

논문 2008-45SP-3-1

스테레오 깊이 정보를 이용한 점유맵 구축 기법과 응용

(A Technique for Building Occupancy Maps Using Stereo Depth Information and Its Application)

김 낙 현*, 오 세 준**

(Nak Hyun Kim and Se Jun Oh)

요 약

점유맵은 3차원 공간상에서 장애물이 놓인 부분과 빈 공간을 구분해서 2차원 평면상에 표현하는 방식으로 자율 내비게이션이나 물체 인식 등을 위해 사용된다. 본 논문에서는 스테레오 영상에서 추출된 깊이 정보를 활용하여 3차원 공간의 점유맵을 구축하고 그 정보를 물체 영역 추출에 활용하는 기법을 제안한다. 스테레오 깊이 영상에서 기반 평면을 추출한 다음, 각 정합 점들을 기반 평면 중심 좌표계로 투사하여 점유맵을 추출한다. 본 연구에서는 이렇게 추출된 점유맵을 활용하여 실내외의 다양한 환경에서 움직이는 물체 영역을 추출하였는데, 실제 실험 영상을 통해 제안된 방식의 유용성을 확인한다

Abstract

An occupancy map is a representation methodology describing the region occupied by objects in 3D space, which can be utilized for autonomous navigation and object recognition. In this paper, we describe a technique for building an occupancy map using depth data extracted from stereo images. In addition, some techniques are proposed for utilizing the occupancy map for the segmentation of object regions. After the geometric information on the ground plane is extracted from a disparity image, the occupancy map is constructed by projecting each matched point to the ground plane-based 3D space. We explain techniques for extracting moving object regions using the occupancy map and present experimental results using real stereo images.

Keywords: Stereo vision, disparity image, occupancy map, segmentation

I. 서 론

3차원 공간 정보 추출은 자율 이동 로봇이나 지능형 자동차 등에서의 자동 내비게이션 기능 구현을 위해 중요한 작업이다. 3D 정보 추출을 위해 레이저나 초음파 등을 이용하는 능동 방식이나 카메라만을 이용하는 수동 방식 등의 다양한 방법이 사용되고 있는데, 각 방식은 각기 고유한 장점들을 가지고 있으며, 각 방식에서의 기술 수준도 계속 향상되고 있다.

이와 같은 센서 기술에 의해 추출된 3차원 정보를 표

현하는 방식 중 점유맵(occupancy map) 기법이 널리 사용되고 있다. 이 방식에서는 공간 정보를 일정한 크기의 셀로 이루어진 2차원 또는 3차원 배열로 표시하고, 물체에 의해 점유되어 있는 셀과 비어있는 셀을 구분함으로써 장애물 위치와 자유공간(free space)을 알 수 있게 한다^[1~3]. 레이저나 초음파를 이용하여 3차원 공간 정보를 추출하는 경우에는 지면(ground plane)으로부터 일정 높이의 위치에 설치되는 스캐닝 센서를 이용하여 전방의 각 물체까지의 거리를 측정한다. 센서를 좌우 방향으로 스캐닝하여 지면상의 일정한 높이에서의 거리정보를 연속적으로 얻게 되고, 이 정보를 그리드로 나뉘어진 2차원 공간상의 점들(point clouds)로 표현하여 점유맵을 구축하게 된다^[2]. 이렇게 구축된 점유맵은 내비게이션시 자유공간을 추출하거나 주행체 자신의 위

* 정회원, ** 학생회원, 한국외국어대학교
디지털정보공학전공
(Dept. of Digital Information Eng., Hankuk
University of Foreign Studies)
접수일자: 2007년10월9일, 수정완료일: 2008년4월14일

치를 파악하기 위한 수단으로 사용될 수 있다. 또한 여러 위치에서 추출된 점유맵을 융합하여 보다 넓은 영역에 대한 점유맵을 구축할 수 있다.

점유맵을 구축하기 위해 거리 센서 대신 스테레오 비전을 이용하는 방식은 여러 연구자들을 통해 이루어져 왔다^[3~4]. 센서 기술 측면에서 본다면 스테레오 방식은 깊이 측 상에서는 상대적으로 낮은 해상도를 보이지만, 영상 평면상에서는 높은 해상도를 가지는 장점이 있다. 이에 따라 스테레오 방식에 의해 구축된 점유맵은 자율주행에 필요한 자유 공간에 관한 기본적인 정보 이외에 각 장애물의 공간상에서의 크기 정보를 제공할 수 있다. 또한 지면에서 일정한 높이에서의 점유 정보를 추출하는 능동적 방식의 거리 센서의 경우와 달리, 스테레오 영상에서 취득된 깊이 정보는 3차원 공간상에서 골고루 분포한다. 따라서 점유맵 상의 각 점에 대해 지면으로 부터의 높이 정보도 추출할 수 있다. 이와 같은 정보는 공간상의 물체 인식이나 움직임 영역 검출에도 활용될 수 있다^[5].

본 논문에서는 스테레오 카메라를 이용하여 취득된 3차원 정보로부터 공간 점유맵을 구축하는 방법을 제시한다. 본 논문에서 고려하는 환경은 주행체가 기반 평면(또는 지면; ground plane)상에서 움직이며, 이 평면이 입력 영상에서 일정한 면적을 차지하는 경우인데, 이러한 조건은 실내외의 이동 주행 환경에서 대부분 만족된다. 제안된 방법은 영역기반 스테레오 정합을 통해 얻은 깊이 영상에서 기반 평면을 추출한 다음, 정합점들을 3차원 좌표값으로 변환하고, 기반 평면상의 점유 그리드(occupancy grid)상으로 투사함으로써 점유맵을 얻는 것이다.

스테레오 카메라를 이용하여 추출된 3차원 정보를 자율주행에 사용하는 방식에 대해 그동안 꾸준한 연구가 이루어져왔다. 본 논문에서는 제시하는 점유맵 구축 기법에서는 고밀도의 영역기반 스테레오 정합 결과를 사용한다. 기존의 연구들과 달리 본 연구에서는 변위 영상에서 기반 평면을 추출하는 과정을 자동화하였고, 카메라와 기반 평면 중심의 두 3D 좌표계간의 변환과정을 명확히 하여 점유맵상의 각 셀에서의 높이 정보가 활용될 수 있게 하였다.

본 논문에서는 또한 구축된 점유맵 정보의 응용 방법으로 물체 영역의 검출 기법을 제시한다. 본 연구에서는 특히 도로 영상에서 자동차 영역을 추출하는 실험을 수행하였는데, 점유맵 상에서 독립적인 물체 영역을 찾은 다음, 각 영역의 높이 정보를 활용하여 자동차 영역

을 탐지한다. 이 방식은 도로와 주차장 영상에서 효과적으로 자동차 영역을 탐지할 수 있게 하였는데, 본 논문에서는 실제 영상을 이용한 실험 결과를 제시한다.

지능형 자동차의 자율주행을 위한 장애물 검출을 위해 스테레오 비전을 활용하는 연구도 최근 여러 연구자에 의해 진행되었는데, 실시간 구현에 보다 유리한 에지 기반 방식으로 에지의 대칭성을 활용하는 방식이 제안된 바 있다^[6]. 장애물 검출에 보다 효과적인 방식으로는 변위(disparity) 영상의 각 수평주사선상에서 변위값들의 히스토그램을 나타내는 영상인 V-변위(V-disparity) 영상에서 직선 성분을 검출하여 도로와 장애물을 검출하는 방식이 제안되었다^[7~8]. 이 방식은 실시간 구현이 가능하고 실제의 자율주행체에서 활용된 바 있다. 또한 V-변위 영상에서 차량을 검출하는 방식이 제안되기도 했다^[8~9]. 본 논문에서 제안된 방식은 점유맵을 활용하여 물체 영역을 검출한다는 점에서 위와 같은 장애물 검출 방식과 차이를 보인다.

본 논문의 구성은 다음과 같다. 우선 II장에서는 본 논문에서의 주행 환경을 설명하고 카메라 좌표계와 기반 평면 중심 좌표계간의 관계를 유도한다. III장에서는 스테레오 정합 결과를 이용하여 기반 평면을 추출하는 방식을 설명하고, 각 정합점들의 3D 좌표를 복원하는 과정을 설명한다. 이후 IV장에서 변위 영상에서 점유맵을 구축하는 과정과 점유맵 상에서 물체 영역을 검출하는 방법을 제시한다. V장에서는 실제 영상에서 점유맵을 구축하고 물체를 추출하는 실험 결과가 제시되는데, 실내와 도로, 그리고 주차장 등 다양한 환경에서 물체 추출 결과를 설명한다.

II. 카메라 구조와 주행 환경

본 연구에서는 스테레오 카메라가 평행 광학축 구조로 배열되어 있다고 가정한다. 좌측 카메라의 3차원 좌표계 F_C 와 영상 좌표계가 각각 $(x, y, z)^T$ 와 (u_l, v_l) 로 주어진다 하므로 두 좌표축간의 관계는 다음과 같다.

$$\begin{pmatrix} u_l \\ v_l \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1)$$

여기서 f 는 초점거리를 나타낸다. 우측 카메라의 경우

$$\begin{pmatrix} u_r \\ v_r \end{pmatrix} = \frac{f}{z} \begin{pmatrix} x - b \\ y \end{pmatrix} \quad (2)$$

로 주어지는데, b 는 baseline 이다. 좌우 영상의 수평좌

표값 간의 차이인 변위는

$$d = u_l - u_r = f \frac{b}{z} \quad (3)$$

로 주어진다. 따라서 변위값 d 를 알면

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{b}{d} \begin{pmatrix} u_l \\ v_l \\ f \end{pmatrix} \quad (4)$$

에 의해 3차원 좌표를 계산할 수 있다.

자율 이동 로봇이나 자동차 등의 주행환경에서는 그림 1에서와 같이 일반적으로 지상 위 일정한 높이에 설치된 카메라가 지면을 내려다보는 형태의 영상 취득 구조를 가지게 된다. 지면상의 한 점을 원점으로 하고, 지면의 수직 방향을 y' 축으로 하는 지면 중심 좌표계 F_G 를 $x' = (x', y', z')^T$ 이라 하고 F_G 와 좌측 카메라 좌표계 $x = (x, y, z)^T$ 와의 관계가

$$x = Rx' + t \quad (5)$$

로 주어진다고 하자. 카메라 좌표계 F_C 로 표현된 점들을 지면 좌표계 F_G 로 표현하려면 두 좌표계간의 변환을 구해야 한다. 우선 회전 행렬 R 을 구하기 위해 지면의 법선 벡터가 F_C 에서 n 으로 주어진다고 하자. 회전 행렬을 결정하려면 추가적인 제한 조건이 필요한데, 알기 쉬운 방법은 법선 n 이 y' 축에 대응되도록 하는 것이다. 또한 새로운 수평축 x' 을 이전 z 축과 y' 에 수직이 되게 만들면, 다음과 같은 회전 행렬을 유도할 수 있다.

$$R = (v \ n \ v \times n) \quad (6)$$

여기서 $v = \frac{n \times u_z}{\|n \times u_z\|}$ 이고, $u_z = (0, 0, 1)^T$ 이다.

F_C 에서 계산된 3차원 좌표값을 F_G 로 변환하려면 R

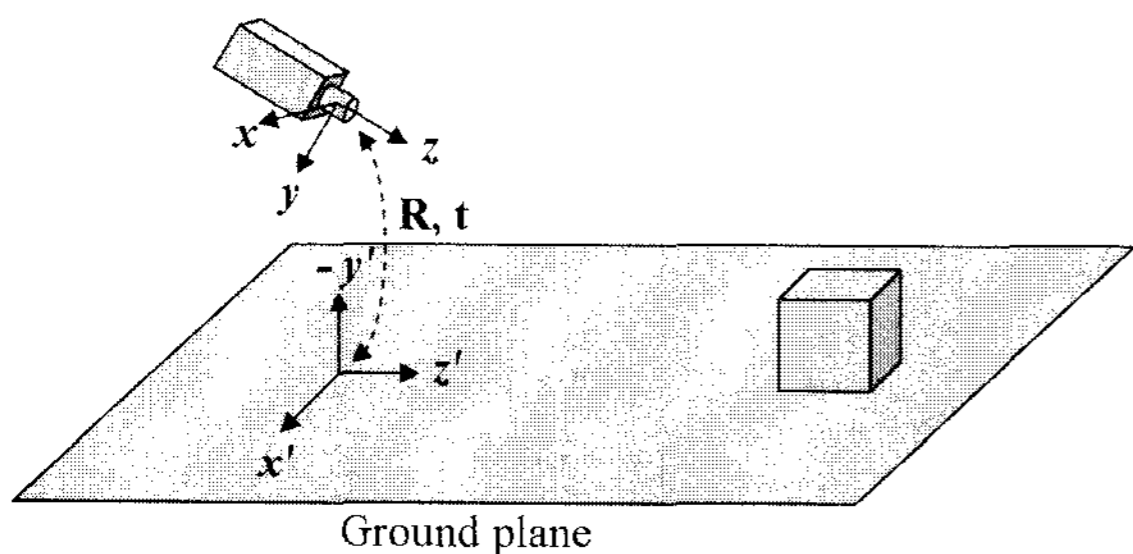


그림 1. 스테레오 영상 취득 구조
Fig. 1. Stereo imaging geometry.

이외에도 천이 벡터 t 를 정해야 하는데, 이 값은 지면상의 임의의 한 점을 원점으로 설정하여 결정할 수 있다. 이 점을 p_0 라 하면, 식 (5)에서

$$x' = R^{-1}(x - p_0) \quad (7)$$

로 설정하여 F_C 에서 F_G 로의 변환을 계산할 수 있다.

III. 기반 평면 추출과 3D 투사

1. 기반 평면 추출

영상에서 지면은 주행체가 움직이는 평면이므로 주행을 위한 장애물 추출이나 물체 검출 등의 응용에서 지면을 추출하는 것은 매우 중요하다. 일반적인 주행 환경에서는 지면은 영상의 하단 부분에서 많은 영역을 차지하는 평면이다. 또한 스테레오 카메라의 수평축은 일반적으로 지면과 평행하게 놓이게 된다. 이러한 성질을 이용하여 깊이 영상에서 지면을 추출하기 위해 V-변위 기법이 제안된 바 있다^[7]. 그림 1에서와 같이 지면이 카메라 축과 일정한 각도를 이루고 카메라의 수평축이 지면과 평행하게 배치되어 있으면 지면의 변위값은 영상의 수직 방향으로만 변화하고 각 수평주사선 상에서는 일정한 값을 유지하게 된다. V-변위는 각 수평주사선에서의 변위값 분포를 구한 것이다.

실험 영상을 통해 V-변위의 사례를 보기로 한다. 그림 2는 Point Gray 사의 Tri-clops 카메라를 이용하여 취득한 320x240 크기의 영상으로 변위값의 범위는 16 픽셀이다. 이 영상에 대해 신뢰전파 (belief propagation) 방식의 스테레오 알고리즘^[10]을 적용하여 얻은 결과가 그림 3 (a)에 나와 있는데, 물체의 경계 부분을 제외하고는 대체로 결과가 정확한 것을 볼 수 있다. 그러나 변위값의 범위가 비교적 작아 지표면의 변위값이 계단 형태로 서서히 변하고 있다. 이 변위 영상을 이용하여 구한 V-변위 영상을 그림 (b)에 표시하였는데, 여기서 수직축은 영상의 수직축과 일치하며, 수평축은 변위값을 나타낸다. 각 점의 밝기는 해당 주사선상에서 그 변위값을 가지는 점들의 개수에 비례하도록 나타냈다. 영상의 하단에서 밝은 점들이 사선으로 나타나는데, 이 부분이 지면(ground plane)에 해당된다. 이 실험 영상의 경우 변위값의 범위가 작아서 지면 부분이 계단 형태의 사선으로 나타나고 있다.

V-변위 영상에서 지면에 해당하는 직선이 수직축과 이루는 각도는 지면에 대한 카메라의 방향 정보를 표현

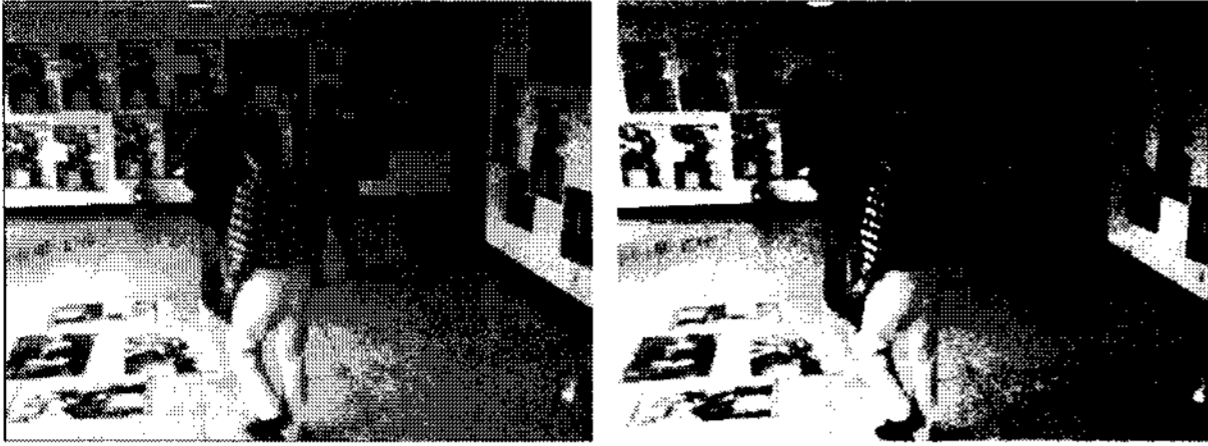


그림 2. 실내 스테레오 영상
Fig. 2. Indoor stereo images.

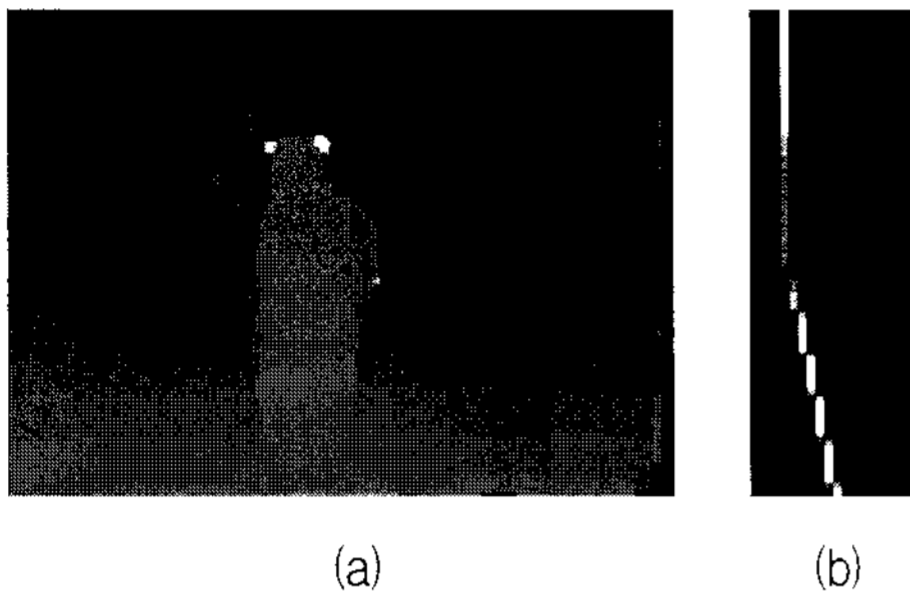


그림 3. (a) 신뢰전파 기법을 이용하여 추출된 변위값
(b) V-변위 영상. 밝기는 각 주사선에서 해당 변위값을 가진 화소들의 숫자에 비례함.
Fig. 3. (a) Disparity image obtained using belief propagation matching algorithm (b) V-disparity image. Intensity is proportional to the number of pixels with the corresponding disparity on each scanline.

하게 된다. 따라서 V-변위 영상에서 지면과 연관된 직선을 추출하여 카메라 좌표계에서의 지면의 평면 방정식을 유도할 수 있다.

V-변위 영상에서 지면에 해당하는 직선 방정식을 추출하기 위해 본 연구에서는 보다 일반적인 상황에서 지표면을 추출할 수 있도록 RANSAC (Random Sample Consensus) 기반의 추출 방식^[11]을 이용하였다. V-변위 영상 중 하단 부분의 각 수평주사선 상에서 일정한 임계치 이상의 값을 가지는 점들 중 가장 왼쪽에 있는 점들을 추출한 다음, 이들 점들에 대한 근사식을 RANSAC 방식으로 구한다. 이 직선의 방정식이

$$d = bv_l + c \quad (8)$$

로 주어진다고 하면, 이는 기반 평면의 방정식이 되는데, 이 식에 대해 식 (5)를 적용하면 3차원 좌표에서의 기반 평면 방정식을 유도할 수 있다. 이와 같이 유도된 평면 방정식이 $\alpha x + \beta y + \gamma z + \delta = c^T x + \delta = 0$ 로 주어진다면 이 평면에 대한 법선 벡터는 $n = \frac{c}{\|c\|}$ 이다. 법선 벡터를 알면 식 (6)을 적용하여 카메라 좌표계에서 기반 평면 좌표계로의 회전행렬을 결정할 수 있다.

2. 3D 복원 및 좌표계 변환

스테레오 정합에 의해 얻어진 각 대응점들에 대해 식 (5)를 적용하여 카메라 좌표계 F_C 에서의 3차원 좌표를 계산하고, 기반 평면 방정식을 통해 계산된 변환 파라미터를 식 (7)을 이용하여 지면 중심 좌표계에서의 좌표값을 계산한다. 이 좌표계에서는 지면이 $y' = 0$ 인 평면에 대응되므로, y' 값을 통해 각 대응점들이 지면에서 얼마나 떨어져 있는지 판단할 수 있으며, 이를 이용하여 장애물 추출이나 자유 공간 (free space) 탐색 등 3차원 공간상에서 필요한 다양한 작업을 수행할 수 있다.

IV. 점유맵 구축

1. 점유맵 구축

스테레오 정합에 의해 복원된 정합점들은 카메라 중심 좌표계로 표현되어 있는데, 점유맵은 기반 평면을 중심으로 한 3차원 좌표계로 표현된다. 스테레오 비전을 이용하여 점유맵을 구축하는 경우 변위값이 화소 단위로 계산되므로 깊이 정보의 해상도가 영상의 해상도에 비해 매우 낮아지게 된다. 일반적으로 320x240 크기의 스테레오 영상에서 얻을 수 있는 변위값의 범위는 30 pixel 이내이므로 깊이 정보는 30 단계 정도로 제한된다.

변위값을 이용하여 복원된 깊이 정보는 카메라에서 멀어질수록 오차 범위가 커지는데, 그림 4는 복원된 좌표값의 오차 범위를 나타낸 것이다. 그림 4에서는 좌측 영상의 한 화소에서의 변위값(d)이 변할 때 복원된 좌표의 오차 범위를 보여주고 있는데, 변위값이 작을수록 오차 범위가 크다는 것을 알 수 있다.

점유맵은 기반 평면상의 각 점들을 물체가 놓인 부분과 빈 부분으로 구분하여 나타낸 것이다. 이러한 점유 정보를 추출하기 위해 우선 기반 평면을 일정한 크기의 셀로 이루어진 그리드로 나타내는데, (x', z') 평면을 일정한 간격으로 양자화한다. 이 평면상에서 점유맵은 정합점들을 기반 평면으로 투사하여 얻게 된다. 점유맵은 일정한 크기의 사각형 셀들로 구성되지만, 각 스테레오 정합점이 나타내는 공간 영역은 그림 4에서 보이는 바와 같이 크기가 일정하지 않은 마름모에 가까운 사각형들이다. 따라서 각 정합점들을 기반 평면상으로 투사할 때 이 문제를 고려하지 않으면 본 영상에서는 인접한 점들에서 투사된 점들이 점유맵 상에서는 서로 분리된 점들로 나타나게 될 수 있다. 이 문제를 해결하기 위해

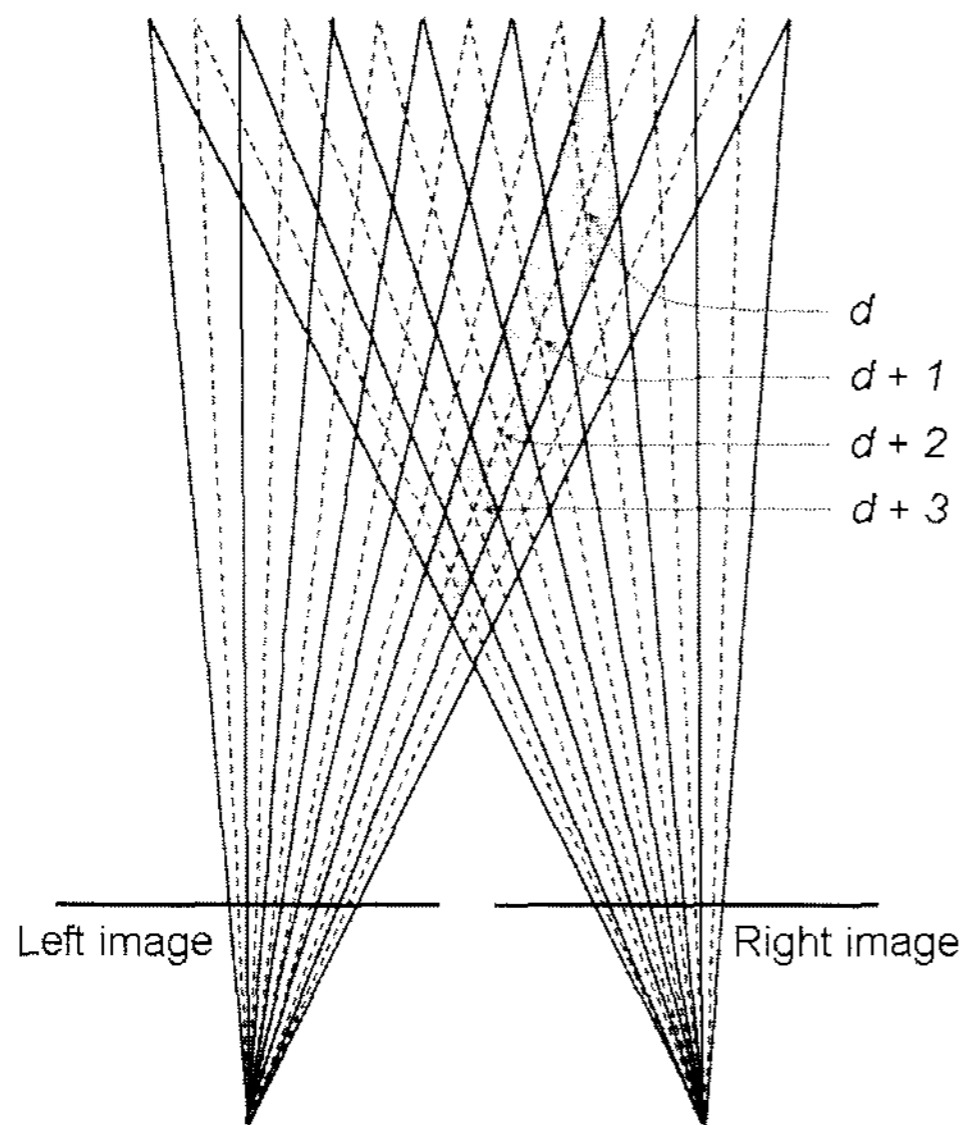


그림 4. 3차원 복원 시의 오차 범위
Fig. 4. Error range of 3D reconstruction.

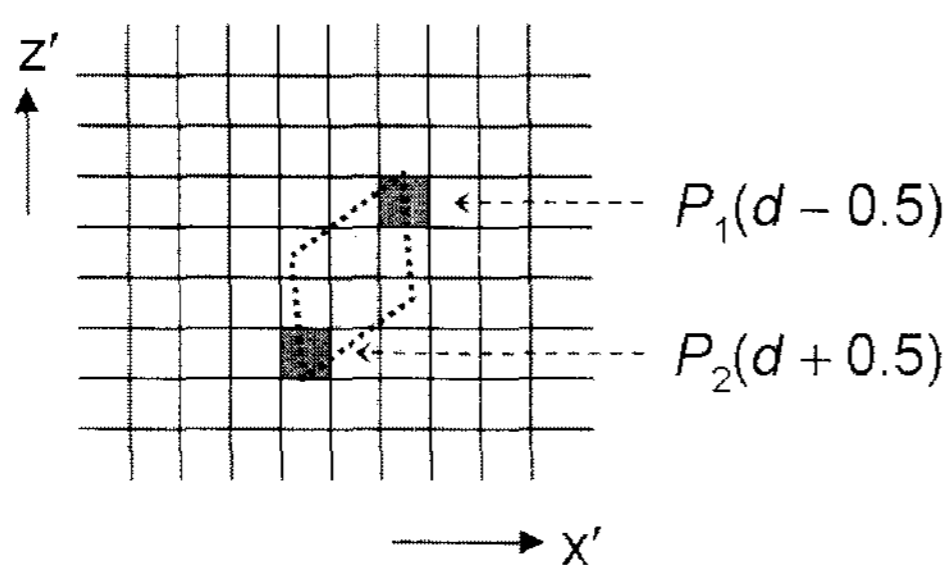


그림 5. 정합점 투사 방식
Fig. 5. Projection of matched point.

본 연구에서는 각 스테레오 정합점들을 점유맵 상의 직선으로 투사하는 방법을 사용하였다. 이 방식을 그림 5에 보였는데, 변위값이 d 인 각 대응점들에 대해 각각 $d-0.5$ 와 $d+0.5$ 에 해당하는 투사점 P_1 , P_2 를 구한다. 이 두 점을 잇는 직선상의 모든 점들로 대응점을 투사하는 것이다. 이 직선은 그림 5에 나타난 바와 같이 3차원 복원 좌표의 오차 범위에 해당하는 점선 사각형과 일치하지 않으나 영상내 물체 영역내의 연결 정보는 보존하게 된다.

점유맵은 물체에 의해 점유된 영역과 자유 공간, 그리고 카메라로 관찰되지 않는 비가시(invisible) 영역으로 구성된다. 자유 공간에 관한 정보는 특히 자율 주행 시의 경로 추출을 위해 사용되므로 자유공간과 비가시 영역을 구분하는 것이 필요한데, 자유 공간은 영상에서 추출된 기반 평면 영역을 점유맵으로 투사하여 얻을 수 있다. 영상에서 기반 평면에 속한 셀을 점유 맵의 (x', z') 평면상으로 투사하여 자유 공간을 얻게 된다.

2. 물체 영역 검출

기반 평면상에 투사된 정합점들은 각 물체영역별로 응집되어 있게 되는데, 이들 정합점들로부터 각각의 물체 영역을 검출하기 위해 점유맵상에서 연결된 영역 (connected regions)들을 찾아낸다. 정합 오차에 의한 오류를 배제하기 위해 점유맵의 각 셀 중 일정한 개수 이상의 정합점이 투사된 셀들을 점유 셀로 판정한다.

점유맵은 (x', z') 평면상에서 표현되는 정보이므로 점유맵에는 평면상에서의 높이에 관한 정보는 기록되지 않는다. 그러나 높이에 관한 정보도 같은 그리드 상에서 저장할 수 있는데, 각 셀에 투사된 점들의 높이의 최소값과 최대값을 기록하면 해당 물체 영역에 대한 추가적인 정보를 얻을 수 있다.

V. 실험 결과

본 논문에서 제안한 점유맵 구축 방식의 유용성을 검증하기 위해 본 연구에서는 실내 환경과 야외 환경에서 실제 스테레오 영상을 이용한 실험을 수행하였다. 실험 영상들은 스테레오 카메라를 이용하여 취득하였는데, 사용된 카메라 장치들은 교정(rectification)이 수행된 영상과 영역 기반 정합 결과를 제공한다. 본 연구에서는 카메라 자체에서 제공되는 깊이 정보 대신 보다 정확도가 높은 신뢰 전파(belief propagation) 방식을 이용한 스테레오 정합 결과를 이용하였다. 사용된 알고리즘은 화소 단위 정확도로 변위값을 계산한다. 이렇게 추출된 변위값에 대해 식 (4)를 적용하여 3차원 좌표값을 계산하였는데, 영상의 교정 과정에서 도입되는 카메라 좌표계의 스케일링과 원점 이동에 관한 정보가 카메라로부터 제공되지 않아서 복원된 3차원 정보에는 투사 왜곡이 포함되어 있다. 그러나 이와 같이 복원된 3차원 정보만으로도 점유 맵 구축과 물체 영역 추출이 가능하고, 정확한 3차원 정보의 복원은 본 연구의 대상이 아니므로 실험에서는 왜곡이 제거된 유클리드 복원을 시도하지 않았다.

1. 실내 환경에서의 물체 추출

우선 실내 환경에서 구축된 점유맵을 이용하여 움직이는 영역을 추출하는 실험을 수행하였는데, 실험 영상은 Point Gray 사의 Triclops 카메라로 취득한 그림 2의 영상이다. 이 영상에 대해 신뢰전파 방식의 스테레오 알고리즘을 적용하여 얻은 결과가 그림 3 (a)에 나와 있는데, 물체의 경계 부분을 제외하고는 대체로 결

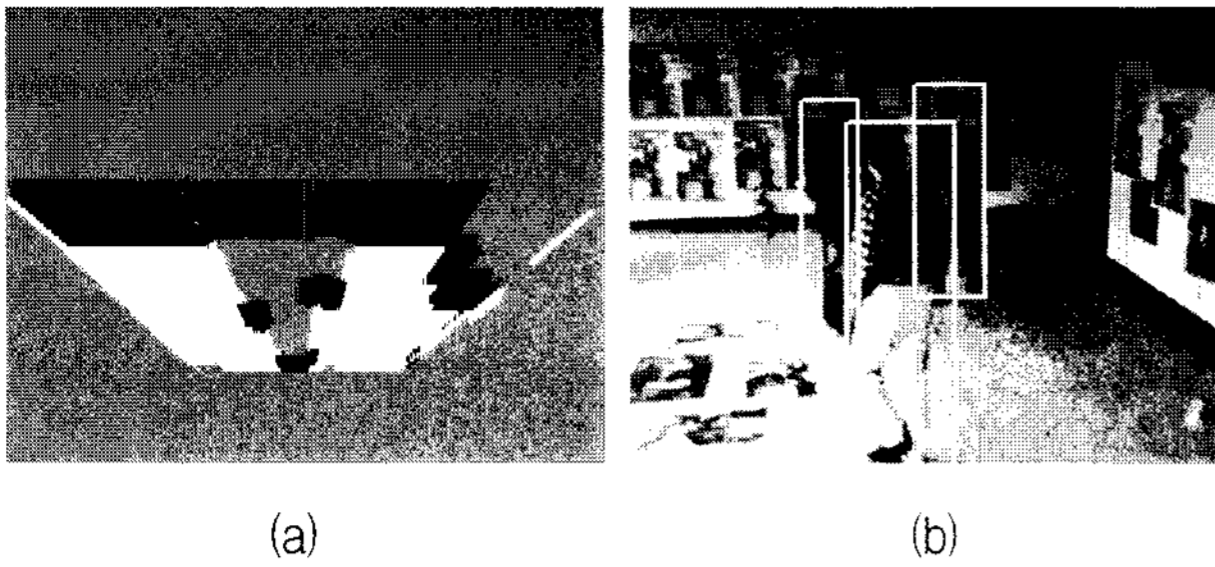


그림 6. (a) 변위 영상에서 추출한 점유맵
(b) 점유맵에서 추출한 물체 영역
Fig. 6. (a) Occupancy map extracted from disparity image (b) Object regions extracted from occupancy map.

과가 정확한 것을 볼 수 있다. 이 영상에서 변위값의 실제 범위는 10 픽셀 정도로 비교적 작아 지표면의 변위값이 계단 형태로 서서히 변하고 있음을 볼 수 있다. 그림 3(b)는 V-변위 영상을 보인 것인데, 지표면에 해당하는 영상 하단 부분을 직선으로 근사하여 지표면을 추출했다.

정합점들의 3차원 좌표를 계산한 후 지표면 상의 정합점들을 제외한 나머지 점들을 지표면으로 투사한 결과를 그림 6(a)에 보였다. 이 그림에서 지표면 상의 점유 공간들은 검은 색으로 표시되었고, 지표면은 흰 색으로, 그리고 공간 정보가 알려지지 않은 점들은 회색으로 표시되었다. 이 그림에서 검게 표시된 점들은 움직이는 물체 영역이거나 벽에 해당된다. 점유맵에서 윗부분으로 갈수록 카메라에서 먼 부분이며 변위값이 상대적으로 작으며, 따라서 거리 정보(z' 값)의 오차가 커지는 것을 볼 수 있는데, 이 현상은 스테레오 방식의 근본적인 한계라 할 수 있다.

본 실험에서는 지표면 상에서 움직이는 물체를 검출하는 실험을 수행하기 위해 각 물체 영역 중 면적이 일정 크기보다 작은 영역을 검출하는 간단한 필터링을 수행했다. 이렇게 검출된 영역에 대해 원래 영상에서의 경계영역을 계산했는데, 그 결과가 그림 6(b)에 표시되었다. 이 그림에서 복도 위를 움직이는 세 명의 윤곽이 정확하게 추출되었는데, 단일 영상에서는 세 명의 영역이 일부분이 겹쳐져 있지만, 그림 (a)의 점유맵 상에서는 서로 분리되어 있어서 사람의 경계선이 그림 (b)에서 정확히 나누어지는 것을 볼 수 있다.

이 실험 영상에서 스테레오 방식을 사용하지 않고 좌측 영상 시퀀스만을 활용하여 움직이는 물체를 검출하는 경우의 정확성을 알아보기 위해 배경 차감(background subtraction) 방식을 이용한 실험을 수행



그림 7. 배경 분리 기법으로 추출된 움직이는 물체들
Fig. 7. Moving object regions extracted using background subtraction.



그림 8. 여러 다른 프레임에서의 움직임 영역 검출 결과
Fig. 8. Moving object regions on different frames.

하였다. 좌측 영상 시퀀스의 미디안(median) 영상을 계산하여 배경 영상을 구한 다음, 그림 2의 좌측 영상과 이 배경 영상의 차이를 계산했다. 그림 7은 이 차이 영상에 대해 두 개의 서로 다른 임계치를 적용하여 얻은 이진(binary) 영상들이다. 그림에서 보는 바와 같이 이 방식에서는 어떤 임계치를 적용하더라도 바닥 면과 벽에서 발생하는 그림자와 물체의 밝기 분포에 따른 오차가 많이 발생함을 볼 수 있다. 따라서 세 명의 영상 영역을 분리하는 것이 쉽지 않을 것임을 알 수 있다.

그림 8은 스테레오 영상 시퀀스의 여러 다른 프레임에서 물체를 검출한 결과를 보이고 있다. 이 그림에서 물체의 윤곽은 모두 정확하게 추출되었는데, 그림 (c)에서는 많이 가려진 사람도 정확히 검출되었으나, (d)에서는 두 사람의 변위값의 차이가 1 픽셀 밖에 되지 않아 영역이 합쳐진 것으로 검출되었다. 이와 같은 결과는 점유맵 상에서 단순히 영역의 면적만으로 물체를 분리했기 때문에 발생한 것으로, 영역의 형상 정보를 활용

하는 보다 정교한 물체 추출 알고리즘을 적용하면 정확성을 높일 수 있을 것이다.

2. 도로 영상에서의 자동차 추출

제안된 방식의 유용성을 검증하기 위한 다른 응용 사례를 위해 도로에서 취득한 스테레오 영상을 이용한 실험을 수행하였다. 이 실험에서의 주안점은 변위영상에서 점유맵을 구축하는 절차의 유용성을 검증하기 위해 점유맵을 활용하여 자동차 영역을 분리하는 것이다. 실험에서는 자동차의 형상 정보와 같은 추가적인 영상 정보를 활용하지 않고 점유맵상에서 단순한 영역분리 방식을 사용하였다. 실험 영상들은 Bumblebee XB3 스테레오 카메라를 이용하여 취득되었는데, 영상의 공간해상도는 400x300, 변위값의 범위는 40 pixel 정도이다.

그림 9는 자동차 도로에서 취득한 스테레오 영상인데, 이 영상에 대해 신뢰 전파 기법의 스테레오 알고리즘을 적용하여 얻은 변위 영상이 그림 10 (a)에 나와 있다. 이전과 마찬가지로 여기에서도 변위값은 픽셀 단위의 정확도로 추출되었다. 이 변위 영상에 대한 V-변위 영상을 그림 10 (b)에 보였는데, 이 그림에서 지면 영역이 흰 직선으로 나타나고 있음을 볼 수 있다. V-변위 영상에서 지면 정보를 추출한 다음, 정합점들을 지



그림 9. 도로 스테레오 영상 1
Fig. 9. Road stereo image 1.



(a) (b)

그림 10. 그림 9의 영상에 대한 정합 결과 (a) 변위 영상 (b) V-변위 영상
Fig. 10. Matching results for Fig. 9 images. (a) Disparity image (b) V-disparity image.

면으로 투사하여 얻은 점유맵이 그림 11에 나와 있다. 이 그림에서는 흰 부분이 지면에 해당되는 영역이고, 검은 부분들이 지면보다 높은 점들을 나타낸다. 그림 9의 영상에서는 도로 바깥 부분에서 자동차 이외에 다른 지형물들이 많이 존재하므로 점유맵 상에 많은 물체 영역이 나타나고 있다. 또한 스테레오 정합에 의해 추출된 변위 영상에 포함된 오차로 인해 점유맵의 물체 영역도 일부분은 부정확하게 나타났다.

자동차 영역을 찾기 위해 앞서 설명한 바와 같이 점유맵의 정보만을 활용하는 단순한 영역분리 방식을 적용하였다. 우선 점유맵에서 영역 추출 알고리즘을 이용하여 독립된 영역들을 추출하여 그 면적이 일정한 범위 내에 들어오는 영역들을 초기 대상 영역으로 설정한다. 이후 이 영역내의 점들에 대한 3차원 정보로부터 그 영역의 지면으로부터의 높이의 범위를 구하고, 그 높이가 자동차의 높이보다 지나치게 낮거나 높은 영역들은 자동차가 아닌 영역으로 구분하였다. 이러한 방식에 의해 검출한 자동차 영역을 그림 12에서 경계영역으로 표시하였다. 자동차 영역 추출 방식이 매우 단순하고 카메라로부터 멀리 떨어진 일부 자동차 영역은 매우 작은데도 불구하고 자동차들이 잘 검출되었음을 볼 수 있다. 카메라에서 먼 영역들은 자동차를 두 대씩 포함하고 있는데, 이는 그림 12 (a)의 점유맵에서 대략 사각형으로 나타나는 자동차 영역들이 서로 닿아 있어서 발생한 현상이다. 보다 정교한 영역 검출 알고리즘을 사용하면

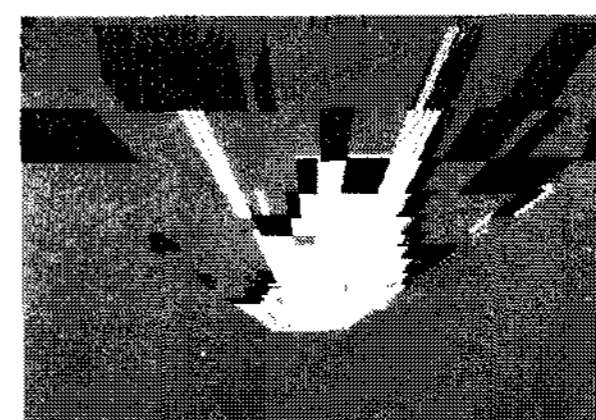
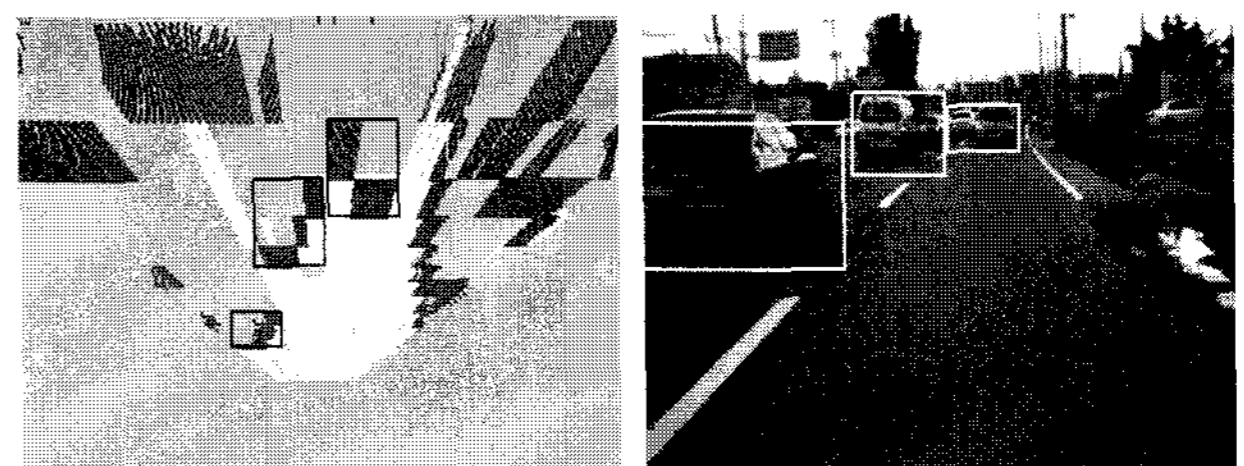


그림 11. 추출된 점유맵
Fig. 11. Extracted occupancy map.



(a) (b)

그림 12. (a) 점유맵 상의 물체 영역 (b) 자동차들의 경계영역
Fig. 12. (a) Object regions on occupancy map (b) Bounding boxes for vehicle regions.

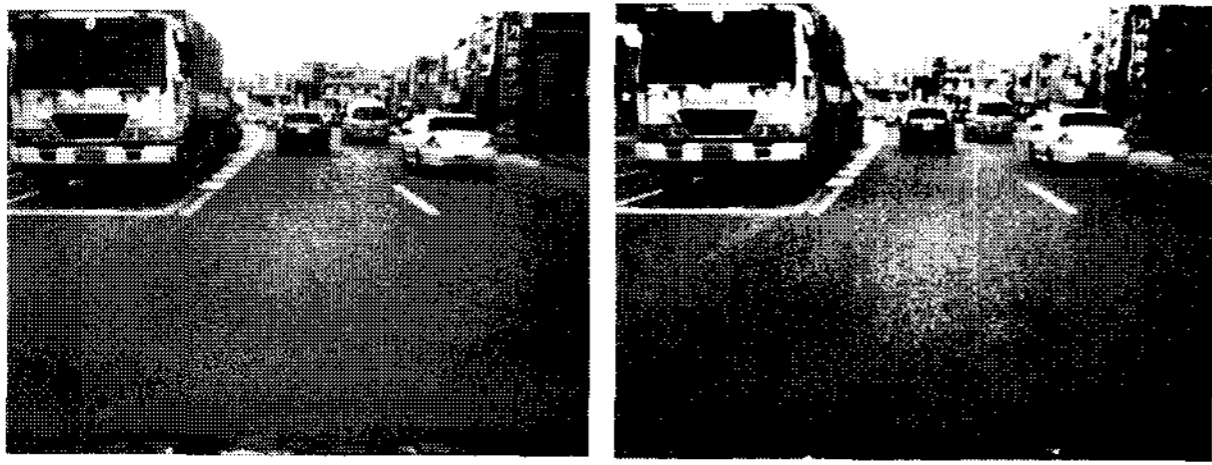


그림 13. 도로 스테레오 영상 2
Fig. 13. Road stereo image 2.

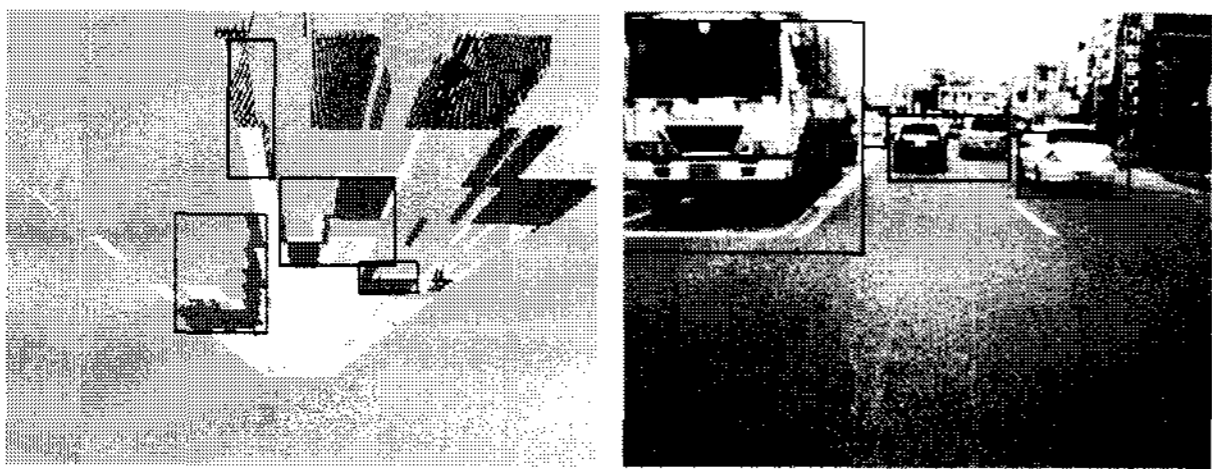


그림 14. (a) 추출된 점유맵 (b) 자동차의 경계영역
Fig. 14. (a) Extracted occupancy map
(b) Bounding boxes for vehicle regions.

각 자동차들을 분리할 수 있을 것이다.

그림 13은 도로 영상의 다른 사례인데, 이 영상에서 추출된 변위맵과 자동차 영역을 그림 14에 보였다. 이 그림에서도 자동차 영역들이 잘 분리되었음을 볼 수 있다. 특히 이 영상에서는 트레일러 영역이 정확히 분리되었는데, 자동차 영역에서의 높이와 길이 정보를 활용하면 자동차의 종류를 인식하는 것도 가능할 것으로 보인다.

3. 주차 구역에서의 자동차 추출

그림 15는 주차장에서 취득한 스테레오 영상이며, 이에 대한 변위 영상과 V-변위 영상이 그림 16에 나와 있다. 스테레오 정합 결과를 보면 자동차의 유리 부분에서의 투명성과 반사로 인해 변위값이 불규칙하게 추출되었지만 대체로 정확한 정합 결과가 얻어졌다. 이 정합 결과로부터 그림 16(a)의 점유맵이 추출되었는데, 지면 영역과 자동차 영역, 뒤의 벽 부분이 대체로 정확



그림 15. 주차장에 대한 스테레오 영상 1
Fig. 15. Parking lot stereo images 1.

하게 추출되었다. 이 점유맵에서는 자동차 영역과 벽이 맞닿아 있어서 영역 추출 알고리즘으로 영역을 분리하는 대신 점유맵의 수직축 상의 점들의 분포를 통해 자동차의 앞 부분을 구하고, 그 점에서 일정한 깊이 만큼

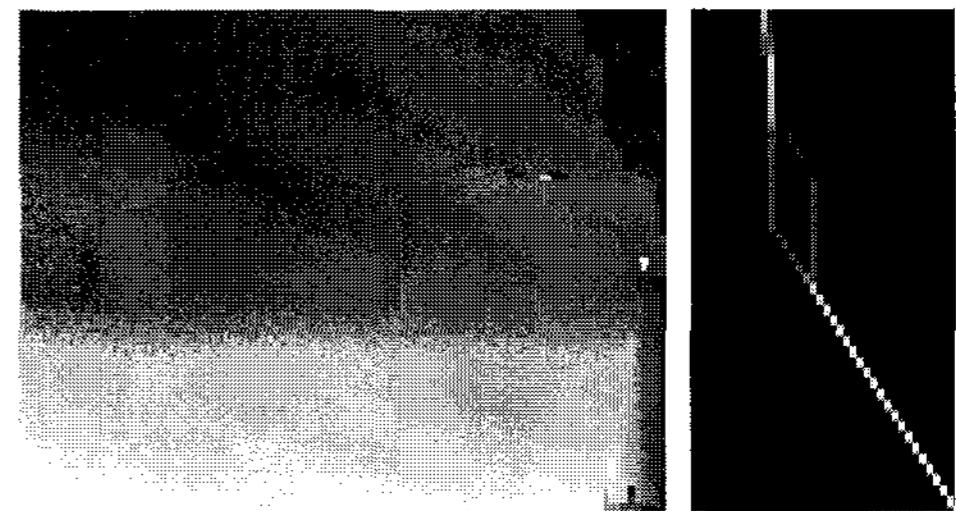


그림 16. 그림 15의 영상에 대한 정합 결과
(a) 변위 영상 (b) V-변위 영상

Fig. 16. Matching results for Fig. 15 images.
(a) Disparity image (b) V-disparity image



그림 17. (a) 추출된 점유맵 (b) 자동차의 경계영역
Fig. 17. (a) Extracted occupancy map
(b) Bounding boxes for vehicle regions.

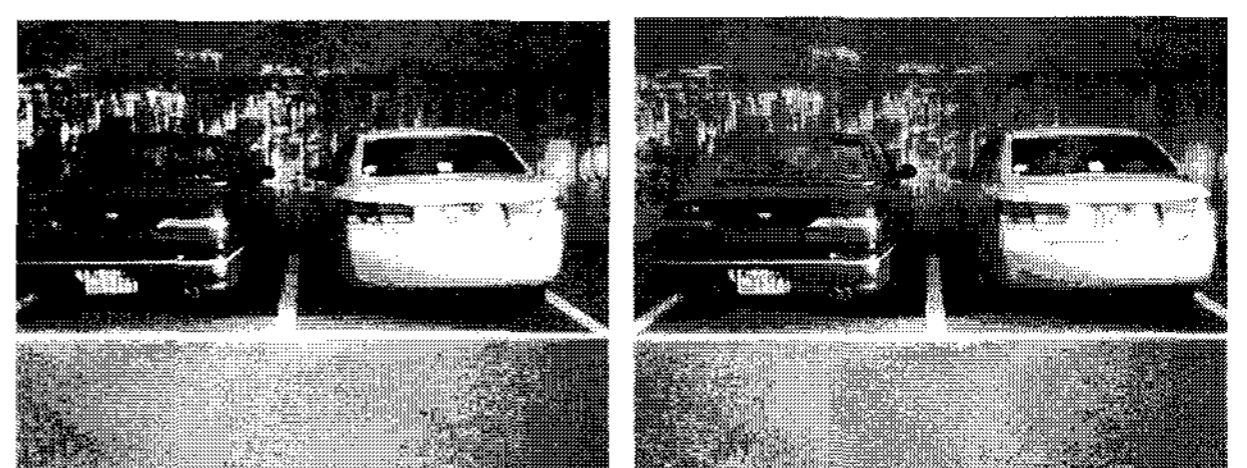


그림 18. 주차장에 대한 스테레오 영상 2
Fig. 18. Parking lot stereo images 1.

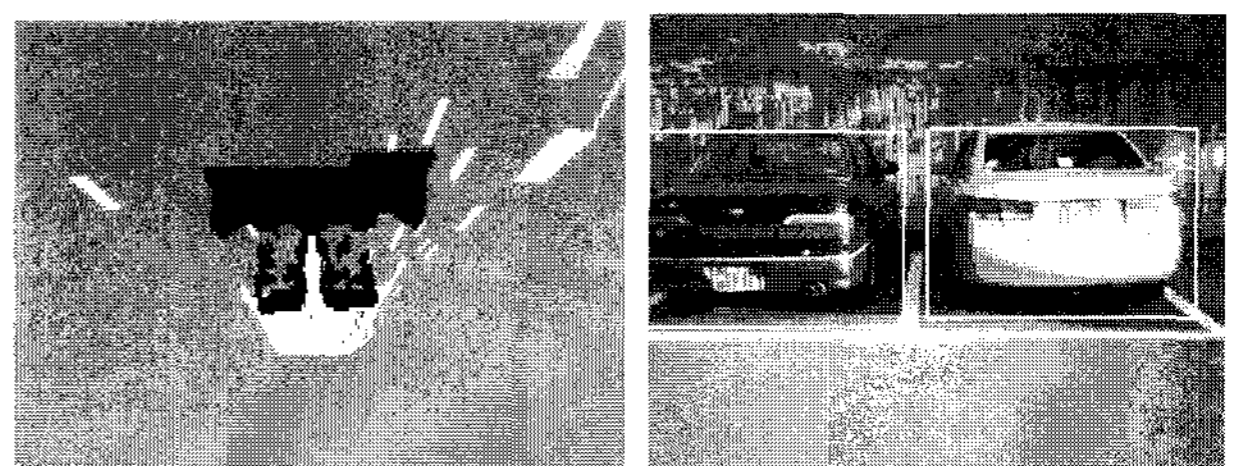


그림 19. (a) 추출된 점유맵 (b) 자동차의 경계영역
Fig. 19. (a) Extracted occupancy map
(b) Bounding boxes for vehicle regions.

의 영역을 분리하는 방식으로 자동차 영역을 추출하였다. 자동차의 경계영역을 그림 16(b)에 표시하였다.

그림 17과 18은 다른 위치에서의 주차장 영상과 추출된 점유맵, 그리고 자동차의 경계영역을 보이고 있다. 자동차 유리 부분의 투명성으로 인해 추출된 점유맵의 자동차 영역에서 빈 공간이 많이 나타나지만, 자동차 영역이 잘 분리되었음을 볼 수 있다.

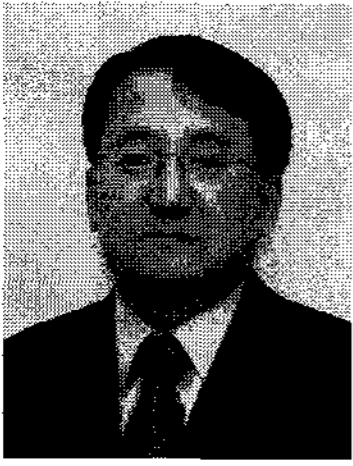
VI. 결 론

본 논문에서는 스테레오 영상에서 추출된 깊이 정보를 활용하여 3차원 공간의 점유맵을 구축하고 그 정보를 물체 영역 추출에 활용하는 방법을 제안하였다. 스테레오 깊이 영상에서 RANSAC 방식을 이용하여 기반 평면을 추출한 다음, 각 정합점들을 기반 평면 중심 좌표계로 투사하여 점유맵을 추출한다. 추출된 점유맵은 주행체가 움직일 수 있는 자유 공간 정보를 기본적으로 제공하지만, 공간상에 놓인 물체에 관한 추가적인 정보를 포함한다. 이러한 공간 정보의 유용성을 검증하기 위해 실내외의 다양한 환경에서 움직임 물체를 추출하는 기법을 설명하였고, 실제 실험 영상에 적용한 결과를 제시하였다. 본 연구에서는 점유맵 자체의 유용성 확인을 위해 물체 영역 분할 과정에서 공간 정보만을 활용하였고 그것만으로도 좋은 결과를 얻을 수 있었다. 각 응용 환경에서 물체의 대칭성이나 색상 정보 등의 추가 정보를 활용하면 더욱 정확한 영역 분할 결과를 얻을 수 있을 것으로 예상된다.

참 고 문 헌

- [1] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," IEEE Computer, pp. 46-57, June 1989.
- [2] H. Moravec and A. Elfes, "High-resolution maps from wide-angle sonar," Proc. IEEE Conf. on Robotics and Automation, March 1985.
- [3] D. Murray and C. Jennings, "Stereo vision based mapping and navigation for mobile robots," Proc. of IEEE Conf. on Robotics and Automation, 1997.
- [4] C. Brailon et al, "Fusion of stereo and optical flow data using occupancy grids," Proc. IEEE Conf. on Intelligent Transportation Systems, 2006.
- [5] N. H. Kim and J. S. Park, "Segmentation of object regions using depth information," Proc. of IEEE Conf. on Image Process., 2004.
- [6] M. Bertozzi, A. Groggi, A. Fascioli, and S. Nichele, "Stereo vision-based vehicle detection," Proc. of IEEE Intelligent Vehicles Symp., 2000.
- [7] R. Labayrade, D. Aubert, and J. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through V-disparity representation," Proc. of IEEE Conf. on Intelligent Transportation Systems, 2002.
- [8] R. Labayrade, D. Aubert, "In-vehicle obstacles detection and characterization by stereovision," Proc. of Int. Workshop on In-Vehicle Cognitive Computer Vision Systems, 2003.
- [9] Z. Hu, F. Lamosa, and K. Uchimura, "A complete U-V-disparity study for stereovision based 3D driving environment analysis," Proc. of IEEE Conf. on 3-D Digital Imaging and Modeling, 2005.
- [10] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision," Proc. of IEEE Conf. Computer Vision and Pattern Recog., 2004.
- [11] D. A. Forsyth and J. Ponce, Computer Vision: A Modern Approach, Prentice Hall, 2003.

저 자 소 개



김 낙 현(정회원)
 1982년 서울대학교 전자공학과
 학사 졸업.
 1984년 서울대학교 전자공학과
 석사 졸업
 1989년 Univ. of Texas at Austin,
 전기공학과 박사 졸업

1993년~현재 한국외국어대학교 전자정보공학부
 교수
 <주관심분야: 컴퓨터 비전, 디지털 비디오 시스
 템>



오 세 준(학생회원)
 1999년 한국외국어대학교 제어
 계측공학과 학사 졸업
 2005년 한국외국어대학교 전자
 정보공학과 석사 졸업
 2008년 현재 한국외국어대학교
 전자정보공학과 박사과정

<주관심분야: 컴퓨터 비전>