복도 주행 로봇을 위한 단일 카메라 영상에서의 사람 검출

Human Detection in the Images of a Single Camera for a Corridor Navigation Robot

김 정 대¹, 도 용 태[†]

Jeongdae Kim¹, Yongtae Do[†]

Abstract In this paper, a robot vision technique is presented to detect obstacles, particularly approaching humans, in the images acquired by a mobile robot that autonomously navigates in a narrow building corridor. A single low-cost color camera is attached to the robot, and a trapezoidal area is set as a region of interest (ROI) in front of the robot in the camera image. The lower parts of a human such as feet and legs are first detected in the ROI from their appearances in real time as the distance between the robot and the human becomes smaller. Then, the human detection is confirmed by detecting his/her face within a small search region specified above the part detected in the trapezoidal ROI. To increase the credibility of detection, a final decision about human detection is made when a face is detected in two consecutive image frames. We tested the proposed method using images of various people in corridor scenes, and could get promising results. This method can be used for a vision-guided mobile robot to make a detour for avoiding collision with a human during its indoor navigation.

Keywords: Single Camera, Human Detection, Mobile Robot, Appearance-based Detection

1. 서 론

이동로봇의 주행은 환경과 위치의 정보를 기반으로 목적지까지 경로를 계획하고 로봇을 제어하는 기술이다. 이동로봇의 주행기술은 청소로봇, 군사로봇, 무인주행 자동차, AGV(Automated Guided Vehicle) 등 가정용에서부터 군사용, 산업용에 이르기까지 다양한 형태로 연구되고 있다.

비정형화된 환경에서 로봇의 주행 경로에는 많은 미지의 장애물들이 존재한다. 특히 동적 환경에서의 안전

한 주행을 위하여 장애물의 효과적인 검출이 필요하다¹. 장애물이 움직이는 동적 환경에서 로봇의 외부 센서로는 초음파 거리 센서가 널리 사용되고 있으나^{2,6}, 해상도가 매우 낮기 때문에 장애물의 존재 여부와 대략적인 거리만을 추출할 수 있을 뿐 장애물의 종류나 움직임 등에 대한 상세한 정보를 얻기 어렵다. 이러한 문제는 대체로 카메라를 사용하여 해결할 수 있으며, 근래 카메라의 가격이 급격히 낮아지고 있어 시각 기반의 이동로봇이 늘고 있다.

고정된 카메라의 경우, 동적 장애물을 검출하기 위해서는 주로 배경과 동적 객체를 분리하는 배경차감법(Background subtraction)이 널리 사용되어 왔다⁷⁻¹⁰. 이 방법은 연속적인 영상 프레임에서 배경 모델을 생성하여 움직이는 객체를 검출한다. 배경을 정적 영역으로 보고, 입력된 영상에서 배경과 어떤 임계치 이상의 차이가 나는 화소들은 동적 객체에 속하는 것으로 간주하여 검출

Received : Mar. 6. 2013; Reviewed : Apr. 4. 2013; Accepted : Jul. 25. 2013

※ This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the Core Technology Development for Breakthrough of Robot Vision Research support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2013-H1502-13-1001).

¹ Dept. of Electronic Engineering, Graduate School, Daegu University (kjd4444@daegu.ac.kr)

[†] Corresponding author: Dept. of Electronic & Electrical Engineering, Daegu University (ytdo@daegu.ac.kr)

하게 되는 것이다.

최근 감시 및 관측을 비롯한 다양한 용도로 카메라의 영상에서 사람을 자동으로 찾는 방법에 관한 연구가 활발하지만, 여전히 동적 객체 중 사람을 검출 및 추적이 어려운 대상에 속한다. 왜냐하면 영상에서 사람은 대체로 크기가 작고, 일정한 형태를 가지지 않으며, 진행 경로 또한 예측할 수 없기 때문이다^[11]. Fujiyoshi 등^[12]은 객체의 경계를 이용한 사람 검출 기법에 대해 제안한 바 있으나 이동로봇에 사용하기는 어렵다. 이는 로봇에 장착된 카메라가 움직이기 때문에 배경 모델을 생성하는 방법으로는 객체를 구분할 수 없기 때문이다.

배경과 상관없이 사람을 검출하는 기법으로 HOG (Histograms of Oriented Gradients) 검출기^[13-17]를 이용한 연구가 근래 활발히 이루어지고 있다. 그러나 HOG 검출기는 많은 학습 데이터를 필요로 하고, 학습에 의해 얻어진 특징들을 입력영상과 비교해야 하기 때문에 실시간성이 떨어진다. 처리시간을 감소시키기 위해 HOG 검출기의 특징을 줄이거나, Mean-Shift 알고리즘을 병행하여 계산량을 줄이는 방법이 연구되었지만^[14-15], 정확도가 낮아지는 단점이 있다. 동적 환경에서 장애물을 실시간으로 검출하는 효과적인 기법으로 Ulrich와 Nourbakhsh^[2]가 제안한 Appearance 기반 장애물 검출 방법이 있다. 이 기법은 이동로봇에 바닥을 향하도록 카메라를 장착하여 영상에서 바닥의 특성과 다른 성질을 가지는 화소들을 장애물에 속하는 것으로 판단하는 방법이다. 간단한 알고리즘만을 사용하기 때문에 앞서 서술한 기법들 보다 수행시간이 짧아 실시간성이 높다. 그러나 바닥에서의 반사광에 의한 오검출률이 높고, 장애물의 정체에 대한 정보를 제공할 수 없기 때문에 사람을 검출하기 위한 보다 효율적인 방법에 대한 연구가 필요하다^[18-20].

본 논문에서는 이동로봇의 안전한 자율 운행을 위하여, 로봇의 근거리 전방에 위치한 사람을 검출하는 영상 처리 기법을 제안한다. 이는 로봇의 실내 작업 공간에서 이동하는 객체가 주로 사람이며, 사람과의 충돌회피에 의한 상호간의 안전보장이 가장 중요한 사항을 고려한 것이다. 제안된 기법은 먼저 영상 내 복도의 바닥에 해당하는 근거리의 영역에서 다리 검출 알고리즘을 수행하여 빠르게 사람 후보 영역을 찾고, 이를 얼굴의 검

출로 확인하되, 비교적 긴 처리 시간이 소요되는 얼굴 검출 알고리즘은 제한된 영역 내에서만 적용하여 전체 처리 시간을 줄인다.

본 논문에서의 연구와 연관된 기존의 연구로서 단일 카메라를 이용하여 사람의 상반신을 모델링하고 추적하는 기법^[21]이 있으나, 상반신만을 찾기 때문에 거리에 관한 정보를 획득하기 어렵다. 또 다른 연구로는 센서를 융합하여 사람을 검출하는 기법들^[22-23]과 적외선 카메라^[24]를 이용하는 방법이 있다. 그러나 이러한 기법은 고가의 장치를 사용해야 하는 문제점이 있다.

2. 기존의 동적 사람 검출 기법

2.1 고정된 카메라에 의한 검출

카메라의 영상에서 사람을 검출하는 기법은 시각에 의한 자동 감시 등의 영역에서 중요하므로, 다양한 방향으로 연구되고 있다. 특히 배경차감법은 대표적인 실시간 동적 객체 검출 기법이다. 고정된 카메라를 이용하여 배경 영상을 지정하고 이와 현재 영상과의 차를 이용해서 동적 객체를 검출할 수 있다. 그러나 정적 객체는 움직임이 없기 때문에 객체임에도 배경으로 분류되고, 동적 객체는 지나간 자리에 꼬리와 같이 다수의 화소가 검출되는 문제가 발생한다^[9]. 이러한 문제는 적응배경차감법^[7]을 이용하여 해결할 수 있다. 그리고 일단 동적 객체가 찾아지면, 사람은 복잡도를 이용하여 차량 등 여타의 객체와 구분될 수 있다^[12]. 그러나 배경차감법은 카메라가 움직이게 되면 기준이 되는 배경을 생성할 수 없기 때문에 이동 카메라에서의 동적 객체 검출 문제는 다른 방향으로 접근해야 한다.

2.2 이동 카메라에 의한 Appearance 기반 검출

로봇이 자율 주행하기 위해서는 로봇의 진행에 방해가 되는 장애물을 검출하는 과정이 필요하다. 이동로봇에서 장애물 검출은 크게 두 가지로 나뉜다. 하나는 바닥에서 돌출되어 있는 물체까지의 거리를 측정하여 판단하는 방법(Range-based obstacle detection)이고, 또 다른 하나는 바닥의 특성과 다른 형태가 나타나면 장애물로 인식하는 방법(Appearance-based obstacle detection)이다^[2].

Ulrich와 Nourbakhsh는 두 번째 방법인 Appearance 기반의 장애물 검출 방법을 제안한 바 있는데, 그 적용이 간단하고 계산량이 적어서 실현이 용이하기 때문에 널리 사용된다^[2, 18-20]. 이 기법에서 장애물의 검출은 다음의 3가지 가정을 사용한다.

- (1) 장애물은 영상에서 바닥과 다른 특성을 갖는다.
- (2) 바닥은 평평하다.
- (3) 검출하고자 하는 장애물 외 여타 장애물은 없다.

이러한 가정을 기반으로 다음 4단계 처리 과정을 수행하면 영상에서 간단하게 장애물을 검출할 수 있다.

- (1) 입력 영상 필터링
- (2) 컬러 공간 변환
- (3) 참조 영역의 히스토그램 획득
- (4) 참조 영역의 히스토그램과 입력 영상의 화소 컬러 비교

먼저, 입력 영상에 가우시안 필터를 이용하여 잡음을 줄인다. 두 번째 단계로 RGB(Red Green Blue) 컬러는 빛의 양이 많거나 적을 때 영향을 크게 받으므로, 색채와 밝기를 구분할 수 있는 HSI(Hue Saturation Intensity) 컬러 영상으로 변환한다. 세 번째 단계로 로봇이 진행하는 전방 바닥을 참조 영역으로 두고, 그 영역의 H와 I값의 히스토그램을 구한다. 이렇게 구해진 히스토그램은 평균 필터를 통하여 잡음을 제거하고, 각각의 히스토그램에서 일정 빈도수 이상을 갖는 값들만 남기고 나머지는 제거한다. 이런 방법을 통해 참조 영역을 규정하는 임계치 집합을 얻을 수 있다. 마지막 단계로 앞 단계에서 얻은 임계치를 이용하여 입력되는 영상의 특정 화소값이 임계치 집합에 포함되면 바닥으로 간주하고, 임계치 집합에 포함되지 않으면 장애물로 간주한다. Appearance 기반 장애물 검출 기법은 off-line으로 처리하여야 하는 작업의 양이 적고 실시간 장애물 검출이 가능한 기법이기 때문에 실용성이 높다.

3. 다가오는 사람 검출을 위해 제안하는 기법

기존의 Appearance 기반의 기법^[2]은 단순하여 이동로봇에의 활용 가능성이 높으나, 밝기와 컬러에 민감하여

오검출의 빈도가 잦다. 그림 1은 그러한 예를 보여주는데, 복도의 바닥에 해당하는 영상의 하단부에서 사람의 발과 다리뿐 아니라 반사된 조명과 같은 불필요한 부분이 검출되었다. 본 논문에서는 이러한 현실적인 문제점을 개선할 수 있는 알고리즘을 서술한다.

제안된 기법은 먼저 다리로 추정되는 부분을 찾고, 찾은 결과를 얼굴 검출을 통하여 확인하는 방식이다. 이는 로봇이 주행하는 복도에서 다가오는 사람을 찾는다면, 먼저 찾을 수 있는 것이 사람의 다리 영역이라는 점과 얼굴이 사람을 특징짓는 중요한 요소라는 점에 착안한다. 특히 탐색의 영역을 적절히 제한하여 영상처리의 효율을 높이면서 검출의 강건성을 증대시킬 수 있도록 하였다. 제안하는 기법의 전체 처리 과정은 그림 2에 요약하였다.

Fig. 1. Example of an appearance-based obstacle detection^[2] in the floor region of the image

Fig. 2. Flow chart of the proposed method

본 논문의 접근법과 관련된 기존의 연구로 Bellotto와 Hu^[23]는 레이저 센서를 사용하여 다리 형태를 찾고, 카메라로 얼굴을 찾는 기법을 제시한 바 있다. 그러나 사용되는 레이저 센서가 고가이기 때문에, 본 논문에서와 같이 저가의 단일 카메라만을 이용한 방법이 현실적으로 장점을 가진다.

3.1 Appearance 기반 다리 영역의 검출

기존의 연구에서는 카메라의 방향을 바닥으로 향하게 하여 주변의 장애물을 검출하지 않거나, 최대한 외부 장애물이 영상에 들어가지 않게 만든 후 HSI컬러를 이용하여 새롭게 나타나는 장애물에 대해 실험을 하였다^[2]. 이러한 제한에도 검출되는 외부 장애물은 소실점과 경계 검출기를 이용하여 로봇의 진행과 관련 없는 장애물을 제거하는 방법을 사용한다^[18]. 그러나 이 경우 바닥의 정보만 이용하기 때문에 장애물 전체의 정보를 이용할 수 없고, 따라서 장애물의 종류나 움직임 등을 알 수 없다.

본 논문에서는 기존의 연구에서와는 달리 장애물의 전체 영상을 얻기 위해 카메라를 바닥이 아닌 전방을 주시하도록 설치하였다. 이렇게 해서 얻은 영상은 대체로 중앙 부근이 소실점이 되고, 하단 부분이 바닥, 상단 부분이 천장이 될 것이다. 이와 같은 설치의 카메라의 영상을 장애물 검출 이외에도 사용자와의 상호작용 등 여러 영역에 활용할 수 있는 장점을 가진다.

로봇에 전방을 향하도록 탑재된 카메라의 영상에서 로봇이 주행하는 복도의 바닥은 영상의 하단이므로, 이 부분에 ROI(Region of interest) 영역을 설정하면 로봇의 진행에 방해가 되는 근거리의 장애물을 검출할 수 있다. 이 때 영상의 중앙 부근에서 소실점이 나타나는 투사변환(perspective transform)을 고려하여 ROI 영역은 사다리꼴 모양으로 작게 정할 수 있다. 이와 같이 설정하면 기존의 기법^[2]에서 전체 영상을 처리하는 것과는 달리 로봇이 곧 이동하여 장애물 회피가 필요한 부분에 해당하는 영상만 다룰 수 있고, 벽이나 천장에 존재하는 잡다한 변화에 의한 탐지 오류를 방지할 수 있다. 설정된 영역의 RGB 컬러 영상에는 먼저 잡음을 줄이기 위해 3×3 가우시안 필터를 적용한다. 또한 바닥에 반사되는 빛에

의한 오류를 없애기 위해 RGB의 최대값과 최소값을 이용하여 다음과 같은 식 (1)을 적용한다. 즉, RGB의 차이가 작고 최대값이 크면 반사된 빛에 해당하는 화소로 간주하고 제거하게 된다.

$$O = \begin{cases} 1, & Th_L < \frac{\max(R,G,B)}{\max(R,G,B)-\min(R,G,B)} < Th_H \\ 0, & \text{Otherwise} \end{cases} \quad (1)$$

다음 단계로 주변 환경의 밝기 변화에 강건한 검출을 위하여 ROI 영역을 HSI 컬러로 변환하여 처리한다. 먼저 HSI 컬러 중 H와 I의 히스토그램을 얻는다. 얻어진 히스토그램은 평균 필터를 사용하여 필터링한다. 그리고 각각의 히스토그램에서 빈도수가 많은 영역만 남기고 나머지는 제거한다. 이로부터 결정되는 임계치를 입력 영상과 비교하여 얻어지는 일정 크기 이상의 화소 집합에 의하여 장애물이 검출된다.

검출된 영역 중 사람의 다리에 해당하는 부분을 찾기 위해 간단한 다리 모델을 설정하여 사용하였다. 복도 바닥에 설정된 ROI 영역에서 발과 다리의 모양은 그 복잡도가 어느 정도 일정하다. 또한 세로 방향이 가로 방향보다 더 길게 나타날 것으로 예측할 수 있다. 이러한 사실을 근거로 복잡도와 가로 대 세로 비를 이용한 두 가지 판정 기법을 적용한다. 먼저 식 (2)의 C는 복잡도를 나타내는데, 객체의 둘레(Obj_c)와 면적(Obj_a)을 이용해서 간단히 계산할 수 있다. 한편 R은 검출된 영역의 가로(Obj_w)와 세로(Obj_h)의 비율을 나타낸다. 일반적으로 사람의 다리는 세로로 길게 나타나므로 그 값이 1보다 작다. 그러나 ROI 영역에서 다리 전체가 검출되지 않고 발이 먼저 보이게 됨을 고려하여 임계값(Th_r)을 1.05정도로 설정하였다. 식 (3)에서 C와 R이 모두 임계치 조건을 만족하면 검출된 영역이 다리에 해당한다고 판정한다.

$$C = \frac{Obj_c}{Obj_a}, \quad R = \frac{Obj_w}{Obj_h} \quad (2)$$

$$L = \begin{cases} 1, & Th_{CL} < C < Th_{CH} \text{ AND } R < Th_r \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

3.2 얼굴 검출 및 사람의 판정

본 논문은 이동하는 로봇의 전방에서 로봇 방향으로

다가오는 사람의 실시간 검출을 주요한 문제로 한다. 앞 절에서 서술한 방법에 의해 검출된 영역이 실제 사람의 다리에 해당한다면 그 수직 상단에서 사람의 얼굴이 나타나게 될 것이다. 우리는 영상에서 사람의 얼굴을 검출하기 위하여 널리 사용되고 있는 Viola와 Jones^[25]의 기법을 사용하여 사람의 존재여부를 확인하였다. 이 알고리즘은 다음 4 단계로 얼굴을 검출한다.

- (1) Haar-like feature 검출
- (2) Integral image 계산
- (3) Adaboost(Adaptive boosting) 알고리즘 적용
- (4) Cascaded classifier 분류

Lienhart와 Maydt^[26]는 얼굴 검출을 위해 Edge 특징, Line 특징, 그리고 Center-surround 특징을 그림 3과 같이 영상에 적용하여 얼굴 영역에 대한 패턴을 찾아내었다. 그러나 이러한 과정을 반복하려면 많은 시간이 요구되는데, Integral image를 이용하여 전체 영상에서 각각의 특징값의 합계를 미리 생성하여 보다 빠르게 얼굴을 검출할 수 있다. Adaboost 알고리즘은 Haar-like feature에 의한 약분류기의 결과를 하나씩 합쳐서 강분류기로 만드는 과정이다. Cascaded classifier는 강분류기를 만드는 과정에서 약분류기를 제거하거나 추가하여 원하는 검출률을 얻을 때까지 반복 학습하는 과정이다^[25]. 그러나 이러한 방법은 얼굴 검출에 소요되는 시간이 대체로 긴 문제가 있다. 그래서 본 논문에서는 탐색 영역을 적절히 제한하여 얼굴 검출 시간을 줄였다.

얼굴 검출을 위한 탐색 영역은 3.1절에서 얻어진 다리 영역의 수직 상단에 위치하지만, 구체적인 사항은 사용된 카메라의 사양에 따라 달라진다. 예를 들어, 본 논문

의 연구를 위해 사용된 카메라는 영상의 크기가 320×240 이었고, 이 카메라로 우리가 실험한 바로는, 그림 4와 같이 검출된 다리 영역의 좌측 하단 좌표를 (x, y) 라고 할 때 얼굴 탐색 영역의 x축 상의 위치는 다리 검출 영역의 폭(width)의 ± 10 화소로 설정하면 적당하였다. 한편 얼굴 탐색 영역은 y값이 커질수록 높이 위치하게 된다. 즉, 사람이 가까워질수록 영상에서 발은 아래로 얼굴은 위로 위치하게 됨을 고려하여, 얼굴 탐색 영역을 y값의 보수를 기준으로 정하였다. 그림 5에서 보인 바와 같이 카메라로부터 2m의 거리에 있는 사람의 키가 170cm라고 할 때 이는 우리가 사용한 카메라에서 약 200화소에 해당하였다. 만약 검출하고자 하는 사람의 키가 최고 200cm라고 한다면, 얼굴 검출 영역의 위치는 상측으로 약 40화소 확장하게 된다. 같은 방식으로 170cm보다 작은 키를 가진 사람의 얼굴을 검출하기 위하여 아래로도 40화소만큼 확장하였는데, 특히 사람이 가까워질수록 ROI 내 검출된 다리의 길이, 즉 height가 커짐까지 고려하였다. 구체적으로 얼굴 검출 영역은 좌측 상단 좌표를 $(x-10, 240-(y+40))$ 로, 우측 하단 좌표를 $(x+width+10, 240-(y-height-40))$ 로 하는 사각형이 된다. 하나의 예로써, $(x, y) = (120, 180)$ 이고 $width = 30, height = 40$ 일 때, 얼굴 탐색 영역의 좌측 상단 좌표는 $(110, 20)$, 우측 하단 좌표는 $(160, 140)$ 이 된다.

이상의 절차와 같이 얼굴 탐색 영역을 제한하면 전체 영역을 검색하는 것 보다 계산 비용이 크게 낮아진다. 만약 탐색 영역 내에서 얼굴이 검출되지 않는다면, 사람이 없다고 판단하고 Flag를 OFF시켜서 다음 영상을 처리한다. 그러나 탐색 영역에서 얼굴이 검출된다면, 그 영역을 사람 후보 영역으로 판단하고, Flag를 ON시킨다. 다음 영상 프레임에서도 앞의 과정들을 반복하고 얼굴이 검출되면 기존의 Flag가 ON인지 확인하여 ON이면 사람 영역으로 최종 판정한다. 즉, 두 영상 프레임 이상에서 사람이 연속적으로 검출되면 최종 판정을 내리게 된다. 일단 사람이 검출되면 로봇은 검출된 사람의 x축 상의 위치를 기준으로 비어 있는 공간으로의 회피구동이 가능해진다.

Fig. 3. Haar-like feature extraction for face detection

Fig. 4. Face detection region specification.

Fig. 5. Image size variation depending on the distance between the camera and a human.

4. 실험 및 결과

제안된 방법은 실험을 통하여 그 성능을 점검하였다. 실험에 사용된 카메라는 저가의 USB용으로 로지텍 pro4000 모델이었으며, 컴퓨터는 Intel i7 코어 기반으로 CPU 클럭은 1.73GHz이고 메모리는 4GB였다. 실험 영상의 크기는 320×240으로 프레임 레이트는 24fps 이었다. 이와 같은 시스템으로 한 장의 영상을 처리하는 데는 평균적으로 80ms 정도 소요되었고, 다리를 검출하는데 20ms, 얼굴을 검출하는데 60ms 정도 소요되었다.

그림 6(a)는 실험의 예를 보여준다. 영상의 중심으로부터 상단은 천장이고, 하단은 바닥, 그리고 좌우는 벽으로 이루어져 있다. 다리 검출을 위한 ROI 영역은 복도의 바닥 안에 존재하면서 로봇의 진행 방향을 고려하여 사다리꼴 모양처럼 설정하였다. 그림 (c)는 ROI 영역에서 다리를 검출하고 이로부터 얼굴까지 검출한 결과를 보여주며, 그 다음 영상 프레임에서 그림 (d)와 같이 얼굴을

연속적으로 검출함으로써 그림 (b)에서 보인 바와 같이 사람을 최종적으로 검출할 수 있었다.

그림 6의 실험 결과를 정량적으로 분석하기 위하여, 카메라와 사람 사이의 거리에 따라 영상의 ROI 내에 사람의 다리가 존재하지 않는 경우, 다리가 ROI에 존재하고 그 상단에 얼굴도 존재하는 경우, 그리고 사람이 너무 근거리에서 있어서 얼굴이 보이지 않는 경우의 3가지로 영상을 구분하고, 각 구간마다 10장의 영상을 무작위로 선택하였다. 단, 제안된 기법은 연속된 두 장의 영상을 처리하므로, 선택된 영상의 이전 영상도 함께 실험에 활용하였다. 각 영상에서 제안된 기법에 의하여 검출된 영역은 실제와 비교하여 2/3 이상 겹치면 성공(True Positive)으로 판단하고, 아니면 오검출(False Positive)로 판단하여 계수하였다. 그리고 사람이나 다리가 있는데도 불구하고 검출하지 못하면 실패(False Negative)로 계수하였다. 이로부터 성능지표는 식 (4)와 같이 검출률(Recall rate)과 구분율(Precision rate)^[26]로 계산할 수 있다. 그림 6의 실험 영상에 대해서 기존의 얼굴 검출 기법^[25]만 영상의 전체에 적용하여 사람을 찾는 경우와 제안한 기법을 비교 실험한 결과는 표 1에 보인 바와 같다. 실험 영상들 중에 얼굴이 존재하는 영상은 10장이고 나머지는 얼굴이 존재하지 않는다. 즉, TP+FN의 합은 10이다. 표 1을 보면 제안된 기법이 얼굴만 찾는 경우에 비해 구분율을

Fig. 6. Human detection by the proposed method: (a) Image at time=t-1, (b) Image at time=t, (c) Detection result at t-1, (d) Detection result at t

Table 1. Quantitative result analysis for the experiment using the video images of Fig. 6.

Methods	TP	FN	FP	Recall rate	Precision rate	Time (ms)
Face detection by [25]	8	2	58	0.80	0.12	1505
Face search after leg detection	9	1	10	0.90	0.47	88
Referencing previous result	8	2	2	0.80	0.80	88

크게 향상시켰으며, 특히 두 장의 연속 영상 처리 결과에 의하여 판정을 내림으로써 구분율이 더욱 상승하였다. 그리고 탐색 영역의 제한으로 인하여 수행시간은 크게 줄었다.

$$\begin{aligned} \text{Recall rate} &= \frac{TP}{TP+FN} \\ \text{Precision rate} &= \frac{TP}{TP+FP} \end{aligned} \quad (4)$$

제안된 기법을 다양한 환경에서 실험해 보기 위하여 그림 7과 같이 여러 가지 비디오를 획득하여 처리하였다. 그림에서 좌측 열은 장면 영상이고, 우측 열은 제안한 기법에 의한 처리 결과이다. 검출의 결과를 상세히 살펴 보면, #1 그림 내 우측 부분에서 복도에 있는 상자의 일부가 ROI 내에 존재하지만 다리 모델의 점검에서 제외되어 검출되지 않았다. #2 영상에서는 다리로 오검출된 부분이 얼굴 탐색에 의한 확인과정에서 수정되었으며, #6과 #7의 영상에서도 비슷한 상황이 나타난다. #8의 영상에서는 다가오는 사람과 두 명의 서 있는 사람이 전방에 존재하며, 단지 다가오는 사람만 검출되었다. 검출되지 않은 정지한 사람의 경우, 얼굴 검출에 실패한 것이 요인이다. #9 영상에서는 두 명의 다가오는 사람들이 모두 잘 검출되었다. 이들 영상을 사용한 실험의 정량적인 성능 지표는 그림 8에 나타난 바와 같다.

Fig. 7. Experiments using various image sets. Left images show experimental scenes and right images show detection results

Fig. 8. Quantitative analysis results for the experiments using the image sets of Fig. 7

5. 결 어

본 논문에서는 실내의 복도를 자율 주행하는 이동로봇에 장착된 저가의 카메라를 이용하여 로봇의 방향으로 다가오는 사람을 효율적으로 검출하는 방법을 제시하였다. 바닥의 컬러 특성과 다리의 간단한 형태적 모델을 이용하여 영상의 일부분에서 사람의 다리를 찾고, 이로부터 한정되는 영역에서 얼굴을 검출함으로써 사람을 판별하였다. 이와 같은 접근법은 기존의 얼굴 검출 기법을 전체 영상에 적용한 경우와 비교하여 소요 시간을 크게 줄일 수 있었을 뿐 아니라, 구분율도 상당히 개선할 수 있었기 때문에, 실내 로봇의 주행에 실제 활용할 수 있을 것으로 기대된다.

제안된 방법과 관련하여 발견된 문제점 중의 하나는 원거리에 있거나 돌아서 있는 사람의 경우 검출되지 않는다는 점이다. 이 문제는 사용된 기존의 얼굴 검출 기법의 성능과 관련되어 있다. 다수의 사람이 존재할 경우 사람의 수에 따라 얼굴 탐색 영역의 크기가 달라지게 되어, 결국 전체 처리 시간이 일정하지 못하게 되는 문제도 있다. 따라서 이와 같은 문제를 개선하는 연구가 향후 필요할 것이다.

References

- [1] W. P. Yu, S. L. Choi, J. Y. Lee, and S. H. Park, "Robot navigation technology and its standardization trends," *Electronics and Telecommunications Trends*, vol. 26, no. 6, pp. 108-119, December, 2011.
- [2] I. Ulrich and I. Nourbakhsh, "Appearance-based obstacle detection with monocular color vision," *Proc. AAAI*, 2000.
- [3] S. Kim, "Robot vision technology trends for intelligent mobile robots," *Magazine of Korea Robotics Society*, vol. 9, no. 1, pp. 26-35, February, 2012.
- [4] N. A. Ogale, "A survey of techniques for human detection from video," *Survey*, University of Maryland, 2006.
- [5] Y. Choi, W. Choi, and J. Song, "Obstacle avoidance of a mobile robot using low-cost ultrasonic sensors with wide beam angle," *Journal of Institute of Control, Robotics and Systems*, vol. 15, no. 11, pp. 1102-1107, November, 2009.
- [6] S. Kim and H. Kim, "Simple and complex obstacle detection using an overlapped ultrasonic sensor ring," *Int. Conf. on Control, Automation and Systems*, pp. 2148-2152, October, 2012.
- [7] R. Collins et al., "A System for video surveillance and monitoring," *Carnegie Mellon University Robotics Institute Technical Report*, CMU-RI-TR-00-12, 2000.
- [8] M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 657-662, 2006.
- [9] C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1337-1342, October, 2003.
- [10] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, 2000.
- [11] Y. Do, and T. Kanade, "Counting people from image sequences," *Proc. Int. Conf. on Imaging Science, Systems and Technology*, pp. 185-190, 2000.
- [12] H. Fujiyoshi, A. Lipton, and T. Kanade, "Real-time human motion analysis by image skeletonization," *IEICE Transactions on Information and Systems*, vol. E87-D, no. 1, pp. 113-120, 2004.
- [13] N. dalal, and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 886-893, 2005.
- [14] Q. Zhu, M. C. Yeh, K. T. Cheng, and S. Avidan, "Fast

- human detection using a cascade of histograms of oriented gradients,” IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 1491-1498, 2006.
- [15] Y. Cui, L. Sun, and S. Yang, “Pedestrian detection using improved histogram of oriented gradients,” VIE 2008 5th Int. Conf. on IET, pp. 288-392, 2008.
- [16] Y. Pang, H. Yan, Y. Yuan, and K. Wang, “Robust CoHOG feature extraction in human-centered image/video management system,” IEEE Trans. on Systems, Man, and Cybernetics, vol. 42, no. 2, pp. 458-568, 2012.
- [17] P. Dollar, C. Wojek, B. Schiele, and P. Perona, “Pedestrian detection: An evaluation of the state of the art,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 34, no. 4, pp. 743-761, 2012.
- [18] N. Tada, T. Saitoh and R. Konishi, “Mobile robot navigation by center following using monocular vision,” SICE Annual Conf., September, 2007.
- [19] R. Derhgawen and D. Ghose, “Vision based obstacle detection using 3D HSV histograms,” 2011 Annual IEEE, India Conf. (INDICON), pp. 1-4, November, 2011.
- [20] J. Kim and Y. Do, “Moving obstacle avoidance of robot using a single camera,” Proc. Int. Symp. on Robotics and Intelligent Sensors, vol. 41, pp. 911-916, 2012.
- [21] C. Hu, X. Ma, X. Dai, and K. Qian, “Reliable people tracking approach for mobile robot in indoor environments,” Robotics and Computer-Integrated Manufacturing, vol. 26, no. 2, pp. 174-179, 2010.
- [22] K. S. Ong, Y. H. Hsu, and L. C. Fu, “Sensor fusion based human detection and tracking system for human-robot interaction,” 2012 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems(IROS), pp. 4835-4840, October, 2012.
- [23] N. Bellotto, and H. Hu, “Multisensor-based human detection and tracking for mobile service robots,” IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 39, no. 1, pp. 167-181, 2009.
- [24] A. Fernández-Caballero, J. C. Castillo, J. Martínez-Cantos, and R. Martínez-Tomás, “Optical flow or image subtraction in human detection from infrared camera on mobile robot,” Robotics and Autonomous Systems, vol. 58, no. 12, pp. 1273-1281. 2010.
- [25] P. Viola, and M. Jones, “Robust real-time face detection,” Int. Journal of Computer Vision, vol. 57, no.

2, pp. 137-154, 2004.

- [26] R. Lienhart and J. Maydt, “An extended set of haar-like features for rapid object detection,” Int. Conf. on Image Processing 2002, vol. 1, pp. 903-903, 2002.

김 정 대

2004 대구대학교 정보통신공학부 제어계측공학과(공학사)

2006 대구대학교 정보통신공학과 제어계측전공(공학석사)

2009~현재 대구대학교 전자공학과 박사과정
관심분야 : 이동로봇, 영상처리, 인공지능망

도 용 태

경북대, 서강대, Univ. of Hull(영)의 전자공학과에서 각각 공학사, 공학석사, PhD. CMU의 로봇연구소, Imperial College London, Nat'l Univ. of Singapore에서 객원교수 및 방문과학자