경고를 알림으로써 안전도를 높이거나, 운전자의 손 동작을 인식하여 에어컨을 켜거나 음악을 재생하는 등의 인터페이스를 구현해 편의성을 높여줄 것으로 기대됨

[ 참고문헌 ]

- [1] IPWatchdog, “Microsoft HoloLens: Will Gamble on Holographic Technology Pay?”, 2018. 2. 20.
- [2] Electronics Weekly, “ISSCC: Deep learning hardware boosts for AI”, 2018. 2. 15.
- [3] Science Daily, “New chip reduces neural networks’ power consumption by up to 95 percent, making them practical for battery-powered devices.”, 2018. 2. 14.
- [4] The MIT Tech, “Energy-efficient encryption for the internet of things”, 2018. 2. 12.