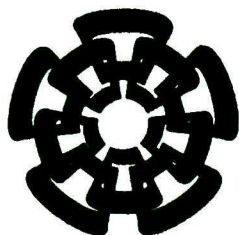


CT-711-SSI
Don, 2013

xx(202793.1)



Centro de Investigación y de Estudios Avanzados
del Instituto Politécnico Nacional
Unidad Guadalajara

Técnicas Geométricas y Aprendizaje Automático para la Reconstrucción 3D y Navegación Robótica

**CINVESTAV
IPN
ADQUISICION
LIBROS**

Tesis que presenta:

Miguel Bernal Marín

para obtener el grado de:

Doctor en Ciencias

en la especialidad de:

Ingeniería Eléctrica

Director de Tesis

Dr. Eduardo José Bayro Corrochano

CT00615
CT-711-SS1
14-03-2013
Dom. 2013
\$

ID: 202581-1001

Técnicas Geométricas y Aprendizaje Automático para la Reconstrucción 3D y Navegación Robótica

**Tesis de Doctorado en Ciencias
Ingeniería Eléctrica**

Por:

Miguel Bernal Marín

Maestría en Ciencias en Ingeniería Eléctrica

Centro de Investigación y de Estudios Avanzados del I.P.N.

Unidad Guadalajara 2002-2004

Becario de CONACyT, expediente no. 42139

Director de Tesis

Dr. Eduardo José Bayro Corrochano

CINVESTAV del IPN Unidad Guadalajara, Febrero de 2012.

Resumen

En la actualidad la robótica móvil es un área de estudio en la cual se conjuntan diferentes técnicas para realizar la exploración de interiores y exteriores. El uso de diferentes algoritmos y aplicaciones facilita el trabajo de búsqueda y reconocimiento de objetos encontrados en el medio ambiente. El presente trabajo estudia las técnicas actuales y propone nuevas soluciones para la exploración, reconocimiento e interpretación de escenarios y objetos.

Los objetos y escenarios encontrados en la vida cotidiana tienen algo en común, son objetos que se pueden representar con primitivas geométricas sencillas. Debido a esto, se propone el uso del álgebra geométrica conformal, la cual, además de representar objetos geométricos de una manera compacta, tiene las herramientas necesarias para implementar la representación de la traslación y rotación de un robot móvil, así como de los objetos que se encuentran a la hora de realizar una exploración.

Además, un objeto tiene ciertas características que lo diferencian de otro objeto de diferente forma. El uso de un método de aprendizaje automático ayuda en el reconocimiento de cierta clase de objetos. Con esto se pueden ubicar ciertos objetos que se encuentren en el ambiente intramuros, como señales de emergencia, extintores, impresoras, etc. los cuales ayudan a la ubicación de un robot móvil dentro de un mapa generado por el mismo. A estos objetos reconocidos dentro

de un mapa se llaman ¿landmarks¿ que su traducción literal es “mojon”, y que aquí les llamamos “marcas espaciales” o simplemente “marcas”. Estas “marcas” son de ayuda a la hora de reubicar el robot móvil en un mapa previamente generado.

Los escenarios, los cuales generalmente pueden ser representados por planos, también tienen características que los distinguen de otros escenarios. Se pueden representar los planos con puntos dentro del espacio de Hough. Estos puntos nos dan la característica que distingue a un lugar de otro y con ello que se puede identificar si estamos en cierto lugar. Cambiar del espacio cartesiano al espacio de Hough para la representación de planos ayuda en la reducción de grados de libertad a la hora de estar buscando entre los diferentes lugares en los cuales ha estado el robot móvil.

En la exploración y navegación de robots móviles se usan diferentes sensores de entrada. En el presente trabajo se utilizan diferentes dispositivos los cuales ayudan a la percepción del medio ambiente. Los dispositivos que se usaron para la realización de los experimentos fueron: un sistema de cámaras estéreo, un sensor de alcance laser y un sensor de profundidad RGB-D. Estos sensores fueron utilizados para la captura de estímulos visuales los cuales después de su pre-procesamiento y empleando las técnicas antes mencionadas permitieron el reconocimiento y exploración del robot móvil.

Abstract

Nowadays mobile robotics is an area of study in which different techniques come together to perform indoor and outdoor exploration. The use of different algorithms for perception and machine learning facilitates greatly the robot task of search and recognition of objects found in the environment. This thesis examines current techniques and proposes new solutions for scanning, recognition and interpretation of scenes and objects.

The objects and scenarios encountered in everyday life have something in common, that they can be represented by simple geometric primitives. Taking into consideration this fact, we propose the use of conformal geometric algebra to represent geometric objects in a compact way. This mathematical framework offers also the necessary tools to implement effectively the representation of translation and rotation of a mobile robot and also the objects found while performing an exploration.

In addition, an object has certain characteristics that differentiate it from other objects in various ways. Using an extended Hough transform for plane and a machine learning method the robot algorithms manages to recognize certain class of objects. With this, objects can be localized in the indoor environment as distress signals, fire extinguishers, printers, etc. In addition, it helps for the localization of a mobile robot in a map generated by itself. A recognized object within a map

is called "landmark". These "landmarks" are very helpful to relocate the mobile robot in a previously generated map.

The scenarios, which usually can be represented by planes, also have features that distinguish them from other scenarios. Planes can be represented by points in Hough space. These points give us enough features that help to distinguish objects or structured places which in turn can be identified even if the robot looks them in a different place. Changing from Cartesian space to Hough space for the representation of planes helps indeed to reduce the degrees of freedom of the perception task at the different places visited by the robot.

In this work we used very effective devices which help in the perception of the environment. The devices used to perform the experiments were: a stereo camera system, a laser range sensor and a depth sensor RGB-D. These were utilized to collect visual cues and together with image processing and machine learning techniques objects and places were recognized in order necessary to facilitate the relocalization and navigation of the robot.

Agradecimientos

Doy gracias a Dios porque me permitió vivir esta hermosa experiencia al realizar esta tesis.

Doy gracias al Prof. Dr. Eduardo Bayro Corrochano por dirigirme en este trabajo de tesis.

Doy gracias al Consejo Nacional de Ciencia y Tecnología (CONACyT) por otorgarme una beca de manutención durante el desarrollo de esta tesis.

Doy gracias a mi familia por haberme apoyado durante estos años; mi madre Ildelisa Marin, mi padre Miguel Bernal, mi hermana Angelica Bernal y su esposo Felipe Muños, y en especial a mi novia/esposa Yolanda Alejo, muchas gracias por aguantarme todo este tiempo con lo de la tesis.

Doy gracias a mis amistades que hice en el CINVESTAV, al descubridor de las patitas Adrian (we we we), las frases celebres de Alejandro (el chino), a Luis Enrique (el Negro) porque si no fuera por el no supiéramos como llamar a las personas sin saber su nombre, a Gaspar (Gasparin) por su valiosa ayuda en el laboratorio, a Heriberto y Alberto (los petrillis) por su sabiduría chilanga, a Oscar (Eleno) por el futbol jajaja, a Adan (landa) por saber que los ciegos necesitan casco, a Carlos (Sr. patitas) por sus sabios consejos de vida, a Alejandro (Pizano) por su sabiduria en Mac OS, a Eduardo (Filemon) por su sabiduría de vida y consejos políticos, a Ulises por su rebeldía, a Efren (el Gerber) por su enseñanza

de empresas dedicadas a la robótica, a Laura (chuchis) por traer alegría al grupo con fiestas, reuniones y demás, a Jose Luis (choche) por su paciencia y su Fe demostrada discretamente (así como de cómo dormir en en cubo), a Luis Eduardo Falcon, Carlos Lopez y Nancy Arana por sus correcciones de tesis, a todos de la generación 2007-2009 de maestría Luis Luque, Lizette, Marcos, Rimann, Lalo, Ross y todos los que me faltan por su ayuda en esos años, a Juan (Ley) por sus correcciones ortográficas, a Julio Zamora y Leo Reyes por acordase de mi cuando necesitaban a alguien en Intel, Jorge Rivera por su apoyo, y a todos aquellos que conocí durante mi estancia en el CINVESTAV.

Índice general

Resumen	I
Abstract	III
Agradecimientos	V
1. Introducción	1
1.1. Motivación .	3
1.2. Objetivo	5
1.3. Organización	6
2. Geometría Algebraica	9
2.1. Álgebra Geométrica N-Dimensional	10
2.1.1. Propiedades de los Multivectores	12
2.2. Geometría Conformal	13
2.2.1. Espacio de Minkowski	13
2.2.2. Proyección Estereográfica	14

2.2.3. Definición del Álgebra Geométrica Conformal	17
2.2.4. Transformaciones Conformes	24
2.2.5. Movimientos Rígidos: Rotores, Trasladadores y Motores	29
3. Reconocimiento de Objetos y Navegación	35
3.1. Introducción	35
3.2. Reconocimiento de Objetos	36
3.2.1. Características Tipo Haar	37
3.2.2. Clasificadores en Cascada	38
3.3. Estimación de la Ubicación de Objetos Dentro del Ambiente Virtual	41
3.4. Representación de Marcas Espaciales	43
3.5. Navegación	46
4. Representación y Reconocimiento de Objetos Usando la Transformada de Hough	49
4.1. La Transformada de Hough	49
4.1.1. Transformada de Hough (2D) Clásica	49
4.1.2. La Transformada de Hough en 3D	52
4.1.3. La Transformada de Hough y el Álgebra Geométrica Conformal	53
4.2. Representación de Objetos 3D en el Dominio de Hough	57
4.2.1. Percepción del Ambiente	59
4.2.2. Variación de Solo un Ángulo en los Objetos	60

4.3. Implementación y Resultados Experimentales	64
4.3.1. Percepción de Formas en 3D	65
5. Reconocimiento de Objetos y Creación de Mapas Usando Imágenes RGB y su Profundidad	71
5.1. Cámaras RGB-D	72
5.2. Reconocimiento de Objetos	75
5.2.1. Transformada de Características de Escala Invariante (SIFT)	76
5.2.2. Representación de Marcas Espaciales	78
5.3. Obtención de Planos en RGB-D	79
5.3.1. Planos en la Imagen de Profundidad	80
5.3.2. Experimentos	85
6. Conclusiones	87
6.1. Trabajo Futuro	88
A. Fusión de Dispositivos para la Creación de un Mapa en 3D	91
A.1. Modelado en 3D y Unificación de Coordenadas .	91
A.2. Uso de <i>Motores Conformales</i> en 3D	92
A.3. Modelado de Dispositivos de Entrada	93
A.3.1. Desplazamiento del Robot Móvil (Odómetro)	93
A.3.2. Unidad de Sensor Láser	94
A.3.3. Unidad Pan-Tilt	95
A.3.4. Sistema de Cámaras Estéreo	96

A.4. Realización del Mapa en 3D	97
B. Publicaciones	99
B.1. Revista .	99
B.2. Capítulo de libro	99
B.3. Conferencia	100
Bibliografía	101

Nomenclatura

Operador del producto interior

e_+ Vector cuyo cuadrado es +1

e_- Vector cuyo cuadrado es -1

e_0 Representación del punto en el origen en Conformal

e_i Elemento de la base del álgebra geométrica

e Representación del punto en el infinito en Conformal

\mathbb{N} Conjunto de los números naturales

$\mathbb{P}(e, e_0)$ Hiperplano que pasa por e_0 y es normal a e

\mathbb{R} Conjunto de los números reales

$\mathcal{G}_{3,0,0}$ Representación del espacio Euclidiano 3D

$\mathcal{G}_{4,1,0}$ Representación del espacio Conformal

$\mathcal{G}_{p,q,r}$ Representación del espacio de álgebra geométrica

\underline{x} Punto en el espacio Conformal

x Punto en el espacio Euclidiano 3D

- E Pseudo-escalar unitario que genera el plano de Minkowski
- I Pseudo-escalar unitario del álgebra geométrica
- \wedge Operador del producto exterior
- x Número escalar

Índice de figuras

1.1. Robot móvil al cual se le pueden agregar diferentes sensores de entrada para su exploración.	2
2.1. Vectores base y líneas nulas del plano de Minkowski.	13
2.2. Proyección estereográfica en 1D	15
2.3. Visualización del modelo homogéneo para la proyección estereográfica en el caso 1D.	18
2.4. La traslación como una reflexión sobre dos planos paralelos	28
2.5. La rotación como una reflexión sobre dos planos que se intersecan	28
3.1. Esquema para el reconocimiento de objetos y su representación en un ambiente virtual 3D.	36
3.2. Cálculo de la integral de una imagen y su área de interés.	38
3.3. Representación de características tipo Haar en cuadros de 20x20 pixeles.	39
3.4. Identificando un objeto y obteniendo su ubicación en el ambiente virtual tridimensional	42

- 3.5. El robot móvil encontrando las marcas espaciales mientras navega en su ambiente. En la parte superior se muestra la vista estéreo (en la imagen izquierda se señala cómo la marca espacial es encontrada y encerrada en un rectángulo blanco) y su representación en el mapa 3D. 44
- 3.6. Robot móvil encontrando una marca espacial y colocando esta en la posición relativa al robot dentro del ambiente virtual. 45
- 3.7. Señales de emergencia como marcas espaciales, ubicados en el laboratorio. 47
- 3.8. Representación de las marcas espaciales y las lecturas del sensor láser. 48
- 4.1. Líneas en \mathbb{R}^2 y su representación 2D en el espacio de parámetros. 50
- 4.2. Planos en \mathbb{R}^3 . 53
- 4.3. Planos en el espacio de parámetros 3D. 54
- 4.4. Planos en \mathbb{R}^3 usando coordenadas polares. 56
- 4.5. Un robot móvil viendo un lugar de trabajo, su representación en el espacio de Hough, y el espacio de Hough con 1 DoF. 58
- 4.6. Una representación de unas escaleras y su transformada de Hough vista desde diferentes ángulos (en la tercera columna φ es la única afectada. 59
- 4.7. Una representación de una entrada a un salón y su representación en el espacio de Hough vista desde diferentes ángulos (en la columna 3 se muestra que solo se afecta el valor de φ). 67

-
- 4.8. (a) Vista de dos escenarios con planos. (b) Imágenes estéreo y representación en el mundo virtual . (c) Los planos representados en \mathbb{R}^3 . (d) El escenario representado en el espacio paramétrico 3D de Hough. 68
- 4.9. Diferentes estructuras donde se muestra al robot detectando el ambiente y los objetos, su representación en coordenadas polares en \mathbb{R}^3 y en el espacio paramétrico 3D de Hough 69
- 5.1. (a) Es la imagen capturada en RGB, y (b) es la información de la profundidad de la imagen (a). 72
- 5.2. (a) El sensor RGB-D (Kinect), y (b) robot móvil con el sensor RGB-D montado en la parte superior. . 74
- 5.3. Esquemático del funcionamiento interno del sensor de profundidad RGB-D. 75
- 5.4. Aplicación de SIFT sobre un robot humanoide. Del lado izquierdo imagen obtenida con el sensor RGB-D. Del lado derecho una imagen de referencia usada como entrenamiento para extraer las características del objeto original. 77
- 5.5. Una marca espacial (extintor) encontrada, en diferentes tomas. 80
- 5.6. Dos vistas del pasillo con planos paralelos al plano de la imagen de la cámara. 81
- 5.7. Segmentación por distancia en la imagen de profundidad. 83
- 5.8. Segmentación de planos paralelos al plano de la imagen de la cámara. 84

5.9. (a) Gráfica del histograma (función de densidad) para la identificación de planos en una escalera. (b) mismo histograma segmentado por colores identificando diferentes planos dentro de la imagen de profundidad.	85
5.10. El robot móvil viendo una escalera, obteniendo los planos que la conforman y su representación en el ambiente virtual 3D.	86
A.1. Origen de coordenadas del sensor láser	94
A.2. Origen de coordenadas del sistema estéreo	96
A.3. Mapa 3D con y sin paredes virtuales	97
A.4. Creación del mapa virtual con primitivas geométricas.	98

Índice de cuadros

2.1. Representación de las entidades y sus duales en Conformal	25
2.2. Transformaciones conformales y sus versores representados	29

Capítulo 1

Introducción

Este documento trata sobre la navegación de robots móviles utilizando técnicas geométricas como el uso del álgebra geométrica conformal y del aprendizaje automático por computadora o *aprendizaje en máquinas* (*machine learning* en inglés) para el reconocimiento de objetos dentro del ambiente donde se encuentra el robot móvil.

La navegación de un robot móvil dentro de un ambiente desconocido es una tarea que se ha estudiado por largo tiempo. Existen varios métodos y aproximaciones que solucionan esta tarea. La utilización de varios sensores de entrada (como sensores ultrasonido, sensores de alcance láser, sensores de tacto, cámaras, etc.) registran el entorno que rodean al robot móvil.

La integración de varios sensores de entrada hacen que la tarea de creación de mapas sea más sencilla. En la figura 1.1 se muestra un esquema de los sensores que son comúnmente usados en los robots móviles. El presente documento propone el uso de un sensor de alcance láser en combinación de un sistema de cámaras estéreo para hacer posible la representación de un ambiente desconocido donde se ubica el robot móvil. Además, al final del documento se expone

un nuevo sensor “el sensor RGB-D”, el cual combina una imagen a color y una imagen de profundidad correspondiente a cada pixel de la imagen en color.

El sensor de alcance láser es un dispositivo que utiliza un haz de luz para medir la distancia hacia un objeto opaco. La medición se hace por medio del reflejo del haz de luz y el tiempo que tarda en regresar al dispositivo. Estos dispositivos suelen tener un dispositivo giratorio donde rota el emisor de luz. El giro suele ser de 180° hasta 310° .

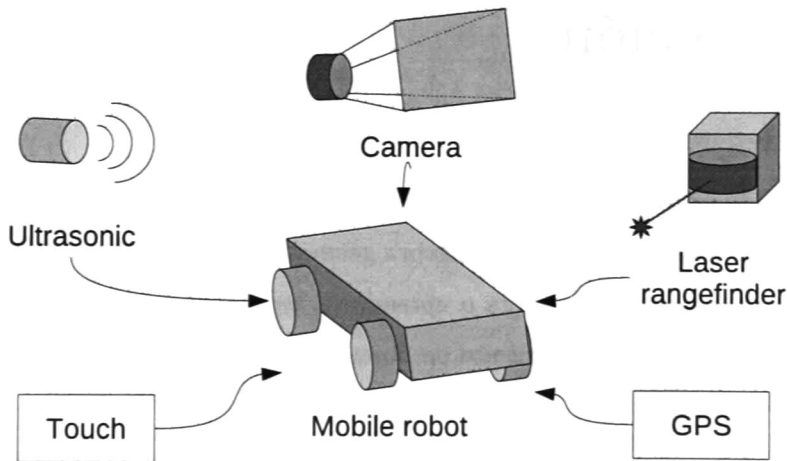


Figura 1.1: Robot móvil al cual se le pueden agregar diferentes sensores de entrada para su exploración.

El uso del sistema de cámaras estéreo en la navegación robótica brinda la oportunidad de capturar los objetos tridimensionales que rodean al robot móvil. Esto se hace por medio del método de triangulación [16]. Además, el robot móvil debe de calcular el lugar en el espacio de los objetos con respecto a la posición actual del robot.

Finalmente el uso de nuevos dispositivos que combinan el sistema de cámaras a color y el sensor de proximidad láser, facilitan el cálculo de la posición de los

objetos y su representación en un ambiente virtual 3D. Nos referimos al sensor RGB-D. Con él se puede extraer una nube densa de puntos con información de profundidad de los píxeles de la imagen a color.

Gracias al aprendizaje automático, es posible el reconocimientos de objetos vistos por las cámaras. Usando el sistema de cámaras estéreo se calcula una representación del objeto en un espacio virtual 3D. Estos objetos servirán como guía en la creación de los mapas.

El *álgebra geométrica* es una herramienta que permite trabajar con objetos geométricos de una manera más fácil, ya que la representación de las entidades geométricas es compacta y las operaciones que se realizan con ellos se interpretan de forma natural en su álgebra. Se trabajará con esta álgebra y en particular en *geometría conformal* para hacer uso de sus entidades. Con estas herramientas se facilita el modelado y representación del robot móvil y los objetos registrados por el mismo. Se obtendrán las propiedades necesarias de cada entidad y su representación en un ambiente virtual.

Este capítulo presenta la motivación, el objetivo y la organización de este trabajo.

1.1. Motivación

Muchas aplicaciones de robótica móvil necesitan navegar en entornos desconocidos. Unos ejemplos son: la BR 700 robots industriales para el transporte; la HELPMATE de Pyxis Corp. también utilizada para tareas de transporte en particular en los hospitales; la limpiadora ROOMBA; la TOURBOT utilizado como un robot guía en museos o Exposiciones; el SOJOURNER utilizados durante las misiones de exploración Pathfinder de Marte; otros robots diseñados para extraer

la madera de los bosques; para aplicaciones agrícolas, el robot KHEPERA de K-Team, la PIONEER de ActivMedia Robotics y el B21 de iRobot Corp. todos para la educación y la investigación, etc. A pesar de amplia gama de las aplicaciones, todos tienen algo en común: la integración de diferentes campos, tales como la dinámica y el control para resolver los problemas de locomoción; el análisis de señal y la visión por computadora, para crear un sólido sistema de percepción, los algoritmos de computadora, teoría de la información, la inteligencia artificial, y la teoría de probabilidad para las tareas de localización y navegación.

Resolver el problema de la navegación involucra contestar las siguientes preguntas:

1. ¿Dónde estoy?
2. ¿Cómo es el mundo que me rodea?
3. ¿Cómo puedo llegar a otro lugar desde donde estoy?

La propuesta mejor conocida para resolver las 2 primeras preguntas es SLAM: *Simultaneous Localization and Mapping* [22] (en español: *Localización y creación de un Mapa Simultáneos*). SLAM resuelve el problema de la localización del robot y creación de un mapa del ambiente que lo rodea basados en la percepción de la información, se hace usando algún método de filtrado, es decir, prediciendo lo que va a ocurrir basados en modelos estocásticos y corrigiéndolos con los datos reales medidos. La tercera pregunta se puede resolver teniendo ya un conocimiento previo de la estructura del ambiente que rodea al robot móvil, el cuál puede ser construido a medida que el robot móvil hace su exploración.

Los mapas son comúnmente usados para la navegación robótica móvil. Para adquirir un mapa, los robots deben de tener sensores que les permita percibir el mundo que los rodea. Sin embargo casi todos los sensores están sujetos a tener

errores, los cuales se refieren, generalmente, a el ruido en las mediciones. Lo más importante es que la mayoría de los sensores de los robots móviles están sujetos a limitaciones de alcance. Estas limitaciones de alcance hacen necesario que el robot navegue dentro de su ambiente, mientras al mismo tiempo, construye un mapa del mismo.

Los sensores que comúnmente son usados en la navegación incluyen cámaras, sensores de alcance como láser, sonares e infrarrojos, radares, sensores táctiles, brújulas, giroscopios y GPS. Inicialmente los sonares fueron usados como un sensor *inteligente*. Sin embargo, basados en el hecho de que los sonares tienen muy poco alcance espacial y son sujetos a ruidos, el uso de sensores del alcance láser fue incrementando. Aunque el sensor láser tiene un alto alcance, también son sujetos al ruido. Además el láser tiene un campo de visión limitado, a menos que se le introduzcan mecanismos tales como espejos de rotación incorporados en el diseño del dispositivo. Debido a esto los investigadores ahora están interesados en el uso de un sistema basado en cámaras (sensores de visión).

Este trabajo pretende cubrir la mayoría de estos problemas para navegar en ambientes parcialmente desconocidos utilizando la combinación del sensor láser y el de visión. Este trabajo se centra principalmente en un robot móvil con ruedas como el Pioneer de ActivMedia, para llevar a cabo el análisis, simulaciones y experimentos.

1.2. Objetivo

Un robot móvil debe ser capaz de navegar en ambientes desconocidos, para hacer esto tiene que construir un mapa 3D del ambiente que lo rodea, con el fin de poder localizarse así mismo dentro del ambiente virtual y decidir hacia donde

dirigirse.

Un estudio, incluyendo la descripción, el análisis e implementación, de un *sistema integrado* que considere la percepción y aprendizaje para un robot móvil en ambientes intramuros, es la principal meta de este trabajo.

Los sub-objetivos son identificados como siguen:

- El estudio y análisis de un método para la localización y creación de un mapa simultáneos, con el fin de enriquecer la navegación y también garantizar una estimación estable de la localización del robot móvil dentro del ambiente creado.

El uso de aprendizaje automático por computadora o *aprendizaje de máquina* para la ubicación de objetos dentro del ambiente, los cuales servirán para la localización del robot móvil.

- Uso combinado de los sensores de entrada del robot móvil (sensor de alcance láser y sistema de cámaras estéreo) para la creación de ambientes 3D.
- Utilizar el *álgebra geométrica conformal* para la representación y modelado del ambiente que rodea al robot móvil.
- Evaluar los resultados de los algoritmos obtenidos tanto en simulación como en una aplicación de *tiempo-real* con un robot móvil equipado con los sensores de visión y láser.

1.3. Organización

La intención de este trabajo es entender el problema de la localización y creación de un mapa simultáneos (SLAM) ayudados por el aprendizaje automáti-

co por computadora. Este documento esta estructurado de la siguiente manera:

Capítulo 1 Introducción a la navegación por medio de robots móviles, los dispositivos de entrada al sistema de un robot, objetivos y organización del documento.

Capítulo 2 Introducción al *álgebra geométrica*, donde se explica cómo son sus entidades y cuál es la forma con que se trabaja con ella; el uso del *álgebra geométrica conformal* con sus entidades y sus representaciones en el espacio, así como las propiedades de algunas de las entidades del álgebra.

Capítulo 3 Se explica la técnica empleada para el reconocimiento de objetos a base del aprendizaje automático por computadora o *aprendizaje en máquinas*. Los objetos reconocidos por las cámaras son representados por una esfera que contienen al objeto en el lugar donde fue registrado.

Capítulo 4 Se desarrolla una nueva técnica para el reconocimiento de objetos o escenarios que este viendo el robot móvil por medio de la transformada de Hough y se usa este espacio para la representación de los mismos. Se usan como entidad principal los planos, ya que estos se encuentran fácilmente en espacios interiores.

Capítulo 5 Se explica cómo el uso de nuevos dispositivos de entrada es usado en la robótica móvil. El uso del sensor RGB-D facilita la representación de un escenario en 3D y con ello la extracción de planos paralelos al plano de la imagen.

Capítulo 2

Geometría Algebraica

Lo que actualmente llamamos *álgebra geométrica*, también es conocida como *álgebra de Clifford*¹, en donde una serie de propiedades y operadores son añadidas al álgebra, para extender el sistema de números reales hacia el cálculo vectorial. El álgebra geométrica es introducida por David Hestenes en la década de 1960.

El álgebra de Clifford (o álgebra geométrica) tiene la propiedad de representar *entidades de orden mayor* en forma simbólica compacta y actuar en ellas de forma lineal. Una *entidad de grado mayor* puede ser visto como un subespacio del espacio vectorial que forma esta álgebra. Esto significa por ejemplo, que líneas y planos son llamados entidades de orden mayor, que son representadas como elementos únicos en el álgebra de Clifford. Esto quiere decir que una línea puede ser diferenciada de otra dentro del álgebra.

El lector puede consultar la siguiente bibliografía para profundizar en el tema [2][24][32][19].

¹William K. Clifford 1845-1879

2.1. Álgebra Geométrica N-Dimensional

Un espacio vectorial n -dimensional puede ser expandido usando las bases ortonormales de sus vectores $\{e_i\}, i = 1, \dots, n$. tal que $e_i e_j = e_{ij}$. Esto nos lleva a una base (2.1) que expande la base ortonormal de vectores del espacio vectorial lineal que corresponde al álgebra geométrica \mathcal{G}_n .

$$1, \{e_i\}, \{e_i \wedge e_j\}, \{e_i \wedge e_j \wedge e_k\}, \dots, I \equiv e_1 \wedge e_2 \wedge \dots \wedge e_n \quad i, j, k \in [1, n] \quad (2.1)$$

En general un álgebra geométrica $\mathcal{G}_{p,q,r}$ ($p, q, r \in \mathbb{N}$) es un espacio lineal de dimensión 2^n , $n = p + q + r$, con un subespacio que llamaremos *blades*, en donde p, q y r corresponden al número de vectores base que cuadran a $+1$, -1 y 0 , respectivamente. La notación $\mathcal{G}_{p,q,r}(\mathbb{R}^{p,q,r})$ es usada para determinar de qué espacio vectorial provienen los elementos del álgebra.

Para aclarar lo anterior, sean e_i y e_j ($e_i, e_j \in \mathbb{R}^{p,q,r}$) dos vectores de la base ortonormal de este espacio vectorial. Entonces el producto geométrico de los vectores de la base, nos da la base del álgebra geométrica $\mathcal{G}_{p,q,r}$ y la cual está definida como

$$e_i e_j := \begin{cases} 1 \in \mathbb{R} & \text{para } i = j \in \{1, \dots, p\} \\ -1 \in \mathbb{R} & \text{para } i = j \in \{p+1, \dots, p+q\} \\ 0 \in \mathbb{R} & \text{para } i = j \in \{p+q+1, \dots, n\} \\ e_{ij} = e_i \wedge e_j = -e_j \wedge e_i & \text{para } i \neq j. \end{cases} \quad (2.2)$$

El producto geométrico (o Clifford) de dos vectores se puede representar como la combinación lineal de un producto interior y un producto exterior y se expresa de la siguiente forma

$$ab = a \cdot b + a \wedge b \quad (2.3)$$

La regla conmutativa $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ junto con la regla anticonmutativa $\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}$ nos da la relación que hay entre \mathbf{ab} y \mathbf{ba} . Así

$$\mathbf{ba} = \mathbf{a} \cdot \mathbf{b} - \mathbf{a} \wedge \mathbf{b} \quad (2.4)$$

Sumando y restando las ecuaciones (2.3) y (2.4), podemos definir

$$\mathbf{a} \cdot \mathbf{b} = \frac{1}{2}(\mathbf{ab} + \mathbf{ba}) \quad (2.5)$$

$$\mathbf{a} \wedge \mathbf{b} = \frac{1}{2}(\mathbf{ab} - \mathbf{ba}) \quad (2.6)$$

La parte anti-simétrica (producto exterior) del producto geométrico es llamado *r-blade* o *blade de grado r*; donde $r \in [2, n]$. Una combinación lineal de los *r-blades* es llamado *r-vector*. El conjunto de *r-vectores* es un subespacio $\binom{n}{r}$ -dimensional de \mathcal{G}_n , denotado por \mathcal{G}_n^r . Todo el espacio \mathcal{G}_n esta dado por la suma directa de sus subespacios

$$\mathcal{G}_n = \sum_{i=0}^n \mathcal{G}_n^i \quad (2.7)$$

Esta suma también se puede representar como una suma de la parte par e impar del álgebra $\mathcal{G}_n = \mathcal{G}_n^+ \oplus \mathcal{G}_n^-$ donde

$$\mathcal{G}_n^+ = \mathcal{G}_n^0 \oplus \mathcal{G}_n^2 \oplus \cdots \oplus \mathcal{G}_n^e \quad \text{parte par} \quad (2.8)$$

$$\mathcal{G}_n^- = \mathcal{G}_n^1 \oplus \mathcal{G}_n^3 \oplus \cdots \oplus \mathcal{G}_n^o \quad \text{parte impar} \quad (2.9)$$

donde $e \leq n$ es par y $o \leq n$ es impar.

Un elemento genérico de \mathcal{G}_n es llamado *multivector*. De acuerdo con (2.7), cada multivector M puede ser escrito en su forma expandida como

$$M = \sum_{i=0}^n \langle M \rangle_i \quad (2.10)$$

donde $\langle M \rangle_i$ denota la parte *i-vector*.

La función exponencial de un multivector M puede ser expresada en series de expansión, como

$$\exp(M) = \sum_{k=0}^{\infty} \frac{M^k}{k!} \quad (2.11)$$

2.1.1. Propiedades de los Multivectores

Un elemento M de \mathcal{G}_n es *invertible* si existe otro elemento N de \mathcal{G}_n tal que $MN = NM = 1$. El elemento N , si este existe, es único. Éste es llamado el *inverso* de M , y es denotado por M^{-1} . Por ejemplo, un vector nulo² de \mathcal{G}_n no es invertible, pero cualquier vector no nulo \mathbf{a} es invertible, con

$$\mathbf{a}^{-1} = \frac{1}{\mathbf{a}} = \frac{\mathbf{a}}{\mathbf{a}^2} \quad (2.12)$$

Esta capacidad del álgebra geométrica para la división de vectores facilita varios cálculos.

El *dual* de un multivector M en \mathcal{G}_n es denotado por

$$M^* = MI_n^{-1} \quad (2.13)$$

donde $I_n = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \dots \wedge \mathbf{e}_n$ es el elemento de grado máximo del álgebra llamado *pseudo-escalar*. I_n^{-1} difiere de I_n por lo menos por un signo. El *dual* de un r -blade es un $(n-r)$ -blade.

La *reversión* de un r -blade $A = a_1 \wedge \dots \wedge a_r$ se denota como

$$\begin{aligned} \tilde{A} &= (a_1 \wedge a_2 \wedge \dots \wedge a_{r-1} \wedge a_r)^\sim \\ &= a_r \wedge a_{r-1} \wedge \dots \wedge a_2 \wedge a_1 \end{aligned} \quad (2.14)$$

y generalizando, la *reversión* para un multivector

²Un vector \mathbf{a} es nulo si $\mathbf{a}^2 = 0$

$$\langle \widetilde{M} \rangle_i = (-1)^{\frac{i(i-1)}{2}} \langle M \rangle_i, \text{ para } M \in \mathcal{G}_n, 0 \leq i \leq n. \quad (2.15)$$

Dado un producto geométrico de varios multivectores $D = ABC$, donde $A, B, C, D \in \mathcal{G}_n$ la reversión de estos está dada por

$$\widetilde{D} = \widetilde{ABC} = \widetilde{C}\widetilde{B}\widetilde{A} \quad (2.16)$$

La *involución principal* de \mathcal{G}_n , también llamada *involución de grado* es denotada por " $\widehat{}$ ", y definida por

$$\langle \widehat{M} \rangle_i = (-1)^i \langle M \rangle_i, \text{ para } M \in \mathcal{G}_n, 0 \leq i \leq n. \quad (2.17)$$

2.2. Geometría Conformal

Es necesario introducir un álgebra que envuelva a varias entidades geométricas (puntos, líneas, planos, esferas, círculos, etc.). El uso de la geometría conformal está motivada por la introducción de proyecciones estereográficas y la inclusión del espacio de Minkowski.

2.2.1. Espacio de Minkowski

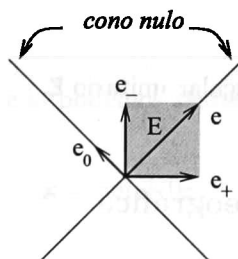


Figura 2.1: Vectores base y líneas nulas del plano de Minkowski.

El plano de *Minkowski* $\mathbb{R}^{1,1}$ tiene una base ortonormal $\{e_+, e_-\}$ que definimos con las propiedades

$$e_+^2 = 1, \quad e_-^2 = -1, \quad e_+ \cdot e_- = 0. \quad (2.18)$$

Una base *nula* $\{e_0, e\}$ la introducimos como

$$e_0 := \frac{1}{2}(e_- - e_+) \quad \text{y} \quad e := e_- + e_+ \quad (2.19)$$

Estos vectores pueden interpretarse como el origen e_0 de nuestro sistema de coordenadas y un punto en el infinito e . Alternativamente, la base nula puede ser definida directamente en términos de las siguientes propiedades

$$e_0^2 = e^2 = 0, \quad e \cdot e_0 = -1 \quad (2.20)$$

Un pseudo-escalar unitario E para $\mathbb{R}^{1,1}$ es definido por

$$E := e_0 \wedge e = e_+ \wedge e_- = e_+ e_- \quad (2.21)$$

donde tenemos las siguientes propiedades

$$\begin{aligned} E^2 = 1, \quad \tilde{E} = -E, \quad Ee = -e, \quad e_0 E = e_0 \\ e_+ E = e_-, \quad e_- E = e_+, \quad e_+ e = 1 + E \\ e_- e = -(1 + E), \quad e \wedge e_- = E, \quad e_+ \cdot e = 1. \end{aligned} \quad (2.22)$$

La base vectorial y líneas nulas de $\mathbb{R}^{1,1}$ son mostradas en la fig 2.1, donde el área sombreada representa el pseudo-escalar unitario E .

2.2.2. Proyección Estereográfica

En pocas palabras, una proyección estereográfica es el medio para generar mapas planos de la tierra. En este caso, imaginemos a la tierra como una esfera

transparente, intersecada en el ecuador con un plano. Ahora imaginemos que en el polo norte n hay un foco que ilumina a través de la esfera. Cada punto de la esfera proyectará un punto en el papel y ahí es donde se dibuja en el mapa. La regla para la proyección estereográfica tiene una clara descripción geométrica y es ilustrada para el caso de 1D en la figura 2.2.

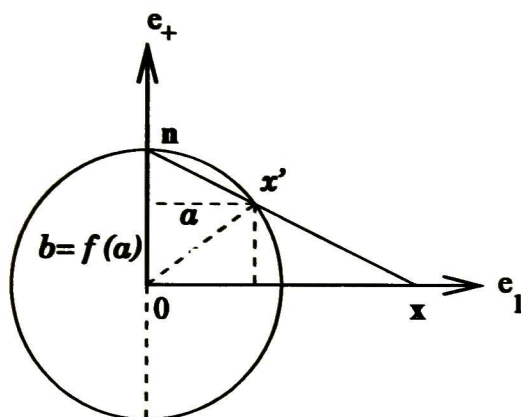


Figura 2.2: Proyección estereográfica en 1D

Antes de introducir formalmente algunos términos del álgebra geométrica conformal, desarrollaremos los cálculos básicos de las propiedades de la proyección estereográfica. Por simplicidad usaremos el caso de 1D (\mathcal{G}_1), y trabajaremos en \mathcal{G}_2 con los vectores base $\{e_1, e_+\}^3$. La proyección la haremos con un círculo de radio 1.

Sea x un punto en e_1 que lo podemos representar en \mathcal{G}_2 como

$$x = xe_1 + 0e_+ \quad (2.23)$$

y n un punto en e_+ ($n = 0e_1 + 1e_+$), que está en el polo norte del círculo unitario.

³ Llamamos e_+ a este vector base porque cuadra a $+1$, al igual que e_1 ($e_1^2 = e_+^2 = 1$).

La ecuación de la recta entre estos dos puntos es

$$f(a) = -\frac{1}{x}a + 1 \quad (2.24)$$

y la ecuación del círculo unitario es

$$a^2 + f(a)^2 = 1 \quad (2.25)$$

Necesitamos encontrar la solución a este sistema de ecuaciones para saber en donde se intersecan las funciones, sustituyendo (2.24) en (2.25)

$$a^2 + \left(-\frac{1}{x}a + 1\right)^2 = 1 \quad (2.26)$$

y desarrollando el cuadrado de (2.26)

$$a^2 + \frac{a^2}{x^2} - 2\frac{a}{x} = 0 \quad (2.27)$$

factorizando a

$$a\left(a + \frac{a}{x^2} - \frac{2}{x}\right) = 0 \quad (2.28)$$

la primer solución de (2.28) es $a = 0$ que es el polo norte. La otra solución sale con

$$a\left(1 + \frac{1}{x^2}\right) = \frac{2}{x} \quad (2.29)$$

despejando a

$$a = \frac{2x}{x^2 + 1} \quad (2.30)$$

sustituyendo el valor de a en (2.24) nos queda

$$f(a) = -\frac{1}{x}\left(\frac{2x}{x^2 + 1}\right) + 1 \quad (2.31)$$

$$= \frac{x^2 - 1}{x^2 + 1} \quad (2.32)$$

con esto obtenemos el punto⁴ $\mathbf{x}' = ae_1 + be_+$ que queda como⁵

$$\begin{aligned}\mathbf{x}' &= ae_1 + be_+ \\ &= \frac{2x}{x^2 + 1}e_1 + \frac{x^2 - 1}{x^2 + 1}e_+\end{aligned}\quad (2.33)$$

Usando coordenadas homogéneas agregamos una tercera base $\{e_-\}$ ⁶, trabajaremos ahora en $\mathcal{G}_{2,1}$, y la ecuación (2.33) queda de la siguiente forma

$$\mathbf{x}'' = \frac{2x}{x^2 + 1}e_1 + \frac{x^2 - 1}{x^2 + 1}e_+ + e_- \quad (2.34)$$

multiplicando (2.34) por $\frac{x^2+1}{2}$ nos queda una representación homogénea del punto sobre el círculo como

$$\mathbf{x}''' = xe_1 + \frac{1}{2}(x^2 - 1)e_+ + \frac{1}{2}(x^2 + 1)e_- \quad (2.35)$$

haciendo una factorización de los x^2 nos queda así

$$\mathbf{x}''' = xe_1 + \frac{1}{2}(x^2)(e_- + e_+) + \frac{1}{2}(e_- - e_+) \quad (2.36)$$

veremos después como ésta representación (2.36) nos será mucho más conveniente en el álgebra conformal. En la figura (2.3) se muestra cómo queda este punto en este espacio.

2.2.3. Definición del Álgebra Geométrica Conformal

El rol que juega el plano de Minkowski de generar *vectores nulos*, y expandir el espacio vectorial Euclidiano \mathbb{R}^n a $\mathbb{R}^{n+1,1} = \mathbb{R}^n \oplus \mathbb{R}^{1,1}$, resulta en el álgebra geométrica conformal $\mathcal{G}_{n+1,1}$. El conjunto N^{n+1} de todos los vectores nulos en \mathbb{R}^n

⁴Sustituyendo $f(a) = b$

⁵El sistema tiene dos soluciones, uno el trivial, donde la línea es intersecada en el polo norte n y la otra en \mathbf{x}'

⁶ Llamamos a esta base e_- porque cuadra a -1 , ($e_-^2 = -1$)

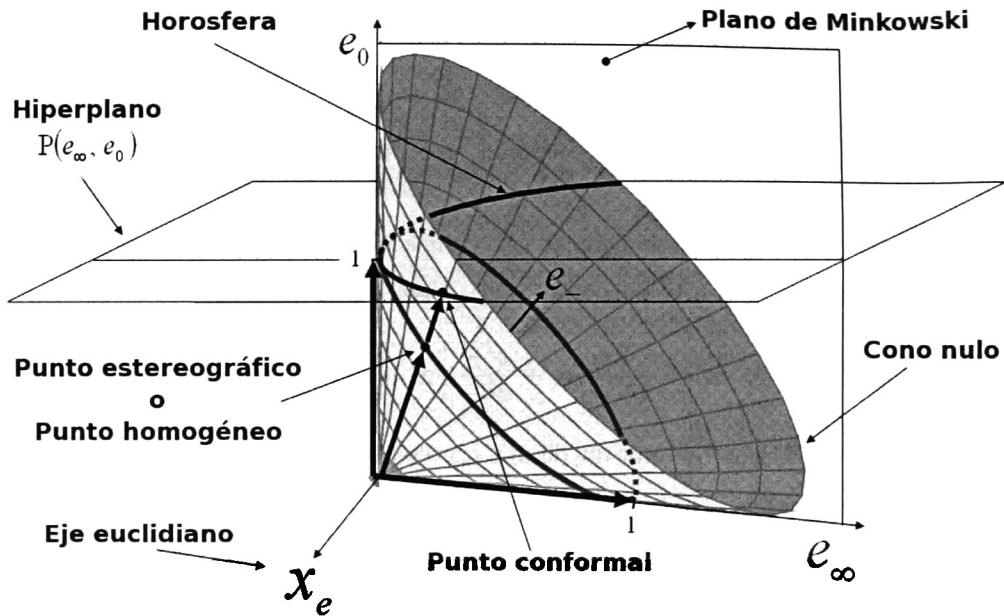


Figura 2.3: Visualización del modelo homogéneo para la proyección estereográfica en el caso 1D.

es llamado *cono nulo*, y el conjunto de los puntos generalizados homogéneos de grado superior es llamado *horósfera*. En la fig. 2.3 se muestra la visualización del modelo homogéneo para la proyección estereográfica en el caso 1D. Todos los puntos estereográficos proyectados viven en el cono, que es el *cono nulo* del espacio de Minkowski. También se notan el Hiperplano $\mathbb{P}(e, e_0)$ y la horósfera del álgebra conformal.

Un vector en este espacio puede ser representado como una *descomposición ortogonal*

$$\underline{a} = P_E(\underline{a}) + P_E^\perp(\underline{a}) \quad (2.37)$$

donde los operadores de proyección P_E y retroproyección o componente perpen-

dicular $P_{\mathbf{E}}^{\perp}$ están dados por

$$P_{\mathbf{E}}(\underline{a}) = (\underline{a} \cdot \mathbf{E})\mathbf{E} = \alpha\mathbf{e}_0 + \beta\mathbf{e} \in \mathbb{R}^{1,1} \quad (2.38)$$

$$P_{\mathbf{E}}^{\perp}(\underline{a}) = (\underline{a} \cdot \mathbf{E}^*)\tilde{\mathbf{E}}^* = (\underline{a} \wedge \mathbf{E})\mathbf{E} = \mathbf{a} \in \mathbb{R}^n \quad (2.39)$$

El espacio vectorial conformal derivado de \mathbb{R}^3 es denotado como $\mathbb{R}^{4,1}$. Una base vectorial esta dada por $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_+, \mathbf{e}_-\}$. El álgebra correspondiente $\mathcal{G}_{4,1}$ contiene $2^5 = 32$ elementos. El pseudo-escalar unitario⁷ esta denotado como

$$\mathbf{I}_c = \mathbf{e}_{+123} = \mathbf{E}\mathbf{I}_{\mathbf{E}} \quad (2.40)$$

En seguida se presentan las entidades del álgebra geométrica conformal, las cuales están en el cuadro 2.1.

Puntos

En esta álgebra los *puntos* son considerados como elementos del llamado cono nulo, y obedecen a las siguientes propiedades

$$\{\underline{x} \in \mathcal{G}_{n+1,1} | \underline{x}^2 = 0, \underline{x} \cdot \mathbf{e} = -1\} \quad (2.41)$$

estas dos restricciones nos llevan a que un punto conformal tiene la forma geométrica

$$\underline{x} = \mathbf{x} + \frac{1}{2}\mathbf{x}^2\mathbf{e} + \mathbf{e}_0 \quad (2.42)$$

donde $\underline{x} \in \mathcal{G}_{3+1,1}$ y $\mathbf{x} \in \mathbb{R}^3$. Sustituyendo los valores de (2.19) en (2.42) tenemos

$$\begin{aligned} \underline{x} &= \mathbf{x} + \frac{1}{2}\mathbf{x}^2\mathbf{e} + \mathbf{e}_0 \\ &= \mathbf{x} + \frac{1}{2}\mathbf{x}^2(\mathbf{e}_- - \mathbf{e}_+) + \frac{1}{2}(\mathbf{e}_- + \mathbf{e}_+) \\ &= \mathbf{x} + \frac{1}{2}(\mathbf{x}^2 - 1)\mathbf{e}_+ + \frac{1}{2}(\mathbf{x}^2 + 1)\mathbf{e}_- \end{aligned} \quad (2.43)$$

⁷El pseudo-escalar unitario de \mathcal{G}_3 se representa como $\mathbf{I}_{\mathbf{E}} = \mathbf{e}_{123}$

Como podemos ver, al desarrollar la ecuación (2.43), es clara la relación que hay con las ecuaciones (2.35) y (2.36). Esta es exactamente la representación homogénea de un punto proyectado estereográficamente sobre el círculo pero ahora en \mathbb{R}^3 . La representación de un punto es más compacta si es escrita usando $\{\underline{e}, \underline{e}_0\}$ en vez de $\{e_+, e_-\}$.

Poniendo $\mathbf{x} = 0$ en (2.42) podemos ver que $\underline{x} = \underline{e}_0$ es el punto homogéneo que corresponde al origen en \mathbb{R}^n . De

$$\frac{\underline{x}}{-\underline{x} \cdot \underline{e}_0} = \underline{e} + 2 \left(\frac{\mathbf{x} + \mathbf{e}_0}{\mathbf{x}^2} \right) \xrightarrow{\mathbf{x}^2 \rightarrow \infty} \underline{e} \quad (2.44)$$

vemos que cuando \mathbf{x}^2 tiende al infinito, obtenemos como resultado \underline{e} . Con esto podemos decir que \underline{e} representa el punto en el infinito.

Usando (2.42) podemos expandir el producto geométrico de dos puntos \underline{a} y \underline{b} como

$$\underline{a}\underline{b} = \underline{a}\underline{b} + (\underline{a} - \underline{b})\underline{e}_0 - \frac{1}{2}[(\underline{a}^2 + \underline{b}^2) + (\underline{b}\underline{a}^2 - \underline{a}\underline{b}^2)\underline{e} + (\underline{b}^2 - \underline{a}^2)\underline{E}]. \quad (2.45)$$

La parte escalar de (2.45) nos da

$$\underline{a} \cdot \underline{b} = -\frac{1}{2}(\underline{a} - \underline{b})^2 = \underline{a}\underline{b} - \frac{1}{2}(\underline{a}^2 + \underline{b}^2) \quad (2.46)$$

Así, el producto interno de dos puntos homogéneos nos da directamente la distancia Euclidiana al cuadrado entre ellos. Como $\underline{a}^2 = \underline{b}^2 = 0$, tenemos que

$$(\underline{a} - \underline{b})^2 = -2\underline{a} \cdot \underline{b} = (\underline{a} - \underline{b})^2 \quad (2.47)$$

Par de Puntos

Tomando la parte bivectorial de (2.45) nos queda una entidad llamada *par de puntos* en su forma dual, la cual, en una sola entidad guarda la localización

de 2 puntos pertenecientes al álgebra y esta definida como

$$\begin{aligned} \underline{PP^*} &= \underline{a} \wedge \underline{b} \\ &= \mathbf{a} \wedge \mathbf{b} + (\mathbf{a} - \mathbf{b}) \wedge \mathbf{e}_0 - \frac{1}{2}[(\mathbf{b}\mathbf{a}^2 - \mathbf{a}\mathbf{b}^2) \wedge \mathbf{e} + (\mathbf{b}^2 - \mathbf{a}^2)\mathbf{E}]. \end{aligned} \quad (2.48)$$

En su forma estándar, el par de puntos es la intersección de 3 esferas.

Esferas

La ecuación de la *esfera* de radio ρ con centro en el punto $\mathbf{p} \in \mathbf{R}^n$ puede ser expresada como

$$(\mathbf{x} - \mathbf{p})^2 = \rho^2 \quad (2.49)$$

Usando (2.46), podemos expresar (2.49) como una ecuación equivalente en términos de puntos homogéneos:

$$\underline{\mathbf{x}} \cdot \underline{\mathbf{p}} = -\frac{1}{2}(\mathbf{x} - \mathbf{p})^2 = -\frac{1}{2}\rho^2 \quad (2.50)$$

Usando la restricción $\underline{\mathbf{x}} \cdot \mathbf{e} = -1$ de (2.41), podemos simplificar (2.50) a

$$\begin{aligned} \underline{\mathbf{x}} \cdot \underline{\mathbf{p}} &= (\underline{\mathbf{x}} \cdot \mathbf{e}) \frac{1}{2}\rho^2 \\ \underline{\mathbf{x}} \cdot \underline{\mathbf{p}} - (\underline{\mathbf{x}} \cdot \mathbf{e}) \frac{1}{2}\rho^2 &= 0 \\ \underline{\mathbf{x}} \cdot (\underline{\mathbf{p}} - \frac{1}{2}\rho^2 \mathbf{e}) &= 0 \\ \underline{\mathbf{x}} \cdot \underline{\mathbf{s}} &= 0 \end{aligned} \quad (2.51)$$

donde

$$\begin{aligned} \underline{\mathbf{s}} &= \underline{\mathbf{p}} - \frac{1}{2}\rho^2 \mathbf{e} \\ &= (\mathbf{p} + \frac{1}{2}\mathbf{p}^2 \mathbf{e} + \mathbf{e}_0) - \frac{1}{2}\rho^2 \mathbf{e} \\ &= \mathbf{p} + \frac{1}{2}(\mathbf{p}^2 - \rho^2) \mathbf{e} + \mathbf{e}_0 \end{aligned} \quad (2.52)$$

es la representación de la esfera en el álgebra geométrica conformal. La entidad \underline{s} tiene las propiedades

$$\underline{s}^2 = \rho^2 > 0 \quad (2.53)$$

$$\mathbf{e} \cdot \underline{s} = -1 \quad (2.54)$$

Con estas propiedades podemos recuperar la forma (2.52) y el centro \underline{p} .

Alternativamente, una esfera puede ser representada en su forma dual de $\underline{s}^* = \underline{s}I_c^{-1}$. La ventaja de la forma dual de la esfera es que puede ser calculada desde 4 puntos conformales (en 3D) que estén en su superficie [4] y puede ser escrita de la forma

$$\underline{s}^* = \underline{a} \wedge \underline{b} \wedge \underline{c} \wedge \underline{d} \quad (2.55)$$

La ecuación (2.51) tiene su forma dual que es

$$\underline{x} \wedge \underline{s}^* = 0 \quad (2.56)$$

tanto (2.51) como (2.56) representan a un punto que esta sobre la superficie de la esfera. Otras propiedades de la esfera se verán mas adelante.

Plano

Si reemplazamos uno de los puntos de (2.55) por el punto en el infinito tenemos

$$\begin{aligned} \underline{P}^* &= \mathbf{e} \wedge \underline{a} \wedge \underline{b} \wedge \underline{c} \\ &= \mathbf{e} \wedge \mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} + \mathbf{E}(\mathbf{b} - \mathbf{a}) \wedge (\mathbf{c} - \mathbf{a}) \\ &= \mathbf{e}I_E d + \mathbf{NE} \end{aligned} \quad (2.57)$$

Geoméricamente podemos interpretar esto como una esfera en su forma dual, ya que es un wedge de 4 puntos, pero uno de los puntos es el infinito, esto nos

da una esfera con radio infinito que equivale a un *plano*. En (2.57) se representa la ecuación del plano pasando por los puntos \underline{a} , \underline{b} y \underline{c} . Podemos reconocer a $d\underline{I}_E = \underline{a} \wedge \underline{b} \wedge \underline{c}$ como el momento del plano con tangente $\underline{N} = (\underline{b} - \underline{a}) \wedge (\underline{c} - \underline{a})$, que es la representación de Hesse del plano con normal \underline{N} y distancia d .

Círculo

Un *círculo* \underline{z} puede ser representado como la intersección de dos esferas \underline{s}_1 y \underline{s}_2 . Esto significa que para cada punto $\underline{x} \in \underline{z}$

$$\underline{x} \in \underline{z} \iff \underline{x} \in \underline{s}_1 \text{ y } \underline{x} \in \underline{s}_2 \quad (2.58)$$

Asumiendo que \underline{s}_1 y \underline{s}_2 son linealmente independientes, podemos escribir usando (2.51)

$$\begin{aligned} \underline{x} \in \underline{z} &\iff (\underline{x} \cdot \underline{s}_1)\underline{s}_2 - (\underline{x} \cdot \underline{s}_2)\underline{s}_1 = 0 \\ &\iff \underline{x} \cdot (\underline{s}_1 \wedge \underline{s}_2) = 0 \\ &\iff \underline{x} \cdot \underline{z} = 0 \end{aligned} \quad (2.59)$$

donde $\underline{z} = \underline{s}_1 \wedge \underline{s}_2$ es la intersección de las esferas. La intersección con una tercera esfera nos regresa la entidad par de puntos [5].

La forma dual del círculo (en 3D) puede ser expresado por tres puntos viviendo sobre él [5], como

$$\underline{z}^* = \underline{a} \wedge \underline{b} \wedge \underline{c} \quad (2.60)$$

donde $\underline{a}, \underline{b}, \underline{c} \in \underline{z}^*$.

Línea

Las *líneas*, así como en el caso de los planos, se pueden definir como un círculo que pasa por el punto en el infinito haciendo un círculo de radio infinito. Su forma dual se define como

$$\begin{aligned}
 \underline{L}^* &= \mathbf{e} \wedge \underline{a} \wedge \underline{b} \\
 &= \mathbf{e} \wedge \mathbf{a} \wedge \mathbf{b} + (\mathbf{b} - \mathbf{a}) \wedge \mathbf{E} \\
 &= \mathbf{eM} + \mathbf{rE}
 \end{aligned} \tag{2.61}$$

donde $\mathbf{M} = \mathbf{a} \wedge \mathbf{b}$ es un bivector representando el momento [2] (que es el plano donde vive la línea) y $\mathbf{r} = \mathbf{b} - \mathbf{a}$ representa la dirección de la línea en coordenadas de Plücker. También se puede ver que esta línea pasa por los puntos \mathbf{a} y \mathbf{b} .

2.2.4. Transformaciones Conformales

Una transformación de figuras geométricas se dice que es *conformal* si preserva su *forma*; esto es, si dicha figura preserva sus ángulos y por lo tanto la forma de líneas, rectas y círculos. Cualquier transformación conformal en \mathbb{R}^n puede ser expresada como una composición de *inversiones* en esferas y *reflexiones* en hiperplanos. Un *versor* es cualquier multivector que puede expresarse como el producto geométrico de vectores invertibles. En general, una transformación conformal [19] tiene la forma

$$g(\underline{x}) = G\underline{x}\widehat{G}^{-1} = \sigma\underline{x}' \tag{2.62}$$

donde $\underline{x} \in \mathbb{R}^{n+1,1}$, G es un versor y σ es un escalar. Como el cono nulo es invariante bajo G , (i.e. $(\underline{x}')^2 = \underline{x}^2 = 0$), es conveniente aplicar un factor escalar σ para asegurar $\underline{x}' \cdot \mathbf{e} = \underline{x} \cdot \mathbf{e} = -1$.

Entidad	Representación	Grado	Representación Dual	Grado
Esfera	$\underline{s} = \mathbf{p} + \frac{1}{2}(\mathbf{p}^2 - \rho^2)\mathbf{e} + \mathbf{e}_0$	1	$\underline{s}^* = \underline{a} \wedge \underline{b} \wedge \underline{c} \wedge \underline{d}$	4
Punto	$\underline{x} = \mathbf{x} + \frac{1}{2}\mathbf{x}^2\mathbf{e} + \mathbf{e}_0$	1	$\underline{x}^* = (-E\mathbf{x} - \frac{1}{2}\mathbf{x}^2\mathbf{e} + \mathbf{e}_0)\mathbf{I}_E$	4
Plano	$\underline{P} = N\mathbf{I}_E - d\mathbf{e}$ $N = (\mathbf{a} - \mathbf{b}) \wedge (\mathbf{a} - \mathbf{c})$ $d = (\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c})\mathbf{I}_E$	1	$\underline{P}^* = \mathbf{e} \wedge \underline{a} \wedge \underline{b} \wedge \underline{c}$	4
Línea	$\underline{L} = \underline{P}_1 \wedge \underline{P}_2$ $= \mathbf{r}\mathbf{I}_E + \mathbf{e}\mathbf{M}\mathbf{I}_E$ $\mathbf{r} = \mathbf{a} - \mathbf{b}$ $\mathbf{M} = \mathbf{a} \wedge \mathbf{b}$	2	$\underline{L}^* = \mathbf{e} \wedge \underline{a} \wedge \underline{b}$	3
Círculo	$\underline{z} = \underline{s}_1 \wedge \underline{s}_1$ $\underline{s}_z = (\mathbf{e} \cdot \underline{z})\underline{z}$ $\rho_z = \frac{z^2}{(\mathbf{e} \wedge \underline{z})^2}$	2	$\underline{z}^* = \underline{a} \wedge \underline{b} \wedge \underline{c}$	3
Par de Puntos	$\underline{PP} = \underline{s}_1 \wedge \underline{s}_2 \wedge \underline{s}_3$	3	$\underline{PP}^* = \underline{a} \wedge \underline{b}, \underline{X}^* = \mathbf{e} \wedge \underline{x}$	2

Cuadro 2.1: Representación de las entidades y sus duales en Conformal

Los versores en \mathcal{G}_n forman un grupo bajo el producto geométrico, llamado *grupo de versores*. Los versores de magnitud unitaria forman un subgrupo, llamado *grupo alfiler o perno* (pin group).

Un *versor par* es aquél que $\mathbf{V} = \widehat{\mathbf{V}}$ y es llamado *spinor* (de giro) o *rotor* si

$$\mathbf{V}\widetilde{\mathbf{V}} = |\mathbf{V}|^2 \tag{2.63}$$

y un *versor impar* es aquel que $\mathbf{V} = -\widehat{\mathbf{V}}$.

Veamos a continuación algunos de estos versores.

Inversiones

La forma general de una reflexión sobre un vector $\underline{s} = -\widehat{s}$ es

$$s(\underline{x}) = -\underline{s}\underline{x}s^{-1} = \underline{x} - 2(\underline{s} \cdot \underline{x})\underline{s}^{-1} = \sigma\underline{x}', \quad (2.64)$$

donde $\underline{s}\underline{x} + \underline{x}\underline{s} = 2(\underline{s} \cdot \underline{x})$ de (2.5). Analizaremos ahora que pasa cuando \underline{s} representa una esfera y tiene la forma de (2.52). Si \underline{s} representa una esfera unitaria centrada en el origen ($\rho = 1$ y $\mathbf{p} = 0$), entonces \underline{s} y \underline{s}^{-1} se reducen a $\mathbf{e}_0 - \frac{1}{2}\mathbf{e}$. Sustituyendo $-2(\underline{s} \cdot \underline{x}) = \mathbf{x}^2 - 1$ en (2.64) obtenemos

$$\sigma\underline{x}' = (\mathbf{x} + \frac{1}{2}\mathbf{x}^2\mathbf{e} + \mathbf{e}_0) + (\mathbf{x}^2 - 1)(\mathbf{e}_0 - \frac{1}{2}\mathbf{e}) = \mathbf{x}^2(\mathbf{x}^{-1} + \frac{1}{2}\mathbf{x}^{-2}\mathbf{e} + \mathbf{e}_0), \quad (2.65)$$

el cual es el mapeo conformal de \mathbf{x}^{-1} , tomando como $\sigma = \mathbf{x}^2$

Veamos ahora cómo una esfera arbitraria invierte un punto, tomando de nuevo (2.52) tenemos

$$\underline{s} \cdot \underline{x} = \underline{p} \cdot \underline{x} - \frac{1}{2}\rho^2\mathbf{e} \cdot \underline{x} = -\frac{1}{2}[(\mathbf{x} - \mathbf{p})^2 - \rho^2] \quad (2.66)$$

sustituyendo en (2.64) y haciendo un poco de álgebra nos queda

$$\sigma\underline{x}' = \left(\frac{\mathbf{x} - \mathbf{p}}{\rho}\right)^2 [g(\mathbf{x}) + \frac{1}{2}g^2(\mathbf{x})\mathbf{e} + \mathbf{e}_0], \quad (2.67)$$

donde

$$g(\mathbf{x}) = \frac{\rho^2}{\mathbf{x} - \mathbf{p}} + \mathbf{p} = \frac{\rho^2(\mathbf{x} - \mathbf{p})}{(\mathbf{x} - \mathbf{p})^2} + \mathbf{p} \quad (2.68)$$

es una inversión en \mathbb{R}^n

Reflexiones

Un hiperplano con normal unitaria \mathbf{n} y una distancia δ (con signo) del origen en \mathbb{R}^n puede ser representada por el vector

$$\underline{P} = \underline{s} = \mathbf{n} + \delta\mathbf{e} \quad (2.69)$$

Insertando $\underline{s} \cdot \underline{x} = \mathbf{n} \cdot \mathbf{x}$ en (2.64), podemos encontrar que

$$g(\mathbf{x}) = \mathbf{n}\mathbf{x}\hat{\mathbf{n}} + 2\delta\mathbf{n} = \mathbf{n}(\mathbf{x} - \delta\mathbf{n})\hat{\mathbf{n}} + \delta\mathbf{n} \quad (2.70)$$

La expresión (2.70) es equivalente a la reflexión $\mathbf{n}\mathbf{x}\hat{\mathbf{n}}$ en el origen, trasladada por δ a lo largo de la dirección de \mathbf{n} . Un punto \mathbf{c} esta en el hiperplano cuando $\delta = \mathbf{n} \cdot \mathbf{c}$, en donde (2.69) puede ser escrita

$$\underline{P} = \underline{s} = \mathbf{n} + (\mathbf{n} \cdot \mathbf{c})\mathbf{e} \quad (2.71)$$

Vía (2.70), este vector representa la reflexión en el hiperplano a través del punto \mathbf{c} .

Traslaciones

Una traslación puede ser modelada como dos reflexiones sobre dos hiperplanos paralelos (ver fig. 2.4). Asumiendo que los dos planos están normalizados esto quiere decir que la magnitud de su normal sea igual a uno y que uno de ellos pasa por el origen (el valor de la distancia $\delta = 0$). Entonces de (2.69) podemos representar el operador de traslación (llamado trasladador) como

$$\begin{aligned} \mathbf{T} &= \underline{P}_1 \underline{P}_2 = (\mathbf{n} + \delta\mathbf{e})(\mathbf{n} + 0\mathbf{e}) \\ &= 1 + \frac{1}{2}\mathbf{t}\mathbf{e} \end{aligned} \quad (2.72)$$

donde $\mathbf{t} = 2\delta\mathbf{n}$ y $\|\mathbf{n}\| = 1$. La distancia de traslación es dos veces la separación entre los hiperplanos.

Rotaciones

Usando (2.71), las rotaciones pueden ser modeladas por la composición de dos reflexiones sobre dos hiperplanos que se intersectan en un punto \mathbf{c} (vea fig.

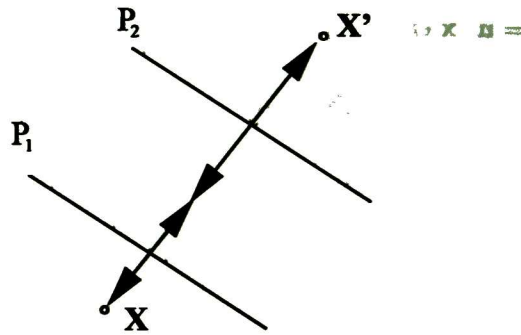


Figura 2.4: La traslación como una reflexión sobre dos planos paralelos

2.5) como

$$\mathbf{R} = (\mathbf{a} + \mathbf{e}(\mathbf{a} \cdot \mathbf{c}))(\mathbf{b} + \mathbf{e}(\mathbf{b} \cdot \mathbf{c})) = \mathbf{ab} + \mathbf{e}(\mathbf{c} \cdot (\mathbf{a} \wedge \mathbf{b})) \quad (2.73)$$

donde \mathbf{a} y \mathbf{b} son las normales unitarias. Escribimos $R = \mathbf{ab}$ y $\mathbf{c} \cdot (\mathbf{a} \wedge \mathbf{b}) = \mathbf{c} \times R$, para simplificar notación. Así se establece que el producto de dos reflexiones en cualquier punto es equivalente a una rotación sobre ese punto.

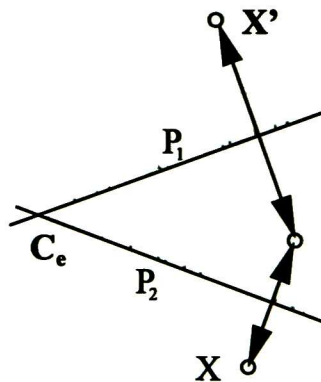


Figura 2.5: La rotación como una reflexión sobre dos planos que se intersecan

Tipo	$g(\mathbf{x})$ en \mathbb{R}^n	Versor en $\mathcal{G}_{n+1,1}$	$\sigma(x)$
Inversión	$\frac{\rho^2}{x-p} + p$	$\underline{s} = p - \frac{1}{2}\rho^2 e$	$\left(\frac{x-p}{\rho}\right)^2$
Reflexión	$-n x n + 2x\delta$	$\underline{P} = n + \delta e$	1
Traslación	$x - t$	$\underline{T} = 1 + \frac{1}{2}te$	1
Rotación	$R(x - c)R^{-1} + c$	$\underline{R} = R + e(c \times R)$	1

Cuadro 2.2: Transformaciones conformales y sus versores representados

2.2.5. Movimientos Rígidos: Rotores, Trasladores y Motores

En esta sección hablaremos de los movimientos rígidos de las entidades en el álgebra conformal. Se ha mencionado anteriormente sobre las transformaciones conformales y su representación. Ahora veremos otra forma de representar algunos de esos versores.

Los bivectores del álgebra geométrica pueden ser usados para representar rotaciones de entidades en el espacio de 3D. Un *rotor*⁸ \underline{R} es un elemento par del álgebra \mathcal{G}_3 que satisface $\underline{R}\tilde{\underline{R}} = 1$. En $\mathcal{G}_{4,1}$ una rotación puede ser expresada por un rotor de la forma

$$\underline{R} = \exp\left(-\frac{\theta}{2}\underline{l}\right) \quad (2.74)$$

donde los componentes de este rotor $\underline{R} \in \mathcal{G}_{4,1}^+$ son similares al rotor de \mathcal{G}_3 , un bivector unitario $\underline{l} \in (\mathcal{G}_3)_2 \subseteq \mathcal{G}_{4,1}$, el cual representa una línea en conformal (de

⁸Rotor es una abreviación de *rotador*

magnitud unitaria) que pasa por el origen (sin momento) que es el eje de rotación, y un ángulo θ , el cual representa el ángulo de rotación. Usando la representación de Euler del rotor (2.74) nos queda como

$$\begin{aligned}\mathbf{R} &= \exp\left(-\frac{\theta}{2}\underline{l}\right) \\ &= \cos\left(\frac{\theta}{2}\right) - \text{sen}\left(\frac{\theta}{2}\right)\underline{l}\end{aligned}\quad (2.75)$$

desarrollamos (2.75) para encontrar también a $\tilde{\mathbf{R}}$

$$\mathbf{R} = r_c - r_s\underline{l} = r_c - r_{s_1}\mathbf{e}_{23} - r_{s_2}\mathbf{e}_{31} - r_{s_3}\mathbf{e}_{12} \quad (2.76)$$

$$\tilde{\mathbf{R}} = r_c + r_s\underline{l} = r_c + r_{s_1}\mathbf{e}_{23} + r_{s_2}\mathbf{e}_{31} + r_{s_3}\mathbf{e}_{12} \quad (2.77)$$

La rotación de una entidad (punto, esfera, línea, plano, par de puntos, círculo) puede ser realizada multiplicando la entidad por \mathbf{R} por la izquierda y por su reverso $\tilde{\mathbf{R}}$ por la derecha. Por ejemplo, la rotación de un punto puede ser escrita como

$$\underline{x}' = \mathbf{R}\underline{x}\tilde{\mathbf{R}} \quad (2.78)$$

Un rotor representa el grupo de $\text{SO}(3)$ en el álgebra geométrica Euclidiana. Como el rotor es un versor en \mathcal{G}_n , pertenece al grupo de versores, entonces el producto geométrico de dos rotores $\mathbf{R} = \mathbf{R}_1\mathbf{R}_2$ resulta en un nuevo rotor. De esto se sigue que

$$\underline{x}' = \mathbf{R}\underline{x}\tilde{\mathbf{R}} = (\mathbf{R}_2\mathbf{R}_1)\underline{x}(\tilde{\mathbf{R}}_1\tilde{\mathbf{R}}_2) \quad (2.79)$$

Para trasladar una entidad respecto a un vector de traslación $\mathbf{t} \in \langle \mathcal{G}_3 \rangle_1$, usamos el *trasladador*, $\mathbf{T} \in \mathcal{G}_{4,1}$,

$$\mathbf{T} = \exp\left(\frac{\mathbf{e}\mathbf{t}}{2}\right) = \left(1 + \frac{\mathbf{e}\mathbf{t}}{2}\right) \quad (2.80)$$

utilizando (2.11) y la restricción $\mathbf{e}^k = 0$ para $k \geq 2$. Un trasladador es un rotor especial actuando en el infinito usando el vector nulo \mathbf{e} . Similar a la rotación, una

entidad puede ser trasladada multiplicando la entidad por \mathbf{T} por la izquierda y por su reverso $\tilde{\mathbf{T}}$ por la derecha.

$$\underline{x}' = \mathbf{T}\underline{x}\tilde{\mathbf{T}} \quad (2.81)$$

Para expresar un movimiento rígido, la aplicación consecutiva de un rotor y un trasladador pueden ser escritos como el producto de ellos. Tal operador es expresado como \mathbf{M} ,

$$\mathbf{M} = \mathbf{TR} \quad (2.82)$$

Es un multivector especial de grado par llamado *motor*, el cual es la abreviación de *momento y vector*. El movimiento rígido por ejemplo de un punto \underline{x} puede ser escrito como

$$\underline{x}' = \mathbf{M}\underline{x}\tilde{\mathbf{M}} = \mathbf{TR}\underline{x}\tilde{\mathbf{R}}\tilde{\mathbf{T}} \quad (2.83)$$

Esta formalización del desplazamiento rígido puede ser utilizada para cualquier entidad en el álgebra conformal (mencionadas en el cuadro 2.1). Esta representación es en contraste para formar una definición de motores en el marco del álgebra de motores [2], el álgebra $\mathcal{G}_{3,0,1}^+$ la cual formula la Cinemática en un espacio compuesto por líneas y el cual es isométrico al álgebra de cuaterniones duales.

Transformaciones de Giro y Tornillo

Siguiendo la definición del motor en el álgebra conformal y basados en el rotor llamado de *giro*, todo movimiento rígido puede ser expresado como un giro o movimiento de tornillo, el cual es una rotación sobre una línea en el espacio (en general, que no pase por el origen)⁹ combinado con una traslación a lo largo

⁹Tal operador también es llamado *rotación general*

de esta línea. En el álgebra geométrica conformal es posible usar los rotores y trasladadores para expresar movimientos de tornillo en el espacio.

Para modelar la rotación de un punto \underline{x} alrededor de una línea \underline{L} arbitraria en el espacio, la idea general es trasladar el punto \underline{x} con el vector de distancia entre la línea \underline{L} y el origen, para realizar la rotación en el origen y trasladar de regreso el punto transformado. Entonces el motor $\mathbf{M} \in \mathcal{G}_{4,1}^+$ describe una rotación general que tiene la forma

$$\mathbf{M} = \mathbf{TR}\tilde{\mathbf{T}} = \mathbf{R}_{\underline{L}} \quad (2.84)$$

donde se puede ver la traslación, la rotación y la traslación de regreso respectivamente. (2.84) se puede expresar como un rotor sobre la línea \underline{L} . Al aplicar $\mathbf{M} = \mathbf{RL}$ a un punto obtenemos

$$\underline{x}' = \mathbf{R}_{\underline{L}}\underline{x}\tilde{\mathbf{R}}_{\underline{L}} = \mathbf{TR}\tilde{\mathbf{T}}\underline{x}\tilde{\mathbf{T}}\tilde{\mathbf{R}} \quad (2.85)$$

donde $\mathbf{TR}\tilde{\mathbf{T}} = \tilde{\mathbf{R}}_{\underline{L}}$. Usando la forma exponencial del trasladador y rotor obtenemos¹⁰

$$\begin{aligned} \mathbf{R}_{\underline{L}} &= \mathbf{TR}\tilde{\mathbf{T}} \\ &= \exp\left(\frac{\mathbf{e}\mathbf{t}}{2}\right) \exp\left(-\frac{\theta}{2}\underline{l}\right) \exp\left(-\frac{\mathbf{e}\mathbf{t}}{2}\right) \\ &= \left(1 + \frac{\mathbf{e}\mathbf{t}}{2}\right) \exp\left(-\frac{\theta}{2}\underline{l}\right) \left(1 - \frac{\mathbf{e}\mathbf{t}}{2}\right) \\ &= \exp\left(\left(1 + \frac{\mathbf{e}\mathbf{t}}{2}\right) \left(-\frac{\theta}{2}\underline{l}\right) \left(1 - \frac{\mathbf{e}\mathbf{t}}{2}\right)\right) \\ &= \exp\left(-\frac{\theta}{2}(\underline{l} + \mathbf{e}(\mathbf{t} \cdot \underline{l}))\right) \end{aligned} \quad (2.86)$$

Esta es la fórmula correspondiente a la rotación general dada en [23]. La representación exponencial de los motores, rotores y trasladadores es usada para un manejo sencillo de ellos y sus derivaciones.

¹⁰En el paso 4 usamos la propiedad $\mathbf{gexp}(\xi)\tilde{\mathbf{g}} = \exp(\mathbf{g}\xi\tilde{\mathbf{g}})$ para $\mathbf{g}\tilde{\mathbf{g}} = 1$.

Es interesante mencionar que la parte exponencial de $R_{\underline{L}}$ consiste directamente en los componentes de la línea para rotar las entidades alrededor de ella. Retomando (2.61) y usando una dirección unitaria $|\mathbf{r}| = 1$ y como segundo punto la representación de la dirección \mathbf{t} de traslación del origen a la línea, podemos representar una línea como

$$\underline{L}^* = \mathbf{e}(\mathbf{t} \wedge \mathbf{r}) + r\mathbf{E} \quad (2.87)$$

la cual en su forma estándar es expresada como:

$$\begin{aligned} \underline{L} &= \mathbf{e}(\mathbf{t} \wedge \mathbf{r})\mathbf{I}_{\mathbf{E}} + r\mathbf{I}_{\mathbf{E}} \\ &= \mathbf{e}(\mathbf{t} \cdot \underline{l}) + \underline{l} \end{aligned} \quad (2.88)$$

Así, dada una línea en álgebra conformal que tenga magnitud unitaria, y usando su forma estándar \underline{L} , se puede hacer que cualquier entidad gire alrededor de ella usando esta línea como eje de rotación. Esta expresión se puede representar como

$$R_{\underline{L}} = \exp\left(-\frac{\theta}{2}\underline{L}\right) \quad (2.89)$$

Los movimientos en tornillo pueden ser representados por movimientos rígidos. Un *movimiento de tornillo* es aquel que para cada movimiento rígido de cuerpos puede ser realizado por una rotación en un eje combinado con la traslación paralela en ese eje. Para modelar el movimiento de tornillo, la entidad tiene que trasladarse durante una rotación general respecto al eje de rotación.

El motor resultante puede ser calculado utilizando (2.82) con (2.86) de la siguiente manera

$$\begin{aligned} \mathbf{M} = \mathbf{T}_{\mathbf{d}\mathbf{r}}R_{\underline{L}} &= \mathbf{T}_{\mathbf{d}\mathbf{r}}\mathbf{TR}\tilde{\mathbf{T}} \\ &= \exp\left(\frac{\mathbf{e}\mathbf{d}\mathbf{r}}{2}\right)\exp\left(-\frac{\theta}{2}(\underline{l} + \mathbf{e}(\mathbf{t} \cdot \underline{l}))\right) \end{aligned}$$

$$\begin{aligned}
&= \exp\left(\frac{\mathbf{e}d\mathbf{r}}{2} - \frac{\theta}{2}(\underline{l} + \mathbf{e}(\mathbf{t} \cdot \underline{l}))\right) \\
&= \exp\left(-\frac{\theta}{2}\left(\underline{l} + \underbrace{\mathbf{e}(\mathbf{t} \cdot \underline{l} - \frac{d}{\theta}\mathbf{r})}_{\mathbf{m}}\right)\right) \\
&= \exp\left(-\frac{\theta}{2}(\underline{l} + \mathbf{e}\mathbf{m})\right) \tag{2.90}
\end{aligned}$$

El bivector en la parte exponencial, $-\frac{\theta}{2}(\underline{l} + \mathbf{e}\mathbf{m})$ es la representación de un tornillo. Si \mathbf{m} es cero, el motor \mathbf{M} actúa como un rotor y si \underline{l} es cero, el motor actúa como un trasladador. Para $\mathbf{m} \perp \underline{l}^*$, el motor actúa como una rotación general y para $\mathbf{m} \not\perp \underline{l}^*$, el motor actúa como un movimiento de tornillo.

Capítulo 3

Reconocimiento de Objetos y Navegación

3.1. Introducción

El aprendizaje, como la inteligencia, cubre una amplia gama de procesos que son difíciles de definir de manera precisa. Una definición de aprendizaje en el diccionario incluye frases tales como “*para adquirir conocimientos o la comprensión de, o habilidad por el estudio, la enseñanza o la experiencia*”, y “*modificación de una tendencia de comportamiento por la experiencia*” Zoólogos y psicólogos estudian el aprendizaje en animales y seres humanos. En este capítulo nos centramos en el aprendizaje utilizando máquinas. Este aprendizaje nos ayudará en la identificación de ciertos objetos que nos serán útiles en la creación de un mapa.

Un mojón (en inglés *landmark*) que de aquí en adelante le llamaremos *marca* en el espacio o *marca espacial*, literalmente es una característica geográfica usada por los exploradores y otros para encontrar su camino de regreso o moverse por un área determinada. En el proceso de creación de mapas para un robot móvil

se pueden utilizar las marcas espaciales para recordar el lugar donde previamente se ha estado, mientras se explora un ambiente desconocido. Además las marcas espaciales se pueden identificar de manera automática mediante una función de costo y pueden ser utilizadas para encontrar la ubicación del robot en un mapa previamente capturado [8], con esto se facilita el trabajo de relocalización. Como se esta usando un sistema de cámaras estéreo, la posición 3D de cualquier objeto puede ser calculada y este puede ser representado en un ambiente virtual tridimensional. Utilizando estos objetos encontrados visualmente por el robot, se puede obtener la ubicación relativa respecto al robot. Un esquema de todo el procedimiento se ilustra en el figura 3.1:

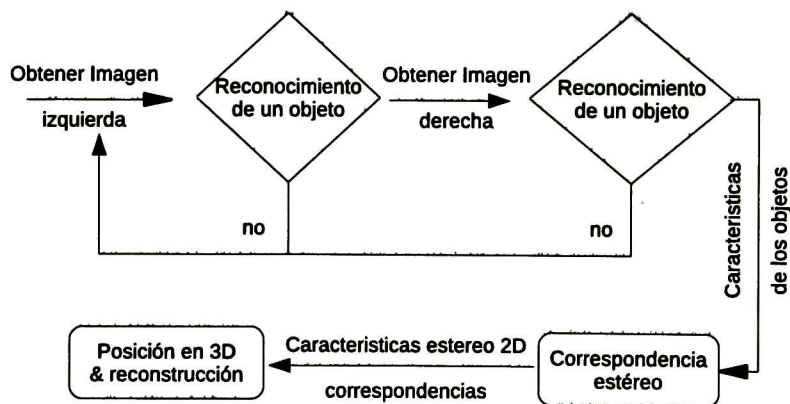


Figura 3.1: Esquema para el reconocimiento de objetos y su representación en un ambiente virtual 3D.

3.2. Reconocimiento de Objetos

Una marca espacial artificial o natural ubicada en el entorno actual, ayuda al robot móvil para conocer su ubicación dentro del mapa. Viola y Jones [36] presentan una propuesta nueva y muy rápida para la detección de rostros

de personas [35], basadas en aprendizaje automático por computadora usando el algoritmo AdaBoost y entrenamientos del tipo Haar [37]. Esta propuesta puede ser usada también para la detección de objetos estáticos las cuales utilizaremos como marcas espaciales.

Un proceso de reconocimiento puede ser mucho más eficaz si se basa en la detección de las características que codifican alguna información acerca de la clase para ser detectado. Este es el caso de las características del tipo Haar que codifican la existencia de contrastes entre las regiones orientadas en una imagen. Un conjunto de estas características puede ser utilizado para codificar los contrastes exhibida por algún objeto y sus relaciones espaciales. Las características del tipo Haar se llaman así porque se calcula de manera similar a los coeficientes en hondaletas (*wavelets*) de la transformada Haar.

3.2.1. Características Tipo Haar

Para realizar la extracción de las características tipo Haar, se necesita calcular la integral de la imagen [11]. Con este proceso se incrementa significativamente su extracción para su análisis y detección de objetos. El concepto de integral de una imagen es muy simple, y se puede definir como una tabla de búsqueda de 2 dimensiones en la forma de una matriz con el mismo tamaño que la imagen original. Cada elemento de la imagen integral contiene la suma de todos los píxeles de la región superior izquierda de la imagen original (en relación a la posición de los elementos). Esto permite el cálculo de una suma de área rectangular S en la imagen, en cualquier posición o escala, usando solo cuatro búsquedas ($S = A - B - C + D$) como se muestra en la figura 3.2, en donde A, B, C, D son puntos de la integral de la imagen.

Las características consisten en cajas de diferentes tamaños y localizacio-

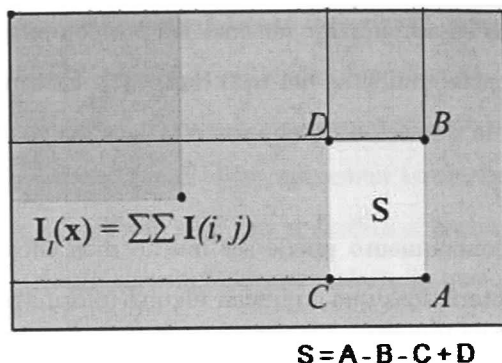


Figura 3.2: Cálculo de la integral de una imagen y su área de interés.

nes [29]. Considere algunos rectángulos de 20x20 píxeles (ver figura 3.3); por ejemplo, se pueden poner dentro de estos: dos rectángulos de 10x20 o cuatro rectángulos con un tamaño de 10x10, etc. La elaboración de ese conjunto más completo de características es una tarea a realizar, donde se deben de calcular la configuración de todas las combinaciones posibles, a su vez. Así, una base de características rectangulares de 20x20, proyectan la imagen a ese conjunto. Teniendo en cuenta que se tiene el cálculo de la imagen integral, donde este paso tiene una cantidad infinitesimal de tiempo en su cálculo. Para una característica que consiste en dos rectángulos de tamaño 10x20, se tiene que calcular la suma de todos los píxeles en los rectángulos de 10x20 como se señaló en anteriormente, así se tiene solo $4 * 2 = 8$ referencias, en lugar de una multiplicación matricial ordinaria de punto flotante que toma $2 * 20 * 20 = 800$ operaciones.

3.2.2. Clasificadores en Cascada

Para acelerar la detección de objetos, se usan los clasificadores en cascada de redes neuronales artificiales, ya que producen una frontera de separación no

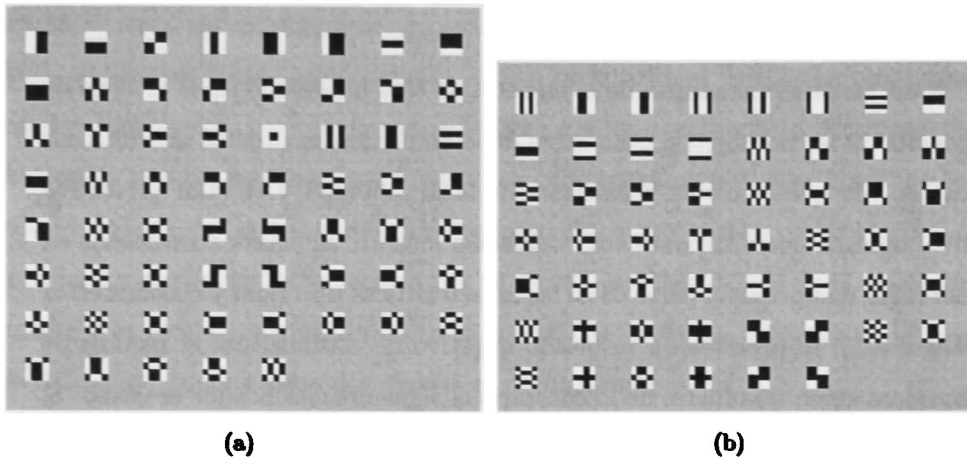


Figura 3.3: Representación de características tipo Haar en cuadros de 20x20 píxeles.

lineal y es posible controlar el tamaño de la red, el número de neuronas ocultas, y el rendimiento de clasificación durante la validación cruzada. Un clasificador (es decir, una cascada de clasificadores trabajando con características similares a Haar) se entrena con unos cientos de puntos de vista de la muestra de un determinado objeto (es decir, una cara o un coche, etc.), llamados ejemplos positivos, que se adaptan a la del mismo tamaño (es decir, una imagen de tamaño de 20x20 píxeles), y ejemplos negativos (no-objetos) - imágenes arbitraria del mismo tamaño.

Después de que un clasificador es entrenado, puede ser aplicado a una región de interés (del mismo tamaño que el utilizado durante el entrenamiento) en una imagen de entrada. La palabra *en cascada* en el nombre del clasificador significa que el clasificador resultante se compone de varios clasificadores simples (etapas) que se aplican posteriormente a una región de interés hasta que en algún momento el candidato es rechazado. Cuando no es rechazado quiere decir que ha pasado

todas las etapas del clasificador.

Para las primeras etapas de la cascada, se está interesado en el rechazo de la mayoría de los no-objetos, pero no se debe perder los objetos que estamos buscando. Así que la media geométrica de la sensibilidad y el valor predictivo positivo tendrán que ajustarse como validación métrica. El clasificador calcula en primer lugar una pequeña parte de las características tipo Haar y clasifica estos rectángulos. Si se produce un resultado negativo de clasificación, el rectángulo se considera como no-objeto, de lo contrario el siguiente clasificador es usado. Si el siguiente clasificador ofrece una decisión negativa, el rectángulo es clasificado como un no-objeto, y otra vez, en caso de una decisión positiva, el rectángulo es pasado a la siguiente etapa. Entonces, para que el objeto sea reconocido como válido todos los clasificadores deben producir una decisión positiva.

Para reducir el cálculo de los objetos que ya fueron clasificados en etapas pasadas, se utilizan las características ya clasificadas para su uso en etapas posteriores. Así para las siguientes etapas, se reducen las características a calcular.

El clasificador regresa “1” si la región puede mostrar el objeto (es decir, una cara, un coche u otro objeto), y “0” de otra manera. Para buscar el objeto en una imagen de conjunto se puede mover la ventana de búsqueda a través de la imagen y comprobar cada lugar utilizando el clasificador. Este clasificador está diseñado para que pueda ser fácilmente “redimensionado” a fin de poder encontrar los objetos de interés en diferentes tamaños, y lo cual es más eficiente que cambiar el tamaño de la imagen. Por lo tanto, para encontrar un objeto de tamaño desconocido en la imagen el procedimiento debe hacerse varias veces en diferentes escalas.

La palabra “aumentar” (en inglés *boosted*) significa que los clasificadores en cada etapa de la cascada son complejas y que ellos mismos se construyen a

partir de los clasificadores base utilizando una de las cuatro diferentes técnicas (voto ponderado), las cuales son Adaboost Discreto, Real Adaboost, Adaboost suave y Logitboost. Los clasificadores básicos son los clasificadores de árboles de decisiones con un mínimo de 2 hojas. Las características del tipo Haar son la entrada básica de las funciones clasificadoras. La característica utilizada en un clasificador particular, se especifica por su forma, posición dentro de la región de interés y la escala (esta escala no es la misma que la escala utilizada en la fase de detección, aunque estas dos escalas se multiplican).

3.3. Estimación de la Ubicación de Objetos Dentro del Ambiente Virtual

Cuando se habla de estimar la ubicación de un objeto dentro del ambiente virtual, no se trata de estimar la posición y orientación del objeto mismo, sino la representación del lugar que ocupa el objeto en el espacio real y su representación en el espacio virtual.

Una vez que las marcas espaciales hayan sido seleccionadas y entrenadas, el robot móvil puede utilizarlos a la hora de estar explorando un área, reconociendo estos, gracias a la aplicación del algoritmo de Viola y Jones.

La búsqueda del objeto, en este caso la marca espacial, se realiza solo en una imagen del sistema estéreo. Si una marca espacial es encontrada en la imagen, entonces se obtiene una sub-imagen I_L de la cámara izquierda. Esta I_L es la región de la imagen donde la marca espacial fue detectada (fig. 3.4).

Cuando una marca espacial es identificada en una imagen (cámara izquierda), se debe de asegurar que el mismo objeto esta siendo detectado por la otra imagen del sistema estéreo (cámara derecha).

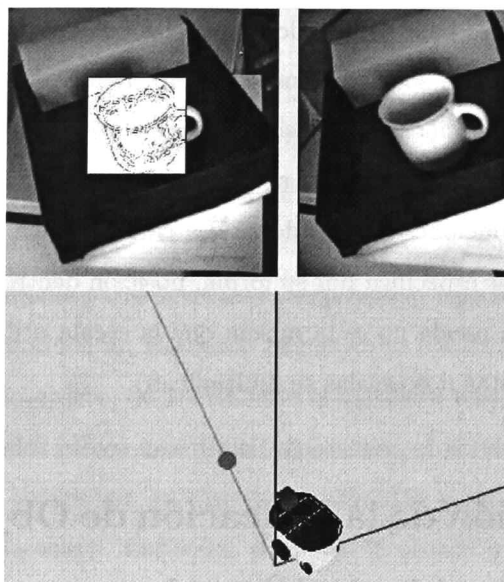


Figura 3.4: Identificando un objeto y obteniendo su ubicación en el ambiente virtual tridimensional

Una vez identificados los objetos en ambas imágenes, se procede a calcular puntos de interés sobre el objeto. Estos puntos nos ayudan a obtener la posición en 3D del objeto. Una manera de caracterizar el objeto encontrado es: detectando los bordes en la imagen donde se encuentra el objeto. Para realizar esto se utiliza el algoritmo de detección de bordes de Canny [10]. La detección de bordes se efectúa sobre la sub-imagen I_L . Utilizando los puntos de los bordes del objeto como puntos de interés se realiza una correlación contra la imagen de la cámara izquierda.

Un número de algoritmos basados en la correlación intentan encontrar puntos de interés en los aplicar dicho método. De hecho, la normalización se consagra en la Correlación Cruzada Normalizada (*Normalized Cross Correlation* (NCC) en inglés) y en la Correlación Cruzada Normalizada con Media Cero (*Zero Mean Normalized Cross Correlation* (ZNCC) en inglés), permitiendo tolerancias a la

variación lineal de brillo en la imagen. Además, gracias a la substracción de la media local, el ZNCC provee una mejor robustez que la NCC [12], ya que esta tolera las variaciones de brillo uniforme.

Correspondencias de una sub-imagen son buscadas a lo largo de la línea epipolar utilizando el cálculo de la ZNCC sólo sobre un intervalo $[d_{min}, d_{max}]$ de supuestas disparidades [14] [1], donde d_{min} y d_{max} son el valor de las disparidades mínimas y máximas respectivamente. El término *disparidad* denota la distancia euclidiana de un punto en la línea epipolar hacia un punto en la otra imagen de la cámara. Un valor pequeño de disparidad representa una distancia larga desde la cámara, un valor grande una distancia pequeña (paralaje, en inglés *parallax*).

Cuando se encuentra la correspondencia de todos los puntos en las dos imágenes se procede a calcular la posición 3D de estos, utilizando la triangulación.

3.4. Representación de Marcas Espaciales

Debido a que se pueden escoger diferentes objetos que sirvan como guía en la navegación de interiores, se buscan objetos comunes que se pueden encontrar en casi cualquier oficina o lugar de trabajo como las señales preventivas e informativas. Una señal de salida de emergencia, extintor, señal de sismos, etc. Estas señales fueron escogidas para actuar como marcas espaciales dentro de nuestro mapa virtual 3D (fig. 3.5).

Al realizar la navegación y habiendo encontrado estas señales. Procedemos a calcular sus puntos en 3D. Una vez que se hallan calculado los puntos en 3D, integramos este conjunto de puntos para obtener su centro de gravedad. El valor obtenido como centro de gravedad lo utilizamos como centro de una esfera con la cual vamos a envolver la mayoría de los puntos pertenecientes a la marca espacial.

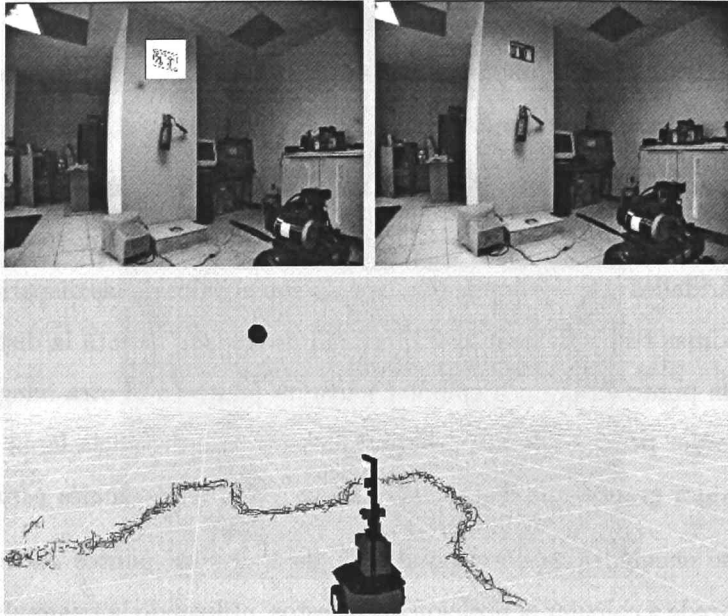


Figura 3.5: El robot móvil encontrando las marcas espaciales mientras navega en su ambiente. En la parte superior se muestra la vista estéreo (en la imagen izquierda se señala cómo la marca espacial es encontrada y encerrada en un rectángulo blanco) y su representación en el mapa 3D.

Esta esfera la representamos con el álgebra geométrica conformal como nos indica (2.52), la cual tiene la cualidad de que la mayoría los puntos identificados estarán dentro de la esfera, cumpliendo con la siguiente propiedad de la esfera.

Sean $\underline{x}, \underline{s} \in \mathcal{G}_{4,1}$ un punto y una esfera respectivamente. Para saber si \underline{x} esta dentro, sobre o fuera de la esfera utilizamos [5][23]

$$\underline{x} \cdot \underline{s} = \begin{cases} > 0 & \text{si } \underline{x} \text{ esta dentro de la esfera} \\ = 0 & \text{si } \underline{x} \text{ esta sobre la esfera} \\ < 0 & \text{si } \underline{x} \text{ esta fuera de la esfera} \end{cases} \quad (3.1)$$

Al final ya que se tiene el radio y centro de la esfera, esta se coloca en el

lugar indicado con respecto a la posición del robot. Un ejemplo de esto se ilustra en la figura 3.6.

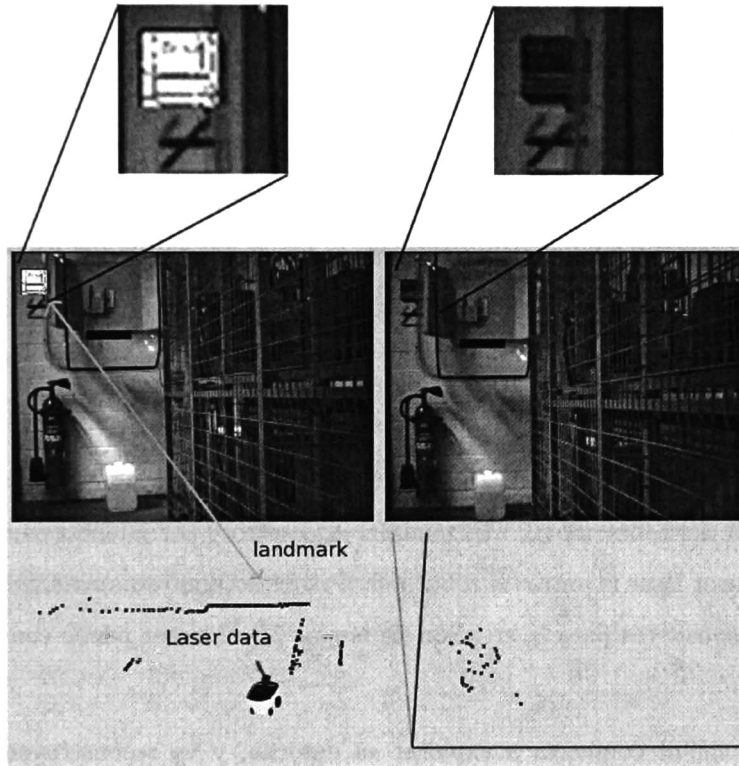


Figura 3.6: Robot móvil encontrando una marca espacial y colocando esta en la posición relativa al robot dentro del ambiente virtual.

Cada una de estas esferas es etiquetada con el tipo de objeto al que pertenece. Así cuando en la exploración volvamos a ver el algún otro objeto quedan tres opciones.

1. El objeto que estamos viendo es de una clase ya conocida, pero es una marca espacial que acabamos de descubrir y lo agregamos al ambiente virtual.

2. El objeto que estamos viendo es de una clase ya conocida, esta marca espacial ya la habíamos registrado con anterioridad, pero esta situado en otro lugar dentro del ambiente. En otras palabras es un objeto de las mismas características pero este es otro objeto no antes visto.
3. El objeto que estamos viendo es de una clase ya conocida, y es un objeto que ya habíamos capturado antes.

Dependiendo en la situación en la que nos encontremos se puede agregar la nueva marca espacial o simplemente rectificar la ubicación actual del robot.

3.5. Navegación

Para la navegación del robot móvil nos basamos en [6], la cual calcula el desplazamiento del robot móvil, utilizando el odómetro y corrigiendo mediante lecturas del sensor láser mientras el robot móvil va explorando su ambiente. Para la fusión de dispositivos para la creación de mapas 3D el lector puede consultar el apéndice A.

El robot móvil comienza a explorar su entorno, y va reconstruyendo el mismo con las lecturas obtenidas por el sensor de alcance láser. Este se ayuda generando un mapa a base de segmentos de línea las cuales son guardadas utilizando la entidad conformal llamada *par de puntos*.

Gracias al uso del álgebra geométrica conformal las entidades del mapa se pueden expresar de forma más compacta y utilizar los operadores de transformación para su traslación y rotación.

Los mapas se van guardando usando las primitivas geométricas como puntos, par de puntos, líneas y esferas.



Figura 3.7: Señales de emergencia como marcas espaciales, ubicados en el laboratorio.

La figura 3.7 muestra la identificación de las señales de emergencia de un laboratorio. En donde se pueden observar que hay una marca espacial repetida que es la número 1. Este demuestra que se puede identificar los objetos vistos desde varios ángulos. El robot móvil detecta el objeto y reconoce que es una marca espacial previamente capturada.

La figura 3.8 muestra la creación de un ambiente virtual por el cual el robot móvil está pasando y en el cual se observan las marcas espaciales ubicadas tridimensionalmente dentro del laboratorio. Asimismo se observa el mapa que se va creando con la ayuda del sensor de distancia láser.

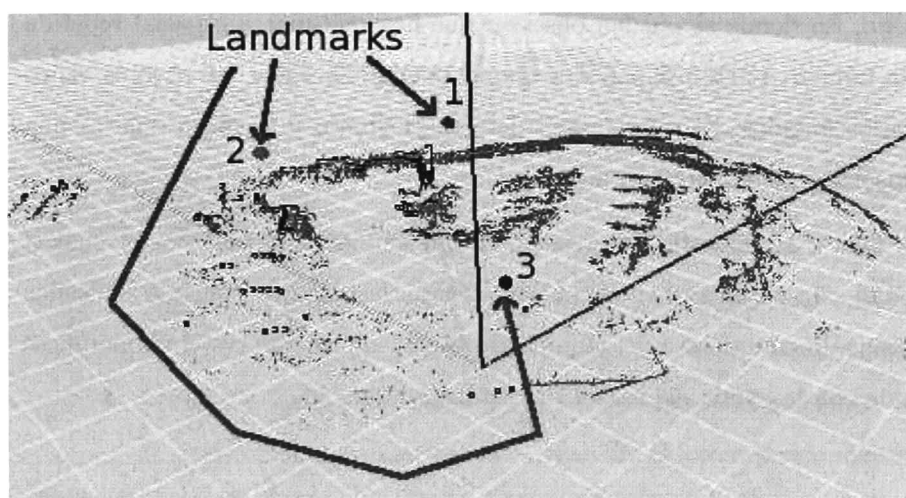


Figura 3.8: Representación de las marcas espaciales y las lecturas del sensor láser.

Capítulo 4

Representación y Reconocimiento de Objetos Usando la Transformada de Hough

La transformada de Hough [20] es un método efectivo para identificar la ubicación y orientación de líneas [13]. En el siguiente capítulo se describe el espacio de Hough en 2D y 3D en relación con el álgebra geométrica conformal.

4.1. La Transformada de Hough

4.1.1. Transformada de Hough (2D) Clásica

La transformada de Hough es la representación de una línea L de \mathbb{R}^2 en un punto en 2D en el espacio de coordenadas de Hough (θ, ρ) , donde $\rho \in \mathbb{R}$ y

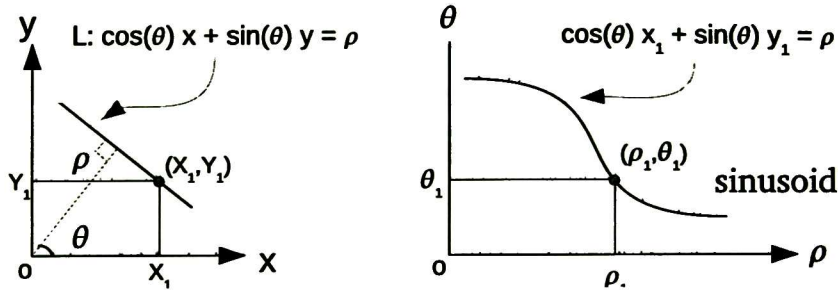


Figura 4.1: Líneas en \mathbb{R}^2 y su representación 2D en el espacio de parámetros.

$\theta \in [0, 2\pi)$. La ecuación que describe a una línea L en \mathbb{R}^2 está dada por

$$L : y = ax + b, \tag{4.1}$$

donde los escalares a, b son los parámetros de la línea. Esta línea L puede ser representada en el espacio de parámetros como un punto con las coordenadas (a, b) . Por otro lado, un punto $(x_i, y_i) \in \mathbb{R}^2$ puede ser representado en el espacio de parámetros como una línea L' de la siguiente forma

$$L' : b = -x_i a + y_i, \tag{4.2}$$

donde $x_i, y_i \in \mathbb{R}$ son los parámetros de la línea L' . Considerando a p_1, p_2, \dots, p_n , como un conjunto de puntos que yacen sobre L . Cada uno de estos puntos representa una línea en el espacio de parámetros. Así, la intersección de estas líneas en el espacio de parámetros corresponde a un punto en a_i, b_i , los cuales a su vez representan los parámetros de la línea L en \mathbb{R}^2 . Para el caso en donde x sea una constante en la ecuación (4.2), el eje y es igual a 0. Una manera de resolver este problema es replanteando la forma de la ecuación de la línea:

$$L : \cos(\theta)x + \sin(\theta)y = \rho, \quad (4.3)$$

donde θ y ρ son los parámetros de la normal hacia la línea que pasa por el origen de coordenadas en \mathbb{R}^2 (Fig. 4.1). De donde, un punto $(x_i, y_i) \in \mathbb{R}^2$ representa una senoidal en el espacio de parámetros (Fig. 4.1). Los puntos que yacen sobre la línea L $(x_i, y_i) \in \mathbb{R}^2$ los cuales representan cada uno de ellos una senoidal, se intersecan en un solo punto (θ, ρ) en el espacio de parámetros. Estos puntos representan la línea L : $\cos(\theta)x + \sin(\theta)y = \rho$.

Hasta aquí, se ha definido la función que lleva a una línea en \mathbb{R}^2 a un punto en el espacio de Hough. Sea $L \in \mathbb{R}^2$ como esta definida en (4.1), la función de la transformada de Hough puede ser definida como

$$HT(L) \rightarrow (\theta, \rho)$$

Una propiedad interesante [21], que será de utilidad en las próximas secciones, es la que hay en relación entre las líneas transformadas en puntos en el espacio de Hough y los movimientos (desplazamientos y rotaciones) que realiza el robot respecto a ellas.

Propiedad 1. Dada la transformación de un conjunto de líneas L al espacio de Hough, capturadas por el robot móvil, y su desplazamiento (traslación y rotación) (T_x, T_y, θ_R) (asumiendo $\theta_R \in [0, 2\pi)$). La transformada de Hough de las líneas L con respecto a la nueva pose del robot serán $HT(L) \rightarrow (\theta', \rho')$ tal que:

$$\text{si } 0 \leq \theta + \theta_R \leq 2\pi$$

$$\theta' = \theta + \theta_R$$

$$\rho' = \rho + T_x \cos(\theta + \theta_R) + T_y \sin(\theta + \theta_R)$$

$$\text{si } \theta + \theta_R > 2\pi$$

$$\theta' = \theta + \theta_R - 2\pi$$

$$\rho' = -(\rho + T_x \cos(\theta + \theta_R) + T_y \sin(\theta + \theta_R))$$

si $\theta + \theta_R < 0$

$$\theta' = \theta + \theta_R + 2\pi$$

$$\rho' = -(\rho + T_x \cos(\theta + \theta_R) + T_y \sin(\theta + \theta_R))$$

Es importante notar aquí que si $\theta_R = 0$ (ej. el robot móvil no rota) entonces $\theta' = \theta$ (θ no tiene cambios), y por el contrario si $T_x = T_y = 0$ (ej. el robot móvil no se desplaza) entonces $\rho' = \rho$ (ρ no sufre cambios). Por medio de la propiedad 1 se puede identificar el desplazamiento del robot en el espacio de Hough, el cual está supeditado a los parámetros T_x, T_y, θ_R .

4.1.2. La Transformada de Hough en 3D

Extendiendo el principio de la línea, se puede representar un plano embebido en \mathbb{R}^3 como un punto (a, b, c) en el espacio de parámetros en 3D.

$$\Pi : z = ax + by + c. \quad (4.4)$$

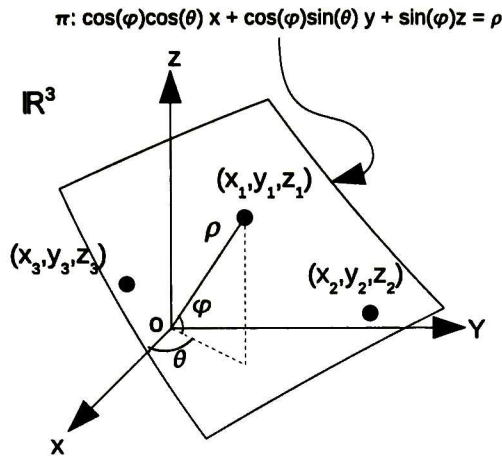
Igualmente, si la ecuación tiene la siguiente forma:

$$ax + by + cz + d = 0, \quad (4.5)$$

pero esta representación del plano no se puede expresar en el espacio de parámetros cuando $x = y = 0$. Por consiguiente se usará la forma normal del plano

$$\Pi : \cos(\theta)\cos(\phi)x + \sin(\theta)\cos(\phi)y + \sin(\phi)z = \rho, \quad (4.6)$$

donde θ, ϕ, ρ son los parámetros de la normal del plano pasando por el origen de coordenadas (Fig. 4.2). Aquí se aprecia como un punto $(x_i, y_i, z_i) \in \mathbb{R}^3$ representa

Figura 4.2: Planos en \mathbb{R}^3

una superficie senoidal en el espacio de parámetros (θ, ψ, ρ) . Los puntos (x_i, y_i, z_i) que yacen sobre el plano en \mathbb{R}^3 representan una superficie senoidal que se interseca en un solo punto $(\theta_i, \phi_i, \rho_i)$, el cual representa el plano original Π (ecuación (4.6)), y se ilustra en la figura 4.3.

4.1.3. La Transformada de Hough y el Álgebra Geométrica Conformal

La transformada de Hough de 2 dimensiones representa el espacio euclidiano \mathbb{R}^2 en el espacio paramétrico 2D de Hough con (θ, ρ) . Debido a que este trabajo se trata de entidades geométricas y para ello se utiliza el álgebra geométrica conformal, se debe de formular la transformada de Hough en términos que se pueda usar cómodamente con este marco de trabajo. Las líneas en 2D o los planos en 3D son de hecho hiperplanos embebidos en un espacio lineal 2D o 3D

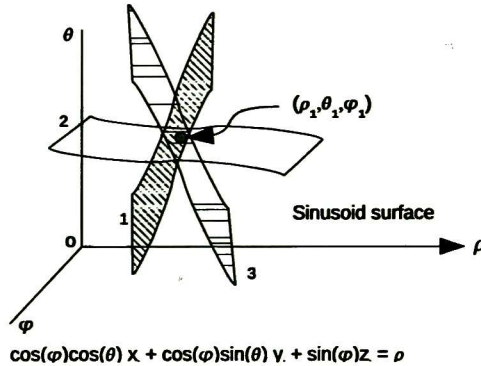


Figura 4.3: Planos en el espacio de parámetros 3D.

respectivamente. En consecuencia, si la transformada de Hough fue desarrollada para líneas, desde el punto de vista matemático, esta puede ser generalizada para cualquier hiperplano de \mathbb{R}^n , donde $n > 2$. Sin embargo, la manera en la que se representan estos hiperplanos es la clave; así, se puede formular la transformada de Hough en forma general para que sea compatible con el álgebra geométrica conformal. Hay solo dos parámetros intrínsecos caracterizando a estos hiperplanos: la distancia de Hesse (un escalar) y la orientación del hiperplano (un n -D vector unitario). Así, un hiperplano en \mathbb{R}^n tiene $n + 1$ grados de libertad (Degree of Freedom (DoF) en inglés) y está dada por la generalización de la ecuación (4.4):

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = \mathbf{a} \cdot \mathbf{x} = \mathbf{a}^T \mathbf{x} = -z, \quad (4.7)$$

donde $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$. Esta ecuación puede ser convenientemente reescrita usando coordenadas homogéneas y un escalar homogéneo $|\mathbf{a}|$ como sigue:

$$[\mathbf{x}, 1]^T[\mathbf{a}, z] = [\mathbf{x}, 1]^T[\mathbf{n}, h] = \mathbf{X}^T\Pi = 0. \quad (4.8)$$

Esto representa una restricción de un punto homogéneo $\mathbf{X} \in \mathbb{R}^{n+1}$ yaciendo en el plano $\Pi = [\mathbf{n}, h]^T$ donde $\mathbf{n} \in \mathbb{R}^n$ y $h \in \mathbb{R}$ son la normal y la distancia de Hesse respectivamente.

Cualquier espacio de Hough es un espacio parametrizado, el cual puede ser puesto en conexión con el álgebra geométrica conformal con el propósito de representar las entidades geométricas. Las cuales surgen usando los hiperplanos representados en el espacio de Hough. Esto significa que un espacio n -dimensional de Hough puede ser utilizado con una álgebra geométrica conformal $\mathcal{G}_{n+1,1}$. Este formalismo esta inspirado por el concepto de dualidad entre los hiperplanos y los puntos [16]. Ambos representados en coordenadas homogéneas. Y como en la ecuación (4.8) claramente se muestra, lo que se nombraría su representación dual dependiente de la relación geométrica que estos guardan; por ejemplo: el punto que yace sobre un plano.

Ahora, imagine una esfera centrada en \mathbb{R}^n conformada por varios hiperplanos de radio variable (un conjunto perfecto), Se puede considerar el radio como la distancia de Hesse en cualquier punto sobre la superficie de cualquier hiperplano, como un único hiperplano; véase la figura 4.4. Entonces, cualquier punto que yace en este hiperplano y que cumpla $\mathbf{P} \wedge \Pi = 0$ incremente un contador de este hiperplano. Este modelo es una extensión sencilla de la transformada de Hough.

Cada contador puede almacenar $(n - 1)$ coeficientes de la coordenada de cada punto unicamente para identificar en cualquier momento el punto con respecto al hiperplano, esto quiere decir 1 DoF (*Degree of Freedom*, grado de libertad en inglés) menos por cada punto. Para el espacio en 3D la ecuación (4.6) muestra

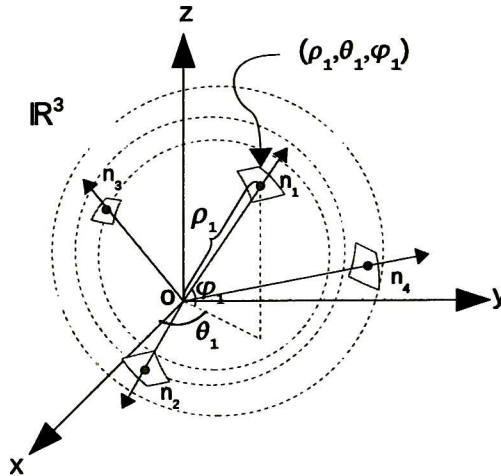


Figura 4.4: Planos en \mathbb{R}^3 usando coordenadas polares.

que se puede representar el plano en términos de 2 coordenadas polares y la distancia de Hesse (θ, ψ, ρ); ver Fig. 4.3 y 4.4.

Usando un algoritmo robusto como el RANSAC (RANDOM SAMPLE CONSENSUS) [33], se buscan los planos en la escena. El algoritmo RANSAC se describe a continuación:

1. Aleatoriamente seleccionar una muestra de p datos del conjunto P e inicializar el modelo $1 - p = (1 - w^n)^k$ con estos datos.
2. Determinar el conjunto de datos que se encuentran a una distancia umbral t del modelo. El conjunto P_i contiene los elementos que mejor se ajustan al modelo.
3. Si el tamaño de P_i es mayor que algún umbral T se vuelve a estimar el

4.2. REPRESENTACIÓN DE OBJETOS 3D EN EL DOMINIO DE HOUGH

modelo utilizando todos los datos del modelo y se terminan.

4. Si el tamaño de P_i es menor que N , se selecciona un nuevo subconjunto y se repite el paso anterior.
5. Después de N iteraciones el mayor conjunto P_i es elegido.

Una vez encontrados los planos se usan sus parámetros (la distancia de Hesse y su orientación) para representar un plano en 3D en el álgebra geométrica conformal $\mathcal{G}_{3+1,1}$ y así encontrar las restricciones entre los puntos, líneas y planos. Para hacer esto, se hace uso del álgebra de incidencia, usando los operadores de *join* y *meet*. El encontrar las restricciones geométricas es extremadamente útil para la navegación de un robot móvil en entornos con obstáculos en su paso. La sección 4.3 contiene interesantes aplicaciones de este enfoque.

4.2. Representación de Objetos 3D en el Dominio de Hough

De los puntos obtenidos del sistema de cámaras estéreo pueden ser utilizados para calcular una representación de un plano en el mundo virtual. Varias técnicas pueden ser utilizadas para obtener las características de los objetos desde las imágenes de las cámaras [31]. Los parámetros de los planos pueden ser calculados a partir de estas características (puntos) implementando el algoritmo de RANSAC [33]. Los parámetros del plano son: el radio o la distancia radial ρ (distancia de Hesse), el plano de inclinación (o ángulo polar) θ , y el azimut (o ángulo azimutal) φ . En el ambiente, la representación del plano está dispuesta por los puntos que yacen en él. Tres de estos puntos son suficientes para representar el plano. En el dominio de Hough, ρ , θ y φ son usados como puntos en

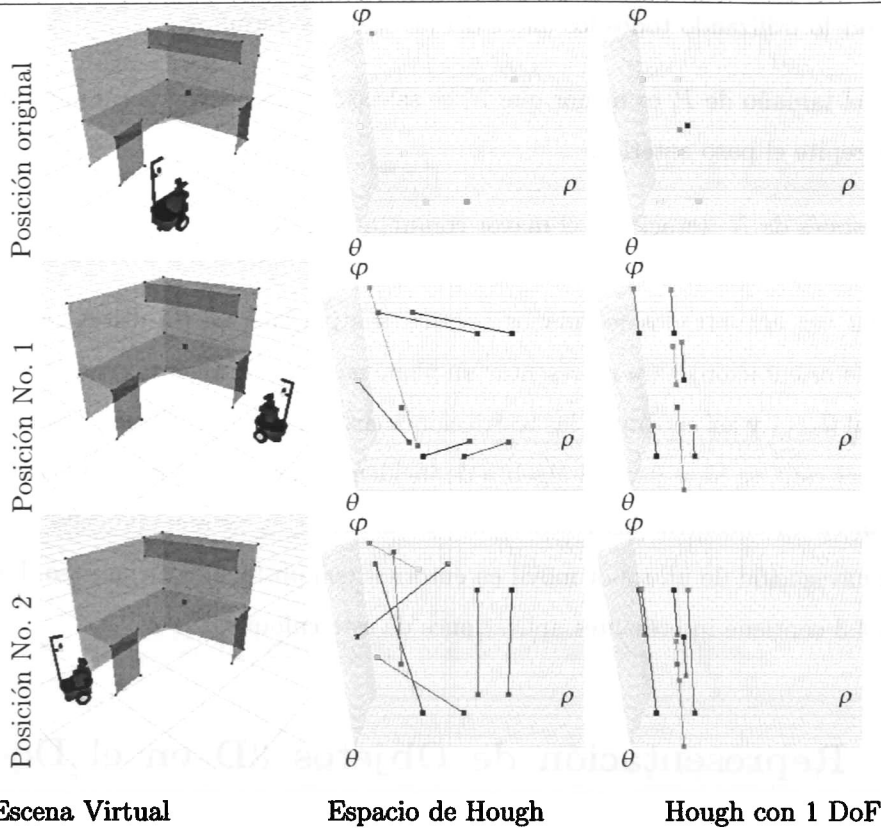


Figura 4.5: Un robot móvil viendo un lugar de trabajo, su representación en el espacio de Hough, y el espacio de Hough con 1 DoF.

este espacio, como se ilustra en la Fig. 4.5. Note que un plano es representado con solo 3 DoF. Usando estos puntos, es fácil manejar los planos gracias a sus valores correspondientes ($\rho > 0$, $\theta \in [0, \pi)$, y $\varphi \in (0, 2\pi]$). Además, cuando el robot móvil está en movimiento, los valores de los planos solo son afectados en los atributos de ρ y φ , solo 2 DoF. En las secciones siguientes se verá cómo reducir la representación de los objetos a 1 DoF cuando cambie de lugar el robot móvil.

4.2. REPRESENTACIÓN DE OBJETOS 3D EN EL DOMINIO DE HOUGH⁵⁹

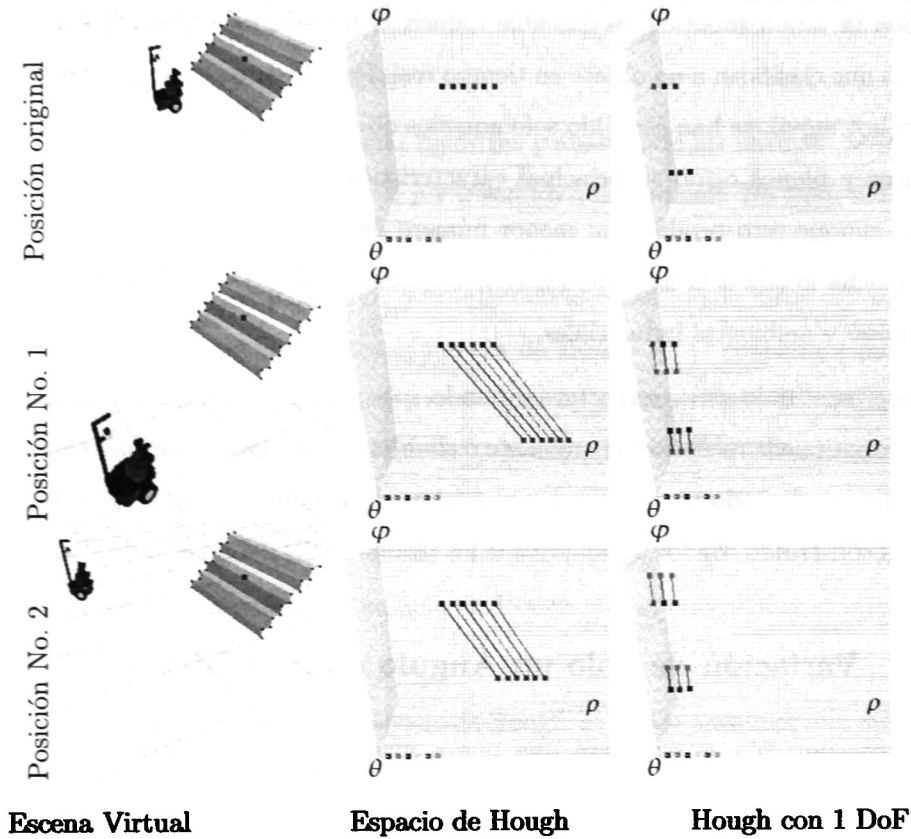


Figura 4.6: Una representación de unas escaleras y su transformada de Hough vista desde diferentes ángulos (en la tercera columna φ es la única afectada).

4.2.1. Percepción del Ambiente

Esta sección presenta la aplicación de la transformada de Hough en 3D para almacenar formas complejas en 3D. En lo que respecta a la complejidad de la representación del objeto, una pregunta natural surge: ¿Qué características del ambiente o de un objeto que los distinguen se deben de considerar relevantes?. Dependiendo del contexto, uno puede tomar en cuenta el color, la textura, los coeficientes de Fourier o de las ondaletas (wavelets en inglés), etc.

Con la visión robótica, es deseable reducir el tamaño del vector de características que clasifican a un objeto en tiempo real. Para la tarea de la navegación de un robot móvil, se han escogido solo aquellos objetos que sean representables con líneas y planos como su principal característica geométrica (ver Fig. 4.5). Incluso, aunque esto produce un menor número de características, ha probado ser de mucha ayuda a la hora de describir objetos complejos en 3D, de oficinas y de equipo y ambientes industriales.

Representando estas estructuras usando solo planos y guardando los mismos, el robot puede reconocer el ambiente o el objeto en cualquier otro momento. Esto se hace aplicando un clasificador al cual se le introduce el vector de características construido con los parámetros de los planos en 3D del espacio de Hough.

4.2.2. Variación de Solo un Ángulo en los Objetos

A continuación se explicará una nueva manera de como clasificar a los objetos representados en el espacio de Hough en 3D. Aquí se explica como reducir los grados de libertad (DoF) a una variable cuando el objeto es visto desde otro ángulo y así reducir la complejidad en la búsqueda de un objeto o ambiente dado.

Sea \mathcal{O} un objeto en 3D, el cual es representado con solo planos (n número de planos), $\mathcal{O} = (\Pi_1, \Pi_2, \dots, \Pi_n)$ donde $\Pi = (a, b, c, d)^T$ y sea $\mathcal{O}^{\mathcal{HT}}$ un vector en el espacio de Hough que contiene estos planos definidos como

$$\mathcal{O}^{\mathcal{HT}} = (\rho_1, \theta_1, \varphi_1, \rho_2, \theta_2, \varphi_2, \dots, \rho_n, \theta_n, \varphi_n)^T$$

La representación en el dominio de Hough es tomada desde la posición inicial del robot, en otras palabras se tiene que capturar el objeto previamente para obtener sus características (los planos) y de ahí se representa el objeto en el espacio de Hough. Si el robot móvil se mueve, se obtiene un nuevo $\mathcal{O}^{\mathcal{HT}'}$ debido al

4.2 REPRESENTACIÓN DE OBJETOS 3D EN EL DOMINIO DE HOUGH

desplazamiento del robot y por consecuencia el origen de coordenadas de donde se toma la referencia para el objeto.

Una vez que se tomaron varias muestras y analizando las mismas. Se puede observar que solo los valores de ρ y φ son los que cambian. De aquí, naturalmente se puede proponer una nueva representación de las características de los objetos; tomando los planos de \mathcal{O} , y obteniendo el centroide de estos (x_c, y_c, z_c) . Con este se puede calcular la transformada de Hough de \mathcal{O} respecto a su mismo centroide. Así realizar otra representación del mismo objeto con referencia al centroide como origen de coordenadas y guardarlo en \mathcal{O}_c^{HT} . Cuando el robot móvil se mueva a otro lugar y vea el objeto desde otro ángulo o distancia, un nuevo vector es obtenido $\mathcal{O}_c^{HT'}$, donde los valores de $\varphi_i \in \mathcal{O}_c^{HT'}$, $i = 1 \dots n$, solo tienen un desplazamiento en el espacio de Hough que se encuentra en el intervalo de $[0, 2\pi)$ (ver pseudo-código 1).

Usando la *Propiedad 1* del espacio de Hough, se puede proponer una métrica para la diferencia de un ángulo en el dominio de Hough [7]. Esta se define como:

$$\Delta\theta(\theta_a, \theta_b) = \begin{cases} \theta_a - \theta_b - 2\pi & \text{caso 1} \\ \theta_a - \theta_b + 2\pi & \text{caso 2} \\ \theta_a - \theta_b & \text{otro} \end{cases} \quad (4.9)$$

donde el caso 1 ocurre si $\theta_a > \theta_b$ y $\theta_a + \theta_\xi \geq 2\pi$ y $\theta_b + \theta_\xi \leq 0$; el caso 2: si $\theta_a < \theta_b$ y $\theta_b + \theta_\xi \geq 2\pi$ y $\theta_a + \theta_\xi \leq 0$. Este resultado da un cálculo de cualquier punto cercano a otro, donde sus ángulos estén cercanos a $0 = 2\pi$.

Se usa (4.9) para obtener $\Delta\varphi$, con la particularidad de que este valor es igual para todos los φ_i . Para clasificar diferentes objetos, ordenamos el vector \mathcal{O}_c^{HT} tomando en cuenta como comparación el valor de ρ de cada plano. Así tendremos los planos ordenados desde el más cercano al origen (el centroide) hasta el más lejano. Una vez ordenado este vector, se compara con cualquier otro ob-

jeto obtenido anteriormente, comparando los valores de ρ y θ para asegurarse de que el plano sea el mismo, y de ahí encontrando el desplazamiento de $\Delta\varphi$ de estos. Si no se encuentran similitudes en los planos el objeto no es encontrado (ver pseudo-código 2).

Pseudocódigo 1 Captura de nuevos Objetos

Salida: vector \mathcal{O}_c^{HT}

- 1: Visualizar el objeto con el sistema de cámaras estéreo.
 - 2: Obtener sus puntos característicos y calcular los planos usando RANSAC.
 - 3: Guardar los planos en \mathcal{O}_i .
 - 4: Obtener el centroide de \mathcal{O}_i y guardarlo en $\mathcal{C}_{\mathcal{O}}^i$
 - 5: Trasladar el origen de coordenadas a $\mathcal{C}_{\mathcal{O}}^i$
 - 6: para todo plano $\in \mathcal{O}_i$ hacer
 - 7: Calcular su representación en el espacio de Hough 3D.
 - 8: Agregar esta representación en \mathcal{O}_c^{HT}
 - 9: fin para
-

En las figuras 4.5, 4.6, y 4.7 tres objetos simulados (cubículo de estudiante, escaleras, entrada a un salón) son ilustrados. La primera fila muestra la posición inicial de donde el objeto fue tomado. La segunda y la tercera fila muestran los objetos tomados desde diferentes posiciones (ángulos y distancia). La primera columna muestra la posición de el robot móvil con respecto al objeto. La segunda columna ilustra el espacio de Hough donde esta representado este objeto tomado desde la posición inicial del robot, y muestra el desplazamiento (por una línea) de la nueva (actual) representación (en color rosa) del objeto en la nueva posición.

La tercera columna ilustra la representación del objeto en el espacio de Hough calculada desde el centroide y se muestra su desplazamiento (por una línea) de la nueva representación en color azul. Como se puede ver en la tercera

4.2. REPRESENTACIÓN DE OBJETOS 3D EN EL DOMINO DE HOUGH3

Pseudocódigo 2 Clasificador de objetos

Entrada: vector \mathcal{VO}_c^{HT}

Salida: posición del objeto encontrado o -1 en caso de fracaso

- 1: Obtener el objeto con el algoritmo 1.
 - 2: para todo objeto $\in \mathcal{VO}_c^{HT}$ hacer
 - 3: Ordenar el de forma ascendente los elementos de \mathcal{O}_c^{HT} usando como referencia el valor de ρ .
 - 4: fin para
 - 5: para todo objeto $\in \mathcal{VO}_c^{HT}$ hacer
 - 6: para todo plano $\in \mathcal{O}_c^{HT}$ hacer
 - 7: Comparar los valores de ρ y θ .
 - 8: Calcular $\Delta\varphi$, con (4.9)
 - 9: si los valores son consistentes de ρ , θ y φ entonces
 - 10: se encontró una correspondencia y se devuelve la posición
 - 11: fin si
 - 12: fin para
 - 13: fin para
 - 14: Si no se encontró correspondencia se devuelve -1
-

columna, solo el valor de φ es afectado (es el único que varia) por un desplazamiento lineal. Esto conlleva a una sencilla clasificación y búsqueda de estos objetos o entornos porque estos solo tienen 1 DoF.

4.3. Implementación y Resultados Experimentales

En los siguientes experimentos, se puede apreciar lo sencillo y descriptivo que son los planos para representar escenarios intramuros almacenados en el espacio paramétrico 3D de Hough. En la Fig. 4.8a, se puede ver a un robot móvil dentro de un ambiente en el cual hay varios planos alrededor de él. Estos planos son detectados gracias a los puntos (esquinas) [31] encontradas en las imágenes que el sistema estéreo registra. En la primera fila se muestra al robot móvil pasando a través de un corredor donde se encuentran cubículos de estudiante. Los cuales tienen una estructura que pueden ser representados por varios planos. Estos planos pueden ser representados como puntos en el espacio 3D de Hough (ver Fig. 4.8d) y pueden ser guardados para su posterior clasificación. En la figura 4.8b se muestran los planos que el robot detectó y estos son representados en un ambiente virtual. Como se puede ver, los planos censados están colocados de acuerdo a su posición espacial respecto a la localización de donde el sistema de cámaras estéreo registró los escenarios. De igual manera el escenario de la segunda fila de la figura 4.8, una esquina del laboratorio, puede ser registrada con algunos planos característicos y estos ser guardados para su reconocimiento futuro.

Todos estos planos tienen su representación en álgebra geométrica conformal, y en el espacio 3D de Hough (ver Fig. 4.8d). Note que se puede pasar de

un plano en álgebra geométrica conformal a un plano en la representación del espacio en 3D de Hough (ver Fig. 4.8c). Además esta clase de representación de escenarios facilita el reconocimiento de ambientes específicos usando la búsqueda y clasificación de objetos en el espacio 3D de Hough.

4.3.1. Percepción de Formas en 3D

Las formas que se detectan usando la transformada de Hough en 3D, pueden tener estructuras complejas las cuales se pueden representar de manera simple en este espacio. Para ello se sacan las esquinas $\mathcal{C} = (x_c, y_c)^T$ de los planos $(\Pi_i, i = 1 \dots n)$, donde n son el número de planos que representan a estos objetos. Con ellos se calculan para cada $\mathcal{C} \in \Pi_i$ la representación del plano característico y los guardamos en el vector \mathcal{O} . Donde este vector guarda los planos que representan al objeto en el espacio cartesiano \mathbb{R}^3 .

En la figura 4.9, el robot observa tres diferentes estructuras las cuales son codificadas usando planos. Los primeros dos objetos (a) y (b) son escogidos para ilustrar el rol que tienen los planos en la descripción de estructuras geométricas puras. El tercero es una clase de estructura metálica que es un poco mas compleja por su variedad en los detalles. Sin embargo, tomando en consideración los planos más relevantes, es posible representar estas estructuras complejas de una manera más económica. Se pueden ver resultados interesantes dentro del espacio paramétrico de Hough ilustrados en la figura 4.9c. Los agrupamientos de puntos de estas tres estructuras son notablemente diferentes. De nuevo, esto facilita la tarea de reconocimiento usando un clasificador. Uno se puede imaginar a un robot móvil explorando ambientes donde estructuras complejas están presentes, en donde este enfoque presentado puede ser de gran utilidad a la hora de codificar y reconocer objetos. Representando objetos en 3D con líneas y planos

*CAPÍTULO 4. REPRESENTACIÓN Y RECONOCIMIENTO DE OBJETOS
66 USANDO LA TRANSFORMADA DE HOUGH*

es una propuesta de compresión de datos en vez de usar una representación de agrupamientos de puntos. En este sentido, este enfoque está inspirado por el uso de ondaletas (*wavelets*) en el procesamiento de señales 1D e imágenes en 2D, las cuales son utilizadas en compresión de datos.

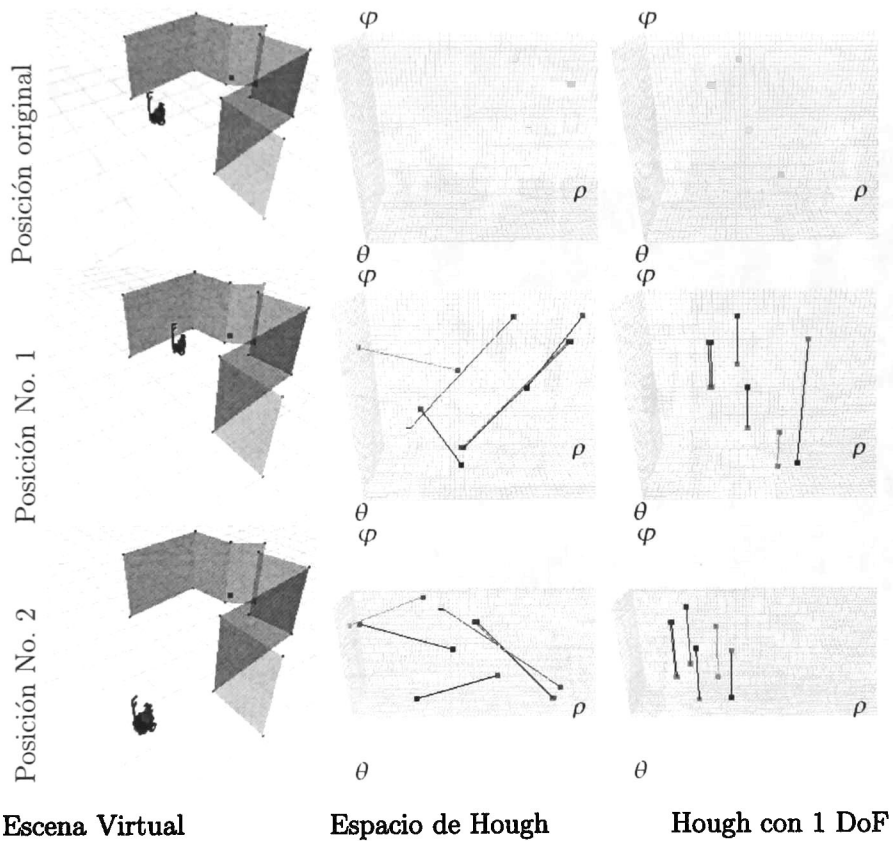


Figura 4.7: Una representación de una entrada a un salón y su representación en el espacio de Hough vista desde diferentes ángulos (en la columna 3 se muestra que solo se afecta el valor de φ).

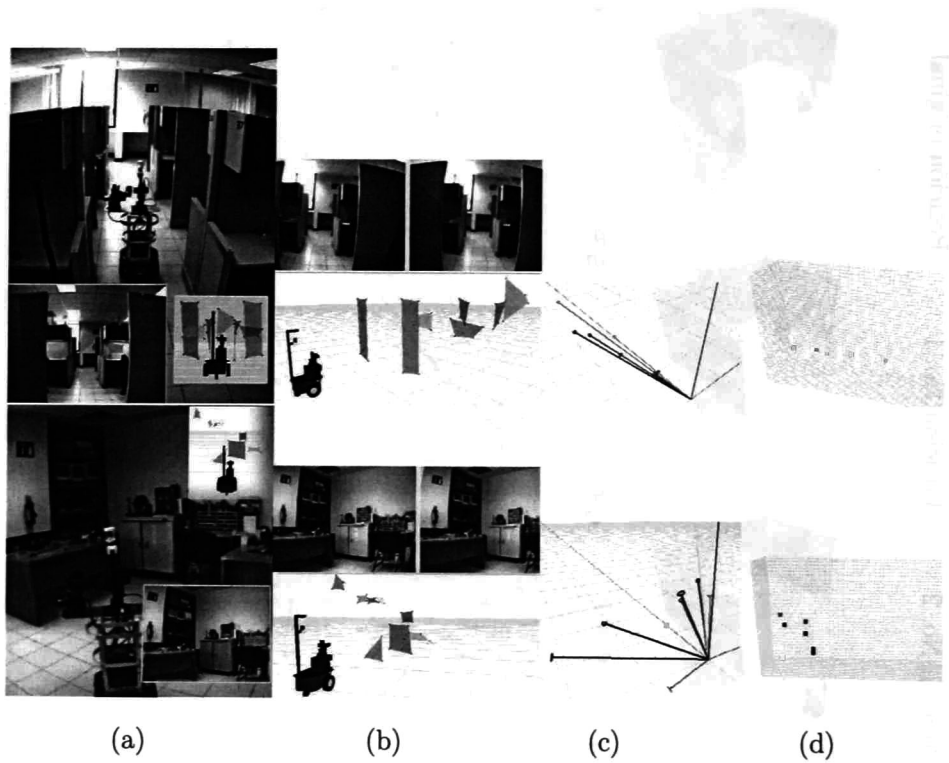


Figura 4.8: (a) Vista de dos escenarios con planos. (b) Imágenes estéreo y representación en el mundo virtual. (c) Los planos representados en \mathbb{R}^3 . (d) El escenario representado en el espacio paramétrico 3D de Hough.

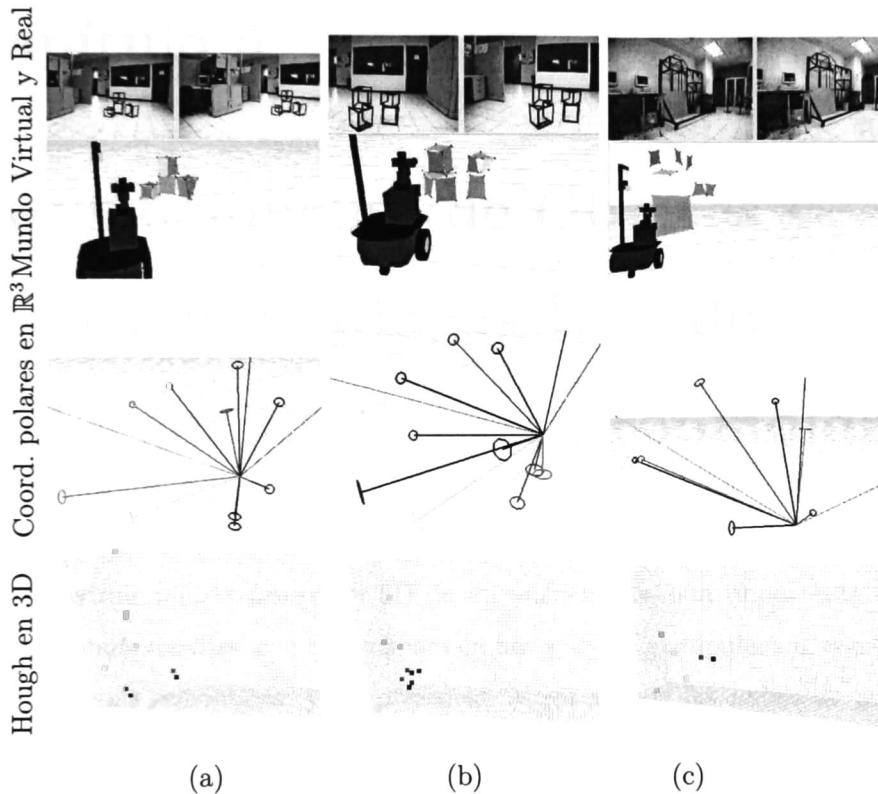


Figura 4.9: Diferentes estructuras donde se muestra al robot detectando el ambiente y los objetos, su representación en coordenadas polares en \mathbb{R}^3 y en el espacio paramétrico 3D de Hough

**CAPÍTULO 4. REPRESENTACIÓN Y RECONOCIMIENTO DE OBJETOS
USANDO LA TRANSFORMADA DE HOUGH**

70

Capítulo 5

Reconocimiento de Objetos y Creación de Mapas Usando Imágenes RGB y su Profundidad

Construir mapas densos en 3D de un ambiente es una importante tarea para los robots móviles, con aplicaciones en navegación, manipulación, construcción de mapas semánticos, y telepresencia. Mientras que las nubes de puntos son extremadamente buenas para el acoplamiento cuadro-a-cuadro y para una reconstrucción 3D densa, estos ignoran la valiosa información que contienen las imágenes. Las cámaras de color por otro lado, capturan una vasta información visual y se están convirtiendo más y más en el sensor escogido en la detección con láser cerrado [28][22]. Sin embargo es extremadamente difícil de extraer una profundidad densa de toda la imagen [18].

En los anteriores capítulos se trabajó con un sistema de cámaras estéreo y un sensor de alcance láser. Ahora, en este capítulo, se expondrá la manera de realizar reconocimiento de objetos y creación de mapas utilizando un dispositivo

que combine una nube densa de puntos 3D y una imagen a color; este dispositivo es el sensor de profundidad RGB-D¹. Con él, se desarrollaran varias técnicas para sacar mayor provecho a este sensor.

5.1. Cámaras RGB-D

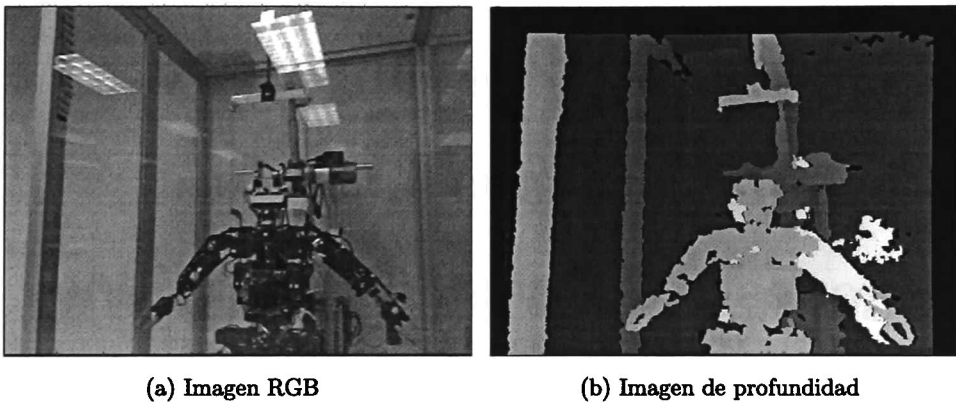


Figura 5.1: (a) Es la imagen capturada en RGB, y (b) es la información de la profundidad de la imagen (a).

Las cámaras RGB-D son sistemas que capturan las imágenes en RGB junto con la información de profundidad por pixel. Con la llegada de estas cámaras que son de bajo costo en el cálculo de la profundidad, los trabajos recientes se han enfocado en la explotación de estos sensores con el propósito de usarlos en la autonomía de los robots móviles. Estas cámaras de profundidad hacen la creación de nubes densas de puntos 3D, los cuales eran solo posibles con sensores mucho mas caros como los sensores de alcance láser 3D. Aplicaciones como la creación

¹RGB son las iniciales Red-Green-Blue (Rojo-Verde-Azul), este es el sistema de colores que se utiliza para las imágenes digitales. D es de Depth (profundidad).

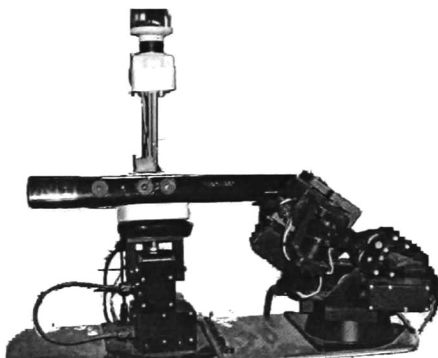
de mapas en 3D y su reconstrucción virtual, el análisis de formas geométricas, obtención de la pose y seguimiento de objetos, pueden ser beneficiados usando esta tipo de sensores. En la figura 5.1 se muestra la captura de una imagen utilizando estas cámaras.

Las cámaras RGB-D permiten la captura de imágenes con una calidad media en resolución, pero esto es suficiente para realizar los cálculos necesarios para la detección de objetos. El equipo que se está usando es un sensor desarrollado por la compañía *Prime Sense* el cual fue integrado en el sensor *Kinect* de Microsoft. La resolución de estas cámaras es de 640x480 píxeles con una captura de 30 cuadros por segundo de manera natural. Los cuadros por segundo decrecen a la hora de realizar algunos cálculos sobre las imágenes ya que estas se deben de alinear. Al decir alinear se refiere a que la imagen que se captura en RGB y la imagen en profundidad se enciman una sobre otra para así tener la correspondencia de profundidad con cada pixel de la imagen en RGB. Como se muestra en la figura 5.1 se puede ver una imagen a colores (RGB) (fig. 5.1a) y otra (fig. 5.1b) en escala de grises (profundidad), esta última por cada uno de sus valores de pixel corresponde a una cierta profundidad dentro de la imagen. Entre más claro sea el color de la imagen en profundidad más cercano está el objeto y entre más oscuro la distancia se incrementa.

Sin embargo estas cámaras tiene un pequeño inconveniente. El alcance de las mediciones en profundidad no es tan bueno comparado con el sensor de alcance láser o con sus semejantes mas caras que se utilizan para la creación de mapas en 3D. El campo visual que tiene el sensor de RGB es de aproximadamente de 60° mientras que otras cámaras alcanzan a tener un campo de visión mayor. El rango de medición de la profundidad esta en un intervalo de treinta centímetros hasta cinco metros aproximadamente.

En sensor RGB-D utiliza un haz de luz en infrarrojo para calcular la pro-

fundidad de los objetos que se encuentren en frente. Debido a esta naturaleza, los objetos que son transparentes (vidrios, vitrinas, etc.) no son detectables por el sensor de profundidad ya que los rayos infrarrojos pasan por ellos. Un ejemplo de esto se puede ver en la figura 5.1. También, los objetos que son de color oscuro (negros mate) no son detectables fácilmente por el sensor de profundidad ya que este, no detecta el reflejo de la luz infrarroja en ellos. Además, si se desea utilizar el sensor en exteriores, los objetos a los cuales les esta irradiando luz solar, no son detectados en su profundidad. Esto es porque la luz solar satura el sensor de profundidad infrarrojo.



(a) Sensor RGB-D



(b) Robot *Geometer* con sensor RGB-D

Figura 5.2: (a) El sensor RGB-D (Kinect), y (b) robot móvil con el sensor RGB-D montado en la parte superior.

En el presente trabajo se trabajará con este tipo de cámaras para la extracción de características de objetos y su futura clasificación así como también representarlos en un ambiente virtual. Para esto se montó uno de estos dispositivos a un robot móvil. La figura 5.2 muestra uno de estos dispositivos montado

sobre el robot móvil "Geometer", en donde se puede ver el sensor RGB-D conformado (de izquierda a derecha) del proyector infrarrojo, cámara RGB, cámara infrarroja. En la figura 5.3 se muestra los dispositivos que cuenta el sensor RGB-D (Kinect).

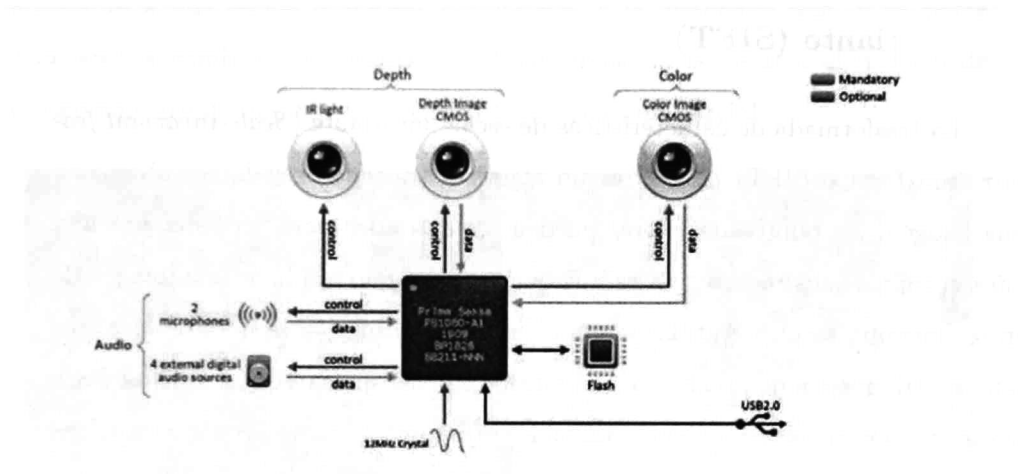


Figura 5.3: Esquemático del funcionamiento interno del sensor de profundidad RGB-D.

5.2. Reconocimiento de Objetos

En capítulos pasados se ha hablado del reconocimiento de objetos como las señales de emergencia localizadas en interiores. Estos objetos se tenían que entrenar para que una red neuronal en cascada pudiera clasificar y reconocer los objetos dentro de una imagen. El inconveniente de realizar esta tarea es el lento procesamiento que lleva el entrenar esta red neuronal para que reconozca un objeto. Existe un algoritmo dentro de visión computacional el cual es capaz de detectar características de objetos ante la variación de su tamaño. Este nos

ayudara a desarrolla una nueva técnica de identificación y reconocimiento de objetos dentro del medio ambiente.

5.2.1. Transformada de Características de Escala Invariante (SIFT)

La trasformada de características de escala invariante (*Scale-invariant feature transform* o SIFT) [25][27] es un algoritmo que para cualquier objeto en una imagen, los puntos de interés pueden ser extraídos para proporcionar una "descripción de sus características". Esta descripción, extraída de una imagen de entrenamiento, se puede utilizar para identificar el objeto a la hora de localizar este en otra imagen de prueba, la cual puede contener muchos otros objetos. Para llevar a cabo un reconocimiento confiable, es importante que las características extraídas de la imagen de entrenamiento sean detectables incluso en los cambios en la escala de la imagen, cuando tenga ruido o cambios de la la iluminación. Dichos puntos se encuentran por lo general en regiones de alto contraste de la imagen, tales como los bordes del objeto.

Otra cualidad importante de estas características es que las posiciones relativas entre ellos en la escena original no debe cambiar de una imagen a otra. Por ejemplo, si sólo las cuatro esquinas de una puerta fueron utilizados como características, el algoritmo funcionaría independientemente de la posición de la puerta, pero si los puntos en un cuadro también fueran utilizados, el reconocimiento daría un error si la puerta estuviera abierta o cerrada. Del mismo modo, las características ubicadas en objetos articulados o flexibles no suelen trabajar si algún cambio en su geometría interna ocurriera entre las dos imágenes a ser procesadas. Sin embargo, en la práctica SIFT detecta y utiliza un número mucho mayor de características de las imágenes, lo que reduce la contribución de los

errores causados por estas variaciones locales en todos los errores de correspondencia de características.

El método SIFT con firmeza puede identificar los objetos, incluso entre el desorden y en la oclusión parcial, ya que su descripción del carácter SIFT es invariante a una escala uniforme, orientación, y la distorsión afín, y en parte invariante a cambios de iluminación. A continuación se resume el método de reconocimiento de objetos en desorden y con oclusión parcial.

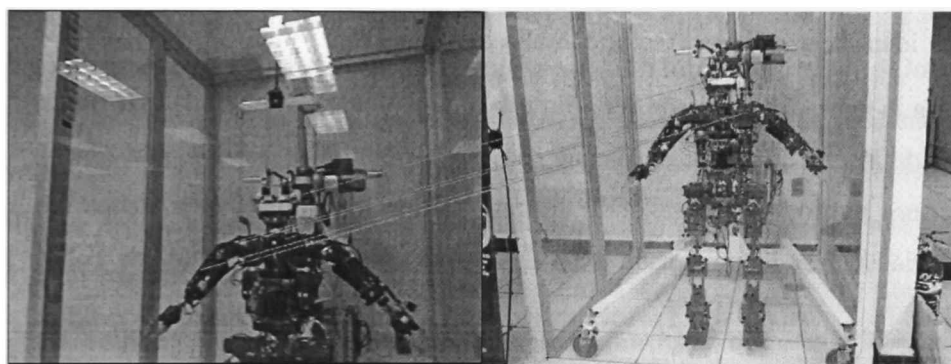


Figura 5.4: Aplicación de SIFT sobre un robot humanoide. Del lado izquierdo imagen obtenida con el sensor RGB-D. Del lado derecho una imagen de referencia usada como entrenamiento para extraer las características del objeto original.

Los puntos clave SIFT [15, 26] de los objetos primero se extraen de un conjunto de imágenes de referencia y se almacena en una base de datos. Un objeto se reconoce en una nueva imagen de forma individual comparando cada característica de la nueva imagen contra las características de la base de datos. La búsqueda de correspondencias de características que pueden ser candidato, está basada en la distancia euclidiana de los vectores de características. De todo el conjunto de correspondencias, los subconjuntos de puntos clave que coinciden sobre el objeto, su ubicación, escala y la orientación de la nueva imagen se identifican para

filtrar las buenas correspondencias. La determinación de los grupos consistentes se realiza rápidamente mediante el uso de una aplicación eficaz tabla hash de la transformada de Hough generalizada. Cada grupo de 3 o más características que coinciden sobre un objeto y su pose, se somete a una verificación más detallada y, posteriormente, se descartan los valores atípicos. Por último, la probabilidad de que un conjunto particular de características indica la presencia de un objeto se calcula, teniendo en cuenta la precisión de ajuste y el número de probables falsos partidos. Las coincidencias de un objeto que pasen todas estas pruebas se pueden identificar como correcta, con una confianza alta.

En la figura 5.4 se puede ver a un robot humanoide, al cual se le extrajeron puntos de interés [15] (imagen izquierda) y estos fueron utilizados para ubicar los mismos puntos en una segunda imagen (derecha), en donde esta última fue capturada utilizando un sensor RGB-D, el cual está montado sobre un robot móvil.

5.2.2. Representación de Marcas Espaciales

Para obtener marcas espaciales, las cuales nos servirán para la ubicación del robot móvil, es necesario extraer las características de los objetos que se requiere reconocer. En el capítulo 3 se habló de una técnica de reconocimiento de objetos, el cual, necesitaba entrenar previamente el objeto que se quiere reconocer para obtener sus características y así obtener su ubicación dentro de la imagen capturada, y finalmente tener su ubicación en el espacio de 3D. Esta técnica tiene la característica de que se necesita entrenar los objetos a reconocer de manera “offline” Esto quiere decir, que el entrenamiento se hace fuera del ambiente de navegación. Por lo tanto, si se quiere reconocer un nuevo objeto, es necesario capturar el objeto fuera de la navegación, entrenarlo e ingresarlo al sistema de

reconocimiento de marcas espaciales del sistema del robot móvil.

Ahora, si se utiliza la técnica de extracción de características como bordes y esquinas de un objeto [15] y teniendo estas características en escala invariante [25], el aprendizaje de este objeto se puede hacer en tiempo real. Basta con que al robot móvil se le diga cual es el objeto que se le quiere extraer las características, y este las guarde para su uso futuro. La utilización de SIFT para la identificación de objetos en el ambiente, da la habilidad al robot móvil de aprender nuevos objetos mientras este está navegando en interiores.

Al igual que con las marcas obtenidas con el método de Viola y Jones, una vez encontrados estos objetos, podemos representar el objeto en el ambiente virtual. En la figura 5.5, se ilustra un extintor el cual es tomado de diferentes ángulos y distancias. Este objeto se puede utilizar como marca espacial, a la hora de estar navegando.

Cualquier objeto identificado en el ambiente puede ser representado en ambiente virtual 3D. Como ahora se esta utilizando un sensor RGB-D, la distancia del objeto con respecto al sensor esta dado en la imagen de profundidad. De esta extraemos el valor de profundidad de los puntos característicos obtenidos. Ya con estos valores podemos representar al objeto en el ambiente virtual.

5.3. Obtención de Planos en RGB-D

Al estar trabajando en interiores es común encontrar planos (ver fig. 5.6), los cuales rodean el ambiente. El sensor de RGB-D regresa una imagen en color y una imagen de profundidad la cual se puede transformar en una nube densa de puntos y estos corresponden a la imagen en color.

Existen diferentes técnicas para encontrar planos como RANSAC o la trans-

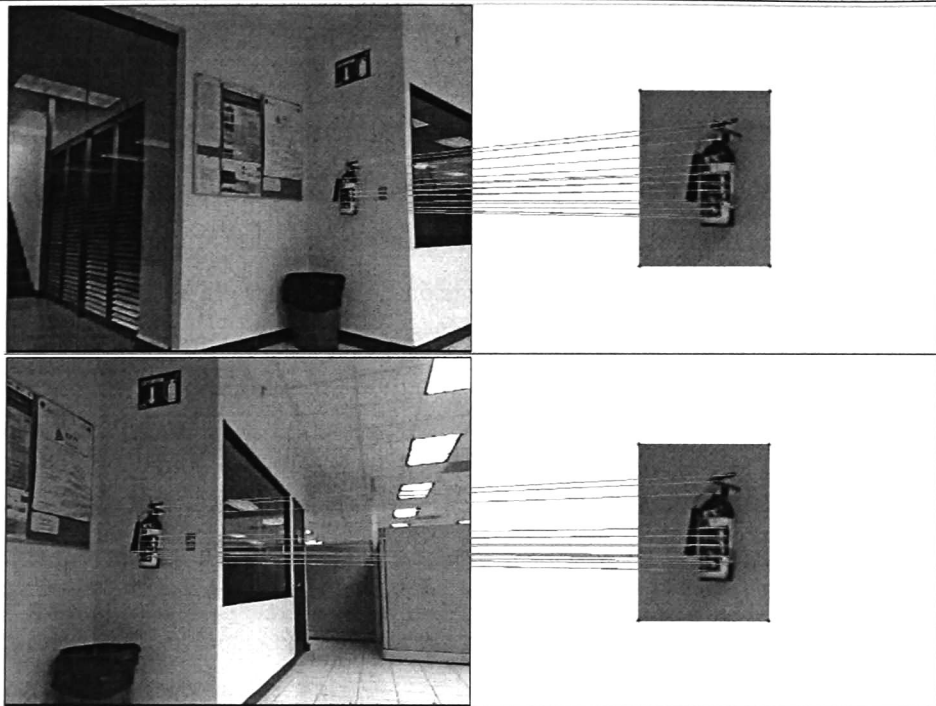


Figura 5.5: Una marca espacial (extintor) encontrada, en diferentes tomas.

formada de Hough. Estos buscan los planos a partir de una nube densa de puntos. A continuación se explicará una nueva técnica para encontrar planos paralelos al plano de la imagen de la cámara.

5.3.1. Planos en la Imagen de Profundidad

La imagen de profundidad del sensor RGB-D, regresa el valor de distancia que le corresponde a cada pixel en la imagen a color. Existe un pequeño detalle con esto, dependiendo de la orientación y posición del sensor, los planos ortogonales al piso, pueden tener diferentes valores de profundidad y con ello hace que

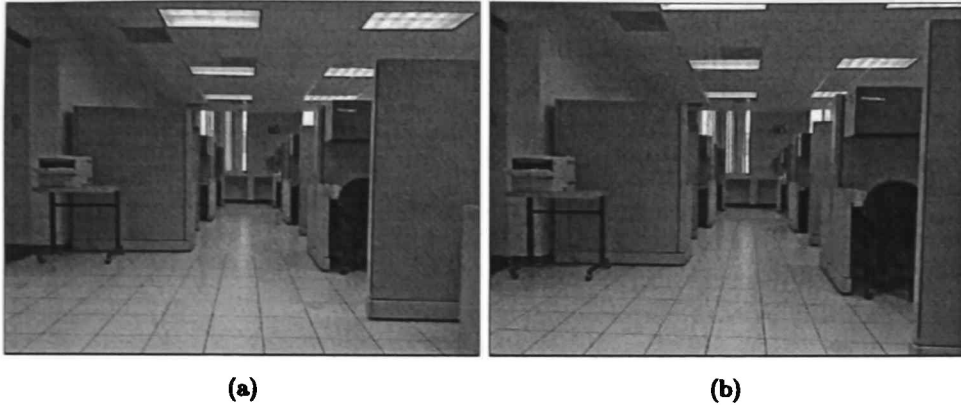


Figura 5.6: Dos vistas del pasillo con planos paralelos al plano de la imagen de la cámara.

estos planos no sean tan fáciles de extraer.

Como estamos buscando extraer información de un conjunto de datos, y los datos que queremos extraer tienen cierta característica. Podemos trasladar nuestro problema a un problema de probabilidad.

Sea D , el conjunto donde se quiere extraer la información de los planos, en nuestro caso la imagen de profundidad. Sea $x \in D$ donde $x \in [0, 10000]$, el valor de profundidad de cada elemento de la imagen. Realizamos un histograma con los valores de D , esto es hacer un conteo de todos estos elementos y los guardamos en $h(x)$, en donde esta función regresa el número de veces que se repite el valor de x . Si pensamos en x como una variable aleatoria continua, entonces $h(x)$ es la función de densidad de probabilidad de D en x .

Si estamos buscando planos en una imagen de profundidad, existirá un gran número de elementos que tengan el mismo valor si el plano se encuentra paralelo al plano de la imagen. Esto quiere decir, que, para encontrar al menos un plano, el valor máximo de $h(x)$ debe de pasar de un umbral ξ .

$$plano = \max(h(x), \xi), \forall x \in D \quad (5.1)$$

Si se hace una búsqueda de varios planos, entonces recorreremos desde el valor máximo de $h(x)$ hasta el número de plano que se desee buscar. Para poderlos distinguir de en la imagen de profundidad, los planos encontrados le asignamos un color. El procedimiento quedaría de la siguiente manera:

1. Obtener la imagen de profundidad D .
2. Calcular el histograma $h(x)$.
3. Ordenar descendentemente el histograma conservando sus respectivos valores.
4. Para distinguir los planos asignarle un color a los n primeros valores del histograma ordenado.
5. Generar la imagen de profundidad teniendo en cuenta si el pixel tiene ya un color pre-establecido.

Esta simple técnica de buscar planos paralelos al plano de la imagen no resulta del todo bien a la hora de usar el sensor de profundidad. La figura 5.7 muestra una segmentación por distancia, en la cual se muestra por colores, los pixeles que corresponden al mismo valor de profundidad. Aquí se puede ver los plano dentro del ambiente y los puntos pertenecientes al mismo, pero que a la hora de ser censados tienen diferente valor de profundidad. Esto es debido a que el sensor y con ello el plano de la imagen de la cámara, no están 100% perpendiculares al plano del piso y además los planos del ambiente y el plano de la imagen no siempre son ortogonales.

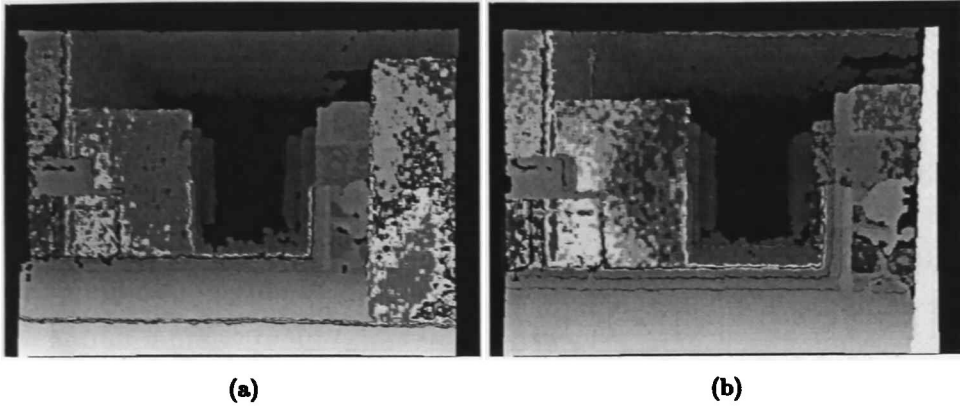


Figura 5.7: Segmentación por distancia en la imagen de profundidad.

Teniendo este pequeño inconveniente, vemos que este problema es un problema de cuantificación, e inspirándonos en [17], donde hablan sobre la cuantificación de una imagen en color, hacemos unos ajustes a la técnica descrita anteriormente.

Como ya habíamos dicho $h(x)$ es el histograma de la imagen de profundidad, y con él se localizaron los valores que tienen mayor probabilidad de ser un plano dados sus valores. Sea x_i un valor de estos planos, y $p = h(x_i)$, el número de veces que se repite x_i dentro de la muestra D . Tomando $\mu = x_i$, el objetivo es encontrar σ_i^2 . Lo que se busca con esto es minimizar la siguiente ecuación:

$$\sigma_i^2 = \sum_k f(x_k)(x_k - x_i)^2 \quad (5.2)$$

en donde $f(x_k)$ es el histograma $h(x_k)$ normalizado para que cumpla

$$\sum_D f(x) = 1.$$

Una vez encontrado σ_i^2 , procedemos a hacer una distribución normal con

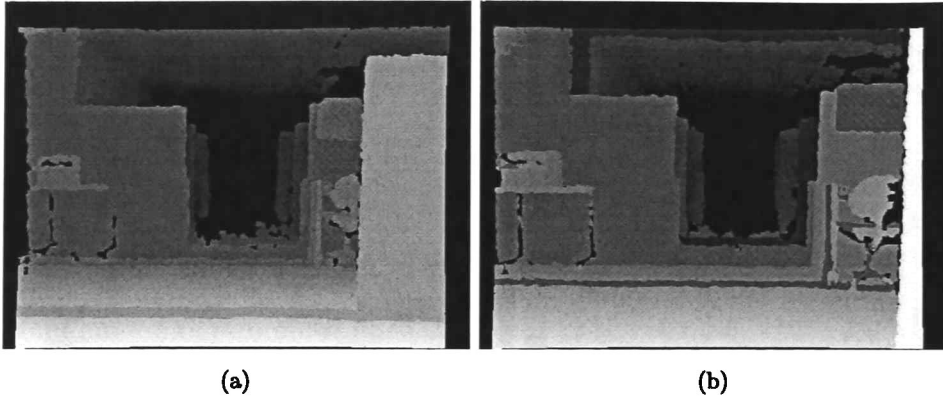


Figura 5.8: Segmentación de planos paralelos al plano de la imagen de la cámara.

$\mu = x_i$ y σ_i^2 . Los valores que caigan dentro de esta distribución serán seleccionados como parte del plano perteneciente a la distancia x_i , y se les asigna su respectivo color para poder ser identificados en la imagen. Este procedimiento se repite para cada plano que se quiera buscar.

Una forma de acelerar la búsqueda de estas distribuciones normales es visitando a los vecinos de x_i tanto del lado izquierdo y derecho hasta ver que el próximo valor es mayor al que se tiene actualmente.

Una vez que ya se reconocieron los planos dentro de la imagen en profundidad se pasa a representarlos en el espacio de Hough como se explica en el capítulo 4. La distancia de estos planos se toma como el valor máximo de la normal que contiene al plano, es decir $x_i = \mu$.

En la figura 5.8 se muestra como los planos de las imágenes ya fueron segmentados (coloreados). Este proceso está restringido a planos que estén paralelos al plano de la cámara.

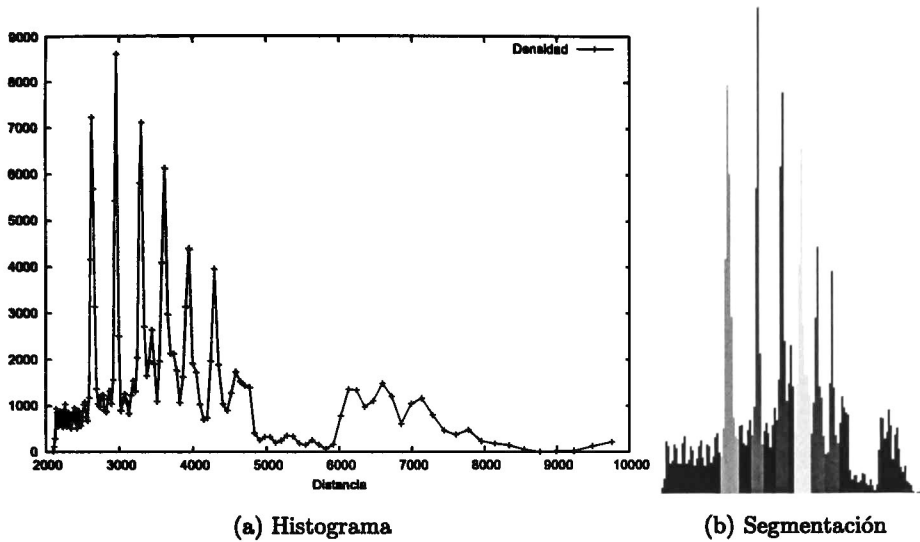


Figura 5.9: (a) Gráfica del histograma (función de densidad) para la identificación de planos en una escalera. (b) mismo histograma segmentado por colores identificando diferentes planos dentro de la imagen de profundidad.

5.3.2. Experimentos

Se hicieron varios experimentos entre ellos los concernientes al capítulo 4. La figura 5.9 muestra el cálculo del histograma y su segmentación. Aquí se puede ver como se encuentran diferentes planos en distancias distintas, los cuales son coloreados e identificados. También en la imagen de segmentación se pueden ver las distribuciones normales de cada plano.

La figura 5.10 muestra el reconocimiento de los planos que conforman una escalera. Del lado superior izquierdo es la imagen de profundidad segmentando los planos de la escalera. Este corresponde a los planos detectados en la figura 5.9, donde estos fueron segmentados por color. Del lado superior derecho es la imagen obtenida de la cámara RGB, de esta se extraen los colores que se deben de



Figura 5.10: El robot móvil viendo una escalera, obteniendo los planos que la conforman y su representación en el ambiente virtual 3D.

utilizar en la representación 3D. En la parte inferior se muestra la representación en 3D de las escaleras combinando la información de profundidad con la imagen de color.

Capítulo 6

Conclusiones

El interés por la navegación de robots móviles ha estado creciendo exponencialmente en los últimos años. El uso de nuevos sensores para el registro del ambiente ha abierto la posibilidad de integrar nuevos objetos para su representación espacial dentro de un mapa.

El uso del aprendizaje automático por computadora o aprendizaje de máquina ayuda a la búsqueda e identificación de objetos dentro de un ambiente común, y los cuales pueden ser utilizados como marcas espaciales para su representación dentro de un mapa virtual 3D. Se puede utilizar casi cualquier objeto ubicado dentro de un ambiente, excepto aquellos que tengan formas similares o características planas que no puedan ser reconocidos por el algoritmo de identificación, como por ejemplo, libros de cubierta lisa, objetos de tamaño pequeño. Los objetos fácilmente distinguibles pueden ser entrenados para que la computadora aprenda sus características y así pueda identificarlos por medio de un sensor de visión.

El uso de un sistema de cámaras estéreo ayuda al cálculo de la ubicación de un objeto. Dichos objetos pueden representar marcas espaciales dentro de un

mapa realizado por el robot móvil.

La utilización del álgebra geométrica conformal ayuda a una sencilla representación de los objetos y estructuras, así como en el modelado del movimiento del robot y de los datos adquiridos por los sensores de visión y alcance láser.

La aplicación de nuevas tecnologías en el área de robótica y visión computacional es importante ya que estas traen nuevas maneras de adquisición de datos. Un ejemplo de estas nuevas tecnologías es el sensor de profundidad RGB-D. Este nuevo dispositivo ayuda en la exploración del ambiente que se encuentra en frente a él. Además de ser de gran utilidad, estos dispositivos son económicos y pueden estar al alcance de cualquier persona que quiera realizar experimentos.

6.1. Trabajo Futuro

Dentro de lo que es la creación de mapas y la navegación robótica, quedan varios temas a tratar al igual que en SLAM, donde se busca que la navegación sea automática y sin necesidad de intervención humana.

El uso de la cámara como medio de medición es algo que se puede realizar, así no es necesario un odómetro el cual diga la posición donde se encuentra el robot móvil.

Dentro del SLAM queda también mucho por desarrollar ya que hay varios factores que entorpecen la creación de estos mapas, como el registro de objetos que se estén moviendo dentro del ambiente, o el registro de objetos que representan un obstáculo pero no pueden ser alcanzados por el sensor láser.

Respecto al sensor RGB-D, las investigaciones de como aprovechar al máximo este dispositivo están empezando. Queda mucho por investigar, sobre las aplicaciones en las que se puede usar este sensor. Gracias a su bajo costo, y su

tamaño puede ser utilizado en casi cualquier lugar. La segmentación dentro de la imagen de profundidad es otra área que se debe de poner atención y más aun la combinación de esta imagen con los datos a color (RGB).

Apéndice A

Fusión de Dispositivos para la Creación de un Mapa en 3D

A.1. Modelado en 3D y Unificación de Coordenadas

Existen varios sensores que registran el ambiente que los rodean. Algunos de estos sensores capturan puntos como los sonares ultrasonido o sensores láser, otros capturan imágenes como las cámaras de vídeo, donde estas imágenes pueden ser de un escenario tal como es o con una perspectiva omnidireccional. Cada uno de estos sensores tiene su propio sistema de coordenadas y registra cada objeto según las mismas.

Al de registrar un mismo objeto por medio de diferentes sensores y combinar estos registros en un mismo ambiente, este objeto aparecería en distintos lugares de acuerdo a la posición física de cada uno de estos sensores. Para que este objeto aparezca dentro del ambiente ubicado en un mismo lugar y que todos las

mediciones tomadas por los diferentes sensores converjan en un mismo punto es necesario unificar las coordenadas de cada uno de éstos en un mismo sistema en común, creando así un ambiente único donde todos los sensores registren un objeto y este se encuentre de acuerdo al lugar y posición donde fue tomado dicho objeto.

El ambiente en el que se van a unificar estos objetos es un ambiente tridimensional creado con el estándar OpenGL. Ahí se ha modelado el robot móvil (geometer), con algunos de sus dispositivos con una escala en la cual todos los puntos registrados queden exactamente 1:1 con el robot. Por lo tanto lo que se registre en los sensores y se plasme en el ambiente virtual será la ubicación exacta en la realidad [9].

A.2. Uso de *Motores Conformales* en 3D

Para poder unificar todos los sensores en un mismo ambiente es necesario conocer la posición del origen del sistema de coordenadas de cada uno de estos sensores con respecto al origen de coordenadas del ambiente unificado. Esto lo podemos lograr gracias al *problema de calibración mano-ojo* [34, 3], y con la ayuda del *álgebra conformal* podemos encontrar el origen del sistema de coordenadas en cada uno de los sensores.

Utilizando los métodos mencionados en [30], modificándolos para el uso de rotores (2.74), trasladadores (2.80) y motores (2.82) según sus definiciones en el álgebra conformal y utilizando las propiedades de cada uno de ellos, podemos extraer los ejes de rotación y vectores de traslación necesarios para llegar al origen de coordenadas de cada uno de los sensores partiendo del origen de coordenadas del ambiente unificado.

Gracias a que estas entidades forman un álgebra [2] podemos realizar operaciones entre los trasladadores, rotores y motores dando como resultado otro motor, el cual tendrá en una sola entidad el desplazamiento y giro de todos los movimientos realizados individualmente.

A.3. Modelado de Dispositivos de Entrada

Llamamos dispositivos de entrada a aquellos dispositivos que ingresan información al sistema interno del robot móvil. Estas entradas pueden ser para saber su ubicación (odómetro), saber qué objetos están a su alrededor mediante medidores de alcance (sonar ultrasonido, sensor láser) o medidores de toque, conocer el ambiente tal cual lo ve una persona (cámaras de vídeo), sonidos en el ambiente, etc.

A.3.1. Desplazamiento del Robot Móvil (Odómetro)

El desplazamiento del robot generalmente es sobre un plano cartesiano con coordenadas x y y , y cuenta también con un ángulo de giro θ sobre un eje determinado (generalmente z). Estas mediciones se registran mediante la odometría, pero esta no es muy confiable debido a que las ruedas del robot pueden patinar y dar mediciones erróneas. Para corregir estas mediciones de desplazamiento utilizamos (A.2), y para la corrección del ángulo de giro del robot (A.1).

$$\theta = \theta_o + \theta_{corr} \quad (\text{A.1})$$

$$T = T_o + T_{corr} \quad (\text{A.2})$$

Declaramos a R_{pos} el rotor creado a partir de la definición (2.76) utilizando

el ángulo (A.1) y el eje de rotación z , y T_{pos} el trasladador definido por (2.80) y utilizando $t = (x, y, 0)^T$ obtenido de las coordenadas de desplazamiento en el plano cartesiano.

Anteriormente ya habíamos mencionado a R_{pos} y a T_{pos} pero solo para 2D y con nombres de R_{robot} y T_{robot} respectivamente. Ahora los usaremos en un ambiente tridimensional.

A.3.2. Unidad de Sensor Láser

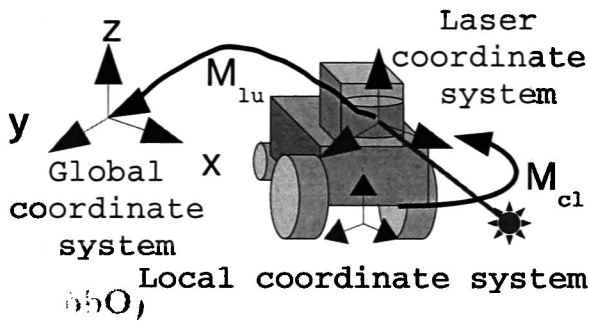


Figura A.1: Origen de coordenadas del sensor láser

Ahora tomando en cuenta el aspecto tridimensional y encontrando a que altura se localiza el sensor láser definimos M_{lsr} que es el motor que lleva desde el origen del sistema de coordenadas unificado al origen del sensor láser. Ahora definimos el motor que lleva los puntos registrados del sensor láser al sistema de coordenadas unificado en cualquier posición en la que se encuentre el robot móvil. Definimos a

$$M_{cl} = R_{pos} M_{lsr} \widetilde{R_{pos}} \quad (A.3)$$

$$M_{pos} = T_{pos} R_{lsr} \quad (A.4)$$

$$M_{lu} = M_{cl} M_{pos} \quad (A.5)$$

donde (A.3) es el motor de traslación y giro hacia el centro del láser; (A.4) es el movimiento del robot usando el láser y (A.5) es el motor que nos lleva al origen del sensor láser en el sistema de coordenadas unificado.

Utilizando (A.5) con cualquier entidad geométrica registrada con el sensor láser la podemos trasladar fácilmente al sistema de coordenadas unificado usando la forma

$$\underline{x}' = \mathbf{M}_{lu} \underline{x} \widetilde{\mathbf{M}}_{lu} \quad (\text{A.6})$$

A.3.3. Unidad Pan-Tilt

La unidad Pan-Tilt es una unidad que tiene dos grados de libertad, la cual hace que se mueva el cabezal tanto vertical como horizontalmente. Sean θ_{pan} y θ_{tilt} los grados de libertad que tiene esta unidad pan-tilt, utilizando *el problema de calibración mano-ojo* [30] encontramos los ejes de rotación $\underline{L}_{pan}, \underline{L}_{tilt} \in \mathcal{G}_{4,1}$ correspondientes al pan y el tilt respectivamente, encontramos el punto que tiene la distancia mínima entre los ejes \underline{L}_{pan} y \underline{L}_{tilt} , y creamos un trasladador a ese punto desde el origen de coordenadas unificado. Sea \mathbf{T}_{aje} el trasladador que manda el origen de coordenadas unificado al punto con distancia mínima entre los ejes del dispositivo pan-tilt. Este punto nos sirve como referencia para encontrar los ejes del pan-tilt. Como el eje del pan es casi vertical y el eje del tilt es casi horizontal, creamos los rotores \mathbf{R}_{pan} y \mathbf{R}_{tilt} en los ejes ¹ z y x respectivamente.

El trasladador a este punto, nos lleva al punto que sirve de partida para ir hacia el origen de coordenadas de las cámaras.

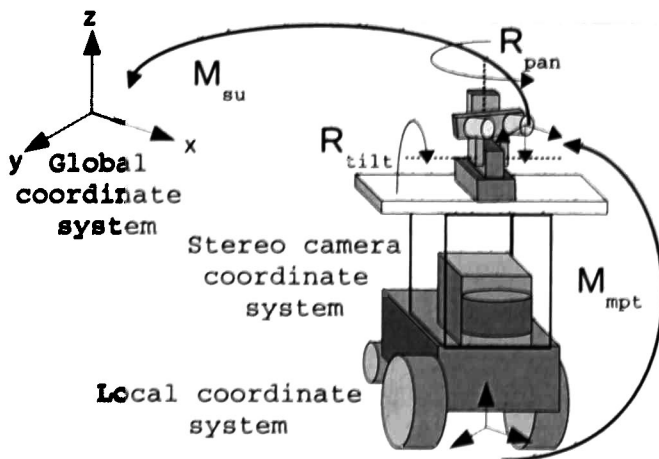


Figura A.2: Origen de coordenadas del sistema estéreo

A.3.4. Sistema de Cámaras Estéreo

El sistema estéreo tiene su origen de coordenadas en la cámara del lado izquierdo según nuestro modelo en el robot móvil (geometer). Para llegar ahí creamos un trasladador T_{ojo} desde el punto que tiene la distancia mínima entre los ejes del pan-tilt hacia el origen de coordenadas del ojo.

El motor que nos lleva desde el origen de coordenadas unificado al origen de coordenadas de la cámara estéreo se desarrolla de la siguiente manera:

$$T_{ep} = R_{pos} T_{ojo} \widetilde{R}_{pos} \quad (A.7)$$

$$R_{pt} = R_{pos} R_{pan} R_{tilt} \quad (A.8)$$

$$T_{opt} = R_{pt} T_{ojo} \widetilde{R}_{pt} \quad (A.9)$$

$$M_{mpt} = T_{pos} R_{pt} \quad (A.10)$$

$$M_{su} = T_{opt} T_{ep} R_{mpt} \quad (A.11)$$

donde (A.7) es la traslación al punto que tiene la distancia mínima a los ejes de

¹Se pueden usar L_{pan} y L_{tilt} como ejes de rotación utilizando (2.89).

pan-tilt, teniendo en cuenta la rotación de la posición del robot. (A.8) es el rotor resultante de todos los giros que sea han hecho tanto en la posición del robot, como en el pan-tilt. (A.9) es la tralación a la cámara izquierda del sistema estéreo teniendo en cuenta todos los giros que ha tenido el sistema. (A.10) es el motor de movimientos del robot, junto con el pan-tilt. (A.11) es el motor de movimiento completo del robot.

Cualquier punto registrado por las cámaras, en cualquier ángulo de θ_{pan} y θ_{tilt} , en cualquier posición del robot se podrá mandar de las coordenadas del sistema estéreo a las coordenadas del sistema unificado con

$$\underline{x}' = M_{rot} \widetilde{M}_{cam} \underline{x} \quad (A.12)$$

A.4. Realización del Mapa en 3D

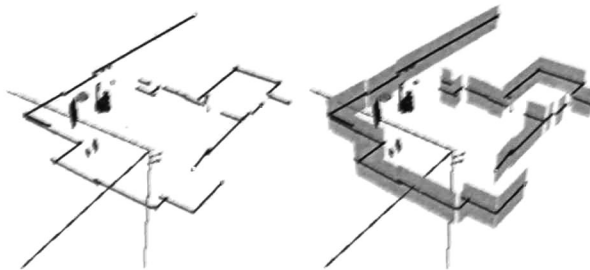


Figura A.3: Mapa 3D con y sin paredes virtuales

El sistema completo del robot móvil geometer (hasta lo que se ha explicado) tiene en total 5 grados de libertad para el registro de información. Estos son:

1. Posición en desplazamiento en el eje x
2. Posición en desplazamiento en el eje y

3. Posición del ángulo de rotación θ
4. Posición del ángulo de rotación de pan θ_{pan}
5. Posición del ángulo de rotación de tilt θ_{tilt}

Si en algún momento cambia el valor de alguno de ellos, es necesario realizar el calculo del motor, rotor, o traslador en el cual intervenga, para mantener actualizados los datos.

Una vez teniendo los datos actualizados se procede a capturar los datos y plasmarlos en el ambiente virtual utilizando (A.6) para las lecturas del láser y (A.12) para las lecturas de las imágenes estéreo (ver Fig. A.4).

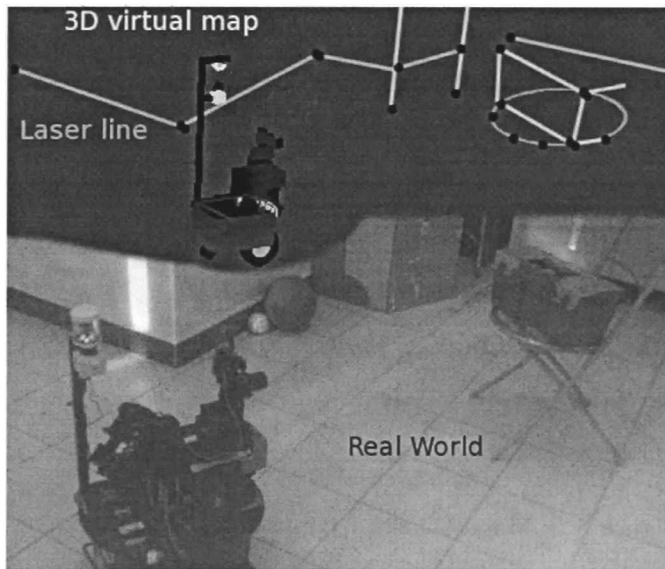


Figura A.4: Creación del mapa virtual con primitivas geométricas.

Apéndice B

Publicaciones

B.1. Revista

- Alberto Petrilli-Barceló, Heriberto Casarrubias-Vargas, Miguel Bernal-Marin, Eduardo Bayro-Corrochano, Rüdiger Dillmann. Geometric Techniques for Humanoid Perception, *International Journal of Humanoid Robotics*, vol. 7, no 3, pp.429-450, 2010
- Bernal-Marin, Miguel and Bayro-Corrochano, Eduardo Integration of Hough Transform of lines and planes in the framework of conformal geometric algebra for 2D and 3D robot vision, *Pattern Recogn. Lett.*, vol. 32, pp. 2213–2223, December, 2011

B.2. Capítulo de libro

- Miguel Bernal-Marin, Adan Landa-Hernández, Eduardo Bayro-Corrochano, Chapter 7 - Visual Geometric Entities for Orientation, Relocalization and

Navigation, in *Autonomous Robots: Control, Sensing and Perception*. pp. 152-183, A. Torres, E. Martínez, Eds. Cuvillier Verlag, 2011. ISBN 978-3-86955-866-0.

B.3. Conferencia

- Eduardo Bayro-Corrochano, Miguel Bernal-Marin. Generalized Hough Transform and Conformal Geometric Algebra to Detect Lines and Planes for Building 3D Maps and Robot Navigation *IEEE/RSJ International Conference on Intelligent RObots and Systems*, October, 2010.
- Miguel Bernal-Marin, Eduardo Bayro-Corrochano. Geometric Hough transform for robot vision, *Applied Geometric Algebras in Computer Science and Engineering*, June 2010.
- Miguel Bernal-Marin, Eduardo Bayro-Corrochano. Machine learning and geometric technique for SLAM, *Iberoamerican Congress on Pattern Recognition*, November, 2009.
- Miguel Bernal-Marin, Eduardo Bayro-Corrochano. Visual and Laser Guided Robot Relocalization Using Lines, Hough Transformation and Machine Learning Techniques, *IEEE/RSJ International Conference on Intelligent RObots and Systems*, October, 2009.
- Miguel Bernal-Marin, Eduardo Bayro-Corrochano. Visual and Laser Guided Robot Relocalization Using Lines and Hough Transformation, *IEEE-RAS International Conference on Humanoid Robots*, December, 2008.

Bibliografía

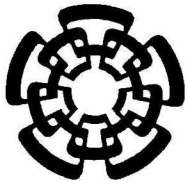
- [1] Pedram Azad, Tilo Gockel, and Rüdiger Dillmann. *Computer Vision: Principles and Practice*. Ed. Elektor Electronics, 2008.
- [2] Eduardo Bayro-Corrochano. *Geometric Computing for Perception Action Systems*. Springer, 2001.
- [3] Eduardo Bayro-Corrochano, Kostas Daniilidis, and Gerald Sommer. Motor algebra for 3d kinematics: The case of the hand-eye calibration. *J. Math. Imaging Vis.*, 13(2):79–100, 2000.
- [4] Edudardo Bayro-Corrochano. *Handbook of Geometric Computing: Applications in Pattern Recognition, Computer Vision, Neuralcomputing, and Robotics*. Springer, first edition, 2005.
- [5] Edudardo Bayro-Corrochano. *Geometric Computing: for Wavelet Transforms, Robot Vision, Learning, Control and Action*. Springer, first edition, 2010.
- [6] Miguel Bernal Marin. Mapas tridimensionales para la navegación robotica. Master's thesis, Centro de Investigación y Estudios Avanzados del IPN, Guadalajara, México, 2007.

-
- [7] Miguel Bernal-Marin and Eduardo Bayro-Corrochano. Visual and laser guided robot relocalization using lines and hough transformation. In *8th IEEE-RAS International Conference on Humanoid Robots, 2008.*, pages 239–244, Daejeon, South Korea, 2008.
- [8] Miguel Bernal-Marin and Eduardo Bayro-Corrochano. Machine learning and geometric technique for slam. In *CIARP '09: Proceedings of the 14th Iberoamerican Conference on Pattern Recognition*, pages 851–858, Berlin, Heidelberg, 2009. LNCS, Springer-Verlag.
- [9] Miguel Bernal-Marin and Eduardo Bayro-Corrochano. Visual and laser guided robot relocalization using lines, hough transformation and machine learning techniques. In *IROS'09: Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, pages 4142–4147, 2009.
- [10] John Canny. A Computational Approach to Edge Detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679–698, November 1986.
- [11] Franklin C. Crow. Summed-area tables for texture mapping. In *SIGGRAPH '84: Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 207–212, New York, NY, USA, 1984. ACM.
- [12] Luigi di Stefano, Stefano Mattoccia, and Federico Tombari. Zncc-based template matching using bounded partial correlation. *Pattern Recognition Letters*, 26(14):2129–2134, 2005.
- [13] Richard O. Duda and Peter E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15:11–15, January 1972.
- [14] Oliver Faugeras et al. Real-time correlation-based stereo: algorithm, implementation and applications. *INRIA Technical Report no. 2013*, 1993.

-
- [15] Chris Harris and Mike Stephens. A Combined Corner and Edge Detection. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [16] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [17] Paul S. Heckbert. Color image quantization for frame buffer display. *Computer Graphics*, 16:297–307, 1982.
- [18] Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren, and Dieter Fox. Rgbd mapping: Using depth cameras for dense 3d modeling of indoor environments. In *In RGB-D: Advanced Reasoning with Depth Cameras Workshop in conjunction with Robotics: Science and Systems Conference, Zaragoza, SPAIN*, 2010.
- [19] David Hestenes, Hongbo Li, and Alyn Rockwood. New algebraic tools for classical geometry. In [32], pages 3–26, 2001.
- [20] Paul Hough. Methods and means for recognizing complex patterns. *U.S. Patent 3,069,654*, 1962.
- [21] Luca Iocchi, Domenico Mastrantuono, and Daniele Nardi. A probabilistic approach to hough localization. In *Proceedings of the IEEE International Conference on Robotics and Automation, ICRA 2001, Seoul, Korea*, pages 4250–4255. IEEE, 2001.
- [22] Kurt Konolige and Motilal Agrawal. FrameSLAM: From Bundle Adjustment to Real-Time Visual Mapping. *Robotics, IEEE Transactions on*, 24(5):1066–1077, 2008.
- [23] Hongbo Li, David Hestenes, and Aalyn Rockwood. Generalized homogeneous coordinates for computational geometry. In [32], pages 27–59, 2001.

-
- [24] Pertti Lounesto. *Clifford Algebras and Spinors*. Cambridge University Press, 1997.
- [25] David G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision*, volume 2, pages 1150–1157, 1999.
- [26] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, pages 91–110, 2004.
- [27] David G. Lowe. Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image. *U.S. Patent 6,711,293*, 2004.
- [28] Paul Newman, Gabe Sibley, Mike Smith, Mark Cummins, Alastair Harrison, Chris Mei, Ingmar Posner, Robbie Shade, Derik Schroeter, Dave Cole, and Ian Reid. Navigating, recognising and describing urban spaces with vision and laser. *The International Journal of Robotics Research*, 28:1406–1433, October 2009.
- [29] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. *Sixth International Conference on Computer Vision*, pages 555–562, 1998.
- [30] Leo Hendrik Reyes Lozano. *Métodos Geométricos para Visión Robótica*. PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, Guadalajara, México, 2004.
- [31] Edward Rosten, Reid Porter, and Tom Drummond. Faster and better: A machine learning approach to corner detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32:105–119, 2010.

-
- [32] Gerald Sommer, editor. *Geometric Computing with Clifford Algebras*. Springer-Verlag, Heidelberg, 2001.
- [33] Fayez Tarsha-Kurdi, Tania Landes, and Pierre Grussenmeyer. Hough-transform and extended ransac algorithms for automatic detection of 3d building roof planes from lidar data. In *Proc. of the International Society for Photogrammetry and Remote Sensing, Workshop on Laser Scanning, Houston, Texas, USA*, volume XXXVI, pages 407–412, 2007.
- [34] Roger Y. Tsai and Reimar K. Lenz. A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. *IEEE Transactions on Robotics and Automation*, pages 5(3):345–358, June 1989.
- [35] Paul Viola and Michael J. Jones. Robust real-time face detection. *Int. J. Comput. Vision*, 57:137–154, May 2004.
- [36] Paul A. Viola and Michael J. Jones. Rapid object detection using a boosted cascade of simple features. In *CVPR (1)*, pages 511–518. IEEE Computer Society, 2001.
- [37] Paul A. Viola and Michael J. Jones. Robust real-time object detection. In *International Journal of Computer Vision*, pages 1,3, 2001.



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL I.P.N. UNIDAD GUADALAJARA

El Jurado designado por la Unidad Guadalajara del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional aprobó la tesis

Técnicas Geométricas y Aprendizaje Automático para la
Reconstrucción 3D y Navegación Robótica

del (la) C.

Miguel BERNAL MARÍN

el día 27 de Febrero de 2012.

Dr. Eduardo José Bayro Corrochano
Investigador CINVESTAV 3D
CINVESTAV Unidad Guadalajara

Dr. Juan Luis Del Valle Padilla
Investigador CINVESTAV 3C
CINVESTAV Unidad Guadalajara

Dr. Alexander Georgievich Loukianov
Investigador CINVESTAV 3C
CINVESTAV Unidad Guadalajara

Dr. Luis Eduardo Falcón Morales
Profesor del Planta
Tecnológico de Monterrey

Dr. Carlos Alberto López Franco
Profesor Investigador
Universidad de Guadalajara



CINVESTAV - IPN
Biblioteca Central



SSIT0011149