

흡연자 검출을 위한 새로운 방법

이 종 석*, 이 현 재*, 이 동 규**, 오 승 준^o

New Scheme for Smoker Detection

Jong-seok Lee*, Hyun-jae Lee*, Dong-kyu Lee**, Seung-jun Oh^o

요 약

본 논문은 흡연으로 인한 화재사고 방지를 위해, 비디오 영상에서 흡연자를 검출하는 알고리즘을 제안한다. 흡연자의 행동을 인식하기 위해 행동 인식 기법의 계층적 방법 중 서술 기반 접근 방법을 기반으로 제안하는 알고리즘은 배경 영역 분리, 객체 검출, 이벤트 탐지, 이벤트 판단 과정으로 구성된다. 배경 영역 분리 과정으로 학습률이 다른 두 개의 가우시안 혼합 모델을 이용하여 입력 영상으로부터 고속 움직임 전경, 저속 움직임 전경 영상을 생성하고, 저속움직임 전경 영상을 chain-rule 기반 외곽선 검출 알고리즘을 통하여 객체의 위치를 추출해낸다. 위치 정보를 기반으로 흡연자의 세 가지 특징인 얼굴, 연기, 손의 움직임을 이벤트 탐지 과정에서 검출한다. Haar-like feature를 이용하여 얼굴을 검출하며, 고속 움직임 전경에서 연기의 발생 빈도수와 방향성을 반영하여 연기를 검출한다. 움직임 추정을 통해 반복적인 손의 움직임을 검출한다. 일정 구간의 비디오 시퀀스 내 객체들에 대하여, 검출된 특징들의 서술적 관계를 반영하여 각각의 객체가 흡연자인지 판단한다. 제안하는 방법은 실시간으로 여러 다른 객체들 사이에서 강인하게 흡연자를 검출한다.

Key Words : Smoker detection, Smoke detection, Object detection

ABSTRACT

In this paper, we propose a smoker recognition algorithm, detecting smokers in a video sequence in order to prevent fire accidents. We use description-based method in hierarchical approaches to recognize smoker's activity, the algorithm consists of background subtraction, object detection, event search, event judgement. Background subtraction generates slow-motion and fast-motion foreground image from input image using Gaussian mixture model with two different learning-rate. Then, it extracts object locations in the slow-motion image using chain-rule based contour detection. For each object, face is detected by using Haar-like feature and smoke is detected by reflecting frequency and direction of smoke in fast-motion foreground. Hand movements are detected by motion estimation. The algorithm examines the features in a certain interval and infers that whether the object is a smoker. It robustly can detect a smoker among different objects while achieving real-time performance.

I. 서 론

최근 국민 안전처 국가 화재정보센터 자료에 따르

면 2015년 10월 한 달 동안 발생한 화재 중 담배꽂초로 인한 화재 건수는 400건으로 약 13%를 차지하고 있다. 담배로 인한 화재는 장소를 가리지 않고 발생하

* 이 논문은 2015년도 광운대학교 교내학술연구비 지원에 의해 연구되었음.

• First Author : Kwangwoon University Department of Electronics, suk2080@media.kw.ac.kr, 학생회원

^o Corresponding Author : Kwangwoon University Department of Electronics, sjoh@kw.ac.kr, 종신회원

* Kwangwoon University Department of Electronics, hickstar@media.kw.ac.kr, 학생회원

** Kwangwoon University Department of Electronics, dongkyu@media.kw.ac.kr, 학생회원

논문번호 : KICS2016-05-095, Received May 14, 2016; Revised September 5, 2016; Accepted September 21, 2016

고 있어 큰 화재로 번질 위험이 있다.

컴퓨터비전(Computer Vision) 분야에서는 이미 지능형 영상 분석 기술의 이벤트 탐지를 활용하여 금연 구역을 감시하고 감시구역에서 발생하는 불꽃과 연기를 검출하여 화재 발생을 알려 빠른 대처를 할 수 있도록 기여하고 있다. 지능형 영상분석 기술이란 영상의 정보를 분석하여 자동으로 이상행위를 탐지하고 관리자에게 정보를 전송하는 기술이다^[1]. 그러나 불꽃과 연기는 이미 화재가 발생한 이후에 검출 될 수 있는 요인들로 화재를 미연에 방지하기에는 여전히 한계가 있기에 화재 발생 이전 시점인 흡연 행위를 포착할 필요가 있다.

흡연 행위는 이미 많은 사람들이 잘 알고 있는 행위이다. 하지만 흡연 행위를 영상으로 검출하기란 쉽지 않다. 그 이유는 연기는 주변 상황에 따라 영상 내에서 어디로 퍼져나갈지 모르고, 담배를 피우면서 흡연자가 계속 움직이고 또한 개인적인 습관도 많이 나타나기 때문이다.

위와 같은 흡연자 검출에 대한 어려움으로 흡연행위 검출을 위해 먼저 얼굴이나 담배 검출 방법을 사용하였다^[2]. 먼저 영상에서 얼굴을 찾고 얼굴에서 입의 위치를 추정 하여 담배가 나타날 영역을 지정하고 그 영역에서 담배의 색을 이용하여 담배를 검출하는 방법이다. 하지만 이 방법은 담배 이외의 다른 물체가 담배의 색과 비슷할 경우 강건한 검출이 어렵다. 담배 이외에 흡연자의 큰 특징으로 연기가 있다. 영상 내에서 흡연으로 인한 연기가 발생하면 전체적인 화면의 밝기가 밝아지는 현상을 이용하여 밝기 변화를 검출하거나^[3] 영상의 장면 전환 검출법을 이용하여 흡연행위를 검출하는 방법이다^[4]. 이 방법의 경우 영상에서 흡연자 또는 흡연 행위가 차지하는 범위가 넓어야 하기 때문에 좁은 영역에서만 흡연자 검출이 가능하다는 문제점이 있다. 또 다른 연구로는 연기를 검출하고 담배, 얼굴, 손을 검출하여 담배를 피기위해 손이 위로 올라가는 행동을 검출한다. 마지막으로 연기, 몸, 손을 올리는 행동 세 가지를 종합하여 도식해법을 이용하여 각 프레임(Frame)마다 흡연 행위의 확률을 구하여 일정 확률 이상이 되면 흡연자로 검출한다^[5].

최근 카메라의 고해상도 지원으로 인하여 CCTV는 더 넓은 범위를 감시할 수 있게 되었다. 그에 따라 영상 내의 물체와 사람의 크기는 점점 줄어들 사람의 얼굴보다 더 작은 담배는 더 이상 강건한 흡연자의 특징점이 아니다. 따라서 흡연 행위는 그림 1과 같이 흡연자의 세 가지 특징 중 연기를 내뿜는 행동과 팔을 반복적으로 움직이는 행동, 두 부분이 행동과 관련된 특징



그림 1. 흡연자가 갖는 세 가지 특징
Fig. 1. Properties of the smokers

이므로 행동 인식의 계층적 기반 서술적 방법으로 접근할 수 있다. 본 논문은 흡연 행위가 계층적 기반 방법 중 서술 기반 접근을 이용하여 인지될 수 있음을 보인다.

제안하는 흡연 행위 인지 과정은 배경 영역 분리, 객체 추출, 이벤트 탐지, 이벤트 판단으로 구성되어 있고 그림 2는 흡연자 인지 과정을 대략적으로 나타낸 것이다. 화면상의 모든 구역은 금연 구역으로 설정하며, 사람이 등장할 경우 인지하여 흡연자의 특징들을 검출하고 종합하여 흡연자로 판단한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 이론들을 설명하고, 3장에서는 흡연자의 특성을 고려한 알고리즘을 제안한다. 4장에서는 제안한 알고리즘을 이용한 흡연자 인지 실험을 통해 성능을 평가하고, 5장에서 결론을 맺는다.

II. 관련 연구

2장에서는 행동인식 기법들을 소개 하고 본 논문에서 제안하는 흡연자 검출 방법에 사용되는 객체 검출, 연기 검출, 손 움직임 검출 관련 연구를 소개한다.

2.1 행동 인식

사람의 행동인식에 대한 접근 방법은 단층 기반의 방법과 계층적 기반의 방법으로 구분할 수 있다^[6]. 단층 기반 방법은 사람의 행동을 이미지 시퀀스를 기반으로 직접 인지하기 때문에 연속적인 특징을 갖는 제스처(Gesture)나 액션(Action) 등의 행동인식에 적합하다. 반면에 계층적 기반 방법은 비교적 단순한 행동들의 조합으로 높은 수준의 행동을 인식한다. 단층 기반의 접근 방법은 시공간 접근법, 연속적 접근법으로 나뉜다. 시공간 접근법은 입력 영상을 3차원의 시공간 볼륨(Volume)으로 보는 것이고, 연속적 접근법은 입력 영상을 일련의 관측들로 해석한다. 계층적 기반 방법은 세 가지로 분류될 수 있는데, 통계적 접근과, 구

문론적 접근, 서술 기반 접근이 있다. 통계적 접근은 계층적으로 연결된 통계적인 상태 기반 모델을 구성한다. 구문론적 접근은 행동을 작은 행동들의 끈 형태로 모델링한다. 서술 기반 접근은 단순한 행동들의 시간적, 공간적, 논리적 관계를 서술함으로써 높은 수준의 행동을 나타내는 방법이다.

2.2 객체 검출

객체 검출을 하기 위해 가장 먼저 진행해야 하는 과정은 배경영역 분리이다. 이전 이미지들을 이용하여 배경영역 모델을 만들고 원본 이미지에서 제거하여 객체 영역과 배경 영역을 분리 하는 것이 일반적이다. 이때, 흔히 사용하는 모델이 가우시안 혼합 모델(GMM : Gaussian Mixture Model)^[7]이다.

GMM은 실시간으로 현재 이미지로부터 일정 비율로 배경 모델을 갱신한다. 이 때, 갱신 하는 비율을 학습률(Learning rate)이라고 한다^[8]. 학습률을 높이면 배경모델의 갱신 속도가 빨라지고 학습률을 낮추면 갱신 속도가 느려진다. 학습률이 높을 때, 움직이던 객체가 영상 내에서 멈추게 되면 빠르게 배경 모델로 갱신되어 배경 모델에 포함되어 객체로 추출되지 않는다. 이와 반대로 학습률을 너무 낮추면 정적 객체, 동적 객체, 배경 노이즈 등 많은 영역이 검출되게 한다. Xuegang는 서로 다른 3개의 학습률을 사용하는 GMM을 이용하여 배경, 움직이는 물체, 노이즈로 구분하여 더 정확하게 객체를 추출하는 방법을 제시하고 있다^[9].

2.3 연기 검출

연기 검출은 화재 탐지에 가장 큰 부분이기 때문에 컴퓨터 비전(Computer Vision)분야에서는 비디오 기반 연기 검출(Video based smoke detection, 이하 VSD)에 대하여 많은 연구를 해왔다.

VSD는 4개의 과정으로 구성된다^[10]. 영상 내에서 연기 후보영역을 분할하고, 영역 내에 특징점들을 추출하여 연기 영역을 검출한 후, 검출한 연기를 확인한다. VSD에 사용하는 특징들은 크게 4개로 분류 할 수 있다. 연기는 흰곳에 머물지 않고 계속 이동하는 특징을 지니고 있어 연기의 움직임을 특징으로 사용한다. 다음은 색을 이용한 방법으로 색성분이 줄어드는 영역을 연기라고 생각한다. 연기는 흰색 또는 회색으로 보일 경우가 많기 때문에 원래의 색성분이 모두 감소하는 경향을 보인다. 세 번째 방법은 연기의 크기와 경계선이 불규칙적인 특징을 이용한다. 마지막으로 연기 영역의 에너지를 이용하는 방법이 있다. 이 방법은

일단 움직이는 영역을 모두 검출한 후에 연기와 움직이는 객체를 구별하게 되는데 움직이는 물체의 경우 픽셀(Pixel) 값의 변화가 심하기 때문에 에너지가 크게 나타난다. 그러나 연기의 경우 거의 투명한 특징을 갖기 때문에 픽셀(Pixel) 값의 변화가 거의 없어서 에너지의 변화도 크지 않다.

VSD는 픽셀(Pixel) 단위의 처리 보다는 영상을 작게 나눈 블록(Block) 단위 처리를 선호 하며 위에서 말한 특징 점들도 블록(Block) 단위로 추출하여 분류한다. 블록(Block) 단위 처리는 픽셀단위 처리보다 공간을 압축 할 수 있기 때문에 분류 과정이 들어가는 알고리즘이라면 일반적으로 사용하는 기법이다.

2.4 손의 움직임 검출

객체 검출 중에서도 신체의 특정 부분만을 검출하는 것은 어렵다. 얼굴의 경우 모양이 크게 변하지 않아 특징점을 찾아내기 유리하지만, 손은 모양이 자주 변하고 시점에 따라 보이는 모습도 다르기 때문에 컴퓨터 비전 분야에서 검출하기 까다로운 객체로 여겨지고 있다.

기존에 수행된 연구들 중에서는 얼굴이나 손과 같은 신체 특정 부분을 검출하기 위한 기본적인 특징으로 모양과 피부색이 고려되었다^[11-13]. 특히, 피부색은 색상 공간에서 고유한 영역을 가지므로 신체라고 판단될 수 있는 후보 객체들을 걸러내기 위해 알고리즘의 도입부에 주로 사용되고 있다. 피부색 검출을 통해 손을 검출하였다면, 의미 있는 손의 동작이나 움직임을 알아내기 위해 손의 모양 정보가 이용될 수 있다. 손의 모양 모델은 다양한 상황에서 손으로 나타낼 수 있는 모양들로 구성되어 있으며, 특수한 의미를 가지는 손 모양을 검출하는 데 활용도가 높다. [12]에서는 수화 인식을 위해 손동작을 인지하는 알고리즘을 제안하였다. 또한, 조명이나 카메라상의 문제로 피부색을 검출하기 어려운 경우, 색상 대비와 모양이 동시에 고려될 수 있다.

손의 움직임 검출 과정은 크게 손 검출, 손 추적 과정으로 나눌 수 있다. 위에서 언급한 대로, 손을 검출하기 위해 고려해야할 것은 손의 고유한 모양과 피부색이다^[14]. 손 모양 모델과 피부색을 기반으로 특징 점들을 추출하여 주변 객체들로부터 분리한다. 손 추적 과정은 이전 프레임에서 검출한 손이 현재 프레임에서 어디로 움직였는지 추론하는 과정으로, 이전 측정값으로부터 현재 상태를 예측을 하기 위해 칼만 필터(Kalman filter)^[15], 입자 필터(Particle filter)^[16], 평균 이동이 내장된 입자 필터(Mean-shift embedded

particle filter, MSEPF)^[17], 베이지안 필터(Bayesian filter)^[18] 등이 사용될 수 있다. 이들은 상태 추정을 위해 필요한 파라미터들이 매 프레임마다 갱신되는 구조이며, 사용한 필터에 따라 옵티컬 플로우(Optical flow)나 확률적 추론 등 다양한 방식으로 접근한다.

제안하는 알고리즘은 위에서 소개한 기법들을 발전 및 조합하여 흡연자를 검출한다. 구체적으로, VSD를 이용하여 연기를 내뿜는 행위와 움직임 추적을 이용하여 흡연 중 손을 움직이는 행위를 각각 검출하고 두 행위 사이의 관계를 공간적, 시간적으로 서술하여 흡연자를 검출한다.

III. 제안하는 흡연자 검출 알고리즘

그림 3은 제안하는 흡연자 검출 알고리즘의 순서도를 나타낸다. 입력 영상에서 배경 모델을 제거하여 객체 영역을 추출하고 영역 내에서 여러 객체들을 각각 추출하고 각각의 객체들마다 이벤트 탐지과정에서 얼굴 위치, 연기 및 움직임을 검출하고 마지막 이벤트 판단과정에서 입력 영상 내 흡연자를 검출한다.

3.1 배경 영역 분리

본 논문은 일반적인 흡연자의 경우 제자리에 서서 흡연 한다고 가정 한다. 따라서 흡연자는 저속 움직임을 보이는 객체일 것이고 연기의 경우 한곳에 머무르지 않고 계속 움직이는 특징이 있기 때문에 고속 움직임을 보이는 객체일 것이다. 그러므로 입력 영상으로부터 배경을 제거한 전경을 저속 움직임 객체와 고속 움직임 객체로 분류할 필요가 있다. 배경 모델로 2.2 절에서 소개한 서로 다른 학습률을 갖는 두 개의 GMM을 사용한다. 학습률은 0에서 1사이 값을 가지며, 일반적으로 움직이는 객체만 검출할 경우 0.01의 학습률을 사용하고 0.001보다 낮은 학습률을 사용할 경우 모든 전경을 검출할 수 있다. 제안하는 알고리즘은 0.005, 0.00001의 학습률을 갖는 GMM을 사용하여 각각 고속 움직임 객체 영상과 전경 영상을 얻는다. 전경 영상에서 고속 움직임 객체를 제거하면 저속 움직임 객체 영상을 얻을 수 있다. 그림 4는 배경 영역 분리 과정을 나타낸다. 그림 4 (a)는 입력 영상이고 (b)와 (c)는 입력 영상에서 배경 모델을 각각 제거한 것이다. (b)는 학습률이 낮은 경우로 서있는 사람, 왼쪽으로 움직이는 자동차 모두 포함하는 전경 영상이며, 반대로 학습률이 높은 (c)는 왼쪽으로 움직이는 자동차만 포함하고 있다. (d)는 (b)와 (c)의 차 영상으로 서있는 사람을 포함한 저속 움직임 영상이다. 마지막

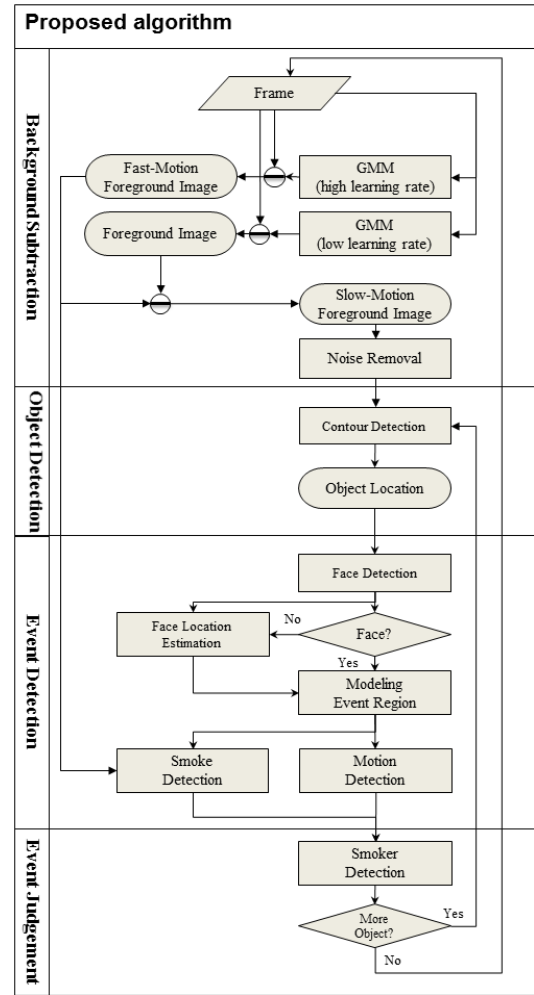


그림 2. 제안하는 흡연자 검출 알고리즘의 순서도
Fig. 2. Flowchart of the proposed smoker detection algorithm

으로 (d)의 이미지로부터 모폴로지(Morphology)연산의 팽창, 침식과정을 통해 얻은 노이즈가 제거된 영상을 객체검출에 이용한다. (c)는 고속 움직임 전경 영상으로 이벤트 탐지 연기 검출 과정에서 사용한다.

3.2 객체 검출

본 논문에서는 저속 움직임 전경 영상에 흡연자가 존재한다고 가정한다. 전경 영상으로부터 객체를 검출하여 각각의 객체에 라벨을 부여하고 다음 과정인 이벤트 탐지에서 ROI (Region of Interest) 설정에 필요한 객체의 위치와 너비, 높이를 추출한다. 제안하는 방법은 2x2 마스크(Mask)를 이용한 chain-rule 기반 외곽선 검출 방법을 이용한다. 라벨링(Labelling)과 동시에 객체의 위치와 넓이, 높이를 얻을 수 있다. 먼저

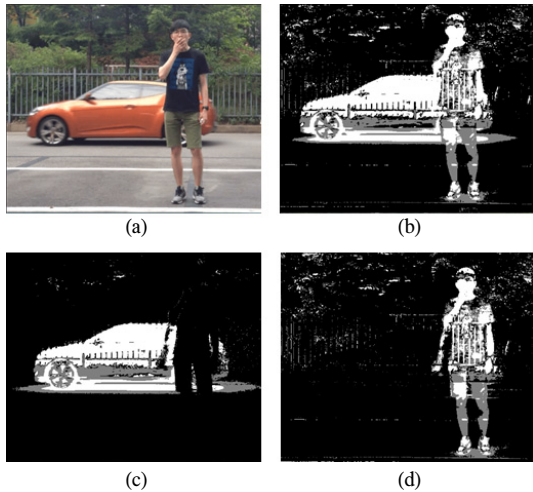


그림 3. 배경 영역 분리. (a)원본 이미지. (b)전경. (c)고속 움직임 전경. (d)저속 움직임 전경
 Fig. 3. Background subtraction. (a) Original image. (b)Foreground. (c)Fast-motion foreground. (d)Slow-motion foreground.

3.1을 통해 얻은 저속 움직임 전경 영역을 2진 영상으로 변환한다. 그 다음 그림 5 (a)의 첫 번째 2x2 마스크를 객체의 시작패턴으로 가정하고 래스터 스캔(Raster scan) 방향으로 이동하면서 객체 외곽선의 첫 시작패턴을 찾는다. 그림 5 (a)는 객체의 경계선에서 검출할 수 있는 모든 패턴을 나타내고 있다. 4비트를 이용하여 모든 패턴을 표현할 수 있으며, 그림 5 (a)의 마지막 두 개의 패턴은 경계가 아닌 배경과 객체에 해당한다. 그림 5 (b)는 외곽선 검출의 예시를 보여주고 있다. 각각의 화살표는 마스크(Mask)가 다음 외곽선 검출을 위하여 이동할 방향을 나타내고 있다. 그림 5 (b)에서 사용하지 않은 6, 9번의 경우는 두 개의 객체가 대각선으로 근접했을 경우를 나타내고 있다. 이 두 개의 경우는 마스크의 이전 이동 방향을 확인하고 방향을 정한다. 6번의 경우 이전 이동 방향이 아래쪽이었다면, 오른쪽 대각선 아래로 이동하고 위쪽이었다면, 왼쪽 대각선으로 위로 이동한다. 9번 마스크의 경우 이전 이동 방향이 오른쪽이었다면 오른쪽 대각선 위로 이동하고, 왼쪽이었다면 왼쪽 대각선 아래로 이동한다. 연결된 점들은 같은 라벨을 부여 받아 하나의 객체로 인지된다. 마스크가 시계 반대 방향으로 이동하면서 객체의 외곽을 검출하고 마지막으로 원래 시작 위치로 돌아오면 하나의 객체 검출이 끝난다. 또 다른 객체를 검출하기 위해 시작 패턴을 검색하여 위 과정을 동일하게 수행한다. 마지막으로 검출된 모든 객체의 넓이, 높이, 위치를 각각 얻는다.

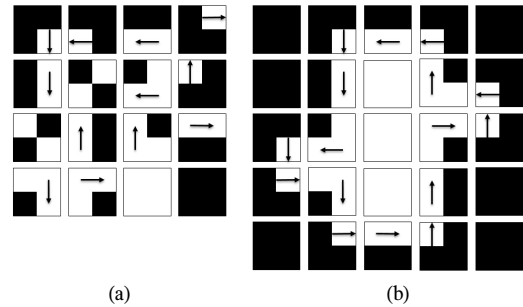


그림 4. Chain-rule기반 외곽선 검출. (a)2x2 마스크(mask). (b)외곽선 검출 예시
 Fig. 4. Chain-rule-based contour detection. (a) 2x2 mask. (b)Example of contour detection

저속 움직임 전경 영상으로부터 검출된 여러 객체 중에서 흡연자인 사람을 검출 할 필요가 있다. 일반적으로 사용하는 HOG(Histogram of Oriented Gradients)와 Haar 같은 특징점을 이용하여 사람을 검출할 수 있지만 복잡도가 높다는 단점이 있다. 본 논문은 복잡도를 줄이기 위해 간단한 가정을 통하여 검출된 객체들 중 사람을 추정하는 방법을 제안한다. 아래 두 조건을 만족하면 검출된 객체를 사람으로 추정한다.

- (1) 높이가 넓이보다 큰 객체.
 - (2) 높이와 넓이가 최소 얼굴크기 보다 큰 객체.
- 조건 (1)은 사람이 서있을 경우 당연한 조건이다. 조건 (2)의 경우 미리 정해진 최소 얼굴 크기보다 높이나 넓이가 작다면, 노이즈일 확률이 높기 때문이다. 미리 얼굴의 크기의 최솟값은 실험적으로 30x30으로 정해두었다.

3.3 이벤트 탐지

이벤트 탐지 과정은 흡연자의 특징인 얼굴, 연기, 움직임을 검출하기 위한 과정으로 각각의 검출된 사람 객체에 대하여 독립적으로 진행된다. 먼저 3.2.절의 과정으로 얻은 객체의 위치와 크기 정보를 기반으로 얼굴영역을 검출하며 다시 얼굴 영역 기반으로 연기와 움직임 검출을 위한 ROI를 설정하고 ROI 내에서 각각의 특징을 검출한다.

3.3.1. 얼굴 위치 검출

연기는 사람의 얼굴로부터 내뿜어지고 손의 움직임도 얼굴 쪽으로 움직이기 때문에 연기와 손의 움직임을 검출하는데 얼굴의 위치는 그만큼 중요하다. 그림 6은 얼굴 검출을 위한 ROI와 얼굴 위치 추정을 나타내고 있다. 3.2에서 검출한 객체가 사람이라면 객체의

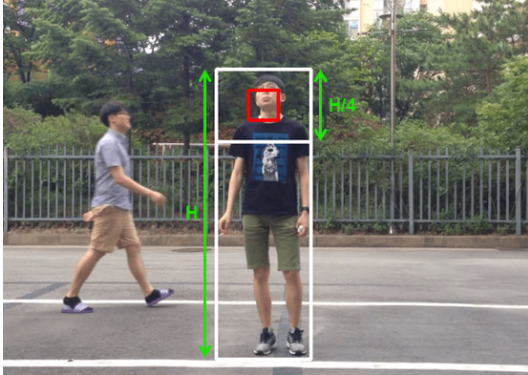


그림 5. 얼굴 영역 ROI와 얼굴 위치 추정.
Fig. 5. Face ROI and estimation of face location

상단부에 얼굴이 있는 것은 당연하다. 따라서 객체 영역의 상단으로 1/4만큼을 ROI로 정하고 Haar feature^[19] 기반 얼굴 검출 알고리즘을 사용한다. 만약 얼굴 검출에 실패 한다면 ROI의 중앙에 이전 프레임에서 찾은 얼굴의 크기를 이용하여 얼굴이라고 추정한다. 본 논문에서 제안하는 방법은 얼굴 오검출의 경우에도 얼굴의 위치를 추정함으로써 모든 프레임에서 흡연자 검출이 가능하다.

3.3.2 후보 영역 모델링

연기와 움직임 영역의 후보 영역은 얼굴 영역으로부터 만들어진다. 그림 7은 검출된 얼굴을 기반으로 각각의 후보 영역을 보여주고 있다. 얼굴 영역은 너비 'L', 높이 'L'를 갖는 적색 정사각형 박스이다. 얼굴을 기준으로 위쪽은 연기 영역이며 인쪽 영역과 바깥 영역으로 나눌 수 있다. 인쪽 영역의 너비는 얼굴영역의 3배고 높이는 2배다. 바깥 영역의 너비는 얼굴 영역의

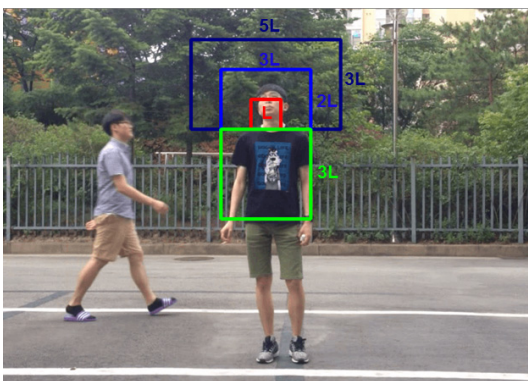


그림 6. 연기, 움직임 영역 모델링. 적색: 얼굴영역, 녹색: 움직임영역, 청색: 연기 영역.
Fig. 6. Modeling smoke and motion region. Red: face region, Green: motion region, Blue: Smoke region

5배이며 높이는 3배이다. 두 개의 영역으로 나눈 이유는 3.3 연기 검출 과정에서 설명한다. 얼굴을 기준으로 아래쪽은 움직임 영역으로 흡연자 손의 움직임을 검출하는 영역으로 너비와 높이 모두 얼굴영역의 3배이다.

연기를 내뿜는 행위와 손을 움직이는 행위는 시간적으로 거리가 있더라도 공간적으로 얼굴이라는 연결고리를 통해 논리적 관계가 형성된다. 따라서, 제안하는 방법은 행동인식의 서술적 접근 방법이라 할 수 있다.

3.3.3 연기 검출

연기는 2.2에서 소개한 많은 특징들을 지니고 있다. 그 중 제안하는 방법은 그림 8과 같이 공간적 특징으로 고속 움직임 전경 영상에서 연기 픽셀(Pixel)의 발생 빈도수를 고려하고 시간적 특징으로 흡연자의 연기는 항상 흡연자 얼굴로부터 밖으로 퍼져나가는 방향성을 고려한다.

먼저 ROI를 블록 단위로 나누고, 각 블록을 서브블록으로 나눠 서브블록의 히스토그램을 기반으로 연기를 검출한다. 그림 8 (a)는 ROI내 발생 할 수 있는 블록들의 예시이다. 첫 블록은 움직임 영역이 거의 없는 배경 블록이라고 할 수 있다. 세 번째 블록의 경우는 ROI 내에 있는 다른 움직이는 객체일 가능성이 높다. 일반적으로 연기는 입자형태로 되어있어 배경을 완전히 덮진 않는다. 따라서 두 번째 블록과 같이 발생 빈도수가 중간인 블록을 연기블록이라고 추정한다. 현재 프레임의 연기 검출은 수식(1)을 사용한다.

$$\Psi(s_t) = \Psi_f(s_t) \times \Psi_t(s_t) \quad (1)$$

s_t 는 t번째 프레임의 연기 후보영역 안에 있는 블록 r 의 집합이다, $s_t = \{r_1, r_2, \dots, r_n\}$. Ψ_f 는 s_t 중에 흡연 블록의 비율이 T_4 이상인지 판단하는 함수로 수식(2)와 같이 정의 된다.

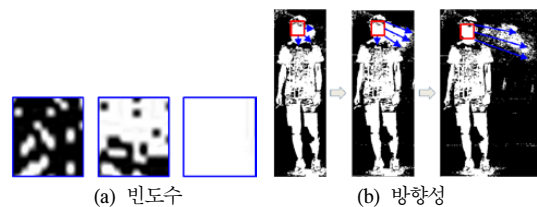


그림 7. 담배 연기의 공간적, 시간적 특징
Fig. 7. Spatial and temporal property of smoke

$$\Psi_f(s_t) = \begin{cases} 1 & \text{if } \frac{1}{m_4} \sum_{r \subset s_t} R(r) > T_4 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

수식(2)의 m_4 은 정규화 상수이고 함수 R 은 블록 r 중 흡연블록의 비율이 T_3 이상인지 판단한다. 함수 R 은 수식(3)으로 정의 된다.

$$R(r) = \begin{cases} 1 & \text{if } \frac{1}{m_2} \sum_{b \subset r} B(b) > T_3 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

블록 r 은 서브블록 b 의 집합이며, 함수 B 는 수식(4)와 같이 정의된다. 함수 B 는 서브블록의 히스토그램을 구하고 m_1 , 서브블록의 픽셀 개수로 정규화시켜 비율이 T_1 과 T_2 사이인지 판단한다. 여기서 i 은 픽셀 값을 나타낸다.

$$B(b) = \begin{cases} 1 & \text{if } T_2 > \frac{1}{m_1} \sum_{i \in b} \delta(i=1) > T_1 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

그림 9는 연기 블록 검출 과정을 보이고 있다. 설정한 ROI를 얼굴 블록 단위로 나눈다. 그 후 각 블록은 다시 9개의 작은 블록으로 나뉘게 되고, 작은 블록의 히스토그램 기반으로 연기블록을 추정하고 큰 블록 단위로 작은 블록의 수를 세어 큰 연기블록으로 추정한다. 마지막으로 큰 연기블록의 수가 미리 정해놓은 문턱값 보다 크다면 ROI 내에 연기가 존재한다고 추정한다. 각 크기 별로 연기 블록을 결정하는 문턱값은 실험적으로 정해놓았으며 블록 크기마다 문턱값이 다르기 때문에 퍼센트로 나타내면, 서브 블록은 픽셀들이 서브 블록의 40 - 70%(T_1 - T_2)를 차지 할 경우 서브 연기 블록으로 추정한다. 블록의 경우는 서브 연기

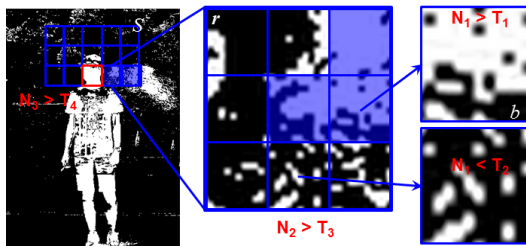


그림 8. 담배 연기의 공간적 특징
Fig. 8. Spatial property of smoke

블록의 수가 30%(T_3) 이상 차지해야 하고 마지막으로 연기블록이 ROI의 10%(T_4) 이상 차지하고 있어야 최종적으로 ROI내에 연기가 있다고 추정한다.

연기의 방향성은 그림 8 (b)와 같이 흡연자가 내뿜는 담배 연기는 항상 얼굴로부터 밖으로 퍼져나가는 성질을 지니고 있다. 영상 내 객체의 이동 방향을 검출 하는 방법에는 옵티컬 플로우(Optical flow)나 움직임 추정이 있다. 그러나 연기의 경우 투명한 성질을 지니기 때문에 모서리가 없고 픽셀 값의 차이가 크지 않기 때문에 정확한 방향이 검출되지 않는다. 연기의 방향성 검출은 연기가 어느 방향으로 움직였는지가 중요한 것이 아니라 얼굴 반대 방향인지가 중요하다. 수식(1)의 Ψ_t 는 연기의 방향성을 고려하는 함수로 수식(5)와 같이 정의되며 함수 D 는 연기 영역의 바깥쪽 연기 블록과 안쪽 연기 블록 수의 차이를 구하는 함수다. 이 함수를 시간으로 미분함으로써 바깥쪽 연기 블록의 점진적으로 증가하는 현상을 검출하고 함수 D 는 수식 (6)으로 정의된다.

$$\Psi_t(s_t) = \begin{cases} 1 & \text{if } \frac{dD(s_t)}{dt} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$D(s) = d_O(s) - d_I(s) \quad (6)$$

d_O 와 d_I 는 검출된 연기 블록 r 중 바깥 블록의 수와 안쪽 블록의 수를 각각 구하는 함수이다. 그림 10과 같은 방법을 이용하여 연기의 방향성을 검출한다. t 번째 프레임은 연기가 얼굴로부터 뿜어져 나온 장면이다. 이 때, 연기 블록의 수는 10%미만으로 연기가 검출되지 않는다. $t+1$ 번째 프레임에서 흡연자는 연기를 더 내뿜어 큰 연기 블록이 2개가 되어 10%를 넘는다. 전 프레임에서는 안쪽 영역에서 1개였지만 현재는 안쪽 1개 바깥쪽 1개가 된다. $t+2$ 프레임은 연기가 더

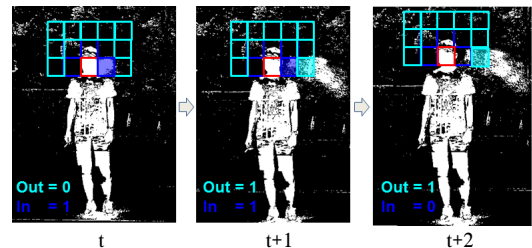


그림 9. 담배 연기의 시간적 특징
Fig. 9. Temporal property of smoke

이상 나오지 않아 바깥영역에서 1개가 검출 되었다. 바깥 영역의 블록 수에서 안쪽 영역 블록 수를 빼주면 각 프레임마다 '-1', '0', '+1'이 나오게 된다. 이렇게 뺀 값이 점진적으로 증가 한다면, 연기를 내뿜는 행동을 하고 있다고 판단한다. 그와 반대로 '+1', '0', '-1'과 같은 값이 나온다면 객체의 방향이 얼굴 쪽이므로 연기가 아니다.

3.3.4 손의 움직임 검출

2.3에서 소개한 방법들과 달리 제안하는 알고리즘은 손 자체를 검출하는 것이 아니라 손의 움직임을 찾는 것이 목적이다. 따라서 피부색을 띠는 객체가 움직였는지의 여부만 확인함으로써 손의 움직임을 검출하여 실시간성을 보장한다. 흡연자는 담배를 태우기 위하여 불가피하게 담배를 쥔 팔을 움직인다. 반복적인 팔의 움직임을 검출하기 위하여 블록 단위 움직임 추정을 ROI영역 내에서 진행한다. 움직임 추정이란 그림 11와 같이 현재 프레임의 블록과 가장 유사한 블록을 이전 프레임의 탐색영역 안에서 추정하는 것이다. 현재 프레임의 블록의 위치와 추정된 블록의 위치 차이를 움직임 벡터(Motion vector)라고 한다. ROI를 5x5 블록으로 분할하고 피부색을 띠는 블록들에 대해서 움직임 벡터(Motion vector)를 찾고, 각 블록이 문턱값 이상의 크기를 가지는지 확인한다. 발생 빈도수가 높은 방향을 손이 움직인 방향으로 설정하며 손이 움직였다고 판단한다. 결과로 얻은 손의 움직임은 그림 12와 같이 위로 올리는 동작과 아래로 내리는 동작

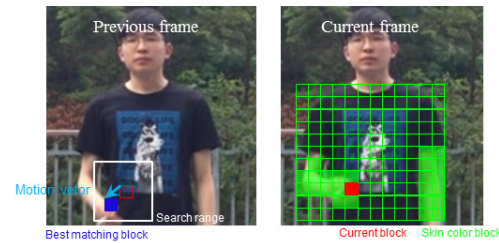


그림 10. 움직임 벡터 탐색 예시.
Fig. 10. Example of motion estimation.

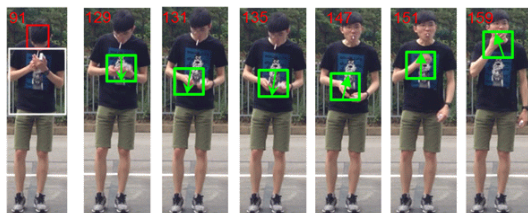


그림 11. 움직임 검출
Fig. 11. Motion detection

이 주로 발생하며, 흡연자 검출을 위한 이벤트 판단 과정에서 사용된다.

3.4 이벤트 판단

이벤트 판단 과정은 위에서 설명한 연기 검출, 손의 움직임 검출을 토대로 최종적으로 객체가 흡연자인지 판단하는 과정이다. 이벤트 탐지과정에서 연기와 손의 움직임 행위는 독립적으로 검출 되었지만 일반적으로 흡연하는 과정은 팔의 움직임이 일어난 후에 연기를 내뿜는 행위가 순차적으로 반복해서 나타나므로 두 행위에 전후관계가 있다고 볼 수 있다. 따라서, 이벤트 판단 과정은 흡연 과정을 다음과 같은 전후관계를 고려하여 흡연 여부를 판단한다.

- (1) 움직임 발생(팔 올림)
- (2) 시간 지연(담배 연기를 흡입)
- (3) 움직임 발생(팔 내림)
- (4) 담배 연기발생

제안하는 알고리즘은 실시간 검출이 목표이기 때문에, 흡연 행위가 시작되고 짧은 시간 내에 반드시 흡연자로 인지해야 한다. 따라서 연속된 비디오 프레임 상에서 이벤트 (1)이 시작되는 프레임을 기점으로 다시 이벤트 (1)이 나타나는 프레임까지를 한 번의 흡연 구간으로 정의하며, 각 흡연 구간이 시작될 때마다 이전 흡연 구간에서 이벤트 (4)가 발생했는지 확인한다. 이는 손의 움직임 사이에 연기가 발생했는지를 확인하기 위함이다. 이 흡연자 검출 과정은 그림 13과 같이 나타낼 수 있다.

연기 검출과 손의 움직임 검출이 한 프레임 내에서 공간적으로 얼굴 영역으로 인해 연결되고, 여러 개의 프레임 상에서 시간적으로 두 행위의 전후 관계를 따져 하나의 큰 행동으로 인식하므로 1장에서 소개한 행동인식 접근법 중 서술적 접근을 통해 흡연 행위를 인식하는 것으로 볼 수 있다.

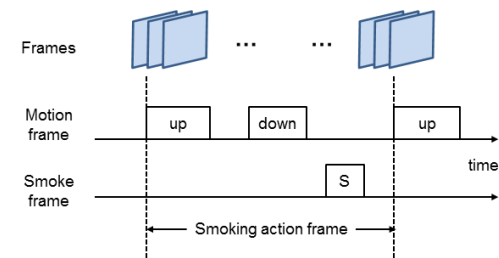


그림 12. 흡연자 검출 과정
Fig. 12. Smoker detection process

IV. 실험 결과

제안하는 알고리즘은 OpenCV 2.4.12 기반으로 구현되었으며, Intel i7-4790K 4.0GHz. 프로세서에서 실시간 성능을 보인다. 알고리즘의 성능 평가를 위해 일반적인 흡연자 영상 2개를 제작하였다. 첫 영상은 한 흡연자가 등장 하며 그 흡연자는 흡연을 하며 얼굴을 돌리고, 움직이며 흡연을 한다. 다음 영상에는 두 객체가 등장하며 객체가 흡연자이다. 이 흡연자는 중간에 등장하여 정면을 바라보며 흡연을 시작한다. 흡연을 하는 도중에 얼굴을 움직이다가 시선을 다른 방향으로 돌린다. 또한 자동차도 뒤로 지나간다. 성능 평가는 각 테스트 영상에 대하여 수식(1)을 이용하여 검출률을 계산하였다.

$$Detection\ Rate = \frac{N_{detected}}{N_{smoking}} \times 100\% \quad (1)$$

표 1은 두 테스트 영상을 실험한 결과이다. 'motion' 영상은 흡연 행위가 4번 발생하며 그 중 3번을 검출 하였고, 'two object' 영상은 흡연행위가 7번 발생하여 7번 모두 검출하였다. 각 테스트 영상에서 오검출은 발생하지 않았다. 위 실험 결과로 'two object' 영상과 같이 영상 내에 여러 객체가 있을 경우

표 1. 제안하는 알고리즘에 대한 실험 결과
Table 1. Result of experiment for proposed algorithm

| Sequence | Detection Rate |
|------------|----------------|
| motion | 75% |
| two object | 100% |
| total | 91% |

에는 흡연자가 강인하게 검출 되지만 'motion' 영상과 같이 흡연자의 움직임이 과할수록 검출률이 감소함을 보인다. 그림 14, 그림 15는 'motion', 'two object' 영상의 실험 결과이다. 흡연 행위의 과정을 나타내고 있으며 제안하는 알고리즘이 흡연자가 많은 움직임을 보일 경우와 두 개의 객체가 있을 때에도 얼굴, 움직임, 연기를 각각 강인하게 검출함을 보인다.

V. 결 론

본 논문은 비디오 영상에서 행동 인식의 계층적 방법 중 서술 기반 접근 방법을 이용하여 실시간으로 흡연자를 검출하는 알고리즘을 제안하였다. 제안 하는 알고리즘에서 흡연자는 서서 흡연한다고 가정하고 흡연자를 검출한다. 비교적 높은 학습률을 갖는 GMM 과 낮은 학습률을 갖는 GMM을 입력 영상에서 빼주어 각각 고속 움직임 전경과 전경 영상을 얻고 두 영

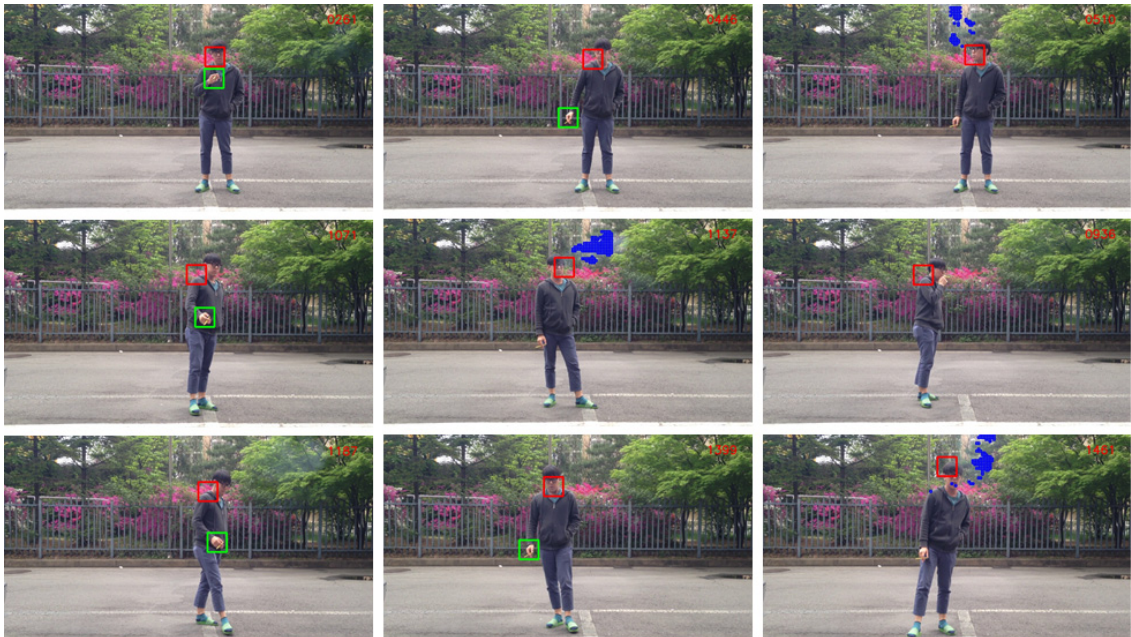


그림 13. 'motion' 영상 실험 결과.
Fig. 13. result of experiment for 'motion' test sequence.



그림 14. 'two object' 영상 실험 결과
 Fig. 14. result of experiment for 'two object' test sequence

상을 빼주어 저속 움직임 전경 영상을 얻는다. 저속 움직임 전경 영상을 노이즈 제거를 하여 흡연자인 사람을 검출하고 상단부터 1/4까지를 얼굴 영역 ROI로 정하여 일반적인 얼굴 검출 알고리즘을 사용하여 얼굴을 검출한다. 얼굴 검출에 실패 한다면 이전 장의 얼굴 크기와 같고 ROI의 중앙에 얼굴이 있다고 추정한다. 얼굴 영역을 기반으로 연기와 손의 움직임 후보 영역을 모델링하여 ROI를 설정하고 각 행위를 검출한다. 마지막으로 프레임 마다 검출 된 결과를 이용하여 두 사건 사이 전후 관계를 따져 흡연 행위로 인식한다. 알고리즘의 성능 평가를 위해 흡연 환경이 다른 두 개의 영상을 직접 제작하고 실험 하여 평균 91%의 검출률을 보였다.

References

- [1] C. Y. Jeong and J. W. Han, "Technical trends of abnormal event detection in video analytics," *Electron. Telecommun. Trends*, vol. 27, no. 4, pp. 114-122, Aug. 2012.
- [2] W. C. Wu and C. Y. Chen, "Detection system of smoking behavior based on face analysis," *IEEE Int. Conf. Genetic and Evolutionary Computing*, pp. 184-187, Aug. 2011.
- [3] S. Y. Shin and S. W. Lee, "Detection of smoking in elevator using average intensity measure," *Int. J. Multimedia and Ubiquitous Eng.*, vol. 10, no. 8, pp. 181-186, Aug. 2015.
- [4] K. H. Lee, S. Y. Shin, and Y. W. Rhee, "Extraction of smoking in elevator using robust scene change detection method," *J. KSCI*, vol. 18, no. 10, pp. 89-95, Oct. 2013.
- [5] P. Wu, et al., "Human smoking event detection using visual interaction clues," *IEEE Int. Conf. Pattern Recognition*, pp. 4344-4347, Aug. 2010.
- [6] J. K. Aggarwal and M. S. Ryoo, "Human activity analysis: a review," *ACM Comput. Surv.*, vol. 43, no. 3, pp. 1-43, 2011.
- [7] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Proc. Int. Conf. Pattern Recognition*, vol. 2, pp. 28-31, 2004.
- [8] H. H. Lin, J. H. Chuang, and T. L. Liu, "Regularized background adaptation: a novel learning rate control scheme for Gaussian mixture modeling," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 822-836, 2011.

- [9] X. Hu and C. He, "Moving object detection algorithm based on gaussian mixture model and HVS space," *British J. Appl. Sci. & Technol.*, vol. 14, no. 6, pp. 1-8, 2016.
- [10] C. Y. Lee, C. T. Lin, C. T. Hong, and M. T. Su, "Smoke detection using spatial and temporal analyses," *Int. J. Innovative Comput., Inf. and Contr.*, vol. 8 no. 7, pp. 4749-4770, 2012.
- [11] H. S. Park and C. S. Bae, "Vision-based hand gesture detection and tracking system," *J. KICS*, vol. 30, no. 12C, pp. 1175-1180, Dec. 2005.
- [12] H. S. Park and C. S. Bae, "Face and hand tracking algorithm for sign language recognition," *J. KICS*, vol. 31, no. 11C, pp. 1071-1076, Nov. 2006.
- [13] J. S. Kim, C. S. Han, and J. S. Yoo, "Real-time face tracking system," in *Proc. KICS Winter Conf.*, pp. 1964-1967, Nov. 2005.
- [14] Y. Fang, et al., "A real-time hand gesture recognition method," in *Proc. IEEE Int. Conf. Multimedia Expo*, pp. 995-998, 2007.
- [15] B. Stenger, P. R. S. Mendonça, and R. Cipolla, "Model-based hand tracking using an unscented kalman filter," in *Proc. British Machine Vision Conf.*, vol. 1, pp. 63-72, 2001.
- [16] C. Shan, Y. Wei, T. Tan, and F. Ojardias "Real-time hand tracking by combining particle filtering and mean shift," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, May 2004.
- [17] C. Shan, T. Tan, and Y. Wei, "Real-time hand tracking using a mean shift embedded particle filter," in *Proc. Int. Conf. Pattern Recognition*, vol. 40, no. 7, pp. 1958-1970, Jul. 2007.
- [18] B. Stenger, et al., "Model-based hand tracking using a hierarchical bayesian filter," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 28, no. 9, pp. 1372-1384, 2006.
- [19] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Computer Vision*, vol. 57, no. 2, 2004.

이 종 석 (Jong-seok Lee)



2016년 2월 : 광운대학교 전자공학과 졸업
2016년 3월~현재 : 광운대학교 전자공학과 석사과정
<관심분야> 영상압축, 컴퓨터 비전

이 현 재 (Hyun-jae Lee)



2016년 2월 : 광운대학교 전자공학과 졸업
2016년 3월~현재 : 광운대학교 전자공학과 석사과정
<관심분야> 영상압축, 컴퓨터 비전

이 동 규 (Dong-kyu Lee)



2012년 2월 : 광운대학교 전자공학과 졸업
2014년 2월 : 광운대학교 전자공학과 석사
2014년 2월~현재 : 광운대학교 전자공학과 박사 과정
<관심분야> 영상압축, 컴퓨터 비전

오 승 준 (Seoung-jun Oh)



1980년 2월 : 서울대학교 전자
공학과 졸업

1982년 2월 : 서울대학교 전자
공학과 석사

1988년 5월 : 미국 Syracuse
University 전기/컴퓨터공학
과 박사

1982년 3월~1992년 8월 : 한국전자통신연구원 멀티
미디어 연구실 실장

1986년 7월~1986년 8월 : NSF Supercomputer
Center 초청 학생연구원

1987년 5월~1988년 5월 : Northeast Parallel
Architecture Center 학생연구원

1992년 3월~1992년 8월 : 충남대학교 컴퓨터공학부
겸임교수

1992년 9월~현재 : 광운대학교 전자공학과교수

2002년 3월~현재 : SC29-Korea 의장 및 MPEG Forum
부의장

<관심분야> 비디오 데이터 처리, 영상압축, 멀티미
디어 시스템