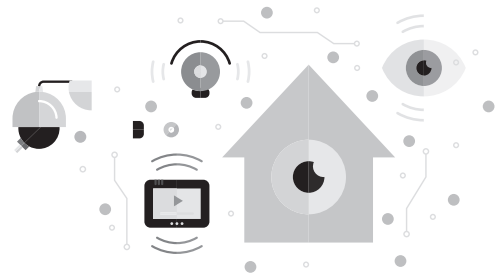


지능형 영상 보안 기술 동향



김건우 한국전자통신연구원 정보보호연구본부 PL

1. 머리말

개인신변과 사회안전을 위협하는 각종 범죄, 위법 행위가 점점 은밀화, 다양화, 복잡화됨에 따라 이를 해결하기 위한 CCTV의 수가 기하급수적으로 증가하고 있는 추세이다. 현재 국내 설치된 CCTV 수는 공공·민간을 포함해서 총 800만대로 추정되며, 이는 CCTV의 천국이라 불리는 영국의 600만대보다도 훨씬 많은 수치이다[1]. 이러한 통계적 현상은 국가 전체를 떠들썩하게 했던 많은 강력범죄 사건의 해결부터 이웃 간 사소한 분쟁 조정까지 CCTV 영상이 결정적인 증거 자료로서 주요 역할을 담당하고 있음을 보여준다.

이러한 사회안전시스템 구축의 일환으로 행정자치부는 2011년부터 전국에 229개 CCTV통합관제센터를 구축함으로써 물리적·운영조직을 통합하고 효율성을 극대화하고 있다[2].

또한, CCTV의 양적 구축 증가와 더불어 전 세계적으로 새로운 패러다임으로 부각되고 있는 이슈는 CCTV의 지능화이다. 즉, CCTV를 통해서 수집되는 영상을 사후, 실시간 자동으로 분석해서 관심상황을 감지하고 대응하기 위한 차세대 기술로 각광받고 있

다. 따라서 본고에서는 이러한 지능형 CCTV, 다양한 센서로 확장된 영상 클라우드 소싱 분석 기술 이슈를 살펴보고 향후 서비스를 전망하고자 한다.

2. 지능형 영상보안 기술

지능형 영상보안 기술은 CCTV 등 다양한 영상 센서를 통해서 수집된 영상을 지능적으로 분석하는 SW/HW 기술로서, 영상을 수집·모니터링·관리하는 VMS는 물론 영상 센서 자체에 탑재되는 모든 지능형 기술을 개념적으로 포괄한다.

지능형 영상보안 기술은 목적에 따라 배경 모델링(세그멘테이션), 영상 보정 등을 포함하는 영상 전처리 기술, 관심 객체를 검출하고 분류하는 탐지 및 분류 기술, 검출된 객체를 연속적으로 추적하는 추적 기술, 위험 상황을 감지하는 이벤트 탐지 기술, 보안 역기능 방지를 위한 영상 보안 기술로 분류할 수 있다.

2.1 지능형 영상보안 기술의 발전

기존 지능형 영상보안 기술은 목적과 적용 방법에 따라 약간의 차이가 있지만 일반적으로 컴퓨터

<표 1> 지능형 영상보안 기술 구분

분류	상세 기술	기술 개요 및 목적
영상 전처리 기술	동적 배경 모델링/세그멘테이션 기술	지속적으로 변하는 영상 내에서 배경을 분리·제거함으로써 움직이는 전경만을 동적으로 검출하는 기술
	영상 보정 기술	입력되는 영상에서 노이즈를 제거하고 정보를 명확화함으로써, 후처리 지능형 기술의 성능을 향상
객체 탐지 및 분류 기술	객체 탐지 기술	사람, 차량, 동물 등 관심 있는 객체를 탐지하는 기술
	객체 분류 기술	탐지된 객체를 상세하게 분류하는 기술로, 사람, 차량, 동물 및 상세 중별, 사람 인식 결과 등을 포함
객체 추적 기술	단일 카메라 객체 추적 기술	단일 카메라 내에서 특정 객체의 움직임을 연속적으로 추적하는 기술
	다중 카메라 객체 추적 기술 (MCT)	다중 카메라를 움직이는 특정 객체를 추적하는 Re-identification 기술
영상 보안 기술	보안 침해방지 기술	영상 센서해킹 방지, 영상 암·복호, 위·변조 방지 등
	프라이버시 보호 기술	영상 내 얼굴, 차량번호판, 창문 등 사생활 영역만 비식별화/복원하는 프라이버시 마스킹 및 De-identification 기술

비전, 이미지 프로세싱, 패턴 인식 등 분야에서 다양한 알고리즘 중심으로 발전되어 왔으나, 2000년대 중반 이후로 기계학습 방식에 기반한 신경망 기술이 급속도로 발전하고 있다. 특히 딥러닝(Deep Learning) 기술은 심화신경망(Deep Neural Network)을 활용한 기계 학습(Machine Learning) 방식을 말하며, deep은 feed-forward 네트워크에서 히든 레이어가 1개 이상을 뜻한다.

딥러닝에서 널리 사용되는 Backpropagation, CNN(Convolutional Neural Network) 등은 이미 1980년대에 이론적인 연구가 끝났지만, 최근 딥-네트워크를 처리할 수 있는 HW 계산 능력의 향상, 대규모 빅데이터의 학습 등으로 인하여 특히 컴퓨터 비전 분야에서 발군의 성장을 이루고 있다.

2.1.1 기술 트렌드

CCTV 시장은 지능정보 기반 지능형 CCTV 중심으로 개편되고 있으며, 세계 CCTV 시장 규모는 2013년 135.2억 달러에서 2018년 236.4억 달러로 5년간 약 75%가 예측되며, 그 중 지능형 CCTV 시장은 연평균 20.6%의 성장을 통해서 2018년도 전체 시장의 65%를 차지할 것으로 예상된다[3].

이러한 시장의 요구사항을 반영하듯 기술적으로

는 딥러닝 기술이 기존 알고리즘 기반 기술을 급속도로 대체하고 있으며, 성능에서도 월등한 결과를 보여주고 있다.

가. 클라우드 기반 인공지능 서비스

구글, IBM, MS 등 글로벌 IT 업체들이 클라우드 방식의 인공지능 서비스 개발에 주력하고 있어 향후 치열한 경쟁이 예상되며, 공유 개념의 클라우드가 분석 및 예측서비스를 제공하는 인공지능 클라우드 서비스로 진화 중에 있다. 기존 ‘클라우드 서비스’는 각종 자료를 외부 서버에 저장한 뒤, 인터넷으로 접속해 다양한 기기로 언제 어디서든 사용할 수 있도록 서비스했던 반면, AI 클라우드 서비스는 기존 클라우드 서비스에 빅데이터 분석, 예측 기능을 추가시켜 사용자 맞춤형 서비스를 제공하고 있다.

- 아마존은 아마존웹서비스(AWS) 클라우드 서비스 컴퓨팅, DB와 분석, AI, 개발자 도구와 지원서비스, 하이브리드-마이그레이션-관리-보안 등 5개 분야에서 29개 신규서비스를 제공
- MS는 머신러닝 기술을 연구하는 ‘옥스퍼드 프로젝트’를 통해 인지(Cognitive) 서비스를 개발하여 머신러닝 API 서비스를 제공하고 있으며 딥러닝 툴킷인 CNTK(Computation Network Toolkit)와 분산 기계학습 툴킷인 DMTK(Distributed Machine Learning Toolkit)를 무료로 배포

<표 2> 클라우드 인공지능 서비스 비교

회사명	클라우드 인공지능 서비스	오픈소스 머신러닝 툴	스타트업 M&A
Amazon	Amazon machine Learning	DSSTNE(Deep Scale Sparse Tensor Network Engine): Library for building deep-learning models	Orbeus
Facebook	None	Tools for deep-learning models released through the open-source Torch library	Wit.AI
Google	Google Cloud Machine Learning	TensorFlow: Library for developing deep-learning models and more general machine-learning models	Dark Blue Labs, DeepMind, DNNresearch, Moodstocks, Vision Factory
IBM	IBM Watson Analytics	IBM SystemML: Optimization platform for general machine-learning models on the open-source Apache Spark library	AlchemyAPI
Microsoft	Microsoft Azure Machine Learning	CNTK(Computational Network Toolkit): Library for building deep-learning models	SwiftKey

- IBM은 인공지능 왓슨을 활용한 '인공지능주택' 서비스를 시작하였고 인공지능 사물인터넷 플랫폼 기반 '왓슨 사물인터넷(Watson IoT)' 개발 계획 발표

나. MCT(Multi-Camera Object Tracking)

MCT는 다중 카메라를 이용해서 특정 객체(사람, 차량 등)의 움직임은 연속적으로 추적하는 기술로, 대상 객체의 ID를 정확하게 식별하기보다는 여러 객체의 유사도(Confidence Level)를 비교해서 동일 객체인지 여부를 판단하는 Re-identification 기술 영역에 속한다.

지금까지 MCT 기술은 제약적 환경(실내, 일정한 조명, 1~2명 객체, 비슷한 수준의 CCTV 등)에서 연구가 진행되다가 2000년 후반에 들어 비교적 제약이 없는 야외 환경에서 연구가 본격적으로 이루어지고 있다.

- ECCV와 공동으로 개최되는 MCT(Multi-Camera Object Tracking) Challenge 2014에서 USC가 최고 79% 성능으로 1위 차지

다. 지능형 도시 감시(Intelligent City Surveillance)

• DAS(Domain Awareness System), 뉴욕, 미국

미국 뉴욕시 경찰국(NYPD)이 마이크로소프트(Microsoft)와 합작으로 개발한 범죄감시시스템으로, 2013년 기준 뉴욕시 전역에 설치된 6,000여 개의 감시 카메라, 방사능 감지기(600여 대), 차량 번호판 인식장치(100여 대), 911전화기 등을 연계하여 대상을 추적하고, 이동 경로 등의 정보를 제공한다. 또한, 범죄지도(Hot spot) 관리를 위한 범죄 예측/예방 효과(예측률: 70%)를 보이고 있으며 10만 명당 FBI 중대범죄가 타도시의 1/3 수준으로 감소하였다. 이는 지능형 영상기술에 기반한 도시 감시시스템의 레퍼런스 모델로 활용되고 있으며, 미국 내 타도시(멤피스, 샌프란시스코 등)에 확대되었고, 상파울루, 싱가포르 등에 수출되기도 하였다.

• NGI(Next Generation Identification), FBI, 미국

미국 FBI의 형사사법정보국(Criminal Justice Information Service Division, CJIS)에서 2008년에 발표하였으며, 지문, 장문, 홍채, 얼굴 등의 인물 정보를 추가한 생체 데이터베이스를 구축하여, 범인의 신원을 확인하고 추적하기 위한 프로젝트로서, 2009년 이후 총 1조 원을 투자하여 미 전역 3,000만대 이상의 CCTV를 실시간 연동한다. 이를 통해서 범죄자/용의자 신원 확인을 현재 2시간에서 10분으로, 일반인 신원조회를 현재 24시간에서 15분으로 단축, 수행한다. 현재 각 주에서 파일럿 프로젝트로 참가하고 있으나, 아직 성능적인 측면에서는 미흡한 부분이 많으며, 특히 프라이버시 침해 이슈가 크게 부각되고 있다.



[그림 1] MCT Challenge 2014의 Dataset 1 화면

2.2 영상 클라우드 소싱

영상 클라우드 소싱은 기존 CCTV뿐만 아니라 차량용 블랙박스, 스마트폰, 드론, 스마트 글래스 등 다양한 영상 장비를 대상으로 지능형 서비스를 제공한다. Xively 플랫폼, Ushahidi 플랫폼, 스트리트 범프, 오픈 스트리트 맵과 같은 클라우드 소싱 플랫폼은 다양한 센서 정보(재난, 에너지, 방사능, 환경, 도로 상황 등)를 수집해서 활용하고 있으며, 크라임 스토퍼나 페이스와치 등은 수동적으로 범죄 정보를 공유하고 공동으로 대응한다.

최근 부산(히든아이), 대구(쉐도우 캡스), 충남(블랙키퍼) 등 지자체에서 시민의 차량용 블랙박스 영상을 범죄 해결에 사용하고 있다.

2.2.1 차량인식 기술

차량인식 기술은 현재 지능형 자동차 안전기술 분야에서 핵심기술로 활용이 되고 있으며 주요 글로벌 완성차 업체들 중심으로 집중적인 투자가 이루어져 상용차량에 탑재된 형태로 출시되고 있다. 주로 LDWS(Lane Departure Warning System), FCW(Forward Collision Warning) 기능 실현을 위해 지능형 영상인식 기술을 사용하고 있으며 완성차에 임베디드 시스템 형태로 설치된다.

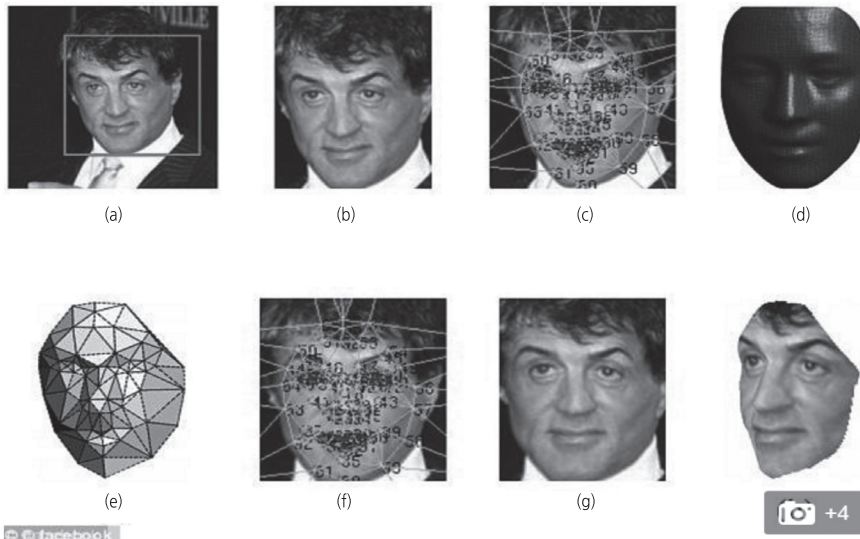
차량용 영상인식 시장(ADAS, Advanced Driver Assistance System)의 선두주자인 이스라엘 모빌아이(최근 인텔에 인수)는 글로벌 차량 안전 시장의 약 80%를 점유하고 있는 업체로 보행자 충돌 경고,

차선 이탈 경고, 전방 충돌 경고, 고속도로 모니터링, 지능형 전조등 제어, 과속 표지판, 사각 지대 인식 기술 등을 개발한다.

보행자 검출 성능 40m 이내 90%, 차량 검출 성능 100m 이내 98%이며, Audi, BMW, Chrysler, GM, Ford, Honda, Volvo, 현대, 기아를 포함한 세계 자동차 제조사에 OEM으로 공급하고 있다.

영국 CitySync Inc.는 매우 높은 번호판 인식 수준을 갖는 ANPR(Automatic Number Plate Recognition) 관련 하드웨어/소프트웨어 설계 및 인식 솔루션을 개발하였으며 자동차 번호판의 글꼴, 주행속도, 빛의 반사수준 및 촬영 각도 등 다양한 번호판 인식조건에서도 정상동작하는 기술을 보유하고 있다. 또한, 미국 vigilant Solutions에서는 고정용 및 이동형 번호판 인식 솔루션을 개발, 차량 위치를 파악하는 솔루션에도 활용하고 있다

네덜란드 Gatsometer사에서는 레이더 센서를 이용하여 다차로 환경에서 여러 차량을 모니터링하며 교통 위반 차량(신호위반, 과속, 차선변경, 불법유턴 등)을 단속할 수 있으며, 3차원 레이더 추적 기술(속도, 거리, 위치 측정)과 카메라 영상 기술, 신호위반, 과속차량, 차간거리, 불법유턴, 차선변경, 보행자 검출 및 추적, 고해상도 카메라를 이용하여 위반 차량의 영상을 촬영, 원거리 번호판 인식을 수행하며, 동시에 22대의 차량과 최대 500m까지 검출할 수 있다.



[그림 2] DeepFace 얼굴 검출

2.2.2 이동 카메라 기반 IVA 기술

대표 평면(dominant scene plane) 기반 IVA 기술은 입력 영상을 대표 평면에 정합하고, epipolar constraint, shape constancy constraint, structure consistency constraint 등을 이용하여 시차(parallax)와 객체의 움직임(object motion)에 의한 정합 오차를 보정한다.

반면, 세그멘테이션(Segmentation) 기반 IVA 기술은 Optical flow, graph-cut 등을 이용하여 객체의 움직임과 카메라의 움직임을 분리하며, 상대적으로 수행 속도가 느리고, 모델의 복잡도로 인해 파라미터에 민감하다.

컴펜세이션(Compensation) 기반 IVA 기술은 카메라의 움직임에 대한 보정을 통해 이전 프레임에서의 배경 모델과 현재 입력 영상이 일치하도록 하지만, 정확한 카메라 모션을 추정하기 어려워 대부분 간단한 카메라 모션 모델(affine or projective)을 사용한다. 영상의 경계에서 오검출이 발생할 수 있으며, 수행 속도가 빠른 것이 장점이다.

2.2.3 컴팩트 딥러닝 기술

주로 CNN(Convolutional Neural Network)을 모바일 기기에서 수행하기 위한 압축 기술들이 연구되고 있으며, 주요 방법론으로는 Low-rank Approximation, Hashing Trick, Vector Quantization, Circulant Project, Pruning, Teacher-student network, Limited numerical precision 등이 있다.

2.3 얼굴인식 기술

최근들어 딥러닝 기술의 도입으로 구글, 페이스북, IBM 등 대형 글로벌 업체를 중심으로 얼굴인식 기술의 성능이 급격히 높아지고 있으나 실시간 처리가 요구되고 상대적으로 열악한 환경에서 동작되는 CCTV 기반 얼굴인식 기술은 아직 사람의 시각수준에는 근접한 기술개발이 이루어지지 못하고 있다.

페이스북에서는 2014년 CVPR에서 사람의 얼굴인식률(97.53%)과 유사한 수준인 97.25%의 얼굴


인식률을 달성하였으며, 딥러닝 학습에 사용된 얼굴 영상은 총 4,030명의 440만장으로 자체 개발한 DeepFace 알고리즘을 사용하였다.

구글에서는 2015년 CVPR에 FaceNet 얼굴인식 알고리즘을 발표, 1만 3천여 명의 얼굴 사진이 들어 있는 LFW DB(Labeled Faces in the Wild)에 적용한 결과 99.96%의 얼굴 인식률을 달성하였으며, 2억 6천만 장의 영상을 대상으로 진행한 실험에서는 86% 이상의 정확도를 보였다.

3. 맺음말

딥러닝 방식을 필두로 인공지능 기술이 급속도로 발전하고, 다양한 영상 산업에 활용됨에 따라 많은 산업적/사회적 요구사항이 발생하고 있다. 지능형 영상분석 기술이 가장 효율적으로 적용될 수 있는 CCTV는 다양한 환경과, 성능, 목적, 요구사항을 모두 만족해야 하기 때문에 1~2년 내에 모든 지능형 기술이 상용화되기는 어렵다. 하지만 단계적인 응용을 통해서 지능형 CCTV의 적용 도메인이 확장되고 있는 추세이며, 이러한 트렌드는 클라우드화, 영

상 다변화에 의한 클라우드 소싱, 타 융합산업으로의 확산 등으로 반영되고 있다. 하지만 이러한 긍정적인 지능형 기술의 확산 이면에는 프라이버시 침해 등과 같은 보안 역기능 발생에 대한 우려도 더불어 증가하고 있다.

따라서 기술의 지능화를 통한 성능 향상은 물론 이에 대한 사회적 요구사항을 만족시키는 동시에 보안 침해 우려를 해소할 수 있는 방향으로 개발이 추진되어야 한다. 

※본 연구는 2017년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행됨[No. 2016-0-00109, 국민참여형 사회안전서비스를 위한 영상 클라우드 소싱 핵심기술 개발].

[참고문헌]

- [1] <http://www.thescoop.co.kr/news/articleView.html?idxno=21576>, 2016.10 '지능형 CCTV, 안전을 업그레이드하다'
- [2] <http://www.cctvnews.co.kr/news/articleView.html?idxno=36039>, 2015.11 'CCTV통합관제센터 229개 지자체 구축으로 대한민국 안전 책임진다!'
- [3] IHS Research, 'World Market for CCTV and Video Surveillance', 2015

스마트폰 좀비 smartphone zombie

스마트폰 화면을 들여다보느라 길거리에서 고개를 숙이고 걷는 사람을 낚는 시체 걸음걸이에 빗대어 일컫는 말.

'스마트폰(smartphone)'과 '좀비(zombie)'를 합성하여 '스몸비(smombie)'라고도 한다. 스마트폰 좀비(또는 스몸비)는 2015년 독일에서 처음 사용되었으며, 스마트폰에 지나치게 매인 세태를 풍자했다. 특히 스마트폰 화면에 눈길을 빼앗긴 탓에 자동차에 치이는 사고가 잦아 문제가 됐다. 서울시와 경찰청이 2016년 6월 시민이 많이 오가는 시청·연세대·홍익대·강남역·잠실역 길바닥에, 걸어가며 스마트폰을 보면 위험하다는 내용을 담은 교통안전표지를 설치해 눈길을 모았다.