카메라-레이저스캐너 상호보완 추적기를 이용한 이동 로봇의 사람 추종

Person-following of a Mobile Robot using a Complementary Tracker with a Camera-laser Scanner

김 형 래, 최 학 남, 이 재 홍, 이 승 준, 김 학 일*
(Hyoungh-Rae Kim¹, Xue-Nan Cui², Jae-Hong Lee², Seung-Jun Lee², and Hakil Kim^{2*})

¹Department of Robotics, Inha University

²Department of Information & Communication Engineering, Inha University

Abstract: This paper proposes a method of tracking an object for a person-following mobile robot by combining a monocular camera and a laser scanner, where each sensor can supplement the weaknesses of the other sensor. For human-robot interaction, a mobile robot needs to maintain a distance between a moving person and itself. Maintaining distance consists of two parts: object tracking and person-following. Object tracking consists of particle filtering and online learning using shape features which are extracted from an image. A monocular camera easily fails to track a person due to a narrow field-of-view and influence of illumination changes, and has therefore been used together with a laser scanner. After constructing the geometric relation between the differently oriented sensors, the proposed method demonstrates its robustness in tracking and following a person with a success rate of 94.7% in indoor environments with varying lighting conditions and even when a moving object is located between the robot and the person.

Keywords: object tracking, person-following, mobile robot, camera-LRF extrinsic calibration

I. 서론

최근 로봇 기술은 다양한 분야에서 발전하고 있다. 특히 인간 중심의 기술로써 서비스 로봇이나 이동 보조 로봇이 개발되는 추세이다. 이런 로봇들은 사람과의 협업을 위해 실내/외 환경을 인식해야 할 뿐만 아니라 사람의 행동 인식이 필요하다. 또한, 인간-로봇 상호작용(human-robot interaction)의 관점에서 로봇이 사람과 가까운 거리를 유지하며 이동하는 것이 다양한 상호작용을 적용하는데 효과적이다.

로봇에서의 컴퓨터비전 기술은 시각을 담당한다. 특히 카메라나 다른 비전 센서로부터 획득된 데이터에서 특정 물체를 추적하는 것은 그 정보를 제어부에 전달함으로써 추종까지 할 수 있게 하는데 큰 역할을 하며 이를 위해 높은 정확성을 가진 추적이 요구된다. 하지만 이동 로봇에서의 추종은 낮과 밤, 실내/외 같은 다양한 조명 조건과 여러 가지 물체 색상, 형태 변화 아래에서 이루어지기 때문에 많은 문제가 발생한다[1].

컴퓨터비전 분야에서 일반적으로 추적에 실패하는 요인은 카메라를 기준으로 내부적인 것과 외부적인 것으로 나뉜다. 먼저 내부적인 요인은 추적하려는 물체의 자세 및 모양 변화와 같은 대상 물체 자체적인 변화이고, 외부적인 요인은 카

메라의 떨림, 시점의 변화, 조명의 변화, 복잡한 배경, 다른 물체에 의한 가려지는 현상과 같이, 대상 물체가 변하는 것 이외에 나타나는 현상이다. 이러한 많은 내부적, 외부적 요인에 의해서 성능이 떨어지는 문제점을 해결하기 위해 다양하게 변화하는 외부 환경에 대해서 강인하게 물체를 추적할 수 있도록 많은 연구가 진행되고 있다[1,2].

기존의 사람 추종 분야에서 사용한 방법 중 하나는 사람에게 발광 장치를 입혀서 위치를 알아내는 연구가 있다[3,4]. 이는 인공적인 장치를 입히는 등 시스템 주변의 인프라와 사전 캘리브레이션(calibration)이 되어있어야 하기 때문에 실제 환경에서 적용의 한계가 있다. [5,6]에서는 카메라를 기반으로 대상의 얼굴을 검출하여 추종하였으나, 얼굴 검출을 적용하는 것은 사람이 로봇을 정면으로 응시해야 하기 때문에 사용자가 매우 불편할 수 있다. 게다가 [5]에서는 조명과 배경 변화에 강인하지 못하고, [6]에서는 레이저스캐너(laser scanner, LRF)를 사용하였음에도 불구하고 단순히 사람의 다리를 검출하는 것으로 이용되었다.

사람 추종의 또 다른 방법은 미리 저장해놓은 지도를 이용하는 것이다[7]. 이는 레이저스캐너 기반 사람 추종을 위하여 주변 환경에 대한 지도를 필요로 한다. 이 방법은 다양한 환경에서 적용하기 힘들고, 작동 환경이 변하면 지도 또한 변경해주어야 하기 때문에 구동에 한계가 있다. 또한 본 논문과 유사한 [8]의 방법은 레이저스캐너를 이용하여 로봇과 사람 사이의 위치와 거리를 구한다. 그러나 이 방법은 물체에 대한 상대 거리만을 이용하여 안전거리를 확보하는데 지나지 않는다.

본 논문의 기초가 되는 형태 특징 및 색상 정보에 기반한 IPCA (Incremental Principal Component Analysis) 물체 추적은 비

* Corresponding Author

Manuscript received August 12, 2013 / revised October 10, 2013 / accepted October 25, 2013

김형래: 인하대학교 로봇공학전공(hrkim@vision.inha.ac.kr)

최학남, 이재홍, 이승준, 김학일: 인하대학교 정보통신공학과

(xncui@vision.inha.ac.kr/jaehong@vision.inha.ac.kr/sjlee@vision.inha.ac.kr/hikim@inha.ac.kr)

※ 본 연구는 지식경제부 ‘산업원천기술개발사업’의 일환인 “노약자를 위한 감성교감 및 이동보조 서비스 로봇 개발” 과제(10038574)의 지원에 의하여 연구되었음.

교적 강인하게 추적이 되는 것을 실험을 통하여 증명하였으나, 추적 대상이 카메라 화각(field-of-view)을 벗어나는 경우에 대한 실험이 부족하다[9]. 이러한 문제를 해결하기 위해 전방향(omni-directional) 카메라를 이용하여 화각에 구속 받지 않는 물체 추적에 관한 연구가 되었다[10]. 전방향 카메라를 이용하면 광범위한 화각을 갖기 때문에 추종하는 대상을 잃지 않고 추적할 수 있는 반면, 각 사람마다 인공 신경망을 이용하여 사전에 학습을 해야 하기 때문에 알고리즘의 일반성을 제약시킨다.

이동 로봇의 추종 주행 시 필요한 목표지점 계산 과정은 추종 대상에 대한 거리와 각도를 입력으로 사용하기 때문에 단일 카메라만으로는 한계가 있다. 따라서 본 논문에서는 카메라를 이용하여 조명과 환경 변화에 강인한 물체 추적을 적용하고, 사람이 카메라 화각에서 벗어날 때는 레이저스캐너를 이용한 추적을 하여 추종의 입력인 상대 거리와 각도를 계속적으로 구할 수 있는 상호보완 추적기(complementary tracker)를 제안한다.

본 논문의 구성은 다음과 같다. II 장에서는 영상 내의 사람에 대해 특징을 검출하고 이를 실시간 학습(online learning)에 적용하는 방법을 기술하고, III 장에서는 파티클 필터(particle filter)를 이용하여 추적하며, IV 장에서 카메라와 레이저스캐너 사이의 기하적 관계를 알 수 있는 캘리브레이션과 추적 대상이 영상 밖으로 벗어났을 때 레이저스캐너를 이용한 추적의 재개에 대하여 소개한다. V 장에서 제작된 플랫폼의 센서 구성에 대한 설명과 다양한 환경에서의 실험을 통한 성능 평가를 하고, VI 장에서 결론을 맺는다.

II. 형태 특징 추출 및 실시간 학습

1. 형태 특징 추출

컴퓨터 비전 분야에서 물체 추적 시 사용하는 다양한 특징은 크게 색상 특징과 형태 특징으로 나눌 수 있다. 물체의 색상 정보는 가장 직관적으로 사용할 수 있는 특징이다. 컴퓨터로 색상을 표현하는 방법은 주로 빛의 3원색인 빨강, 초록, 파랑을 이용한 RGB 색공간(color space)이지만 사람이 인지하는 색과는 차이가 있다[11]. 또한 색상 기반의 물체 추적은 조명의 변화나 배경의 변화에 민감하다[12]. 특히 로봇에서의 물체 추적은 배경과 조명 조건이 쉽게 바뀌기 때문에 더 강인한 특징을 필요로 한다. 따라서 본 논문에서는 색상 기반의 특징보다 조명의 영향을 덜 받는 형태 특징으로 사람을 추적하고, 형태 특징의 한 종류인 물체의 외곽선을 히스토그램화하여 비교적 강인한 특징을 추출한다.

대부분의 경우, 추적하고자 하는 대상 물체의 외곽선은 배경과 뚜렷한 밝기 차이를 보인다. 이 차이를 이용하여 영상에서의 외곽선을 추출하는 작업은 1차 미분인 기울기(gradient)를 구하는 것과 같다. 최근 많은 연구에서 이용된 HOG (Histogram of Oriented Gradient) 방법은 이름에서 의미하듯이 방향성을 가진 기울기 값을 히스토그램화 하여 특징으로 사용한 것이다. 이 방법은 특히 보행자 검출에 뛰어난 성능을 보이는 것으로 알려져 있다[13]. 본 논문에서는 HOG의 단순화된 방법을 이용하고, 이는 다음의 순서로 계산된다.

그림 1. 8개의 방향성을 가지는 24(6x4)개의 HOG 서술자.

Fig. 1. 24 HOG descriptor with 8 orientation bin.

- 1) 입력된 gray 영상 I 에서의 각 픽셀 $I(x, y)$ 의 X, Y 방향의 기울기 값을 계산

$$\begin{aligned} g_x(x, y) &= I(x+1, y) - I(x-1, y) \\ g_y(x, y) &= I(x, y+1) - I(x, y-1) \end{aligned} \quad (1)$$

- 2) 계산된 기울기를 이용하여 크기(magnitude) $m(x, y)$ 와 각도(angle) $\theta(x, y)$ 를 계산

$$\begin{aligned} m(x, y) &= \sqrt{g_x(x, y)^2 + g_y(x, y)^2} \\ \theta(x, y) &= \tan^{-1} \left(\frac{g_y(x, y)}{g_x(x, y)} \right) \end{aligned} \quad (2)$$

- 3) $n \times n$ 격자(grid) 안에서 8개의 방향성 저장소 u 로 누적과 정규화(normalization)를 이용한 서술자(descriptor) 형성

$$D_u = \frac{1}{n^2} \sum_{i=1}^{n^2} m_i(x, y) \delta(\theta_i(x, y) - u), \quad u = 1, \dots, 8 \quad (3)$$

식 (1)에서의 값은 X, Y 방향에서의 1D 마스크를 이용한 기울기 값을 계산한 것이고, 식 (2)는 각 픽셀에서의 기울기의 크기와 방향을 계산한다. 식 (3)에서의 D_u 는 각 저장소에 정규화된 기울기의 값을 의미하며 $\delta(\theta_i(x, y) - u)$ 는 픽셀 (x, y) 에서의 기울기 각도의 값이 해당 저장소의 범위에 속할 때만 누적함을 나타낸다. 그림 1은 실제 영상에서 기울기 히스토그램을 얻는 과정을 나타낸 것이다. 이를 실시간 학습 및 파티클 필터를 이용한 추적에 적용한다.

2. IPCA를 이용한 실시간 학습

이동 로봇에 장착된 카메라를 통해 추적 알고리즘을 적용할 때 대상 물체는 매 프레임마다 다른 형태를 취하고 있으며, 배경과 조명의 변화로 인한 특징 변화가 추적 실패의 원인이 된다. 이를 효과적으로 해결하기 위해 실시간으로 물체의 정보를 갱신해야 한다. 본 논문에서는 이러한 문제를 해결하기 위해 패턴인식 분야에서 많이 응용되는 PCA (Principal Component Analysis) 기반의 접근 방법을 적용하였다. 기존에 PCA를 이용한 추적은 사전에 특징을 학습하여야만 했다. 반면에 Incremental-PCA 방법은 실시간 학습하여 물체 추적에 적용하기 위한 방법으로 연구되었다[14-16].

IPCA는 PCA와 같이 특정 값들의 주성분을 분석하는 것이 주요 목표이다. 각 특징들 D_u 를 d 차원 벡터 형태인 I_i 로

정의한다면 $d \times n$ 의 행렬 $A = [(I_1 - \bar{I}) \cdots (I_n - \bar{I})]$ 를 만들 수 있고, 여기에서 $\bar{I} = \frac{1}{n} \sum_{i=1}^n I_i$ 이다. 행렬 A 의 고유벡터(eigenvector)와 고유값(eigenvalue)을 구하기 위해 SVD (Singular Value Decomposition)를 이용하면 $A = U \Sigma V^T$ 형태로 표현할 수 있다. B 는 A 이후의 프레임에서 관측된 특징들로 만든 $d \times m$ 행렬이다. 그리고 두 행렬 A 와 B 에서의 주요 성분을 분석하기 위해서 $[A \ B] = U \Sigma' V'^T$ 로 나타낼 수 있으며, 주적이 진행될 때 고유벡터 U' 와 고유값 Σ' 를 구하는 것이 최종 목표이다. 하지만 프레임이 증가함에 따라 $[A \ B]$ 의 크기가 점점 커질 것이고, 이는 매우 큰 저장 공간과 많은 계산 시간을 필요로 하기 때문에 SVD를 연속적으로 적용할 수 없다. 따라서 다음과 같이 표현한다.

$$[A \ B] = [U \ \tilde{B}] \begin{bmatrix} \Sigma & U^T B \\ 0 & \tilde{B}^T B \end{bmatrix} \begin{bmatrix} V^T & 0 \\ 0 & I \end{bmatrix} \quad (4)$$

식 (4)에서 $\tilde{B}$ 는 QR decomposition을 통해 구해진 U 와 직교하는 B 의 성분이고, 가운데 행렬을 R 이라 정의하여 식 (5)같이 전개한다.

$$R = \begin{bmatrix} \Sigma & U^T B \\ 0 & \tilde{B}^T B \end{bmatrix} = \tilde{U} \tilde{\Sigma} \tilde{V}^T \quad (5)$$

식 (5)의 행렬은 크기가 $k+m$ 인 정방행렬이며 k 는 Σ 에서 사용된 고유값의 개수이다. 여기에 SVD를 취하여 식 (6)을 얻는다.

$$[A \ B] = ([U \ \tilde{B}] \tilde{U}) \tilde{\Sigma} \left(\tilde{V}^T \begin{bmatrix} V^T & 0 \\ 0 & I \end{bmatrix} \right) \quad (6)$$

구하고자 하는 최종 목표인 U' 과 Σ' 은 각각 식 (6)에서의 $[U \ \tilde{B}] \tilde{U}$, $\tilde{\Sigma}$ 와 같다. 따라서 B 가 입력됨에 따라 갱신된 평균값과 고유벡터가 계산되기 때문에 연속되는 프레임에서 물체의 특징이 계속적으로 업데이트되어 다른 PCA 기반의 방법에 비해 성능이 우수한 것으로 알려져 있다[15]. 또한 native PCA 방법과 비교하였을 때 시간복잡도는 $O(d(k+m)^2)$ 에서 $O(d(k+m))$ 로 감소된다.

III. 영상 기반의 파티클 필터를 이용한 물체 추적

컴퓨터 비전에서 물체를 추적하는 문제는 먼저 대상에 대하여 현재 상태와 이전 상태의 관계를 수학적으로 모델링 해야 한다. 상태변수는 추적하고자 하는 물체의 현재 움직임을 나타낸 것이며 상태변수를 정의하는 것에 따라 다른 시스템이 될 수 있다. 현재 시점 t 의 상태변수를 $\mathbf{x}_t$ 라 했을 때 식 (7)과 같이 나타낼 수 있다.

$$\mathbf{x}_t = [x, y, \dot{x}, \dot{y}, h] \quad (7)$$

위 식에서 x 와 y 는 영상에서의 해당 샘플의 위치를 나타내고 $\dot{x}$ 와 $\dot{y}$ 는 속도를 의미한다. 각 샘플을 사각형 형태로 사용하기 때문에 사각형의 높이를 의미하는 h 를 포함한다. 또한, 본 연구에서 사용된 사람 검출 알고리즘에서 결과가

고정된 사각형 비율(aspect ratio)을 가지고 있기 때문에 사각형의 가로의 길이는 실험적으로 $w=1.1h$ 로 설정하였다.

파티클 필터 역시 재귀적 베이저안 필터(recursive Bayesian filter)의 한 종류이기 때문에 예측 단계와 갱신 단계로 이루어진다. 각 단계는 식 (8)과 (9)로 표현할 수 있다.

$$p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1}) d\mathbf{x}_{t-1} \quad (8)$$

$$p(\mathbf{x}_t | \mathbf{z}_t) = \frac{p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{1:t-1})}{p(\mathbf{z}_t | \mathbf{z}_{1:t-1})} \quad (9)$$

위 식에서 $\mathbf{z}$ 는 측정변수이며 식 (8)에서 이전 상태 $p(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1})$ 와 현재 상태는 오직 이전 상태에만 영향을 받는다는 1차 마르코프 과정(first-order Markov process) $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ 에 의해 $p(\mathbf{x}_t | \mathbf{z}_{1:t-1})$ 을 예측하고, 식 (9)는 예측 값을 기반으로 새로운 $p(\mathbf{z}_t | \mathbf{x}_t)$ 을 관측하여 현재 상태 $p(\mathbf{x}_t | \mathbf{z}_t)$ 를 갱신한다.

식 (10)은 식 (8)을 행렬로써 표현한 것이다. t 시간에의 상태변수를 구하기 위해 이전 시간 $t-1$ 에서의 상태변수에 측정 행렬을 곱하고 잡음을 더하여 만든다.

$$\begin{bmatrix} x_t \\ y_t \\ \dot{x}_t \\ \dot{y}_t \\ h_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ \dot{x}_{t-1} \\ \dot{y}_{t-1} \\ h_{t-1} \end{bmatrix} + \begin{bmatrix} N(0, \sigma_x) \\ N(0, \sigma_y) \\ N(0, \sigma_{\dot{x}}) \\ N(0, \sigma_{\dot{y}}) \\ N(0, \sigma_h) \end{bmatrix} \quad (10)$$

위 식에서 Δt 는 프레임 간의 시간을 나타내나 계산의 편의를 위해 Δt 의 값을 1로 설정하였다. 또한, 잡음 σ 는 각 변수의 잡음을 발생시키기 위한 난수의 분산 값이다.

현재 상태를 예측한 후 우도(likelihood) 계산을 통해 사후 확률을 계산한다. 여기에서는 앞서 작성한 HOG 서술자를 이용하여 우도를 계산하며, 각 샘플의 상태벡터와 학습된 벡터의 유사성을 판별하기 위해 식 (11)과 같이 Bhattacharyya distance를 이용하여 저장된 히스토그램 H_i 와 현재 샘플에 해당하는 히스토그램 H_s 를 비교한다. 이때 두 히스토그램이 완벽하게 일치하면 0, 불일치하면 1이다.

$$d_B(H_i, H_s) = \sqrt{1 - \sum_i \frac{\sqrt{H_i(i)H_s(i)}}{\sum_j H_i(j) \sum_j H_s(j)}} \quad (11)$$

$$\text{where, } H_i = U U^T \bar{I}, \quad H_s = D_u$$

식 (11)에서 얻어진 각 샘플의 신뢰도를 식 (12), (13)과 같이 정규화를 통해 0~1 범위를 갖는 확률 값으로 변환할 수 있다.

$$w(i) = 1 - d_B(H_i, H_s), \quad i = 1, \dots, N \quad (12)$$

$$w_n(i) = \frac{w(i)}{\sum_{j=1}^N w(j)}, \quad i = 1, \dots, N \quad (13)$$

IV. 카메라-레이저스캐너 상호보완 추적

1. 카메라-레이저스캐너 캘리브레이션

레이저스캐너는 레이저를 물체에 반사시키고 되돌아 오는 시간을 계산하여 장애물과의 거리를 측정하는 센서로써 로봇 및 다양한 분야에 많이 사용되고 있다. 하지만 레이저스

그림 2. 체스보드를 이용한 카메라-LRF extrinsic calibration[18].
Fig. 2. Camera-LRF extrinsic calibration using chessboard.

캐너의 설치 위치에 따라 스캐닝(scanning)하는 부분이 다르기 때문에 3차원 실세계를 분석하기에는 제한적이다. 따라서 본 연구에서는 Q. Zhang의 알고리즘[17]을 이용하여 레이저스캐너와 카메라의 기하 관계를 구하고 카메라 영상 내의 레이저스캐너 데이터의 좌표를 계산한다[18]. 그림 2는 레이저스캐너에서 카메라로의 3×3 회전 행렬 Φ 와 3차원 이동 벡터 Δ 를 구하는 과정이다. 체스보드 위의 임의의 레이저 데이터의 좌표를 P_i 이라고 하고, 이 좌표가 카메라 좌표계에서 P_c 로 표현된다. 식 (14), (15)을 이용하여 식 (16)을 얻을 수 있고 이는 영상 내의 픽셀 좌표 p 를 표현한다. K 는 카메라의 내부 파라미터를 나타낸다.

$$P_i = \Phi P_c + \Delta \quad (14)$$

$$P_c = \Phi^{-1}(P_i - \Delta) \quad (15)$$

$$p = KP_c = K\Phi^{-1}(P_i - \Delta) \quad (16)$$

영상 내 픽셀 좌표와 기하학적 제약 조건을 이용하여 Φ 와 Δ 를 구할 수 있다. 그림 2와 같이 N 은 카메라에서 체스보드에 수직으로 닿는 벡터로 아래의 식 (17)로 표현된다.

$$N = -R_3(R_3^T t) \quad (17)$$

식 (17)에서 R_3 는 카메라에서 체스보드로의 회전 행렬 R 의 3번째 열 벡터이고 t 는 이동 벡터이다. 즉, R_3 는 크기가 1인 체스보드의 법선 벡터이고 $R_3^T t$ 는 카메라에서 체스보드까지의 거리를 나타내는 것이기 때문에 방향성을 고려하여 N 을 계산할 수 있다.

그림 2에서의 P_c 와 N 의 내적을 구하면 $\|N\|^2$ 과 같다는 걸 알 수 있다. 따라서 식 (18)을 얻는다.

$$N \cdot \Phi^{-1}(P_i - \Delta) = \|N\|^2 \quad (18)$$

카메라 캘리브레이션의 과정을 통해 R 과 t 가 결정되면 N 을 계산할 수 있다. 결국 식 (18)은 Φ 와 Δ 를 구하기 위한 제약 조건식이다. 이 식을 조금 바꿔보면 레이저 데이터는 레이저스캐너 좌표계 상에서 Y 좌표는 모두 0의 값을 갖는다. 따라서 $\hat{P}_i = [X \ Z \ 1]^T$ 로 표현한다면 식 (18)을 식 (19)와 같이 선형 방정식으로 다시 쓸 수 있다.

$$N \cdot H\hat{P}_i = \|N\|^2 \quad \text{where, } H = \Phi^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 & -\Delta \\ 0 & 1 \end{bmatrix} \quad (19)$$

그림 3. Extrinsic 캘리브레이션 후의 재투영 결과.

Fig. 3. Re-projection to image after extrinsic calibration.

또한, 다양한 각도의 체스보드 영상에서의 N 과 P_i 를 여러 번 구하여 최소 자승법을 통해 H 를 계산할 수 있고, 아래의 유도 식 (20)을 통해 Φ, Δ 를 구할 수 있다.

$$\begin{aligned} H &= [H_1 \ H_2 \ H_3] = [\Phi_1^T \ \Phi_2^T \ -\Phi^T \Delta] \\ \Phi_1^T &= H_1, \quad \Phi_2^T = -H_1 \times H_2, \quad \Phi_3^T = H_2 \\ \Phi &= [H_1 \ -H_1 \times H_2 \ H_2]^T \\ -\Phi^T \Delta &= H_3, \quad \Delta = -\Phi H_3 \end{aligned} \quad (20)$$

로봇과 사람의 동반 주행을 위하여 레이저스캐너는 지면과 수평으로 놓고, 카메라는 레이저스캐너 위에서 오른쪽으로 90도 정도 회전시켜 장착하였다. 레이저 데이터가 카메라 화각에 들어오게 하기 위해 한 변이 100mm인 격자무늬를 이용하여 10×7 의 체스보드를 이용하였다. 본 실험에서 구한 Φ, Δ 는 각각 아래와 같다.

$$\Phi = \begin{bmatrix} -0.0319 & -0.0221 & -0.9992 \\ -0.0004 & 0.9998 & -0.0221 \\ 0.9995 & -0.0003 & -0.0319 \end{bmatrix}$$

$$\Delta = [0.1006 \ 0.8006 \ -0.1717]^T$$

그림 3은 식 (16)을 이용하여 레이저 데이터를 영상에 투영해 본 결과이다. Residual error가 23mm로 비교적 정확하게 기하 관계를 구하였다.

2. 레이저 도메인에서의 Mean-shift 추적

Mean-shift 알고리즘은 데이터 집합의 밀도 분포에서 지역 극값(local extrema)을 찾는 강인한 방법으로 알려져 있다[19]. 이는 연속 분포에서 제대로 적용되며, 이는 데이터의 밀도 히스토그램에 적용된 “언덕 오르기(hill climbing)” 방법이라고 볼 수 있다. 그러나 이산 데이터 집합에서는 커널(kernel)을 이용하여 윈도우 안에 존재하는 데이터 점들만을 사용하여 움직이는 형태로 동작한다. 그림 4는 본 논문에서 적용한 단순화된 Mean-shift 알고리즘의 의사 코드(pseudo-code)이다. 여기에서 n 은 682이고 r 과 ε 는 각각 270mm와 10mm이다. M_{00} 은 영상에서의 0차 모멘트, M_{10} 과 M_{01} 은 각각 x 와 y

그림 4. Mean-shift를 이용한 추적 알고리즘 의사 코드.

Fig. 4. Pseudo-code for mean-shift tracking algorithm.

그림 5. 2D 레이저 좌표계에서의 mean-shift 추적.

Fig. 5. Mean-shift tracking in 2D laser coordinates.

방향의 1차 모멘트이다. 이를 이용하여 무게 중심인 $\mathbf{p}_{curr}^c$ 을 구할 수 있고, 한 위치에 수렴할 때까지 2~12 줄의 과정을 반복하게 된다. 그림 4의 알고리즘을 레이저 좌표계에서 적용하면 그림 5와 같이 이동 물체에 대한 추적이 가능하다. 그림 5의 중앙에 있는 원은 Mean-shift 추적 윈도우를 나타낸다. 여기에서 물체의 움직임이 레이저스캐너의 데이터 입력 속도보다 빠르지 않다고 가정한다. 이는 현재 프레임에서의 물체 위치가 이전 프레임의 윈도우 내에 있다는 것을 의미한다.

3. 카메라-레이저스캐너 간의 적응적 추적 기법

카메라를 이용한 사람 추종 분야 중 대다수의 연구가 카메라를 로봇의 진행방향과 같은 방향으로 향하여 설치하여 추종 대상이 화각을 벗어나는 경우가 적다. 사람이 카메라 화각에서 벗어나기 전에 단순히 로봇의 헤딩(heading)각을 조정하면 추종 대상을 영상 중심으로 위치시킬 수 있기 때문이다. 하지만 본 논문에서 적용한 센서 구성은 카메라를 측면

그림 6. 카메라-레이저스캐너 간의 상호보완 추적 기법.

Fig. 6. Camera-LRF complementary tracking.

에 설치하여 추종 대상이 카메라 화각에서 벗어나는 경우가 종종 발생한다. 또한, 레이저스캐너만을 이용하여 추종 알고리즘을 구현할 수 있으나 추종 대상의 선택과 물체의 가려짐 문제로 어려움이 있다. 따라서 본 논문에서는 카메라와 레이저스캐너의 관계를 도출하여 물체의 위치를 카메라 화각 안팎에서 알 수 있도록 하고, 상대 거리와 각도 정보를 기반으로 추종한다.

앞서 설명한 카메라-레이저스캐너 extrinsic 캘리브레이션을 통해 카메라와 레이저 좌표계 간의 관계를 알 수 있다. 식 (16)을 이용하여 실시간 레이저 데이터를 영상에 투영시킬 수 있다. 레이저스캐너와 카메라는 같은 축에 설치되어 영상과 레이저 데이터 사이의 시차(parallax)가 발생하지 않는다고 가정하며, 식 (14)를 이용하여 영상의 좌표에서도 레이저 좌표를 얻을 수 있다. 이를 이용하면 추적 대상이 카메라 화각에서 벗어난 때와 영상 밖의 추적 대상이 영상 안으로 이동하는 것을 알 수 있고, 그림 6의 흐름과 같이 온라인 학습과 추적을 정지/재개 할 수 있다.

V. 실험결과 및 분석

1. 시스템 구축 및 DB 설정

실험을 위하여 그림 7과 같은 로봇의 프레임에 설계 및 제작하였다. 표 1은 구성된 이동 로봇의 제원을 나타내며, 각 장비의 특징이 기술되어 있다. 카메라는 해상도를 640x480으로 설정한 후에 실험을 위한 데이터를 취득하였다.

그림 7의 실험에 사용된 로봇 시스템은 카메라와 레이저스캐너가 각각 우측과 정면을 향하여 다른 방향으로 설치되어 있기 때문에 영상에서 추종 대상이 벗어나는 경우가 있다. 기존 알고리즘과의 비교를 위하여 추적 대상이 화각 안팎으로 움직이는 경우와 유사 물체가 존재하는 경우에 대한 DB를 취득하였다. 총 8개의 DB를 동일한 시나리오로 획득하였으며 추적 실험을 위한 시나리오는 우측과 같이 작성되었다. 시나리오에서 로봇의 속도는 등속도로 하며 유사 물체는 임의의 색상의 옷을 입고 있는 사람으로 하였다. 조명 조건은 낮과 밤으로, 배경은 벽면과 창가로 정한 후 DB를 취득하였

그림 7. 로봇 프레임 설계 및 제작.

Fig. 7. Design and manufacture of the robot.

다. 또한 기존의 영상기반 추적과 비교를 위해 추적 대상이 화각에서 벗어나는 경우를 추가하였다.

- 1) 로봇의 우측에 사람이 서있고, 그때 영상 내에 사람을 추적 대상으로 한다.
- 2) 로봇과 추적대상이 같은 속도로 움직인다.
- 3) 추적대상은 로봇을 추월하여 화각 밖으로 벗어난 후 정지한다.
- 4) 카메라 화각 내에 유사물체가 등장한다.
- 5) 유사물체가 로봇을 추월하여 화각 밖으로 벗어난다.

표 2. 추적 결과 정성적 분석 (#1).

Table 2. Qualitative analysis of tracking (#1).

알고리즘	Sample frame of DB_02							
CAM-shift [20]								
IPCA PF [9]								
Complementary Tracker								

표 3. 추적 결과 정성적 분석 (#2).

Table 3. Qualitative analysis of tracking (#2).

알고리즘	Sample frame of DB_06							
CAM-shift [20]								
IPCA PF [9]								
Complementary Tracker								

표 1. 로봇의 제원.

Table 1. Specification of the robot.

품목	모델명/제조사	특징
로봇 플랫폼	Tetra-DS III / (주)동부로봇	차륜구동방식, 최대속도: 2m/s
레이저스캐너	URG-04LX-UG01 / Hokuyo	각도 분해능: 1° 감지 범위: 240°
카메라	C920 / Logitech	최대해상도: 1280x720
노트북	SVS13A26PK / Sony	CPU: Intel i7 (2.9GHz) RAM: 8GB

6) 유사물체와 추적대상이 화각 내에 존재하며, 유사물체는 정지해있다.

7) 유사물체가 화각 밖으로 벗어난 후 추적대상도 정지하여 화각 밖으로 벗어났다가 다시 들어온 후 취득을 끝낸다.

2. 추적 실험

표 2와 3은 각각 DB_02, 06에 대한 추적 결과의 정성적 분석표이다. 비교 분석을 위하여 프레임 샘플링을 통해 표현하였다. 표 2, 3의 첫 번째 행은 색상 기반의 추적을 할 때 많이 사용하는 방법인 CAM-Shift [20], 두 번째 행은 형태 특징 및 색상 정보에 기반한 IPCA 및 파티클 필터 물체 추적 방법 (IPCA PF) [9], 마지막 행은 본 논문에서 제안한 방법인 상호 보완 추적기(complementary tracker)이다.

CAM-Shift는 모델이 되는 색상 히스토그램을 구하고 각 픽셀을 확률로 표현한 back-projection 영상을 얻은 후 윈도우

표 4. 추적 결과 정량적 분석 (횟수).

Table 4. Quantitative analysis of tracking.

Data Set		Complementary Tracker				IPCA PF [9]			
DB	frames	TP	FP	TN	FN	TP	FP	TN	FN
01	690	352	1	330	7	264	425	0	1
02	790	428	14	319	29	215	574	0	1
03	826	514	10	268	34	293	532	0	1
04	873	318	3	467	20	209	663	0	1
05	812	378	14	409	11	254	557	0	1
06	788	369	15	384	20	279	508	0	1
07	745	401	23	288	33	242	502	0	1
08	780	475	8	265	32	229	550	0	1

그림 8. 알고리즘의 성능 평가 그래프.

Fig. 8. Performance evaluation graph.

가 확률 밀도가 높은 쪽으로 수렴하는 방법이다. 시간 변화에 따른 모델을 업데이트하여 비교적 적응적인 물체 추적 방법이지만, 채도(saturation)와 색조(hue)가 높은 색상에만 강인한 것으로 알려져 있다.

IPCA PF는 물체에서 추출된 색상 및 형태 특징을 히스토그램화하고 IPCA 실시간 학습 방법을 이용하여 물체의 자세 변화와 외부 조명 변화에 강인하게 모델을 학습한 후 파티클 필터를 적용한 물체 추적 방법이다. 실시간 학습을 이용하기 때문에 배경이나 조명 변화에 비교적 강인하지만, 실시간 학습에서 종종 문제가 되는 drift 현상이 발생하였고 이는 잘못된 모델을 학습하여 잘못된 배경이나 다른 물체를 추적하는 오류를 가져온다.

기존 알고리즘과 제안한 알고리즘의 객관적인 비교 분석을 위하여 위의 DB에서 각각의 알고리즘을 정량적으로 평가하고 이를 표 4와 같이 표현하였다. TP, FP, TN, FN 중 P와 N은 각각 positive와 negative로써 화각 내 물체 존재 여부에 따라 표현된 것이고, T와 F는 각각 true와 false로써 정답 유무에 따른 표현이다. 즉, TP는 정확히 추적이 되는 횟수이고 TN은 물체가 화각 밖에 있어서 추적이 일시정지된 횟수이다. 성능을 비교하기 위하여 식 (21)과 같은 방법으로 정확성을 계산하였다. 또한, 각 DB에 대한 정확성을 그래프로 그림 8과 같이 표현한다.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (21)$$

그림 8에서 기존 알고리즘보다 제안한 방법의 정확성이

표 5. 추종 결과표.

Table 5. Results of person-following.

DB	평균거리(m)	표준편차
01	1.35	0.27
02	1.29	0.63
03	1.32	0.64

그림 9. 복도 환경에서의 추종 궤적(위: 직진, 아래: 회전구간).

Fig. 9. Person-following trajectory in a corridor.

평균적으로 31.6%에서 94.7%로 향상되었다. 표 2와 3에서 보듯이 CAM-Shift와 IPCA는 추적 대상이 화각 밖으로 벗어나는 경우에도 추적이 중지되지 않고 계속적으로 진행되기 때문에 FP 수치가 높다. 따라서 로봇을 위한 물체 추적에는 기존 알고리즘을 적용하기 어렵다는 결과를 보인다. 반면에 본 논문에서 제안하는 상호보완 추적기는 drift 현상이 발생하지 않도록 하기 위하여 레이저스캐너를 이용한다. 카메라와 레이저스캐너의 기하적 관계를 이용하여 물체의 위치를 계산하고 있기 때문에 영상 내 추적 대상의 유무에 관계없이 추적이 가능하다.

3. 추종 실험

본 논문에서는 상호보완 추적기를 이용하여 로봇이 사람과 일정거리를 유지하며 추종 주행을 하게 하였다. 거리센서를 이용한 이동로봇의 추종을 시뮬레이션한 기존 연구[21]를 참고하여 실제로 로봇에 적용하여 보았다. 추종을 하기 위해서는 목표 값 두 개를 입력해야 하는데, 하나는 거리이고 다른 하나는 각도이다. 목표 값을 이용하여 최종적으로 양쪽의 바퀴 속도가 정해진다. 따라서 사람이 빨리 걷게 되면 로봇도 빠르게 추종하게 되며 로봇과 사람 사이의 각도에 따라 로봇이 회전하게 된다.

그림 9는 실내 복도 환경에서의 사람 추종 결과이다. 또한 표 5는 사람과 로봇의 평균거리 및 표준편차를 나타낸다. 세 번의 실험을 통하여 목표거리 1m와 목표각도 90도를 추종하는 것을 확인하였다.

VI. 결론 및 향후 연구

본 논문에서는 측방 카메라와 전방 레이저스캐너가 장착된 로봇을 이용하여 물체 추적 및 추종에 관한 알고리즘을 제안하였다. 카메라가 측방으로 설치되어 있기 때문에 추적

대상이 카메라 화각 밖으로 벗어나기 쉽고 이에 따라 추적이 종종 실패한다. 하지만 카메라와 레이저스캐너의 기하적 관계를 이용하면 화각에서 벗어나는 것을 쉽게 감지할 수 있고, 카메라에서의 추적을 레이저스캐너에서의 추적으로 계속 진행할 수 있다. 영상에서는 IPCA와 파티클 필터를 기반의 추적을 하여 가려짐과 조명변화에 강인하게 추적하였고, 레이저도메인에서는 Mean-shift를 이용하여 상대 각도와 거리를 쉽게 구하고 추적을 할 수 있었다. 낮과 밤, 복도와 창가에서의 실험을 통하여 94.7%의 정확성을 확인하였다.

본 논문에서 제안한 방법은 잠재적으로 충돌의 위험성을 내포하고 있고 정적인 환경에서 단일 물체의 추적에 적합하다. 로봇의 사람 추종은 동적 장애물을 고려하고 있기 때문이다. 따라서 동적 장애물의 이동 벡터를 추정하여 충돌 방지 경로 계획이 필요하다. 이로써 능동적 장애물 회피 주행이 가능하고, 향후 능동적인 장애물 회피를 포함한 추종은 건물 로비나 공항과 같은 곳에서 응용할 수 있을 것으로 전망된다.

REFERENCES

- [1] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: a survey," *ACM Computing Surveys*, vol. 38, no. 4, Dec. 2006.
- [2] Y. Salih and A. S. Malik, "Comparison of stochastic filtering methods for 3D tracking," *Pattern Recognition*, vol. 44, no. 10, pp. 2711-2737, Oct. 2011.
- [3] R. Bianco, M. Caretti, and S. Nolfi, "Developing a robot able to follow a human target in a domestic environment," *Proceedings of the First RoboCare Workshop*, Rome, Italy, 2003.
- [4] O. Gigliotta, M. Caretti, S. Shokur, and S. Nolfi, "Toward a person-follower robot," *Proceedings of the second RoboCare Workshop*, Rome, Italy, 2005.
- [5] C. Schlegel, J. Illmann, K. Jaberg, M. Schuster, and R. Worz, "Vision based person tracking with a mobile robot," *Proceedings of the Ninth British Machine Vision Conference*, pp. 418-427, 1998.
- [6] M. Kobilarov, G. Sukhatme, J. Hyams, and P. Batavia, "People tracking and following with mobile robot using an omnidirectional camera and a laser," *IEEE International Conference on Robotics and Automation*, pp. 557-562, 2006.
- [7] M. Montemerlo, S. Thrun, and W. Whittaker, "Conditional particle filters for simultaneous mobile robot localisation and people-tracking," *IEEE International Conference on Robotics and Automation*, pp. 695-701, 2002.
- [8] S. Shaker, J. J. Saade, and D. Asmar, "Fuzzy inference-based person-following robot," *International Journal of Systems Applications, Engineering and Development*, vol. 2, no. 1, pp. 29-34, 2008.
- [9] H. Lee, X. Cui, H. Kim, S. Ma, J. Lee, and H. Kim, "Robust object tracking in mobile robots using object features and on-line learning based particle filter," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 18, no. 6, pp. 562-570, 2012.
- [10] G. Cielniak, M. Miladinovic, D. Hammarin, L. Goransson, A. Lilienthal, and T. Duckett, "Appearance-based tracking of persons with an omnidirectional vision sensor," *Proceedings of the Fourth IEEE Workshop on omnidirectional vision*, p. 84, 2003.
- [11] G. Paschos, "Perceptually uniform color spaces for color texture analysis: an empirical evaluation," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 932-937, 2001.
- [12] K. Song, J. Kittler, and M. Petrou, "Defect detection in random colour textures," *Image and Vision Computing*, vol. 14, no. 9, pp. 667-683, 1996.
- [13] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
- [14] D. A. Ross, J. W. Lim, R.-S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking," *International Journal of Computer Vision*, vol. 77, no. 1, pp. 125-141, 2007.
- [15] P. Hall, D. Marshall, and R. Martin, "Adding and subtracting eigenspaces with eigenvalue decomposition and singular value decomposition," *Image and Vision Computing*, vol. 20, no. 13, pp. 1009-1016, 2002.
- [16] A. Levy and M. Lindenbaum, "Sequential Karhunen-Loeve basis extraction and its application to images," *IEEE Transaction on Image Processing*, vol. 9, no. 8, pp. 1371-1374, 2000.
- [17] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, pp. 2301-2306, 2004.
- [18] S.-W. Ma, X. Cui, H.-H. Lee, H.-R. Kim, J.-H. Lee, and H. Kim, "Robust elevator door recognition using LRF and camera," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 18, no. 6, pp. 601-607, 2012.
- [19] D. Comnaiciu and P. Meer, "Mean Shift Analysis and Applications," *Proceedings Seventh International Conference on Computer Vision*, pp. 1197-1203, 1999.
- [20] G. R. Bradski and S. Clara, "Computer vision face tracking for use in a perceptual user interface," *Intel Journal*, 1998.
- [21] N.-Y. Ko, D.-J. Seo, and Y. Moon, "A method for real time target following of a mobile robot using heading and distance information," *Korean Institute of Intelligent Systems*, vol. 18, no. 5, pp. 624-631, 2008.

김 형 래

2011년 목원대학교 지능로봇공학과 졸업. 2013년 인하대학교 로봇공학전공 석사. 2013년~현재 인하대학교 로봇공학 전공 박사과정 재학중. 관심분야는 로봇비전, 물체인식 및 추적.

최 학 남

2004년 연변대학교 응용수학(이학사). 2007년 상명대학교 컴퓨터과학과 석사. 2011년 인하대학교 정보과학과 박사. 2012년~현재 인하대학교 정보통신공학부 조교수. 관심 분야는 로봇비전, 의료영상처리, 머신비전, 병렬영상처리.

이 재 홍

2012년 인하대학교 정보통신공학부 졸업. 2012년~현재 인하대학교 대학원 정보통신공학과 석사과정 재학중. 관심분야는 로봇비전, 위치인식.

이 승 준

2013년 인하대학교 정보통신공학부 졸업. 2013년~현재 인하대학교 대학원 정보통신공학과 석사과정 재학중. 관심분야는 로봇비전, 얼굴인식, 머신비전.

김 학 일

1983년 서울대학교 제어계측 공학과 졸업(공학사). 1985년 Purdue Univ. 전기/컴퓨터 공학과 석사. 1990년 Purdue Univ. 전기/컴퓨터 공학과 박사. 1990년~현재 인하대학교 공과대학 교수. 관심분야는 바이오인식, 바이오인식 표준화, 컴퓨터 비전.