

스테레오 비전을 이용한 실시간 인간형 로봇 궤적 추출 및 네비게이션

(Real-time Humanoid Robot Trajectory Estimation and Navigation with Stereo Vision)

박지환[†] 조성호^{**}
(Jihwan Park) (Sungho Jo)

요약 스테레오 카메라를 갖춘 인간형 로봇이 자율적으로 주변 상황을 인지하면서 목적지까지의 경로를 실시간으로 생성 및 수정하는 간단한 알고리즘을 제시한다. 특징점들을 시각적 이미지에서 추출함으로써 주위의 장애물들을 인지한다. 인간형 로봇의 뒤통거리기는 보행 움직임을 모델링함으로써 로봇의 중심부 기준에서의 실제 경로를 유추하여 계획된 경로와 비교함으로써 시각적 피드백 제어를 구현하고 성공적인 네비게이션을 수행한다. 실제 인간형 로봇의 네비게이션 실험을 통해 제안된 알고리즘의 가능성을 입증한다.

키워드 : 스테레오 비전 네비게이션, 휴머노이드 로봇 네비게이션, 칼만필터 기반 자기 위치 표현 알고리즘

Abstract This paper presents algorithms for real-time navigation of a humanoid robot with a stereo vision but no other sensors. Using the algorithms, a robot can recognize its 3D environment by retrieving SIFT features from images, estimate its position through the Kalman filter, and plan its path to reach a destination avoiding obstacles. Our approach focuses on estimating the robot's central walking path trajectory rather than its actual walking motion by using an approximate model. This strategy makes it possible to apply mobile robot localization approaches to humanoid robot localization. Simple collision free path planning and motion control enable the autonomous robot navigation. Experimental results demonstrate the feasibility of our approach.

Key words : Stereo vision navigation, Humanoid robot navigation, Kalman-filter based localization algorithm

1. 서론

로봇의 실시간 경로 생성은 지능형 로봇 공학의 주요

- 본 연구는 지식경제부/한국산업기술진흥원 융복합형로봇전문인력양성사업의 지원으로 수행되었음
- 이 논문은 2009주제 인공지능연구회 워크샵에서 '스테레오 비전을 이용한 인간형 로봇 네비게이션'의 제목으로 발표된 논문을 확장한 것임

† 비 회 원 : 한국과학기술원 전산학과
jhpark83@kaist.ac.kr

** 정 회 원 : 한국과학기술원 전산학과 교수
shjo@kaist.ac.kr

논문접수 : 2010년 2월 2일

심사완료 : 2010년 6월 9일

Copyright©2010 한국정보과학회 : 개인 목적이거나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 소프트웨어 및 응용 제37권 제8호(2010.8)

연구 주제 중에 하나이다. 바퀴 형태의 모바일 로봇을 위한 경로 생성 알고리즘은 많은 연구가 이뤄져 오고 있지만[1], 인간형 로봇을 위한 경로 생성은 쉽지 않은 연구 주제로 고려된다. 인간형 로봇의 보행 움직임은 모바일 로봇들처럼 단순하지 않고 반복적인 뒤통거리과 덜커덕거리음이 나타나고 있기 때문에 분석하기가 용이하지 않기 때문이다. 더욱이, 그 동안의 대부분의 연구는 시뮬레이션 중심으로 이루어져 왔다.

성공적인 네비게이션을 구현하기 위해서는 주변 상황의 인식이 필요하다. 이를 위한 센서들 중에 가격면에서나 실제 장착의 용이성 등을 고려할 때 시각센서가 관심의 대상이 된다. 시각정보는 쉽게 얻을 수 있으면서도 풍부한 정보를 포함하고 있기 때문에 로봇 센서로서 자주 이용된다. 특히 스테레오 비전은 3차원적 깊이 정보를 얻는데 용이하다.

스테레오 비전을 이용한 인간형 로봇 네비게이션을

실증적으로 연구해온 결과들이 있다[2-5]. 하지만 대개의 경우, 복잡한 알고리즘을 이용하고 있다. 본 연구는 스테레오 비전을 갖춘 인간형 로봇의 실시간 네비게이션을 가능케 하는 간단한 알고리즘 제안을 목표로 한다.

휴머노이드 형태의 이족보행 로봇의 경우 걸음걸이 과정에서 넘어지지 않기 위해 무게중심을 움직이는 과정에서의 흔들림이나 발걸음 자체의 진동이 강하기 때문에 이미지에서 회전성분이 강하게 나타난다. 본 논문에서는 이러한 이미지 특성을 고려하여 특징점을 선정하고 추적하여 자기 위치를 결정하고 이족보행 인간형 로봇의 걸음걸이 동작을 분석하여 바퀴형태의 모바일 로봇과 같은 경로생성 알고리즘을 적용할 수 있음을 보이고자 한다.

2. 알고리즘

2.1 실시간 로봇 위치 계산 및 장애물 인식 알고리즘

SIFT(Scale-invariant feature transform) 알고리즘을 이용하여 카메라 이미지에서부터 주요 특징점들을 뽑아낼 수 있을 뿐만 아니라 스테레오 카메라의 양쪽 이미지에서 동일한 특징점들을 매칭을 통해 추출할 수 있다[6]. SIFT 알고리즘은 Laplacian을 근사화한 DoG(Difference of Gaussian) 이미지 피라미드를 사용하기 때문에 크기변화에 강인하며 descriptor의 표현에 있어서 특징점 주변의 Gradient vector의 방향을 사용하기 때문에 회전에도 강하다. 인간형 로봇은 무게 중심을 좌우로 옮겨가며 전진하기 때문에 앞으로 전진하는 움직임은 여러 프레임에 걸쳐 천천히 일어나며 그 과정에서 좌우로 심하게 rolling하는 특성이 있다. 양쪽 이미지 2차 평면 상에서 동일한 점의 위치를 안다면 camera projection model을 활용하여 아래와 같이 그 점의 실제 공간상 3차원 위치 정보를 계산할 수 있다[7].

$$\begin{pmatrix} x & y & z \end{pmatrix}^T = \begin{pmatrix} -\frac{B(u_l - p_x)}{u_l - u_r} - \frac{B(v_l - p_y)}{u_l - u_r} - \frac{Bf}{u_l - u_r} \end{pmatrix}^T$$

B 는 스테레오 카메라의 양 렌즈간 거리를 의미하며, (u_l, v_l) 와 (u_r, v_r) 는 각각 왼쪽과 오른쪽 이미지에서부터의 픽셀 단위 2차원 위치 정보, f 는 렌즈의 초점거리, 그리고 (p_x, p_y) 는 기준 카메라의 principal point를 나타낸다.

카메라 렌즈 관련 파라미터들은 calibration 과정을 통해 실험적으로 구한다. 위의 방정식을 이용하여 각 특징점들의 카메라(로봇) 위치 기준좌표계에서의 상대적 3차원 위치정보를 계산한다. 유효한 특징점들만 선택하기 위해서 연속된 이미지 시퀀스 상에서 추출된 두 점 사이의 거리가 시간에 대해 일정한 경우의 특징점들만 고려한다. 특징점들은 실제 공간상에 고정된 위치만을 표현하도록 한다. 또한, 렌즈의 초점거리와 이미지 해상도

를 고려해서 0.3~2미터 사이의 거리 정보를 보이는 특징점들만 유효한 것으로 고려한다.

시간상에서 연속되는 두 이미지 상의 유효한 동일 특징점의 3차원 위치 정보를 계산함으로써 그 점의 상대적 움직임 궤적이 추출된다. 연속되는 두 이미지상에서 같은 특징점을 찾기 위해 SIFT알고리즘이 역시 사용된다. 각 유효 특징점들의 카메라 위치 기준의 상대적 3차원 움직임은 아래와 같은 관계를 만족시킨다.

$$\bar{x}_i(t) = \mathbf{R}(t)\bar{x}_i(t-1) + \bar{p}(t)$$

$\bar{x}_i(t)$ 는 시간 t 에서의 특징점 i 의 상대적 3차원 위치 정보를 나타내며, $\mathbf{R}(t)$ 는 회전행렬, $\bar{p}(t)$ 는 평행이동을 나타낸다.

위 공식에서 각 시간마다 $\mathbf{R}(t)$ 와 $\bar{p}(t)$ 를 계산함으로써 로봇의 움직임 궤적을 계산할 수 있다. 이는 실제 특징점들은 공간상에 정지되어 있음을 고려하기 때문이다.

$$\operatorname{argmin}_{\mathbf{R}(t), \bar{p}(t)} \frac{1}{n} \sum_{i=1}^n \|\bar{x}_i(t) - \mathbf{R}(t)\bar{x}_i(t-1) - \bar{p}(t)\|^2$$

n 은 시간 t 에 유효한 특징점들의 총 개수를 나타낸다. 위와 같이 나타난 최적화 문제에서의 $\mathbf{R}(t)$ 와 $\bar{p}(t)$ 는 SVD(singular value decomposition) method를 이용하여 풀 수 있다[8].

다음 장에서 설명되는 로봇 위치를 결정하는 과정에서 카메라 기준 좌표계에서의 유효한 특징점들의 3차원 위치정보를 알 수 있다. 시간 t 에서 세계 고정 좌표계에서 카메라 좌표계로의 homogeneous transformation matrix를 H_0^t , 카메라 기준 특징점들의 3차원 위치정보를 P_c 라고 하면 그때마다 세계 좌표계에서의 3차원 점 P_w 으로 다음과 같이 변환할 수 있다.

$$P_w = H_0^t P_c$$

위 변환결과를 매 시간 t 마다 누적하면 특징점들의 집합을 얻을 수 있으며 그 집합으로 장애물을 표현한다.

2.2 로봇 중심 경로 계산 알고리즘

앞서 계산된 로봇의 움직임 궤적은 네 가지 state 정보로 표현하기로 한다(그림 1). 평평한 지면 위에서의 움직임이라고 가정하면 평면 위에 위치 정보 (x_r, z_r) 와 z 축과 진행 방향 사이의 회전각 θ , 그리고 한 보행 스텝에서의 위상을 나타내는 각 ϕ 로 나타낸다. 인간형 로봇의 보행 움직임에는 두 가지 요소가 크게 두드러진다. 그림 1에서 보여주는 것과 같이 진행방향 기준 좌우로의 흔들림(wiggling)과 기준 사인함수 형태의 궤적 관점에서 끌림(dragging) 현상이 존재한다. 먼저 이족보행 로봇은 자신의 무게중심을 앞으로 뺀 발의 반대쪽으로 옮겨서 넘어지지 않도록 한다. 그 성분을 흔들림 성분이라고 정의하면 걸음걸이는 주기성을 가지므로 sine

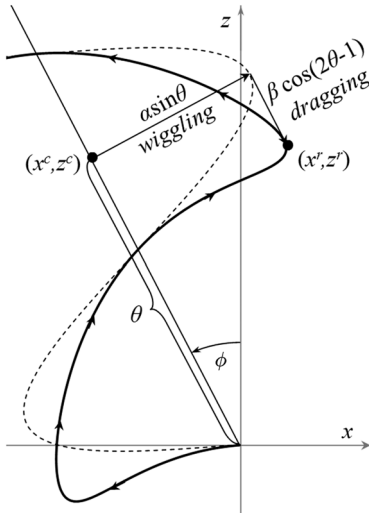


그림 1 인간형 로봇 보행 모델

wave로 표현할 수 있다. 흔들림 표현 계수를 α 라고 하면 로봇 움직임 중심에서 로봇의 머리부분이 벗어난 정도는 $\alpha \sin \theta$ 이다. 또한 앞으로 전진하기 위해 발을 앞으로 뺄 경우 무게 중심을 지지하는 발의 안쪽으로 두기 위해 몸체를 뒤로 당겨야 하는 효과가 있다. 이 성분을 끌림 성분이라고 하면 이 역시 걸음걸이에 의해 영향을 받으므로 끌림 표현 계수를 β 라고 할 때 $\frac{\beta}{2}(\cos 2\theta - 1)$ 라고 표현할 수 있다. 이를 고려하여 로봇 움직임을 모델링할 수 있다. 현재 휴머노이드 로봇의 가상 걸음걸이 중심을 $\bar{x}_r(t) = [x^c(t) \ z^c(t) \ \phi(t) \ \theta(t)]^T$ 로 표현할 수 있고, 휴머노이드 로봇의 선속도를 $v(t)$, 회전 각속도를 $\dot{\phi}(t)$, 발걸음속도를 $\dot{\theta}(t)$, 휴머노이드 로봇의 제어 입력을 $\bar{u}(t) = [v(t) \ \dot{\phi}(t) \ \dot{\theta}(t)]^T$ 라 하자.

$$\begin{aligned} \bar{x}_r(t) &= f(\bar{x}_r(t-1), \bar{u}(t), \bar{w}) \\ &= \begin{bmatrix} x^c(t-1) - Tv(t)\sin\phi(t-1) \\ z^c(t-1) + Tv(t)\cos\phi(t-1) \\ \phi(t-1) + T\dot{\phi}(t) \\ \theta(t-1) + T\dot{\theta}(t) \end{bmatrix} + \bar{w}(t) \end{aligned}$$

$\bar{x}_r(t) = [x^c(t) \ z^c(t) \ \phi(t) \ \theta(t)]^T$ 는 로봇 상태정보이며 \bar{w} 는 가우시안분포를 따르는 프로세스 노이즈를 나타낸다. 한편, 이미징로부터 추출되는 로봇 움직임은 아래와 같이 모델링된다.

$$h(\bar{x}_r(t), \bar{v}) = \begin{bmatrix} x^c(t) \\ z^c(t) \\ \phi(t) \end{bmatrix} + \bar{v}(t)$$

$$= \begin{bmatrix} x^c(t) + \alpha \sin \theta(t) \cos \phi(t) - \frac{1}{2} \beta (\cos 2\theta(t) - 1) \sin \phi(t) \\ z^c(t) + \alpha \sin \theta(t) \sin \phi(t) + \frac{1}{2} \beta (\cos 2\theta(t) - 1) \cos \phi(t) \\ \phi(t) \end{bmatrix} + \bar{v}(t)$$

\bar{v} 는 가우시안 분포를 따르는 센서 노이즈를 나타낸다. 관련 파라미터들 값은 실험적으로 추출한다.

이들 방정식을 Extended Kalman filter에 적용함으로써 로봇의 실제 움직임 경로를 나타내는 (x^c, z^c) 및 ϕ 의 시간에 따른 궤적을 추정할 수 있다[9].

2.3 로봇 보행 제어 및 경로 생성 알고리즘

앞서 로봇의 실시간 경로와 주변 환경을 묘사하는 특징점들의 로봇과 상대적인 공간상 정보를 획득할 수 있다. 평면상의 로봇 움직임만을 가정하기로 한다면, 높이 정보를 크게 고려하지 않고, 공간상의 특징점들의 무리가 로봇 주변의 장애물들을 묘사하는 것으로 이해할 수 있다. 경로 생성 알고리즘을 구현하기 위해 각 특징점을 중심으로 하는 가상의 원들을 생각한다. 각 원의 반지름은 로봇너비의 절반에 안전성을 고려하여 1.2배 한 길이로 결정한다. 가까운 특징점들의 원은 서로 겹치므로 전체적으로 한 덩어리의 장애물로 간주할 수 있다. 이들 정보를 활용하여 목적지점까지의 로봇 경로 생성을 다음과 같이 수행한다. 현재 로봇 위치에서 목적지까지를 직선으로 연결하였을 때 가로막는 장애물이 있다면 장애물의 양쪽 끝에 접선을 긋는다. 두 접선 중에 목적지에 가까운 쪽을 경로로 선택함으로써 그 장애물을 돌아갈 수 있는 경로를 생성한다. 장애물을 지나치면 다시 목적지를 향하는 직선을 경로로 판단한다. 앞서 설명한 과정을 장애물이 인식될 때마다 반복함으로써 실시간으로 경로를 수정하면서 로봇의 보행을 실현할 수 있다.

실시간으로 계산하는 경로와 현재 추정되는 로봇 중심 움직임 경로를 비교함으로써 로봇이 계산된 경로를 잘 따를 수 있도록 하는 시각적 피드백 제어가 가능하다. 로봇의 회전각 θ 를 조정함으로써 로봇이 생성한 경로에서 벗어나지 않고 보행하도록 한다.

3. 실험

그림 2에서 보여주는 것과 같이 로봇의 실시간 정보를 나타내주는 소프트웨어를 개발하였다. 로봇의 위치 정보를 실시간으로 보여주고 생성된 경로와 실제 로봇의 움직임 궤적을 보여주게 된다. SIFT 알고리즘은 많은 양의 특징점 정보를 계산하고 저장해야 하기 때문에 계산 속도를 향상하기 위해 GPU를 이용한다[10]. 제한한 알고리즘을 검증하기 위해 실제 인간형 로봇을 이용하여 실험을 하도록 한다. 우리는 로보티즈에서 개발한 인간형 로봇 키트 바이올로이드의 머리 부분을 스테레오 카메라(STH-MDCS3, Videre design)로 대체하여



그림 2 실험에 사용한 로봇 플랫폼

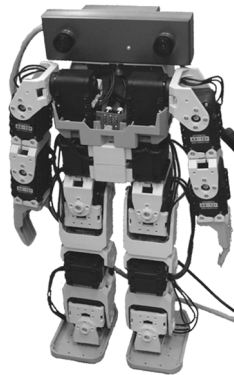


그림 3 실험 환경 구성

로봇 플랫폼을 구성하였다(그림 3). 본 실험은 2.83GHz의 처리속도를 가진 Intel® Core™2 Quad CPU Q9550를 갖춘 PC에서 Windows Vista 운영체제 기반으로 진행된다. 스테레오 카메라는 양쪽 렌즈에서 15 frames/sec으로 해상도 320×240의 이미지들을 받아들인다.

3.1 실험 구성

그림 4는 인간형 로봇 네비게이션을 실험하기 위한 주변 환경의 구성을 보여준다. 로봇이 출발점에서 있고 전방 약 1.5미터 앞에 있는 목표지점까지 보행하게 된다. 주변에 장애물들이 놓여져 있다. 두 가지 환경 시나리오를 실험한다. 목적지까지 도달하기 위해 우측으로 돌아가야 하는 경우와 두개 이상의 장애물이 놓여 있을 때 이들 사이로 진행되는 경우를 실험한다

3.2 실험 결과

그림 5는 두 가지 실험 시나리오(왼쪽: 실험1, 오른쪽: 실험 2)에서 로봇이 움직이는 동안 일정 시간마다의 스냅샷을 보여준다. 각 그림에서 왼쪽에 배치된 스냅샷들

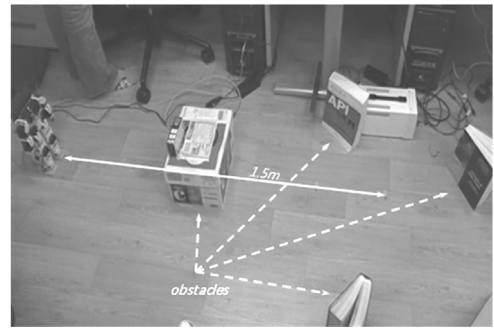


그림 4 실험 환경 구성

은 실제 로봇의 움직임을 보여주고 있으며 오른쪽에 배치된 것들은 실시간으로 시뮬레이션되고 있는 경로와 인지된 장애물들을 나타내고 있다. 빨간색 선은 적용적으로 변경되고 있는 생성 경로이며 하얀색 선은 로봇의 실제 움직임을 나타내고 있다. 시간이 감에 따라 주변 장애물들이 인지되는 것을 볼 수 있고 그에 따라 움직임 경로가 실시간으로 변경됨을 알 수 있다. 이미지들을 받아들여 특징점을 획득하는 과정이 약 0.19초가량 소요되며, 시각정보에서 위치들을 계산하고 움직임 궤적을 판단하는 과정이 약 0.03초 정도 걸린다. 그림 6은 두 실험 동안, 생성 경로와 실제 로봇 움직임의 오차를 보여주고 있다. 오차가 제일 클 때(로봇경로가 갑작스럽게 방향을 바꾸는 순간) 12cm 미만의 오차를 보이며 로봇 너비가 18cm임을 고려하면 오차가 로봇 양다리 사이 간격 정도 안에 있음을 알 수 있다. 실험을 통해 제안된 알고리즘이 실시간으로 충분히 만족스럽게 작동함을 확인하였다.

4. 결론

본 연구는 실시간 스테레오 비전 센서에 기반한 인간형 로봇의 자기 위치 결정 및 경로 계획 알고리즘을 제시했다. 사전에 환경에 대한 정보없이 로봇은 스스로 일련의 알고리즘의 결과로 스스로 환경을 인지하여 경로를 계획하고 목적지까지 무사히 도달함을 실제로 확인할 수 있었다. 기존의 유사 연구 결과에 비추어 보았을 때 본 논문의 가장 큰 목적은 다음과 같다. 첫 번째로 본 연구는 인간형 로봇의 자기위치 결정 및 자기위치 표현에 대한 간단한 방법을 제시한다. 본 논문에서 제시한 표현방법을 통해 인간형 로봇의 경로계획에 있어서도 기존의 바퀴형태의 모바일 로봇에 적용되어 온 간단한 알고리즘을 적용할 수 있다. 샘플링이나 복잡한 추론 알고리즘을 사용하지 않고 실시간 구현이 용이한 Extended Kalman filter를 이용하면서 인간형 로봇의 뒤통거림을 묘사하는 모델을 통해 이를 가능하게 하였다.

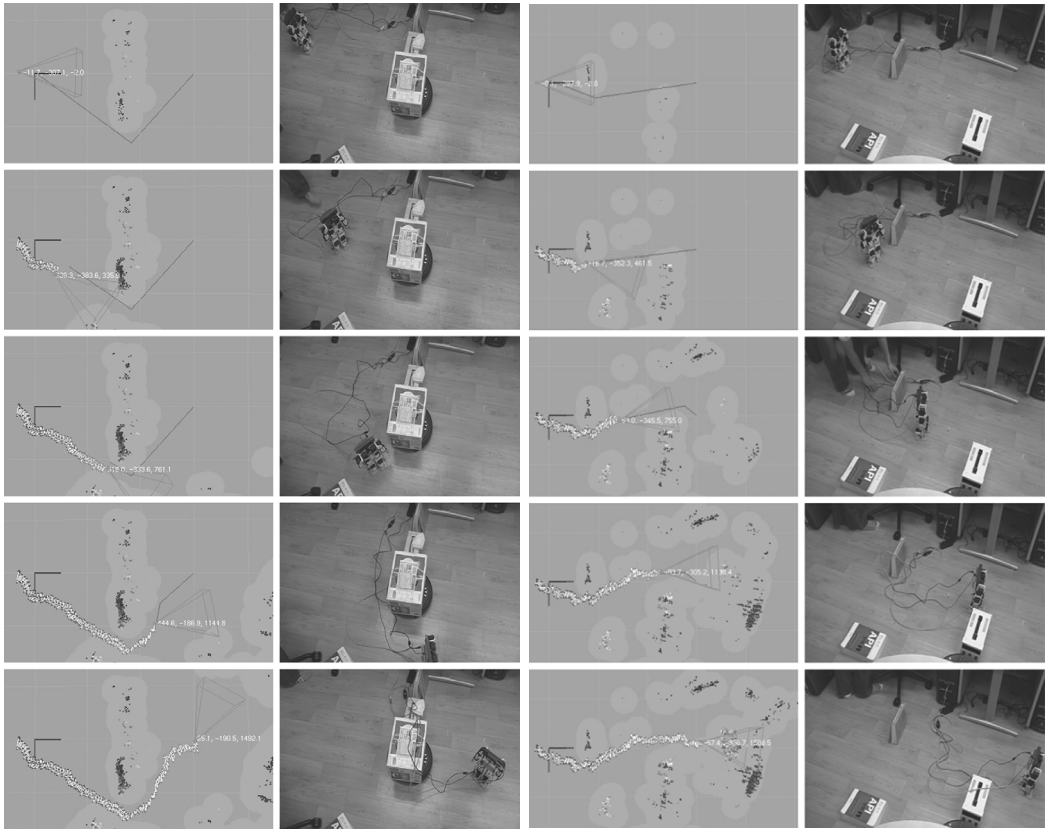


그림 5 로봇 네비게이션 실험 결과

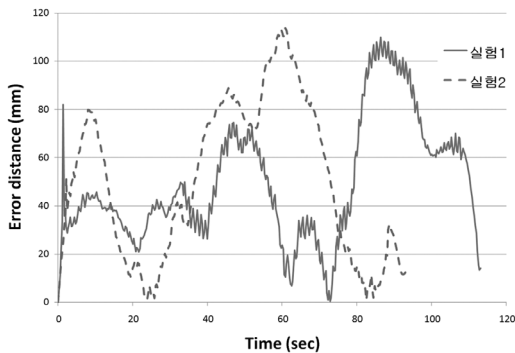


그림 6 계획 및 실제 경로 사이의 오차 궤적

특히 다른 센서 정보는 없이 스테레오 비전을 이용하여 충분히 로봇의 실시간 자기 위치 결정 알고리즘이 문제 없이 동작함을 확인하였다. 두 번째로 본 논문은 특징점 기반의 자기 위치결정 알고리즘을 통해 실시간으로 인간형 이족보행 로봇 주변의 환경정보를 인식해내는 동시에 경로를 생성해내고 피드백 컨트롤을 할 수 있음을 보인 것이다. Takaoka 연구팀[5]이 KLT 추적 알고리

즘을 사용한 특징점 기반 visual odometry system을 이용해 보인 적이 있으나 계산량이 복잡한 ICP(Iterative Closest Point) 알고리즘에 의존함으로 본 논문의 접근 방법이 실시간에 있어 더 유리하다. 실험결과는 이 접근방법이 가능함을 보였다.

본 알고리즘은 인간형 이족보행 로봇이라면 로봇 자체의 정보를 이용하는 것은 없기에 그 형태나 크기를 가리지 않고 쉽게 적용 가능하다. 추후로 움직이는 물체에 대해 더 정교한 알고리즘을 통해 움직임을 피해가고, 사전 정보가 없는 목적지에 대해 더 효율적으로 환경을 탐색하는 알고리즘에 대해 연구해야 할 것이다.

참고 문헌

[1] G.N. DeSouza and A.C. Kak, "Vision for mobile robot navigation: a survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.24, no.2, pp.237-267, Feb. 2002.
 [2] K. Sabe, M. Fukuchi, J.-S. Gutmann, T. Ohashi, K. Kawamoto, and T. Yoshigahara, "Obstacle avoidance and path planning for humanoid robots

- using stereo vision," in *Int. Conf. on Robotics and Automation*, New Orleans, USA, vol.1, pp.592-597, 2004.
- [3] J. -S. Gutmann, M. Fukuchi, and M. Fujita. "Real-time path planning for humanoid robot navigation," in *Proc. of Int. Joint Conf. on Artificial Intelligence*, Edinburgh, Scotland, pp.1232-1237, 2005.
- [4] S. Thompson and S. Kagami, "Humanoid robot localization using stereo vision," in *Proc. IEEE Int. Conf. on Humanoid Robots*, Tsukuba, Japan, pp.19-25, 2005.
- [5] Y. Takaoka, Y. Kida, S. Kagami, H. Mizoguchi, and T. Kanade, "3D map building for a humanoid robot by using visual odometry," in *Proc. of the IEEE Int. Conf. on Systems, Man & Cybernetics*, Paris, France, pp.4444-4449, 2004.
- [6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol.60, pp.91-110, 2004.
- [7] N. Sünderhauf and P. Protzel, "Stereo odometry - a review of approaches," *Technical report*, Chemnitz University of Technology, 2007.
- [8] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. on Pattern Analysis Machine Intelligence*, vol.PAMI-13, no.4, pp.376-380, 1991.
- [9] G. Welch, and G. Bishop, "An Introduction to the Kalman Filter," *Dept. of Computer Science, Univ. of North Carolina at Chapel Hill*. TR-95-041, 1995.
- [10] C. Wu, SiftGPU: A GPU Implementation of David Lowe's Scale Invariant Feature Transform (SIFT). Available: <http://cs.unc.edu/~ccwu/siftgpu/>.



박지환

2008년 서강대학교 컴퓨터공학부 졸업(학사)
2010년 한국과학기술원 전산학과 졸업(석사)



조성호

1999년 서울대학교 기계공학부 졸업(학사)
2001년 MIT 기계공학과 졸업(석사). 2006
년 MIT 전자전산학과 졸업(박사). 2006
년~2007년 MIT 미디어랩 박사후 연구
원. 2008년~현재 한국과학기술원 전산학
과 조교수. 관심분야는 지능형 로보틱스