

# 산업용 자율 주행 로봇에서의 격자 지도를 사용한 강화학습 기반 회피 경로 생성기 개발

## Development of Reinforcement Learning-based Obstacle Avoidance toward Autonomous Mobile Robots for an Industrial Environment

양정연  
목원대학교

Jeong-Yean Yang(jyyang@mokwon.ac.kr)

### 요약

자율 주행은 이동 로봇의 핵심적 기술로써, 측정된 센서 정보를 토대로 불확실한 위치 정보를 이용한 지도 작성 및 수정 기능과, 불확실한 지도 및 센서 정보를 이용한 로봇의 위치 인식 기능으로 구성된다. 자율주행은 이러한 주행 위치의 불확실성에 기반한 확률론적 방법론과 함께 주행 시 장애물의 감지 및 회피 경로의 생성, 반복적 주행 패턴에 따른 경로 관리 기능이 필수적 요소이다. 거리 기반의 스캐너를 통해 관측된 센서 입력은, 지도 구성에 사용된 벽과 같은 정적 물체와 주행 시의 사람처럼 움직이는 동적 물체와의 구별이 필요하기 때문에 장애물 감지에 어려움이 있다. 본 논문에서는, 이러한 자율 주행 환경에서 기존의 정적, 동적 개체의 판별 방식과 비교하여, 장애물 회피를 위한 저해상도 격자 공간의 생성 및 강화학습을 이용한 경로 생성을 다루고자 한다. 최종적으로 실험을 통해 제안된 방법론의 실효성을 검증하고자 한다.

■ 중심어 : | 장애물 회피 | 격자 지도 | 강화 학습 | 자율 주행 | 경로 생성 |

### Abstract

Autonomous locomotion has two essential functionalities: mapping builds and updates maps by uncertain position information and measured sensor inputs, and localization is to find the positional information with the inaccurate map and the sensor information. In addition, obstacle detection, avoidance, and path designs are necessarily required for autonomous locomotion by combining the probabilistic methods based on uncertain locations. The sensory inputs, which are measured by a metric-based scanner, have difficulties of distinguishing moving obstacles like humans from static objects like walls in given environments. This paper proposes the low resolution grid map combined with reinforcement learning, which is compared with the conventional recognition method for detecting static and moving objects to generate obstacle avoiding path. Finally, the proposed method is verified with experimental results.

■ keyword : | Obstacle Avoidance | Grid Map | Reinforcement Learning | Autonomous Locomotion | Path Generation |

## 1. 서론

이동 로봇은, 바퀴 또는 캐터필러와 같은 구름 접촉을 이용한 지면에서의 이동 방식 및 드론과 같은 3차원 공간 상의 비행에 따른 위치 변경이 가능한 로봇을 말한다. 이러한 이동성에 대한 제어 기능을 토대로, 물류 이송, 농업 및 해양 환경, 보안 설비, 위험한 환경의 탐사에 활용되며, 주행 방식에 따라 크게 원격지 조종에 의한 원격 제어와 로봇 내 부착된 센서를 이용하여 로봇이 스스로 움직이는 자율 주행 방식으로 구별된다.

자율 주행은, 부착된 거리 기반 센서를 이용하여 작업 공간 내 실제 주행을 통한 지도 작성 기능과, 생성, 수정된 지도와 센서 정보를 융합하여 현재 위치를 보정하는 위치 추정 기능이 요구된다. 이러한 방법론은, 현재 위치가 불확실한 상태에서 관측을 통해, 위치의 추정 및 보정과 함께 지도를 수정하는 방식으로 구현되며, 이러한 불확실성 때문에 확률 기반 방법론이 필수적으로 사용되고 있다.

초기 자율 주행 로봇의 경우, 이러한 지도 작성과 위치 추종이 별개의 작업으로 구성되었으나, 자율 주행 작업 수행의 편의성 및 성능 향상을 위해 지도 작성과 위치 추종이 동시 수행되는 동시적 위치추정 및 지도 작성(Simultaneous Localization and Mapping, SLAM) 기법이 보편적이다[1][2].

지도 작성의 경우, 로봇 내 부착된 거리 기반 센서, 영상 기반 센서, 또는 포인트 클라우드 추출 센서를 이용하여, 현재 추정된 위치를 기반으로 주위 벽과 같은 고정된 물체의 센서 측정값을 지도로 활용하기 위해 확률값을 반복, 수정하는 방식을 사용하고 있다.

위치 추종은, 이동 로봇의 제어 명령에 따른 바퀴 회전량과 로봇의 방향각, 위치에 대한 기구적 모델을 사용하여 예상 위치를 추정하고, 이후 지도 및 센서 측정값의 비교를 통해 위치 추정 결과를 역시 반복, 수정하는 방식을 사용하고 있다.

자율 주행 로봇은, 지능형 서비스 로봇 분야에서 활발히 진행되어 왔으나, 다양한 불확실성이 잠재하고 있는 실 환경 서비스 측면에서 다소 미흡한 실정이다. 가정, 도로, 공공 장소 등의 환경은, 지형의 변화에 따른

지속적인 지도 수정이 빈번하며, 환경 내의 동적 물체의 출현 빈도가 높아 감지 및 회피에 어려움이 있다.

반면 산업용 자율 주행 로봇의 경우, 주어진 작업 공간 환경에 대한 오염 방지, 외부 요소의 접근 제한, 이동 경로의 최적화 등 환경의 유지 관리가 가능하여, SLAM에 의해 확률기반의 지도를 먼저 작성하고, 이후 위치를 추정하여 속도 제어에 의한 경로 추종 운동을 선호한다. 대부분의 산업용 로봇은, 주어진 경로의 추종과 반복적 작업을 수행하는 것이 목적이기에, 이러한 지도 작성과 위치 추정의 분리를 통해, 해당 기능의 복잡성 및 연산량 감소를 추구하고 있다.

이러한 경향성에 따라, 장애물 회피는 SLAM기반의 주행을 전제로, 부가적 기능으로 처리하는 방식을 사용하게 된다. 지도 작성은 움직임이 없는 벽과 같이 고정된 정적 요소에 한정하고, 동적 물체는 임시적인 장애물로 고려하여 지도에 포함하지 않고 경로 생성을 통해 회피하는 방식을 말한다[3].

지도 작성은, 측정된 센서값을 이용한 역측정 모델(Inverse Measurement Model)을 통해 확률 값으로 표현되고, 이를 지도 위에 투영하는 방식을 사용하고 있다[2]. 따라서, 반복적 움직임이 관측되는 동적 물체의 경우, 확률로 표기되는 지도 위의 잠재적 장애물 출현 지역으로 간주될 가능성이 높아진다.

이는 특정 공간에 움직임이 많은 물체가 존재할 때, 해당 지역의 확률적 분포가 정적 물체보다는 낮고 빈 영역 보다는 높게 되어, 동적 물체의 출현 빈도가 높음을 의미하는 정보로도 활용 가능하다. 하지만, SLAM은 지도 생성을 포함하고 있고, 회피 경로의 생성 자체는 포함되어 있지 않기 때문에, 자율 주행에서는 장애물 인식에 따른 회피 경로의 생성이 추가로 필요하다.

일반적으로, SLAM의 복잡성과 비교하여 장애물 회피는 비교적 용이한 기능으로 여겨지나, 이는 개방형 공간에 국한된다. 산업용 자율 주행과 같은 실 환경에서는, 밀폐 공간 내 회피, 장애물의 크기 예측, 3차원 공간 감지의 비용 및 기술적 어려움, 주기적 경로 스케줄링과의 연계성이 필요하여, 장애물 회피는 산업용 자율 주행 로봇에서의 성능 향상의 필수 요소로 간주된다.

본 논문에서는, 이러한 장애물 회피를 위해 강화 학

습 기반의 회피 경로 생성을 다루고자 한다. 기존의 장애물 감지에 따른 강화 학습 기반 회피와, 제안된 장애물 감지용의 격자 지도 생성 및 강화 학습 기반 회피의 방법론을 비교하고자 한다.

## II. 확률 기반의 지도 작성 기법

SLAM은 마르코프 체인상의 센서 측정값과 위치 상태량의 변화를 토대로 다음의 확률값을 구하고자 한다.

$$p(x_{0:t}, m | z_{1:t}, u_{1:t})$$

이 때,  $x, z, u$  는 각각 시작부터 현 시점,  $t$  까지의 위치 벡터, 관측값, 제어 입력으로 구성되어 있고,  $m$  은 지도를 의미한다.

이처럼 위치 벡터  $x$  와 지도 작성  $m$  은 이동 로봇의 움직임과 관측 결과에 의존하는 조건부 확률로 구성되어 있으며, 각 요소의 결합 확률을 직접 구하기 어려워 Rao-Blackwellization에 의해 다음처럼 독립적 요소로 분리한다[4].

$$p(m | x_{1:t}, z_{1:t}) p(x_{0:t} | z_{1:t}, u_{1:t}) \quad (1)$$

앞부분 확률은 위치벡터와 관측 결과를 토대로 지도  $m$  의 상태를 나타내는 확률이며, 뒷부분 확률은 관측 결과와 제어 입력에 따른 현재 위치에 대한 확률을 의미한다. 이에 따라, 앞부분 확률은 지도 작성, 뒷부분 확률은 위치 추정으로 표현하기도 한다. 두 가지 기능의 동시성을 의미하는 SLAM 기법은 Rao-Blackwellization에 의해 독립된 두 프로세스의 곱으로 위치 추정과 지도 인식의 기능을 동시 수행하는 것을 의미한다.

지도 작성은 현재의 불확실한 위치 추정값,  $x$  을 토대로 벽과 같은 정적인 물체까지의 거리를 측정한 값인,  $z$  를 토대로 지도를 반복 수정하는데, 격자 기반 지도 생성의 경우 다음과 같은 식으로 표현 가능하다.

$$p(m) = \prod_{i=1}^N p(m_i)$$

격자 기반 지도의 경우, 각 격자가 비어 있는지 채워져 있는지를 의미하는 점유 격자 (Occupancy grid) 방식이 적용되는데, 이 때 확률  $p(m)$  은 각각의 격자,  $m_i$

의 확률의 곱으로 표현가능하다.

앞서 (1)에서 언급된 지도 생성 확률은 베이시안 확률 및 마르코프 정의에 의해 다음처럼 변형가능하다.

$$\begin{aligned} p(m_i | x_{1:t}, z_{1:t}) &= p(m_i | x_{1:t}, z_{1:t-1}, z_t) \\ &= p(m_i | z_t, x_t) \frac{p(z_t | x_t) p(m_i | x_{1:t-1}, z_{1:t-1})}{p(m_i) p(z_t | x_{1:t}, z_{1:t-1})} \end{aligned}$$

각각의 격자는 점유 또는 비어 있는 경우에 해당하여 다음처럼 확률 연산이 간단하게 변경가능하다.

$$\frac{p(m_i | x_{1:t}, z_{1:t})}{p(\neg m_i | x_{1:t}, z_{1:t})} = \frac{p(m_i | x_{1:t}, z_{1:t})}{1 - p(m_i | x_{1:t}, z_{1:t})}$$

확률 연산에서는 작은 확률값의 반복적 곱 연산에 의한 수치 오류를 감소하기 위해 로그연산을 사용하는 것이 일반적이다. 따라서, 다음처럼 log odds 표기를 사용하면 확률 연산을 간결하게 표현가능하다.

$$\log \text{Odds} : l(a) \triangleq \log \frac{p(a)}{1 - p(a)}$$

이 때, 기호  $l$  은 log odds를 의미한다. 측정한 값에 대한 각 격자의 점유 확률을 log odds로 표현하면,

$$\begin{aligned} l(m_i | x_{1:t}, z_{1:t}) &= \log \left( \frac{p(m_i | x_{1:t}, z_{1:t})}{p(\neg m_i | x_{1:t}, z_{1:t})} \right) \\ &= \log \left( \frac{p(m_i | x_{1:t}, z_{1:t})}{1 - p(m_i | x_{1:t}, z_{1:t})} \right) \end{aligned}$$

이고, 이를 원래의 지도 작성 확률에 대입하면,

$$\begin{aligned} l(m_i | x_{1:t}, z_{1:t}) &= l(m_i | x_t, z_t) + l(m_i | x_{1:t-1}, z_{1:t-1}) - l(m_i) \quad (2) \end{aligned}$$

와 같은 지도 상의 각 격자에 대한 반복 수정식을 얻게 된다. 이때 식(2)의 현재 시각  $t$  에서 위치 및 관측값을 토대로 격자를 업데이트 하는 요소를 역 측정 모델 (Inverse measurement model)로 정의한다[5]. 역 측정 모델은, 격자 방식에서 지도를 수정, 변경하는 식(2)를 구성하는 주요 모델로 사용되며, 거리 기반 센서의 값에 따른 격자 내 점유 확률값을 토대로 매번 수정을 반복하여 지도를 작성하게 된다[6-8].

### III. 이동 로봇의 장애물 회피 기능

SLAM이 적용된 자율 주행 로봇의 경우, 앞서의 식 (2)를 이용하여 작업 공간에 대한 지도를 작성하는 것이 일반적이다. 이러한 지도는 작업 공간 내의 해상 능력에 따라 최소 1cm에서 최대 10~15cm 정도의 격자 단위로 구성하는 것이 일반적이다.

주행 시의 위치 추종을 위한 파티클 필터(Particle filter)의 적용 시, 예측된 다수 파티클의 이동 결과를 통해 정규 분포로 구성된 근사치인 제안(Proposal) 분포와 측정값과의 가중치를 계산하고 이를 통해 실제 목표(Target) 분포를 추정, 샘플링을 반복하는 과정을 거친다. 이는 중요도 샘플링(Importance sampling) 방법론에 근거하여 적은 수의 샘플에 대한 가중치를 토대로 목표 분포를 얻어내는 방식이다.

위치 추정 방식의 경우, 각 파티클에 의한 예측 위치를 토대로 생성된 지도의 각 격자 확률값과 측정값과의 관계로 가중치를 결정하기 때문에, 지도의 해상 능력인 격자 갯수에 따라 위치 오차의 최소치가 결정된다. 따라서 많은 양의 격자 사용과 역 측정 모델링에 의한 갱신 과정과 다수의 파티클 사용이 일반적이다.

산업용 로봇은 무거운 하중과 노면의 미끄러짐에 의해 위치 제어 성능이 떨어지기에, 최종적인 위치 정확도는 파티클의 수량과 격자 해상도의 설정에 비례한다. 하지만 이러한 다수의 파티클과 고해상도의 격자 사용은, 연산량을 증가시키는 단점이 있어, 일반적으로 실 환경의 상황에 따라 적합한 수량을 결정한다.

#### 1. 정적 및 동적 물체 구별에 의한 장애물 감지

개방형 공간에서 장애물 회피의 대표적인 방법론은 정면에 장애물이 감지되면, 좌측 또는 우측 방향으로 이동하여 장애물과의 충돌을 회피하는 것이 일반적이다. 따라서 정면 거리 센서에 변화가 감지되면 장애물로 인식하고 좌우 이동을 통해 회피하는 간단한 방법이 사용되고 있다.

하지만, 밀폐된 공간에서의 SLAM 기반 주행은, 벽까지의 원거리 측정을 통해 지도를 작성하며, 사용되는 거리 기반 센서의 경우, 180도 또는 360도에 해당하는

최대 반경 15~20m 내의 장애물을 1도 이하의 해상도로 측정 가능하다. 이러한 넓은 영역의 감지 기능 때문에 오히려 장애물 인식에 부가적인 어려움이 있다.

주행 시, SLAM을 위한 거리 기반 센서를 장애물 감지용으로 같이 사용하는 경우, 원거리에 위치한 벽과 같은 정적 물체와 근거리 안에 위치한 일시적 장애물인 동적 물체와의 구별이 반드시 필요하게 된다.

[그림 1]의 상단 그림은, 정적인 요소에 대해 형성된 지도를 의미하며, 아래는 주행 중에 장애물이 나타난 경우로, 관측 정보를 지도 위에 재구성한 경우이다. 이는, 만약 로봇의 정면에 사람이 등장하면, 상단 지도와 하단 센서 정보를 비교하여, 지도상에 존재하지 않는 물체가 나타났음을 찾아내는 과정이 필요함을 말한다.

이러한 기존 방식은, 지도와 센서 정보사이의 위치 차이가 큰 대상을 동적인 장애물로 취급하는 방식이다. 하지만 해당 방식은 현재의 위치 및 각도 정보가 필요하기 때문에, 위치 추정에 따른 오차가 반영되는 방식이다[3].

[그림 1]과 같이 이를 실 환경에 적용한 경우, 위치 추정 오차, 특히 로봇의 진행 각도(Heading angle)의 오차에 의해 두 개의 그림이 정확히 정렬되지 않음을 알 수 있다. 이러한 경향성은 공간의 모양에 따라 달라지는데, 복도와 같이 길쭉한 환경에서 두드러지게 나타난다.

해당 환경처럼 각도에 의한 위치 추정 오차가 발생하는 경우, 작은 진행 각도 차이에 의해서도 그림 상의 위, 아래 양쪽 벽과 원거리의 전방 벽은 지도가 아닌 장애물로 오인되는 경우가 발생하게 된다.

센서 정보가 지도상의 데이터에 해당하는지 여부는 다음과 같은 평가 함수로 판별한다.

$$\begin{aligned}
 m' &= {}^R H_x d \bullet u \\
 b &= \min_{m \in M} \sum \| m - m' \|^2 \\
 \therefore b &> b_{th}
 \end{aligned}
 \tag{3}$$

각도  $\theta$ 에 따른 방향 벡터  $u$ 와 거리 값,  $d$ 를 현재 추종 위치,  $X$ 에 대한 동차 행렬,  $H$ 로 변환하면, 각도에 따른 장애물 위치,  $m'$  로 표현된다. 이를 실제 지도상의 벽에 해당하는 격자 위치,  $m$ 과 거리 비교하여, 특정 문턱값보다 큰 경우 지도상에 존재하지 않는 동적 장애물로

평가한다. 식 (3)은 전체 정렬 오차에 민감하여, 지도 위의 벽이 움직이는 장애물로 오인되는 단점이 있다.

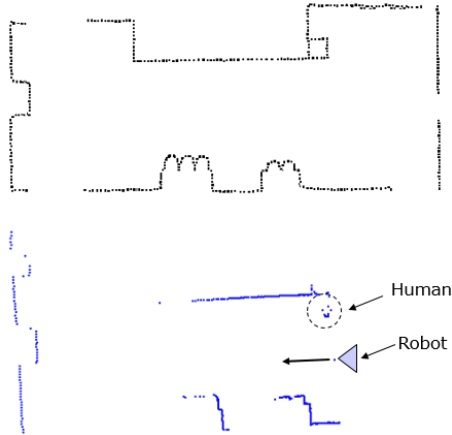


그림 1. 생성된 지도 데이터(상)와 주행 중의 로봇의 추정 위치를 토대로 복원된 거리 정보(하)의 관계. 원형 안에 동적 장애물인 사람이 이동 중이며, 진행 위치, 각도 오차에 따른 지도와 센서 정보간의 정렬 오차를 표현함.

## 2. 격자 기반 장애물 판단 지도 작성

SLAM의 지도 작성 기능은 위치 추정의 오류를 감소하기 위해 다수의 조밀한 격자 공간으로 이루어진다. 이와 달리, 장애물 감지를 위한 격자 지도를 강화학습과 결합하여, 장애물 회피와 함께 처리하고자 한다.

장애물 감지를 위해 격자 기반 지도를 작성하는 경우, 부정확한 로봇의 움직임을 고려하기 때문에 높은 해상 능력이 반드시 필요하지 않다. 로봇의 하중이 크고, 이동 속도에 따른 동역학적 효과가 클수록, 생성된 회피 경로를 토대로 최소한의 움직임으로 원거리에서부터 장애물을 회피하는 것이 유리하다.

따라서, SLAM을 위한 격자 지도는 조밀하게 구성하고, 장애물의 감지 및 회피 경로 생성을 위해서는 별도의 저해상도 격자 공간을 사용하고자 한다.

[그림 2]와 같이, 산업용 로봇의 크기를 고려하여 30cm의 격자를 장애물 회피용 지도로 사용하고, 이를 토대로 식 (2)에 의한 장애물 감지용의 지도 공간을 구현하였다. 이는, 30cm의 격자 공간을 이용하여 전방의 장애물 또는 주위 벽을 추종한 결과로, 왼쪽은 장애물

출현 전의 상태이고, 공간 내 상태 변화에 따라 지도 정보가 수정되고 있는 상태를 의미한다.

정면에 나타난 장애물이 우측으로 이동 중에 있고, 이에 의해 장애물 감지를 위한 저해상도 격자 공간이 장애물의 이동에 의해 수정, 갱신되고 있다.

이처럼 장애물이 이동하거나, 또는 이동 로봇의 움직임에 의해 장애물 주위의 지도의 확률 수정을 통해 장애물의 크기와 위치를 추정할 수 있다. 장애물 감지를 위한 지도가 저해상도를 가지고 있어, 위치 추정을 위한 고해상도 지도와 달리 비교적 적은 연산량으로 처리가 가능한 장점을 갖는다.

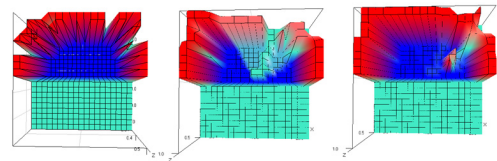


그림 2. 장애물 감지 및 회피를 위한 저해상도 격자 공간의 생성 및 장애물 이동에 따른 확률 수정 과정

## 3. 강화 학습 기반의 회피 경로 설계

산업용 자율 주행 로봇과 같이 제한된 작업 공간과 해당 공간내의 경로 생성을 위해서는, 보로노이 지도(Voronoi's map)와 같은 경계면의 최단 거리 분할 기법, 또는 강화 학습이 주로 사용된다.

보로노이 지도는 영역을 거리에 따라 분할하여, 공간을 다각형으로 나누는 방식이다. 공간 내 두 지점의 최단 거리를 빠르게 계산할 수 있는 장점이 있으나, 동적 장애물과 같은 동적 특성이 필요한 경우, 공간의 재분할 및 트리 구조의 변화가 필요하며, 또한 작은 내각으로 구성된 경계면을 통과할 때 이동 로봇의 급격한 진행 각도 변화를 발생시키는 단점이 있다.

이와 달리, 강화 학습은 미래에 가능한 보상의 기대 값을 사용하는 확률론적 방법론에 해당한다. 장애물의 판단 여부가 불확실한 경우에도, 미래의 보상 값을 추정하는 방식으로 장애물 회피에 적합한 특징을 가지고 있다. 산업용 자율 주행 로봇이 일반적으로 제한된 작업 공간내의 움직임으로 제한되어 있어 다양한 기하학적 공간 내의 회피 경로 생성에 장점을 가진다.

Q 강화 학습은 상태량, s와 행동, a로 구성된 Q(s,a) 값을 사용하는 것이 일반적이다. Q 공간내의 Q 값 학습 기법(Q-learning)은 다음 수식과 같다[9].

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha \left[ r(s, a) + \gamma \max_{a'} Q(s', a') \right]$$

이 때, s, s'은 각각 현재와 천이 이후의 상태량, a, a'은 각각 현재와 천이 이후의 행동, r은 보상값에 해당하며  $\gamma$ 는 천이과정의 미래 보상의 감소율,  $\alpha$ 는 추계적 모델에 따른 갱신률을 의미한다.

격자 공간내의 행동은, 산업용 로봇의 비교적 느린 동특성을 고려하여, 45도 단위의 총 8개 방향으로 구성하고, 앞서 제시한 장애물 격자 해상도에 따라 30cm의 격자 이동을 수행하는 것으로 정의하였다[10].

#### IV. 장애물 회피 기능의 성능 테스트

장애물 회피의 성능 검토를 위해, 가상 환경 기반의 모의 실험 환경을 구축하고, 실 환경에서 산업용 로봇을 이용하여 제안된 방법의 회피 성능을 테스트하였다.

강화 학습의 경우, 탐색에 따른 반복적 에피소드(episode)의 학습이 필요하다. 각 에피소드의 경우, 최대 천이 과정은 3000회 이내로 제한하여 최대 4000번을 반복 탐색하였다. 이는 실시간 주행 중 처리되어야 하므로, 로봇의 이동 속도를 고려하여 0.3초 내에 학습 가능한 최대 회수로 결정하였다.

먼저, 기존의 정적 및 동적 물체 구별에 의한 장애물 감지 방식의 경우, 식 (2)를 이용하여 지도와 비교하여 장애물 여부를 감지하고, 이후 강화 학습으로 경로를 생성하는 방식을 테스트하였다.

이 경우의 강화 학습의 보상 값은, 정적 물체에 대해 음수의 정규 분포 함수로 구성하고, 장애물 감지 시 해당 위치에 음의 보상을 사용하는 방식으로 구성하였다. [그림 3]의 경우, 실제 작업 공간과 동일한 지도를 이용한 모의환경 내에서의 강화 학습 결과이며, (6, 2.5) 좌표에 장애물 감지 시 강화 학습에 의한 회피 경로 생성 결과를 나타내고 있다.

하단부 그림은 언급한 좌표에 동적 장애물이 출현한 경우에, 음수의 보상값에 의해 상태 함수가 오목하게 파인 것을 표현하고 있다.

식 (3)을 적용한 경우, 해당 공간 내에서 코너를 돌 때마다 장애물 감지 오류가 빈번히 발생하였다. 로봇이 해당 공간을 직사각형으로 순회하는 경우, 각 코너에서 90도의 회전이 발생하며 이 때마다 진행방향각에 대한 추정 오차에 의해 [그림 4]와 같은 지도와 현재 센서값 사이의 정렬 오차가 크게 발생한다.

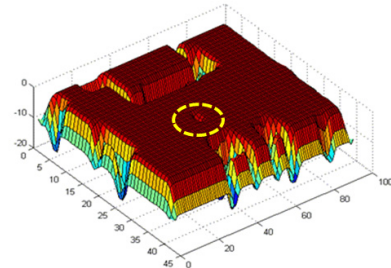
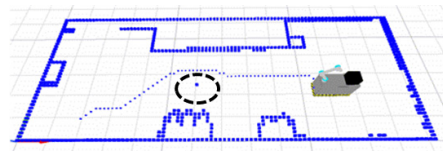


그림 3. 강화 학습 방식으로 구성된 장애물 회피 경로 생성 (상) 및 장애물 감지에 따른 상태함수 분포 결과(하)

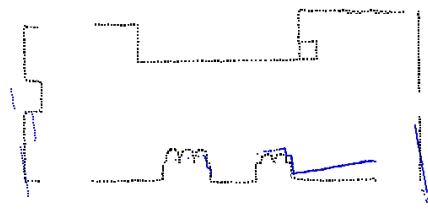


그림 4. 코너링에 의한 급격한 회전 시, 지도와 센서 값의 큰 정렬 오차 발생에 의한 장애물 판단 불가 예시

이러한 한계성 때문에, 제안된 저해상도의 격자 공간을 토대로 장애물을 감지하고자 하였다. 장애물을 감지하는 식 (3)의 방식과 달리, 장애물용 격자를 이용하기에 강화 학습의 보상 함수는 지도상의 확률값을 이용하여 설계가능하다. 벽과 같은 정적인 물체는 최대 확률값을 가지게 되며, 비어 있는 공간은 최저 확률값을 가



지게 된다. Q 학습은, 가능한 보상의 최대 기대치를 추정하기 때문에 경로 생성을 위해 다음과 같이 음의 부호로 보상 함수를 정의한다.

$$r(s, a) \cong r(s) = \begin{cases} -p(m(s)), & s \neq g \\ 1, & s = g \end{cases}$$

이에 따라, 목표 지점, g에 도달하기 전에는 격자 상 위치 s에 해당하는 장애물 지도의 음수 확률 값을 보상 함수로 사용하고, 목표 지점에서는 최대 보상 값인 1을 가지게 된다. 또한, 벽면에서는 최저, 빈 공간에서는 확률값에 의해 벽보다 더 낮은 보상을 받고, 목표점에서 최대의 보상을 받도록 설계하였다.

또한 장애물 격자와 강화 학습의 조합은, [그림 5]의 구성으로 실시간 사용된다. 목표 위치가 주어지면, 강화 학습에 따른 최단 및 회피 경로를 생성하고, 이를 실제로 로봇 경로에 적용하는 방식이다.

강화 학습의 결과물은 [그림 6]에 명시되어 있다. 장애물 발생 빈도가 높은 지역의 상태 함수값이 [그림 3]과 비교하여 보다 더 넓게 분포되어 있다, 새로 장애물이 등장하여 이동하는 경우, 역 측정 모델에 의해 초기에는 약간 높은 확률로 변경되나, 지도 수정 과정을 거치며 해당 지역의 확률이 점차로 증가하고 이는 강화 학습 상의 음의 보상 함수로 고려되어 해당 지역을 회피하는 경로를 만들게 된다.

실제 로봇에 적용 시의 시간 지연, 격자 지도의 수정 빈도가 중요한 요소이다. [그림 7]은 실제 로봇에 적용 시의 장애물 회피 사례이다. 제안된 방식은 장애물 격자 지도의 수정과 동시에 실시간 강화 학습을 적용하기 때문에, 기존 방식 대비 연산량이 다소 높다.

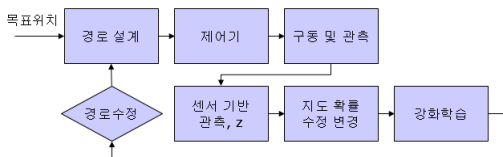


그림 5. 장애물 지도 및 강화학습을 연계한 회피 방식

또한, 파티클 필터를 이용하여 초당 10회의 공간 위치 추정을 수행하고 있기에, 장애물 출현에 따른 최대 초당 10회의 확률 수정이 가능하다. 그러나, 연산량은

고려하여, 지도의 수정 결과에 큰 차이가 있을 때만 최대 초당 1회 이하의 강화학습을 수행하도록 강제하여 적은 연산량으로도 회피 주행이 가능하게 구성하였다.

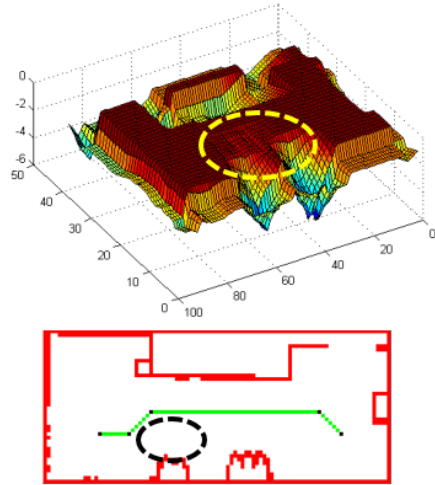


그림 6. 저해상도의 장애물 감지용 지도 작성을 통해, 장애물 출몰지역에 대한 강화학습 결과 및 생성 회피 경로



그림 7. 실 환경에서 불특정 장애물을 회피하는 테스트 장면

## V. 결론

본 논문에서는 산업용 자율 주행 로봇을 위한 장애물 회피 기법의 개발 사례를 다루고 있다. 이를 위해, 두 가지의 서로 다른 방법을 통해, 연산량의 특징과 작업 공간에 대한 의존성을 설명하려 한다.

기존 연구의 정적 및 동적 물체 구별은, 적은 연산량을 가지나 위치 정렬 상태의 검증 및 위치 추정 결과에

민감하여 정렬 오류에 대한 보정이 필수적이다. 따라서, 진행각이 큰 코너 주행 시에 장애물 감지에 문제점을 가지며, 장애물의 크기를 정확히 예측하기 어렵다.

반면, 제안된 장애물 판단을 위한 저해상도의 격자 지도를 사용하는 경우, 장애물의 위치 판단 및 크기 예측, 위치 추정치에 민감하지 않은 장점을 갖지만 지도 작성 및 강화 학습의 잦은 적용에 따른 연산량 증가의 단점을 갖는다.

두 가지 방법 중, 전자의 경우는 사무실과 같은 좁은 환경에서 적합하다. 협소한 공간적 한계 때문에 거리 센서의 최대 유효거리를 제한하여 정렬 오차 발생을 억제하는 경우를 의미한다. 후자의 경우는, 공장, 홀과 같은 넓은 공간에 적합한 방식이다. 실험 결과, 장애물 크기에 대한 예측이 가능하여 장애물 주위를 자연스럽게 회피하는 경우의 성능 향상이 우수하였다.

장애물 회피는, SLAM 기반의 자율 주행과 같은 정해진 경로를 반복 추종하는 환경에서 복잡성이 크게 요구되는 분야로, 안전 사고 예방 등 해당 기술의 중요도가 높기 때문에 향후 제안된 방법론의 하이브리드 적용에 대한 추가 연구가 필요할 것으로 사료된다.

**참 고 문 헌**

[1] S. Thrun, "A Probabilistic Approach to Concurrent Mapping and Localization for Mobile Robots," *Machine Learning and Autonomous Robots*, Vol.31, No.5, pp.1-25, 1998

[2] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, pp.171-172, MIT Press, 2005.

[3] D. Holz, D. Droschel, S. Behnke, M. May, and H. Surmann, "Fast 3D perception for collision avoidance and SLAM in domestic environment," In *Mobile robots navigation*, In-Tech Education and Publishing, Vienna, Austria, 2010.

[4] G. Grisetti, C. Stachniss, and W. Burgard, "Improved Technique for Grid Mapping with Rao-Blackwellization Particle Filters," *IEEE Trans. on Robotics (T-RO)*, Vol.23, pp.34-46, 2007.

[5] J. Civera, A. J. Davison, and J. M. M. Montiel, "Inverse depth parametrization for monocular SLAM," *IEEE Trans. Robot.*, Vol.24, No.5, pp.932-945, 2008(10).

[6] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse, "MonoSLAM: Real-time single camera SLAM," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.29, No.6, pp.1052-1067, 2007(6).

[7] D. Marzorati, M. Matteucci, D. Migliore, and D. G. Sorrenti, "Monocular SLAM with inverse scaling parametrization," In *Proc. of the British machine vision conference*, Leeds, pp.945-994, 2008.

[8] J. Sola, T. Vidal-Calleja, J. Civera, and J. M. M. Montiel, "Impact of landmark parametrization on monocular EKF-SLAM with points and lines," *International Journal of Computer Vision*, Vol.97, pp.339-368, 2012.

[9] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, Cambridge Univ. Press, pp.157-159, 1998.

[10] J. Roberts, I. Manchester, and R. Tedrake, "Feedback controller parameterizations for reinforcement learning," In *2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, 2011.

**저 자 소 개**

양 정 연(Jeong-Yean Yang)

정희원



- 2000년 2월 : KAIST 기계공학 과(공학석사)
- 2011년 8월 : KAIST 기계공학 과(공학박사)
- 2016년 3월 ~ 현재 : 목원대학교 지능로봇공학과 교수

<관심분야> : 지능 로봇