

Tabla de Contenido

1. Introducción	1
1.1. Motivación y definición del problema	1
1.2. Hipótesis	3
1.3. Objetivos	3
1.3.1. Objetivo general	3
1.3.2. Objetivos específicos	4
1.4. Alcances de este trabajo	4
1.5. Estructura del documento	4
2. Antecedentes	5
2.1. Procesos de Decisión de Markov	5
2.1.1. Observabilidad parcial	6
2.2. Aprendizaje reforzado	6
2.2.1. Funciones de valor	8
2.2.2. Algoritmos de aprendizaje reforzado	9
2.3. Aprendizaje reforzado profundo	11
2.3.1. Aproximación funcional y redes neuronales artificiales	12
2.3.2. Deep Q-Network (DQN)	17
2.3.3. Deep Deterministic Policy Gradient (DDPG)	19
3. Evasión de colisiones mediante aprendizaje reforzado multimodal	21
3.1. Motivación	21
3.2. Trabajos relacionados	22
3.3. Evasión de colisiones para el robot Pepper	24
3.3.1. Modelamiento del problema	25
3.3.2. Procesamiento de imágenes de profundidad	27
3.3.3. Procesamiento de mediciones de rango	30
3.3.4. Algoritmo y parametrización de la política	31
3.4. Evaluación experimental	32
3.4.1. Entrenamiento y evaluación en simulaciones	33
3.4.2. Validación en el mundo real	37
3.5. Discusión	39
4. Planificación local usando nubes de puntos 2D como observaciones	41
4.1. Motivación	41
4.2. Trabajos relacionados	42

4.3.	Propuesta	43
4.3.1.	Modelamiento del problema	43
4.3.2.	Nubes de puntos 2D como observaciones	46
4.3.3.	Algoritmo y parametrización de la política	47
4.4.	Evaluación experimental	48
4.4.1.	Entrenamiento y evaluación en simulaciones	49
4.4.2.	Validación en simulaciones	54
4.4.3.	Validación en el mundo real	61
4.5.	Discusión	64
5.	Conclusión y trabajo futuro	65
	Bibliografía	66