

스테레오 비전을 이용한 이동로봇의 실시간 지도 작성을 위한 연구

성용원, 김태민, 이민기, 윤태성
국립창원대학교 전기전자제어공학과

A Study on the Real-Time Map Building of Mobile Robot Using Stereo Camera

Yong-Won Sung, Tae-Min Kim, Min-Ki Lee, Tae-Sung Yoon
Electrical, Electronics and Control Engineering, ChangWon National Univ.

Abstract - In this paper, we studied on the real-time environment map building for the mobile robot navigation using the stereo camera system. Distance measurement are necessary to build the environment map. We used a area-based stereo matching for the distance measurement with the stereo camera system. To reduce the computation time, we used DSP processor on the vision board, took a suitable area size for stereo matching, and used hierarchical search method. Using the fast acquired distance values, the environment map was built.

1. 서 론

환경정보를 인식하는데 필요한 센서는 일반적으로 초음파 센서, 레이저 센서, 적외선 센서, 자이로 센서, 비전 센서 등이 활용된다. 초음파센서는 센서의 활용범위가 넓고 가격이 싼 편이나 인식범위가 한정되어 있고 환경에 대한 영향을 많이 받는다. 적외선 센서나 자이로 센서도 초음파와 같이 비슷하게 인식범위가 한정되는 센서이다. 레이저 센서는 비교적 정확하고 활용범위가 넓은 반면 가격이 정밀도에 비례하여 비싸지는 편이 있어서 쉽게 적용하기는 힘들다. 비전 센서는 활용범위가 넓고 적용하는 분야에 따라 프로 그램 적인 작업으로 쉽게 다양한 응용이 가능하다. 즉, 비전 센서의 특징은 하드웨어에 종속적인 것이 아니라 얻어진 데이터를 어떻게 처리하는 가 하는 것을 설정하는 소프트웨어에 의해 그 기능이 결정이 되므로 다른 기능의 추가나 변경이 용이한 면이 있다. 일반적으로 컴퓨터 비전에서 스테레오 기법에 의한 거리 측정은 삼각 측량법의 원리에 근거를 두고 있고 삼각측량의 원리(triangulation)를 이용하여 근접한 위치에 있는 좌, 우 두 개의 카메라로부터 얻은 영상에서 3차원 상의 좌표를 구할 수 있으나, 이때 왼쪽 영상의 특정한 점이 오른쪽 영상의 어느 점에 대응되는지를 결정해야 되는 문제점(correspondence problem)이 발생한다.[1][2]

이러한 대응점을 찾는 방법에는 일반적으로 특징기반과 영역기반의 스테레오매칭이 쓰여진다. 첫 번째로 특징기반의 매칭방법은 밝기 정보의 이미지로부터 매칭 기본단위 집합을 추출하고, 만족할 만한 기본 단위를 가지고 매칭을 수행하는 방법이다. 양안차(disparity)를 나타내는 결과가 이산 적으로 나타나기 때문에 보간이 필수이고 기본적인 이미지 처리에 따라 특징이 상반되게 나타나는 특징이 있다. 두 번째로 일반적인 영역기반의 스테레오 매칭은 매칭 기본단위표면 전체를 매칭하는 방법으로 밝기 이미지 자체를 기본단위로 한다.

이러한 스테레오 매칭으로 얻어진 거리정보를 이용하여 이동 로봇의 환경지도를 작성하는 방법은 일반적으로 격자맵이나 확률지도에 의한다. 격자맵은 거리단위를 일정한 간격으로 표시를 하여 일정크기의 위치에 물체의 존재유무를 판별하고 확률지도는 이러한 존재유무가 애매 모호한 경우 이러한 단점을 확률적인 분포로 나타내게 된다.[4]

본 논문은 스테레오 비전을 이용하여 이동로봇의 실시간

간 지도작성을 하기 위한 방법을 제시한다. 영역기반 스테레오 매칭을 이용하여 거리를 측정하며 이에 요구되어지는 계산 시간 및 양을 DSP가 내장된 영상처리 전용 영상보드의 활용과 영역기반에 사용되는 모델의 크기설정과 탐색영역의 크기설정, 로봇의 환경을 고려한 탐색방법 등에 대한 연구를 통하여 감소시킴으로써 이동로봇의 실시간 지도작성을 수행하는 방법에 관하여 논의하고자 한다.

2. 본 론

2.1 영역기반 스테레오 매칭[3]

영역 기반 스테레오 매칭은 왼쪽 및 오른쪽 영상의 영역간의 유사성에 근거하여 대응점을 찾는다. 일반적으로 왼쪽 영상의 작은 윈도우(그림 1의 모델)내의 화소들의 특성을 조사하여 유사한 특성을 갖는 윈도우를 오른쪽 영상의 탐색 영역에서 찾아, 각 윈도우의 중심점을 대응점으로 한다.

한쪽 영상에 있는 특징 점에 대응하는 정합 점(matching point)을 다른 쪽 영상에서 찾기 위하여 영역간의 유사성을 확인한다. 각 특징 점을 중심으로 하는 작은 영역을 모델(model)이라 하고, 두 점간의 상관성을 조사할 때 모델의 중심을 각 점에 위치시킨다. 유사성 측정을 위하여 자주 사용되는 척도로써 좌측 모델과 우측 영역에 있는 이미지 사이의 명암 상관함수(correlation)값이 사용된다. 윈도우는 직사각형이고, $M \times N$ 화소로 구성되며, 좌·우 영상의 i 번째 행과 j 번째 열 화소의 명암 값을 각각 M_{ij} 와 I_{ij} 라 한다면 상관함수 값 γ 는 다음 식(1)과 같이 정의된다.

$$\gamma = \frac{\sum_{i=1}^M \sum_{j=1}^N M_{ij} I_{ij}}{\sum_{i=1}^M \sum_{j=1}^N M_{ij}^2} \quad (1)$$

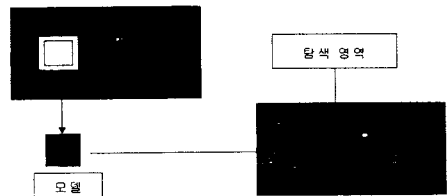


그림 1. 모델과 탐색영역

그러나, 일반적인 스테레오 카메라에서 결과 값은 이미지의 전체적인 밝기 값이 커지면 증가한다. 이미지가 흰색에 가까워지면 상관함수 값이 최고에 도달하고 모델을 찾기가 힘들어진다. 이 문제에 대한 해법으로써 식(2)와 같은 정규화된 상관함수를 이용한다.

$$\gamma = \frac{N \sum I M - (\sum I)(\sum M)}{\sqrt{[N \sum I^2 - (\sum I)^2][N \sum M^2 - (\sum M)^2]}} \quad (2)$$

식(2)의 결과 값은 모델과 이미지가 정확하게 일치할 때 최대 값인 1이 되며, 모델과 이미지가 상관관계가 없을 때 0이 된다.

2.2 모델의 크기 설정

유사성의 척도를 이용하여 대응점을 발견하는 문제를 생각하자. 주어진 좌측 영상의 특징 점에 대응되는 우측 영상에서의 점은 만약 이러한 점이 존재한다면 반드시 에피폴라 선(epipolar line)상에 위치하여야 한다. 대응점을 찾는 기본 전략은 모든 후보 점에 대하여 유사성을 계산하여 유사성이 최대가 되는 점을 선택하는 방법이다. 대응점 발견을 위해서는 여러 가지 문제점들을 해결하여야 하는데 바로 모델 크기 결정 및 대응점 존재유무를 확인하는 작업이다.

일반적으로 큰 원도우는 개략적인 정보를 얻기에 적당하다. 그림2에서 보는 것처럼 좌측 영상의 특징 점 L은 에피폴라 선상에 있는 후보 점들과 정합이 시도된다. 유사성의 척도로서 식(2)에 의하여 정의된 γ 를 사용하여 에피폴라 선상에 위치한 임의의 점 P에서의 γ 의 값을 $\gamma(P)$ 라 하자. 모델의 크기가 작으면 $\gamma(P)$ 는 대응점 P_0 에서 날카로운 정상을 가진다. 좁은 정상은 대응점의 정확한 위치 확인을 가능하게 하지만, 반면에 잡음에 민감하고 깊이가 깊지 않아 후보 점들 간의 $\gamma(P)$ 값의 차이가 크지 않을 수 있으며 여러 개의 정상이 발생할 수 있다. 한편, 그림 2에서 보는 바와 같이 모델의 크기가 클 때는 $\gamma(P)$ 가 완만하게 변한다. P_0 점 근방에서 넓은 정상을 가지므로 대응점이 정확히 결정되지 않는다. 또 다른 단점은 $\gamma(P)$ 를 얻기 위한 계산 시간이 더 많이 요구된다. 그러나 잡음에 덜 민감하게 된다.

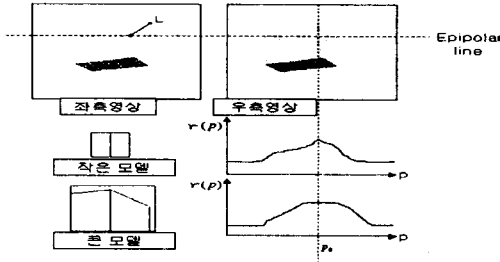


그림 2. 모델 크기의 효과

2.3 영역 탐색 방법

선택된 영역은 대부분 영상보드에서는 640*480 픽셀(pixel)의 해상도를 갖는다. 이러한 영상은 고유한 정보를 가지고 있지만 그것을 모두 활용하는 것은 어렵고 그중 관심 영역이나 특징들을 이용하게 된다. 본 연구에서는 실시간이라는 조건에 맞게끔 계산량을 줄일 목적으로 관심 영역을 640*100 픽셀로 줄였다. 우리가 탐색하고자 하는 영역은 국부적으로 존재하게 된다. 물론 임의의 물체나 사람이 있는 경우 다른 관심 영역이 필요하기도 하지만 실제 실내 공간이나 복도와 같은 단조로운 공간에서의 탐색 영역이 상하로 크게 증가할 이유는 거의 없기 때문에 이동하는 로봇의 높이에 비례하여 관심 영역의 폭이 설정된다.

계산량을 줄이는 효과적인 방법은 계층적인 탐색방법(hierarchical search)을 사용하는 것이다. 처음에는 이미지와 모델 두 가지 모두 작은 크기 및 저해상도로 만들고 이로부터 탐색을 시작한다. 저해상도에서 탐색한 위치를 참조하여 더 고해상도에서는 그 위치를 감안한 허용범위 안에서만 탐색을 하기 때문에 계산시간을 효과적으로 줄일 수 있다.

2.4 좌표변환

이동로봇에서 카메라에서 획득한 거리정보는 카메라 중심 좌표계에서 얻어진 정보이기 때문에 실제적으로 이동로봇을 이동시키기 위한 위치값은 이동로봇중심의 좌표계로 변환하여야 한다. 좌표 변환 기본 행렬식은 평행이동,

회전이동으로 구분된다. 이러한 좌표변환 식을 이용하여 카메라 좌표계의 거리 정보를 이동로봇의 좌표계로 바꾸어 지도를 작성할 때 사용한다.

1) 평행이동

3차원 공간에서는 평행이동은 각 축에 하나씩 모두 세 개의 병진 성분이 있다. 병진을 나타내는 벡터 $p_x + p_y + p_z$ 에 해당하는 일반 변환 행렬은 다음과 같다.

$$Trans(p_x, p_y, p_z) = \begin{bmatrix} 1 & 0 & 0 & p_x \\ 0 & 1 & 0 & p_y \\ 0 & 0 & 1 & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

2) 회전이동

회전의 경우는 세 개의 좌표축 중에 하나를 중심으로 한 회전이 가능하므로 더욱 복잡하다. 메커니즘을 이루는 링크에 적당한 좌표계를 설정하고, 일반 벡터에 대한 회전을 x, y, z 좌표축에 대한 순차적인 회전으로 분해함으로써 일반 축에 대한 회전을 대체할 수 있다.

$$Rot(x, \phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi & 0 \\ 0 & \sin \phi & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$Rot(y, \phi) = \begin{bmatrix} \cos \phi & 0 & \sin \phi & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \phi & 0 & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$Rot(z, \phi) = \begin{bmatrix} \cos \phi & -\sin \phi & 0 & 0 \\ \sin \phi & \cos \phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

3) 좌표변환

카메라 좌표 상에 표현된 장애물의 좌표를 이동로봇 중심의 좌표계로 표현해야 한다. 식 (3)-(6)을 이용하여 다음과 같은 실제 변환에 적용된 식 (7)을 얻을 수 있다. 한쪽 벽의 영상을 많이 받기 위해 카메라를 오른쪽으로 15도 각도로 회전 변환을 주었으므로 식 (7)에서와 같이 회전변환을 적용하였다.

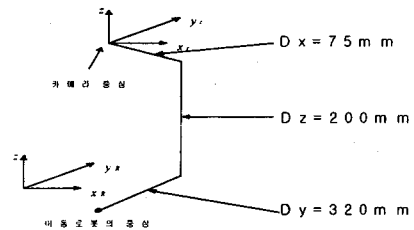


그림 3. 카메라좌표계와 로봇좌표계의 관계

카메라 좌표상의 장애물 위치 : (x_c, y_c, z_c)

이동로봇을 중심으로 한 좌표계 : (x_R, y_R, z_R)

$$\begin{bmatrix} x_R \\ y_R \\ z_R \\ 1 \end{bmatrix} = Trans(Dx, -Dy, -Dz) \cdot Rot(z, 15^\circ) \cdot \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \quad (7)$$

2.5 지도 작성[5]

점유 격자는 확률 예측값을 저장하고 있는 각 셀의 모자이크로 2D 또는 3D로 공간을 표현한다. 점유 격자에서 각 상태의 변수 s(C)는 'occupied'인 OCC 또는 'empty'인 EMP의 2개 상태인 임의의 이산 변수로서 정의된다.

$$P[s(C) = OCC] + P[s(C) = EMP] = 1 \quad (8)$$

