



Universidad Politécnica
de Madrid

**Escuela Técnica Superior de
Ingenieros Informáticos**



Máster Universitario en Inteligencia Artificial

Trabajo Fin de Máster

**Navegación Reactiva Aplicada a Agentes
Físicos en Entornos Domésticos usando
Habitat Sim**

Autor(a): Luna Jiménez Fernández
Tutor(a): Martín Molina Gómez

Madrid, Septiembre - 2021

Este Trabajo Fin de Máster se ha depositado en la ETSI Informáticos de la Universidad Politécnica de Madrid para su defensa.

Trabajo Fin de Máster
Máster Universitario en Inteligencia Artificial

*Título: Navegación Reactiva Aplicada a Agentes Físicos en Entornos Domésticos
usando Habitat Sim*

Septiembre - 2021

Autor(a): Luna Jiménez Fernández
Tutor(a): Martín Molina Gómez
Computer Vision and Aerial Robotics (CVAR)
ETSI Informáticos
Universidad Politécnica de Madrid

Resumen

«Aquí va el resumen del TFM. Extensión máxima 2 páginas.»

Abstract

«Abstract of the Master Project. Maximum length: 2 pages.»

Tabla de contenidos

1. Introducción	1
1.1. Introducción	1
1.2. Motivación	2
1.3. Estructura	2
2. Descripción del problema	5
2.1. Problema	5
2.2. Antecedentes	5
2.3. Objetivos	5
3. Revisión de técnicas	7
3.1. <i>Deep Learning</i>	7
3.2. Aprendizaje por refuerzo	7
3.2.1. Algoritmos de aprendizaje por refuerzo clásicos	7
3.2.2. Algoritmos de aprendizaje por refuerzo profundos	7
3.3. Algoritmos de navegación automática	7
4. Simulador: <i>Habitat Sim</i> y <i>Habitat Lab</i>	9
4.1. <i>Habitat Sim</i> y <i>Habitat Lab</i>	9
4.1.1. <i>Habitat Sim</i>	9
4.1.2. <i>Habitat Lab</i>	9
4.1.3. Principales conceptos de <i>Habitat Lab</i>	9
4.1.3.1. Entornos	9
4.1.3.2. Tareas	9
4.1.3.3. Conjuntos de datos	9
4.1.3.4. Episodios	10
4.1.3.5. Sensores	10
4.1.3.6. Ficheros de configuración	10
4.1.3.7. Entrenadores	10
4.1.3.8. Agentes	10
4.1.3.9. <i>Benchmarks</i>	10
4.2. Instalación del simulador	10
4.2.1. Requisitos y versiones	10
4.2.2. Proceso de instalación	10
5. Diseño del agente	11
5.1. Caracterización del conocimiento	11
5.1.1. Descripción del problema	11

5.1.2. Estado	11
5.1.3. Acciones	11
5.1.4. Recompensas	11
5.2. Arquitectura del agente	11
5.2.1. Propuesta 1: Red convolucional (CNN)	11
5.2.2. Propuesta 2: Red híbrida (CNN + MLP)	11
5.3. Actuación del agente	11
5.4. Entrenamiento del agente	11
5.4.1. <i>Experience Replay</i>	11
5.4.2. Memorización de experiencias	11
5.4.3. Aprendizaje a partir de las experiencias	11
5.5. Otros agentes propuestos	11
5.5.1. Agentes heurísticos	12
5.5.2. Agente PPO	12
6. Experimentación	13
6.1. Experimentos realizados y parametros utilizados	13
6.2. Resultados durante el entrenamiento	13
6.3. Resultados durante la evaluación	13
7. Conclusiones	15
7.1. Conclusiones	15
7.2. Trabajo futuro	15
7.3. Agradecimientos	15
Bibliografía	15
Anexos	18

Capítulo 1

Introducción

En este capítulo se realizará una breve introducción a los contenidos que serán expuestos posteriormente a lo largo de la memoria. Tras ésta presentación, se expondrá la motivación que ha propiciado el desarrollo de este trabajo. Finalmente, se describirá la estructura seguida por la memoria.

1.1. Introducción

La **navegación autónoma** de robots en entornos desconocidos y complejos es un problema de gran interés en la actualidad para el que se ha propuesto una amplia gama de soluciones, buscando que éstas sean a la vez eficientes durante su entrenamiento y capaces de navegar entornos de forma exitosa. Una de las familias de algoritmos más relevantes para este propósito son los **algoritmos de aprendizaje por refuerzo**, capaz de aprender de forma autónoma a navegar entornos desconocidos a partir de experiencia previa, con gran éxito.

Además, la **simulación virtual** tanto de estos robots como de otros problemas es un campo en crecimiento, especialmente durante la pandemia del CoVID-19, al verse limitadas las capacidades de experimentación en entornos físicos. Por tanto el objetivo de este trabajo es aunar el **desarrollo de un algoritmo híbrido eficiente** para la navegación en entornos complejos (como el interior de un domicilio) con el **estudio y uso de *Habitat Sim***, un simulador novedoso para el entrenamiento y evaluación de agentes robóticos físicos.

Esta memoria comienza con una revisión de las principales técnicas usadas actualmente tanto en *Deep Learning* como en aprendizaje por refuerzo y en navegación autónoma de robots (centrándose en los algoritmos de *Artificial Potential Field (APF)*). Tras esto, se realiza un estudio en detalle del simulador *Habitat Sim* y su principal librería / API para *Python*, *Habitat Lab*; centrándose en los principales componentes de la librería, su funcionamiento y su uso. Posteriormente, se describe el diseño e implementación que se ha realizado en el trabajo, haciendo hincapié en la representación del conocimiento (problema a resolver y definiciones de estado, acción y recompensa), las arquitecturas del agente propuestas y su funcionamiento tanto durante el entrenamiento como la actuación.

Se procede después a la explicación de los experimentos realizados (tanto los parámetros usados como los experimentos a realizar), analizando el rendimiento de las

variantes propuestas del agente durante el entrenamiento y en una evaluación posterior, analizando éstos resultados y comparándolos con otros agentes usados como *benchmarks*. Finalmente, se interpretarán estos resultados, extrayendo unas conclusiones y ofreciendo futuras líneas de trabajo a partir del conocimiento adquirido.

1.2. Motivación

Este trabajo se puede entender como una continuación del trabajo realizado por C. Sampedro *et al.* en 2018 [1], en el que se desarrolla con buenos resultados un sistema de navegación autónomo para drones aéreos usando aprendizaje por refuerzo profundo con campos de potenciales artificiales y láseres para percibir el entorno. Una de las metas de este trabajo es estudiar si la implementación de un algoritmo de características similares pero aplicado a robots terrestres usando cámaras de profundidad sería a su vez efectivo.

Además, la situación de pandemia actual ha dejado en evidencia la necesidad del uso de simuladores, especialmente para algoritmos que necesiten un entrenamiento largo y que puedan necesitar equipamiento especializado para ello (como robots, drones, instalaciones especializadas...). Por eso, otra de las principales metas del trabajo es el estudio de la herramienta *Habitat Sim* [2] [3], viendo su viabilidad de cara a futuros trabajos.

Para acabar, otra razón no despreciable para la elección de esta temática de trabajo es el propio interés de la alumna por el campo del aprendizaje por refuerzo profundo. Ya se realizó un trabajo previo estudiando la aplicación de estas técnicas a juegos reales como Tetris [4], y este trabajo sirve para ampliar más el conocimiento y aplicarlo a tecnologías modernas y a otros campos de interés.

1.3. Estructura

Esta memoria está dividida en un total de 7 capítulos, que serán descritos brevemente a continuación.

- **Capítulo 1:** En este capítulo se introduce el trabajo desarrollado, la motivación que ha llevado a éste y la estructura general de la memoria.
- **Capítulo 2:** En este capítulo se describe en profundidad el problema a resolver, presentando los antecedentes previos al trabajo realizado y detallando los objetivos que se esperan cumplir.
- **Capítulo 3:** En este capítulo se realiza una revisión de las principales técnicas en los campos relacionados con el trabajo: *deep learning* y redes neuronales convolucionales, aprendizaje por refuerzo (estudiando tanto las técnicas clásicas como las técnicas de aprendizaje por refuerzo profundo) y algunos de los principales algoritmos de navegación automática.
- **Capítulo 4:** En este capítulo se presentan tanto *Habitat Sim* como *Habitat Lab*, las principales herramientas usadas durante el desarrollo del trabajo. Tras esto, se exponen los principales componentes de Habitat Lab, explicando su funcionamiento y uso. Finalmente, se habla sobre la instalación y las dependencias necesarias del simulador.

- **Capítulo 5:** En este capítulo se detalla el diseño del agente de navegación reactiva propuesto. Se describe tanto la representación del conocimiento (estado, acciones y recompensas) como la arquitectura, el método de actuación y el entrenamiento llevado a cabo por el agente. Finalmente, se realiza una breve explicación del funcionamiento y la arquitectura del resto de agentes usados como *benchmarks* y comparativas ofrecidos por *Habitat Lab*.
- **Capítulo 6:** En este capítulo se detalla la experimentación realizada, indicando los parámetros utilizados. Además, se presentan los resultados y el rendimiento obtenido por los agentes tanto durante el entrenamiento como durante la evaluación posterior.
- **Capítulo 7:** Finalmente, en este capítulo se presentan las conclusiones alcanzadas tras el desarrollo del trabajo, proponiendo posibles líneas de trabajo futuro para continuarlo.

Además, al final de la memoria se incluye una bibliografía en la que se encuentra la lista de fuentes y referencias usadas a lo largo de ésta.

Capítulo 2

Descripción del problema

2.1. Problema

[HABLAR DEL PROBLEMA ESPECIFICO A RESOLVER, MENCIONAR QUE COVID MOTIVA A SIMULADOR]

2.2. Antecedentes

[HABLAR POR ENCIMA DEL PROPIO ARTICULO Y DE OTROS ALGORITMOS DE POINT NAV]

2.3. Objetivos

[LA VIEJA CONFIABLE PERO MENCIONANDO HABITAT-SIM Y HABITAT-LAB]

Capítulo 3

Revisión de técnicas

3.1. *Deep Learning*

[SOBRE DEEP LEARNING, REDES NEURONALES PROFUNDAS Y CNNs]

3.2. Aprendizaje por refuerzo

[APRENDIZAJE POR REFUERZO, Q LEARNING (y quizás otros clásicos?), DEEP Q LEARNING (y mejoras), AGENTE / CRITICO, PPO...]

3.2.1. Algoritmos de aprendizaje por refuerzo clásicos

3.2.2. Algoritmos de aprendizaje por refuerzo profundos

3.3. Algoritmos de navegación automática

[PROBABLEMENTE CENTRARSE MAS EN ALGORITMOS DE ATRACCION REPULSION]

Capítulo 4

Simulador: *Habitat Sim* y *Habitat Lab*

4.1. *Habitat Sim* y *Habitat Lab*

[UNA INTRODUCCION MAS BREVE AL SIMULADOR COMO TAL]

4.1.1. *Habitat Sim*

[HABLA POR SEPARADO DE HABITAT 1 Y HABITAT 2]

4.1.2. *Habitat Lab*

[HABLA DE HABITAT LAB Y HABITAT BASELINES]

4.1.3. Principales conceptos de *Habitat Lab*

[DE CADA CONCEPTO, PROBABLEMENTE EXPLICAR QUE ES, EXPLICAR QUE OFRECE POR DEFECTO HABITAT / BASELINES]

[NO SE SI SERÁ NECESARIO HABLAR DEL PROPIO AGENTE AQUÍ O EN UN ANEXO]

[EL ORDEN ES TENTATIVO]

4.1.3.1. Entornos

[ENV, RLENV, NAVRLENV Y LOS METODOS QUE HAY QUE IMPLEMENTAR]

4.1.3.2. Tareas

[PROBABLEMENTE MENCIONAR LAS PRINCIPALES TAREAS OFRECIDAS POR ENCIMA]

4.1.3.3. Conjuntos de datos

[FOTOS DE CADA CONJUNTO DE DATOS] [MENCIONAR LOS 4 DATASETS DISPONIBLES, Y COMO USARLOS (ESTRUCTURA)]

4.1.3.4. Episodios

4.1.3.5. Sensores

4.1.3.6. Ficheros de configuración

[PROBABLEMENTE MENCIONAR SECCIONES CLAVES Y ELEMENTOS CLAVES QUE SIGNIFICAN]

4.1.3.7. Entrenadores

[AQUI UN PSEUDOCODIGO DEL METODO DE ENTRENAR QUEDARÍA DE LUJO]

4.1.3.8. Agentes

4.1.3.9. Benchmarks

4.2. Instalación del simulador

4.2.1. Requisitos y versiones

[INDICAR LAS VERSIONES USADAS DE TODO, Y ESPECIFICAR QUE NO SE GARANTIZA QUE FUNCIONE CON TODO]

4.2.2. Proceso de instalación

[INDICAR TAMBIEN CUDA]

Capítulo 5

Diseño del agente

[SE HABLAN DE TODOS LOS AGENTES, PERO NOS CENTRAMOS EN EL PROPIO (que para eso esta)]

5.1. Caracterización del conocimiento

5.1.1. Descripción del problema

5.1.2. Estado

5.1.3. Acciones

5.1.4. Recompensas

5.2. Arquitectura del agente

5.2.1. Propuesta 1: Red convolucional (CNN)

5.2.2. Propuesta 2: Red híbrida (CNN + MLP)

5.3. Actuación del agente

[exploracion explotacion]

5.4. Entrenamiento del agente

5.4.1. *Experience Replay*

5.4.2. Memorización de experiencias

5.4.3. Aprendizaje a partir de las experiencias

5.5. Otros agentes propuestos

[AQUI DENTRO SE INDICAN LOS AGENTES PROPUESTOS Y BREVEMENTE SU ARQUITECTURA]

5.5.1. Agentes heurísticos

5.5.2. Agente PPO

Capítulo 6

Experimentación

[PONEMOS EL ANALISIS EN UN CAPITULO APARTE?]

6.1. Experimentos realizados y parametros utilizados

[EXPERIMENTOS A REALIZAR, PARAMETROS A USAR, ORDENADOR USADO, ETC]
[PARA REPRODUCIBILIDAD VAMOS]

6.2. Resultados durante el entrenamiento

[MEDIDAS TIPO TIEMPO PARA ENTRENAR, RECOMPENSA MEDIA POR EPOCH, ACCIONES POR EPOCH, TASA DE EXITO...]

6.3. Resultados durante la evaluación

[TASA DE EXITO, CUAL ES MEJOR, ETC]

Capítulo 7

Conclusiones

7.1. Conclusiones

7.2. Trabajo futuro

7.3. Agradecimientos

[YA QUE NO LO TIENEN ANTES...]

Bibliografía

- [1] C. Sampedro, H. Bavle, A. Rodriguez-Ramos, P. De La Puente y P. Campoy, "Laser-Based Reactive Navigation for Multirotor Aerial Robots using Deep Reinforcement Learning," *IEEE International Conference on Intelligent Robots and Systems*, págs. 1024-1031, 2018, ISSN: 21530866. DOI: 10.1109/IROS.2018.8593706.
- [2] M. Savva, A. Kadian, O. Maksymets y col., "Habitat: A Platform for Embodied AI Research," en *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.
- [3] A. Szot, A. Clegg, E. Undersander y col., "Habitat 2.0: Training Home Assistants to Rearrange their Habitat," *arXiv preprint arXiv:2106.14405*, 2021.
- [4] L. Jimenez, *Aplicación de Deep Reinforcement Learning a un juego real - Tetris*, Albacete, 2020.

Apéndice A

[NO TENGO MUY CLARO COMO SE ESCRIBIAN LOS APENDICES NO TE VOY A ENGAÑAR] [EN LATEX ME REFIERO]

[APENDICES PROBABLES:] [MANUAL DE USUARIO, COMO GENERO GRAFICAS, CONTENIDOS DEL CD]