

# 스마트폰 기반의 무인 영상 추적 시스템 연구

안병태  
안양대학교 교양대학 교수

## A Study on Unmanned Image Tracking System based on Smart Phone

Byeong-tae Ahn  
Professor of Liberal & Arts College, Anyang University

**요약** 최근 스마트폰 기반의 영상 이미지 추적을 통한 무인 녹화 시스템은 급속히 발전하고 있다. 기존의 제품 중 적외선 신호를 이용하여 촬영 대상을 자동으로 추적 및 회전하여 녹화하는 시스템은 일반 사용자가 사용하기에는 매우 고가이다. 따라서 본 논문에서는 스마트폰을 사용하는 사용자라면 누구나 자동 녹화가 가능한 모바일용 무인 녹화 시스템을 제안한다. 본 시스템은 상용 Mobile 카메라, 좌우로 카메라를 움직이는 서보모터(Servo Motor), 모터를 제어하는 마이크로 컨트롤러 그리고 동영상 오디오 입력을 담당할 상용 무선 블루투스 이어셋(Wireless Bluetooth Earset)으로 구성된다. 본 논문에서는 스마트 폰을 이용하여 영상 추적을 통해 무인 녹화가 가능한 시스템을 설계하였다.

**주제어** : 영상 추출, 딥러닝, 이미지 추출, 무인 영상, 블루투스

**Abstract** An unattended recording system based on smartphone based image image tracking is rapidly developing. Among the existing products, a system that automatically tracks and rotates the object to be photographed using an infrared signal is very expensive for general users. Therefore, this paper proposes a mobile unattended recording system that enables automatic recording by anyone who uses a smartphone. The system consists of a commercial mobile camera, a servomotor that moves the camera from side to side, a microcontroller to control the motor, and a commercial wireless Bluetooth Earset for video audio input. In this paper, we designed a system that enables unattended recording through image tracking using smartphone.

**Key Words** : Video extraction, Deep running, Image extraction, Unattended moving, Bluetooth

### 1. 서론

이미지 영상 제작을 위한 카메라의 대중화는 모바일 폰이나 디지털 카메라 등을 통해서 쉽게 나타나고 있다. 특히 카메라 응용의 분야로써 감시용 카메라 또는 강의용 카메라 분야는 더욱 폭넓게 사용되고 있다. 본 논문에서

서는 감시용 또는 강의용 카메라에 활용될 수 있는 이미지 영상 추적 무인 녹화 시스템을 설계하였다. 이미지 영상 추적 시스템은 영상처리 분야의 Face Detection, CAM-Shift 그리고 FFmpeg 기법 등 소프트웨어 집약적 기술로 복잡하지만 기존의 적외선 신호처리를 이용한 하

\*Corresponding Author : Byeong-tae Ahn (ahnbt@anyang.ac.kr)

Received February 14, 2019

Revised March 6, 2019

Accepted March 20, 2019

Published March 28, 2019

드웨어 집약적 무인 녹화 시스템에 비해 성능과 가성비가 매우 높다[1,2]. 그리고 유튜브 동영상 및 SNS를 이용한 1인 개인 인터넷 방송이 증가하면서 자가 셀프 동영상의 필요성이 대두되고 있다. 그러나, 셀카봉 같은 1인 개인 사진 촬영 기술은 많이 연구 개발 되어 대중화 되었지만 자가 셀프 동영상 촬영 기술은 아직까지 개발되어 있지 않다. 따라서 본 논문에서는 영상처리 기술을 이용하여 조작자 없이 촬영 대상을 인식하고 추적하기 위한 모바일용 무인 녹화 시스템을 제안한다. 시스템은 얼굴 검출 및 추적 시스템의 처리 프로세스를 이용하여 물체 자동 추적 시스템을 설계하였다. 물체 자동 추적을 위한 제어장치로는 아두이노를 이용하였으며 아두이노는 프레임 내에서 물체의 움직임을 추적할 뿐만 아니라 물체가 바깥으로 벗어나는 상황까지도 모터의 회전 장치를 이용해 실시간 추적이 가능하도록 하였다[3,4]. 본 논문의 2장에서는 관련 연구에 대해서 알아보고 3장에서는 모바일용 무인 녹화 시스템을 설계하였다. 끝으로 4장에서는 결론 및 향후과제를 제안한다.

## 2. 관련연구

### 2.1 무인 녹화 시스템

적외선 신호 처리 기술을 이용한 무인 강의 녹화 시스템은 'Swivi'이 있다.[1]



Fig. 1. Unattended Moving System 'Swivi'

그리고 공공시설, 군부대 경계, 기업보안용으로 잠재적 위협을 예측하는 자동 추적 기능이 포함된 무인 감시용 카메라 'PTZ Camera'가 있다[2,3]. 본 제품은 가격이 매우 비싸 비용 상승이라는 단점을 보완해주지 못하고 있다. 무인 감시용 카메라의 대표적인 예로는 Fig. 1이 있다.

일반적으로 움직이는 물체를 추적하기 위한 방법으로

는 3D 모델 기반, 영역 기반, 능동 윤곽선 기반, 특징 기반 방법이 활용되고 있다. '3D 모델 기반' 방법은 소수 차량에 대해 정확성이 높은 모델과 궤적을 복원하는 물체 중심의 방법이다.[5,6] '영역 기반' 방식은 연속 영상에 연결된 영역을 확보하고 상관관계 추정을 통해 움직이는 물체를 추적하는 것으로, 촬영시점의 배경을 추정해 입력되는 영상과의 차이에 따라 물체를 검출하게 된다.[7] 또한, 물체의 경계 윤곽선을 표현하고 이를 동적으로 갱신하면서 추적하는 '능동 윤곽선 기반' 방법이 있다.[8] 마지막으로, 물체 전체를 추적하는 것이 아닌 국부적인 특징을 추적하는 '특징 기반' 방식도 대표적인 물체 추적 방법으로 손꼽힌다.[9,10] 그러나 각 방법에는 단점도 존재하기 때문에 각각의 단점을 적절하게 보완할 필요가 있다. 최근 연구되는 물체 추적 방법은 하나의 특정방법에 의존하지 않고 이들 방식이 갖는 장점만을 선별해 적용하고 있다[11]. 즉 감시지역 내 움직임의 감지, 배경화면으로부터 움직임 영역의 분리 및 격리, 특징점 추출, 연속영상의 특징점 추적을 통한 움직임 방향 벡터 추적 등의 과정으로 구성되는 것이 일반적이다. 본 논문에서는, '영역 기반', '능동 윤곽선 기반', '특징 기반'의 3가지 기반 방식 각각의 장점만을 뽑아내어 본 시스템에 적용하였다. Face Detection 기법과 CAM-Shift 기법을 차례로 적용하여 추적 대상을 검출하고, 확정하여 추적하게 하도록 하여 단점은 최소화하고 장점은 최대화하였다.

### 2.2 Face Detection 관련 연구

Face Detection의 일반적인 접근방법은 4가지가 있다. 첫째, 평면의 단색 배경 이미지를 이용하거나 사전에 정의된 고정 배경 이미지를 이용하는 통계된 배경 이미지에서의 얼굴 검출 방법이다. 둘째, 얼굴색에 의한 검출 방법이다. 얼굴영상을 찾기 위해 전형적인 피부색을 이용하므로 실시간 구현이 가능하고 제한된 환경에서 가장 널리 사용된다. 셋째, 움직임에 의한 얼굴 검출 방법이다. 실시간 비디오를 이용하여 움직이는 얼굴 영역을 단순히 계산하는 방법이다. 그러나 배경에서 다른 대상물이 움직이는 경우 문제가 발생한다. 넷째, 제약 없는 장면에서의 얼굴 검출 방법이다. 흑백 이미지에서 인간처럼 얼굴을 정확히 탐지하는 알고리즘으로 통계적 클러스터 정보를 이용하여 신경회로망에 접근하는 방법 등이 있다.

본 논문에서는 가장 널리 사용되고 있는 얼굴 색에 의한 탐지 방법을 이용하였다. Face Detection 과정 이후

촬영 대상을 추적하기 위해 얼굴이 검출된 영역의 색 정보를 기반으로 CAM-Shift가 이용되고 현재 Face Detection 방법 중 가장 안정적인 검출율을 보이고 있다.

### 3. 설계

#### 3.1 시스템의 구성

본 시스템은 촬영 디바이스, 무선 이어셋, 거치대로 구성된다. Fig. 2는 시스템의 구성도를 나타낸 것이다. 카메라를 적재한 블루투스 연결이 가능한 촬영 디바이스(스마트폰)와 마이크 기능을 탑재한 무선 블루투스 이어셋 그리고 및 촬영 디바이스가 배치되어 자유롭게 회전기능을 가지는 거치대로 분류된다. 촬영 디바이스는 화면을 사용자에게 표시하는 디스플레이와 대상을 촬영할 수 있는 카메라 및 프로그램 명령을 수행하기 위한 프로세서 등으로 구성된다. 일반 안드로이드 스마트폰은 전면 카메라로써 디스플레이와 동일한 방향으로 배치되어 있어 스마트폰의 후면 카메라가 촬영 수단으로 이용된다. 거치대는 서보 모터 등이 몸체와 모터에 연결되어 회전이 가능하도록 받침대를 포함하고 있으며 받침대에는 촬영 디바이스가 배치될 수 있는 홈이 형성되어 있다. 무선 이어셋은 촬영 대상자에게 장착되어 촬영 대상자의 음성을 블루투스로 촬영 디바이스로 전달된다. 블루투스 통신은 무선 이어셋을 통한 음성과 영상을 데이터 채널이 아닌 음성 채널을 통해 음성과 화상의 시간 동기화가 이루어지도록 하였다.

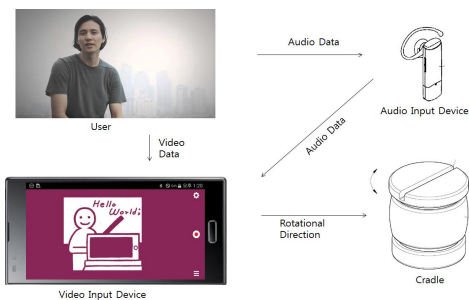


Fig. 2. Structure of System

Fig. 3은 이미지 영상 추적을 위한 프로세스를 도식화 하였다. 첫 번째, 스마트폰의 디스플레이 창에서 추적 대상의 스크린 상에서 좌표값을 입력받아 중앙처리장치에 전송한다. 중앙처리장치는 스크린 상의 추적 대상 좌표값과 스크린 화면 중앙 좌표값의 차이를 비교하여 회전유무를 결정하고 회

전시 연산처리 과정을 통해 좌.우측을 결정한다. (X, Y) 좌표의 차이 값이 미리 결정된 임계값보다 작으면 회전하지 않고 임계값보다 크면 좌우상하로 회전하도록 회전 방향을 결정한다. 두 번째, 안드로이드폰의 블루투스 모듈에서 거치대 내부 아두이노와 연결된 블루투스 모듈로 결정된 회전 방향을 전송한다. 세 번째, 블루투스 모듈은 아두이노로 결정된 회전 방향을 전송한다. 마지막으로 아두이노는 수신한 회전 방향에 근거하여 서보모터를 회전시키게 된다. 따라서 최종적으로 서보모터가 거치대의 본체 상부를 회전시킴으로써 거치대 본체 상부 홈에 꽂힌 스마트폰이 제어된 방향에 따라 회전하면서 촬영 대상의 실시간 추적이 가능하다.

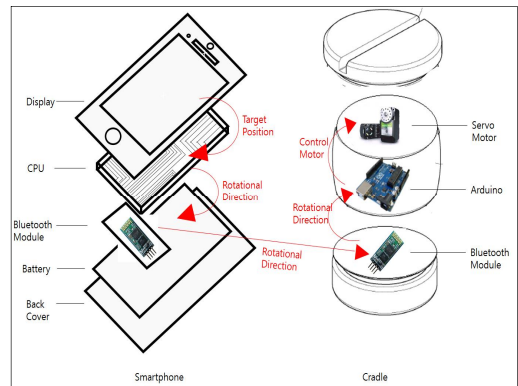


Fig. 3. Process for Image Video Extraction

본 논문의 영상 이미지 추적 무인 녹화 시스템은 Fig. 4에서 보는 바와 같이 총 7단계로 구성된다. 첫째, 디바이스에 연결된 카메라를 통해 영상에서 얼굴 형태를 인식하는 단계이다. 둘째, 얼굴 형태 인식에 기초하여 촬영 대상을 디바이스에 연결된 디스플레이에 표시하는 단계이다. 셋째, 표시된 부분이 촬영 대상임을 선택하는 선택 신호를 수신하는 단계이다. 넷째, 선택된 촬영 대상의 영상 이미지 색 정보를 취득하는 단계이다. 다섯째, 상기 취득된 색 정보를 가지는 영역 내의 일정 좌표를 촬영 대상의 타겟 좌표로 설정하는 단계이다. 여섯째, 타겟 좌표와 카메라로부터 입력되는 영상의 특정 영역과의 차이를 판단하는 단계이다. 일곱 번째, 상기 차이를 감소시키기 위한 회전 신호를 회전 장치에 전송하는 단계이다.

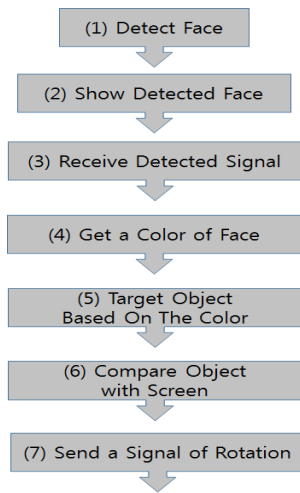


Fig. 4. 7 Step of Target Extraction Method

얼굴 검출 및 추적 과정 외에도 녹화 과정에서는 카메라를 통해 입력되는 영상을 녹화하고 촬영 대상에 장착된 오디오 송신 장치로부터 오디오 신호를 수신하여 오디오 신호와 녹화되는 영상을 시간 동기화하여 메모리에 저장한다.

3.2 얼굴 검출 및 영상 추적

얼굴 검출은 안드로이드 스마트폰 카메라를 통해 획득한 영상에서 사람의 얼굴이라고 추측할 수 있는 몇 가지 주요 특징점을 이용하여 사람의 얼굴을 검출하고 그 영역을 획득한다. 얼굴 검출 과정에 관해서는 RGB 영상 값이 아닌 Gray 영상 값을 이용하였다. 그러나 화면에 보여져야 할 것은 RGB 값이기에 영상을 입력받을 때에는 Gray 이미지와 RGBA 이미지 두 개 데이터 모듈을 가져올 수 있어야 한다. 따라서 본 논문에서는 OpenCV Library의 Face Detection 기법을 이용하여 Gray 이미지와 RGBA 이미지 처리를 가능하도록 하였다.



Fig. 5 Face-Detection

Fig. 5는 Face-Detection 기법을 이용하여 얼굴 검출이 가능하도록 한 것이다. 일반적으로 사람의 얼굴에는 두 눈은 명암이 어둡고 코는 명암이 밝다. Face-Detection 기법은 이런 명암을 이용해 패턴을 구하는 것이다[10]. 사람의 얼굴 위에 흑백의 사각형을 겹쳐 놓은 다음 밝은 영역에 속한 픽셀 값들의 평균에서 어두운 영역에 속한 픽셀 값들의 평균의 차이를 구한 후 그 차이가 문턱값(threshold)를 넘으면 사람 얼굴에 대한 특징이 있는 것이다. 사람의 얼굴은 다양하지만 생김새의 패턴은 비슷하므로 임의의 얼굴 위에서의 특정 위치, 특정 분포에 따른 명암의 차이는 거의 없기 때문이다.

```

    if(mCascade != null){
        int height = mGray.rows();
        int faceSize = Math.round(height *
        FdActivity.minFaceSize);
        List<Rect> faces = new
        LinkedList<Rect>();
        mCascade.detectMultiScale(mGray, faces, 1.1,
        2, 2,
        new Size(faceSize, faceSize));

        for(Rect r : faces)
            Core.rectangle(mRgba, r.tl(),
            r.br(),
            new Scalar(0, 255, 0, 255), 3);
    }
    
```

Fig. 6 Face Detection Algorithm

Fig. 6은 얼굴 검출 알고리즘을 나타낸 것이다. Gray Matrix의 열로부터 이미지의 높이를 구하고 전체 이미지 크기에 대한 얼굴 검출 범위를 지정한다. 얼굴들을 저장하기 위해 List를 선언하고 얼굴을 찾는다. 찾아진 얼굴의 사각 둘레(좌측 위, 우측 아래)를 지정하고 레벨 3 두께의 녹색 선으로 검출된 얼굴 영역의 테두리를 표시한다. 검출된 얼굴 영역 내부 색 정보를 획득하여 색 정보를 기반으로 관심 영역을 축소 혹은 확장해가며 탐색 윈도우 사이즈를 조절한다.

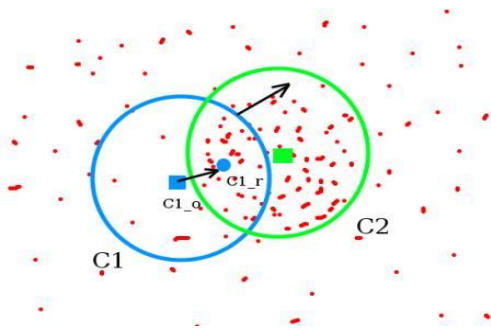


Fig. 7 CAM-Shift Method

Fig. 7은 영상 추적 중 CAM-Shift(Continuously Adaptive Mean Shift) 알고리즘 기법을 나타낸 것으로 기존의 Mean Shift 알고리즘을 개선한 것이다. Mean-Shift 알고리즘은 데이터 집합의 밀도분포(특징점, 코너, 색상)를 기반으로 관심영역(ROI) 객체를 고속으로 추적하는 알고리즘이다. 이 알고리즘은 초기의 검색 영역의 크기와 위치를 지정하면 반복되는 색 분할 계산에 의해서 색상 클러스터가 발생되고 초기 지정한 색 영역에 기반하여 경계를 결정하여 관심 물체를 추출하는 알고리즘이다.[12] CAM-Shift는 Color Segmentation 방법의 Mean Shift 알고리즘을 Streaming(효과, 추적) 환경에서 사용하기 위해 개선한 것으로 탐색 윈도우의 크기를 스스로 조정하는 기법을 사용하여 Mean Shift의 단점을 보강한 것이다. 초기의 윈도우는 파란색 원 C1이다. 이 윈도우의 원래 중심 좌표는 파란색 네모 박스 C1\_o로 표시되었다. 하지만 윈도우 영역 내에서 빨간 점들의 중심을 찾는다면 윈도우 내 빨간 점들의 실제 중심은 파란색 동그라미 C1\_r이 될 것이다. 확실히 C1의 좌표값에 근거한 절대중심값(C1\_o)과 빨간 점들의 실제 분포도에 따른 상대중심값(C1\_r)에는 차이가 있는 것을 알 수 있다. 그렇다면, 다시 윈도우 내의 절대 중심값을 빨간 점들의 실제 분포도에 따른 상대중심값에 일치하도록 이동해보자. 이 과정을 절대 중심값과 상대 중심값이 일치할 때까지(혹은 미리 결정된 임계값 범위 안에 들어올 때까지) 반복하면, 윈도우는 빨간 점들의 영역을 최대한 반영할 수 있고, 윈도우의 절대중심값도 실제 빨간 점들의 분포도에 가장 유사한 곳에 자리잡을 수 있다. 마침내 최대 픽셀 분포를 가진 초록색 동그라미 영역의 윈도우 C2를 얻을 수 있다. Fig. 7에서 볼 수 있듯이, 윈도우 C2는 빨간점을 가장 많이 포함하는 영역이다.

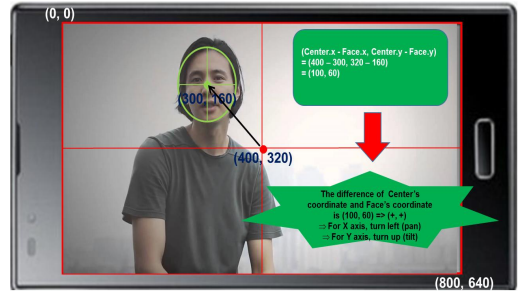


Fig. 8 Determine the presence and direction of rotation(1)

$$\text{diff}_x = \text{ScreenCenter}_x - \text{Target}_x$$

$$\text{diff}_y = \text{ScreenCenter}_y - \text{Target}_y$$

Turn left if  $\text{diff}_x > \text{threshold}_x$   
 Stop if  $-\text{threshold}_x < \text{diff}_x < \text{threshold}_x$   
 Turn right if  $\text{diff}_x < -\text{threshold}_x$

Tilt up if  $\text{diff}_y > \text{threshold}_y$   
 Stop if  $-\text{threshold}_y < \text{diff}_y < \text{threshold}_y$   
 Tilt down if  $\text{diff}_y < -\text{threshold}_y$

Fig. 9 Determine the presence and direction of rotation(2)

Fig. 8과 Fig. 9는 회전 유무 및 회전 방향 판단 예시와 알고리즘을 나타낸 것이다. 실시간으로 얻어지는 탐색 윈도우의 중심점 좌표와 카메라 프레임 중심 좌표를 비교하여 매순간 그 차이값이 임계값을 넘으면 좌우로 회전하도록 판단해야 한다. Fig. 8에서와 같이 카메라 프레임의 중심 좌표 (X1, Y1)가 (400, 320)이고, 탐색 윈도우의 중심 좌표 (X2, Y2)가 (300, 160)이라고 할 때, (X1-X2, Y1-Y2) = (400-300, 320-160) = (100, 160)으로 X, Y값 각각 양수(+)임을 알 수 있다. 이 경우에는 카메라 프레임 중앙에 해당 물체가 오게 하려면 카메라 본체의 좌회전과 상향회전이 필요하다. 그러나 차이값이 0일 때에만 회전을 멈추도록 설계한다면 대부분의 시간을 회전하는데 보내게 될 것이고 끊임없이 좌우로 회전하며 녹화가 진행되어 녹화의 질이 떨어지게 된다. 따라서 임계값을 정하여 X, Y값 각각의 차이값이 임계값을 넘지 않으면 회전을 멈추도록 한다. 예를 들면, 임계값을  $\pm 50$ 으로 정하고 카메라 프레임의 중심 좌표(X1, Y1) = (380, 200), 탐색 윈도우의 중심 좌표(X2, Y2) = (400, 320)가 일 때, 두 좌표의 차이값(X1-X2, Y1-Y2) = (20, -120) 이므로 X

축에 대해서는 회전하지 않고, Y축에 대해서만 상향 회전하게 된다.

#### 4. 결론

본 논문에서는 기존에 사용되고 있는 고비용 무인 감시 시스템 혹은 무인 촬영 시스템을 대신하여 접근성이 쉽고 저비용의 영상 이미지 추적 무인 녹화 시스템을 설계하였다. 본 시스템의 특성은 영상의 검출된 얼굴 영역 정보와 그 영역 내부 색 정보 및 이동 물체의 중심점을 이용하여 상시적으로 이동 물체 영역으로 축소 및 확장하며 탐색윈도우를 이동시키는 것이다. 그리고 물체의 움직임이 미리 결정한 임계값의 범위를 넘어서면 서보모터가 움직임의 방향을 따라 회전하며 물체를 추적한다. 본 연구에서 제안하는 무인 녹화 시스템은 기존에 상용화된 적외선 신호 등 영상 처리와는 전혀 무관한 기술을 이용하여 물체 추적 및 녹화 시스템을 설계한 것으로 기술 집약적이며 저비용으로 대량 생산이 가능하다. 향후 연구과제로는 본 연구에서 설계한 무인 녹화 시스템은 조명이 매우 어둡거나 매우 밝은 상태에서는 CAM-Shift을 이용한 물체 추적 정확도가 감소하였다. 따라서 이를 효과적으로 개선하기 위한 새로운 특징점 검출 연구가 필요하다.

#### REFERENCES

- [1] Y. J. He, M. Li, J. L. Zhang & J. P. Yao. (2015). Infrared target tracking via weighted correlation filter. *Infrared Physics & Technology*, 73, 103-114.
- [2] M. Davide, S. Raimondo, D. Gasparella Sturzenegger, J. Lygeros & M. Morari. (2010). *A tracking algorithm for PTZ cameras*, IFAC Proceedings Volumes, 43(19), 61-66.
- [3] P. Zhang, T. Zhuo, L. Xie & Y. Zhang. (2016). Deformable object tracking with spatiotemporal segmentation in big vision surveillance. *Neurocomputing*, 204, 87-96.
- [4] S. Yasukawa, H. Okuno, K. Ishii & T. Yagi. (2016). Real-time object tracking based on scale-invariant features employing bio-inspired hardware. *Neural Networks*, 81, 29-38.
- [5] G. Edelman & J. Bijhold. (2010). Tracking people and cars using 3D modeling and CCTV. *Forensic Science International*, 202(1-3), 26-35.
- [6] J. Liu, D. Liu, J. Dauwels & H. S. Seah. (2015). 3D Human motion tracking by exemplar-based conditional particle

filter. *Signal Processing*, 110, 164-177.

- [7] J. Wang, Y. Lu, L. Gu, C. Zhou & X. Chai. (2014). Moving object recognition under simulated prosthetic vision using background-subtraction-based image processing strategies. *Information Sciences*, 277, 512-524.
- [8] J. Yin, C. Fu & J. Hu. (2012). Using incremental subspace and contour template for object tracking. *Journal of Network and Computer Applications*, 35(6), 1740-1748.
- [9] F. Liu, C. Shen, I. Reid & A. van den Hengel. (2016). Online unsupervised feature learning for visual tracking. *Image and Vision Computing*, 51, 84-94.
- [10] X. Shen, X. Sui, K. Pan & Y. Tao. (2016). Adaptive pedestrian tracking via patch-based features and spatial-temporal similarity measurement. *Pattern Recognition*, 53, 163-173.
- [11] S. Chen & C. Liu. (2015). Eye detection using discriminatory Haar features and a new efficient SVM. *Image and Vision Computing*, 33, 68-77.
- [12] H. Zheng, X. Mao, L. Chen & X. Liang. (2015). Adaptive edge-based mean shift for drastic change gray target tracking. *Optik - International Journal for Light and Electron Optics*, 126(23), 3859-3867.
- [13] M. Gentili, R. Sannino & M. Petracca. (2016). BlueVoice: Voice communications over Bluetooth Low Energy in the Internet of Things scenario. *Computer Communications*, 89, 51-59.

안 병 태 (Byeongtae Ahn)

[중신회원]



- 1999년 2월 : 국민대학교 컴퓨터과 학부(이학사)
- 2006년 8월 : 경상대학교 컴퓨터과 학부(공학박사)
- 2012년 3월 ~ 현재 : 안양대학교 교양대학 컴퓨터전공 교수

- 관심분야 : 딥러닝, 영상처리, 멀티미디어 시스템, 블록체인
- E-mail : ahnbt@anyang.ac.kr