**실내 밀집 구역에서 마스크 착용 유무를 감지하는 지능형 CCTV**

**(기초조사서)**

팀: 2017103962 김동규, 2019102198 오인제, 2017105715 LI ENGUANG

목 차

**Ⅰ. 서론**

1) 지능형 CCTV

2) 지능형 CCTV의 원리

**Ⅱ. 프로젝트 내용**

1) 프로젝트 목표

2) 구조

3) 핵심 기능

**Ⅲ. 기능 구현**

1) Main Camera

2) Sub Camera

**Ⅳ. 결론 및 기대효과**

**Ⅴ. 개선점**

**1. 서론**

**1.1. 지능형 CCTV**

최근 몇년간 스마트시티 프로젝트 등이 추진되면서 공공기관이 운영하는 CCTV 대수는 더욱 늘어나고 있다. 또한 민간시장에서도 홈-security 등을 위해 소형 CCTV 디바이스 설치가 확대되는 추세다. 이처럼 CCTV가 늘어나면서 관제 인력이 부족해지는 문제가 발생하고 있다. CCTV 설치 대수가 기하급수적으로 늘어나면서 모든 영상을 사람이 실시간으로 감시하는 것은 불가능에 가까워졌다. 이에 따라 최근에는 영상 분석 기술을 활용한 지능형 CCTV 도입이 일반화되는 추세다.

지능형 CCTV는 실시간으로 영상을 분석해 움직임이 있는 물체를 감지·분류하고, 사전에 정의된 사건을 감지해 자동 식별하고 이를 감시자에게 알려 사람이 24시간 영상을 감시할 필요없이 선별적인 관제가 가능하다. 이는 실제로도 식별하는 것이 중요한 상황이나 행위를 놓치는 것을 방지하는 효과가 있어 실시간으로 위험 상황에 대한 대응이 가능하게 돕는다. 또한 선별적으로 유의미한 상황이 담긴 영상만을 저장하고 관리할 수 있기에 더 효율적이다.

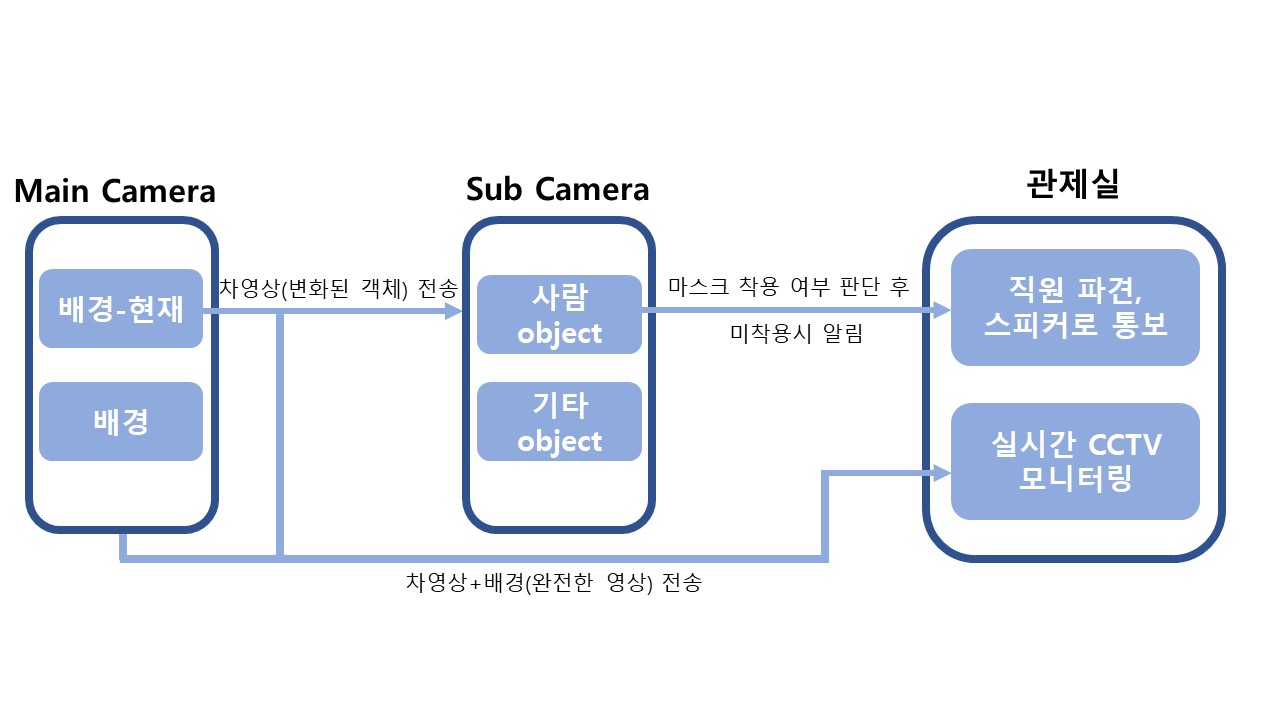
**1.2. 지능형 CCTV의 원리**

딥 러닝 기반의 '행동인식 기술', 행동인식 기술이란 인간의 신체 동작을 자동으로 인식하는 기술이다. 딥 러닝으로 영상 기반의 행동을 인식하기 위해서는 CNN(Convolution Neural Network) 기반의 방법과 LSTM(Long Short Term Memory) 기반의 방법이 있다. 이 두 방법은 모두 움직이는 모션 정보를 추출하기 위해 비디오 영상의 여러 프레임을 종합적으로 분석하는 데에 중점을 두는데, CNN 기반 방법은 연속적인 프레임을 쌓아 두고 3D 커널을 적용함으로써 움직임 정보를 잘 축적한 특징맵을 생성하고, LSTM 기반 방법은 영상을 프레임별로 분석하되 과거 프레임에 대한 기억을 Cell이라는 구조를 통해 조절한다. 즉, 이전 프레임이 행동인식에 있어 중요한 정보를 가지고 있다고 판단된다면 이 정보를 유지하고, 이전 프레임에 중요한 정보가 없다고 판단된다면 해당 프레임의 정보를 잊는다. 이렇게 CNN과 LSTM 방법 모두 연속된 비디오 프레임에서 정보를 분석하여 특징 벡터를 추출한 후 행동에 대한 스코어 값을 계산하여 행동을 식별한다.

**2. 프로젝트 내용**

**2.1. 프로젝트 목표**

동영상 압축 기능은 스마트폰 등 모든 IT 정보가전 제품군에 탑재된 매우 중요한 기능이며, 특히 CCTV 등의 Surveillance System에는 필수적인 기능이다. 지능형 CCTV에는, Moving Object Tracking, Synchronized Main and Sub PTZ-Camera System 등이 주요 필수 요소이며 Killer-Application으로서 부각되기 위해서는, 필히 Cost-effective하여야만 한다. 특히 Synchronized Main & Sub PTZ-Camera System에서, Main Camera는 Wide-viewed angle로 Event를 기록하고, Sub-Camera는 Main Camera에서 포착된 사물의 상세 영상을 Close-up해서 기록하는 방식이 필요하며, 이를 위해서 Object Tracking 기능이 필수적으로 필요하다. 본연구에서는 동영상 압축 과정에서 발생되는 정보를 이용하여 (별도의 Object Tracking 기능을 탑재하지 않거나 해당 기능을 최소화시킨 상태에서, Sub PTZ-Camera에 필요한 Object Tracking 기능을 제공하여 Cost-Down을 실현하고, 동영상 압축과정에서 발생되는 정보를 효율적이고 지능적으로 이용하여 우수한 성능을 지닌 Intelligent CCTV System을 구현한다.

**2.2. 구조**

(그림-1) 지능형 CCTV 구조

**2.3. 핵심 기능**

**-** Main Camera

**>** 건물 CCTV는 감시하는 배경이 실내에 고정되어 있으므로 배경은 변하지 않음을 활용

> (현재 영상) - (기존 배경)으로 차영상을 구하게 되면 움직임이 생긴 객체를 얻을 수 있다.

> Sub Camera에게는 차영상만, 모니터링이 필요한 관제실엔 완전한 영상을 보낸다.

- Sub Camera

> 전달받은 객체를 사람과 기타 필요 없는 객체로 구분한다.

> 사람 객체에서 마스크를 tracking하여 마스크 착용 여부를 판단한다.

> 미착용 적발 시 관제실에 알람을 보내 이벤트 발생을 알린다.

- 관제실

> Main Camera에서 오는 실시간 CCTV 영상을 모니터링 할 수 있다.

> Sub Camera에서 적발한 마스크 미착용 사람에게 통보, 직원 파견 등 조치를 취한다.

**3. 기능 구현**

**3.1. Main Camera**

**-** 역할: 영상의 압축과 관제실에서 관측하는 Wide View 영상을 전송

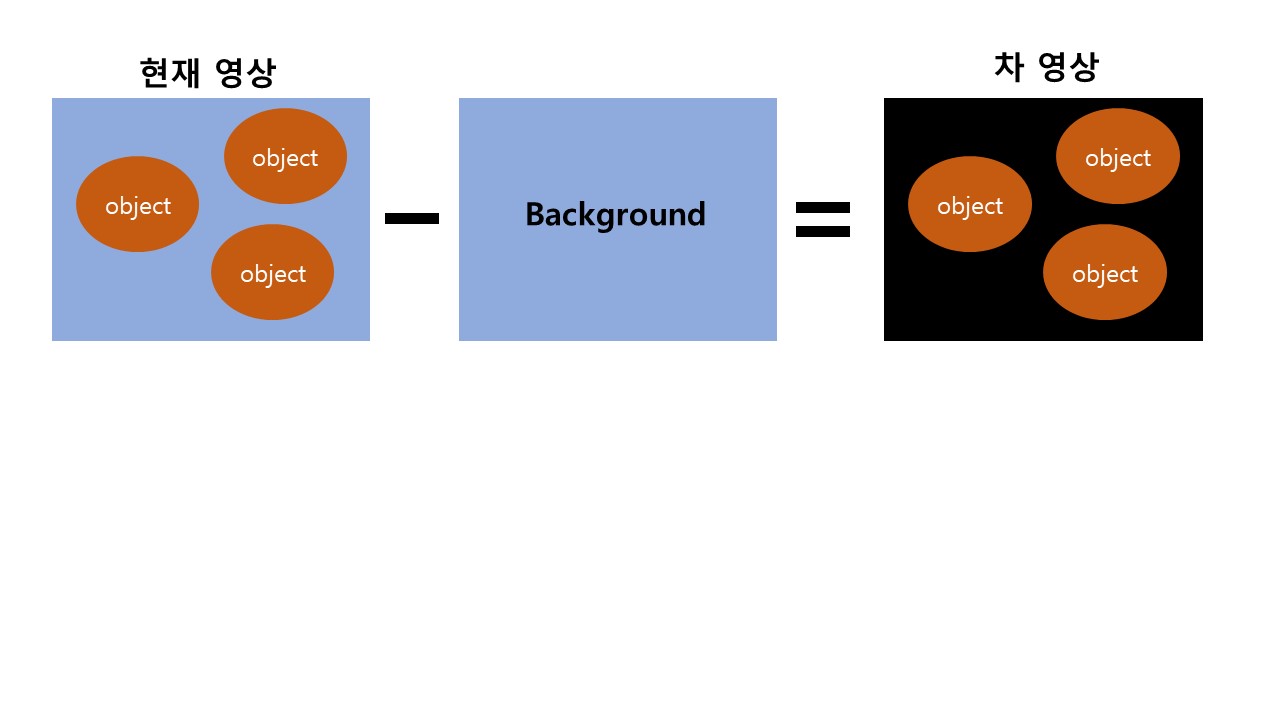
- 필요 도구: CCTV 역할을 할 고정된 카메라

> 카메라: 라즈베리파이 사용(여건이 안될 경우 웹캠 사용)

- 구현 방법: 차영상 기법 활용

> 기대 효과: Sub Camera가 모든 영상을 처리하는 것이 아닌 객체만 추출된 영상을 사용하여 Object Tracking의 효율을 증대시킨다.

> 객체 분할: 차영상 기법을 사용해 배경은 mask되고 객체만 있는 영상을 구한다



(그림-2) 차영상

**3.2. Sub Camera**

- 역할: Main Camera에서 전송 받은 영상에서 객체 검출, 추적, 상황인식

> 사람과 그 외의 객체를 검출

> 사람 객체에서 마스크를 검출하여 착용 여부 판단

> 마스크 착용을 하지 않은 사람을 추적, 관제실에 전송

- 목표로 정한 CCTV에서 Sub Camera가 줌인 하는 기능이 있지만 팀에서 활용할 수 있는 카메라로는 한계가 있을 경우 카메라 한 대로 Main, Sub Camera에서 처리하는 영상을 띄워서 과정을 확인

- 구현 방법: 특징점(feature)을 활용해 학습 알고리즘을 이용

**4. 결론 및 기대효과**

지능형 CCTV를 도입하면 적은 인원으로도 많은 CCTV를 동시에 감시할 수 있어 업무효율이 크게 증가한다. 모든 영상을 실시간으로 감시하기보다 확인이 필요한 장면만을 알려줌으로써 업무 부담을 줄이는 효과도 있다. 또한 중요한 영상만을 저장하는 것으로 데이터 용량 부족의 문제도 해결할 수 있다. 오작동이나 허위 알람도 줄어든다.

**5. 개선점**

기초 조사서에서 설계했던 내용 이외에 프로젝트 진행 중, 꾸준히 개선해야할 점을 찾아야한다.

- Main Camera에서의 압축 효율

- Sub Camera의 Object Tracking 정확도

- 마스크 미착용 사람에 대한 대응, 조치