

# 지능형 영상분석 이벤트 탐지 기술동향

Technical Trends of Abnormal Event Detection in Video Analytics

정치윤 (C.Y. Jeong) 영상감시연구팀 선임연구원  
한종욱 (J.W. Han) 영상감시연구팀 팀장

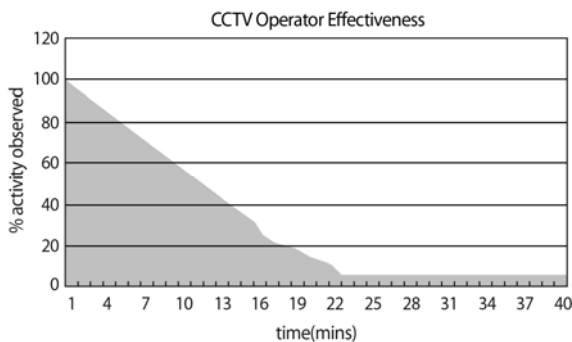
- I. 서론
- II. 지능형 영상분석 기술
- III. 이벤트 탐지 기술 발전 동향
- IV. 결론

최근 CCTV(Closed Circuit Television)의 설치가 증가하면서 효율적인 모니터링을 위하여 지능형 영상분석 기술에 대한 관심이 높아지고 있다. 지능형 영상분석 기술은 영상의 정보를 분석하여 자동으로 이상 행위를 탐지하고 관리자에게 경보를 전송하는 기술로써, 사고를 사전에 예방하고 사고가 발생한 경우에는 신속하게 대응하여 피해를 줄일 수 있게 해준다. 본고에서는 지능형 영상분석 기술이 탐지할 수 있는 이상 행위, 즉 이벤트를 그 목적에 따라서 보안, 비즈니스 인텔리전스, 객체인식으로 구분하여 현재 기술 수준을 살펴볼 것이다. 그리고 앞으로 지능형 영상분석에서 이벤트 탐지 기술의 발전 방향을 사람의 행동인식, 행위 기반 이상 현상 탐지, 군중 환경에서 이벤트 탐지, 지능형 영상분석 구조의 변화 등의 관점으로 구분하여 살펴보고자 한다.

## I. 서론

최근 기사에 따르면, 전국에 설치된 방범용 CCTV (Closed Circuit Television)는 3만 500여 대, 버스와 택시 등 대중교통에 설치된 CCTV는 20만여 대에 이르며, 민간이 설치한 CCTV까지 계산하면 최소 300만 대 이상의 CCTV가 설치되어 운용되고 있다고 한다[1]. 하지만, 2011년 경찰청 자료를 보면 서울 지역 경찰서 CCTV 관제센터에 근무하는 감시 인력은 1인당 평균 45대를 모니터링하는 것으로 조사된 것과 같이, 현재 설치된 CCTV의 수에 비해서 감시할 인력은 크게 부족한 것이 현실이다[2]. 또한, (그림 1)에서 보듯 2대 이상의 CCTV를 감시하는 사람의 경우에는 12분이 지나면 위험 상황의 45%, 22분이 지나면 위험 상황의 95%를 놓치기 때문에 사람을 통한 CCTV의 감시에는 한계가 있다[3].

이와 같은 문제를 해결하기 위하여 지능형 영상분석 기술이 개발되었다. 지능형 영상분석 기술은 영상분석을 통하여 사전 정의된 이벤트가 발생될 때 감시자에게 정보를 생성하여 알려주기 때문에, 모든 영상을 24시간 365일 쉼 없이 감시하지 않고, 정보가 발생할 때 해당 화면을 보고 실시간으로 상황을 판단하고 대처하면 된다. 따라서 지능형 영상분석 기술을 사용하면 한 사람이 다수의 CCTV를 감시할 수 있으며, 사고를 사전에 예방하고 사고가 발생할 경우에는 신속하게 인지하여 대응함으로써 피해를 줄일 수 있도록 해준다.



(그림 1) CCTV 감시자의 감시 효율[3]

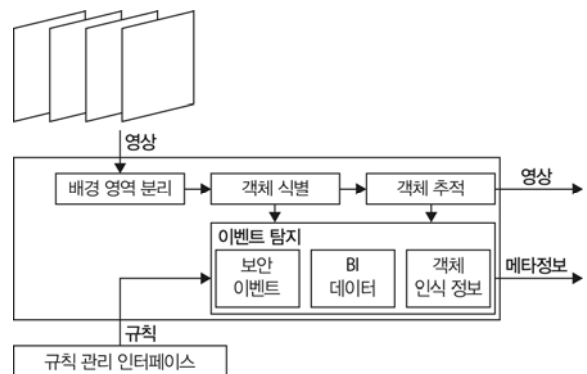
지능형 영상분석 기술의 결과로써 이상 현상에 대한 이벤트가 생성되고, 기술을 사용하는 사용자는 탐지된 이벤트만을 보기 때문에 본고에서는 지능형 영상분석 기술의 이벤트 탐지 기술에 초점을 맞출 것이다. 본고에서는 지능형 영상분석 기술에 대해서 소개하고, 현재 지능형 영상분석 기술에서 탐지할 수 있는 이벤트를 살펴볼 것이다. 그리고 이벤트 탐지 기술의 발전 방향을 살펴본 후, 결론을 내릴 것이다.

## II. 지능형 영상분석 기술

### 1. 지능형 영상분석 기술 소개

지능형 영상분석 기술이란 영상의 정보를 분석하여 자동으로 이상 행위를 탐지하는 기술로써, 일반적인 구조는 (그림 2)와 같이 배경 영역 분리 단계, 객체 식별 단계, 객체 추적 단계 및 사전 정의된 규칙을 기반으로 이벤트를 탐지하는 이벤트 탐지 단계로 구성된다.

배경 영역 분리 단계는 입력되는 영상에서 관심이 있는 전경 영역과 그 외의 배경 영역을 구분하여 활성 객체를 탐지하는 과정으로 초기에서 주로 이전 영상과의 현재 영상의 밝기 차이를 계산하여 분리하는 방법이 사용되었다. 하지만, 최근에는 배경 영역에 대한 특성 정보를 가우시안 또는 가우시안 믹스처 모델(GMM: Ga-



(그림 2) 지능형 영상분석 기술 구조도

ussian Mixture Model) 등을 사용하여 정교하게 모델링한 후, 전경 영역과 배경 영역을 구분하는 방법들이 많이 사용되고 있다.

객체 식별 단계는 배경 영역에서 전경 영역으로 판단한 객체 중에서 탐지된 물체가 사람인지 사물인지 여부를 구분하는 과정으로 현재 지능형 영상분석 기술은 사람(한 명, 그룹), 동물, 자동차 등을 주로 구분하고 있다.

객체 추적은 연속되는 영상에서 식별된 객체의 이동 경로를 찾는 과정으로, 칼만 필터, 파티클 필터, CAM-SHIFT(Continuously Adaptive Mean Shift) 등의 다양한 알고리즘을 바탕으로 추적하는 물체의 특징을 정의하는 방법, 추적하는 알고리즘의 조합 방법 등에 따라서 많은 방법들이 사용되고 있다.

이벤트 탐지 단계에서는 객체의 식별 정보 및 객체의 이동 정보를 바탕으로 보안 관리자가 정의한 규칙을 위반하는지 여부를 판단하여 이벤트를 탐지하고, 탐지된 정보를 메타 데이터 형태로 VMS(Video Management System)나 기타의 다른 보안 관리 서버로 전송하게 된다.

## 2. 지능형 영상분석 기술의 탐지 이벤트

현재 지능형 영상분석 기술에서 탐지하는 이벤트에 대한 정의 및 명칭은 업체마다 다르기 때문에, 본고에서는 탐지 이벤트를 보안 이벤트, 비즈니스 인텔리전스 이벤트, 객체인식 이벤트로 구분하여 현재 탐지 가능한 이벤트의 수준을 살펴보고자 한다. <표 1>은 본고에서 구분한 지능형 영상분석 기술의 탐지 이벤트 종류 및 해당 이벤트를 나타낸다.

### 가. 보안 이벤트

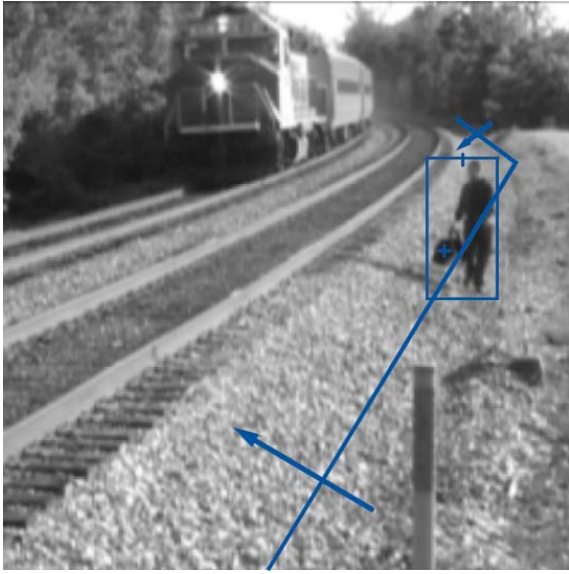
보안 이벤트는 보안을 목적으로 탐지되는 이벤트로 정의될 수 있으며, 지능형 영상분석 기술을 개발하게 된 동기를 제공하였다. 보안 이벤트는 영상에서 객체들의

<표 1> 지능형 영상분석 기술의 탐지 이벤트

이벤트 유형	설명	해당 이벤트
보안 이벤트	영상에서 객체의 움직임, 상태 정보를 분석하여 사용자가 정의한 규칙에 위반되는 행위 탐지	<ul style="list-style-type: none"> <li>- Regional entrance</li> <li>- Tripwire</li> <li>- Fence trespassing</li> <li>- Unattended baggage</li> <li>- Object Removal</li> <li>- Loitering</li> <li>- Stooped vehicle</li> <li>- Counting in a crowd</li> <li>- Crowd density analysis</li> <li>- Overcrowding</li> <li>- Traffic congestion</li> <li>- Removed objects (crowded area)</li> </ul>
BI 이벤트	영상에서 분석을 통하여 비즈니스에 유용한 통계 정보 생성	<ul style="list-style-type: none"> <li>- People counting</li> <li>- Vehicle counting</li> <li>- Dwell time</li> <li>- Queue line analysis</li> <li>- Room/area occupancy</li> </ul>
객체 인식	영상에서 객체를 검출하고 인식하여 객체의 식별 정보 생성	<ul style="list-style-type: none"> <li>- Face recognition</li> <li>- License plate recognition</li> </ul>

움직임, 상태 정보를 기반으로 관리자가 정의한 규칙에 위배되는지 여부를 검사하여 탐지되며, 대표적인 이벤트로 트립와이어 기반의 침입 탐지 이벤트가 있다. 트립와이어 기반의 침입 탐지는 감시 구역에서 관리자가 설정한 가상의 선과 움직이는 방향을 객체의 움직임과 비교하여 설정된 규칙에 해당되는 경우 비정상 행위로 판단하고 이벤트를 생성하게 된다. (그림 3)은 철길에 사람이 들어가는 것을 인지하기 위해서 관리자가 철길 주변에 가상의 선과 철길로 들어가는 방향을 규칙으로 설정했을 때 사람의 침입을 탐지한 모습이다. 이 외에 다른 보안 이벤트로는 설정된 영역으로 객체의 출입 및 불법 침입, 존재하지 않던 객체가 새로 생겨나거나, 존재하던 객체가 사라지는 등의 상태 변화, 잘못된 방향으로의 움직임, 동반되지 않는 짐, 잘못 주차된 차, 배회 등이 있다. 보안 이벤트의 경우 대부분의 지능형 영상분석 제품들이 지원하고 있다.

최근에는 비교적 사람들의 수가 많지 않은 단순한 장



(그림 3) 트립와이어 기반 이상 행위 탐지[4]

면에서 벗어나 군중이 있는 환경에서 이벤트를 탐지하려는 기술들이 개발되고 있다. 군중들이 움직이는 상황에서 비정상 행위를 탐지하는 것은 객체들의 겹침 현상으로 인하여 각 객체의 정보를 정확하게 얻기가 힘들기 때문에 현재 활발히 연구되고 있는 분야이다. (그림 4)는 기존의 지능형 영상분석 기술들이 분석하고 있는 일반적인 상황과 군중 상황에서 사람의 수를 세는 모습을 나타내고 있으며, 군중 상황에서의 보안 이벤트를 탐지하는 것이 쉽지 않음을 알 수 있다.



(a) 일반적인 상황

(b) 군중 상황

(그림 4) 피플카운팅[5]

#### 나. 비즈니스 인텔리전스 이벤트

비즈니스 인텔리전스 분석은 객체 식별, 객체 추적 등과 같은 영상분석을 통하여 영상에 존재하는 객체의 수,



(그림 5) BI를 위한 큐라인 분석[6]

출입자 수, 체류 시간, 공간 점유율 등과 같이 비즈니스에 도움이 될만한 정보를 생성하는 것을 의미하며, 최근 많이 개발되어 적용되고 있다.

쇼핑몰, 카지노, 경기장과 같이 출입자 수에 대한 정보가 중요한 곳에서는 지능형 영상분석을 통하여 생성되는 사람 수에 대한 정보가 유용할 수 있다. 또한 사람들이 많이 다니는 곳의 이동 패턴, 시간대별 사람의 수 등을 분석하여 상권을 분석할 때도 사용될 수 있다. (그림 5)는 큐라인 분석의 한 예를 나타내고 있으며, 지능형 영상분석을 통하여 줄이 일정 수준 이상으로 길어지는 경우 직원을 더 배분하여 효율성을 높일 수 있다.

BI(Business Intelligence)를 위한 지능형 영상분석 기술을 사용하면 기존에 분석하지 못했던 유용한 정보들을 생성할 수 있으며, 이는 회사의 자원을 최적으로 배분할 수 있게 해준다. 따라서, 지능형 영상분석 기술에서 BI 분석이 차지하는 비중은 점점 더 커질 것으로 예상된다.

#### 다. 객체인식

얼굴 인식, 자동차 번호 인식과 같은 객체 신원을 확인하는 인식 기술들은 인식 거리의 제한, 영상 내 인식 가능한 객체의 크기 제한 등 제약 조건이 많은 환경에서 개발되어 지능형 영상분석과 독립된 분야로 인식되어 왔다. 하지만 최근 객체인식 기술들이 기존에 가지던 다양한 제한 조건을 극복하고 비제한적인 환경에서 객체를 인식할 수 있는 기술로 발전하고 있다. 또한 지능형

영상분석 기술에서 객체인식 기술을 포함하여 제품들이 출시되고 있으며, 향후 다수의 제품들이 객체인식 기술을 지능형 영상분석 기술의 한 기능으로 포함할 것으로 예상된다.

### III. 이벤트 탐지 기술 발전 동향

초기의 지능형 영상분석의 이벤트 탐지 기술은 많은 오탐으로 인하여 사용자들의 외면을 받았으며, 현재에는 사용자의 요구사항과 기술 수준의 격차를 줄여가고 있다. 하지만 여전히 사용자들은 이벤트 탐지 기술이 사람이 보고 판단하는 것과 같은 능력을 가지길 요구하고 있다. 따라서 사용자의 요구사항을 만족시키기 위해서는 사람의 행동인식, 행위 기반 이상 현상 탐지, 군중 환경에서 이벤트 탐지, 지능형 영상분석 구조의 변화 등이 필요하며, 다음 절에서 이와 같은 관점에서 이벤트 탐지 기술의 발전 방향에 관해서 살펴보고자 한다.

#### 1. 사람의 행동인식

사람의 행동인식에서 인식하는 동작은 신체 일부의 움직임에 의미하는 제스처, 한 사람에 대한 다양한 제스처의 조합으로 정의되는 액션, 두 객체 사이에서 발생하는 행동인 인터랙션, 다수의 객체로 구성된 그룹에서 발생하는 그룹 행동으로 구분될 수 있다. 현재 사람의 행동인식에 대한 접근 방법은 싱글 레이어드 기반의 방법

과 계층적 기반의 방법으로 구분할 수 있다[7].

싱글 레이어드 기반의 방법은 영상에서 사람의 행동을 직접 구분하는 방법으로써 상대적으로 간단하고 짧은 사람의 동작을 인식하는 것이 목적이기 때문에 걷기, 뛰기 등의 비교적 단순한 제스처, 액션 등의 행동인식에 주로 사용된다.

계층적 기반의 방법은 상대적으로 인식하기 쉬운 단순한 행동 정보를 인식한 후, 행동 정보들을 조합함으로써 고수준의 행동을 판단하는 방법이기 때문에 복잡한 사람의 행동을 인식하는 데 주로 사용된다. 예를 들어 싸움과 같은 행동의 경우 영상에서 연속적인 펀칭과 킥의 행위를 탐지함으로써 인식할 수 있다.

최근의 영상분석 기술은 (그림 6)과 같은 테스트데이터 세트에서 사람의 액션을 95% 정도의 정확도로 인식할 수 있는 수준에 접근하였으며[8], 최근에는 인터랙션, 그룹 행동에 대한 인식을 하려는 시도를 하고 있다. 하지만 현재 학계에서 진행되고 있는 연구는 객체들 간의 겹침이 많이 존재하지 않고 배경의 변화가 심하지 않은 제약적인 환경에서 이루어지고 있기 때문에 실제 영상 감시 환경에 바로 적용하기에는 어려움이 있다. 또한 위에서 분류된 방법들의 경우 오프라인으로 수행되는 방법들이 대부분이며, 위의 방법을 개선하여 실시간으로 적용하려는 시도들도 일어나고 있다. 또한 GPU (Graphic Processing Unit)의 병렬처리 성능을 통하여 실시간으로 행동인식을 수행하려는 연구도 있었다[9]. 최근 사람의 행동인식에서 또 다른 시도는 사람의 행동



(그림 6) 사람의 행동인식 데이터 세트 (위) KTH, (아래) UT-interaction[8]

을 분류하는 것에서 벗어나 사람의 행동을 예측하는 기술에 대한 연구도 진행되고 있다[10].

기존의 객체의 크기, 위치 정보만으로는 현재 이슈가 되고 있는 폭력 등과 같은 이상 행위를 탐지하지 못하기 때문에, 앞으로의 이벤트 탐지 기술은 사람의 행동을 인식 기술을 사용하여 탐지하는 이벤트의 종류를 다양화하고, 탐지된 이벤트의 정확도를 향상시키는 방향으로 발전할 것이다.

## 2. 행위 기반의 이상 현상 탐지

현재 상용화된 대부분의 지능형 영상분석 기술은 관리자가 사전에 설정한 규칙에 위반되는 행동을 탐지하는 규칙 기반의 이상 행위 탐지 방법이다. 규칙 기반의 탐지 방법은 기술을 사용하는 사용자가 규칙을 설정하고 지속적으로 관리해야 하며, 규칙을 관리하는 동안 설정의 오류 등이 발생할 수 있다. 또한 관리자가 설정한 규칙을 회피하는 경로가 존재하는 경우에는 탐지할 수 없는 단점이 있다.

이런 단점을 보완하기 위해서 영상에서 일어나는 객체들의 행동 패턴을 학습하여 행위 기반으로 이상 현상을 탐지하는 방법들에 대한 연구가 진행되고 있다. BRS Labs사에서 출시한 AISight가 대표적인 행위 기반의 이상 현상 탐지 솔루션으로 영상분석을 통하여 사람들의 행위를 관찰하고 일반적인 행위에서 벗어나는 경우 경보를 생성하는 방식이다. AISight는 관리자가 규칙을 정의할 필요가 없으며 영상 내 특정 장소에서의 행위, 하루의 시간대별 변화, 한 주에서 요일별 변화 등을 자율 학습(unsupervised learning)을 통하여 정상 행위를 기억하고, 객체의 위치와 움직임이 정상 행위에서 벗어나는 경우 이상 행위로 인식하게 된다. (그림 7)은 AISight에서 탐지하는 이벤트의 한 예로써 왼쪽의 영상은 차들이 정상적으로 운행하는 방향을 인지하여 중앙선을 넘어서 역방향으로 차단기를 통과하려는 차를 탐지하는 모습이다. 오른쪽 영상의 경우 보안 검색대에서 운전자



(a) 잘못된 방향의 트래픽 (b) 보안 검색대에서 이상 행동  
(그림 7) 행위 기반 이벤트 탐지[11]

가 내리는 것은 일반적인 상황이지만 조수석에 앉은 사람이 내리는 것은 비정상 행위이기 때문에 조수석에서 내리는 사람을 탐지한 모습이다. 이와 같은 이벤트들은 관리자의 정교한 규칙 설정에 의해서도 탐지할 수 있지만, 행위 기반의 탐지 기법은 관리자의 개입 없이 자동으로 탐지할 수 있다는 장점이 있다.

기존의 IT 보안 분야에서 규칙 기반의 IDS(Intrusion Detection System)/IPS(Intrusion Prevention System) 등의 네트워크 보안 장비가 출시된 이후, 행위 기반의 탐지 제품들로 이동했듯이, 영상감시 분야에서도 행위 기반의 이상 현상 탐지 방법들에 대한 연구와 제품의 출시가 앞으로 이루어질 것으로 생각된다.

## 3. 군중 환경에서의 이벤트 탐지

현재의 이벤트 탐지 기술들은 영상에서 객체를 찾은 후 객체를 추적하면서 설정된 규칙을 위반하는지 여부를 판단하는 객체 중심의 방법들이 주를 이루었다. (그림 8)과 같이 군중들이 움직이는 환경에서는 객체들 간

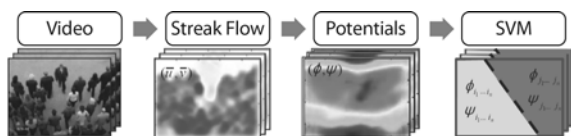


(그림 8) 시공간 기반 군중 속에서의 이상 행위 탐지[12]

의 겹침 현상이 많이 발생하기 때문에 개별 객체에 대한 정확한 정보를 얻는 것은 힘들기 때문에 기존의 영상분석 기술을 적용하기가 어렵다. 따라서 군중들이 움직이는 환경에서 이상 행위를 탐지하기 위한 지능형 영상분석 기술에 대한 연구가 진행되고 있으며, 그 목적에 따라서 군중의 밀도 추적 및 사람 수 카운팅과 집단의 행동인식으로 구분될 수 있다. 군중이 움직이는 환경에서 밀도 추적, 카운팅과 같은 기술은 현재 호주의 iOmniscient사에 의해서 제품화되어 출시 되고 있으며, 다른 업체들도 해당 기술을 개발하여 적용할 것으로 예상된다. 하지만 집단의 행동인식에 대한 기술은 상용화되지 않고 학계에서 연구하는 수준에 머물러 있다.

집단의 행동을 인식하는 분야는 현재 활발히 연구되는 분야로서 시공간 기반의 분석 방법과 유체 역학 기반의 방법이 있다. 시공간 기반의 분석 방법[12]은 영상의 X, Y축, 시간 축으로 구성된 3차원 공간을 큐브로 분할한 후, 각 큐브에 대해서 모션의 변화와 같은 특징 정보를 모델링하고 큐브의 연관관계에 대해서 통계적 분석을 통하여 이상 행위를 찾는 방법이다. (그림 8)에서 파란색 사각형이 이상 행위로 탐지된 모습을 나타내며, 주로 다수의 사람이 움직이는 방향과 다르게 움직이는 보행자, 보행자의 불규칙한 움직임, 군중들의 흐름을 방해하는 사람 등의 이벤트를 탐지할 수 있다.

유체 역학 기반의 방법은 영상에서 시간에 따른 특징 점들의 흐름을 추적하여 모션 정보를 분석함으로써 화면에서 발생하는 이상 행위를 인식하는 방법이다. (그림 9)는 최근 제안된 유맥선(streaklines)을 이용한 군중 속에서의 이상 행위 탐지 방법을 순서도를 나타낸다. 제안된 방법[13]의 경우 영상으로부터 군중의 지역적인 모션 변화, 시간에 따른 모션의 변화를 잘 반영하는 스트



(그림 9) 유체역학 기반 군중 속에서 이상 행위 탐지[13]

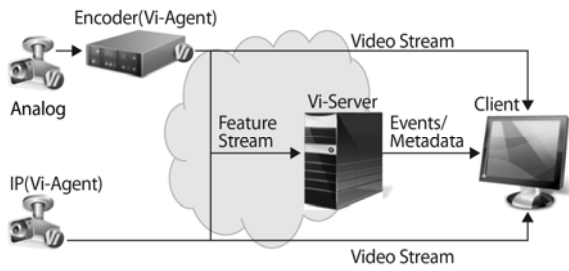
릭 플로우를 계산한 후, 군중의 행동 변화를 유선 함수와 속도 함수로 정의하여 모델링하였다. 두 개의 함수로 모델링된 값들은 비정상 행위를 구분하는 지지벡터 기반 분류기의 입력 값으로 사용되며, 이를 통하여 군중 속에서 발생하는 비정상 행위를 판단하게 된다.

사람들이 화재를 피해서 움직이고, 사람들 간의 싸움으로 군중들이 모여드는 것처럼 군중 속에서 발생하는 비정상 행위의 경우 재난이나 범죄 등과 관련성이 많고, 기존 기술의 적용이 어렵기 때문에 앞으로 이에 대한 연구가 많이 진행될 것이다.

#### 4. 지능형 영상분석 구조의 변화

지능형 영상분석 기술의 적용 구조는 영상을 분석하여 이벤트를 탐지하는 지점에 따라서 서버 기반의 방법과 에지 기반의 방법으로 구분된다. 에지 기반의 방법은 카메라 또는 카메라와 연결된 영상 인코더에서 영상을 분석하는 방법으로 영상을 전송하기 위해서 압축하는 과정에서 발생하는 신호의 왜곡 없이 영상을 분석할 수 있다는 장점이 있지만, 임베디드 장비의 연산 능력이 제한되기 때문에 복잡도가 높은 영상분석 기법을 적용할 수 없고 추후 기능 업그레이드가 어려운 문제점이 있다. 서버 기반의 방법은 카메라로부터 영상을 전송받아서 지능형 영상분석을 수행하는 집중형 구조로서 카메라에 대한 의존성이 없기 때문에 기존 시스템에 적용하기 용이하지만, 서버 한 대가 처리할 수 있는 카메라 수가 제한적이기 때문에 확장성에 문제가 있다.

최근 서버 기반의 방법의 문제점인 높은 네트워크 대역폭 사용율, 서버당 처리할 수 있는 채널의 수 문제, 에지 기반의 방법의 디바이스 연산 능력으로 고수준 영상분석 기법의 적용 제한 등의 문제점을 해결하기 위하여 두 방법의 장점만을 결합한 하이브리드 기반의 영상분석 방법이 등장하였다. 하이브리드 기반의 지능형 영상분석의 대표적인 방법으로는 AgentVi사의 IPoIP(Image Processing over IP)가 있으며, 그 구조는 (그림 10)과



(그림 10) AgentVi사의 IPoIP 구조[14]

같다.

IPoIP의 구조에서 카메라는 영상 정보를 전송하는 모드와 영상의 특징 정보를 전송하는 모드로 구성되며, 모드의 변화는 서버에 의해서 제어된다. 특징 정보를 전송하는 모드에서 카메라가 서버로 전송하는 특징 데이터는 이벤트가 발생된 특정 영역의 영상 정보, 검출된 객체의 형태, 크기, 속도, 방향, 색상, 명도 등의 다양한 형태가 될 수 있다. 만약, 트립와이어 기반의 이벤트 탐지를 서버에서 수행할 경우에는 검출된 객체의 위치, 방향, 크기 정보만을 필요하기 때문에 객체의 영상, 색상, 명도 등의 불필요한 정보를 전송하지 않으며, 이를 통하여 전송되는 데이터의 양을 줄일 수 있다. 또한 모든 프레임 전송하지 않고, 특징 데이터를 전송하기 전에 영상에서의 모션 정보를 계산하여 서버에서 설정한 임계값을 넘는 모션이 발생한 경우에만 특징 정보를 전송함으로써 네트워크 데이터의 사용률을 더 낮출 수 있다.

서버에서는 카메라에서 전송할 특징 정보 및 카메라의 전송 모드를 설정하게 되며, 카메라에서 전송된 특징 정보를 사용하여 이벤트를 탐지하게 된다. 서버에서 이벤트를 탐지하게 되면 해당 카메라의 전송 모드를 영상 전송 모드로 변경하게 되며, 이를 통하여 관리자는 이벤트에 대한 검증 수행하게 된다. 이와 같은 방법을 통하여 서버 기반의 방식이 일반적으로 4~16대의 카메라를 수용할 수 있는데 반하여, IPoIP의 구조에서 한 대의 서버가 200대의 카메라를 수용할 수 있다[14].

IPoIP와 같은 하이브리드 기반의 방법은 영상에서의 특징 추출과 이벤트 탐지 기능을 카메라와 서버로 분배

함으로써 서버의 카메라 수용 능력을 높이고, 카메라 연산 능력의 한계를 극복할 수 있는 장점이 있다. 하이브리드 기반의 방법의 장점을 활용하기 위해서는 서버와 카메라의 협업이 중요하기 때문에, 서버에서 카메라의 연산 기능 및 전송 모드를 정교하게 제어하기 위한 방법에 대한 연구가 많이 진행될 것으로 기대된다.

#### IV. 결론

본고에서는 지능형 영상분석 기술에 대해서 소개하고, 현재 지능형 영상분석이 탐지하는 이벤트들에서 살펴보았다. 그리고 사용자들의 요구사항을 만족시키기 위해서 이벤트 탐지 기술이 발전해야 할 방향에 관하여 기술하였다.

앞으로 지능형 영상분석 기술에서 사람의 행동을 인식하여 개념적인 이벤트(폭력, 납치 등)를 탐지하기 위한 방향으로의 연구가 진행될 것으로 기대된다. 또한 기존의 규칙 기반 이벤트 탐지 방법에서 벗어나 사람의 개입 없이 행위 기반의 이벤트 탐지 방법이 주류로 등장할 것이다. 그리고 군중들이 움직이는 환경에서 영상의 전역 정보를 분석하여 이상 행위를 인식하는 이벤트 탐지 기술이 시장에 보급될 것이다. 또한 기존의 서버/에지 기반의 지능형 영상인식 구조에서 벗어난 새로운 구조를 가지는 기술들이 속속 개발될 것이다.

#### 용어해설

**트립와이어(Tripwire)** 전선에서 침입하는 적들이 건드리면 폭발물이나 신호탄 등을 터뜨려서 적의 침입을 알 수 있게 해주던 철선을 의미하며, 지능형 영상분석에서는 영상에서 가상으로 설정한 선을 의미하며 해당 선을 지나치거나 넘는 경우 이벤트가 발생함.

#### 약어 정리

BI	Business Intelligence
CCTV	Closed Circuit Television
CAMSHIFT	Continuously Adaptive Mean Shift

GMM	Gaussian Mixture Model
GPU	Graphic Processing Unit
IPoIP	Image Processing over IP
IDS	Intrusion Detection System
IPS	Intrusion Prevention System
VMS	Video Management System

## 참고문헌

- [1] 뉴스스, "CCTV 300만 시대," 2011. 10. 18.
- [2] YTN, "CCTV 감시인력 크게 부족," 2011. 9. 28
- [3] S. Fleck and W. Straber, "Privacy Sensitive Surveillance for Assisted Living-A Smart Camera Approach," *Handbook of Ambient Intelligence and Smart Environments*, H. Nakashima, H. Aghajan, and J.C. Augusto, Ed., Springer, 2010, pp. 985-1014.
- [4] Pelco, "ObjectVideo<sup>®</sup> Analytics for Cameras with Sarix<sup>TM</sup> Technology," 2012, p. 1. [http://kondorsecurity.com/store/media/pdf/Pelco\\_ObjectVideo\\_Analytics\\_Sarix\\_Technology\\_spec.pdf](http://kondorsecurity.com/store/media/pdf/Pelco_ObjectVideo_Analytics_Sarix_Technology_spec.pdf)
- [5] iOmniscient, "Comprehensive Features," 2010. [http://www.iomniscient.com/index.php?option=com\\_content&view=article&id=107&Itemid=79](http://www.iomniscient.com/index.php?option=com_content&view=article&id=107&Itemid=79)
- [6] 3VR, "3VR VIP ANALYTICS," 2011, p. 2. [http://www.3vr.com/sites/default/files/3VR\\_VIP\\_AnalyticsD\\_atasheet\\_110209.pdf](http://www.3vr.com/sites/default/files/3VR_VIP_AnalyticsD_atasheet_110209.pdf)
- [7] J.K. Aggarwal and M.S. Ryoo, "Human Activity Analysis: A Review," *ACM Comput. Surveys*, vol. 43, no. 3, Apr. 2011.
- [8] T.-H. Yu, T.K. Kim, and R. Cipolla, "Real-time Action Recognition by Spatiotemporal Semantic and Structural Forests," *Proc. British Mach. Vision Conf.*, 2010.
- [9] M. Rofouei, M. Moazeni, and M. Sarrafzadeh, "Fast GPU-based Space-Time Correlation for Activity Recognition in Video Sequences," *Proc. IEEE/ACM/IFIP Workshop Embedded Syst. Real-Time Multimedia*, 2008.
- [10] M.S. Ryoo, "Human Activity Prediction: Early Recognition of ongoing Activities from Streaming Videos," *IEEE Int. Conf. Comput. Vision*, Nov. 2011.
- [11] BRS Labs, "Alert Examples," <http://www.brslabs.com/product-demo>
- [12] L. Kratz and K. Nishino, "Anomaly Detection in Extremely Crowded Scenes using Spatio-Temporal Motion Pattern Models," *IEEE Conf. Comput. Vision Pattern Recognit.*, 2009.
- [13] R. Mehran, B.E. Moore, and M. Shah, "A Streakline Representation of Flow in Crowded Scenes," *Eur. Conf. Comput. Vision*, 2010.
- [14] AgentVi, "Video Analytics Architectures," 2012. <http://www.agentvi.com/20-Technology-58-Video-Analytics-Architectures>