

로컬 영상 분석과 클라우드 서비스를 활용하는 하이브리드 지능형 침입 감지 시스템

박유환, 정영빈, 이정훈, 황광일
 인천대학교 임베디드시스템 공학과
 yubak351@naver.com, 201601791@inu.ac.kr, junghoon450@gmail.com, hkwangil@inu.ac.kr

Hybrid Intelligent Intrusion Detection System Utilizing Local Image Analysis and Cloud Services

Yu-Hwan Park, Yeong-Bin Jeong, Jeong-Hun Lee, Kwang-il Hwang
 Dept of Embedded System Engineering, Incheon National University

요 약

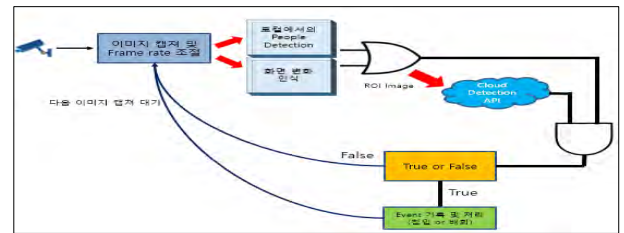
각종 테러 및 범죄 예방을 위한 지능형 영상분석 시스템을 구현하는데 기본적으로는 로컬 기반의 처리를 진행하고 필요에 따라 클라우드 기반의 처리가 병렬적으로 처리됨과 동시에 기존의 시스템에서 추가적인 구축, 설비비용 없이 설치되어있던 DVR 또는 NVR 장치를 활용하여 간단한 소프트웨어를 추가함으로써 영상분석 서비스를 이용 가능하도록 하여 영상분석 결과에 대한 신뢰성이 높고 비용 절약적인 하이브리드 지능형 침입 감지 시스템을 제안.

1. 서론

작년(2017)의 범죄 발생건수는 1,662,341건에 달하고[1] 침입/절도 범죄는 끊임없이 발생하고 있다. 범죄의 수준은 날이 갈수록 발전하고 있고 이에 따라 자체적으로 가지고 있는 CCTV는 무용지물이 되어 가고 있다. 이로 인해 녹화 영상을 저장만 하는 일반적인 CCTV에 그치지 않고 CCTV를 이용한 지능형 서비스가 점차 확산 되고 있다. 기존의 지능형 CCTV의 시스템구조의 경우 로컬 기반에서 러닝 알고리즘[2~3]을 이용한 식별 기술을 이용하고 있는데[4] 이는 클라우드 영상분석과 같은 방대한 양의 학습 데이터를 가지고 있지 않기 때문에 분석 결과에 대한 신뢰성이 검증되지 않는다. 따라서 본 논문에서는 기존의 DVR, NVR을 활용하여 로컬과 클라우드 기반의 영상분석을 병렬적으로 처리하는 지능형 영상 분석이 가능한 침입 감지 시스템을 제안한다.

제안된 방법은 카메라로부터 들어온 이미지를 로컬 영역에서는 관심 영역에 침입했는지, 배회하고 있는지 알기 위한 사람 검출과 이전 장면에서 현재 장면이 얼마나 변화하였는지 화면변화를 확인한다. 이후 하나의 이미지에 대해서 사람이 검출되었거나 화면 변화가 감지되었다면 이 이미지를 클라우드로 전송하여 Detection API를 이용해 최종 판단을 하게 된다. 최종 판단이 사람으로 판별이 되었을 경우 이벤트를 기록하고 해당하는 이벤트에 대한

처리가 이루어진다. 만약 사람으로 판별 되지 않은 경우는 다음 이미지가 캡처되기 까지 대기하게 된다.



(그림 1) 전체 구성도

2. 기존 영상 분석 시스템의 문제점

기존의 영상 분석시스템은 로컬기반의 영상분석 시스템과 클라우드 기반의 영상분석 시스템으로 나눌 수 있다.

로컬 기반의 영상분석 시스템의 경우 전문적인 영상분석을 위하여 별도의 서버가 필요한데 이는 추가적인 네트워크 설비와 영상을 저장할 수 있는 저장장치가 필요하다는 것을 말한다. 따라서 초기 설비비용이 많이 지출된다는 점이 단점이고, 학습된 데이터양이 많지 않기 때문에 이미지를 이용해 영상분석을 한 뒤 분석결과에 대한 객관적인 검증이 어렵다.

클라우드를 기반으로 한 영상분석 시스템의 경우 로컬 기반으로 한 영상분석에서 초기에 발생하는 설비, 구축비용이 발생하지 않고 방대한 데이터를 가지고 학습을 진행하기 때문에 신뢰성 높은 영상분석 결과를 얻을 수 있다는 장점을 가지고 있지만, 영상이 클라우드로 저장되기 때문에 여기서 발생하는 저장비용이 부담 될 수 있다. 만약 이전 30일에 대한 영상을 저장한다는 기능을 제공하게 된

본 연구는 2017년도 한국산업단지공단 생산기술사업화지원사업으로 지원되었음. (과제명: 클라우드 비전처리 기반의 스마트 PC DVR 시스템개발)

다면 저장 비용으로 인한 심각한 문제가 발생할 것이다. 그리고 네트워크만을 통해서 실시간 영상을 전송하여 이를 분석하기 때문에 네트워크 상태에 대한 의존성이 굉장히 높아 예를 들어 자연재해로 인한 네트워크 장애가 발생하였을 경우 서비스를 안정적으로 유지시킬 방안이 존재하지 않는다. 또한 실시간으로 영상분석을 요청하고 수신 받아 클라우드에서 분석한 응답에 따른 처리를 위한 별도의 로컬 소프트웨어가 추가적으로 필요하다.

3. 기존의 DVR, NVR 정의 및 문제점

DVR(Digital Video Recorder)은 아날로그 카메라에서 동축케이블로 영상 신호를 받은 후 이를 저장장치에서 디지털화하여 저장하는 방식으로 현재 가장 많이 사용되고 있는 CCTV 기술이다. 단점으로는 별도의 동축 케이블을 설치해야한다는 부담감과 구형의 경우 27만, 41만 화소를 가지고 있기 때문에 영상품질이 열약해 화질 보장에 어려움이 있다.

NVR(Network Video Recorder)의 기본적인 구성은 DVR과 같으나 IP카메라를 사용한다는 점에서 다르다. IP 카메라에서 디지털 인코딩을 한 후 전송하게 되며 인터넷망을 이용해 영상을 수신 받는다. 단점으로는 기존의 아날로그 카메라와 연동을 할 수 없다는 점과 LAN선이나 인터넷을 통한 해킹 시도에 대해서 취약하고 네트워크에 대한 의존도가 높기 때문에 네트워크가 다운되거나 날씨, 장애물, 자연재해와 같은 방해요소에 의해 영상이 기록되지 않거나 간섭을 받을 수 있다.

4. 하이브리드 지능형 침입 감지 시스템

4. 1 하이브리드 지능형 침입 감지 시스템의 정의

하이브리드 지능형 침입 감지 시스템은 로컬과 클라우드 둘 중 하나의 환경만을 사용하지 않고 필요에 따라 두 개의 환경을 모두 사용하는 침입 감지 시스템이다. 기본적인 환경은 로컬기반으로 진행되며, 사람 감지 또는 화면 변화가 발생하게 되면 클라우드 기반의 처리를 동시에 진행함으로써 클라우드 영상분석 서비스를 통한 신뢰성 높은 영상분석 결과를 얻게 된다. 실시간 영상을 계속하여 클라우드로 보내 영상분석을 하는 것이 아닌 침입 감지 이벤트가 발생한 것과 같은 이상상황이 발생할 때만 클라우드로 영상분석을 요청하는 것이기 때문에 사용한 만큼 지불해야하는 클라우드 사용료에 대한 부담을 최소화 시킬 수 있고 로컬기반에서 최소의 영상분석 전처리를 수행하기 때문에 기존 방식인 추가적인 서버 설치와 같은 설비, 구축비용이 불필요하다. 또한 기존의 DVR, NVR에서 간단한 소프트웨어를 추가함으로써 영상분석 처리가 가능하다. 추가적인 기능으로 로컬에서의 하드웨어 자원 (메모리, 프로세서)을 최소화하기 위한 적응형 프레임 캡처 기능(Adaptive frame capture rate)이 존재하며 클라우드 분석 결과 활용을 위한 다이나믹한 영상처리 알고리즘 등

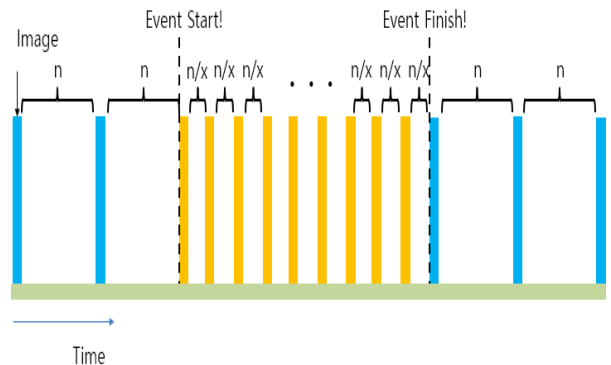
작과 로컬에서 다양한 기능으로서의 확장이 용이해진다.

4. 2 적응형 프레임 캡처 (Adaptive frame capture rate)

하이브리드 지능형 침입 감지 시스템은 클라우드 영상 분석 요청 횟수를 최소화, 하드웨어 자원규모를 최소화하기 때문에 상황에 따라 프레임 레이트를 조절하여 실시간 영상 분석을 진행한다.

기존 영상 프레임 레이트 만큼 실시간 영상분석을 진행한다면 하드웨어의 규모가 커지게 되고 만약 로컬 기반에서 이상상황이 아닌 상황을 오검출하여 이상상황으로 판별하게 된다면 불필요한 클라우드 영상분석요청을 하게 되기 때문에 클라우드 영상분석에 대한 지불비용이 증가하게 된다.

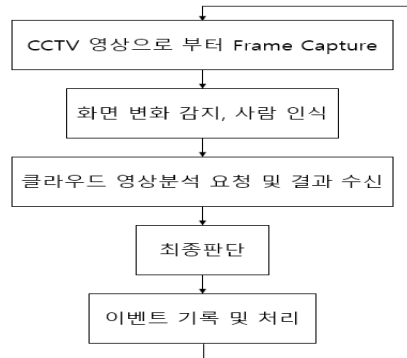
따라서 하이브리드 지능형 침입 감지 시스템은 평상시에는 프레임 레이트를 낮추어 Idle 모드로 들어가 지속적인 감시를 진행하며, 로컬기반에서의 판단하에 이상상황이 발생하면 이벤트 발생 유무를 판단하여 이벤트 상황이 발생하였을 경우 프레임 레이트를 증가시켜 이벤트에 대한 신속한 처리를 진행시키게 된다. 이후 이벤트가 종료하게 되면 다시 프레임 레이트를 낮추고 Idle 모드로 들어감을 반복한다.



(그림 2) 이상상황 발생시 적응형 프레임 캡처 기능을 이용한 프레임 레이트 조절

4. 3 하이브리드 지능형 침입 감지 시스템 알고리즘

하이브리드 지능형 침입 감지 시스템은 다음과 같은 과정으로 진행되게 된다.



(그림 3) 침입 감지 시스템의 전반적인 과정

4. 3. 1 CCTV 동영상으로부터 Frame capture

이벤트가 발생하기 전엔 비교적 낮은 Frame capture rate로 영상분석을 진행하며 화면변화 또는 사람 검출을 하기 전까지의 기능을 한다. 이후 Local 기반에서 이상상황이 감지되어 Cloud 기반의 Detection API를 사용해 신뢰성 있는 최종 판단으로 사람이 감지되었다는 이벤트가 발생하게 되면 이전보다 높은 Frame capture rate로 영상분석을 진행한다. 이벤트 종료 시 다시 기존의 Frame capture rate로 낮추어 검출함을 반복한다.

4. 3. 2 화면변화 감지과정

이 과정은 현재의 이미지와 직전 이미지를 비교하는 동작이다. 즉 현재의 이미지와 직전의 이미지를 비교함으로써 일정 수치 이상의 변화가 일어났을 경우 새로운 객체가 등장하였거나 움직이는 것으로 판단할 수 있다. 대표적인 예로는 현재의 프레임 장면과 객체를 추출하기 위한 배경 모델 장면의 뺄셈을 통해 얻은 영상을 이진화 함으로서 전경 마스크를 계산하는 배경추출(Background Subtraction) 알고리즘을 이용할 수 있다. 이러한 방식을 이용해 관심영역에 일정 수치이상의 변화가 감지되었을 경우 화면변화 감지로 판단하게 된다.

4. 3. 3 사람 인식 과정

화면변화 감지만을 이용할 경우 객체가 등장하였는데 움직이지 않고 정지해있는 경우는 인식을 하지 못한다는 단점을 가지고 있기 때문에 이 문제를 보완하기 위하여 로컬기반에서의 검출식을 이용한 사람 인식을 진행한다. 대표적인 사람을 검출하는 계산식에는 Haar feature와 Hog descriptor가 존재하는데 Haar feature를 기준으로 비교하면 Haar feature가 정확도는 비교적 낮지만, 보다 처리속도가 빠르고 64*128 크기의 보행자 템플릿으로 학습된 Hog descriptor에 비해 작은 크기의 템플릿으로 학습이 되어있기 때문에 더 작은 크기의 사람을 인식할 수 있는 장점을 가지고 있다. 로컬에서 사람을 검출하는 것이 정확도가 완벽한 수준은 아니므로 낮은 정확도는 클라우드의 영상분석을 통해 보완한다.

4. 3. 4 클라우드 분석요청 및 결과 수신

로컬 기반의 화면변화와 사람 인식만으로는 사람이라고 확신하기엔 신뢰성이 부족하다. 따라서 방대한 데이터로 학습되어 높은 정확도와 신뢰성을 가진 클라우드 영상분석을 이용한다. 화면변화 또는 사람 인식이 발생하게 되면 해당 이미지의 관심영역을 클라우드로 영상분석에 대한 요청하게 된다. 그에 따른 응답으로 요청한 이미지에 포함된 여러 객체 목록과 식별자를 수신으로 받을 수 있다.

4. 3. 5 최종 판단

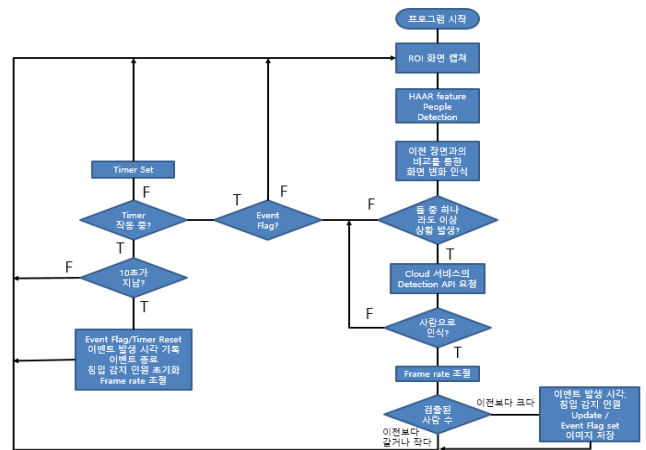
로컬 기반 두 가지 메소드의 결과를 이용한 클라우드

영상분석을 통해 얻은 결과 목록에서 사람이라고 판단이 되었을 경우 해당 관심영역 이미지에 사람이 존재한다고 최종적으로 판단한다.

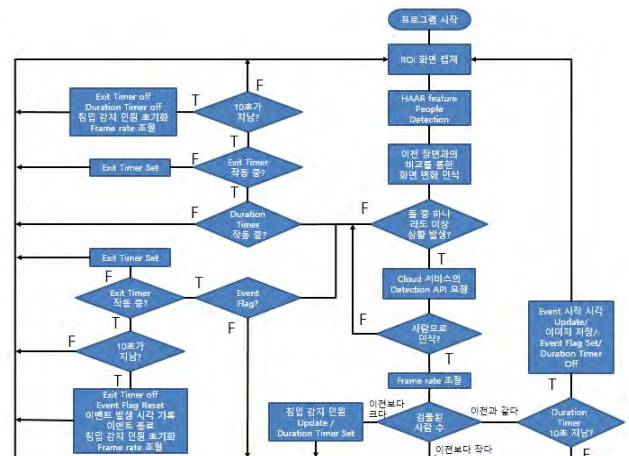
사람으로 인식이 되었을 경우 침입과 배회 이벤트에 따라 각각 다른 처리가 이루어진다.

4. 3. 6 이벤트 기록 및 처리

침입과 배회 이벤트 중 어떤 이벤트가 발생 하였으나에 따라 각각 다른 처리가 이루어진다. 침입의 경우 관심영역에 사람이 들어와 클라우드 영상분석 결과 침입이라고 판단되면 침입 이벤트가 발생하게 되어 적응형 프레임 캡처 기능(Adaptive frame capture rate)을 이용해 Frame rate를 높여 정확한 검출을 하게 될과 동시에 이벤트 발생 시각을 기록, 추가적인 인원이 관심영역에 들어올 것을 대비하여 침입 감지 인원을 업데이트하고 이미지를 저장한다. 이후 관심영역에서 사람들이 모두 나가게 되면 이벤트가 종료되어 Frame rate가 다시 낮아지고 또 다른 침입 상황을 준비한다. 배회의 경우 관심 영역에 사람이 들어온 후 10초가 경과하면 배회 이벤트가 발생하게 되고 마찬가지로 Frame rate를 높여 정확도를 높이고 침입이벤트와 같은 방법으로 이벤트를 기록 및 처리하게 된다.



(그림 4) 침입 이벤트에 대한 알고리즘 Flow chart



(그림 5) 배회 이벤트에 대한 알고리즘 Flow chart

5. 하이브리드 기반에서의 사람 감지 성능 평가

침입, 배회이벤트의 정의와 평가방법은 한국인터넷진흥원(KISA)의 지능형CCTV 솔루션 평가 시험 인증 안내서를 참고하였다.[5]

침입은 1명 또는 다수의 인원들이 관심영역에 침입한 순간을 정의하며 배회의 경우 1명 또는 다수의 인원들이 관심영역에 10초 이상 머물러 있는 경우로 정의한다.

침입은 관심영역에 사람몸 전체가 침입한 시각, 배회는 관심영역에 사람몸 전체가 들어온 후 10초가 경과한 시각을 이벤트 시작 시각으로 하며 이벤트 상황 발생 이전 2초, 이후 10초 이내에 이벤트 시작시각이 존재하여야 정상 검출로 판별한다.

특이 사항으로는 한명이 아닌 다수의 인원이 관심영역에 들어올 경우 마지막에 들어온 인원을 기준으로 정상 검출을 판별한다.

평가에 대한 점수로는 검출률, 정밀도를, 그리고 이 둘을 통합한 점수로 구성된다.

산출식	세부내용
$\text{점수} = \left(\frac{rp}{r+p} \right) \times 2 \times 100$	$\text{검출률}(r) : r = \frac{\text{정상검출}}{\text{정상검출} + \text{미검출}}$
	$\text{정밀도}(p) : p = \frac{\text{정상검출}}{\text{정상검출} + \text{오검출}}$

(표 1) 성능 평가 점수 산출식



(그림 6) 하이브리드 기반에서의 자체 사람 감지 성능평가 (좌)침입 이벤트, (우) 배회 이벤트



(그림 7) 배회 이벤트 도중 사람 인식 과정

※ 본 성능평가에서의 테스트 영상은 침입, 배회 각각 30개로 구성되었으며, 영상의 출처는 한국인터넷진흥원(KISA)에서 지능형 CCTV 솔루션 성능시험 인증을 위해 배포하는 배포용 영상DB를 이용.

6. 결론

하이브리드 지능형 침입 감지 시스템을 이용한 사회적 효과로는 클라우드 분석기술 기반의 영상 분석에 대해 신뢰도가 향상되고 로컬 기반에서의 영상분석환경을 벗어남으로서 추가적인 하드웨어(서버) 구축이 불필요해진다. 그리고 이상상황이 발생할 때만 클라우드에 분석요청을 하기 때문에 클라우드 사용료에 대한 부담 최소화가 범용적으로 이루어지고 기존의 DVR, NVR 시스템에 간단한 S/W 추가만으로 활용가능하기 때문에 자원의 재활용 효과 또한 기대할 수 있다.

참고문헌

[1] 경찰청, 『2017 범죄개요 1. 전체범죄 발생 및 검거 추세』, 2018, pp. 20
 [2] Viola, Paul, and Michael J. Jones. “Robust real-time face detection.” International journal of computer vision 57.2 (2004): 137-154.
 [3] Zhang, Liliang, et al. “Is faster r-cnn doingwell for pedestrian detection?.” arXiv preprintarXiv:1607.07032 (2016).
 [4] 이상훈, 추성권, 권기범, 조남익. (2017). 지능형 CCTV 시스템을 위한 보행자 검출과 재식별 기술. 한국통신학회지(정보와통신), 34(7), 40-47.
 [5] 한국인터넷진흥원, 『2018년 지능형CCTV 솔루션 성능시험인증안내서』, 2018, pp. 30-33