



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2021년06월01일

(11) 등록번호 10-2259932

(24) 등록일자 2021년05월27일

(51) 국제특허분류(Int. Cl.)  
*G06N 3/08* (2006.01) *B60R 21/0134* (2006.01)  
*B60W 30/08* (2006.01) *B60W 30/14* (2006.01)  
*B60W 40/02* (2006.01)  
(52) CPC특허분류  
*G06N 3/084* (2013.01)  
*B60R 21/0134* (2013.01)  
(21) 출원번호 10-2019-0102225  
(22) 출원일자 2019년08월21일  
심사청구일자 2019년08월21일  
(65) 공개번호 10-2021-0022888  
(43) 공개일자 2021년03월04일  
(56) 선행기술조사문헌  
KR1020190076628 A\*  
이홍석 외 2명, “자율주행을 위한 멀티에이전트  
심화 강화학습”, KIISE Transactions on  
Computing Practices, Vol. 24, No. 12, pp.  
670-674(2018.12.31.) 국내논문 사본 1부.\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
**연세대학교 산학협력단**  
서울특별시 서대문구 연세로 50 (신촌동, 연세대  
학교)  
(72) 발명자  
**김신덕**  
서울특별시 서대문구 수색로 100, 212동 804호(북  
가좌동, DMC래미안e편한세상)  
**신새벽**  
경기도 군포시 삼성로20번길 8-15, 302호 (   
부곡동)  
(74) 대리인  
**민영준**

전체 청구항 수 : 총 5 항

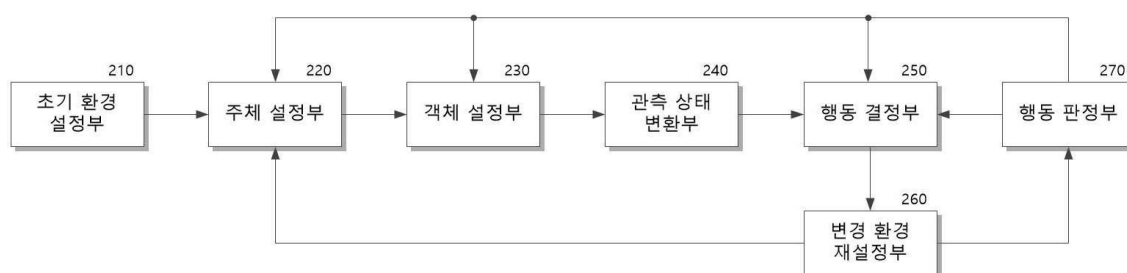
심사관 : 송근배

(54) 발명의 명칭 자율 이동체를 위한 사고 유발자 공동 학습 장치 및 방법

### (57) 요약

본 발명은 기설정된 시뮬레이션 환경에서 자율 이동체인 주체와 우호적 객체 및 적대적 객체를 포함하는 다수의 객체의 배치 위치를 설정하는 배치 설정부, 상기 주체의 관점에서 상기 주체와 다수의 객체가 배치된 시뮬레이션 환경을 관측한 관측 상태 정보를 획득하는 관측 상태 변환부, 학습되는 패턴에 따라 상기 관측 상태 정보에 기반 (뒷면에 계속)

### 대표도



하여 상기 주체의 동작을 결정하고, 상기 주체의 동작과 다른 객체의 동작에 기반하여 다수의 객체의 각각의 동작을 결정하는 행동 결정부, 상기 주체 및 다수의 객체 각각의 결정된 동작 결과를 상기 시뮬레이션 환경에 반영하여 상기 시뮬레이션 환경을 변경하는 변경 환경 재설정부 및 상기 변경된 시뮬레이션 환경을 분석하여 사고 피해에 대비례하는 보상을 계산하고, 계산된 보상을 상기 주체 및 상기 우호적 객체에 대해서는 동일한 부호로 역전과하고, 상기 적대적 객체에 대해서는 반대 부호로 역전과하여 상기 행동 결정부를 학습시키는 행동 판정부 를 포함하는 자율 이동체를 위한 학습 장치 및 방법을 제공할 수 있다.

(52) CPC특허분류

**B60W 30/08** (2013.01)

**B60W 30/14** (2013.01)

**B60W 40/02** (2013.01)

**B60W 2554/00** (2020.02)

**B60Y 2300/08** (2013.01)

**B60Y 2300/14** (2013.01)

## 명세서

### 청구범위

#### 청구항 1

기설정된 시뮬레이션 환경에서 자율 이동체인 주체와 우호적 객체 및 적대적 객체를 포함하는 다수의 객체의 배치 위치를 설정하는 배치 설정부;

상기 주체의 관점에서 상기 주체와 다수의 객체가 배치된 시뮬레이션 환경을 관측한 관측 상태 정보를 획득하는 관측 상태 변환부;

학습되는 패턴에 따라 상기 관측 상태 정보에 기반하여 상기 주체의 동작을 결정하고, 상기 주체의 동작과 다른 객체의 동작에 기반하여 다수의 객체의 각각의 동작을 결정하는 행동 결정부;

상기 주체 및 다수의 객체 각각의 결정된 동작 결과를 상기 시뮬레이션 환경에 반영하여 상기 시뮬레이션 환경을 변경하는 변경 환경 재설정부; 및

상기 변경된 시뮬레이션 환경을 분석하여 사고 피해에 대비례하는 보상을 계산하고, 계산된 보상을 상기 주체 및 상기 우호적 객체에 대해서는 동일한 부호로 역전파하고, 상기 적대적 객체에 대해서는 반대 부호로 역전파하여 상기 행동 결정부를 학습시키는 행동 판정부를 포함하되,

상기 행동 결정부는

상기 행동 판정부에서 기지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 관측 상태 정보에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 최소화되도록 상기 주체의 동작을 결정하는 주체 행동 결정부;

상기 행동 판정부에서 기지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 시뮬레이션 환경과 상기 주체 및 다른 객체의 동작에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 최소화되도록 상기 우호적 객체의 동작을 결정하는 우호적 객체 행동 결정부; 및

상기 행동 판정부에서 부호가 반대인 기지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 시뮬레이션 환경과 상기 주체 및 다른 객체의 동작에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 증가되도록 상기 적대적 객체의 동작을 결정하는 적대적 객체 행동 결정부를 포함하는 자율 이동체를 위한 학습 장치.

#### 청구항 2

삭제

#### 청구항 3

삭제

#### 청구항 4

제1 항에 있어서, 상기 행동 결정부는

상기 시뮬레이션 환경의 기지정된 위치에 배치되어 상기 주체 및 다른 객체의 동작에 무관하게, 고정되거나 미리 지정된 동작을 수행하는 중립적 객체의 동작을 결정하는 중립적 객체 행동 결정부를 더 포함하는 자율 이동체를 위한 학습 장치.

#### 청구항 5

제1 항에 있어서, 상기 배치 설정부는

상기 시뮬레이션 환경에 배치되는 다수의 객체 중 상기 적대적 객체의 비율이 기지정된 비율 이하가 되도록 배치하는 자율 이동체를 위한 학습 장치.

## 청구항 6

초기 환경 정보를 획득하여 시뮬레이션 환경을 설정하는 단계;

설정된 시뮬레이션 환경에서 자율 이동체인 주체와 우호적 객체 및 적대적 객체를 포함하는 다수의 객체의 배치 위치를 설정하는 단계;

상기 주체의 관점에서 상기 주체와 다수의 객체가 배치된 시뮬레이션 환경을 관측한 관측 상태 정보를 획득하는 단계;

학습되는 패턴에 따라 상기 관측 상태 정보에 기반하여 상기 주체의 동작을 결정하고, 상기 주체의 동작과 다른 객체의 동작에 기반하여 다수의 객체의 각각의 동작을 결정하는 단계;

상기 주체 및 다수의 객체 각각의 결정된 동작 결과를 상기 시뮬레이션 환경에 반영하여 상기 시뮬레이션 환경을 변경하는 단계; 및

상기 변경된 시뮬레이션 환경을 분석하여 사고 피해에 반비례하는 보상을 계산하고, 계산된 보상을 상기 주체 및 상기 우호적 객체에 대해서는 동일한 부호로 역전파하고, 상기 적대적 객체에 대해서는 반대 부호로 역전파하는 단계를 포함하되,

상기 동작을 결정하는 단계는

기 지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 관측 상태 정보에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 최소화되도록 상기 주체의 동작을 결정하는 단계;

기 지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 시뮬레이션 환경과 상기 주체 및 다른 객체의 동작에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 최소화되도록 상기 우호적 객체의 동작을 결정하는 단계; 및

부호가 반대인 기 지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 시뮬레이션 환경과 상기 주체 및 다른 객체의 동작에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 증가되도록 상기 적대적 객체의 동작을 결정하는 단계를 포함하는 자율 이동체를 위한 학습 방법.

## 청구항 7

삭제

## 청구항 8

삭제

## 청구항 9

삭제

## 청구항 10

제6 항에 있어서, 동작을 결정하는 단계는

상기 시뮬레이션 환경의 기 지정된 위치에 배치되어 상기 주체 및 다른 객체의 동작에 무관하게, 고정되거나 미리 지정된 동작을 수행하는 독립적 객체의 동작을 결정하는 단계를 더 포함하는 자율 이동체를 위한 학습 방법.

## 발명의 설명

## 기술 분야

[0001] 본 발명은 자율 이동체를 위한 학습 장치 및 방법에 관한 것으로, 자율 이동체를 위한 사고 유발자 공동 학습 장치 및 방법에 관한 것이다.

## 배경 기술

- [0002] 근래에 자율주행 자동차 등 유/무인 자율 이동체의 자율 행동 인공지능 개발이 지속적으로 이루어지고 있으나, 사고 사례가 지속적으로 등장하고 있다. 특히 사람이 회피할 수 있었던 사고 상황에 대한 책임 소지가 지속적으로 논란이 되고 있다.
- [0003] 도 1은 기존의 자율 이동체를 위한 학습 장치의 개략적 구조를 나타낸다.
- [0004] 도 1에 도시된 바와 같이, 자율 이동체를 학습시키기 위해 시뮬레이션을 수행하는 학습 장치는 초기 환경 설정부(110), 관측 상태 변환부(120), 행동 결정부(130) 및 변경 환경 재설정부(140)를 포함한다.
- [0005] 초기 환경 설정부는 자율 이동체가 이동해야하는 초기 환경 정보를 획득하여 설정한다. 즉 자율 이동체가 활동하는 초기의 주변 환경 정보를 설정한다. 그리고 관측 상태 변환부(120)는 초기 환경 설정부(110)에서 설정된 초기 환경 정보를 자율 이동체가 관측한 상태 정보로 변환한다. 초기 환경 설정부(110)에서 설정된 초기 환경 정보는 자율 이동체가 활동을 수행하는 전체 주변 환경에 대한 정보이다. 그리고 자율 이동체가 실제로 동작하는 상태에서는 주변 환경에 대한 모든 정보를 획득하는 것은 불가능하며, 주변 환경에서 관측으로 획득되는 제한된 관측 정보를 기반으로 동작을 수행한다. 이에 관측 상태 변환부(120)는 자율 이동체의 실제 동작을 모의해야 하는 시뮬레이션에도 초기 환경 설정부(110)에서 제공된 초기 환경 정보를 자율 이동체가 관측하여 획득한 관측 정보의 형태로 변환한다.
- [0006] 행동 결정부(130)는 자율 이동체에 설치되어 자율 이동체의 동작을 제어하는 구성으로, 관측 상태 변환부(120)에서 획득한 관측 정보를 기반으로 자율 이동체의 동작을 결정한다.
- [0007] 변경 환경 재설정부(140)는 행동 결정부(130)에 의해 결정되어 시뮬레이션 내에서 수행되는 자율 이동체의 동작에 기반하여, 자율 이동체의 주변 환경에 대한 환경 정보를 변경하여 재설정한다. 그리고 재설정된 환경 정보는 관측 상태 변환부(120)로 다시 제공하여 자율 이동체의 관측 정보를 변경한다.
- [0008] 또한 변경 환경 재설정부(140)는 자율 이동체의 동작을 평가하고, 평가 결과를 행동 결정부(130)로 전달하여, 행동 결정부(130)가 결정한 동작의 적절성을 판단하고 이후 동일한 동작으로 선택할지 다른 동작을 탐색할지 학습할 수 있도록 한다.
- [0009] 상기한 바와 같은 기존의 자율 이동체 학습의 가장 큰 문제점은 초기 환경 정보의 설정이다. 기존의 자율 이동체 학습에서 초기 환경은 주로 실제 관측된 데이터셋을 기반으로 생성된다. 그러나 이렇게 관측된 데이터 셋에서는 돌발적 상황에 대한 정보가 포함되기 어렵다. 일례로 자율 주행 차량에 대한 초기 환경 정보를 설정하기 위한 데이터 셋을 획득할 때, 대부분 다른 차량의 운전자나 보행자와 같은 객체가 기본적으로 교통 질서를 준수하여 올바른 행동을 수행할 뿐만 아니라, 사고 발생 가능성이 있는 경우, 사고를 회피하는 방향의 행동을 수행한다. 즉 의도적이든 비의도적이든 객체가 사고를 유발하는 방식으로 행동을 수행하는 돌발 상황에 대한 정보가 초기 환경 정보에 포함되지 않는다.
- [0010] 그리고 이러한 초기 환경 정보를 기반으로 학습이 수행된 행동 결정부(130)를 기반으로 동작을 수행하는 자율 이동체는 인간의 입장에서는 손쉽게 대처 가능한 돌발 상황에 대해서도 제대로 대처하지 못하는 문제점이 발생하게 된다.

## 선행기술문헌

### 특허문헌

- [0011] (특허문헌 0001) 한국 공개 특허 제10-2019-0054374호 (2019.05.22 공개)

## 발명의 내용

### 해결하려는 과제

- [0012] 본 발명의 목적은 자율 이동체가 다양한 돌발 상황에서도 적합한 동작을 수행하도록 학습시킬 수 있는 자율 이동체를 위한 학습 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0013] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 자율 이동체를 위한 학습 장치는 기설정된 시뮬레이션

환경에서 자율 이동체인 주체와 우호적 객체 및 적대적 객체를 포함하는 다수의 객체의 배치 위치를 설정하는 배치 설정부; 상기 주체의 관점에서 상기 주체와 다수의 객체가 배치된 시뮬레이션 환경을 관측한 관측 상태 정보를 획득하는 관측 상태 변환부; 학습되는 패턴에 따라 상기 관측 상태 정보에 기반하여 상기 주체의 동작을 결정하고, 상기 주체의 동작과 다른 객체의 동작에 기반하여 다수의 객체의 각각의 동작을 결정하는 행동 결정부; 상기 주체 및 다수의 객체 각각의 결정된 동작 결과를 상기 시뮬레이션 환경에 반영하여 상기 시뮬레이션 환경을 변경하는 변경 환경 재설정부; 및 상기 변경된 시뮬레이션 환경을 분석하여 사고 피해에 반비례하는 보상을 계산하고, 계산된 보상을 상기 주체 및 상기 우호적 객체에 대해서는 동일한 부호로 역전파하고, 상기 적대적 객체에 대해서는 반대 부호로 역전파하여 상기 행동 결정부를 학습시키는 행동 판정부부를 포함한다.

[0014] 상기 행동 결정부는 상기 행동 판정부에서 기지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 관측 상태 정보에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 최소화되도록 상기 주체의 동작을 결정하는 주체 행동 결정부; 상기 행동 판정부에서 기지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 시뮬레이션 환경과 상기 주체 및 다른 객체의 동작에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 최소화되도록 상기 우호적 객체의 동작을 결정하는 우호적 객체 행동 결정부; 및 상기 행동 판정부에서 부호가 반대인 기지정된 가중치가 가중되어 역전파되는 보상에 따라 학습되고, 상기 시뮬레이션 환경과 상기 주체 및 다른 객체의 동작에 기반하여 상기 시뮬레이션 환경에서 사고 피해가 증가되도록 상기 적대적 객체의 동작을 결정하는 적대적 객체 행동 결정부를 포함할 수 있다.

[0015] 상기 행동 판정부는 상기 행동 결정부에서 주체와 우호적 객체 및 적대적 객체가 특정 행동을 결정할 확률을 나타내는 확률 함수를 인가받고, 주체가 수행하는 행동과 객체가 수행하는 행동에 의한 결과에 따라 미리 지정된 보상 함수와 상기 확률 함수의 곱의 적분값으로 나타나는 주체 및 우호적 객체 행동의 기대값을 계산하고, 상기 기대값이 최대가 되도록 하는 주체와 우호적 객체의 행동을 상기 행동 결정부로 전달하고, 상기 보상 함수의 부호를 반전하여 계산되는 적대적 객체 행동의 기대값이 최대가 되도록 하는 적대적 객체의 행동을 상기 행동 결정부로 전달할 수 있다.

[0016] 상기 행동 결정부는 상기 시뮬레이션 환경의 기지정된 위치에 배치되어 상기 주체 및 다른 객체의 동작에 무관하게, 고정되거나 미리 지정된 동작을 수행하는 중립적 객체의 동작을 결정하는 중립적 객체 행동 결정부를 더 포함할 수 있다.

[0017] 상기 배치 설정부는 상기 시뮬레이션 환경에 배치되는 다수의 객체 중 상기 적대적 객체의 비율이 기지정된 비율 이하가 되도록 배치할 수 있다.

[0018] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 자율 이동체를 위한 학습 방법은 초기 환경 정보를 획득하여 시뮬레이션 환경을 설정하는 단계; 설정된 시뮬레이션 환경에서 자율 이동체인 주체와 우호적 객체 및 적대적 객체를 포함하는 다수의 객체의 배치 위치를 설정하는 단계; 상기 주체의 관점에서 상기 주체와 다수의 객체가 배치된 시뮬레이션 환경을 관측한 관측 상태 정보를 획득하는 단계; 학습되는 패턴에 따라 상기 관측 상태 정보에 기반하여 상기 주체의 동작을 결정하고, 상기 주체의 동작과 다른 객체의 동작에 기반하여 다수의 객체의 각각의 동작을 결정하는 단계; 상기 주체 및 다수의 객체 각각의 결정된 동작 결과를 상기 시뮬레이션 환경에 반영하여 상기 시뮬레이션 환경을 변경하는 단계; 및 상기 변경된 시뮬레이션 환경을 분석하여 사고 피해에 반비례하는 보상을 계산하고, 계산된 보상을 상기 주체 및 상기 우호적 객체에 대해서는 동일한 부호로 역전파하고, 상기 적대적 객체에 대해서는 반대 부호로 역전파하는 단계를 포함한다.

### 발명의 효과

[0019] 따라서, 본 발명의 실시예에 따른 자율 이동체를 위한 학습 장치 및 방법은 주체에게 우호적인 우호적 객체뿐만 아니라 사고를 유발하는 적대적 객체를 포함하고, 주체와 함께 우호적 객체 및 적대적 객체를 공동 학습시켜, 자율 이동체가 다양한 돌발 상황에서도 적합한 동작을 수행하도록 할 수 있다.

### 도면의 간단한 설명

[0020] 도 1은 기존의 자율 이동체를 위한 학습 장치의 개략적 구조를 나타낸다.

도 2는 본 발명의 일 실시예에 따른 자율 이동체를 위한 학습 장치의 개략적 구조를 나타낸다.

도 3은 도 2의 객체 설정부의 상세 구성을 나타낸다.

도 4는 도 2의 행동 결정부의 상세 구성을 나타낸다.



도 5는 본 발명의 일 실시예에 따른 자율 이동체를 위한 학습 방법을 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

- [0021] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0022] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.
- [0023] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0024] 도 2는 본 발명의 일 실시예에 따른 자율 이동체를 위한 학습 장치의 개략적 구조를 나타내고, 도 3은 도 2의 객체 설정부의 상세 구성을 나타내며, 도 4는 도 2의 행동 결정부의 상세 구성을 나타낸다.
- [0025] 도 2를 참조하면, 본 실시예에 따른 자율 이동체를 위한 학습 장치는 초기 환경 설정부(210), 주체 설정부(220), 객체 설정부(230), 관측 상태 변환부(240), 행동 결정부(250), 변경 환경 재설정부(260) 및 행동 관정부(270)를 포함한다.
- [0026] 초기 환경 설정부(210)는 도 1의 초기 환경 설정부(110)와 마찬가지로, 자율 이동체가 이동해야 하는 초기 환경 정보를 획득하여 시뮬레이션 환경을 설정한다. 초기 환경 설정부(210)는 자율 이동체가 이동하는 시뮬레이션을 수행하기 위한 주변 환경에 대한 초기 정보를 획득하고, 획득된 초기 환경 정보 따라 초기 시뮬레이션 환경을 생성한다. 일례로 자율 이동체가 자율 주행차인 경우, 초기 환경 설정부(210)는 자율 주행 차량이 운행되어야 하는 도로의 상태, 차선 개수 및 기후 정보 등과 같은 도로 상황에 대한 초기 시뮬레이션 환경을 생성할 수 있다.
- [0027] 주체 설정부(220)는 초기 환경 설정부(210)에 의해 생성된 초기 시뮬레이션 환경에서 주체인 자율 이동체의 위치를 설정하여 배치한다. 즉 초기 시뮬레이션 환경 내에서 주체의 초기 위치를 설정한다.
- [0028] 객체 설정부(230)는 주체가 배치된 초기 시뮬레이션 환경에서 다수의 객체의 위치를 설정하여 배치한다. 여기서 객체는 자율 이동체인 주체의 이동을 방해할 수 있는 모든 종류의 장애물을 의미한다. 특히 본 실시예에서 객체 설정부(230)는 다수의 객체를 객체 특성을 고려하여 구분하고, 구분된 객체 각각을 독립적으로 배치할 수 있다.
- [0029] 객체 설정부(230)는 도 3에 도시된 바와 같이, 우호적 객체 배치부(231), 중립적 객체 배치부(232) 및 적대적 객체 배치부(233)를 포함하여, 다수의 객체를 적대적 객체, 중립적 객체 및 우호적 객체로 분류하고, 분류된 적대적 객체, 중립적 객체 및 우호적 객체들의 위치를 서로 다른 방식으로 설정하여 배치할 수 있다.
- [0030] 여기서 우호적 객체는 주체인 자율 이동체의 이동에 우호적인 행동을 취하게 되는 객체로서 주체의 동작에 응답하여 가능한 사고가 발생되지 않도록 동작하는 객체를 의미하며, 일반적으로 이동 또는 행동 가능한 모든 종류의 객체를 포함할 수 있다. 일례로 자율 이동체가 차량인 경우, 주체와 동일 차선의 전후에서 주행하는 차량이나 인접 차선을 주행하는 차량과 보도 또는 횡단 보도를 보행하는 보행자, 이동 중인 자전거 등과 같은 다양한 객체가 우호적 객체로 포함될 수 있다.
- [0031] 그리고 중립적 객체는 주체의 이동에 무관한 객체들로서 이동되지 않는 각종 장애물과 주체의 움직임을 고려하지 않고 완전히 독립적으로 움직이는 객체들이 포함될 수 있다. 일례로 건물이나 가로수, 가로등과 같이 주체의 움직임에 무관하게 초기에 배치된 위치에 고정되는 객체가 중립적 객체에 포함될 수 있다. 또한 경우에 따라서는 도로 공사 장치와 같이 이동 또는 행동 가능한 객체일지라도 주변 환경을 감지하지 않고 움직여서 주체의 이동에 응답하지 않고 지정된 동작을 수행하는 객체가 중립적 객체에 포함될 수 있다.
- [0032] 이러한 우호적 객체와 중립적 객체는 기존의 자율 이동체를 위한 학습 장치에서 이용되는 객체 중 이동 또는 동작 가능한 객체와 위치가 고정된 객체들과 동일할 수 있다.

- [0033] 마지막으로 적대적 객체는 주체의 움직임에 대해 응답하여 동작하지만, 우호적 객체와 달리 사고가 발생되도록 적대적 행동을 취하는 객체를 의미한다.
- [0034] 일반적으로 대부분의 객체는 우호적 객체 또는 중립적 객체로 포함된다. 그러나 현실의 상황에서도 보복 운전이나 음주 운전, 운전자의 착오, 운전 조작 실수와 같은 다양한 이유로 인해, 비록 그 빈도가 낮다고는 하나 적대적으로 행동하는 객체가 존재한다. 이러한 적대적 객체는 기존의 자율 이동체를 위한 학습 장치에서는 고려되지 않은 객체이므로, 기존의 자율 이동체를 위한 학습 장치에 의해 학습된 자율 이동체는 사람이 조작하는 경우에 간단하게 회피 가능한 사고일지라도 회피하지 못하고 사고를 유발할 수 있다.
- [0035] 이에 본 실시예의 객체 설정부(230)는 우호적 객체 배치부(231), 중립적 객체 배치부(232) 뿐만 아니라 적대적 객체 배치부(233)를 구비하여, 주체에 적대적으로 행동하는 객체를 시뮬레이션 환경에 배치한다. 다만 현실에서 적대적 객체가 존재할 확률은 매우 희박하며, 이에 객체 설정부(230)의 우호적 객체 배치부(231), 중립적 객체 배치부(232) 및 적대적 객체 배치부(233)는 미리 지정된 비율에 따라 우호적 객체와 중립적 객체 및 적대적 객체의 수를 조절하여 시뮬레이션 환경에 배치할 수 있으며, 적대적 객체의 비율(예를 들면, 1% 이하)을 우호적 객체와 중립적 객체에 비해 매우 작게 설정할 수 있다.
- [0036] 또한 적대적 객체의 비율을 작게 설정하는 것은 자율 이동체를 위한 학습 장치가 현실적으로 가능한 시뮬레이션을 수행할 수 있도록 하기 위함이다. 적대적 객체의 비율이 높아진다는 것은 주체의 이동을 방해하는 방향으로 행동하는 객체의 수가 증가하는 것을 의미한다. 그리고 제한된 공간에서 제한된 형태로만 동작 가능한 주체가 이동하는 동안 적대적 객체의 개수가 증가하여 주체의 이동을 방해하게 되면, 현실적으로 주체가 이동이 불가능해지는 상황이 빈번하게 발생하게 된다. 그리고 이러한 상황은 시뮬레이션을 통한 학습 시간을 크게 증가시킬 뿐만 아니라 정상적인 학습이 수행되지 못하도록 한다. 즉 학습이 완료되지 못하도록 하는 요인이 될 수 있다. 이에 적대적 객체의 비율은 현실에서 발생한 사고 발생 유형 중 적대적 사고 발생 비중을 분석하여 설정할 수 있다.
- [0037] 여기서는 설명의 편의를 위하여 주체 설정부(220)와 객체 설정부(230)를 구분하여 도시하였으나, 주체 설정부(220)와 객체 설정부(230)는 배치 설정부로 병합되어도 무방하다.
- [0038] 객체 설정부(230)에 의해 시뮬레이션 환경에 우호적 객체와 중립적 객체 및 적대적 객체가 배치되면, 관측 상태 변환부(240)는 주체가 배치된 시뮬레이션 환경을 주체의 관점에서 관측한 주체 관측 정보의 형태로 변환한다. 즉 주체가 획득 가능한 시뮬레이션 환경의 정보로 변환한다. 이때, 시뮬레이션 환경에 배치된 우호적 객체와 중립적 객체 및 적대적 객체에 대한 관측 정보를 함께 획득한다. 다만 주체가 시뮬레이션 환경의 모든 정보를 관측할 수 없는 것과 마찬가지로 다수의 객체 중 주체가 관측 가능한 객체에 대한 정보만을 관측 정보에 포함시킬 수 있다.
- [0039] 경우에 따라서 관측 상태 변환부(240)는 우호적 객체 및 적대적 객체의 관점에서 관측되는 우호적 객체 관측 정보와 적대적 객체 관측 정보를 더 획득할 수도 있다.
- [0040] 행동 결정부(250)는 시뮬레이션 환경에 배치된 주체 및 다수의 객체 각각의 행동을 결정한다. 본 실시예에서 행동 결정부(250)는 주체의 행동뿐만 아니라 다수의 객체의 행동을 함께 결정할 수 있다. 즉 우호적 객체와 중립적 객체 및 적대적 객체의 행동을 함께 결정할 수 있다.
- [0041] 도 4를 참조하면, 행동 결정부(250)는 주체 행동 결정부(251), 우호적 객체 행동 결정부(252), 중립적 객체 행동 결정부(253) 및 적대적 객체 행동 결정부(254)를 포함할 수 있다.
- [0042] 주체 행동 결정부(251)는 관측 상태 변환부(240)에서 획득된 관측 정보에 기초하여 주체의 행동을 결정한다. 관측 정보에는 시뮬레이션 환경 정보뿐만 아니라, 관측된 우호적 객체와 중립적 객체 및 적대적 객체에 대한 정보가 포함되어 있으므로, 주체 행동 결정부(251)는 관측된 모든 시뮬레이션 환경과 객체에 대한 정보를 고려하여 주체의 행동을 결정한다. 여기서 주체 행동 결정부(251)는 가능한 사고 발생을 회피하고, 불가피하게 사고가 발생되더라도 피해를 최소화될 수 있도록 주체의 행동을 결정한다.
- [0043] 그리고 우호적 객체 행동 결정부(252)는 시뮬레이션 환경 정보와 주체 행동 결정부(251)에서 결정된 주체의 행동, 다른 우호적 객체의 행동과 중립적 객체 및 적대적 객체의 행동에 기초하여 우호적 객체의 행동을 결정한다. 우호적 객체 행동 결정부(252) 또한 주체 행동 결정부(251)와 유사하게 가급적 사고를 회피하거나 사고로 인한 피해가 최소화되도록 우호적 객체의 행동을 결정한다.
- [0044] 그리고 중립적 객체 행동 결정부(253)는 중립적 객체 배치부(232)가 시뮬레이션 환경에 배치한 중립적 객체 중



이동 가능한 객체의 행동을 결정한다. 상기한 바와 같이 중립적 객체는 주체나 다른 객체의 행동에 영향을 받지 않는 객체이므로 미리 지정된 패턴의 행동을 수행하도록 결정될 수 있다. 또한 중립적 객체 중 이동 가능한 중립적 객체가 존재하지 않는 경우, 중립적 객체 행동 결정부(253)는 생략될 수 있다.

- [0045] 적대적 객체 행동 결정부(254)는 우호적 객체 행동 결정부(252)와 유사하게 시뮬레이션 환경 정보와 주체 행동 결정부(251)에서 결정된 주체의 행동과 우호적 객체, 중립적 객체 및 다른 적대적 객체의 행동에 기초하여 적대적 객체의 행동을 결정한다. 그러나 적대적 객체 행동 결정부(254)는 우호적 객체 행동 결정부(252)와 반대로, 가급적 사고를 유발하도록, 그리고 사고로 인한 피해가 극대화되도록 적대적 객체의 행동을 결정한다.
- [0046] 변경 환경 재설정부(260)는 행동 결정부(250)에서 결정된 주체와 우호적 객체, 중립적 객체 및 적대적 객체의 행동 결과를 시뮬레이션 환경에 반영하여 변경하고, 변경된 시뮬레이션 환경 정보를 주체 설정부(220)와 행동 판정부(270)로 전달한다.
- [0047] 행동 판정부(270)는 변경 환경 재설정부(260)로부터 변경된 시뮬레이션 환경 정보를 인가받아 분석하여, 사고 발생 여부와 사고 발생으로 인한 피해 정도를 분석하고, 분석 결과에 따른 보상을 주체 설정부(220)와 객체 설정부(230) 및 행동 결정부(250)로 전달한다.
- [0048] 행동 판정부(270)는 사고 발생으로 인한 피해 수준을 기지정된 방식으로 수치적으로 계산하고, 계산된 피해 수준에 반비례하는 보상을 결정하여 주체 설정부(220)와 객체 설정부(230) 및 행동 결정부(250)로 역전파할 수 있다.
- [0049] 본 실시예에서 주체 설정부(220)와 객체 설정부(230)의 우호적 객체 배치부(231), 적대적 객체 배치부(233) 및 행동 결정부(250)의 주체 행동 결정부(251)와 우호적 객체 행동 결정부(252) 및 적대적 객체 행동 결정부(254)는 학습이 가능한 인공 신경망으로 구현될 수 있다.
- [0050] 즉 학습에 의해 주체와 우호적 객체 및 적대적 객체의 배치 위치 및 행동이 조절될 수 있다. 이에 행동 판정부(270)는 인공 신경망으로 구현되는 주체 설정부(220)와 우호적 객체 배치부(231), 적대적 객체 배치부(233), 주체 행동 결정부(251), 우호적 객체 행동 결정부(252) 및 적대적 객체 행동 결정부(254)가 정상적으로 학습이 될 수 있도록 보상을 역전파할 수 있다.
- [0051] 이때, 행동 판정부(270)는 주체 설정부(220), 우호적 객체 배치부(231), 주체 행동 결정부(251) 및 우호적 객체 행동 결정부(252)에 대해서 동일 부호의 서로 다른 가중치를 가중하여 보상을 역전파하는 반면, 적대적 객체 배치부(233) 및 적대적 객체 행동 결정부(254)에 대해서는 주체 및 우호적 객체와 반대 부호의 가중치를 가중하여 보상을 역전파 할 수 있다.
- [0052] 즉 주체 설정부(220), 우호적 객체 배치부(231), 주체 행동 결정부(251) 및 우호적 객체 행동 결정부(252)의 경우, 피해가 최소화될수록 높은 보상이 역전파되는 반면, 적대적 객체 배치부(233)와 적대적 객체 행동 결정부(254)에는 패널티가 주어진다.
- [0053] 반대로 피해가 커질수록 주체 설정부(220), 우호적 객체 배치부(231), 주체 행동 결정부(251) 및 우호적 객체 행동 결정부(252)는 패널티를 받는 반면, 적대적 객체 배치부(233)와 적대적 객체 행동 결정부(254)에는 보상이 주어지도록 보상이 역전파된다.
- [0054] 행동 판정부(270)는 행동 결정부(250)에서 주체 및 다수의 객체에 대한 행동이 결정될 때마다 반복적으로 보상을 계산하여 역전파할 수 있으며, 이로 인해, 주체 및 우호적 객체는 사고가 발생되지 않도록 행동하게끔 학습이 수행되는 반면, 적대적 객체는 사고를 발생하도록 학습이 수행된다.
- [0055] 이렇게 적대적 객체가 사고를 발생하도록 학습이 수행됨에 따라 사고를 회피하도록 학습되는 주체 및 우호적 객체는 다양한 돌발 상황에서도 사고를 회피하거나 피해가 최소화되는 행동을 도출할 수 있게 된다.
- [0056] 그리고 학습이 완료된 주체 행동 결정부(251)가 실제의 자율 이동체에 설치됨으로써, 자율 이동체는 현실의 돌발 상황에서도 사고 피해가 최소화되도록 행동할 수 있게 된다.
- [0057] 한편 주체 설정부(220)와 객체 설정부(230)의 우호적 객체 배치부(231), 적대적 객체 배치부(233) 및 행동 결정부(250)의 주체 행동 결정부(251)와 우호적 객체 행동 결정부(252) 및 적대적 객체 행동 결정부(254)가 인공 신경망으로 구현되지 않고 기계 학습이 수행되도록 구성된 경우, 주체 행동 결정부(251)와 우호적 객체 행동 결정부(252) 및 적대적 객체 행동 결정부(254)는 주체와 우호적 객체 및 적대적 객체가 특정 행동을 수행하도록 결정할 확률을 확률 함수( $p_s(x)$ ,  $p_o(x)$ )로 획득할 수 있다.

- [0058] 이 경우 행동 판정부(270)는 주체 및 각 객체가 행동에 대한 보상을 보상 함수의 형태로 정의할 수 있다. 예를 들어 주체가 수행한 행동을  $x$ 라 하고 객체가 수행한 행동을  $y$ 라 할 때, 행동 판정부(270)는 보상 함수( $r(x, y)$ )를 미리 정의할 수 있으며, 주체 및 객체의 행동에 대한 확률 함수( $p_s(x)$ ,  $p_o(x)$ )와 보상 함수의 곱을 적분하여 보상에 대한 주체 및 우호적 객체에 대한 기대값( $E_s(x) = \int r(x, y)p_o(y)dy$ ,  $E_o(y) = \int r(x, y)p_s(x)dx$ )을 계산할 수 있다. 여기서 적대적 객체에 대한 보상 함수는  $(-r(x, y))$ 로 설정될 수 있다.
- [0059] 그리고 기대값이 최대가 되도록 주체 행동 결정부(251)와 우호적 객체 행동 결정부(252) 및 적대적 객체 행동 결정부(254)의 행동을 결정할 수도 있다.
- [0060] 도 5는 본 발명의 일 실시예에 따른 자율 이동체를 위한 학습 방법을 나타낸다.
- [0061] 도 2 내지 도 4를 참조하여, 자율 이동체를 위한 학습 방법을 설명하면, 우선 주체를 학습시키기 위한 초기 환경 정보를 설정한다(S11). 여기서 초기 환경 정보는 주체가 활동해야 하는 주변 환경을 시뮬레이션한 시뮬레이션 환경 정보이다.
- [0062] 초기 환경 정보가 설정되면, 설정된 시뮬레이션 환경 내에 주체를 배치한다(S12). 그리고 시뮬레이션 환경 내에 다수의 객체를 배치한다(S13). 여기서 다수의 객체는 우호적 객체, 중립적 객체 및 적대적 객체가 포함될 수 있다.
- [0063] 시뮬레이션 환경 내에 주체 및 객체가 배치되면, 주체의 관점에서 시뮬레이션 환경과 객체를 관측한 관측 상태 정보를 획득한다(S14). 그리고 관측 상태 정보를 기반으로 주체의 행동을 결정하고, 시뮬레이션 환경에서 주체의 행동과 다른 객체의 행동을 기반으로 각 객체의 행동을 결정한다(S15). 이때, 주체와 객체 중 우호적 객체는 사고를 회피하거나 사고로 인한 피해가 최소화되도록 되도록 행동을 결정한다. 반면, 적대적 객체는 사고를 유발하고 사고의 피해가 증가되도록 행동을 결정한다. 그리고 중립적 객체는 주체나 다른 객체에 무관하게 지정된 방식에 따라 행동하도록 결정될 수 있다.
- [0064] 각 객체의 행동이 결정되면, 결정된 행동을 시뮬레이션 환경에 반영하여 시뮬레이션 환경 정보를 변경한다(S16). 그리고 변경된 시뮬레이션 환경 정보를 분석하여 주체 및 각 객체의 행동 결과를 판정한다(S17). 여기서 각 객체의 행동 결과는 사고 발생으로 인한 피해 수준을 기지정된 방식으로 수치적으로 계산하고, 계산된 피해 수준에 반비례하는 보상을 계산하여 판정될 수 있다.
- [0065] 판정 결과로서 보상이 계산되면, 보상을 역전파하여 주체와 우호적 객체 및 적대적 객체의 배치 위치와 행동이 조절되도록 학습시킨다(S18). 이때 주체와 우호적 객체에는 동일한 부호의 서로 다른 가중치가 가중된 보상을 역전파하는 반면, 적대적 객체에는 반대 부호의 가중치가 가중된 보상을 역전파하여 적대적 객체가 주체 및 우호적 객체와 상이하게 사고를 유발하도록 학습시킬 수 있다.
- [0066] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.
- [0067] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.
- [0068] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

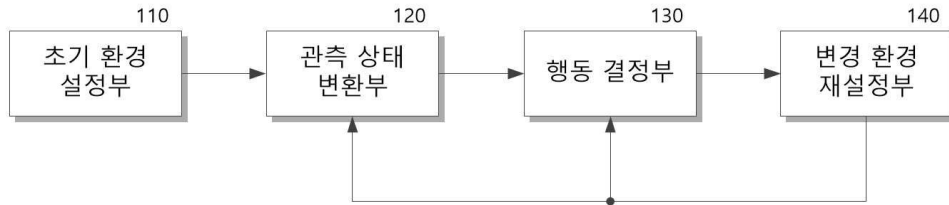
## 부호의 설명

- |        |                |                 |
|--------|----------------|-----------------|
| [0069] | 210: 초기 환경 설정부 | 220: 주체 설정부     |
|        | 230: 객체 설정부    | 240: 관측 상태 변환부  |
|        | 250: 행동 결정부    | 260: 변경 환경 재설정부 |
|        | 270: 행동 판정부    | 231: 우호적 객체 배치부 |

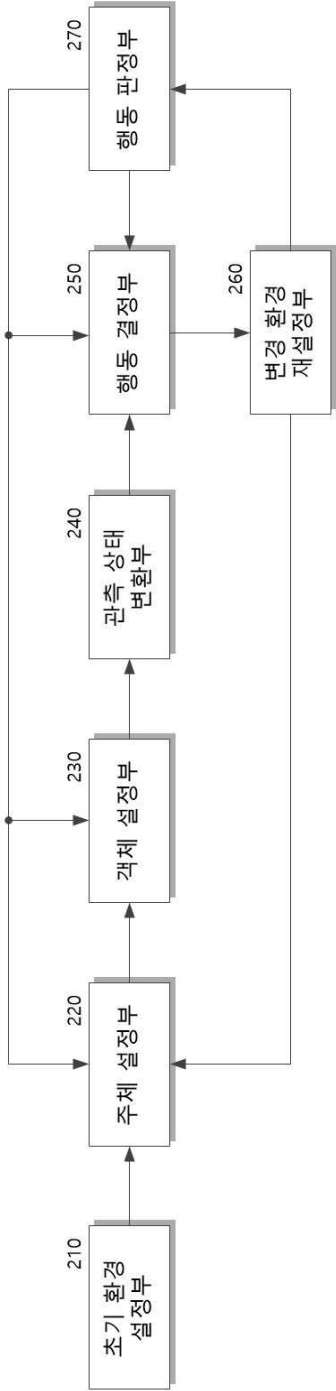
232: 중립적 객체 배치부                      233: 적대적 객체 배치부  
251: 주체 행동 결정부                      252: 우호적 객체 행동 결정부  
232: 중립적 객체 행동 결정부              233: 적대적 객체 행동 결정부

도면

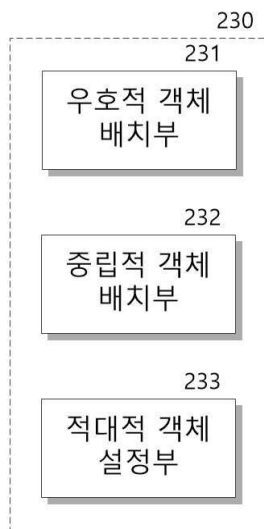
도면1



도면2



도면3



도면4



도면5

