



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**CONCEPTUALIZACIÓN DEL ESPACIO DE
NAVEGACIÓN DE UN ROBOT MÓVIL,
UTILIZANDO INFORMACIÓN SEMÁNTICA.**

T E S I S

QUE PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERÍA

P R E S E N T A:

NOE ABRAHAM ROMERO RUIZ

DIRECTOR DE LA TESIS: DR. JESUS SAVAGE CARMONA

MÉXICO, D.F.

2011.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

DEDICATORIA

A mi familia y amigos por su apoyo incondicional.

A mis profesores y sinodales por sus aportaciones a este trabajo.

Al CONACYT por la beca recibida.

ÍNDICE GENERAL

Índice de Figuras	4
Índice de Tablas	5
Resumen	6
Organización de la tesis	7
CAPÍTULO 1 INTRODUCCIÓN	8
1.1 Objetivo general	8
1.2 Contribución y relevancia	8
1.2.1 Relevancia	9
1.2.2 Contribuciones	10
1.3 Objetivos específicos	11
1.4 Metodología de Investigación	12
1.5 Detalles del sistema	13
1.5.1 El robot	13
1.5.2 El ambiente	13
1.5.3 Implementación	13
1.5.4 Suposiciones	14
1.5.5 Sistema de reconocimiento	14
1.5.6 Representación del ambiente y clasificación de lugar	15
CAPÍTULO 2 MARCO TEÓRICO	16
2.1 Robótica	16
2.2 Esquema general del sistema robot	16
2.3 Robots Móviles	17
2.3.1 Planificación en robots móviles	18
2.3.2 Sistema de percepción en un robot móvil	19
2.4 Creación de mapas en robótica móvil	19
2.4.1 Mapas métricos	20
2.4.1.1 Mapas de rejilla	20
2.4.1.2 Mapas geométricos o basados en características.	20
2.4.2 Mapas topológicos	20

2.4.3 Mapas híbridos	21
2.4.4 Mapas cognitivos	21
2.4.5 Mapas semánticos	24
2.5 Escáner láser de medición de distancias	26
2.5.1 Escáner Láser URG-04LX	27
2.6 La transformada de Hough	28
2.6.1 Descripción del algoritmo de la transformada de Hough	31
2.7 Clustering de barrido láser	31
CAPÍTULO 3 EXTRACCIÓN DE MARCAS NATURALES USANDO SENSOR LÁSER.	32
3.1 Detección de discontinuidades	32
3.2 Detección de paredes	34
3.3 Detección de Esquinas	35
3.4 Algoritmo General para la extracción de características	35
3.5 Pseudocódigo del algoritmo	36
CAPÍTULO 4 CLASIFICACIÓN DE LUGARES USANDO SENSOR LÁSER.	38
4.1 Algoritmo AdaBoost	38
4.2 Aplicación del algoritmo AdaBoost para clasificar lugares.	40
4.3 Multiclasificación de Lugares	42
4.4 Características del barrido láser	44
CAPÍTULO 5 CONSTRUCCIÓN DE UNA REPRESENTACIÓN MÉTRICO-SEMÁNTICA DEL AMBIENTE	47
5.1 Propiedades del sistema	47
5.2 Propiedades del entrenamiento	50
5.3 Experimentos	52
5.3.1 Experimentos sobre ambientes simulados	52
5.3.2 Experimentos sobre ambientes reales	58
5.4 Experimentos de clasificación y mapeo en tiempo real.	61
5.5 Clasificación en ambientes desconocidos	66

CAPÍTULO 6 CONCLUSIONES Y TRABAJO FUTURO	71
6.1 Conclusiones	71
6.2 Acerca de la contribución de la investigación	72
6.3 Acerca de las limitaciones	72
6.4 Acerca de líneas futuras de investigación	73
REFERENCIAS BIBLIOGRÁFICAS	75
ANEXO I TABLA DE CARACTERÍSTICAS EXTRAIDAS CON EL SENSOR LÁSER DE PROXIMIDAD	79

ÍNDICE DE FIGURAS

Figura 2.1 Esquema general de un robot	17
Figura 2.2 Barrido del láser URG-04LX	28
Figura 2.3 Transformada de Hough	30
Figura 3.1 Gráfica de diferencia de distancias del barrido láser	32
Figura 3.2 Gráfica del filtrado de las discontinuidades en el barrido láser	33
Figura 3.3 Aplicación de la transformada de Hough	34
Figura 3.4 Detección de esquinas	35
Figura 3.5 Aplicación del algoritmo general para extraer características	37
Figura 4.1 Algoritmo AdaBoost	39
Figura 4.2 Algoritmo AdaBoost Generalizado	40
Figura 4.3 Ejemplo de una posible interacción Humano-robot utilizando la información semántica adquirida	41
Figura 4.4 Ejemplos posibles estructuras de un barrido láser que pueden ser categorizadas semánticamente	41
Figura 4.5 Algoritmo AdaBoost adaptado a la clasificación de lugares usando sensor láser	43
Figura 4.6 Ejemplos de características obtenidas mediante el sensor láser	44
Figura 5.1 Cascada de clasificadores débiles para formar un clasificador fuerte de Lugares	48
Figura 5.2 Características de alto y bajo nivel	48
Figura 5.3 Robot Pac-Ito del Laboratorio de Biorrobótica de la UNAM	49
Figura 5.4 Sistema de control integrado al robot Pac-Ito	49
Figura 5.5 Ambientes de simulación	50
Figura 5.6 Un ejemplo difícil de clasificar	51
Figura 5.7 Lectura en el ambiente de simulación 2	53
Figura 5.8 Lectura en el ambiente de simulación 3	55
Figura 5.9 Trayectoria recorrida en el experimento 7	62
Figura 5.10 Laboratorio de Biorrobótica de la UNAM	62
Figura 5.11 Representación métrico-semántica del recorrido del experimento 7	64
Figura 5.12 Trayectoria recorrida en el experimento 8	65
Figura 5.13 Representación métrico-semántica del recorrido del experimento 8	66
Figura 5.14 Trayectoria recorrida en el experimento 9	67
Figura 5.15 Ejemplo de un ambiente desconocido para el robot	67
Figura 5.16 Representación métrico-semántica del recorrido del experimento 9	68
Figura 5.17 Trayectoria recorrida en el experimento 10	69
Figura 5.18 Representación métrico-semántica del recorrido del experimento 10	70
Figura 6.1 Comparación de Mapa de Activación de árbitro de comportamientos	74

ÍNDICE DE TABLAS

Tabla 4.1 Características del conjunto B	45
Tabla 4.2 Características del conjunto P	45
Tabla 4.3 Características del conjunto M	46
Tabla 5.1 Porcentaje de error de clasificación semántica del experimento simulado 1	52
Tabla 5.2 Porcentaje de error de clasificación semántica del experimento simulado 2	53
Tabla 5.3 Porcentaje de error de clasificación semántica del experimento simulado 3	54
Tabla 5.4 Porcentaje de error de clasificación semántica del experimento simulado 4	55
Tabla 5.5 Porcentaje de error de clasificación semántica del experimento simulado 5	55
Tabla 5.6 Porcentaje de error de clasificación semántica del experimento simulado 6	56
Tabla 5.7 Porcentaje de error de clasificación semántica del experimento simulado 7	56
Tabla 5.8 Porcentaje de error de clasificación semántica del experimento simulado 8	57
Tabla 5.9 Porcentaje de error de clasificación semántica del experimento simulado 9	57
Tabla 5.10 Porcentaje de error de clasificación semántica del experimento simulado 10	57
Tabla 5.11 Porcentaje de error de clasificación semántica del experimento simulado 11	58
Tabla 5.12 Porcentaje de error de clasificación semántica del experimento 1	59
Tabla 5.13 Porcentaje de error de clasificación semántica del experimento 2	59
Tabla 5.14 Porcentaje de error de clasificación semántica del experimento 3	59
Tabla 5.15 Porcentaje de error de clasificación semántica del experimento 4	60
Tabla 5.16 Porcentaje de error de clasificación semántica del experimento 5	60
Tabla 5.17 Porcentaje de error de clasificación semántica del experimento 6	61
Tabla 5.18 Porcentaje de error de clasificación semántica del experimento 7	64
Tabla 5.19 Porcentaje de error de clasificación semántica del experimento 8	66
Tabla 5.20 Porcentaje de error de clasificación semántica del experimento 9	68
Tabla 5.21 Porcentaje de error de clasificación semántica del experimento 10	69

Resumen

Una de las capacidades fundamentales para un robot móvil autónomo es la de tener una representación del ambiente. Los modelos de ambiente métricos, topológicos e híbridos son suficientes para proporcionar a un robot la mayoría de sus funcionalidades básicas de localización y navegación, sin embargo estos modelos no sirven para codificar la semántica del medio ambiente, lo que puede limitar su uso en tareas más complejas como la interacción humano-robot donde se puede decir que se requieren capacidades cognitivas.

En contraste, en el proceso de mapeo de los seres humanos (y animales), referido como mapeo cognitivo [1], se tiene la capacidad de codificar la información semántica del ambiente. Kuipers [2] define al mapa cognitivo como: *“El conjunto de conocimiento de un ambiente de gran escala que se adquiere mediante la integración de observaciones que se reunieron con el tiempo, y se utiliza para encontrar las rutas y determinar las posiciones relativas de los lugares”*. En ese sentido, en este documento se hace énfasis en *“la determinación de posiciones relativas de los lugares”*, mediante la inclusión de descripciones de más alto nivel o conceptos semánticos que tradicionalmente han recibido poca atención en el modelado del ambiente, y que podrían complementar a otro tipo de información, como la métrica o topológica y así poder ampliar capacidades como la localización, la navegación y exploración de un robot móvil, además de facilitar la interacción humano-robot.

En la actualidad, la motivación principal que impulsa gran parte de la investigación en robótica cognitiva [25], es la de construir en un futuro “robots cognitivos autónomos” capaces de realizar tareas similares a los humanos. En ese sentido este trabajo toma como inspiración las formas en que creamos los humanos para representar el espacio. La representación propuesta tiene el potencial de habilitar a los robots no solo a navegar en sus alrededores sino también de conceptualizar el espacio. Típicamente los humanos percibimos el espacio en términos de información de alto nivel tal como objetos, estados y descripciones, así se intenta realizar el mecanismo cognitivo con el cual el robot adquiere y almacena información sensorial, para obtener conceptos espaciales compatibles con los humanos, con el fin de obtener una representación espacial de sus alrededores. Tal representación es métrico-semántica compuesta de las observaciones obtenidas a través de algún algoritmo de mapeo estándar, aumentada con las categorías o clases de lugares. Es así como en este trabajo se piensa en la incorporación de información semántica sobre lugares, como un primer paso hacia una cognición espacial en sistemas de robots móviles.

Organización de la tesis

Capítulo 1 Introducción: Es aquí donde se presentan los elementos necesarios para entender el objetivo de la tesis, la metodología a seguir, así como la contribución y la relevancia del tema desarrollado.

Capítulo 2 Marco Teórico: En este capítulo se realiza una revisión bibliográfica para presentar el estado del arte en modelado del ambiente en robots móviles, además de fundamentar temas o conceptos importantes que se van a utilizar a lo largo de este documento.

Capítulo 3 Extracción de Marcas Naturales Usando Sensor Láser: En este capítulo se aborda la extracción de características del entorno del robot, haciendo uso únicamente de un sensor láser de medición de distancias, dichas características o “marcas naturales” extraídas contribuirán posteriormente a la construcción de una representación del ambiente para el robot móvil.

Capítulo 4 Clasificación de Lugares Usando Sensor Láser: En este capítulo se aborda el tema de la construcción de un clasificador de lugares en base a las características del ambiente, utilizando el algoritmo de aprendizaje supervisado AdaBoost.

Capítulo 5 Construcción de una Representación Métrica-Semántica del Ambiente: Este capítulo está orientado a la construcción de una representación métrico-semántica del ambiente utilizando la información semántica obtenida con el clasificador de lugares, en combinación con un algoritmo de mapeo estándar. Además también se muestran los resultados de experimentos de clasificación, con el fin de demostrar que la representación propuesta es confiable, capaz de generalizar en nuevos ambientes y que esta puede mejorar las capacidades de localización y navegación del robot.

Capítulo 6 Conclusiones y Trabajo Futuro: Finalmente aquí se exponen las conclusiones generales de la tesis presentando propuestas y recomendaciones necesarias para implementar con éxito el sistema propuesto, así como también sugerencias de nuevas líneas de investigación en las que se pueda desarrollar trabajo en un futuro basándose en este documento.

CAPÍTULO 1 - INTRODUCCIÓN

1.1 Objetivo general

Adaptándose a las fortalezas y debilidades de un robot móvil, se realizará el mecanismo cognitivo por el cual el robot adquiere y almacena conocimiento de su ambiente y los procesos por los cuales determina posiciones relativas.

Con el fin de interpretar y describir el ambiente en una forma eficiente y consistente de forma que sea comprensible a los humanos, se sugiere una representación del espacio basada en las características del ambiente, tal representación y la consiguiente conceptualización deben usarse para permitirle al robot tener conocimiento de sus alrededores.

1.2 Contribución y relevancia

Durante ya varias décadas el proceso de mapeo se ha convertido en un área de investigación muy activa en la robótica e inteligencia artificial, en la que se aborda el problema de la adquisición de modelos espaciales de los ambientes físicos a través de robots móviles. El problema de mapeo es generalmente considerado como uno de los problemas más importantes en la búsqueda de la construcción de verdaderos robots móviles autónomos [26].

Para modelar el ambiente generalmente se representa el espacio libre y el espacio ocupado (obstáculos) mediante una representación geométrica: *un mapa*. La construcción de mapas tradicionalmente se ha clasificado en dos categorías: métrico y topológico. Recientemente se ha hecho popular el mapa híbrido, el cual usa información métrica para la precisión de navegación e información topológica para moverse entre lugares. Los mapas métricos, topológicos e híbridos son suficientes para proporcionar a un robot la mayoría de sus funcionalidades básicas para navegar, sin embargo estos mapas no sirven para codificar la semántica del medio ambiente, lo cual puede limitar su uso en tareas más complejas como la interacción humano-robot donde se puede decir que se requieren capacidades cognitivas [25]. En contraste, el proceso de mapeo de los seres humanos (y animales), referido como mapeo cognitivo, es capaz de codificar la información semántica del ambiente.

Kuipers [2] define al mapa cognitivo como: *“El conjunto de conocimiento de un ambiente de gran escala que se adquiere mediante la integración de observaciones que se reunieron con el tiempo, y se utiliza para encontrar las rutas y determinar las posiciones relativas de los lugares”*. Es en ese sentido, que en este documento se hace énfasis en *“la determinación de posiciones relativas de los lugares”*, mediante la inclusión de descripciones de más alto nivel o conceptos semánticos que tradicionalmente han recibido poca atención en el modelado del ambiente, y que podrían complementar a otro tipo de información, como la métrica o topológica y así poder ampliar capacidades como la localización, la navegación y exploración de un robot móvil; además de facilitar la interacción humano-robot. Dichas capacidades son especialmente relevantes para robots diseñados para funcionar en entornos cotidianos para los humanos, que por lo general implican un gran número de lugares y acciones posibles.

El uso de la información semántica es un campo reciente en la robótica, que ha surgido debido a que el nivel de conocimiento espacial en sistemas de robots actuales es bastante modesto. En este trabajo se piensa en la incorporación de información semántica sobre lugares, como un primer paso hacia una cognición espacial en sistemas de robots móviles.

1.2.1 Relevancia

Una de las capacidades fundamentales para un robot móvil autónomo es conocer su posición en el mundo. Esto puede ser representado en términos, localización topológica, coordenadas métricas crudas, o incluso con descripciones semánticas. Debido a ello recientemente, ha habido un creciente interés en aumentar representaciones métricas o topológicas del espacio con información semántica de lugares.

La motivación que impulsa gran parte de la investigación robótica cognitiva en la actualidad es la de construir en un futuro “robots cognitivos autónomos” capaces de realizar tareas similares a los humanos. Algunos ejemplos de aplicaciones de este tipo son los robots para atención a los ancianos, robots de rescate, juguetes interactivos robot para los niños y en general todo tipo de robot de servicio. Para cada una de estas aplicaciones, a parte de la navegación y manipulación el robot tendrá que entender, interpretar y representar el ambiente en una forma eficiente, además de interactuar y comunicarse en forma efectiva y compatible con los humanos.

En resumen, muchas tareas complejas de un robot dependen de que tan bien estén representados sus alrededores [7,8].

Se piensa que el tipo de representación presentada aquí podría en un futuro ser utilizada tanto para localización y navegación como para tareas donde el operador humano sería capaz de transmitir órdenes al robot en términos semánticamente significativos para una interacción humano-robot más natural: por ejemplo, "Robot, tráeme el desayuno a la habitación".

La propuesta presentada en esta tesis puede ser entendida como una solución ingenieril aplicable a robots móviles al problema general de mapeo del entorno en robótica, haciendo uso de la semántica del ambiente. La representación propuesta tiene el potencial para conceptualizar el espacio y por lo tanto se considera como un primer paso hacia una cognición espacial en robots móviles.

1.2.2 Contribuciones

La idea de incluir información semántica en mapas de robot ha sido aceptada durante mucho tiempo, y algunas de las primeras obras importantes fueron realizadas por Chatila y Laumond [27] y Kuipers [28], sin embargo este tema en la investigación todavía se considera una tendencia emergente [29]. De manera general se puede decir que la información semántica permite a un robot inferir nuevo "conocimiento" de su entorno, mejorar la comunicación entre humanos y robots utilizando conceptos, términos comprensibles para las personas [30].

En la presente tesis se aborda el problema de mapeo semántico, que consiste en la utilización de robots móviles para crear mapas que representan no sólo ocupación métrica, sino también otras propiedades del medio ambiente. En contraste con otros trabajos de mapeo semántico [22], se piensa que un enfoque basado en características simples (de bajo nivel) puede ser mejorado con características tales como el número de clusters, líneas rectas (pueden ser paredes), esquinas, discontinuidades detectadas en el barrido láser, porque se considera que es más adecuado al mapeo semántico, debido a que la marcas naturales utilizadas pueden representar objetos físicos reales (o partes de objetos).

Los robots de servicio están diseñados para apoyar a las personas en su vida cotidiana. Esto implica que operan en cercanía de usuarios humanos y comparten el ambiente con ellos. Métodos de localización y mapeo simultáneo (SLAM) permiten a un robot navegar en un entorno y adquirir un mapa al mismo tiempo [23,26]. Este mapa se puede utilizar posteriormente para tareas como la navegación, sin embargo, al ser un mapa puramente métrico difiere a la comprensión del entorno que un usuario pudiera aportar.

Así, el mapa de un robot tiene que integrarse con la representación humana, por lo que se propone utilizar una técnica de aprendizaje supervisado para especificar por medio del entrenamiento la clasificación del medio ambiente que refiere a la opinión de un usuario. Es decir, para agregar la información semántica de las clases de lugares consideradas por los usuarios a los mapas obtenidos a partir de datos sensoriales, se considera aquel escenario en el que un usuario como parte de una instalación inicial del sistema, le da el robot un recorrido del medio ambiente y proporciona orientación a los lugares en el medio ambiente para ser etiquetados, por medio de afirmaciones del usuario como, "este es el pasillo". A partir del aprendizaje de áreas geométricas en base al etiquetado del usuario y en combinación con una técnica de mapeo tradicional se obtiene:

1. La proposición de una representación métrica-semántica del espacio para robots móviles.
2. Una conceptualización del espacio, para robots móviles desde una representación aprendida.
3. Validación y/o corroboración de la compatibilidad del enfoque presentado con nosotros los humanos.

1.3 Objetivos específicos

Se plantean las siguientes metas enfocadas a lo que el robot deberá realizar con el fin de alcanzar el objetivo general de la tesis.

1. **Desarrollar un método efectivo de detección de marcas naturales y objetos que limitan y/o dividen el ambiente (paredes, esquinas, discontinuidades, clusters),** utilizando sensor láser 2D de largo alcance, basándose en un algoritmo de extracción de línea.
2. **Aprender la estructura espacial del ambiente:** Es decir se debe explorar el espacio con el robot y construir un mapa, reuniendo "observaciones" con las cuales se reconocerán lugares en base a características del ambiente (de acuerdo al punto anterior).
3. **De acuerdo a la estructura del ambiente se clasificará en una de cuatro posibles categorías:** Puerta, Pasillo, Espacio Libre y Habitación.

- 4. Realizar experimentos de clasificación y reconocimiento de lugares:** Con el fin de demostrar la aplicabilidad de la representación semántica-métrica, la comprensión del espacio y por consiguiente la obtención de una cognición espacial.

1.4 Metodología de Investigación

La representación del mundo es probablemente uno de los temas más importantes abordados en la literatura robótica debido a la necesidad de contar con un modelo adecuado del medio ambiente para que un robot pueda planificar, razonar, y ejecutar sus tareas. Recientemente, la tendencia en esta área, es hacer uso de información semántica del ambiente para aumentar las capacidades del robot, de tal forma, que en un futuro esto permita el desarrollo de sistemas inteligentes con un entendimiento del medio ambiente común o compatible con el de los humanos, lo cual es fundamental para tareas como la navegación y localización (a través de la conexión de conceptos espaciales humanos a determinados lugares), así como para una posible interacción humano-robot.

El objetivo del mapeo semántico es la creación de mapas que incluyen un significado, tanto a los robots y los seres humanos. Debido a que los seres humanos tienen una representación topológica y jerárquica de su medio ambiente [31], el mapeo semántico en robótica a menudo se presenta con representaciones jerárquicas del medio ambiente, con al menos una capa semántica de información. Uno de los primeros trabajos relacionados con jerarquía en mapeo semántico fue el desarrollado por Kuipers [2], en donde definió el concepto de jerarquía espacial semántica (SSH por sus siglas en inglés). La jerarquía espacial semántica está inspirada en las propiedades de los mapas cognitivos, incluyendo los principios que los seres humanos utilizan para almacenar el conocimiento espacial de áreas de gran escala. Para Kuipers [2,28] el conocimiento espacial describe ambientes y es esencial para ir de un lugar a otro.

Por todo lo anterior se piensa que el “conocimiento espacial” es fundamental para un robot móvil y por lo tanto el primer paso de esta investigación está orientado a desarrollar una representación del espacio para robots móviles, encaminado en particular a resolver estas dos preguntas:

- ¿Cómo puede un robot formar una representación del espacio?
- ¿Cómo un robot puede entender un lugar?

La primera pregunta se dirige a los problemas de extracción de características de alto nivel, mapeo y formación de lugar. En “lugar” se refiere a una abstracción espacial en ambientes interiores, los cuales pueden ser habitaciones. La segunda pregunta puede ser considerada como un problema de cognición espacial.

El segundo paso en la investigación es demostrar que dicha representación del espacio puede ser usada tanto para navegar a través del ambiente, como para habilitar al robot a conceptualizar el espacio, clasificar sus alrededores y reconocerlos cuando sea posible. Finalmente restaría comparar la representación espacial adquirida con otro sistema de mapeo para determinar los beneficios alcanzados usando propiedades o descripciones semánticas del entorno en los robots reales.

1.5 Detalles del sistema

1.5.1 El robot

El robot cuenta con sonares, cámaras y sensores láser, los cuales le permiten detectar y evitar posibles obstáculos en su trayectoria.

1.5.2 El ambiente

El ambiente con el que se experimentara básicamente consiste de tres habitaciones y un pasillo. La primera habitación corresponde al laboratorio de biorrobótica de la UNAM, la segunda habitación corresponde al cubículo Dr. Jesús Savage, jefe del laboratorio de biorrobótica y la tercera habitación corresponde a una oficina.

1.5.3 Implementación

La propuesta de mapeo semántico presentada en esta tesis es el resultado de una combinación de técnicas de aprendizaje maquina con algoritmos de mapeo estándar. Es decir se utilizarán métodos de aprendizaje supervisado para asociar automáticamente las características o propiedades del espacio a los patrones de clasificación deseados. El robot no tiene un conocimiento a priori de un particular arreglo de pasillos, puertas y habitaciones del mapa. Para realizar la implementación practica en el robot real habría que hacer algunas suposiciones acerca de lo que el ambiente del robot tendrá que tener.

1.5.4 Suposiciones

Para la implementación del sistema sobre el robot es necesario introducir algunas limitaciones del ambiente del robot.

- Todos los experimentos son realizados en ambientes interiores, es decir la implementación será válida para ambientes de oficina, simplificando la exploración y la navegación.
- El robot será inicialmente guiado a través de las rutas del ambiente en lugar de permitirle explorarlo autónomamente.
- Información general acerca del ambiente es conocida por el robot tales como tamaño de las puertas. El arreglo particular de pasillos, puertas y habitaciones y la total extensión del ambiente es desconocida por el robot.

En general dichas suposiciones son simplemente hechas para facilitar la implementación a través de los experimentos realizados, limitaciones adicionales en el ámbito del sistema en sí mismo (e.g. que marcas el robot puede reconocer) y sobre las particularidades del ambiente serán discutidas conforme vayan surgiendo.

1.5.5 Sistema de reconocimiento

Este trabajo ofrece una representación de espacio usando distintas características del medio ambiente, para la clasificación de lugares. En particular dos tipos de características son usadas aquí:

- 1.- Características geometrías básicas.
- 2.- Marcas naturales y objetos que limitan y/o dividen el ambiente (paredes, esquinas, discontinuidades).

Dichas características serán detectadas adaptando el barrido láser 2D con algunas técnicas provenientes de áreas como el procesamiento digital de imágenes, el procesamiento de señales (véase capítulo 3), la estadística y la geometría analítica. Durante este documento se referirá al primer tipo de características como de bajo nivel y al segundo grupo como características de alto nivel (véase anexo I).

1.5.6 Representación del ambiente y clasificación de lugar

La representación será métrico-semántica compuesta de las observaciones obtenidas a través de algún algoritmo de mapeo estándar (SLAM), aumentada con las categorías o clases de lugares obtenidas mediante el aprendizaje supervisado.

CAPÍTULO 2 – MARCO TEÓRICO

Con el objeto de fundamentar aspectos importantes para el desarrollo del proyecto, en el presente capítulo se refleja la revisión bibliográfica realizada para describir el marco teórico de este trabajo, ofreciendo así una conceptualización adecuada de los términos que se van a utilizar a lo largo de este documento. Este capítulo comienza con una introducción en temas fundamentales sobre robótica móvil, posteriormente se presenta el estado del arte en modelado del ambiente en robots móviles, haciendo énfasis en el mapeo semántico. Por otro parte se exponen algunas de las herramientas utilizadas en este trabajo y para finalizar se explican brevemente algunas técnicas de procesamiento digital de imágenes, de extracción de características y de procesamiento de señales que serán abordadas en detalle en posteriores capítulos para cumplir con los objetivos de la tesis.

2.1 Robótica

El término robot agrupa las imágenes de maquinas para la realización de trabajos productivos y de imitación de movimientos y comportamientos de seres vivos [9]. La robótica puede ser considerada como otra rama de la ciencia y la tecnología, que ha ido creciendo y evolucionado con el tiempo, y que hoy en día combina diversas ramas de la ingeniería, tales como la mecánica, la informática, la eléctrica y la electrónica; y de la ciencia, tales como física, anatomía, psicología, entre otras [25]. La finalidad de todo esto es la creación de robots inteligentes y autónomos, capaces de razonar, comportarse, evolucionar y actuar como las personas, y que por lo general son concebidos para producir bienes y servicios para los seres humanos.

2.2 Esquema general del sistema robot

En la figura 2.1 se muestra el esquema básico de un robot [9]. En ella se observa un sistema mecánico, actuadores, sensores y un sistema de control. Las tareas y funciones de un robot involucran sistemas de control de movimientos, percepción y planificación.

El sistema de control involucra realimentación de la información suministrada por los sensores internos y externos. Los sensores internos se encargan de la medición de lo referente a desplazamientos, velocidades y fuerzas.

La información suministrada por los sensores externos puede ser usada por el sistema de percepción del robot, para adaptar su comportamiento en función de acontecimientos no esperados producidos en el entorno.

El desarrollo de sistemas de percepción en robótica surge a partir de los progresos tecnológicos en sensores como el láser de proximidad, las cámaras de video, entre otros. Por definición la percepción no sólo involucra la captación de la información sensorial, sino también su tratamiento e interpretación. La complejidad de la percepción artificial depende de la estructura del entorno.

Finalmente la tarea de la planificación de caminos en el entorno de navegación de un robot tiene como objetivo encontrar una trayectoria desde la posición inicial a una posición objetivo.

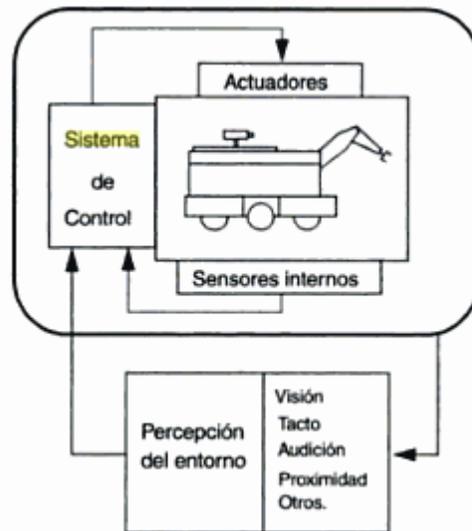


Figura 2.1 Esquema general de un robot.

2.3 Robots Móviles

El desarrollo de robots móviles extiende el campo de la aplicación de la robótica, tratando de incrementar la autonomía limitando todo lo posible la intervención humana. Se trata de que el robot sea lo suficientemente “inteligente” como para reaccionar y tomar decisiones con base en las observaciones de su entorno, sin suponer que se encuentra en un entorno conocido [9].

2.3.1 Planificación en robots móviles

Básicamente la planificación en robots móviles de ambientes interiores puede descomponerse en cuatro tareas principales: planificación de la misión, planificación de la ruta, planificación de la trayectoria y evasión de obstáculos. La misión consiste en determinar a qué habitación hay que desplazarse, la ruta establece el camino desde la posición inicial hacia una posición objetivo en la habitación, definiendo puntos intermedios de paso.

En la actualidad existen diversos métodos de planificación de ruta para robots móviles, algunos de ellos suponen un entorno conocido y estático, donde los movimientos son lentos y ejecución perfecta de la trayectoria. Otros métodos buscan caminos libres de obstáculos, ya sea modelando el espacio libre mediante celdas y/o modelando el entorno mediante polígonos.

La planificación de trayectoria puede ser dinámica, considerando la posición actual del robot y los puntos intermedios de paso definidos en la planificación de la ruta. La trayectoria se corrige debido a variaciones en el entorno no consideradas. Una vez realizada la planificación de la trayectoria, es necesario planificar movimientos concretos y controlar estos movimientos en función de la posición y orientación actual del robot, para obtener la velocidad y el ángulo de dirección aproximados para seguir la trayectoria planificada.

Métodos como el de campos potenciales que permiten integrar la planificación con el control del robot móvil. La idea consiste en determinar la resultante de fuerzas que atraen el robot hacia el objetivo y otro conjunto de fuerzas que lo repelen de los obstáculos [9].

El control del robot móvil requiere la determinación de su posición y orientación en un tiempo mínimo, para lo cual, lo más simple consiste en la utilización de la odometría a partir de las medidas suministradas por los encoders situados en los ejes de movimiento. Sin embargo la inevitable acumulación de error proporcional al desplazamiento, hace necesario el uso del sistema de percepción para reducir significativamente la incertidumbre de la posición. Por ejemplo detectar marcas (naturales o artificiales) en el ambiente, puede ayudar a estimar la posición actual del robot, aunque en entornos complejos, puede aumentar la dificultad del cálculo.

2.3.2 Sistema de percepción en un robot móvil

El sistema de percepción de un robot es imprescindible para modelar el entorno, estimar de forma precisa la posición del robot, y al mismo tiempo, detectar obstáculos lo suficientemente rápido para mejorar la navegación; por lo tanto para su diseño deben tomarse en cuenta la velocidad, las características de precisión y alcance de los sensores utilizados, además de diversas condiciones de navegación de acuerdo a la estructura del entorno.

Con respecto a los sensores utilizados en el sistema de percepción, debido a la posibilidad de interpretación errónea de datos, es necesario tener en cuenta su inmunidad ante vibraciones, ruido o efectos originados por el robot y el entorno. A continuación se describen brevemente algunas características de tres de los principales sensores utilizados en un robot móvil

- La percepción activa mediante láser es un método alternativo que ha cobrado gran importancia en robots móviles [13]. Son dispositivos que obtienen un barrido de imágenes de distancia y reflectancia a las superficies interceptadas por el haz. La ventaja fundamental de un láser frente a otro tipo de sensores es la velocidad de propagación y precisión de sus medidas de distancia.
- Por otra parte las cámaras de video tienen la desventaja de la sensibilidad a las condiciones de iluminación y otros problemas de calibración y fiabilidad, además del costo computacional para la interpretación de las imágenes captadas.
- Los sensores de ultrasonido se basan en la determinación del tiempo de vuelo de un pulso de sonido entre 30 kHz y 1 MHz. Sin embargo, un inconveniente de este tipo de sensores es su decaimiento ante la influencia de las condiciones ambientales, lo que hace necesario una calibración adecuada. A su vez, la relación señal/ruido es normalmente inferior a la de los otros dos sensores, lo cual puede hacer necesario el uso de técnicas de filtrado que implican mayor complejidad computacional.

2.4 Creación de mapas en robótica móvil

Durante ya varias décadas el proceso de mapeo se ha convertido en un área de investigación muy activa en la robótica e inteligencia artificial, en la que se aborda el problema de la adquisición de modelos espaciales de los ambientes físicos a través de robots móviles.

El problema mapeo es generalmente considerado como uno de los problemas más importantes en la búsqueda de la construcción de verdaderos robots móviles autónomos [26]. Para modelar el ambiente generalmente se representa el espacio libre y el espacio ocupado (obstáculos) mediante una representación geométrica: *un mapa*. La construcción de mapas tradicionalmente se ha clasificado en tres categorías: el mapa métrico, el mapa topológico y un híbrido de los anteriores.

2.4.1 Mapas métricos

En este tipo de mapas intenta representar el ambiente del robot usando las características geométricas presentadas en el medio ambiente. Existen dos formas básicas, los mapas de rejilla y los mapas geométricos o basados en características.

2.4.1.1 Mapas de rejilla

Los mapas de rejilla son probablemente la representación del entorno más común usada en aplicaciones interiores de robots móviles. Este tipo de mapa se representa como una matriz en donde cada celda representa una cantidad de espacio en el mundo real y cada celda tiene una probabilidad de estar ocupada.

La principal característica de estos mapas es que son fáciles de construir, sin embargo uno de los principales inconvenientes que se presenta en este tipo de mapas, es que requieren de gran cantidad de almacenamiento. Además se requiere la determinación precisa de la posición del robot para poder utilizarla en la navegación, lo cual es un problema debido a que el ruido de los sensores puede afectar en la localización [32].

2.4.1.2 Mapas geométricos o basados en características

Se representan mediante figuras geométricas básicas en 2 ó 3 dimensiones como pueden ser triángulos, círculos o características que representan marcas en el ambiente del robot, como pueden ser bordes, planos y esquinas.

2.4.2 Mapas topológicos

En los mapas topológicos el ambiente del robot se representa como un *grafo* en donde los nodos corresponden a los lugares y los arcos a las conexiones entre lugares, de tal forma que no dependen de información métrica, sino simplemente de la codificación de datos del lugar para poder llegar de un nodo a otro.

Una de las ventajas de los mapas topológicos es que son compactos, debido a que no almacenan el ambiente completo sino ciertos lugares en específico, lo que permite una planificación eficiente y por lo tanto una disminución de las necesidades de cómputo. Una desventaja con este tipo de mapas es que en ambientes complejos o en ambientes muy grandes, la falta de información métrica o información ambigua proveniente de los sensores puede afectar el reconocimiento del lugar [32].

2.4.3 Mapas híbridos

Recientemente se han hecho populares debido a las deficiencias presentadas al solo usar un tipo de representación, se usa información métrica para “precisión” de la navegación e información topológica para moverse entre lugares.

2.4.4 Mapas cognitivos

Los mapas cognitivos parten de un modelo cognitivo, dicho modelo constituye la representación de un sistema cognitivo. La palabra cognitivo tiene relación con el proceso de cognición. Este proceso se refiere a actividades mentales que: 1) lidian con información abstracta que provienen del mundo real, 2) su representación, y 3) la forma en que se tiene acceso a esta información desde la memoria. Por lo anterior, un sistema cognitivo debe dar un marco que una los fragmentos dispersos de nuestro conocimiento, formando una unidad cohesionada. Un sistema cognitivo debe ser un cuerpo de información estructurado, organizado en consonancia con principios taxonómicos y explicativos que unan esa información en un todo coordinado con bases racionales [50].

Un sistema cognitivo logra simular la realidad, adaptándose a ella en relación a objetivos. El objetivo, por tanto, de un modelo cognitivo es desarrollar una efectiva simulación de la solución de un problema en un determinado dominio desde el punto de vista del humano. En esta técnica el conocimiento se divide en componentes que guardan una relación directa con la forma en que el humano los clasifica y los utiliza [51].

Particularmente, los mapas cognitivos proceden de la teoría del aprendizaje de signos de Edward C. Tolman, quien en su análisis de la conducta del hombre, considera el aprendizaje como la interconexión de conceptos que de alguna manera se relacionan entre sí, donde el cerebro del educando se considera como una sala de control de mapas donde éstos se reorganizan constantemente a fin de producir nuevas respuestas de conducta (cogniciones) [50].

La teoría de Edward C. Tolman se denominó conductismo propositivo en su principal trabajo sistemático, conductismo propositivo en animales y en el hombre, la idea principal de Tolman, consiste en proponer que la conducta está dirigida a una meta, es dócil y varía de acuerdo a las circunstancias ambientales en la búsqueda de una meta dada. Se refiere al tiempo de aprendizaje factor que es central para la conducta. En otras palabras el organismo aprende qué conduce a qué. La docilidad está en términos de aprendizaje, esto es, el organismo aprende qué respuesta ejecutar con el fin de alcanzar su objetivo. Con base en lo anterior elegirá el camino más corto hacia su objetivo. Otra idea de Tolman es que los organismos adquieren conocimientos acerca de su medio ambiente, en donde están localizadas sus metas importantes, y de la forma de ir de un lugar a otro.

El estudio de los mapas cognitivos podría considerarse que solo le atañe al campo de la psicología, sin embargo, con el paso del tiempo se ha observado que es un tema interdisciplinario. En base al trabajo de Tolman en psicología los mapas cognitivos son aquellas representaciones del ambiente que los animales (incluidos los humanos) crean en sus mentes, que denotan relaciones espaciales entre objetos. Por ejemplo las personas se hacen una idea de las distancias entre dos puntos con base en sus mapas cognitivos internos.

En inteligencia artificial tienen un significado más amplio, son redes capaces de adquirir, aprender, codificar y decodificar conocimiento/información, con respecto a eventos causales y la forma en la que es activada [50].

Si bien el nombre de mapa cognitivo se debe al psicólogo Edward Tolman en 1948 [1], se sabe que en el año de 1913 el geógrafo C. Trowbridge definió un “mapa imaginario”, al que se le considera un precedente del mapa cognitivo [46]. Es a partir de 1960 con la escritura de la obra “La imagen de la ciudad” de Kevin Lynch [47], cuando el estudio de la cognición ambiental toma fuerza y es así que junto a los psicólogos, geógrafos, antropólogos y otro grupo de científicos sociales, se incorporan a este tipo de estudios.

Por otro lado diversos investigadores en el área de inteligencia artificial se han enfocado en la aplicación de los mapas cognitivos para representar el ambiente de navegación de un robot móvil. Tomando como base el trabajo de Lynch, el científico estadounidense Benjamin Kuipers a partir de 1978 [5] incorpora una teoría de cognición ambiental al campo de la inteligencia artificial en sus diversos artículos, donde sugiere la existencia de 5 diferentes tipos de información en los mapas cognitivos (topológica, métrica, rutas, características fijas y observaciones) cada una con su propia representación.

Es en el año de 1992 cuando Kortenkamp [4] logró implementar una teoría de mapeo cognitivo en un robot móvil, el cual estaba equipado por 16 sonares y una sola cámara. En el trabajo de Kortenkamp se puede destacar que además de definir nodos en puntos de referencia (comúnmente llamados marcas) y rutas entre los nodos, se formalizó cuando y donde se deben buscar dichos puntos de referencia, y al mismo tiempo contribuyó a crear mapas regionales del ambiente espacial.

Otro trabajo que incorpora la cognición ambiental a la inteligencia artificial es el desarrollado por Yeap et al [6], en donde se refieren a cada espacio local calculado como una “representación de espacios absolutos” (ASR) y el mapa cognitivo construido es una red de espacios locales. El robot adquiere lecturas utilizando sonares hasta que encuentra un obstáculo, en donde aplica un algoritmo de evasión de obstáculos para poder seguir adelante, posteriormente con esto se crean mapas geométricos simplificados con el camino recorrido. Es decir usan un robot equipado con sonares para calcular una descripción de cada espacio visitado. Sin embargo la información obtenida con los sonares no es suficiente, por ello implementan dos estrategias: uso de información de distancia y uso de información de orientación relativa. La primera estrategia se basa en la idea de usar la distancia viajada del robot (en base a los ASR) desde el punto de regreso a la posición actual, precisamente como los humanos tienen una noción basta de que tan lejos han caminado. En lo que se refiere a la estrategia de localización, el robot se basa en un mapa generado previamente.

Por otra parte Vasudevan [7] es pionero en usar información de alto nivel tal como objetos para crear mapas cognitivos, análogamente a la percepción del espacio de las personas, propone una representación jerárquica probabilística para crear dichas representaciones, haciendo uso de SIFT para reconocer objetos y algoritmos con modelos ocultos de Markov para clasificar lugares. Finalmente cabe mencionar que además de los mapas cognitivos relacionados con la representación del medio ambiente en robótica, recientemente han surgido los mapas cognitivos difusos, que sirven para codificar conocimiento/información del mundo real utilizando lógica difusa para representar razonamiento por medio de un Sistema Experto¹ en la toma de decisiones, un ejemplo de la aplicación de este tipo de mapas puede verse en [50].

¹ Un Sistema Experto (SE) es un sistema basado en computadora que integra bases de datos, memorias, mecanismos de razonamiento, agentes, algoritmos, heurísticas, para adquirir, representar, almacenar, generar y difundir conocimientos, inicialmente adquiridos a través de varios expertos humanos dentro de un dominio específico llamado “nube”. Con un Sistema Experto, se pueden dar recomendaciones y/o tomar acciones en las áreas de análisis, diseño, diagnóstico, planeación y control o dar solución a problemas o aplicar técnicas de enseñanza o en general recomendar, actuar y explicar las acciones que hay que tomar en actividades en las cuales normalmente, se requiere del conocimiento o saber de expertos humanos dentro de una nube específica [52].

2.4.5 Mapas Semánticos

En robótica los modelos de ambiente métricos, topológicos e híbridos son suficientes para proporcionar a un robot la mayoría de sus funcionalidades básicas de localización y navegación, sin embargo estos modelos no sirven para codificar la semántica del medio ambiente, lo que puede limitar su uso en tareas más complejas como la interacción humano-robot donde se puede decir que se requieren capacidades cognitivas [25].

Actualmente, existe un creciente interés en aumentar representaciones métricas o topológicas del espacio con información semántica, con el fin de obtener un punto de vista del ambiente o del entorno del robot compatible o “similar” con lo humano, permitiendo el desarrollo de “robots inteligentes” capaces de llevar a cabo una interacción humano-robot de más alto nivel.

Uno de los primeros trabajos relacionados con la inclusión de la semántica en la representación del ambiente fue el desarrollado por Kuipers [2], en donde definió el concepto de jerarquía espacial semántica (SSH por sus siglas en inglés).

La jerarquía espacial semántica está inspirada en las propiedades de los mapas cognitivos, tomando los principios que los seres humanos utilizan para almacenar el conocimiento espacial de áreas de gran escala. Para Kuipers [2,28] el conocimiento espacial describe ambientes y es esencial para ir de un lugar a otro.

En la actualidad han surgido gran cantidad de investigaciones que abordan el tema del mapeo semántico, y por lo general tienen el objetivo de mejorar las capacidades localización del robot, identificando la clasificación del lugar a partir de datos obtenidos de los sensores, utilizando algoritmos de aprendizaje. Un trabajo ampliamente reconocido en este campo es el realizado por Mozos y Burgard [22]. Este trabajo aborda el problema de utilizar solo datos provenientes del barrido láser para obtener una clasificación del lugar en una de tres posibles categorías: habitación, pasillo y puerta. En lugar de requerir la extracción de un conjunto de características de alto nivel, se transformaron los datos crudos obtenidos del barrido láser en características geométricas simples, que sirven como entrada a un algoritmo de boosting, con el fin de obtener un clasificador de lugar fuerte o robusto. Posteriormente en Rottmann [33], este trabajo fue mejorado añadiendo clasificadores provenientes de la visión, además se aplicó un modelo oculto de Markov (HMM) para aumentar la robustez de la clasificación.

Otro trabajo que usa un sensor láser 2D y un robot móvil para crear mapas con anotaciones semánticas se presentó en [34]. En este trabajo se utilizaron dos tipos de clasificadores: un clasificador ad-hoc y un clasificador entrenado con AdaBoost.

Dichos clasificadores se utilizaron para generar representaciones del ambiente métricas aumentadas con la información semántica obtenida por los clasificadores, y se hace una comparación de ambas representaciones.

Una aplicación interesante y diferente de clasificación de lugares fue presentada por Ulrich Weiss y Peter Biber [35], en donde se planteó una clasificación semántica de lugar para mejorar la navegación de robots agrícolas equipados con un sensor láser 3D. Aquí las distintas clases de lugares al aire libre (líneas de siembra, campo abierto, etc.) y las transiciones entre ellas se codifican utilizando un autómata de estados probabilísticos. La aportación de este trabajo consiste en determinar el estado semántico utilizando un autómata que describe las transiciones entre las diferentes clases de lugares.

Por otro lado varias investigaciones en mapeo semántico se han enfocado en la creación de vínculos entre la información geométrica y objetos que limitan y/o dividen el espacio, por ejemplo al etiquetar clases como accesos, puertas, paredes o suelo. Limketkai y sus colaboradores [36] usan redes relacionales de Markov para detectar objetos como puertas basándose en solo en el barrido láser.

Althus y Christensen [37] aplican la transformada de Hough a las lecturas de un sonar para extraer un par de líneas paralelas con el fin de detectar pasillos, a su vez las puertas son detectadas mediante la detección de brechas o discontinuidades entre líneas rectas. Otro trabajo relacionado con la obtención de líneas o planos en el ambiente es el realizado por Alexander J. y B. Trevor [38] en donde el mapa semántico es en realidad un mapa geométrico que utiliza características que se consideran semánticamente significativas en el medio ambiente (paredes y puertas).

Otros investigadores también han considerado la construcción de mapas semánticos basándose en localización de objetos, por medio de “sistemas autónomos e inteligentes” equipados con sofisticados sistemas de visión o de láser, como por ejemplo los presentados por [39] y [40].

El trabajo realizado en [39] describe un sistema robótico móvil que tiene la capacidad de explorar su entorno y reconocer objetos utilizando las imágenes recogidas desde su cámara, para construir un mapa semántico, es decir el mapa es creado a partir de un algoritmo de SLAM y es aumentado con las ubicaciones de los objetos a la vista.

En [40] se adquieren mapas semánticos para uso doméstico, particularmente en cocinas, a partir de los objetos detectados en la nube de puntos 3D. Tal representación de objetos relevantes en el mapa (como aparatos de cocina, armarios, mesas y cajones) es útil para que un asistente robótico pueda realizar tareas complejas de manipulación.

En la actualidad se está trabajando en sistemas robóticos que utilizan multijeraquías semánticas para representar el ambiente por medio de objetos y sus relaciones, la misma que se suele utilizar para mejorar la localización, navegación y la planificación, y así poder realizar tareas cada vez más complejas. Ejemplos de trabajos este tipo de representaciones de alto nivel son [41,42,43].

También se han encontrado algunas otras aplicaciones relacionadas con el mapeo semántico que se han aislado un poco de las demás por ser menos comunes o innovadoras al representar otras propiedades del entorno del robot, que no son tomadas en cuenta por otros investigadores, por ejemplo Denis F. Wolf et al [44] desarrollan técnicas para crear mapas semánticos que representan la actividad y la navegabilidad del medio ambiente. Por otro lado Bipin Suresh et al [45] describen un sistema que construye un mapa con un algoritmo de SLAM y este es aumentado con etiquetas de texto, tales como números de habitación y los nombres de ocupantes de las oficinas detectando el texto de los letreros en las paredes a lo largo de un edificio.

Muchos trabajos en robótica se han enfocado en la obtención de semánticas del ambiente, sin embargo es un área de investigación relativamente nueva [29]. El trabajo presentado en esta tesis está relacionado con otros documentos relativos al tema de mapeo semántico. Está ligado con todos aquellos trabajos que aumentan o anotan un mapa métrico con información semántica para etiquetar lugares, como por ejemplo los presentados por [34,39,45]. Otra referencia clave en este trabajo para la determinación de las etiquetas semánticas, es el uso de características geométricas simples para construir un clasificador fuerte [22], sin embargo, el enfoque de mapeo semántico que se construye aquí, también hace uso de características o marcas semánticamente significativas como lo refieren en [38,40] para obtener una representación del ambiente del robot compatible con los humanos.

2.5 Escáner láser de medición de distancias

La palabra láser responde a las siglas *Light Amplification by Stimulated Emission of Radiation* (amplificación de luz mediante la emisión inducida de radiación). El láser utiliza el principio de tiempo de vuelo para medir distancias, es decir el tiempo de ida y vuelta en que tarda en reflejarse un objeto.

La ventaja fundamental de un láser frente a otro tipo de sensores es la velocidad de propagación. En general, se pueden citar como ventajas de los sensores láser frente a otro tipo de sensores las siguientes: [10]

- En la mayoría de los casos, las medidas pueden ser consideradas como instantáneas. Esto significa que no hace falta tener en cuenta si el robot se desplaza mientras se efectúa la medición, debido a que la velocidad de movimiento del robot es despreciable en comparación con la velocidad de la luz. El sónar, por el contrario, puede requerir compensaciones de movimiento del robot para que la medida de distancia sea más exacta.
- La precisión de un láser es mucho mayor que la de un sónar. Por ejemplo, existen modelos que tienen una desviación estándar de error en la medida de tan sólo 1 mm, la distancia máxima que pueden medir es de 80 m, y la resolución angular es de 0,25°.
- Los datos de un escáner láser pueden interpretarse directamente, puesto que representan distancias a objetos. Esto también se cumple para un sónar, pero no para los sensores de visión, cuyas imágenes son mucho más difíciles de analizar.
- Con respecto a los sensores de visión, el tiempo de procesamiento de los datos de un láser es mucho menor. Una imagen de una cámara de visión está compuesta por millones de píxeles, mientras que en un barrido completo, un láser obtiene del orden de 300 a 600 medidas de distancia.

Como principales desventajas se pueden contar su elevado precio y el hecho de que algunos materiales como el cristal son invisibles al sensor, debido que los pulsos de luz lo atraviesan. Otro inconveniente es que normalmente los datos de un láser se limitan a un plano, si bien esto se puede solventar colocando el sensor sobre un dispositivo pan & tilt, de modo que se puedan obtener datos balanceando y girando el láser en diferentes planos. Debido a sus ventajas con respecto al resto de sensores, el uso de un escáner láser de medición de distancia respalda su uso en este trabajo.

2.5.1 Escáner Láser URG-04LX

El dispositivo láser utilizado para el desarrollo de este trabajo ha sido un URG-04LX [12]. Este modelo escanea su entorno en un plano 2D y está especialmente diseñado para poder trabajar tanto en ambientes interiores, como en ambientes exteriores.

El URG-04LX basa su funcionamiento en la medida del tiempo de vuelo. Para determinar la distancia a la que se encuentra un objeto, el láser emite un pulso de luz infrarroja, cuando el pulso incide sobre el objeto más cercano, regresa hacia el sensor y se determina el tiempo transcurrido.

Conocido el tiempo de ida y vuelta del pulso (tiempo de vuelo), se calcula fácilmente la distancia al objeto detectado. Si los pulsos emitidos no inciden sobre ningún objeto cercano, el láser devuelve la distancia máxima. Al conjunto de medidas que obtiene el láser se le denomina barrido láser. La amplitud del barrido en este modelo es siempre de 240°, como se ilustra en la figura 2.2.

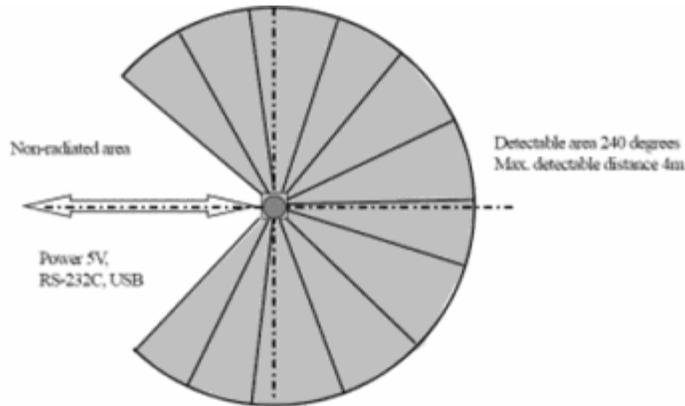


Figura 2.2 Barrido del láser URG-04LX. Aquí se muestran algunas características del barrido del URG-04LX

2.6 La transformada de Hough

Una de las técnicas más populares de extracción de características en el campo de procesamiento digital de imágenes es la transformada de Hough. La transformada de Hough original (1962) [15] es capaz de extraer líneas en un conjunto de puntos en el plano xy , de tal manera que encuentra ecuaciones de la recta que unen secuencias de puntos o píxeles en una imagen de la forma:

$$y = ax + b \quad (1)$$

Considérese un punto (x_i, y_i) , existe un número infinito de rectas que pasan por dicho punto y que satisfacen la ecuación $y_i = ax_i + b$ para valores variables de a y b . Si escribimos la ecuación como $b = -ax_i + y_i$ y consideramos el plano ab , también llamado espacio de parámetros, tendríamos la ecuación de una única recta para un valor fijo de (x_i, y_i) . Un segundo punto (x_j, y_j) , tendrá también una línea en el espacio de parámetros asociada y esta línea interceptará con la línea asociada a (x_i, y_i) , en (a', b') donde a' es la pendiente y b' la altura a la que la recta que contiene a los dos puntos (x_i, y_i) y (x_j, y_j) corta al eje xy . De hecho todos los puntos contenidos en esta línea tendrán rectas asociadas en el espacio de parámetros que se cortaran en (a', b') .

El algoritmo de la transformada de Hough subdivide el espacio de parámetros en **celdas de acumulación**, donde la celda de acumulación $A(i,j)$ esta asociada a la coordenada (a_j, b_j) en el espacio paramétrico, estableciendo los intervalos (a_{\min}, a_{\max}) y (b_{\min}, b_{\max}) dentro de los cuales estarán la pendiente a y la ordenada en el origen b . Para cada punto (x_k, y_k) del plano, se toma cada una de las subdivisiones del eje a y se calcula el valor de b correspondiente, usando la ecuación $b = -x_k a + y_k$. Los valores de b obtenidos se redondean al valor más cercano del eje b . Sí al tomar a_p obtenemos una distancia b_q , se incrementa en uno el valor de la celda $A(p,q)$ en el acumulador, es decir $A(p,q) = A(p,q) + 1$. Cuando acaba este proceso, un valor M en la celda $A(i,j)$ corresponde con M puntos del plano xy , que se encuentran en la recta $y = a_i x + b_j$. Un problema importante surge a la hora de fijar los límites a_{\max} , a_{\min} para recorrer el rango definido por estos valores, ya que a priori no son conocidos.

Otro inconveniente de la transformada original es que utiliza la definición de la recta en la forma punto-pendiente, $y = ax + b$, con lo que el espacio de transformación queda ilimitado, puesto que la pendiente a puede tomar un valor infinito cuando una recta se aproxima a posiciones verticales [16].

Una forma de evitar este problema es utilizar una variante de la transformada de Hough que utiliza la forma normal de la recta, cuya ecuación viene dada por:

$$x \cos \theta + y \sen \theta = \rho \quad (2)$$

Ahora el espacio de parámetros viene determinado por θ y ρ . Esto posibilita barrer todo el rango de valores para el ángulo θ desde -180° a $+180^\circ$ para transformar cada punto (x_i, y_i) ubicado en el plano cartesiano xy , en una curva sinusoidal en el plano $\rho\theta$.

Es decir si en la ecuación 2 se sustituyen los valores de cada punto (x_i, y_i) del plano xy , y por cada punto se sustituye el valor de θ recorriéndolo sobre el intervalo de -180° a $+180^\circ$ para obtener valores de ρ , se puede comprobar haciendo una gráfica con estos valores de ρ y θ , que el resultado es una curva sinusoidal (véase figura 2.5).

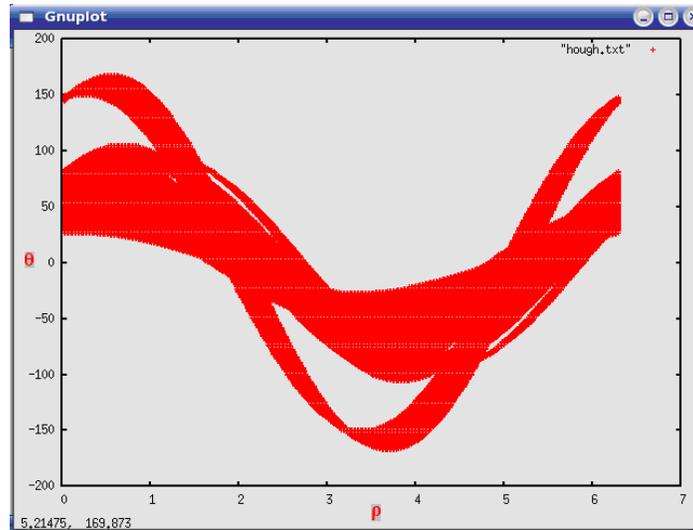


Figura 2.3. Transformada de Hough. En esta grafica muestra que cada punto en el plano xy se transforma en una línea sinusoidal en el plano $\rho\theta$

La peculiaridad de esta transformación está en que si los puntos analizados en el plano xy forman una recta, entonces las curvas sinusoidales generadas por la transformación en el plano Hough $\rho\theta$, se cortarán en un único punto (ver figura 2.5). Se pueden resumir las propiedades de la transformada de Hough [10] utilizando la variante que hace uso de la forma normal de la recta como sigue:

1. Un punto en el plano xy corresponde a una curva sinusoidal en el plano de Hough $\rho\theta$.
2. Un punto en el plano de Hough $\rho\theta$ corresponde a una línea recta en el plano xy .
3. Puntos de una misma recta en el plano xy corresponden a curvas sinusoidales que se cortan en un único punto en el plano Hough $\rho\theta$.
4. Puntos de una misma curva en el plano Hough $\rho\theta$ corresponden a rectas que se cortan en un único punto en el plano xy .

Aunque la transformada originalmente se utilizó para extraer rectas se puede generalizar para identificar formas más complejas, siempre y cuando exista una ecuación en forma paramétrica para dichas formas. La complejidad de la transformada de Hough depende en gran medida, del número de coordenadas y coeficientes de la función a la que se quiera aplicar.

2.6.1 Descripción del algoritmo de la transformada de Hough

- 1) Discretizar el espacio paramétrico (ρ, θ) de la transformada de Hough: identificar la resolución angular para θ , la resolución para la distancia ρ , así como los valores máximos y mínimos para los dos parámetros.
- 2) Crear la matriz Acumulador $A(\rho, \theta)$ e iniciarla en cero.
- 3) Para cada punto (x_i, y_i) hacer
 - 4) Para cada ángulo θ_j hacer
 - 5) Calcular $\rho_i = x_i \cos(\theta_j) + y_i \sin(\theta_j)$
 - 6) Incrementar el acumulador $A(\rho_i, \theta_j)$
 - 7) Fin para
- 8) Fin para
- 9) Encontrar el máximo local en la matriz acumulador $\max(A(\rho_i, \theta_j))$
- 10) Con los valores de ρ y θ correspondientes al máximo del acumulador describir la recta en forma normal.

2.7 Clustering de barrido láser

Se entiende por clustering a un proceso de organización de un conjunto de objetos en grupos cuyos miembros tienen una característica en común.

Un clúster es una colección de objetos que son similares entre sí y son diferentes a los objetos pertenecientes a otros grupos.

En un barrido láser, se puede considerar un cluster como un conjunto de medidas de distancia que se encuentran lo suficientemente cerca entre sí como para determinar que inciden sobre el mismo objeto.

Existen diversos algoritmos de clustering, uno de los más simples y conocidos, es el algoritmo de K-medias [49], el cual utiliza un proceso iterativo para el cálculo de k centroides, organizando los datos con respecto a ellos en k clusters.

La búsqueda de centroides no es comúnmente utilizada en la etapa de clustering de aplicaciones con láser 2D, porque los datos que conforman el barrido láser están ordenados secuencialmente, y no es necesario utilizar un proceso iterativo, sino lo que normalmente se busca es subdividir el conjunto de medidas de distancia en subconjuntos más pequeños, tomando en cuenta las discontinuidades de distancia y con esto obtener los límites de los objetos, en una sola etapa o iteración [10].

CAPÍTULO 3 – EXTRACCIÓN DE MARCAS NATURALES USANDO SENSOR LASER

Este capítulo se centra en abordar la extracción de características del entorno del robot, haciendo uso únicamente de un sensor láser de medición de distancias, dichas características o “marcas naturales” extraídas contribuirán posteriormente a la construcción de una representación del ambiente para el robot móvil. Existen varias marcas que se pueden detectar, sin embargo este apartado se enfoca a detectar discontinuidades, paredes y esquinas. Dichas marcas son detectadas con el barrido láser mediante el uso de algunas técnicas provenientes de áreas como el procesamiento digital de imágenes y el procesamiento de señales, tales como la transformada de Hough que fueron introducidas en el capítulo anterior.

3.1 Detección de discontinuidades

Los enfoques de detección de marcas, se fundamentan en la idea de ubicar puntos o regiones específicas dentro de los datos, que sean representativos del conjunto completo [11]. En el caso de la detección de discontinuidades, la idea básica consta en realizar una división del entorno por medio de un apropiado agrupamiento de las lecturas del barrido láser, identificando aquellos puntos iniciales o finales de cada cluster que pueden ser tomados como discontinuidades en el ambiente (véase figura 3.1).

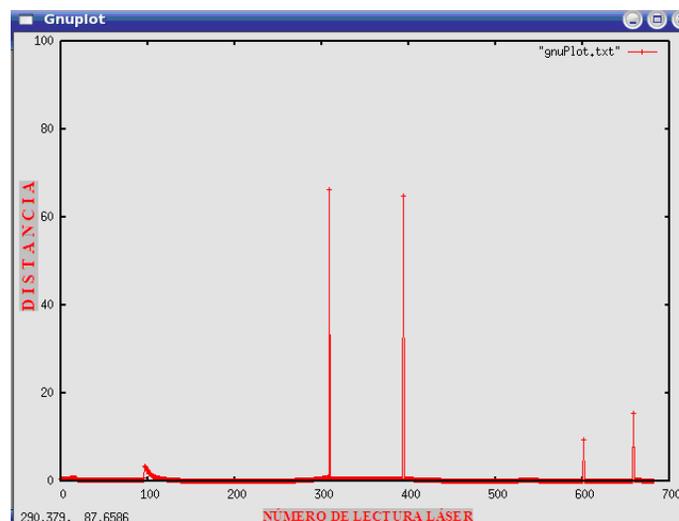


Figura 3.1 Gráfica de diferencia de distancias del barrido láser. Aquí se puede percibir las discontinuidades dentro del barrido láser.

Otra estrategia encontrada en la literatura [10] para encontrar estas divisiones entre clusters puede ser llevada a cabo por medio de un filtro paso alto, es decir se atenúan las frecuencias bajas y se filtran las frecuencias altas, que en este caso corresponden a las discontinuidades. Por lo tanto para obtener estas discontinuidades en el ambiente del robot, se puede operar una señal que se obtiene del barrido láser con un filtro por medio de una suma de convolución [14].

$$y[i] = \sum_{j=0}^{M-1} f[j] x[i-j] \quad (4)$$

Donde el vector $\mathbf{x}=[\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N]$ de tamaño N , corresponde a las lecturas obtenidas por el láser y el vector $\mathbf{f}=[\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_M]$ de tamaño M corresponde a un filtro laplaciano que generalmente es usado para la detección de bordes en imágenes digitales. La figura 3.2 muestra el resultado de aplicar este proceso a las lecturas mostradas en la figura 3.1. Una vez terminado el proceso anterior solo resta buscar los valores distintos de cero en el vector resultante de la convolución $\mathbf{y}[\mathbf{i}]$, dichos valores corresponden a las discontinuidades en el ambiente.

Al extraer este tipo de marcas se determina el punto de partida para aplicar otras técnicas de detección de características como lo puede ser una transformada de Hough para la detección de paredes en el ambiente.

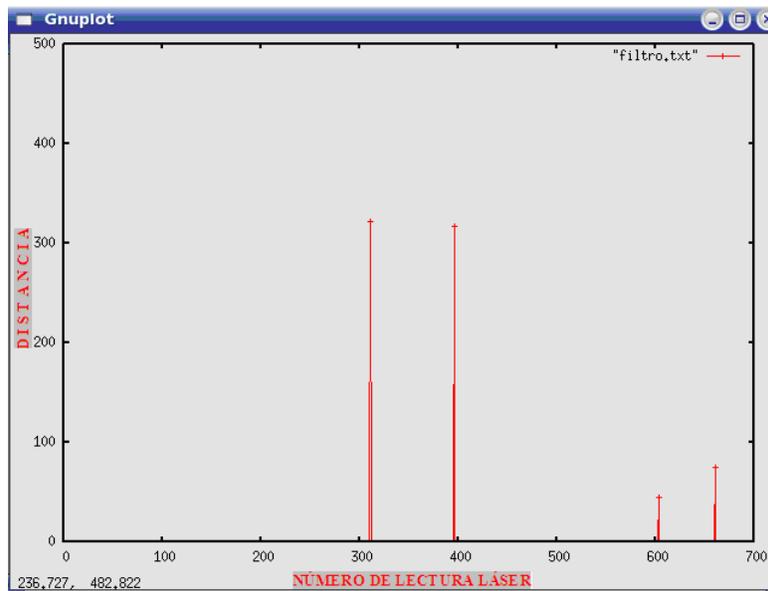


Figura 3.2 Gráfica del filtrado de las discontinuidades en el barrido láser.

3.2 Detección de paredes

Para la detección de este tipo de marcas en el ambiente se utilizó la transformada de Hough, la cual es comúnmente utilizada para detectar líneas rectas en imágenes, por lo que aplicar la transformada de Hough para detectar paredes con un barrido láser es un proceso muy similar y menos costoso computacionalmente, ya que solo hay que transformar cada uno de los puntos del barrido, que generalmente no son muchos. Los pasos que se siguieron para implementar el algoritmo de la transformada de Hough (véase capítulo 2) en el barrido láser se detallan a continuación.

El primer paso consta en discretizar el espacio paramétrico $\theta\rho$, definiendo la resolución angular para θ de 1° , debido a que son 682 lecturas del sensor láser se obtuvieron 682 valores discretos. Para discretizar ρ hay que decidir cual es la máxima distancia desde el origen a donde se espera encontrar una recta, por lo que se optó por una distancia de 4m a una resolución de 1 cm obteniendo así 400 valores discretos para la distancia. Una vez hecho esto el espacio paramétrico quedó definido en el arreglo acumulador $Acc[682][400]$ donde cada elemento del arreglo corresponde con un valor particular de $\theta\rho$. Posteriormente se calcula para cada lectura del láser (x_i, y_i) sus respectivos valores de ρ y θ usando la ecuación en forma normal de la recta (capítulo 2) y para finalizar solo resta buscar en el acumulador las celdas con mayor valor, que corresponden con las rectas encontradas en el ambiente (véase figura 3.3).

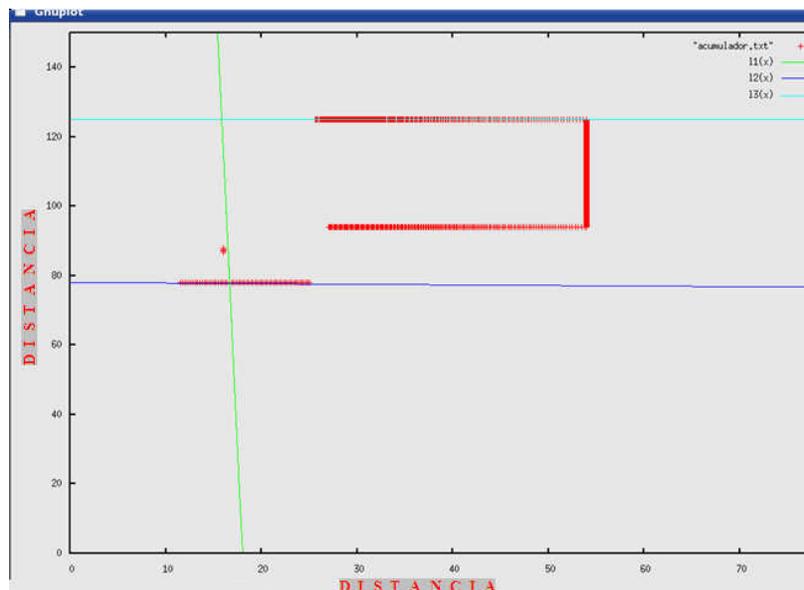


Figura 3.3 Aplicación de la transformada de Hough. Aquí se puede observar la aplicación de la transformada de Hough sobre el entorno simulado.

3.3 Detección de Esquinas

La detección de esquinas se realiza por cada cluster encontrado por el método de detección de discontinuidades. Supongamos que tenemos un cluster que va desde la lectura La_1 hasta La_n , y se traza una línea imaginaria entre estas dos lecturas (ver figura 3.4), posteriormente se recorren todas las lecturas que conforman dicho cluster, buscando la lectura La_k que corresponda al punto más alejado de dicha línea y si la distancia D entre este punto y la recta es mayor cierta distancia D , entonces La_k corresponde a una esquina en el ambiente y se procede a partir el cluster. De manera general se puede seguir este procedimiento hasta que no exista la posibilidad de dividir más el barrido láser.

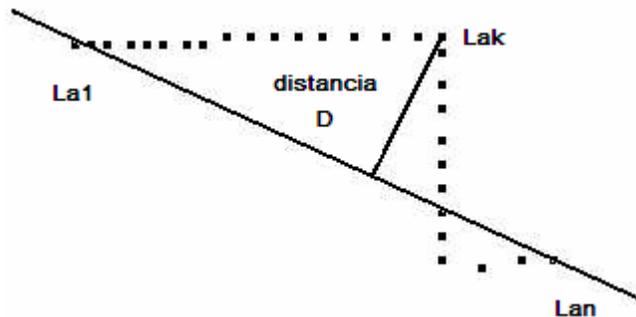


Figura 3.4 Detección de esquinas. Si la distancia D es mayor al umbral el punto La_k corresponde a una esquina.

3.4 Algoritmo General para la extracción de características

Aquí se explica brevemente el algoritmo general para extraer características del ambiente basándose en las tres secciones anteriores.

1. Lo primero es obtener las lecturas del láser y con base en ellas obtener un vector de distancias, es decir dicho vector tendrá las diferencias de distancias entre la lectura i y la lectura $i+1$.
2. Posteriormente este vector de distancias se opera con un filtro paso alto por medio de convolución y así filtrar todas las posibles discontinuidades del sistema.
3. Se obtiene solo la información relevante del vector resultante de la convolución.
4. Con el vector resultante de la convolución se procede a contar el número de discontinuidades detectadas en el ambiente.

5. Una vez que se sabe el número de discontinuidades se pasa al ciclo principal del algoritmo, el cual dependiendo de las discontinuidades encuentra nuevos clusters en el barrido láser.
6. Por cada nuevo cluster mayor al umbral (en este caso 50 puntos) encontrado en el barrido láser se procede a buscar esquinas recursivamente.
7. Por cada esquina encontrada se divide el cluster en uno más pequeño, hasta que no se puede dividir más se procede a aplicar el proceso de la transformada de Hough sobre dicho cluster, que solo encontrará líneas rectas mayores a cierto umbral.
8. Una vez terminado el proceso iterativo anterior se procede a dibujar los resultados del Algoritmo como se muestra en la figura 3.5.

3.5 Pseudocódigo del algoritmo

- 1) Para cada lectura i del láser hacer
 - 2) Calcular la distancia entre dicha lectura i y la lectura $i+1$
 - 3) Guardar el resultado de la diferencia en `vector_distancias`
 - 4) Fin Para
 - 5) Calcular la convolución del `vector_distancias` con un kernel de convolución
`convolucion(vector_distancias[L],kernel_convolucion[M],vector_resultante)`
 - 6) Obtener la información relevante del vector resultante y guardarla en `vector_filtro`
 - 7) Calcular el número de discontinuidades en el `vector_filtro`
 - 8) Para cada valor del `vector_filtro`
 - 9) Si encuentra una discontinuidad
 - 10) Nuevo cluster encontrado
 - 11) Si el tamaño del cluster $>$ `umbral_tamaño`
 - 12) `BuscaEsquina(laser, inicio_cluster, fin_cluster);`
 - 13) Para cada Esquina encontrada
 - 14) Calcular la transformada de Hough
`transformadaHough(laser, posEsquina, fin_cluster)`
 - 15) Fin Para
 - 16) Fin Si
 - 17) Fin Si
 - 18) Fin Para

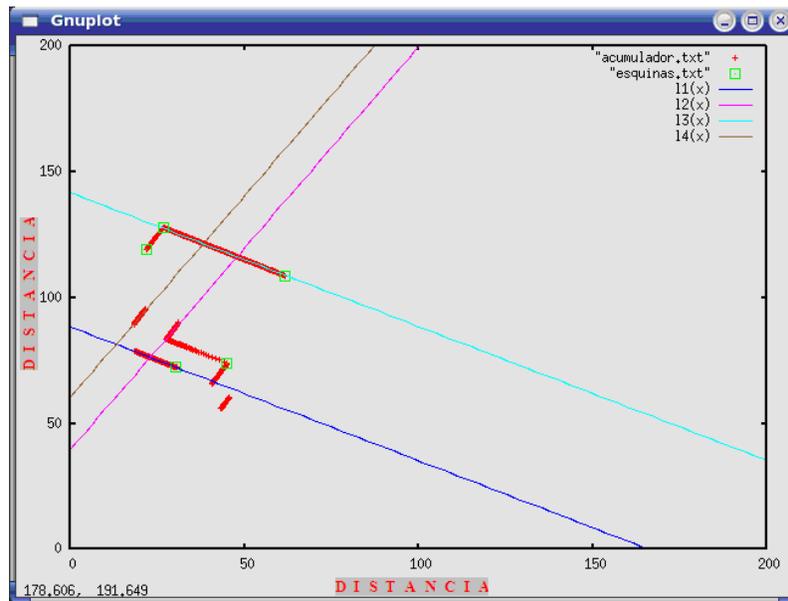


Figura 3.5 Aplicación del algoritmo general para extraer características. Aquí se puede observar el reconocimiento de de paredes y esquinas, en base a las discontinuidades percibidas en un entorno simulado.

CAPÍTULO 4 – CLASIFICACIÓN DE LUGARES USANDO SENSOR LÁSER

En este capítulo se aborda el tema de la construcción de un clasificador de lugares, que será fundamental para la posterior construcción de una representación del ambiente del robot. Para realizar dicha clasificación se ha utilizado el algoritmo de aprendizaje supervisado AdaBoost [17], la idea básica para aplicar dicho algoritmo a este problema es tomar las características extraídas del ambiente con el sensor láser, para construir clasificadores débiles y combinarlos para crear un clasificador fuerte. Así de acuerdo a la estructura del ambiente, puede ser clasificado en cuatro categorías semánticas: puerta, pasillo, habitación y espacio libre.

4.1 Algoritmo AdaBoost

De manera general el algoritmo trabaja de la siguiente forma: dado un conjunto de ejemplos de entrenamiento $E = \{(x_1, y_1), \dots, (x_N, y_N)\}$, donde y_i es la etiqueta de clase asociada a una instancia x_i de un dominio X , AdaBoost llama a un algoritmo de aprendizaje repetidamente (weak learning) durante T iteraciones, por cada iteración se crea una distribución D_t o conjunto de pesos sobre el conjunto de ejemplos de entrenamiento y se regresa una hipótesis débil h_t con un error mínimo ϵ_t con respecto a D_t . Inicialmente, todos los pesos se establecen por igual, pero en cada iteración, los pesos de ejemplos clasificados de forma incorrecta se incrementan, de manera que el algoritmo de aprendizaje se ve obligado a concentrarse en los ejemplos difíciles de clasificar en el conjunto de entrenamiento. Por lo tanto D_t especifica la importancia relativa de cada ejemplo en la iteración actual. Después de T iteraciones, el algoritmo combina las hipótesis débiles en una sola regla de predicción. El pseudocódigo del algoritmo se presenta en la figura 4.1.

Una versión generalizada del algoritmo AdaBoost fue propuesta por Schapire [17, 18] que mejora el algoritmo original básicamente en dos aspectos.

1. El rango de valores para la hipótesis final h_f puede tomar un rango de valores sobre \mathbf{R} y no solo dos valores.
2. La introducción del parámetro α_t que corresponde al peso de la hipótesis h_t , que depende del error cometido por ese clasificador.

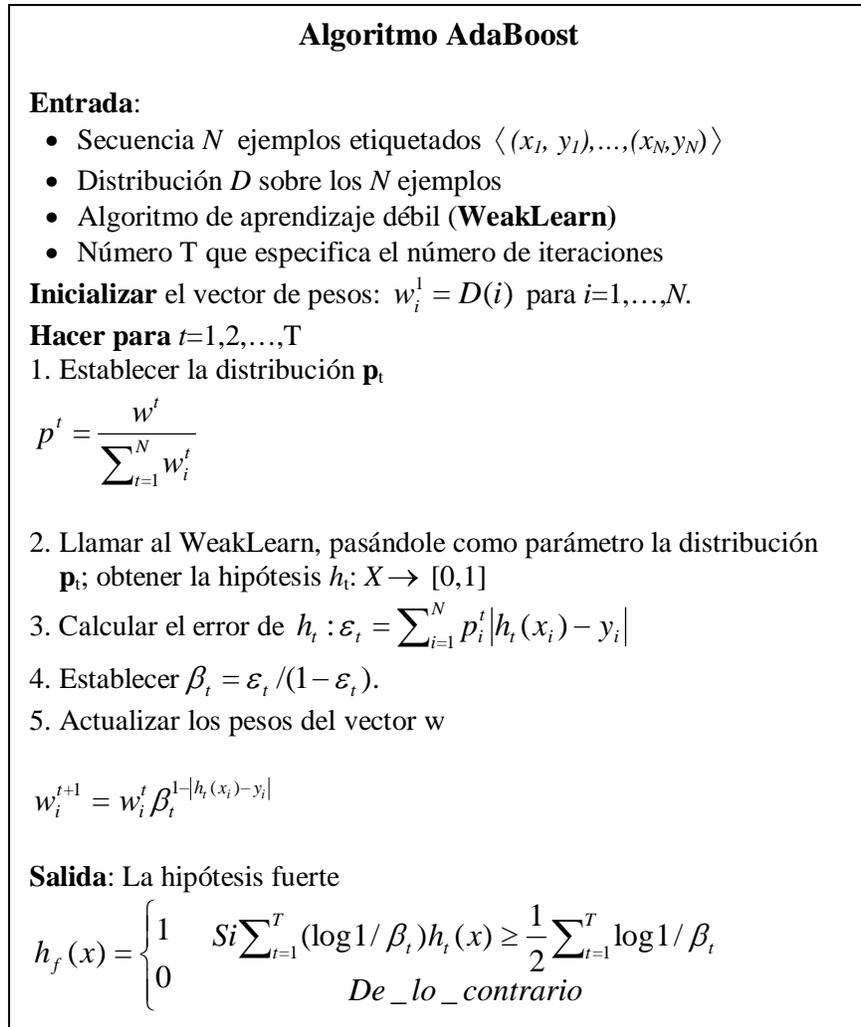


Figura 4.1 Algoritmo AdaBoost.

Por lo tanto α_t mide la importancia que es asignada a la hipótesis h_t , tal como se puede observar en el algoritmo (véase Figura 4.2) α_t aumenta cuando ε_t disminuye, y es así como se le asigna mayor peso a las hipótesis que obtienen menor error. A su vez α_t sirve para aumentar los pesos $\mathbf{D}_t(\mathbf{i})$ de los ejemplos mal clasificados como se puede observar en la etapa de actualización de los pesos al potenciar $\mathbf{D}_t(\mathbf{i})$ mediante la expresión $D_t(i) \exp(-\alpha_t y_i h_t(x_i))$.

Entrada:

- Conjunto de N ejemplos etiquetados $(x_1, y_1), \dots, (x_N, y_N)$,
- Donde $y_i=1$ si el ejemplo x_i es positivo, y $y_i=-1$ si el ejemplo x_i es negativo;
- Weak learner
- Número T de iteraciones

Inicializar los pesos $D_1(i) = 1/N$ para $i=1, \dots, N$.

Para $t=1, \dots, T$

1. Entrenar el weak learner usando la distribución D_t , y obtener la hipótesis $h_t: X \rightarrow \mathbf{R}$.
2. Escoger $\alpha_t \in \mathbf{R}$.
3. Actualizar:

$$D_{t+1}(i) = \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

Donde Z_t es un factor de normalización tal que D_{t+1} sea una distribución

- La hipótesis fuerte final esta dada por:

$$hs(x) = \text{sign} \left(\sum_{i=1}^T \alpha_i h_i(x) \right).$$

Figura 4.2 Algoritmo AdaBoost Generalizado.

4.2 Aplicación del algoritmo AdaBoost para clasificar lugares

Se cree que la capacidad del robot de almacenar información semántica para conceptualizar el espacio es fundamental para que el robot en un momento dado pueda realizar tareas cotidianas e interactuar con los humanos (ver figura 4.3). Como se ha mencionado anteriormente este capítulo esta dirigido al desarrollo de un clasificador que permita categorizar el ambiente del robot de forma semántica, con el fin de incorporar dicha clasificación en una etapa posterior en donde se construirá una representación del entorno o ambiente del robot.

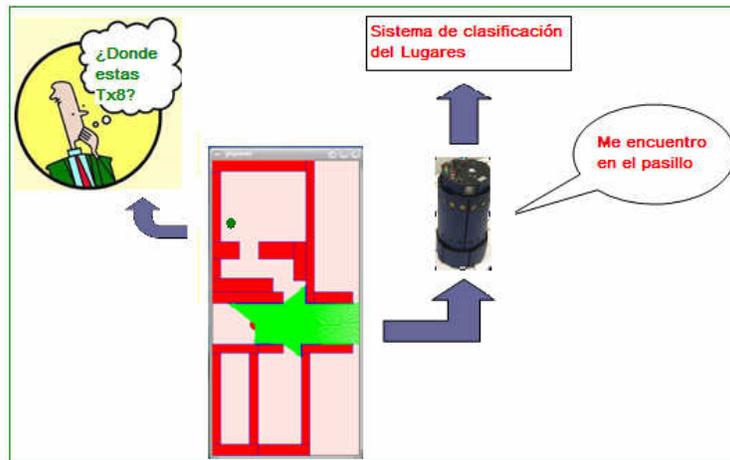


Figura 4.3. Ejemplo de una posible interacción Humano-robot, utilizando la información semántica adquirida.

Los ambientes interiores generalmente están compuestos por lugares o áreas que tienen diferentes funcionalidades tales como oficina, pasillo y puerta; cada uno de ellos cuenta con una estructura diferente por lo que se puede categorizar con un sensor láser.

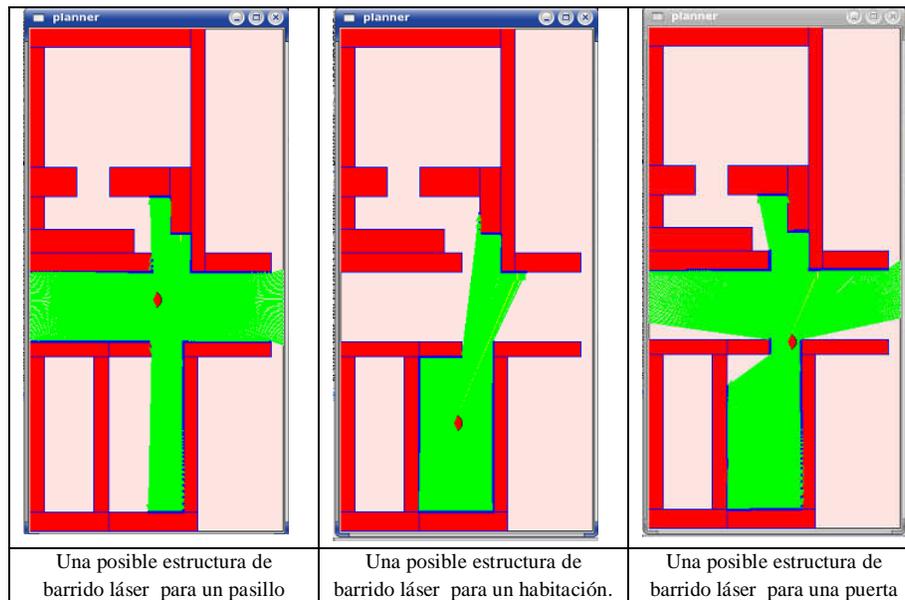


Figura 4.4 Ejemplos posibles estructuras de un barrido láser que pueden ser categorizadas semánticamente.

En la figura 4.4 se muestran ejemplos de que los barridos láser pueden tener cierta estructura característica que los identifica con una posible categoría semántica [21].

Para construir el clasificador de dichas estructuras en el ambiente se utiliza el algoritmo AdaBoost aplicado sobre distintas características extraídas del barrido láser, es decir se utilizan las características para construir clasificadores débiles para combinarlos y encontrar una hipótesis fuerte (véase figura 4.6). Por lo que el algoritmo de aprendizaje (weak learner) debe estar restringido a un conjunto de funciones de clasificación en donde cada una de ellas dependa de una simple característica, para que posteriormente al combinar dichas funciones, se seleccione una función única que mejor separe los ejemplos positivos y negativos del entrenamiento. Dicho de otra forma, el algoritmo de aprendizaje debe determinar el umbral óptimo de la función de clasificación (véase ecuación 4.2), de modo que se minimice el error de ejemplos clasificados erróneamente. La representación de clasificadores débiles usando características es hecha en base a las ideas presentadas por Viola y Jones [19].

Un clasificador $h_j(x)$ consiste en una característica f_j , un umbral θ_j y una polaridad p_j que indican la dirección de la inecuación signo (ecuación 4.1). El algoritmo AdaBoost modificado para la tarea de clasificación de lugares es el mostrado en la figura 4.5.

$$h_j(x) = \begin{cases} +1 & \text{Si } p_j f_j(x) < p_j \theta_j \\ -1 & \text{De_lo_contrario} \end{cases} \quad (4.1)$$

$$(p_j, \theta_j) = \arg \min_{(\theta_j, p_j)} \sum_{n=1}^N |h_j(x_n) - y_n| \quad (4.2)$$

Como lo menciona Paul Viola [19] los clasificadores que utilizan este tipo de características (simples con umbral) pueden ser vistos como simples nodos de árboles de decisión (decision stumps).

4.3 Multclasificación de Lugares

Para realizar la clasificación de lugares se necesitan varias categorías y como se sabe el algoritmo AdaBoost fue diseñado inicialmente para resolver problemas binarios, sin embargo Schapire [17] propuso posteriormente varias versiones para habilitar el AdaBoost a clasificar en múltiples clases entre las que destacan: AdaBoostM1, AdaBoostM2 y AdaBoostMH [20].

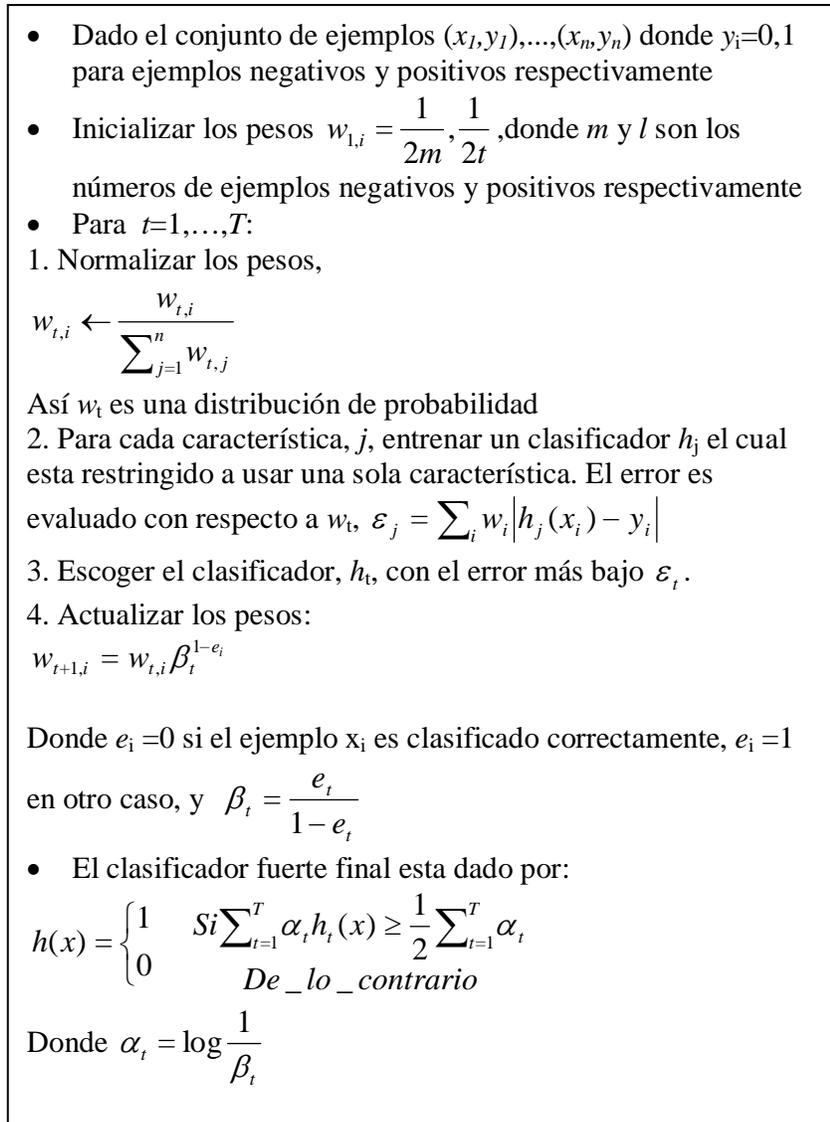


Figura 4.5 Algoritmo AdaBoost adaptado a la clasificación de lugares usando sensor láser.

En este caso la versión del algoritmo utilizada para resolver el problema de categorización semántica de lugares fue AdaBoostMH [20], que enfatiza que para problemas con varias etiquetas de clasificación, es conveniente mantener un conjunto de pesos sobre los ejemplos de entrenamiento y las etiquetas, así durante el progreso del algoritmo los ejemplos de entrenamiento y sus correspondientes etiquetas que son mas difíciles de predecir correctamente obtienen un aumento de pesos, mientras los ejemplos y las etiquetas que son fáciles de clasificar obtienen pesos muy bajos. El efecto que se produce con lo anterior es forzar al algoritmo de aprendizaje a concentrarse sobre ejemplos y etiquetas que son más benéficos para llegar al objetivo de encontrar una regla de clasificación precisa.

4.4 Características del barrido láser

Para el desarrollo del método de clasificación de lugares se ha considerado un conjunto de características geométricas simples [22], además de incluir aquellas extraídas en la etapa anterior (véase capítulo 3) con el fin de hacer todavía más robusto el clasificador. En la siguiente figura se pueden observar algunas de estas características que son utilizadas por el algoritmo AdaBoost para realizar la clasificación semántica.

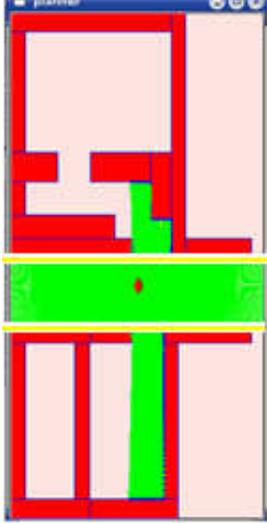
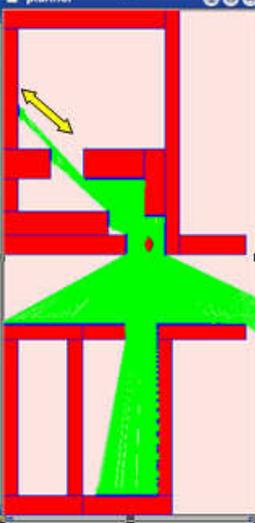
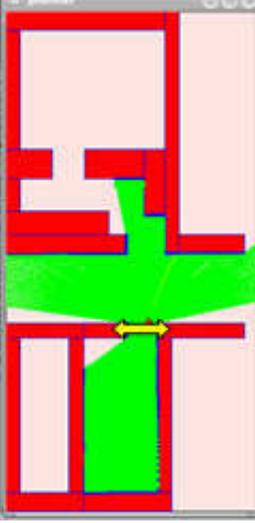
		
<p>Rectas Paralelas:</p> <p>La extracción de líneas rectas de la etapa anterior puede servir para determinar si la estructura del barrido láser pertenece a la clase pasillo.</p>	<p>Distancia entre lecturas consecutivas:</p> <p>Características como estas son combinadas para formar un clasificador fuerte que determine la clase de lugar.</p>	<p>La distancia mínima entre dos mínimos locales:</p> <p>Puede ayudar a los clasificadores débiles a clasificar una estructura tipo puerta</p>

Figura 4.6 Ejemplos de características obtenidas mediante el sensor láser.

Por cada una de las características utilizadas, es calculado un valor numérico a partir del barrido del sensor láser (observación) de 360° . Asimismo cada observación $\mathbf{z}=\{\mathbf{z}_0,\dots,\mathbf{z}_{M-1}\}$ contiene un número de lecturas \mathbf{z}_i , y cada lectura \mathbf{z}_i consiste en una tupla (θ_i, \mathbf{d}_i) donde θ_i es en ángulo de la lectura y \mathbf{d}_i es la distancia con respecto al robot de dicha lectura.

Por lo tanto el conjunto de entrenamiento $T = \{(z_i, y_i), \dots, (z_{M-1}, y_{M-1})\}$ contiene las $M-1$ tuplas ejemplo formadas por una observación z_i y una etiqueta y_i que corresponde a una categoría semántica del ambiente del robot, es decir $y_i \in Y = \{\text{Pasillo, Habitación, Puerta, Espacio Libre}\}$. Aquí se han utilizado algunas de las características definidas en dos conjuntos por Mozos [22]. Las del conjunto **B** (tabla 4.1) corresponden a las calculadas directamente por medio de los rayos x_i en la observación z , mientras que las del conjunto **P** corresponden a características extraídas de una aproximación poligonal del área y perímetro cubierto por z (tabla 4.2).

Características del conjunto B	
1.	El promedio de longitud de las lecturas.
2.	El promedio de la diferencia entre la longitud de dos lecturas consecutivas.
3.	La desviación estándar de la longitud de las lecturas.
4.	La desviación estándar de la diferencia entre la longitud de lecturas sucesivas.
5.	El promedio de la relación entre 2 lecturas.
6.	La desviación estándar de la relación entre dos lecturas consecutivas.
7.	El promedio normalizado de las lecturas.
8.	La desviación estándar normalizada.
9.	La distancia mínima entre dos mínimos locales.
10.	Curtosis de las lecturas.

Tabla 4.1 Características del conjunto **B**. Son obtenidas a partir de las medidas de distancia del barrido láser.

Para complementar las características usadas por Mozos se ha definido un nuevo conjunto **M** de características basadas en el procesamiento de la etapa anterior (véase capítulo 3), que se considera que pueden ayudar en la categorización semántica.

Características del conjunto P	
1.	Área de $P(z)$.
2.	Perímetro de $P(z)$.
3.	Área de $P(z)$ dividida por el perímetro de $P(z)$.
4.	Factor de Forma de $P(z)$.
5.	Circularidad de $P(z)$.
6.	Circularidad normalizada de $P(z)$.

Tabla 4.2 Características del conjunto **P**. Son calculadas a partir de una aproximación poligonal del barrido.

El detalle del cálculo de las características utilizadas para la construcción del clasificador semántico se puede observar en el Anexo I.

Características del conjunto M	
1.	Número de líneas rectas encontradas en el barrido láser.
2.	Número de configuraciones de dos rectas paralelas.
3.	Número de configuraciones de dos rectas perpendiculares.
4.	El promedio de la distancia entre líneas paralelas.
5.	Número de clusters encontrados en el barrido láser.
6.	Número de discontinuidades encontradas en el barrido láser.
7.	Promedio de la diferencia entre discontinuidades.
8.	Número de esquinas encontradas.
9.	Promedio de la distancia entre esquinas.

Tabla 4.3 Características del conjunto **M**. Son calculadas a partir de un proceso de extracción de marcas naturales.

CAPÍTULO 5 –CONSTRUCCIÓN DE UNA REPRESENTACIÓN MÉTRICO-SEMÁNTICA DEL AMBIENTE

Un problema fundamental en el diseño de un robot móvil es la percepción de su medio ambiente. Algunas veces el desempeño de un robot móvil depende básicamente de la precisión, duración y fiabilidad de sus percepciones y el proceso de interpretación en cuestión, sin embargo hay que considerar que se cuenta con recursos limitados. En este caso para el proceso de percepción elemental como el de categorización semántica de lugares, se ha realizado la identificación, aprendizaje y clasificación de patrones observados mediante un dispositivo láser de proximidad. La idea principal es emplear una serie de clasificadores débiles con base en características extraídas del ambiente por medio del láser, con el objetivo de construir un clasificador fuerte lo suficientemente robusto para realizar una clasificación correcta de lugares, lo que posteriormente le brindará la posibilidad al robot de construir una representación semántica de su ambiente.

5.1 Propiedades del sistema

Para el aprendizaje de categorías de lugares, se utiliza el algoritmo de aprendizaje supervisado AdaBoost descrito en el capítulo anterior. Con base en las ideas de Viola y Jones [19], cada clasificador se compone de la combinación de varios clasificadores débiles, formando una cascada, que también se puede ver como un árbol de decisión, en donde cada clasificador débil corresponde a una característica extraída del láser (véase figura 5.1).

A diferencia de otros trabajos [22] las características utilizadas para la construcción del clasificador de lugares, no solo se han utilizado características simples, sino se ha complementado dicho clasificador con un conjunto de características de alto nivel como líneas rectas, esquinas, discontinuidades, número de clusters y conjuntos compuestos de éstos patrones (véase capítulo 3), con el propósito de obtener mejores resultados en el momento de la categorización semántica del ambiente. Incorporando las características de alto y bajo nivel extraídas mediante el láser, el sistema de clasificación quedaría gráficamente ilustrado como se muestra en la figura 5.2.

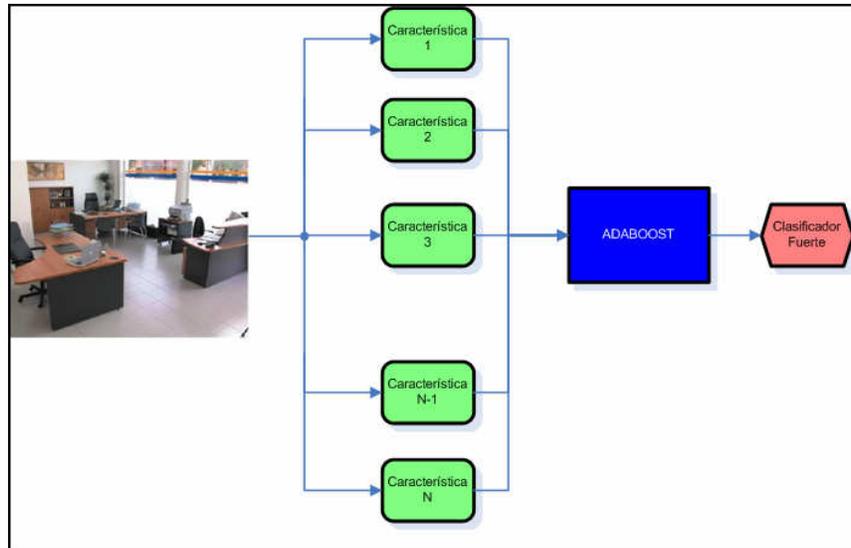


Figura 5.1. Cascada de clasificadores débiles para formar un clasificador fuerte de lugares. Cada clasificador débil corresponde con una característica del ambiente a ser evaluada por el algoritmo AdaBoost, para construir una hipótesis o clasificador fuerte.

Para llevar a cabo los experimentos de clasificación de lugares se empleó el robot móvil de servicio Pac-Ito del laboratorio de biorrobótica de la UNAM, el cual está equipado con un sensor láser de proximidad Hokuyo URG04LX (ver figura 5.3). El movimiento del robot y la adquisición de datos extraídos del barrido láser se controlaron por medio de un joystick analógico de manera remota (véase figura 5.4).

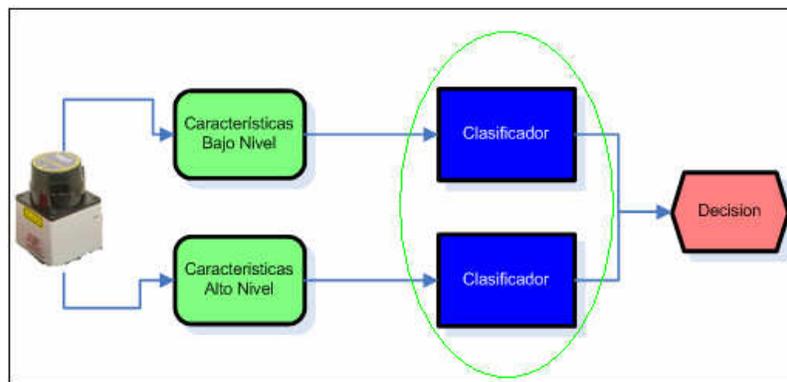


Figura 5.2. Características de alto y bajo nivel. Características de alto nivel, extraídas en una etapa previa de este trabajo se han añadido con el fin de hacer más robusta la clasificación semántica final de los lugares percibidos.



Figura 5.3. Robot Pac-Ito del Laboratorio de Biorrobótica de la UNAM. El robot utiliza un sensor láser de proximidad para percibir información de su ambiente.

El láser Hokuyo tiene un alcance de 240° con una resolución de 0.3515° , por lo tanto para cada lectura del láser se obtienen 682 valores. Con la finalidad de conseguir que cada muestra del barrido láser adquirido fuera de 360° , se optó por tomar el barrido final en dos tiempos, es decir se limitó el alcance de cada barrido a 180° , y se dio media vuelta sobre el mismo eje del robot entre cada toma del barrido, obteniendo un total de 1028 valores para procesar.

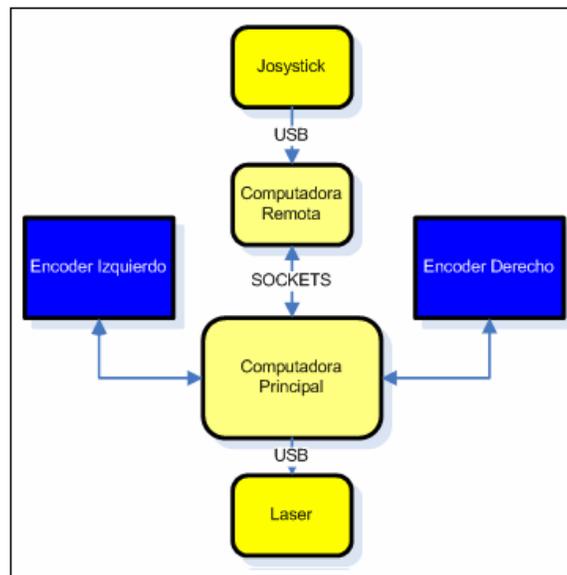


Figura 5.4. Sistema de control integrado al robot Pac-Ito. Por medio de un joystick conectado a una computadora remota se controla el movimiento y el registro de observaciones registradas por el sensor láser, conectado todo a la computadora principal.

5.2 Propiedades del entrenamiento

El ambiente que se destinó para el entrenamiento y prueba de la clasificación semántica, primordialmente consistió del laboratorio de biorrobótica, oficinas y pasillos dentro de las instalaciones del edificio de posgrado de ingeniería de la UNAM. Cabe mencionar que antes de llevar a cabo un entrenamiento del robot en un entorno real, se hicieron pruebas sobre un simulador; los ambientes de simulación utilizados se pueden observar en la figura 5.5, siendo el *Ambiente 1* el que se destinó para el entrenamiento por ser el más parecido al ambiente de navegación real del robot. El entrenamiento consistió en guiar al robot (véase figura 5.4) en el ambiente (simulado o real), obteniendo una secuencia de observaciones de láser, a las cuales se les fue asignando una etiqueta de la categoría semántica de lugar correspondiente a su posición y siempre tomando en cuenta varias situaciones o ejemplos difíciles de clasificar, es decir, someter al robot a situaciones ambiguas en las cuales podría tomar una mala decisión en el momento de clasificar, como por ejemplo una situación en el pasillo donde se encuentre cerca de una o más puertas abiertas (véase figura 5.6) o diferenciar entre una puerta y un espacio entre dos muebles (un acceso) que podría tomarse como un ejemplo de la clase puerta.

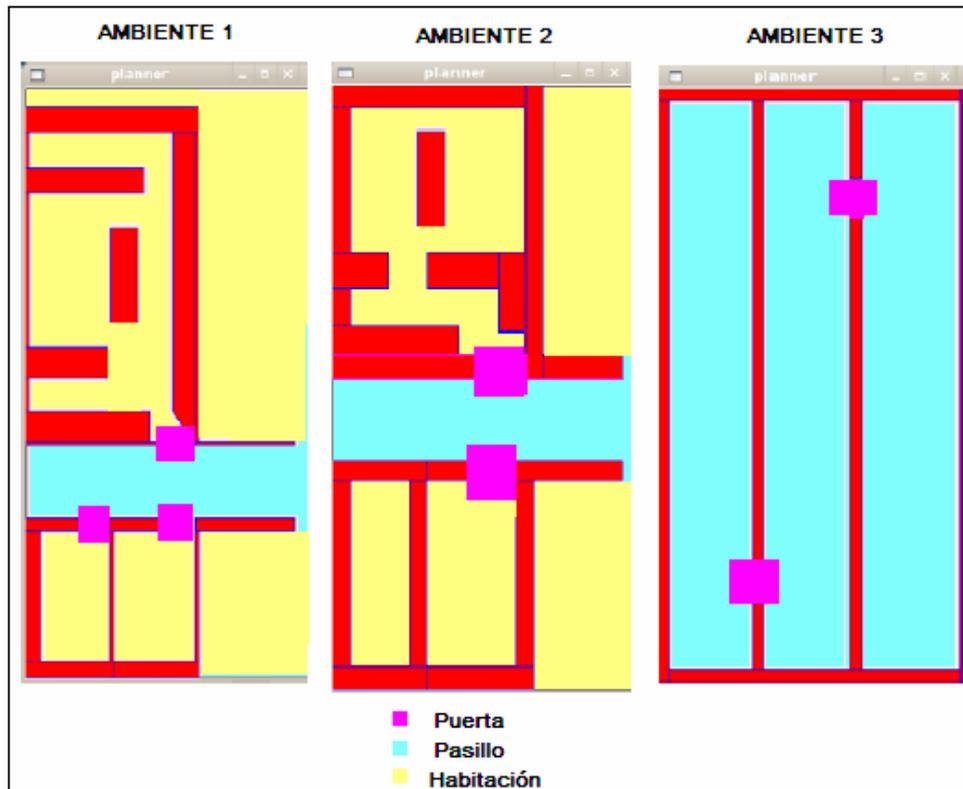


Figura 5.5 Ambientes de simulación. Aquí se muestran los ambientes de simulación utilizados para probar la clasificación de lugares, el ambiente 1 es el más parecido al ambiente real y es el ambiente de entrenamiento.

Cabe mencionar que para este trabajo, se entrenó el reconocimiento de la clase “puerta” suponiendo que el robot se encuentra en frente del marco de la puerta a una distancia no mayor de 50 cm. Hay que mencionar que para el entrenamiento de la clase “espacio libre” se han utilizado observaciones simuladas, provenientes de ambientes donde es casi nula la existencia de puertas, paredes u objetos dentro del rango de incidencia del láser con respecto al robot. Debido a que no es una clase comúnmente encontrada en ambientes interiores, sin embargo es una clase que difiere de las demás y se considera que en un momento dado el robot pueda encontrarse en tal caso y agregando dicha clase se puede evitar una mala clasificación, un ejemplo de esto se puede dar al estar en una habitación muy amplia fuera del rango del láser o simplemente al salir de un edificio a un estacionamiento o quizás al llegar a un campo abierto.

A pesar que en un principio se pensaba en suponer algunas otras propiedades que el ambiente del robot debía cumplir durante el entrenamiento con el fin de corroborar la validez de la clasificación, se han eliminado varias de estas suposiciones cuando se agregaron nuevas características de alto nivel al clasificador. Un ejemplo de lo anterior es que el ancho del pasillo ya no es una variable que afecte al reconocimiento en general de dicha clase, debido a que basta reconocer que existen dos líneas rectas paralelas que lo definen.



Figura 5.6. Un ejemplo difícil de clasificar. En la imagen se muestra al robot ante una situación en donde podría cometer un error en el momento de realizar la clasificación, ya que podría haber dos opciones para hacerlo: puerta y pasillo.

Para llevar a cabo el entrenamiento, el número total de lecturas tomadas por experimento fue dividido en dos conjuntos: 80% se utilizó como el conjunto de ejemplos de entrenamiento (ejemplos categorizados), mientras que se tomó el otro 20 % y se usó como conjunto de prueba (sin categorizar), además se mostrará en otros experimentos los resultados de usar nuevos conjuntos de prueba, es decir se sometió al robot al reconocimiento de lugares no vistos durante el entrenamiento.

5.3 Experimentos

A continuación se presentan los resultados de los experimentos realizados con la finalidad de comprobar que los resultados del entrenamiento realizado para clasificar lugares, hecho a partir de las combinaciones de características de alto y bajo nivel extraídas del láser, es confiable y capaz de generalizar en nuevos ambientes. Los experimentos se han realizado tanto en entornos de navegación reales como en simulados.

5.3.1 Experimentos sobre ambientes simulados

Los siguientes experimentos muestran los resultados de la clasificación de lugares en entornos simulados (véase figura 5.5).

5.3.1.1 Experimento Simulado 1

Para este experimento se utilizó el *ambiente 1* (véase figura 5.5), se entrenó con 100 iteraciones el algoritmo de aprendizaje, teniendo un conjunto de entrenamiento de 200 lecturas y un conjunto de prueba de 40 lecturas. Se ha clasificado sobre el conjunto de prueba usando características de bajo nivel y de alto nivel, y en ambos casos los resultados son satisfactorios, obteniendo un 0% de error en la tasa de clasificación (véase tabla 5.1).

<i>Error Habitación</i>	<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
0%	0%	0%	0%

Tabla 5.1 Porcentaje de error de clasificación semántica del experimento simulado 1.

5.3.1.2 Experimento Simulado 2

Una vez que se han obtenido resultados satisfactorios sobre el clasificador entrenado en el *ambiente 1* se optó por utilizar tal clasificador bajo nuevos ambientes.

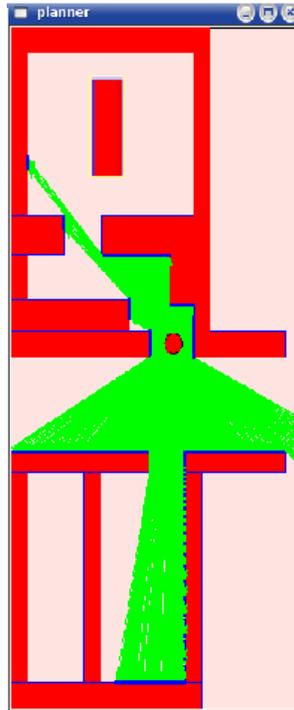


Figura 5.7 Lectura en el ambiente de simulación 2. Aquí se muestra como se obtuvo una lectura correspondiente a la clase puerta en el ambiente 2.

Con el clasificador obtenido en el experimento anterior se clasificó el *ambiente 2* (véase figura 5.7) tomando un conjunto de prueba de 100 lecturas, correspondiendo 25 a la clase habitación, 35 a la clase pasillo y 40 a la clase puerta. Se ha obtenido la clasificación solamente tomando en cuenta características de bajo nivel y los resultados se muestran en la tabla 5.2.

<i>Error Habitación</i>	<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
0%	7.5%	42.85%	18%

Tabla 5.2 Porcentaje de error de clasificación semántica del experimento simulado 2.

5.3.1.3 Experimento Simulado 3

Con el clasificador obtenido en el *experimento simulado 1* se clasificó el *ambiente 2* tomando un conjunto de prueba de 100 lecturas, correspondiendo 30 a la clase habitación, 40 a la clase pasillo y 30 a la clase puerta. Esta vez se han utilizando características de bajo nivel y alto nivel para construir el clasificador, se puede observar en la tabla 5.3 que los resultados mejoran al clasificar el pasillo.

<i>Error Habitación</i>	<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
0%	6.6%	27.5%	13%

Tabla 5.3 Porcentaje de error de clasificación semántica del experimento simulado 3.

5.3.1.4 Experimento Simulado 4

Hasta el momento en la clasificación de la habitación se han obtenido resultados satisfactorios, teniendo un porcentaje de error más significativo en la clase puerta y en la clase pasillo. Por lo tanto en este experimento se utiliza el *ambiente 3* (véase figura 5.8) que solo cuenta con dos categorías semánticas: pasillo y puerta. Se ha tomado un conjunto de prueba de 100 lecturas, correspondiendo 50 a la clase pasillo y 50 a la clase puerta. Se han utilizando características de bajo nivel para construir el clasificador y los resultados se muestran en la tabla 5.4.

<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
36%	46%	41%

Tabla 5.4 Porcentaje de error de clasificación semántica del experimento simulado 4.

5.3.1.5 Experimento Simulado 5

Con el clasificador obtenido en el *experimento simulado 1* se clasificó el *ambiente 3* tomando un conjunto de prueba de 100 observaciones, correspondiendo 60 a la clase pasillo y 40 a la clase puerta.

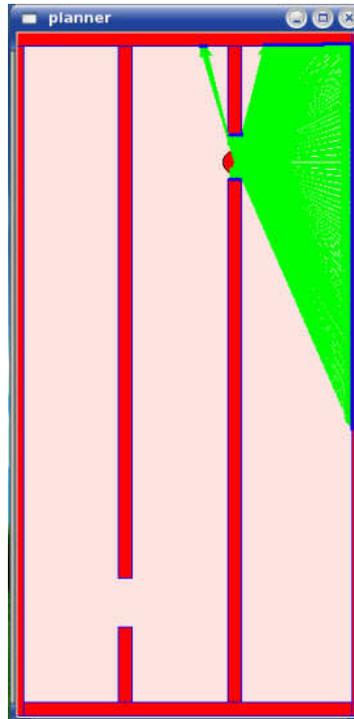


Figura 5.8 Lectura en el ambiente de simulación 3. Aquí se muestra como se obtuvo una lectura correspondiente a la clase puerta en el ambiente 3.

Esta vez se han utilizando características de bajo nivel y alto nivel para construir el clasificador, se puede observar en la tabla 5.5 que los resultados del mejoran un poco con respecto al experimento anterior.

<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
27.5%	28.33%	38%

Tabla 5.5 Porcentaje de error de clasificación semántica del experimento simulado 5.

5.3.1.6 Experimento Simulado 6

En los experimentos anteriores se han utilizado entornos simulados y en general se han obtenido buenos resultados, se piensa que estos entornos están bien estructurados y por lo tanto la complejidad de la clasificación se ha visto reducida, por lo que a partir de este experimento se ha considerado tener en cuenta en la simulación las posibles vibraciones, ruido o efectos originados por el robot y el medio ambiente que se pudieran generar en un ambiente real.

Con el clasificador obtenido en el *experimento simulado 1* se clasificó el *ambiente 1*, tomando un conjunto de prueba de 100 lecturas, a las que se les ha añadido ruido, intentado con esto que las observaciones realizadas sean más parecidas a las de un entorno de navegación real. Se tomaron 40 observaciones para la clase pasillo, 25 para la clase habitación y 35 de la clase puerta, clasificando con base en características de bajo nivel. Los resultados se muestran en la tabla 5.6.

<i>Error Habitación</i>	<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
4%	5.71%	22.5%	13%

Tabla 5.6 Porcentaje de error de clasificación semántica del experimento simulado 6.

5.3.1.7 Experimento Simulado 7

Con el clasificador obtenido en el *experimento simulado 1* se clasificó el *ambiente 1*, tomando un conjunto de prueba de 100 lecturas a las que se les ha añadido ruido, de las cuales 40 observaciones corresponden a la clase pasillo, 25 para la clase habitación y 35 de la clase puerta, y se han utilizado características de bajo y alto nivel para construir el clasificador. Los resultados se muestran en la tabla 5.7.

<i>Error Habitación</i>	<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
16%	5.71%	12.5%	11%

Tabla 5.7 Porcentaje de error de clasificación semántica del experimento simulado 7.

5.3.1.8 Experimento Simulado 8

Con el clasificador entrenado del *experimento simulado 1* se clasificó el *ambiente 2*, tomando un conjunto de prueba de 100 lecturas, a las que se les ha añadido ruido, intentado con esto que las observaciones realizadas sean más parecidas a las de un entorno de navegación real. Se tomaron 35 observaciones para la clase pasillo, 35 para la clase habitación y 30 de la clase puerta, clasificando con base en características de bajo nivel. Los resultados se muestran en la tabla 5.8.

<i>Error Habitación</i>	<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
11.42%	16.42%	37.14%	22%

Tabla 5.8 Porcentaje de error de clasificación semántica del experimento simulado 8.

5.3.1.9 Experimento Simulado 9

Con el clasificador entrenado en el *experimento simulado 1* se clasificó el *ambiente 2*, tomando un conjunto de prueba de 100 lecturas a las que se les ha añadido ruido, de las cuales 35 observaciones corresponden a la clase pasillo, 35 a la clase habitación y 30 a la clase puerta, pero esta vez se ha clasificando tomando en cuenta características de bajo y alto nivel. Los resultados se muestran en la tabla 5.9.

<i>Error Habitación</i>	<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
31.42%	6.6%	25.71%	22%

Tabla 5.9 Porcentaje de error de clasificación semántica del experimento simulado 9.

5.3.1.10 Experimento Simulado 10

Aquí se prueba el clasificador de lugares obtenido en *experimento simulado 1* en el *ambiente 3* (véase figura 5.5) que solo cuenta con dos categorías semánticas: pasillo y puerta. Se ha tomado un conjunto de prueba de 100 lecturas a las que se les ha añadido ruido, correspondiendo 35 a la clase puerta y 65 a la clase pasillo, que hasta el momento es en donde se ha obtenido un mayor índice de error. Las características utilizadas para construir el clasificador son solo de bajo nivel. Los resultados se muestran en la tabla 5.10.

<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
14.28%	49.23%	37%

Tabla 5.10 Porcentaje de error de clasificación semántica del experimento simulado 10.

5.3.1.11 Experimento Simulado 11

De la misma forma que en el experimento anterior se ha utilizado el clasificador entrenado en el *experimento simulado 1* y se clasificó el *ambiente 3* tomando un conjunto de prueba de 100 lecturas, añadiendo ruido y clasificando con base en características de bajo y alto nivel. Las observaciones se han dividido en 65 para la clase pasillo y 35 para la clase puerta. Se puede observar en la tabla 5.10 que la tasa de error de clasificación de pasillo mejora, sin embargo aumenta el error de la clasificación de puerta.

<i>Error Puerta</i>	<i>Error Pasillo</i>	<i>Error Global</i>
28.57%	38.46%	36%

Tabla 5.11 Porcentaje de error de clasificación semántica del experimento simulado 11.

5.3.2 Experimentos sobre ambientes reales

Hasta el momento los experimentos se han realizado con entornos simulados, obteniendo un error de clasificación aceptable, se ha tomado en cuenta en la simulación las posibles vibraciones o efectos generados por el ambiente y el robot en un ambiente real, añadiendo *ruido* a las observaciones adquiridas con el simulador. A su vez también se ha probado que los clasificadores semánticos son capaces de generalizar en nuevos ambientes.

A partir de aquí las pruebas se llevarán a cabo tomando muestras de observaciones en el ambiente de navegación real, los primeros 6 experimentos tienen el objetivo de obtener un clasificador que sea lo suficientemente robusto, en los últimos experimentos están orientados a utilizar dicho clasificador para construir o aumentar con etiquetas semánticas un mapa en tiempo real, dicho mapa será llamado semántico o representación métrica-semántica del ambiente.

5.3.2.1 Experimento 1

Los resultados que se muestran en la tabla 5.12 corresponden a un entrenamiento de 25 iteraciones del algoritmo de aprendizaje con un conjunto de entrenamiento de 300 lecturas del ambiente utilizado y un conjunto de prueba de 60 lecturas, teniendo 15 ejemplos para cada clase y usando solo características de bajo nivel para la clasificación semántica. Se presentan los porcentajes de error por clase y se obtiene un error global de la clasificación.

<i>Error Pasillo</i>	<i>Error Espacio Libre</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
46.66%	0%	33.33%	20%	23.33%

Tabla 5.12 Porcentaje de error de clasificación semántica del experimento 1.

5.3.2.2 Experimento 2

En este experimento se han tomado el mismo conjunto de lecturas que el del ejemplo anterior, pero aumentando el número de características, tomando en cuenta también las de alto nivel. Los resultados se muestran a continuación.

<i>Error Pasillo</i>	<i>Error Espacio Libre</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
20%	0%	40%	13.33%	18.33%

Tabla 5.13 Porcentaje de error de clasificación semántica del experimento 2.

5.3.2.3 Experimento 3

Aquí el número de ejemplos de entrenamiento fue de un total de 400, por lo tanto se tomaron 80 lecturas para el conjunto de prueba, el cual esta dividido en 25 ejemplos para pasillo, 25 para puerta, 20 para habitación y 10 espacio libre. El entrenamiento se realizó durante 50 iteraciones, tomando en cuenta solo las características de bajo nivel.

<i>Error Pasillo</i>	<i>Error Espacio Libre</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
8%	0%	12%	5%	7.5%

Tabla 5.14 Porcentaje de error de clasificación semántica del experimento 3.

5.3.2.4 Experimento 4

Aquí el número de ejemplos de entrenamiento fue de un total de 400, por lo tanto se tomaron 80 lecturas para el conjunto de prueba, el cual esta dividido en 30 ejemplos para pasillo, 20 para puerta, 25 de habitación y 5 de espacio libre. El entrenamiento se realizó durante 50 iteraciones, tomando en cuenta características de bajo y alto nivel.

<i>Error Pasillo</i>	<i>Error Espacio Libre</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
3.33%	0%	10%	4%	5%

Tabla 5.15 Porcentaje de error de clasificación semántica del experimento 4.

5.3.2.5 Experimento 5

De la misma forma aquí se han usado 400 ejemplos para el conjunto de entrenamiento, y 80 para el conjunto de prueba, sin embargo esta vez se hizo énfasis en añadir más ejemplos de la clase puerta, para minimizar el error de clasificación de dicha clase, tomando 35 ejemplos para dicha clase, 20 para la clase pasillo, 20 para la clase habitación y 5 para espacio libre. También se ha aumentado el número de iteraciones a 80 y solo se han utilizado características de bajo nivel. Los resultados se muestran en la tabla 5.16.

<i>Error Pasillo</i>	<i>Error Espacio Libre</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
5%	0%	8.57%	5%	6.25%

Tabla 5.16 Porcentaje de error de clasificación semántica del experimento 5.

5.3.2.6 Experimento 6

En el presente experimento el número de ejemplos de entrenamiento fue de de 400, el conjunto de prueba de 80 lecturas se dividió en 25 ejemplos para pasillo, 40 para puerta, 10 habitación y 5 espacio libre. El entrenamiento se realizó durante 80 iteraciones, tomando en cuenta características de bajo y alto nivel. En la tabla 5.17 se muestran los resultados.

<i>Error Pasillo</i>	<i>Error Espacio Libre</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
8%	0%	7.5%	0%	6.25%

Tabla 5.17 Porcentaje de error de clasificación semántica del experimento 6.

5.4 Experimentos de clasificación y mapeo en tiempo real

Una vez que se han tomado las lecturas correspondientes al ambiente del robot y se han probado diferentes tipos de clasificadores semánticos, lo que resta es seleccionar el clasificador más conveniente para construir o aumentar con etiquetas semánticas un mapa en tiempo real, dicho mapa será llamado semántico o representación métrica-semántica del ambiente.

Este tipo de representación es próxima a como los humanos perciben y clasifican el entorno, en el sentido de que por medio de la percepción de información proveniente del láser, el robot reconoce su alrededor o el lugar en donde se encuentra, dándole así una cierta capacidad de localización y almacenamiento de su ambiente, lo cual le provee la posibilidad para poder llevar a cabo una interacción con el humano, que posteriormente le será útil para otras realizar otras actividades cotidianas.

Para verificar la capacidad del clasificador se ha probado en varias rutas o trayectorias y sobre de ellas se ha realizado la clasificación semántica, aunque también se podría extraer rutas o trayectorias entre etiquetas semánticas y en un futuro construir más rutas dentro del mapa, útiles para la navegación del robot, como lo propone Martínez Mozos [21]. En lo que refiere a la propuesta presentada en este trabajo, se ha optado por una representación espacial más reducida, más enfocada a la localización y la categorización semántica del ambiente. Para realizar el mapa semántico se ha combinado la técnica del etiquetado semántico con SLAM [23] y se ha optado por implementar un algoritmo pequeño y eficiente propuesto por [24].

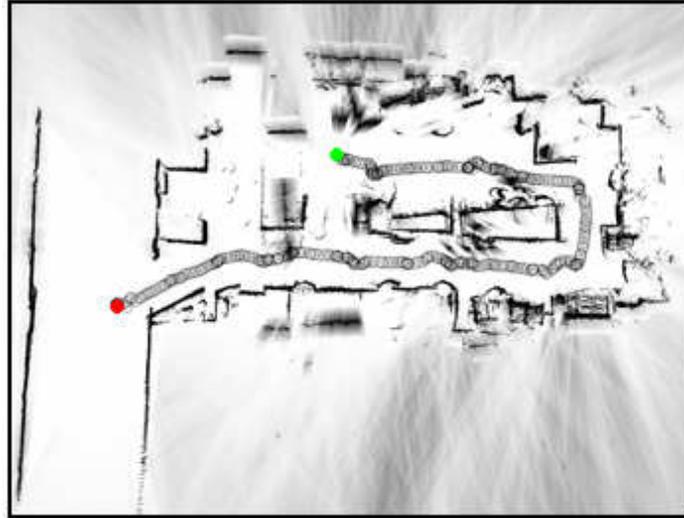


Figura 5.9. Trayectoria recorrida en el experimento 7. Se muestra la trayectoria, iniciando el laboratorio (punto verde) y terminando en el pasillo (punto rojo).

5.4.1 Experimento 7

Para mostrar que la clasificación semántica funciona de manera eficiente y que puede esta puede ser ejecutada en tiempo real, se procedió a guiar al robot a través de varias habitaciones, puertas y pasillos del edificio de posgrado de ingeniería de la UNAM, empezando por un ambiente “conocido” o más bien dicho el utilizado para extraer los ejemplos de entrenamiento del clasificador semántico. Este recorrido se muestra en la figura 5.9 y como se puede observar se realizó teniendo como punto de partida el laboratorio de biorrobótica de la UNAM (véase figura 5.10) y terminando en pasillo del segundo piso del edificio de posgrado.



Figura 5.10. Laboratorio de Biorrobótica de la UNAM.

Hay que mencionar que se intenta obtener una representación del ambiente aproximada con lo humano, por lo tanto se considera que lo principal es demostrar que dicha representación es eficiente para una localización espacial “global” en base a las características percibidas del ambiente y que cuando se habla de “punto de partida como el laboratorio” no se esta refiriendo específicamente a un punto en el eje de coordenadas, si no es simplemente una expresión para indicar cualquier punto ubicado dentro de la habitación, tal y como lo entienden los humanos, es decir los humanos no saben en que posición exacta en el eje de coordenadas se localizan, sin embargo se puede expresar semánticamente utilizando frases similares a *“me encuentro en frente de la puerta”* o *“te espero en el pasillo”*.

Se ha manifestado que se ha empleado SLAM para obtener una representación métrica del entorno de navegación del robot, en particular se utilizó el algoritmo CoreSlam [24], con el cual básicamente se obtiene un mapa de proyecciones de cada haz del láser que colisiona con los obstáculos presentes en el entorno y esta basado en el método de Monte Carlo para estimar la posición del robot a partir de las desplazamientos medidos por la odometría y las observaciones del espacio obtenidas mediante el láser de proximidad. Sin embargo en este trabajo la parte de localización inmersa en el algoritmo se ha dejado de lado en el sentido de que la localización se hace semánticamente, quedándose solo con lo referente al mapeo de lo percibido durante el recorrido, para ser integrado con el etiquetamiento semántico del ambiente y construir de esta forma lo que se ha llamado una representación métrica semántica del ambiente o mapa métrico del entorno aumentado con etiquetas semánticas. Solo esta vez y por la representación semántica se ha hecho de esta forma, sin embargo como se mencionara más adelante, para investigaciones futuras y para desarrollar sistemas autónomos más robustos se puede integrar la localización de la odometría con la localización semántica.

En la figura 5.11 se puede observar la representación métrico-semántica que el robot “construyo” durante el recorrido del experimento 7. Las formas geométricas dibujadas en el mapa que se han utilizado pertenecen a las distintas categorías semánticas percibidas por el robot en cada momento que el robot examino las características del ambiente extraídas con el sensor láser; los círculos corresponden a una habitación, los cuadrados rellenos a pasillo y las cruces a puertas abiertas. Para este recorrido se ha utilizado el clasificador generado durante el experimento número 4. Como se puede ver en el mapa, durante el recorrido adentro del laboratorio existen varios errores en donde el robot confunde la clase *habitación* con la clase *puerta*.

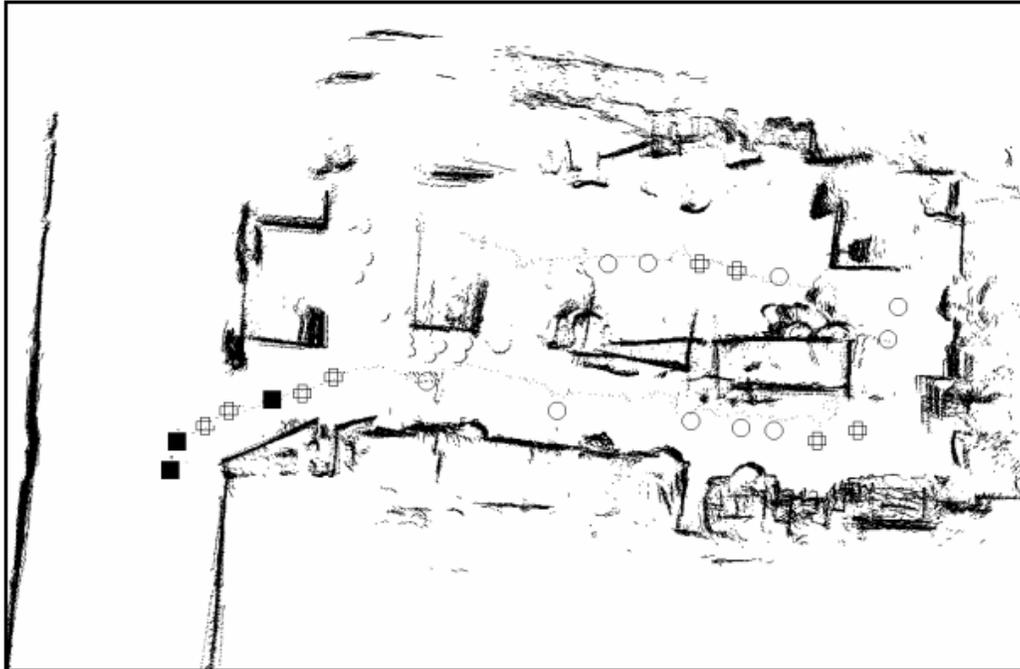


Figura 5.11 Representación métrico-semántica del recorrido del experimento 7. Se ilustra lo percibido por el barrido láser durante el recorrido, además se han agregado etiquetas semánticas. Un cuadrado representa a la clase pasillo, un círculo a la clase habitación y una cruz a la clase puerta.

Por otra parte también se aprecia que las observaciones del pasillo y la puerta han resultado parcialmente favorables, debido a que se puede notar que existe una etiqueta correspondiente a *pasillo* en el interior del laboratorio, donde la clase correcta es *puerta*. El total de los resultados de este experimento se resumen en la tabla 5.18.

<i>Error Pasillo</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
0%	20%	28.57%	23.80%

Tabla 5.18 Porcentaje de error de clasificación semántica del experimento 7.

5.4.2 Experimento 8

Para realizar este experimento se guió al robot para que explorara el ambiente cubriendo la ruta presentada en la figura y la clasificación obtenida durante el recorrido es presentada en la figura 5.12. Una vez más los círculos corresponden a la etiqueta de una habitación, los cuadrados rellenos a pasillo y las cruces a puertas abiertas.

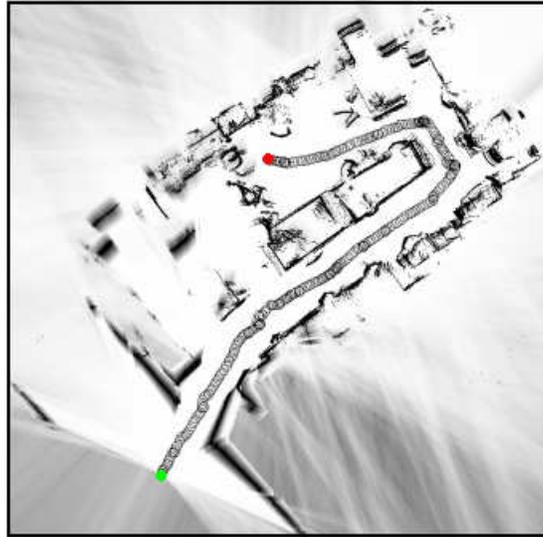


Figura 5.12. Trayectoria recorrida en el experimento 8. Se muestra la trayectoria, iniciando el pasillo (punto verde) y terminando en el laboratorio (punto rojo).

Para este recorrido se ha utilizado el clasificador generado durante el experimento número 6, la prueba se desarrollo iniciando en el pasillo y terminado en el laboratorio de biorobótica. Se puede observar en la parte inferior del mapa (ver figura 5.13), se empezó tomando observaciones del pasillo y la puerta, obteniendo resultados parecidos a los del experimento 7, debido a que se sigue cometiendo errores de confusión entre la clase *pasillo* en el interior del laboratorio, aunque esta vez se puede distinguir que durante el recorrido en el interior del laboratorio se han eliminado aquellos errores cometidos en el experimento 7, en donde el robot confundía la clase *habitación* con la clase *puerta* . El total de los resultados de este experimento se resumen en la tabla 5.19.

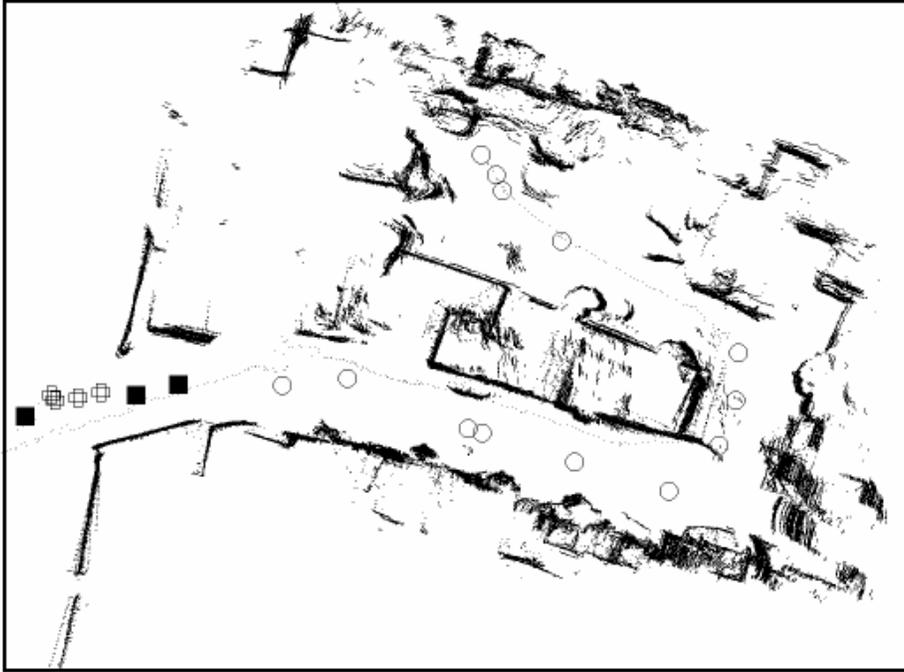


Figura 5.13 Representación métrico-semántica del recorrido del experimento 8. Se ilustra lo percibido por el barrido láser durante el recorrido, además se han agregado etiquetas semánticas. Un cuadrado representa a la clase pasillo, un círculo a la clase habitación y una cruz a la clase puerta.

<i>Error Pasillo</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
0%	33.33%	0%	10%

Tabla 5.19 Porcentaje de error de clasificación semántica del experimento 8.

5.5 Clasificación en ambientes desconocidos

En esta sección se presentan experimentos que tratan de mostrar la validez del clasificador semántico en ambientes “desconocidos”, esto quiere decir que al robot se le presentan observaciones de ambientes diferentes al ambiente de entrenamiento. Por lo tanto con los resultados de estos experimentos se tendrá una idea de que tan capaz es el clasificador de generalizar en cada una de las clases o categorías semánticas para las que fue entrenado.

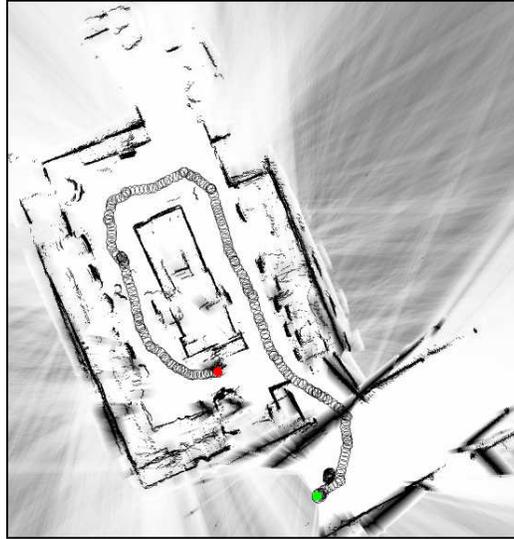


Figura 5.14. Trayectoria recorrida en el experimento 9. Se muestra la trayectoria para un ambiente desconocido, el punto verde indica el punto de partida y el rojo el punto donde termino el recorrido.

5.5.1 Experimento 9

Para realizar este experimento se guió al robot sobre el pasillo y posteriormente se metió por la puerta de un laboratorio no observado antes (véase figura 5.15). La ruta del robot se presenta en la figura 5.14 y la representación métrico-semántica que el robot “construyo” durante el recorrido es presentada en la figura. 5.16.



Figura 5.15 Ejemplo de un ambiente desconocido para el robot. Aquí se ilustra el nuevo ambiente utilizado para el experimento 9.

Para este recorrido se ha utilizado el clasificador del experimento 6, por ser el que ha arrojado mejores resultados.

En el mapa generado para este recorrido (figura 5.16) se puede notar que el clasificador semántico fue preciso en todas las ocasiones, lo cual es un resultado ideal para demostrar que se esta generalizando bien, sin embargo hay que señalar que el ambiente al que fue sometido el robot es muy parecido al ambiente de entrenamiento, pero sin tantos objetos a su alrededor, por lo que se considera que debido a ello el ambiente fue “fácil” de clasificar y lo que falta es observar como se comporta el clasificador en una “habitación” con características distintas a las presentadas hasta ahora.

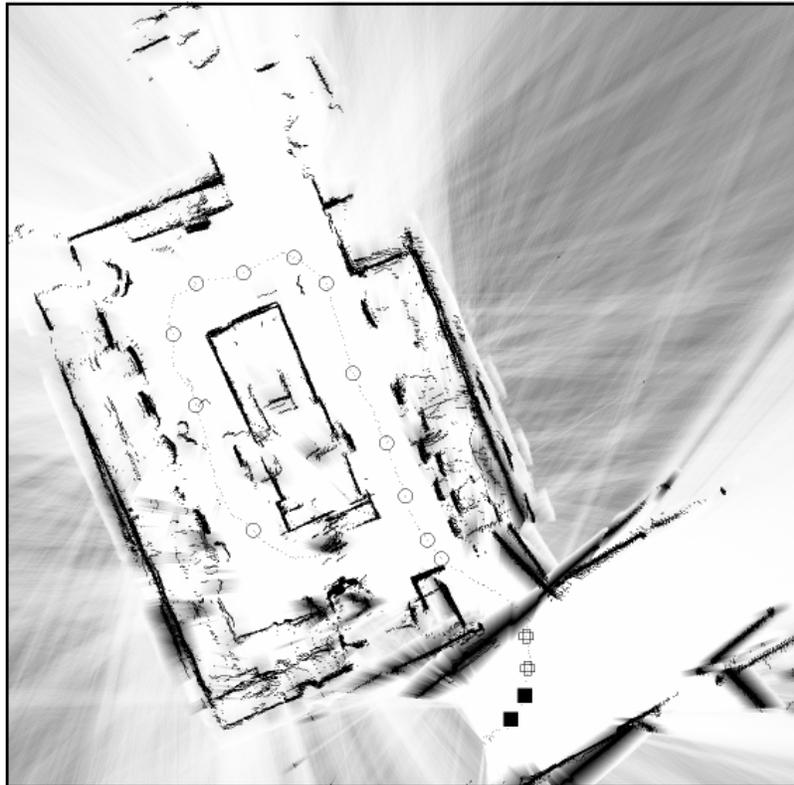


Figura 5.16 Representación métrico-semántica del recorrido del experimento 9. Se ilustra lo percibido por el barrido láser, además se observa que las etiquetas semánticas (cuadrado representa a la clase pasillo, un círculo a la clase habitación y una cruz a la clase puerta) que se han agregado para este experimento clasifican correctamente en todo el recorrido.

<i>Error Pasillo</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
0%	0%	0%	0%

Tabla 5.20 Porcentaje de error de clasificación semántica del experimento 9.

5.5.2 Experimento 10

Con base en los valores obtenidos en el experimento anterior, aquí se ha decidido aumentar el grado de dificultad para corroborar que el clasificador es lo suficientemente apto para generalizar en otro tipo de ambientes con configuraciones de espacio y mobiliario distintas a las presentadas hasta ahora. Para llevar a cabo este experimento se guió al robot en un nuevo ambiente empezando desde el interior (figura 5.17) y la representación que el robot “construyo” se muestra en la figura 5.18.

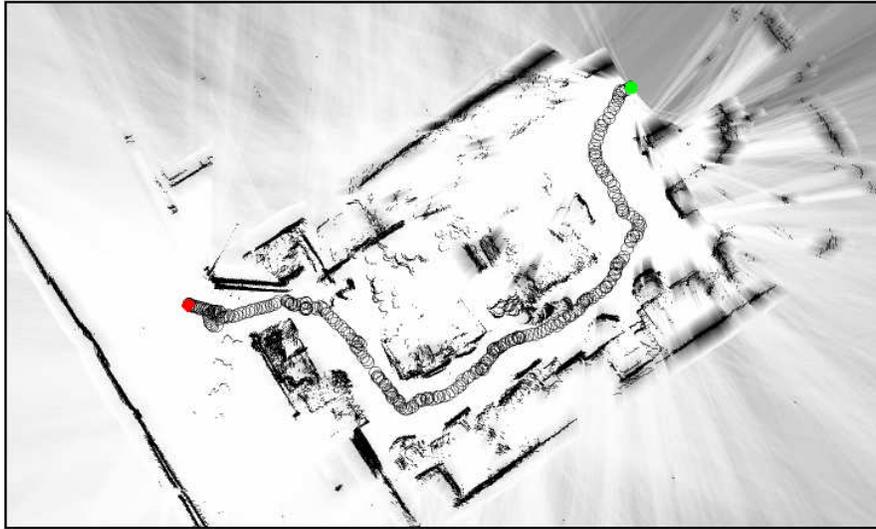


Figura 5.17. Trayectoria recorrida en el experimento 10. Se muestra la trayectoria para un ambiente desconocido, esta vez se inicio (punto verde) desde el interior de la habitación y se termino en el pasillo (punto rojo).

Una vez más se ha utilizado el clasificador del experimento 6 y como se puede observar en la representación generada (figura 5.18), se cometieron algunos errores durante el recorrido, clasificando en ciertos punto de la habitación como pasillo, sin embargo no se consideran significativos para no decir que se ha entrenado al robot para que clasifique y generalice correctamente en ambientes desconocidos. El resumen de resultados de este experimento se expone en la tabla 5.21.

<i>Error Pasillo</i>	<i>Error Puerta</i>	<i>Error Habitación</i>	<i>Error Global</i>
<i>0%</i>	<i>50%</i>	<i>11.76%</i>	<i>15.78%</i>

Tabla 5.21 Porcentaje de error de clasificación semántica del experimento 10.

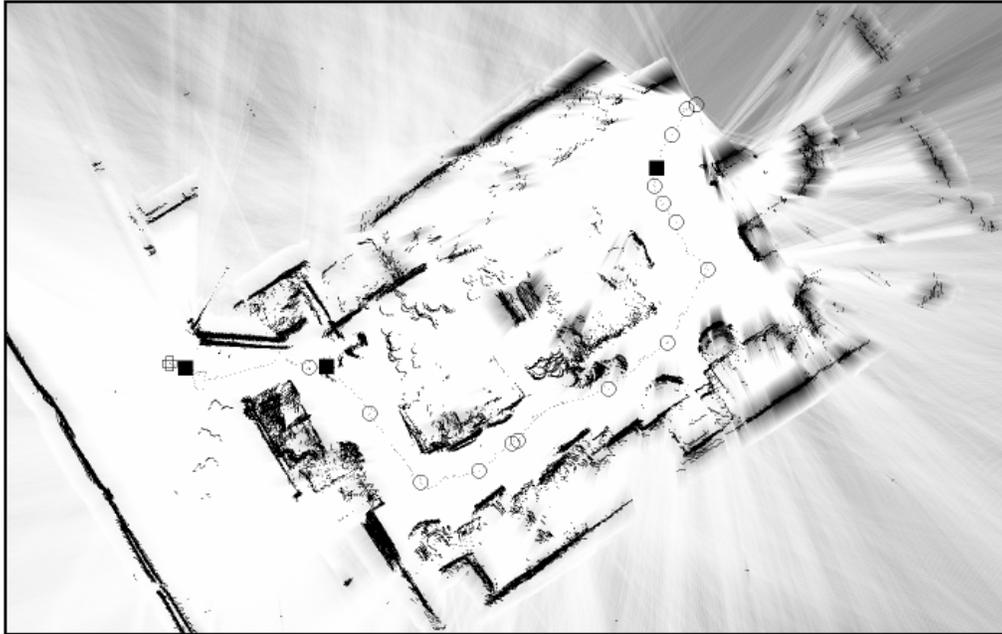


Figura 5.18 Representación métrico-semántica del recorrido del experimento 10. Un cuadrado representa a la clase pasillo, un círculo a la clase habitación y una cruz a la clase puerta. A diferencia del resultado anterior se ilustra que se cometieron errores en la clasificación semántica, pero estos no se consideran significativos para juzgar que se ha hecho una mala clasificación en este nuevo ambiente.

CAPÍTULO 6 – CONCLUSIONES Y TRABAJO FUTURO

6.1 Conclusiones

El trabajo presentado se ha enfocado en la conceptualización de la posición de un robot móvil en su ambiente, mediante la inclusión de conceptos o etiquetas de lugares, para complementar la información métrica de la representación del entorno. Para aprender dichos conceptos semánticos se ha utilizado el algoritmo AdaBoost, tomando características (de alto y bajo nivel) extraídas del ambiente con el sensor láser, para construir un clasificador de lugares que de acuerdo a la estructura del ambiente lo clasificó en: puerta, pasillo, habitación o espacio libre.

En este documento se hecho énfasis en indicar que la información semántica acerca de los lugares puede mejorar el rendimiento de un robot para llevar a cabo sus acciones. La idea principal es que los robots móviles pueden utilizar la información del ambiente, con el fin de interpretarlo y describirlo en una forma eficiente y consistente de forma que sea comprensible a los humanos, y de esta manera poder ampliar capacidades en tareas cotidianas que puedan requerir de una interacción humano-robot.

En particular, en base a la experimentación realizada con el robot, se observa que la clasificación del tipo de lugar puede ser vista como una localización de alto nivel, que puede complementar la localización del robot obtenida por otros sistemas (i.e. odometría y visión), lo cual podría ser útil para llevar a cabo otras tareas como la navegación, la exploración y la planeación de acciones.

Asimismo se han llevado a cabo experimentos en donde se puso a prueba el clasificador bajo observaciones de ambientes desconocidos o no incluidos en el entrenamiento y se ha comprobado que el clasificador de lugares es capaz de generalizar en cada una de las clases o categorías semánticas para las que fue entrenado en tales ambientes.

6.2 Acerca de la contribución de la investigación

A partir del aprendizaje de características del ambiente adquiridas con el sensor láser 2D, en base al etiquetado del usuario y en combinación con una técnica de mapeo tradicional se obtiene una representación métrica-semántica del espacio que tiene el potencial para conceptualizar el espacio, y por lo tanto se considera como un aspecto importante de la cognición espacial en robots móviles, que en un futuro permita aspirar al desarrollo de “robots cognitivos inteligentes” capaces de interactuar con los humanos de forma más natural.

En comparación con otros trabajos de mapeo semántico utilizando láser 2D, se ha mejorado un enfoque basado en características simples (de bajo nivel) tales como aquellas basadas solo en los valores de distancias regresados por el sensor láser, con otro tipo de características (aquí consideradas de alto nivel) tales como el número de clusters, líneas rectas (pueden ser paredes), esquinas y discontinuidades detectadas en el barrido láser. Lo anterior se ha considerado más adecuado para el mapeo semántico, porque dichas características o marcas naturales utilizadas pueden representar objetos físicos reales (o partes de objetos).

6.3 Acerca de las limitaciones

Este trabajo contribuye a la investigación de mapeo semántico en robótica, es decir, se aumenta una representación del ambiente de navegación del robot con información semántica. Esta representación es compatible con el punto de vista humano y es capaz de formar una conceptualización del espacio, pero es insuficiente para estimar la similaridad con la representación mental del espacio que tienen los humanos. Para codificar información semántica se requiere el trabajo de expertos en psicología, de tal forma que el robot pueda estar equipado con una representación más completa, suficientemente eficiente para entender, interpretar y representar el ambiente. Una vez que se tenga la información semántica que se necesita codificar, el trabajo en esta área se debe orientar a la constante mejora de los sistemas de percepción existentes con el fin de mejorar la representación del medio ambiente

6.4 Acerca de líneas futuras de investigación

Una motivación fundamental en este trabajo es poder desarrollar en un futuro “robots cognitivos autónomos e inteligentes” capaces de construir mapas semánticos de multijerarquía, es decir mapas aumentados con objetos clave a la vista y sus relaciones, lo que permitiría al robot poder realizar tareas cada vez más complejas, por ejemplo, determinar la categoría de lugar de manera más específica (cocina, baño, oficina) a partir del número y tipo de objetos presentes en el entorno.

Por lo anterior se han considerado abordar varios temas como trabajo a corto y largo plazo con el fin de mejorar aspectos que el sistema presentado aquí no cumple debido a algunas de sus limitaciones.

- Lo primero que se podría hacer es mejorar la clasificación de lugares utilizando técnicas basadas en modelos ocultos de Markov, para que posteriormente en base a esta clasificación más robusta se puedan establecer regiones y rutas para obtener de esta manera información topológica, y de esta forma complementar aun más la representación del entorno adquirida por el robot.
- Debido a que el uso de láser 2D no es suficiente para obtener la totalidad de la geometría del entorno del robot, una opción natural para la mejora del sistema de percepción presentado aquí, es el uso de un sensor láser 3D para clasificar el ambiente en base a objetos detectados en la nube 3D tales como paredes, piso, techo y puertas, además de muebles como sillas, mesas, escritorios.
- Por otra parte otro reto que surge es codificar la información semántica directamente con información visual. Debido a los crecientes avances en el área de visión en la actualidad, ha surgido como una alternativa viable que proporciona una entrada sensorial más completa, permitiendo una mejor distinción de los objetos en el ambiente. Sin embargo, la información visual tiende a ser dinámica, ruidosa (por iluminación) y difícil de interpretar, por lo que se considera que lo ideal sería emplear una combinación de clasificadores de lugares provenientes de sensores láser y de la visión.
- Otro aspecto importante de la clasificación de lugares que se debe tomar en cuenta es la dinámica del ambiente, debido a que si constantemente se muevan los muebles y objetos, además de la constante actividad de las personas puede afectar al sistema de percepción para la clasificación de los lugares en el tiempo.

- Sin embargo se pueden crear mapas semánticos que den significado a otras propiedades de los objetos y del medio ambiente, como asignarles estados a los objetos (i.e. puerta abierta, cerrada etc.) o representar la actividad y la navegabilidad como se ha propuesto en otros trabajos.
- En lo que se refiere a trabajo a corto plazo para mejorar la navegación del robot, se propone integrar lo presentado aquí con un mapa de activación de árbitro de comportamientos [48], pero haciendo uso de las semánticas de ambiente adquiridas en lugar de regiones definidas por el usuario, y ejecutar un arbitro deseado, a fin de que el robot vaya de un lugar a otro del ambiente (véase figura 6.1).

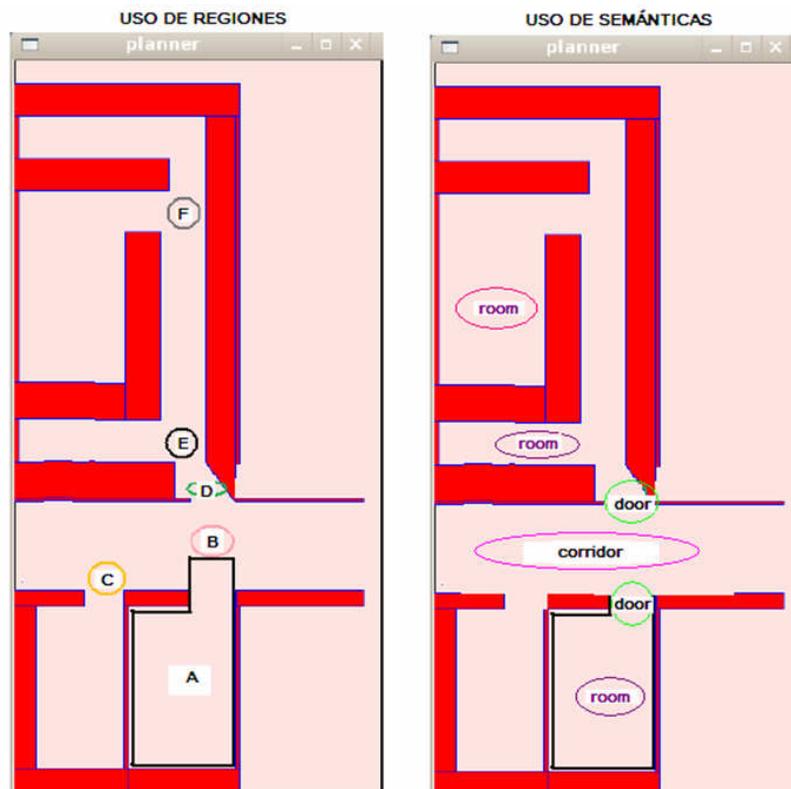


Figura 6.1 Comparación de Mapa de Activación de árbitro de comportamientos. Se muestra como se podría aplicar este mapa usando solo semánticas del ambiente en lugar de regiones.

- Finalmente se cree que se puede extender el método de clasificación de lugar por medio de aprendizaje no supervisado, para obtener sistemas capaces de aprender continuamente categorías de lugares utilizando múltiples sensores.

REFERENCIAS BIBLIOGRÁFICAS

- [1] E.C. Tolman, Cognitive maps in rats and men, *Psychological Review* 55 (1948) 189–208
- [2] B. Kuipers, The spatial semantic hierarchy, *Artificial Intelligence* 119 (2000) 191–233.
- [3] D. Kortenkamp, Cognitive maps for mobile robots: A representation for mapping and navigation, Ph.D. Thesis, University of Michigan, 1993.
- [4] D. Kortenkamp, "Applying Computational Theories of Cognitive Mapping to Mobile Robots," *Working Notes: AAAI Fall Symposium on Applications of Artificial Intelligence to Real-World Mobile Robots*, 1992.
- [5] B.J. Kuipers, The cognitive map: Could it have been any other way? in: *Spatial Orientation: Theory, Research, and Application*, Plenum Press, New York, 1983, pp. 345–359.
- [6] Wong, C.K., Schmidt, J., Yeap, W.K.: Using a Mobile Robot for Cognitive Mapping. In: *International Joint Conference on Artificial Intelligence (IJCAI)*, Hyderabad, India (2007).
- [7] S. Vasudevan, S. Gachter, V.T. Nguyen and R. Siegwart. Cognitive Maps for Mobile Robots - An object based approach. *Robotics and Autonomous Systems*, Volume 55, Issue 5, From Sensors to Human Spatial Concepts, 31 May 2007, Pages 359-371.
- [8] S. Vasudevan, V.T. Nguyen and R. Siegwart (2006) Towards a Cognitive Probabilistic Representation of Space for Mobile Robots. In the Proceedings of the IEEE International Conference on Information Acquisition (ICIA) 2006, Shandong, China.
- [9] Aníbal Ollero Baturone. "Robotica. Manipuladores y Robots Móviles". Ed. Alfaomega México D.F (2007)
- [10] Carlos Fernández Caramés. "Técnicas de Navegación de Robots Basadas en Sistemas de Medición por Láser". Departamento de Informática y Automática Universidad de Salamanca (Septiembre, 2007).
- [11] Sergio Francisco Hernández Alamilla "Navegación de un Robot Móvil en Ambientes Interiores Usando Marcas Naturales del Ambiente". ITESM Campus Cuernavaca (Mayo 2005).
- [12] Range Finder Type Laser Scanner URG 04LX Specifications. Hokuyo Automatic Co. Ltd. 2005.
- [13] V.T. Nguyen, A. Martinelli, N. Tomatis, R. Siegwart, A comparison of line extraction algorithms using 2D laser rangefinder for indoor mobile robotics, in: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS, Edmonton, Canada, 2005.
- [14] Steven W. Smith. *The Scientist and Engineer's Guide to Digital Signal Processing* California Technical Publishing San Diego, California. 2 Ed. (1999).
- [15] *Hough, P.V.C. 1962. Method, and means for recognizing complex patterns. U. S. Patent 3069654.*
- [16] Gonzálo Pajares Martínsanz, Jesús M. de la Cruz "Visión por Computador: Imágenes digitales y aplicaciones". México D.F. Ed. Ra-ma (2008).

- [17] Yoav Freund Robert E. Schapire. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of computer and system sciences* 55, 119_139 (1997).
- [18] Yoav Freund Robert E. Schapire. A Short Introduction to Boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(5):771-780, September, 1999.
- [19] Paul Viola and Michael J. Jones. Robust real-time object detection. In *Proceedings of IEEE Workshop on Statistical and Theories of Computer Vision*, 2001.
- [20] Robert E. Schapire, Yoram Singer. BoosTexter. A Boosting-based System for Text Categorization. *Machine Learning*, 39(2/3):135-168, 2000.
- [21] Martínez Mozos, C. Stachniss A. Rottmann, W Burgard. Using AdaBoost for Place Labeling and Topological Map Building. University of Freiburg, Department of Computer Science, Freiburg, Germany.
- [22] Martínez Mozos, C. Stachniss W Burgard. Supervised Learning of Places from Range Data using AdaBoost. *Proceedings of the 2005 IEEE International Conference on Robotics and Automation* Barcelona, Spain, April 2005.
- [23] Søren Riisgaard and Morten Rufus Blas. SLAM for Dummies. A Tutorial Approach to Simultaneous Localization and Mapping.
- [24] Bruno Steux, Oussama El Hamzaoui. CoreSLAM : a SLAM Algorithm in less than 200 lines of C code Mines ParisTech - Center of Robotics, Paris, Francia.
- [25] Cebrecos del Castillo, Lozano Prieto, Nieto Domínguez. Robots inteligentes autónomos de nueva generación. Universidad Carlos III. Madrid España, 2008.
- [26] S. Thrun. Robotic Mapping: A Survey. In G. Lakemeyer and B. Nebel, eds., “Exploring Artificial Intelligence in the New Millenium”, Morgan Kaufmann, 2002.
- [27] R. Chatila, J. Laumond, Position referencing and consistent world modeling for mobile robots, in: *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 1985.
- [28] B. Kuipers, Modeling spatial knowledge. *Cognitive Science* 2, 1978.
- [29] ICRA-07 Workshop on Semantic Information in Robotics IEEE Int Conf on Robotics and Automation. Rome, Italy, April 2007.
- [30] C. Galindo and J. González and J.A. Fernández-Madrigal. A Control Architecture for Human-Robot Integration: Application to a Robotic Wheelchair. *IEEE Transactions on Systems, Man, and Cybernetics Part B*, vol. 36, 2006.
- [31] T.P. McNamara. Mental Representations of Spatial Relations. *Cognitive Psychology*, 18:87–121, 1986.
- [32] S. Thrun. Learning metric-topological maps for indoor mobile robot navigation. *Artificial Intelligence*, 99(1):21–71, 1998.
- [33] A. Rottmann, M. O. Mozos, C. Stachniss, and W. Burgard. Semantic place classification of indoor environments with mobile robots using boosting. In *Proceedings of the National Conference on Artificial Intelligence*, pages 1306–1311, 2005.
- [34] N. Goerke and S. Braun. Building semantic annotated maps by mobile robots. In *Proceedings of the Conference Towards Autonomous Robotic Systems*, Londonderry, UK, 2009.

- [35] Ulrich Weiss, Peter Biber. Semantic Place Classification and Mapping for Autonomous Agricultural Robots. IROS 2010 Workshop: Semantic Mapping and Autonomous Knowledge Acquisition.
- [36] B. Limketkai, L. Liao, and D. Fox. Relational object maps for mobile robots. in Proc. of the Int. Conf. on Artificial Intelligence (IJCAI), Edinburgh, Scotland, 2005, pp. 1471-1476.
- [37] P. Althaus and H. Christensen, "Behaviour coordination in structured environments," *Advanced Robotics*, vol. 17, pp. 657–674, Nov. 2003.
- [38] Alexander J. B. Trevor, John G., Carlos Nieto, I. Christensen. Tables, Counters, and Shelves: Semantic Mapping of Surfaces in 3D. IROS 2010 Workshop: Semantic Mapping and Autonomous Knowledge Acquisition.
- [39] David Meger, Per-Erik Forssen, Kevin Lai, Scott Helmer, Sancho McCann, Tristram Southey, Matthew Baumann, James J. Little, David G. Lowe, Bruce Dow. Curious George: An Attentive Semantic Robot. *Robotics and Autonomous Systems* Volume 56, Issue 6, 30 June 2008, Pages 503-511 From Sensors to Human Spatial Concepts.
- [40] R.B. Rusu, Z.C. Marton, N. Blodow, A. Holzbach, and M. Beetz. Model-based and learned semantic object labeling in 3D point cloud maps of kitchen environments. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, St. Louis, MO, USA, 2009.
- [41] C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J. Fernandez- Madrigal, and J. Gonzalez, "Multi-hierarchical semantic maps for mobile robotics," in *International Conference on Intelligent Robots and Systems (IROS)*, Aug. 2005, pp. 2278–2283.
- [42] A. Pronobis, O. Martínez Mozos, B. Caputo, P. Jensfelt. Multi-modal Semantic Place Classification. *The International Journal of Robotics Research* Vol. 29, No. 2–3, February/March 2010, pp. 298–320.
- [43] A. Nuchter and J. Hertzberg, "Towards semantic maps for mobile robots," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 915–926, 2008.
- [44] Wolf, Sukhatme, *gs Semantic Mapping Using Mobile Robots*. In *Proceedings of IEEE Transactions on Robotics*. 2008.
- [45] Bipin Suresh, Carl Case, Adam Coates, Andrew Y. Ng. Autonomous Sign Reading for Semantic Mapping. Department of Computer Science, Stanford University.
- [46] Trowbridge, C. On fundamental methods of orienting and "imaginary Maps" *Science*, 38, 888-897, 1913.
- [47] Lynch, K. *The image of City* (1960). Cambridge, Mass: MIT. Press. versión en castellano: *la imagen de la ciudad*. Buenos Aires, ed. Infinito, 1976.
- [48] Sánchez Carmona Iván, Savage Carmona Jesús, Kuri Morales Ángel. Evolución de la organización de los comportamientos de un robot usando algoritmos genéticos. Universidad Nacional Autónoma de México. México D.F., 2010.
- [49] Garre, M., Cuadrado-Gallego J.J., Sicilia, M.A., Rodríguez, D., Rejas, R. Comparación de diferentes algoritmos de clustering en la estimación de coste en el desarrollo de software. *Revista Española de Innovación, Calidad e Ingeniería del Software*, Abril 2007

[50] Mora Torres Martha, Laureano Cruces. Sistema experto en la toma de decisiones de un escenario de riesgo: loca pequeño en una planta nucleoelectrica. Universidad Nacional Autónoma de México. México D.F., 2007. <http://ce.azc.uam.mx/profesores/clc/>. Recuperado el 29 de junio de 2011.

[51] Laureano Cruces, A. 2000. *Interacción dinámica en sistemas de enseñanza inteligentes*. Tesis de Doctorado en Investigación Biomédica Básica, Instituto de Investigaciones Biomédicas - UNAM. <http://ce.azc.uam.mx/profesores/clc/>. Recuperado el 29 de junio de 2011.

[52] Marcellin Jacques Sergio, Notas del Curso: “*Construcción de Sistemas Expertos*”, Posgrado en Ciencia e Ingeniería de la Computación, Universidad Nacional Autónoma de México, 2010.

ANEXO I

TABLA DE CARACTERÍSTICAS EXTRAIDAS CON EL SENSOR LÁSER DE PROXIMIDAD

En esta tabla se describe brevemente el cálculo de las características extraídas con el sensor láser de proximidad usadas para construir el clasificador semántico de lugares. Se han utilizando algunas de las características definidas por Mozos [22], es decir aquellas características que aquí se han considerado como de bajo nivel (conjuntos B y P). Las del conjunto B (b1-b10) corresponden a las calculadas directamente por medio de los valores de distancia obtenidos por el láser, mientras que las del conjunto P (p1-p6) corresponden a características obtenidas de una aproximación poligonal del área y perímetro cubierto por la observación del láser.

Finalmente las características del conjunto M (m1-m9) corresponden a las características de alto nivel aquí propuestas para mejorar la clasificación semántica, las cuales están basadas primordialmente en una etapa de extracción de marcas naturales en el ambiente.

ID	DESCRIPCIÓN	CÁLCULO
b1	Promedio de longitud de las lecturas.	$avg = \frac{1}{M} \sum_{i=0}^{M-1} d_i$ <p><i>Donde: d_i corresponde a la distancia devuelta por el láser en la lectura i.</i></p>
b2	Promedio de la diferencia entre la longitud de dos lecturas consecutivas	$avgdif = \frac{1}{M-1} \sum_{i=0}^{M-2} d_i - d_{(i+1)} $
b3	Desviación estándar de la longitud de las lecturas.	$std = \sqrt{\frac{1}{M} \sum_{i=0}^{M-1} (d_i - avg)^2}$
b4	Desviación estándar de la diferencia entre la longitud de lecturas	$stddif = \sqrt{\frac{1}{M-1} \sum_{i=0}^{M-2} (d_i - d_{(i+1)} - avgdif)^2}$

b5	Promedio de la relación entre dos lecturas consecutivas.	$avgrel = \frac{1}{M-1} \sum_{i=0}^{M-2} \left \frac{d_i}{d_{(i-1)}} \right $
b6	Desviación estándar de la relación entre dos lecturas consecutivas.	$stdrel = \sqrt{\frac{1}{M-1} \sum_{i=0}^{M-2} \left(\left \frac{d_i}{d_{(i-1)}} \right - avgrel \right)^2}$
b7	Promedio normalizado de las lecturas.	$avgnorm = \frac{1}{M} \sum_{i=0}^{M-1} \frac{d_i}{d_{\max}}$
b8	Desviación estándar normalizada	$stdnorm = \sqrt{\frac{1}{M} \sum_{i=0}^{M-1} \left(\frac{d_i}{d_{\max}} - avgnorm \right)^2}$
b9	La distancia mínima entre dos mínimos locales.	$dist \min = \sqrt{(x_{\min L} - x_{\min R})^2 - (y_{\min L} - y_{\min R})^2}$ <p>Donde:</p> <p>$(x_{\min L}, y_{\min L})$ punto en el mínimo local del lado izquierdo</p> <p>$(x_{\min R}, y_{\min R})$ punto en el mínimo local del lado derecho</p>
b10	Curtosis del barrido láser	$kurtosis = \frac{\sum_{i=0}^{M-1} (d_i - std)^4}{M \cdot avg^4} - 3$
p1	Área de la aproximación poligonal	$areap = \frac{1}{2} \sum_{i=0}^{M-1} (x_i y_{(i+1)} - x_{(i+1)} y_i)$ <p>Donde:</p> <ul style="list-style-type: none"> • $x_i = d_i \cos(\theta_i)$, $y_i = d_i \sin(\theta_i)$ • d_i Distancia devuelta por el láser en la lectura i. • θ_i Ángulo correspondiente a la lectura i del laser
p2	Perímetro de la aproximación poligonal	$perimetrop = \sum_{i=0}^{M-1} \sqrt{(x_i - x_{(i+1)})^2 - (y_i - y_{(i+1)})^2}$

p3	Factor de forma de la aproximación poligonal.	$factorf = \frac{4\pi \cdot areap}{\sqrt{perimetrop}}$
p4	Circularidad de la aproximación poligonal.	$circularidad = \frac{(perimetrop)^2}{areap}$
p5	Circularidad normalizada de la aproximación poligonal.	$circularidadn = \frac{4\pi \cdot areap}{(perimetrop)^2}$
p6	Relación área perímetro	$areaper = \frac{areap}{perimetrop}$
m1	Número de líneas rectas encontradas en el barrido láser.	<i>Se calculan con el algoritmo de la Transformada de Hough sobre el barrido láser</i>
m2	Número de configuraciones de dos rectas paralelas.	<i>Se comparan los ángulos de las rectas devueltas por el proceso de la transformada de Hough, buscando pendientes iguales.</i>
m3	Número de configuraciones de dos rectas perpendiculares.	<i>Se comparan los ángulos de las rectas devueltas por el proceso de la transformada de Hough, buscando configuraciones que formen un ángulo recto.</i>
m4	El promedio de la distancia entre líneas paralelas.	<i>Se toma el número de configuraciones de rectas paralelas, se estima la distancia entre rectas y se saca un promedio de dicho valor</i>
m5	Número de Clusters encontrados en el barrido láser.	<i>Número de clusters encontrados después de extraer discontinuidades y esquinas en el proceso de extracción de marcas</i>
m6	Número de discontinuidades encontradas en el barrido láser.	<i>Número de discontinuidades calculadas al aplicar un filtro paso alto al barrido láser por medio de convolución</i>
m7	Promedio de la diferencia entre discontinuidades.	<i>Se promedia la distancia entre puntos correspondientes a discontinuidades consecutivas.</i>

m8	Número de esquinas encontradas.	<i>Se obtienen mediante un proceso recursivo en la extracción de marcas.</i>
m9	Promedio de la distancia entre esquinas.	<i>Se promedia la distancia entre puntos correspondientes a discontinuidades consecutivas.</i>