



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0080179
(43) 공개일자 2023년06월07일

(51) 국제특허분류(Int. Cl.)

G06T 7/246 (2017.01) B64C 39/02 (2023.01)
G06N 20/00 (2019.01) G06T 19/00 (2011.01)
G06T 7/11 (2017.01)

(52) CPC특허분류

G06T 7/246 (2017.01)
B64C 39/024 (2023.01)

(21) 출원번호 10-2021-0167658

(22) 출원일자 2021년11월29일

심사청구일자 2021년11월29일

(71) 출원인

부경대학교 산학협력단

부산광역시 남구 용소로 45, 부경대학교내 (대연동)

(72) 발명자

권기룡

부산광역시 해운대구 좌동순환로 117, 201동 903호 (좌동, 롯데2차아파트)

박진혁

부산광역시 동구 충장대로289번길 45(좌천동)

(74) 대리인

유성원, 임종우

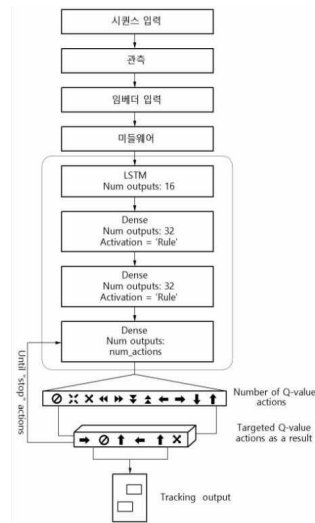
전체 청구항 수 : 총 4 항

(54) 발명의 명칭 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법

(57) 요약

본 발명은 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법에 관한 것으로, 보다 상세하게는 공간 및 시간과 같은 조건이 통제된 가상 환경에서 드론을 에이전트로 설정하여 딥러닝 모델이 가상 공간을 자동으로 학습하여 객체 검출을 위한 액션을 결정하는 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법에 관한 것이다.

대표도 - 도4



(52) CPC특허분류

G06N 20/00 (2021.08)

G06T 19/003 (2013.01)

G06T 7/11 (2017.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711125957
과제번호	1711125957
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	지역지능화혁신인재양성(Grand ICT연구센터)
연구과제명	IoT 및 지능정보 기반 동남권 제조 IT 기술 혁신 및 인재양성
기 여 율	1/2
과제수행기관명	부산대학교 산학협력단
연구기간	2020.01.01 ~ 2021.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	0004797
과제번호	P0004797
부처명	산업통산자원부
과제관리(전문)기관명	한국산업기술진흥원
연구사업명	혁신도시 공공기관연계 육성사업
연구과제명	첨단 해양산업 오픈랩 구축 및 실감형 융합 콘텐츠 개발
기 여 율	1/2
과제수행기관명	부경대학교 산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

3차원 가상의 도시환경을 구축하고, 상기 가상의 도시환경에서 상기 드론을 에이전트로 설정하는 단계;

상기 드론을 통해 촬영한 연속 이미지를 입력으로 심층 강화 학습 모델을 트레이닝하여 도시환경 이미지에 기초한 객체 추적 모델을 구축하는 단계;

상기 객체 추적 모델을 기반으로 수신 받은 영상 내에서 객체를 추적하기 위한 상기 드론의 액션을 결정하는 단계; 및

상기 객체 추적 모델은 상기 결정된 드론의 액션에 따른 다음 상태($t+1$)에 대한 예상 보상을 계산하고, 상기 객체 추적 모델의 훈련 에피소드 결과에 따라 t 단계에서 예상 보상을 업데이트하는 단계;를 포함하는 것을 특징으로 하는 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법.

청구항 2

제 1 항에 있어서,

상기 학습 단계는,

상기 가상의 도시환경에서 가상 객체 및 주변 도시 환경의 구성과 위치에 변화를 준 연속 이미지 세트들로 파인 튜닝을 실시하는 것을 특징으로 하는 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법.

청구항 3

제 1 항에 있어서,

상기 심층 강화 학습 모델은,

순환 신경망을 사용하는 DQNN을 포함하여 구성되는 것을 특징으로 하는 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법.

청구항 4

제 1 항에 있어서,

상기 예상 보상을 업데이트하는 단계는,

상기 훈련 에피소드 결과에 따라 긍정적인 보상을 1로, 모든 부정적인 보상을 -1로 클리핑하고, 변화가 없을 경우 보상을 0으로 남겨두는 것을 특징으로 하는 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법.

발명의 설명

기술 분야

[0001] 본 발명은 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법에 관한 것으로, 보다 상세하게는 공간 및 시간과 같은 조건이 통제된 가상 환경에서 드론을 에이전트로 설정하여 딥러닝 모델이 가상 공간을 자동으로 학습하여 객체 검출을 위한 액션을 결정하는 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법에 관한 것이다.

배경 기술

[0002] 시각적 객체 추적은 촬영 장치로 획득한 영상의 장면에서 특정 항목을 식별하고 모든 프레임에서 이를 인식하여 비디오 시퀀스에서 정적 및 동적이더라도 다른 클래스와 구별해야 하는 고전적인 컴퓨터 비전 문제이다.

- [0003] 영상 내 객체 추적에는 폐색, 흐림, 배경 클러터, 촬영 환경의 조명 변화 등과 같이 영상 품질 및 촬영 환경과 관련하여 고려되어야 하는 많은 요인들이 있다. 이러한 문제를 해결하기 위해 갖아 일반적인 추적 접근 방식은 다양한 기능 학습 알고리즘을 사용하여 특정 객체 클래스를 모니터링 한다. State-of-the-art로 대표되는 최신 객체 추적 방법 중에서 높은 효율성과 경쟁력 있는 추적 결과를 보여주는 기법이 있음에도 불구하고 영상적으로 객체 추적이 어려운 환경에서는 높은 정확도와 더 빠른 추적 속도를 얻기 위해서 극복해야 할 한계가 존재한다.
- [0004] 심층 인공 신경망 기반의 방법 중에 기존 객체 추적 기법이 가지고 있는 문제를 해결할 수 있는 많은 추적 필터 및 객체 감지 기반 모델이 있지만 해당 방법들에도 몇 가지 단점이 있다.
- [0005] Deep CNN(Deep CNN) 기반 시각적 객체 추적 모델은 오랜 기간 연구된 추적 알고리즘으로, 최근 몇 년 동안 다양한 객체 추적에 활용되고 있다. CNN 기반 추적기는 객체 추적 시 추적 성공률이 높고 객체 추적 프로세스 중 감지 장치에서 영상의 피쳐(Feature) 표현이 대폭 개선된다는 장점이 있다.
- [0006] 해당 연구는 대부분의 표적 추적 상황에서 객체를 찾고 클래스(class)를 분류하기 위해 자르기(crop)나 회귀 방법(Regressing methods) 또는 사전 훈련된 CNN 분류기를 사용한다. 하지만, 분류를 위한 CNN 기반 특징 표현과 추적 알고리즘 간의 불일치는 추적 결과의 출력에 영향을 미치기도 한다. 또한 사전 훈련된 CNN 분류기는 추적 프로세스의 까다로운 환경에서 제대로 작동하지 않는다.
- [0007] 그럼에도 불구하고, 탐지를 이용한 추적은 객체 추적이 까다로운 환경에서 CNN을 사용하여 훈련된 추적기보다 효과적인 것으로 입증되었으며 표준 표적 추적기보다 우수한 추적 결과를 얻을 수 있었다.
- [0008] 불행히도, 시각적으로 혼잡한 장면은 가려진 프레임이 많고, 상관거리가 작은 경우도 있기 때문에 추적 대상을 놓칠 수도 있다. CNN 기반 모델의 주요 전략 및 성과는 모양 분류 모델을 사용하여 배경 혼란을 해결하는 것 외에도 목표 간의 객체 클래스를 생성한다는 데 있다.
- [0009] 따라서, 앞서 설명한 문제점들을 극복하고 보다 빠르고 정확하게 객체를 추적할 수 있도록 하는 기술이 개발될 필요가 있다.

발명의 내용

해결하려는 과제

- [0010] 따라서, 본 발명은 상기한 바와 같은 문제점을 해결하기 위하여 제안된 것으로, 공간 및 시간과 같은 조건이 통제된 가상 환경에서 드론을 에이전트로 설정하여 딥러닝 모델이 가상 공간을 자동으로 학습하는 방법과 장기 기능 학습으로 전처리하고 제안된 엔드-투-엔드(end-to-end) 모델을 사용하여 동일한 객체 유형으로 분류하는 방법을 개시한다.
- [0011] 또한, 본 발명의 동기는 심층 강화 학습을 사용하여 대상을 학습하고 추적할 수 있는 객체 추적 알고리즘과 복잡하고 다양한 미션에서 다른 모델과 경쟁하기 위해 다양한 기술과 능력을 가진 인공지능 네트워크 모델을 개발하기 위한 것이다.
- [0012] 본 발명의 목적은 이상에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자에게 명확히 이해될 수 있을 것이다.

과제의 해결 수단

- [0013] 상기와 같은 목적을 달성하기 위한 본 발명에 따른 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법은 3차원 가상의 도시환경을 구축하고, 상기 가상의 도시환경 안에서 조정에 의해 움직일 수 있는 드론과, 임의의 알고리즘에 의해 이동하는 자동차를 구현하는 단계; 상기 가상의 도시환경에서 상기 드론을 에이전트로 설정하는 단계; 상기 드론을 통해 촬영한 연속 이미지를 입력으로 심층 강화 학습 모델을 트레이닝하여 도시환경 이미지에 기초한 객체 추적 모델을 구축하는 단계; 상기 객체 추적 모델을 기반으로 수신 받은 영상 내에서 객체를 추적하기 위한 상기 드론의 액션을 결정하는 단계; 및 상기 객체 추적 모델은 상기 결정된 드론의 액션에 따른 다음 상태(t+1)에 대한 예상 보상을 계산하고, 상기 객체 추적 모델의 훈련 에피소드 결과에 따라 t 단계에서 예상 보상을 업데이트하는 단계;를 포함한다.

발명의 효과

- [0014] 본 발명에 의하면, 객체 추적 성능 개선을 위해 별도의 장비나 환경 설정 없이 AirSim API를 사용하여 코드 또

는 알고리즘을 실행하고 테스트할 수 있다는 장점이 있다.

- [0015] 본 발명의 효과는 이상에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자에게 명확히 이해될 수 있을 것이다.

도면의 간단한 설명

- [0016] 도 1은 가상 환경과 AirSim 시뮬레이션 플랫폼의 통합 계획을 보여주는 제안의 전체 프레임 워크를 도시한 도면.
- 도 2는 객체 추적을 위해 집중된 작업 Q-값을 제공하는 DQN 네트워크가 있는 가상 환경의 학습 프로세스 네트워크 구조를 도시하는 도면.
- 도 3은 저장 값 모니터링 프로세스가 표시된 응답 메모리 장치의 구성을 도시하는 도면.
- 도 4는 순환 신경망을 기반으로 하는 추적과 DQN 에이전트 학습 아키텍처를 통합한 방법을 도시하는 예시도.
- 도 5는 본 발명의 실시예에 따른 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법을 설명하기 위한 순서도.

발명을 실시하기 위한 구체적인 내용

- [0017] 본 발명의 목적 및 효과, 그리고 그것들을 달성하기 위한 기술적 구성들은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 본 발명을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들은 본 발명에서의 기증을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다.
- [0018] 그러나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있다. 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [0019] 한편, 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 부재를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 구비할 수 있다는 것을 의미한다.
- [0020] 오늘날 객체 추적 모델은 하드웨어의 발전과 여러가지 요구사항에 따라서 많은 활용 분야에 적용되고 있으며, 최근에는 활용 초기에 비해서 요구되는 추적의 대상 및 환경, 방법 등이 복잡해지는 추세이다. 하지만 객체 추적 모델의 성능 확인을 위한 학습, 검증, 테스트 과정을 진행하기 위해서는 고려되어야 하는 여러가지 요건들 중 실제 환경에서는 현실적으로 반영하기 어려운 요건들이 몇 가지 있다.
- [0021] 본 발명에서는 AirSim(Aerial Informatics and Robotics Simulation - AirSim, CityEnvironment)을 이용하여 가상의 도시 환경을 구축하고, 가상 환경 내에서 Deep Reinforcement Learning Model Deep Q-Learning을 사용하여 사람이나 자동차 같은 특정 객체를 추적한다.
- [0022] 본 발명에서 제안하는 객체 추적 네트워크는 가상 드론의 동작을 제어하기 위해 가상 환경 시뮬레이션 시스템에서 촬영한 연속 이미지를 입력으로 받는 deep reinforcement learning model을 활용하여 환경을 관찰한다. Deep reinforcement learning model은 기존의 다양한 연속 이미지 세트를 사용하여 사전 학습되어 있는데 기존의 다양한 연속 이미지 세트는 실제 환경과 사물들을 촬영한 영상 데이터들로 가상 환경과 그 안에서 움직이는 객체들을 추적하기 위해서는 3D로 구현된 환경과 객체들이 촬영된 이미지들의 학습이 필요하다.
- [0023] 가상 공간에서 객체 추적을 하기 위해서 가상 객체 및 주변 도시 환경의 구성과 위치에 변화를 준 연속 이미지 세트들로 파인 튜닝(fine-tuning)을 실시한다.
- [0024] 가상 시뮬레이션 환경에서 제안 추적 모델의 실시간 객체 추적 실험을 하였고, 심층 신경망(deep neural network; DNN)을 이용한 state-of-art(SOTA) 기법과 비교하였을 때 속도와 정확도 면에서 더 뛰어난 성능을 가진다. 또 제안된 추적 모델의 테스트는 가상 시뮬레이션 환경에서 실행되기 때문에 속도 측면에서 GPU를 사용하

면 더 좋은 성능을 낼 수 있다.

[0026] < DQN 모델 >

[0027] 본 발명에서는 영상의 전체 비디오 프레임을 열고 영상 내에서 추적 대상을 획득하기 위해 가상 환경에 연결된 네트워크를 구축할 것을 제안한다. 도 1은 가상 환경과 AirSim 시뮬레이션 플랫폼의 통합 계획을 보여주는 제안의 전체 프레임 워크이다.

[0028] 시뮬레이션 플랫폼은 AirSim Python을 사용하고, 3D로 구현된 가상의 도시와 가상의 도시 안에서 움직일 수 있는 드론(Drone, Multirotor)과 자동차(Car)가 구현되어 있다. 자동차는 임의의 알고리즘을 이용해 도로에서 이동하며, 드론은 사용자가 원하는 대로 조종이 가능하도록 구현된다.

[0030] 본 발명에서 DQN 모델은 드론 에이전트의 움직임과 관련된 기술을 배우기 시작하면서 가상 도시 환경을 빠르게 탐색하게 된다. 본 발명에서 제안하는 아키텍처는 수신 받은 영상을 학습 및 가치 주정을 위한 입력 값으로 받는다.

[0031] 이 입력 영상은 다양한 시뮬레이션 환경에서 에이전트의 행동 제어를 위한 정책을 학습하는데 사용된다. 제안 모델의 환경 시뮬레이션 코드는 학습과 추적이 가능한 드론 시뮬레이션을 실행하기 위해 AirSim 파이썬 클라이언트로 가상 도시 환경과 DQN 에이전트 시뮬레이션 네트워크 알고리즘을 연결한다. 객체 추적 성능 개선을 위해 별도의 장비나 환경 설정 없이 AirSim API를 사용하여 코드 또는 알고리즘을 실행하고 테스트할 수 있다는 것이 장점이다.

[0033] < DQN network architecture >

[0034] 권장되는 네트워크 모델 아키텍처는 강화 학습과 순환 신경망을 통합한 것으로, 순환 네트워크는 연속적인 상황에 적용하고 환경 목표 속성을 예측하는 데 효과적이다.

[0035] 도 2는 객체 추적을 위해 집중된 작업 Q-값을 제공하는 DQN 네트워크가 있는 가상 환경의 학습 프로세스 네트워크 구조를 보여준다. 도 3의 다이어그램에서 학습 절차의 목표 조치 및 상태 값이 있는 DQN 네트워크 모델 처리 단계의 표현을 볼 수 있다. 여기서 취해진 조치와 상태 값은 드론 에이전트가 동일한 위치를 두 번 탐색하지 않도록 하는 폴 타임 훈련 작업을 위한 초기값으로서, 다음 상태 명령을 위해서 유지된다. 이 동작은 반복되는 절차로 보일 수 있지만 가상 환경에서 동작 및 상태 값은 이전과 동일하다.

[0037] 도 2는 학습 반복 에피소드의 샘플 또는 미니 배치를 업데이트하기 위한 Q-러닝의 일부를 나타낸다. 이 과정에서 업데이트 Q-값은 에이전트의 행동을 결정하는데, 가장 높은 목표 Q-학습 값을 가진 행동이 선택되고 업데이트는 벨만 방정식을 사용하여 수행된다.

[0038] DQN 네트워크의 에이전트는 동적 모드에서 환경을 추적하는 재생 메모리 클래스 단위를 포함하며, 모든 state, action, new state, reward, done transition이 모두 기억된다. 이 재생 메모리 접근 방식을 사용하면 보존된 값에서 미니 배치를 효과적으로 샘플링하고 정확한 상태 표현을 생성할 수 있다.

[0039] 이 상황에서 우리는 정확한 묘사를 위한 값을 얻기 위해 이전 프레임의 수를 요구할 것이다.

[0040] 도 3은 저장 값 모니터링 프로세스가 표시된 응답 메모리 장치의 구성을 보여준다. 버퍼 메모리는 가상 환경에 직접 연결되어 메모리 장치에 필요한 전환이 추가될 것이다. 샘플링 프로세스 동안 다양한 크기의 지도 인덱스가 무작위로 메모리에 생성되고 반환된 인덱스는 AirSim python 클라이언트의 상태 가져오기 기능을 사용하여 검색할 수 있다. 이 상황에서 상태는 마지막으로 기록된 길이 인식으로 구성된다. 훈련 활동의 수동으로 제공된 크기 매개변수로 표시된 샘플 수를 사용하여 미니배치 프로세스를 생성한다.

[0041] 재생 메모리는 QN 에이전트의 가장 중요한 개별 핵심 구성 요소 중 하나로, 대상 Q 네트워크를 분리하고 성능에 부정적인 영향을 미친다.

- [0043] 도 3과 같이 DQN 네트워크 모델의 데이터 흐름도는 AirSim Python 클라이언트를 통해 Q-네트워크와 순환 네트워크 설계(예측용 네트워크)를 가상 시뮬레이션 플랫폼에 연결한다. 누산기는 에이전트 평가에 사용할 프레임 N의 추적을 유지한다. 또한 예측 및 목표 Q 값을 결합하여 DQN 손실 함수를 계산할 수 있으며 기울기 손실 출력도 얻을 수 있다.
- [0045] < Deep Q-Agent with tracking unit >
- [0046] 합성곱 신경망 모델을 사용하는 DQN 에이전트 모델을 제안한 방법과 달리, 본 발명에서는 순환 신경망을 사용하는 DQNN 모델을 제안한다. 종래 기술에서는 가상 시뮬레이션 환경의 정보와 현재 프레임 이후 상황의 예측을 위한 학습에 순환 계층을 사용하는 대신 합성곱 신경망 계층을 사용하는데, 본 발명의 실시예에서는 추가적인 특징 기반 접근 방식을 사용하여 최종 액션 값의 마지막 결과를 결정한다. 순차 환경 학습 프로세스에서 순환 계층은 정책에 의해 일반화된 데이터를 생성하여 상태 및 동작 값에 대한 정확한 예측 결과를 제공한다. 도 4는 순환 신경망을 기반으로 하는 추적과 DQN 에이전트 학습 아키텍처를 통합한 방법을 나타낸다.
- [0048] 구현의 첫 번째 단계에서는 초기 설정으로 DQN 에이전트 모델의 매개변수를 구성한다. 여기서 액션 값 모델은 에이전트가 가상 시뮬레이션 환경과 상호 작용하기 위해 사용한다. 목표 모델은 훈련에서 목표 Q-값을 계산하는데 사용되며 DQN 에이전트의 학습 안정성을 높이기 위해 업데이트 빈도를 줄인다.
- [0049] 순차적 모드에 구축된 네트워크 모델은 활성화 relu 함수와 벡터의 차원을 64에서 32로 변경하기 위해 조밀한 계층과 깊게 연결된 LSTM 레이어에 의해 적용된다. 네트워크는 상태, Q-값 및 네트워크 모델 출력의 출력 값을 액션 값의 수로 제공한다. 많은 Q-값 액션을 통해 에이전트는 현재 환경의 상태와 관련하여 수행할 다음 액션 (행동)을 선택할 수 있다. 다음 단계의 액션을 위해 드론 에이전트는 단기 기억에 상태를 추가하는 정보를 사용하고 카운터에서 간격 액션의 추적을 유지하여 시간에 따라 선행적으로 애플론을 어닐링하는 동안 네트워크를 사용하여 무작위로 액션을 선택하여 최상의 액션을 얻는다.
- [0050] 네트워크 모델의 관찰 장치를 통해 에이전트는 이전 상태에서 작업 기능을 통해 작업을 수행한 결과를 관찰할 수 있다. 완료되면 프로세스는 네트워크의 단기 기억에서 재설정되며 학습 절차의 일회성 탐색 에피소드 요약에 대한 요약 결과를 제공하고 네트워크의 장기 기억에 추가한다.
- [0052] < Training the DQN network model >
- [0053] 훈련 과정을 통해 에이전트는 환경 역학을 더 잘 이해하고 다음 상태(t+1)에 대한 예상 보상을 계산하고 네트워크 모델의 첫 번째 훈련 에피소드 결과에 따라 t 단계에서 예상 보상을 업데이트하도록 스스로 훈련하게 된다. 목표 기대치는 훈련 안정성을 높이기 위한 행동 가치 네트워크의 더 안정적인 버전인 목표 네트워크를 통해 계산된다.
- [0054] 실제로, 목표 네트워크는 정기적으로 업데이트 되는 행동 가치 네트워크의 동결된 복사본이다. 훈련 과정 후, 네트워크는 모든 긍정적인 보상을 1로, 모든 부정적인 보상을 -1로 클리핑하고, 보상이 변화가 없을 경우 0으로 남겨둔다. 그리고 다음 에피소드를 위해 네트워크를 다시 훈련하고, 대상 네트워크를 업데이트하고, 최종적으로 네트워크 출력 파일을 고정 경로에 저장한다. 재생 메모리 장치로부터 전환기의 미니배치를 무작위로 샘플링하는 배치 생성과 그래프 계산을 호출하는 훈련 함수의 확장이 있다. 네트워크를 훈련하기 전에 deep Q-agent의 하이퍼 파라미터와 기본값을 설정한다. 그리고 마지막에 총 보상, 평균 최대 Q, 지속 시간 및 평균 손실 값의 모든 학습 값을 요약하여 조정된 경로 위치에 저장한다.
- [0056] < Tracking baseline of the DQN network model >
- [0057] 추적 구현 단계에서는 객체 클래스, 정보 및 속성을 식별하기 위해 지도 학습 방식을 활용했으며, 이를 DQN 에이전트 시뮬레이션 네트워크 아키텍처와 결합했다. 제안된 추적기를 테스트하는 동안 추적 접근 방식은 사전 훈련된 객체 분류기 모델을 사용하여 가상 시뮬레이션 플랫폼에서 대상을 감지한다.
- [0058] 제안하는 추적기는 가상 시뮬레이션 환경에서 드론 에이전트와의 순차적 의사결정 절차로 DQN 네트워크와 통합

하여 구현하며, 네트워크 모델 관찰부는 가상 환경 시퀀스를 나타내고 순환 네트워크 기반 아키텍처 계층은 위의 그림 5에서처럼 각 프레임의 예측된 경계 상자 위치를 제공한다. 주어진 가상 시뮬레이션 모델 상태에 적절한 동작으로 대상 클래스를 예측하기 위해 네트워크를 훈련하고 딥 q-러닝 접근 방식으로 업데이트되어 드론 에이전트가 엔드 투 엔드 강화 학습을 통해 고차원 가상 시뮬레이션 환경 입력에서 계속 효과적으로 학습할 수 있도록 한다. 네트워크는 학습 과정에서 출력 예측을 제공하고 프레임 오른쪽 및 상단 위치의 원점을 양의 축으로 간주하는 데카르트 좌표계에서 왼쪽 상단 모서리와 오른쪽 하단 모서리에 위치한 경계 상자의 IoU(교집합)를 계산하여 추적 단위를 통합한다. 그런 다음 교차 직사각형의 좌표는 최대값과 최소값을 식별하여 결정된다. 두 축 정렬 경계 상자(the two axis-aligned bounding boxes)의 교차 영역은 항상 축 정렬 경계 상자 값(axis-aligned bounding boxes, AABB)으로 간주된다. 그런 다음 두 축 정렬 경계 상자의 면적을 계산한다. 합집합에 대한 교차는 계산된 교차 영역을 취하여 계산되고 예측값 더하기 정답 영역(ground truth areas)에서 교차 영역을 뺀 값의 합으로 나누어 0과 1 값 사이의 값으로 도출한다. 다음 단계로, 시뮬레이션 환경에서 액션의 해석과 해석된 액션 시퀀스에 추적 계산 활동이 적용된다. 보상 함수는 경계 상자의 중심과 프레임 중심 사이의 유클리드 거리, 경계 상자의 합집합에 대한 교차점 및 매개 변수 임계값 높이와 무게가 있는 프레임 중심을 중심으로 하는 가상 상자의 스케일 합으로 계산되며, 완성된 부분은 미리 정해진 간격으로 보상 값을 취하여 결정된다. 마지막 단계로, 강화 에이전트를 생성하여 지정된 매개변수를 구성하고 가상 환경 시뮬레이션 모델 입력을 사용하여 알고리즘을 테스트한다.

[0060] 도 5는 본 발명의 실시예에 따른 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법을 설명하기 위한 순서도이다. 도 5를 참조하면, 심층 강화 학습 기반 가상 환경 시뮬레이션에서의 객체 추적 방법은 3차원 가상의 도시환경을 구축하고, 상기 가상의 도시환경 안에서 조정에 의해 움직일 수 있는 드론과, 임의의 알고리즘에 의해 이동하는 자동차를 구현하며, 상기 가상의 도시환경에서 상기 드론을 에이전트로 설정하는 단계(S110), 상기 드론을 통해 촬영한 연속 이미지를 입력으로 심층 강화 학습 모델을 트레이닝하여 도시환경 이미지에 기초한 객체 추적 모델을 구축하는 단계(S120), 상기 객체 추적 모델을 기반으로 수신 받은 영상 내에서 객체를 추적하기 위한 상기 드론의 액션을 결정하는 단계(S130) 및 상기 객체 추적 모델은 상기 결정된 드론의 액션에 따른 다음 상태($t+1$)에 대한 예상 보상을 계산하고, 상기 객체 추적 모델의 훈련 에피소드 결과에 따라 t 단계에서 예상 보상을 업데이트하는 단계(S140)를 포함한다.

[0061] 여기서, 학습 단계(S130)는, 상기 가상의 도시환경에서 가상 객체 및 주변 도시 환경의 구성과 위치에 변화를 준 연속 이미지 세트들로 파인 튜닝을 실시한다.

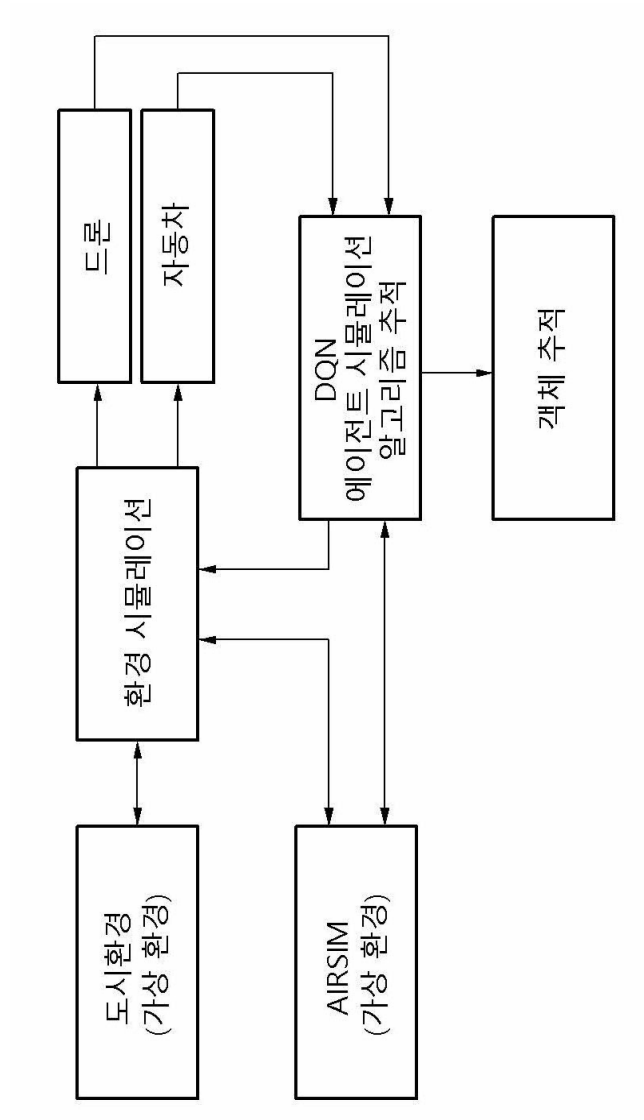
[0062] 그리고, 심층 강화 학습 모델은, 순환 신경망을 사용하는 DQNN을 포함하여 구성될 수 있다.

[0063] 예상 보상을 업데이트하는 단계(S150)는, 상기 훈련 에피소드 결과에 따라 긍정적인 보상을 1로, 모든 부정적인 보상을 -1로 클리핑하고, 변화가 없을 경우 보상을 0으로 남겨둘 수 있다.

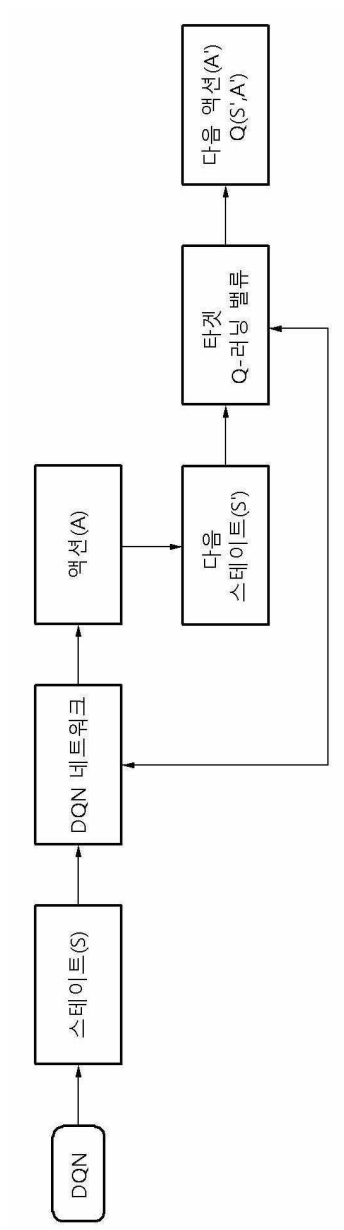
[0065] 본 명세서와 도면에는 본 발명의 바람직한 실시예에 대하여 개시하였으며, 비록 특정 용어들이 사용되었으나, 이는 단지 본 발명의 기술 내용을 쉽게 설명하고 발명의 이해를 돕기 위한 일반적인 의미에서 사용된 것이지, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시예 외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형예들이 실시 가능하다는 것은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명한 것이다.

도면

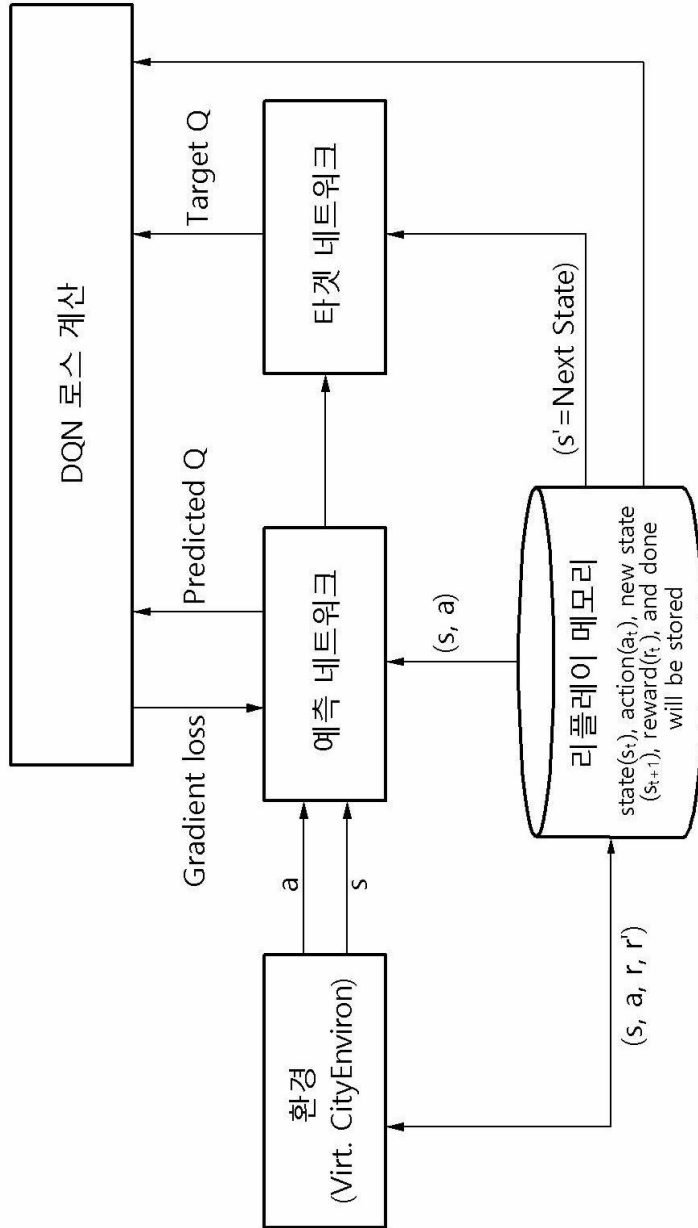
도면1



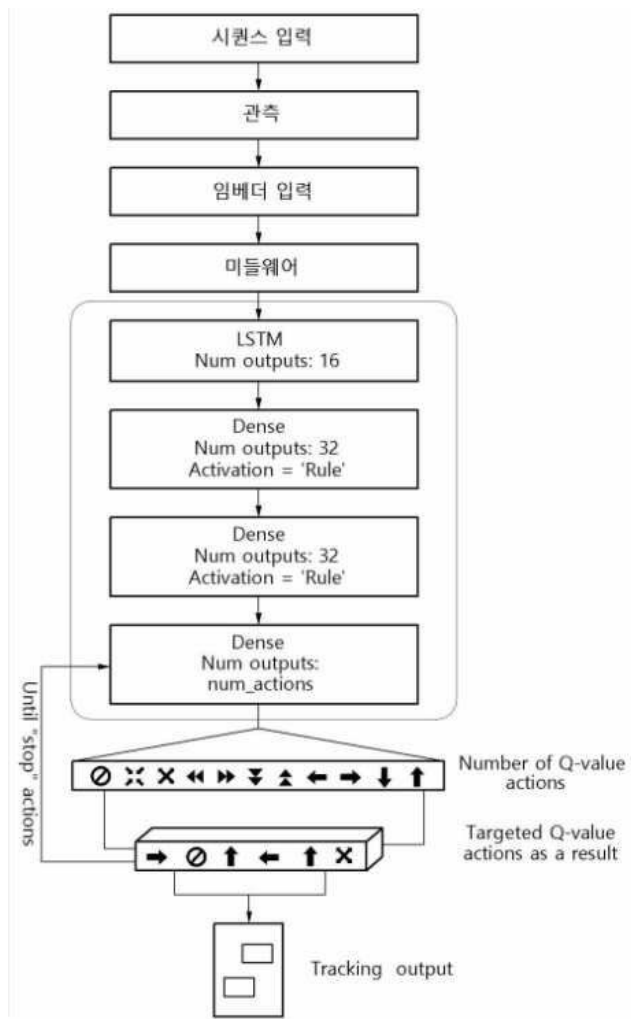
도면2



도면3



도면4



도면5

