

# **ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

## **Facultad de Ingeniería en Electricidad y Computación**

“Hacia la navegación autónoma de robots a partir de la implementación de un método de localización y mapeo simultáneos (SLAM) mediante el uso de un sistema de visión 3D”

### **TESIS DE GRADO**

Previa a la obtención del Título de:

### **INGENIERO EN COMPUTACIÓN**

### **ESPECIALIZACIÓN SISTEMAS TECNOLÓGICOS**

Presentado por:

**MIGUEL REALPE ROBALINO**

**GUAYAQUIL – ECUADOR**

**AÑO**

**2006**

# AGRADECIMIENTO

Quisiera agradecer a todas las personas que, de una u otra manera, han contribuido al desarrollo de la presente tesis.

A mis padres, Miguel Realpe y Azalia Robalino, por su apoyo incondicional, confianza, ejemplo y amor que han forjado mi carácter y me han ayudado a seguir adelante ante cualquier adversidad.

A Daniel y Cristina, por ser cómplices y amigos más que hermanos.

Al Dr. Boris Vintimilla, por su aporte, paciencia, confianza y dedicación en la dirección de esta tesis.

Al Dr. José Gaspar por la orientación y sugerencias en el desarrollo de las pruebas experimentales de esta tesis y por el tiempo y dedicación invertidos en el laboratorio y frente al pizarrón.

Finalmente, quisiera agradecer a mis compañeros del CVR y a todos aquellos que han demostrado su amistad en momentos difíciles.

# DEDICATORIA

A mis padres y mis  
hermanos

# **TRIBUNAL DE GRADUACIÓN**

---

Ing. Holger Cevallos  
SUB-DECANO DE LA FIEC

---

Dr. Boris Vintimilla  
DIRECTOR DE TESIS

---

Dr. Enrique Peláez  
VOCAL

# **DECLARACIÓN EXPRESA**

“La responsabilidad del contenido de esta Tesis de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”

(Reglamento de Graduación de la ESPOL)

---

Miguel Andrés Realpe Robalino

# RESUMEN

El presente trabajo ha sido desarrollado como un proyecto de tesis de graduación de la FIEC. El principal objetivo es generar mapas tridimensionales de ambientes reales a partir de imágenes de profundidad que son adquiridas desde una cámara de estéreo visión montada sobre un robot móvil que explora dichos ambientes y, a la vez, estimar la localización del robot dentro del ambiente.

Para un mejor entendimiento, se ha dividido el proyecto en 3 sistemas que son: sistema de adquisición 3D, sistema de localización, sistema de mapeo 3D. El sistema de adquisición 3D tiene como objeto obtener información del medio por el que navega el robot, el sistema de localización determina en que posición se encuentra el robot dentro del ambiente de navegación, y el sistema de mapeo 3D se refiere a representar tridimensionalmente el ambiente de navegación del robot.

Durante el desarrollo de los capítulos que conforman este trabajo se describe cada uno de los tres sistemas en los que se ha dividido la tesis.

# ÍNDICE GENERAL

<b>RESUMEN</b> .....	<b>V</b>
<b>ÍNDICE GENERAL</b> .....	<b>VI</b>
<b>ÍNDICE DE FIGURAS</b> .....	<b>IX</b>
<b>ÍNDICE DE TABLAS</b> .....	<b>XI</b>
<b>1. INTRODUCCIÓN</b> .....	<b>1</b>
1.1 .Descripción del proyecto .....	1
1.2 Justificación del proyecto .....	4
1.3 Objetivos.....	6
1.3.1 Objetivo general.....	6
1.3.2 Objetivos específicos .....	7
1.4 Análisis de costos .....	7
1.5 Organización del proyecto .....	8
<b>2. ANTECEDENTES BIBLIOGRÁFICOS</b> .....	<b>10</b>
2.1 Introducción .....	10
2.2 Análisis bibliográfico .....	13
2.2.1 Sistemas de adquisición .....	13
2.2.1.1 Estimadores explícitos.....	14
2.2.1.1.1 Estimadores internos.....	14
2.2.1.1.1.1 Odometría .....	15
2.2.1.1.1.2 Navegación inercial.....	16
2.2.1.1.2 Sensores externos de posición.....	18
2.2.1.1.2.1 Sistema de posicionamiento global (GPS) .....	19
2.2.1.1.2.2 Balizas.....	20
2.2.1.2 Estimadores basados en la percepción del entorno .....	21
2.2.1.2.1 Sonar.....	22
2.2.1.2.2 Adquisición 3D .....	23
2.2.1.2.2.1 Escáner láser... ..	23
2.2.1.2.2.2 Adquisición 3D mediante visión .....	25
2.2.2 Registro del movimiento. (Sistemas de localización 3D). .....	26

2.2.2.1 Registro de características .....	27
2.2.2.2 Registro de puntos .....	27
2.2.2.2.1 Etapas del ICP... .....	28
2.2.2.2.1.1 Selección de puntos.....	29
2.2.2.2.1.2 Correspondencia de puntos .....	29
2.2.2.2.1.3 Asignación de peso de correspondencias .....	30
2.2.2.2.1.4 Rechazo de pares .....	30
2.2.2.2.1.5 Asignación y minimización de la métrica de error .....	31
2.2.3 Integración de imágenes.....	31
2.2.3.1 Métodos poligonales.....	32
2.2.3.2 Métodos basados en funciones implícitas.....	33
2.3 Análisis de resultados .....	34
2.4 Conclusiones.....	38
<b>3. SISTEMA DE ADQUISICIÓN 3D .....</b>	<b>40</b>
3.1 Introducción .....	40
3.2 Sistema de visión estéreo.....	42
3.2.1 Geometría epipolar. ....	43
3.2.1.1 Cámara de estéreo visión Digiclops.....	44
3.2.1.2 Correspondencia y profundidad de puntos .....	46
3.3 Resultados.....	51
3.4 Conclusiones.....	54
3.4.1 Trabajos paralelos .....	55
<b>4. LOCALIZACIÓN Y MAPEO 3D .....</b>	<b>58</b>
4.1 Introducción .....	58
4.2 Registro de imágenes 3D .....	60
4.2.1 Registro del Movimiento. (Sistemas de Localización 3D) .....	60
4.2.1.1 Selección de puntos .....	61
4.2.1.2 Correspondencia de Puntos.....	62
4.2.1.3 Rechazo de pares.....	64
4.2.1.4 Asignación y minimización de la métrica de error .....	67
4.2.2 Alineación de Imágenes. (Creación de Mapas 3D) .....	69
4.2.2.1 Translación .....	70

4.2.2.2 Rotación.....	71
4.2.2.3 Alineación.....	72
4.3 Resultados.....	73
4.4 Conclusiones.....	82
<b>5. RESUMEN Y CONTIBUCIONES.....</b>	<b>84</b>
5.1 Introducción.....	84
5.2 Resumen.....	86
5.3 Contribuciones.....	89
5.3.1 Filtro en el espacio de color HSI.....	89
5.3.2 Creación de LUT (Look Up Table) para espacio HSI.....	90
5.3.3 Modelado de superficies de objetos reales.....	90
5.3.4 Publicaciones Científicas.....	91
5.3.5 Logros adicionales obtenidos con la ejecución de este proyecto.....	92
<b>CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>93</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>95</b>

## ÍNDICE DE FIGURAS

Figura 1.1	Sistema de visión estéreo montado sobre un robot móvil .....	3
Figura 1.2	Esquema del proyecto.....	4
Figura 2.1	Componentes de un sistema de odometría básico.....	15
Figura 2.2	Desventajas del sistema de odometría.....	16
Figura 2.3	Esquema de un acelerómetro omnidireccional .....	17
Figura 2.4	Interferencia entre sonares (Crosstalk) .....	22
Figura 2.5	Escáner láser: método indirecto .....	24
Figura 3.1	Geometría de un sistema estéreo.....	43
Figura 3.2	Sistema de visión estéreo constituida por 3 cámaras .....	45
Figura 3.3	Geometría epipolar y líneas epipolares .....	46
Figura 3.4	Zona donde se puede obtener profundidad en geometría epipolar .....	48
Figura 3.5	Imagen de disparidad.....	48
Figura 3.6	Sistema de coordenadas para la cámara estéreo.....	49
Figura 3.7	Nube de puntos.....	51
Figura 3.8	Sistema de visión estéreo montado sobre robot móvil .....	52
Figura 3.9	Nubes de puntos obtenidas experimentalmente .....	53
Figura 3.10	Procedimiento para la creación del modelo 3D de una superficie .....	56
Figura 3.11	Ilustración de un proceso de triangulación 2½D.....	56
Figura 3.12	Modelado 3D para objetos reales.....	57
Figura 4.1	Selección uniforme de puntos.....	62
Figura 4.2	Espacio de color RGB .....	64

Figura 4.3	Espacio de color HSI.....	65
Figura 4.4	Rechazo de pares de puntos.....	67
Figura 4.5	Nubes de puntos sin alinear.....	70
Figura 4.6	Nubes de puntos alineadas.....	73
Figura 4.7	Cámara Digiclops montada sobre el robot TRC.....	74
Figura 4.8	Trayecto lateral realizado por el robot en Prueba 1.....	75
Figura 4.9	Nubes de puntos resultantes en Prueba 1.....	75
Figura 4.10	Mapa 3D de ambiente de navegación del robot en Prueba 1.....	76
Figura 4.11	Trayecto lateral realizado por el robot en Prueba 2.....	77
Figura 4.12	Nubes de puntos resultantes en Prueba 2.....	77
Figura 4.13	Mapa 3D de ambiente de navegación del robot en Prueba 2.....	78
Figura 4.14	Trayectoria circular realizada por el robot.....	79
Figura 4.15	Trayectoria calculada de forma iterativa cada 5 imágenes.....	80
Figura 4.16	Nubes de puntos resultantes en Prueba 3.....	80
Figura 4.17	Segmentos del mapa 3D del ambiente de navegación del robot en prueba 3.....	81

## ÍNDICE DE TABLAS

Tabla 1.1	Costos de equipos requeridos para implementación de proyecto .....	8
Tabla 2.1	Cuadro comparativo de los sistemas de adquisición .....	35
Tabla 2.2	Precios de sensores de adquisición 3D .....	36
Tabla 2.3	Cuadro comparativo de los métodos de registro .....	37
Tabla 2.4	Cuadro comparativo de los métodos de integración de imágenes .....	37

Para tener éxito, la planificación sola es insuficiente.  
Uno debe improvisar también.  
Issac Asimov

# CAPÍTULO I

## INTRODUCCIÓN

Este capítulo presenta una descripción global del proyecto, así como la importancia que puede tener su aplicación en el campo de la navegación autónoma de robots. Se exponen además las bases sobre las cuales se justifica el proyecto y el aporte que este brinda al avance de la investigación en el campo de la robótica. Así mismo, describe los objetivos generales y específicos del proyecto.

Al final del capítulo se resume la manera cómo este documento está organizado para alcanzar las metas propuestas.

### 1.1 DESCRIPCIÓN DEL PROYECTO

Para el presente proyecto se plantea el tema “*Hacia la navegación autónoma de robots a partir de la implementación de un método de localización y mapeo simultáneos (SLAM) mediante el uso de un sistema de visión 3D*”. Este proyecto desea resolver dos problemas muy importantes en el campo de la navegación autónoma de robots: (1) *la localización de un robot móvil en un ambiente desconocido*, y (2) *la generación de mapas 3D de estos ambientes*.

La información que se obtenga de estas soluciones servirá, entre otras cosas, para que el robot pueda, mediante el uso de técnicas de navegación, desplazarse autónomamente en estos ambientes.

Según la literatura, los problemas de localización y mapeo están estrechamente ligados. En ambos problemas es necesario conocer la trayectoria del robot. Si no se conocen las transformaciones de la trayectoria no se podrá localizar al robot [1]. Los errores en la estimación de la posición del robot generan un error en la creación del mapa, es por esto que, no se puede estimar con precisión un mapa si no se conoce con precisión la trayectoria del robot.

La finalidad de los métodos de *localización y mapeo simultáneos (SLAM)* es que un robot que use un sistema de visión pueda construir de forma incremental un mapa de su medio de navegación, y mientras utiliza este mapa estime la trayectoria realizada por el robot al recorrer el ambiente desconocido sin ninguna información previa.

El desarrollo de la *localización y mapeo simultáneos (SLAM)* presenta sus principales aplicaciones en robots de navegación autónomos. Esto debido a que, gracias a la información proporcionada por los métodos SLAM, los robots pueden desplazarse con seguridad evadiendo los posibles obstáculos que se presenten en el terreno.

Para el presente proyecto se plantea usar un robot móvil provisto de un sistema de visión estéreo que permite obtener imágenes 3D del ambiente de navegación del robot, tal como se muestra en la Figura 1.1.

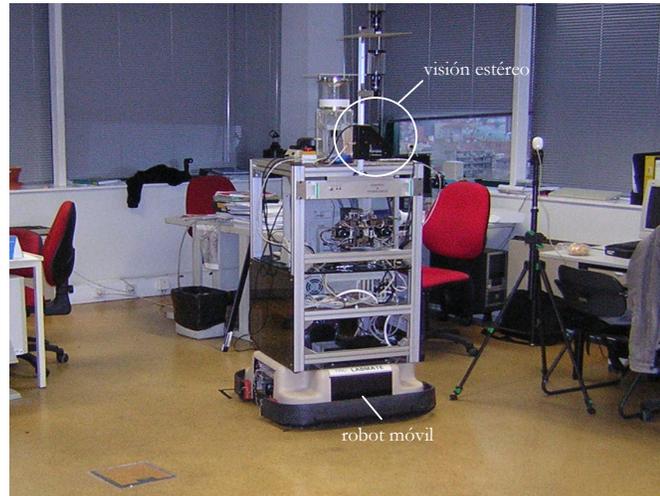


Figura 1.1.- Sistema de visión estéreo montado sobre un robot móvil

Empezando en un punto al azar, dentro del ambiente de navegación del robot, este realiza un recorrido de acuerdo a una trayectoria determinada por un operador. Durante el recorrido, se adquieren una serie de imágenes a través del sistema de visión estéreo, como se muestra en el ejemplo de la Figura 1.2 (*fila superior*). Para cada par de imágenes consecutivas adquiridas, el robot calcula el movimiento que ha realizado compuesto por la translación y la rotación, obteniendo una matriz de movimiento, Figura 1.2 (*inferior izquierda*). Luego aplica la matriz de movimiento a las imágenes 3D obtenidas para generar el mapa del entorno, tal como se presenta en la Figura 1.2 (*inferior derecha*).

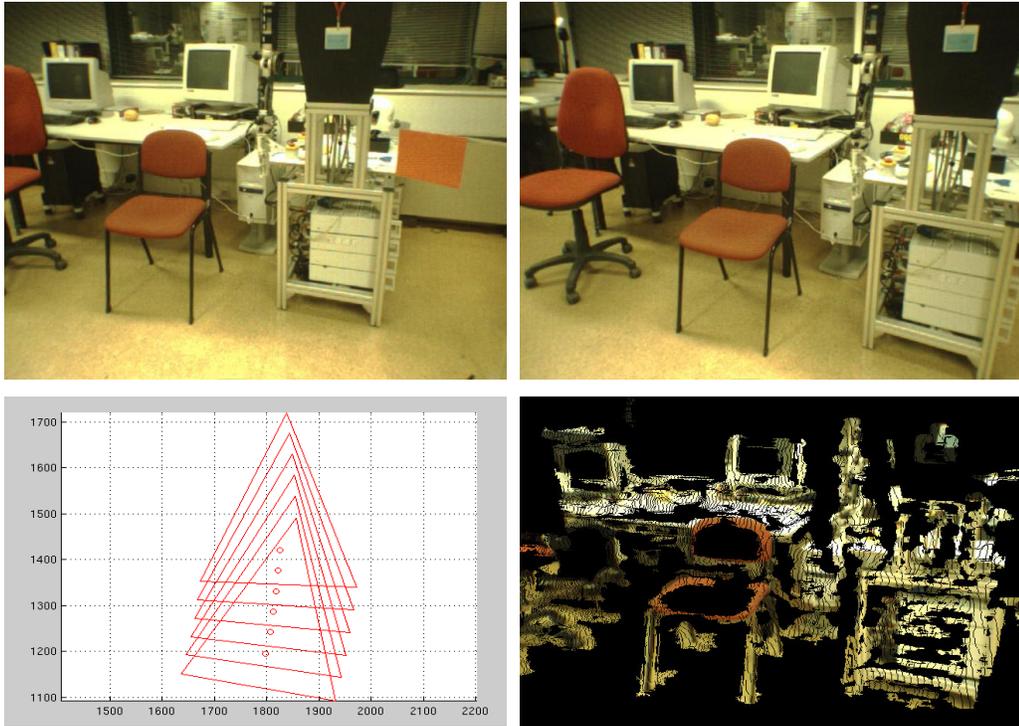


Figura 1.2.- *fila superior*) Par de imágenes consecutivas del ambiente de navegación del robot; *inferior izquierda*) Estimación del movimiento del robot; *inferior derecha*) Mapa generado del ambiente.

## 1.2 JUSTIFICACIÓN DEL PROYECTO

El proyecto prevé fortalecer los trabajos de investigación y desarrollo de la ESPOL en los campos de robótica y visión por computador, incentivando el desarrollo de proyectos de investigación enfocados al área de navegación autónoma de robots, implementando algoritmos de localización y modelado 3D de ambientes.

Al finalizar el proyecto los resultados obtenidos serán la: *Generación de un mapa 3D* correspondiente al entorno donde el robot móvil se desplaza, y la *Estimación de la posición del robot* dentro de este entorno.

En la actualidad existen múltiples aplicaciones que se están dando a los robots móviles equipados con sistemas de visión, entre las principales se pueden mencionar: asistencia en hospitales, exploración y transporte en minas, transporte automatizado en industrias, supervisión de ductos marinos, búsqueda y rescate, entre otros. En Ecuador aun no existen aplicaciones que involucren robots móviles que incluyan sistemas de visión, pero si existen varios proyectos de investigación y prototipos que se han desarrollado en esta área, principalmente con fines educativos. [37] [42].

Adicionalmente, con el proyecto se promueve el intercambio de experiencias y conocimientos entre investigadores a través de publicaciones de artículos y participación en congresos y eventos de carácter científico.

Se debe destacar que este proyecto es un modelo ideal de un proyecto semilla, que además de servir para la implementación de la localización y mapeo simultáneos (SLAM) abre también nuevos campos de investigación dado que servirá como base para futuras generaciones de proyectos en áreas relacionadas como: navegación autónoma de robots, modelado tridimensional de ambientes, y otras áreas afines.

Cabe mencionar que el desarrollo de las pruebas experimentales de adquisición para este proyecto se han llevado a cabo en el *Laboratorio de Visión por Computador - Vislab* del *Instituto de Sistemas e Robótica – ISR*, perteneciente a la *Universidade Técnica de Lisboa* (Portugal), como parte de las actividades de los Proyectos de Investigación: “Programación Asistida de Robots para Tareas Industriales (PARTI)” y de la “Red Iberoamericana de Robótica (RIBERO)”, financiados por la Agencia Española de Cooperación Internacional (AECI) a través del Programa CYTED (Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo), Subprograma VII: *Electrónica e Informática Aplicadas*, en los que el Centro de Visión y Robótica de la ESPOL participa.

### **1.3 OBJETIVOS**

El objetivo general y los objetivos específicos de este proyecto son detallados a continuación:

#### **1.3.1 Objetivo General**

Generar mapas tridimensionales de ambientes reales a partir de imágenes de profundidad adquiridas desde una cámara de estéreo visión montada sobre un robot móvil que explora dichos ambientes, y estimar la localización del robot dentro del ambiente.

### **1.3.2 Objetivos Específicos**

- Implementar un sistema de adquisición de imágenes del entorno de navegación del robot que permita adquirir información 3D de forma automática mientras el robot realiza un recorrido determinado.
  
- Diseñar e implementar un sistema que permita estimar el movimiento de una cámara de visión estéreo, montada sobre un robot móvil, a partir de dos imágenes consecutivas.
  
- Diseñar e implementar un sistema que permita modelar el ambiente que ha recorrido el robot, basándose en las imágenes obtenidas y en el movimiento que el robot ha realizado.

## **1.4 ANÁLISIS DE COSTOS**

Los costos de los equipos y materiales que se utilizaron en la implementación del proyecto se describen en la Tabla 1.1. En el Capítulo II se muestra detalladamente las distintas opciones que se pudieron utilizar en el desarrollo de este proyecto y se evalúan los costos que esto implica.

Ítem	Descripción	Cant.	P.U.	Total
1	Cámara 3D Digiclops *	1	6450	6450
2	Computador Pentium 4	1	700	700
3	Soportes de metal (10 cm.), tornillos y tuercas para soportes	4	5.30	21.2
<b>TOTAL</b>				7171.2
* Precio cotizado en USA P.U.= Precio Unitario, Cant. = Cantidad				

Tabla 1.1 Costos de equipos requeridos para implementación de proyecto

## 1.5 ORGANIZACIÓN DEL PROYECTO

Para resolver los problemas de localización y mapeo, se prevé dividir el proyecto en 3 fases: 1.) *Obtención de imágenes 3D de un entorno de navegación*; 2.) *Registro de movimiento de la cámara*; y 3.) *Mapeo del entorno de desplazamiento del robot*.

La descripción detallada sobre la implementación de cada una de estas tres fases está organizada como se menciona abajo:

- En el capítulo I se presenta una descripción general del proyecto y se realiza un análisis de costos de los equipos necesarios para la ejecución del mismo.

- El capítulo II realiza un estudio bibliográfico para analizar los diferentes métodos que actualmente se utilizan para solucionar los problemas propuestos. Las ventajas y desventajas de estos métodos son también descritos en este capítulo.
  
- En el capítulo III se describe la implementación de la fase de Obtención de imágenes 3D en el cual se indica los procesos que se han seguido para obtener la información tridimensional del ambiente de navegación del robot.
  
- El capítulo IV presenta el proceso para estimar el movimiento del robot, donde se describe la obtención de la matriz de movimiento, con la cual se puede conocer la localización del robot. Describe además la implementación de la fase de Mapeo del entorno de desplazamiento del robot, en el que se detalla el proceso de creación de un mapa 3D a partir de la matriz de movimiento.
  
- Un resumen general del trabajo realizado, así como las diferentes contribuciones del proyecto son presentados en el capítulo V.

## **CAPÍTULO II**

# **ANTECEDENTES BIBLIOGRÁFICOS**

En el presente capítulo se realiza un análisis bibliográfico de los distintos métodos existentes para obtener la ubicación de un robot dentro del medio por el cual se desplaza y para la creación de un modelo correspondiente a la forma física del medio. Primero se realiza una introducción general a la navegación autónoma, donde se resalta la importancia del proyecto propuesto. Luego se realiza el análisis de los antecedentes bibliográficos, en el cual se explican brevemente los sistemas de adquisición, localización del robot y la integración de imágenes para crear el mapa del ambiente de navegación del robot. Después se muestra un análisis de resultados en el que se exponen las ventajas y desventajas en el uso de cada una de las opciones expuestas anteriormente. Finalmente, varias recomendaciones para la implementación del proyecto son presentadas.

### **2.1 INTRODUCCIÓN**

La robótica móvil es un área de la robótica que se encarga de lograr que cualquier vehículo se pueda mover con alguna inteligencia propia, orientándose en el espacio de manera similar a como lo haría un humano:

esquivando obstáculos, moviendo articulaciones, etc. Dentro de esta área, un aspecto de vital importancia es el sistema de percepción por medio del cual se puede proveer del sentido de la vista a los robots para que estos puedan interactuar de forma más eficiente en ambientes o entornos complejos.

En los últimos años, los robots móviles guiados autónomamente han venido ganando importancia debido a la variedad de aplicaciones en diferentes campos en los que pueden ser utilizados [38] [39]. La esencia de estos robots es la capacidad de desplazarse autónomamente en terrenos, algunas veces desconocidos, usando técnicas de navegación apoyados con información de sensores como por ejemplo una cámara tridimensional o un láser para hacer el reconocimiento físico del entorno. De esta manera, el robot puede desplazarse con seguridad evadiendo los posibles obstáculos que se presenten en el trayecto, haciendo una reconstrucción de un modelo correspondiente a la forma física del terreno.

Uno de los aspectos importantes en la navegación autónoma de robots es el conocimiento de su posición y orientación en un determinado momento. Usualmente para la obtención de este conocimiento se utilizan algoritmos de mapeo para representar a los entornos de navegación del robot en forma de modelos virtuales. Esto permite al robot realizar tareas como planificar caminos, generar trayectorias, evitar obstáculos, entre otras cosas.

Los mapas a priori del entorno están raramente disponibles y cuando lo están, estos pueden presentar imprecisiones, estar incompletos o pueden no estar actualizados. De esta forma, lo deseable sería una construcción automática del entorno donde el robot va a navegar. Hoy en día, este desafío se ha convertido en un importante campo de investigación de la robótica móvil [2].

El objetivo principal de este capítulo es realizar una revisión bibliográfica de los distintos métodos existentes para crear la reconstrucción de un modelo correspondiente a la forma física del medio por el cual se desplaza un robot, y la obtención de su ubicación dentro de este medio. Esta información con la visión de crear un modelo virtual que, en un futuro, pueda ser utilizado por el robot para navegar autónomamente. Para esto, diferentes técnicas relacionadas con la adquisición de datos y modelamiento de ambientes para la navegación de robots serán revisadas.

Se pretende que la revisión bibliográfica además de ayudar a entender mejor los distintos métodos utilizados para la obtención de información y la creación de un modelo del medio de navegación de un robot, permita también escoger la mejor alternativa de solución que sirva como base para la futura implementación de la navegación autónoma de un robot en un entorno desconocido. Algunas posibles soluciones son propuestas, las cuales muestran las ventajas y desventajas frente a su costo económico y a las facilidades que prestan al cumplimiento del objetivo principal.

## 2.2 ANÁLISIS BIBLIOGRÁFICO

Para un mejor entendimiento, se ha dividido el proyecto en 3 partes principales que son: *Sistema de Adquisición*, *Registro del Movimiento* e *Integración de Imágenes*. El sistema de adquisición tiene como objeto obtener información del medio, el registro del movimiento determina el movimiento que ha realizado el robot en cada adquisición de datos consecutivos, y la integración de imágenes se refiere a la unificación de las distintas imágenes adquiridas durante el recorrido del robot.

Un análisis de las diversas técnicas existentes por los cuales se puede optar para la implementación de estos subsistemas es descrito en los siguientes apartados.

### 2.2.1 Sistemas de adquisición

Los sistemas de adquisición son aquellos que permiten al robot móvil obtener información propia o del medio por el cual navega. En la literatura el sistema de adquisición para la navegación autónoma de robots se divide en dos grupos: *estimadores explícitos* y *estimadores basados en la percepción del entorno* [3].

### 2.2.1.1 Estimadores explícitos

Los estimadores explícitos proporcionan la posición y orientación del robot a partir de medidas más o menos directas con sensores del robot (acelerómetros, giroscopios). Estos estimadores a su vez se dividen en dos grandes grupos que son: *estimadores internos* y *estimadores basados en estaciones de transmisión o sensores externos de posición*. [40]

#### 2.2.1.1.1 Estimadores internos

Los estimadores internos se refieren a sensores integrados al robot que proporcionan medidas internas y no trabajan con información del medio donde se desenvuelve el robot. El problema principal con este tipo de métodos se da en fallas mecánicas y en pequeñas desviaciones e imprecisiones que se acumulan. Esto último, produce como resultado estimaciones de falsas posiciones del robot.

Basándose en las medidas realizadas por los sensores internos se puede calcular la posición del robot por medio de la integración de las distancias recorridas ( $\Delta s$ ) y ángulos girados ( $\Delta\theta$ ) en cortos periodos (*Dead reckoning*) [4]

Las dos técnicas de mayor uso de los estimadores o sensores internos son la *odometría* y la *navegación inercial* [41].

#### 2.2.1.1.1 Odometría

La odometría es la técnica que calcula la posición y orientación del robot midiendo el giro de las ruedas. Para saber cuales han sido los movimientos angulares que han realizado las ruedas se utilizan encoders, que son codificadores ópticos montados en los ejes de giro del robot. Generalmente, un encoder consiste en un disco codificado montado en el eje de la rueda por medio del cual pasa un haz de luz proveniente de un led emisor a un fotorreceptor, al girar el disco, el haz de luz se corta intermitentemente creando una secuencia que permite calcular el número de vueltas que ha dado la rueda. En la siguiente gráfica se muestra el montaje de los componentes de un sistema de odometría. [7]

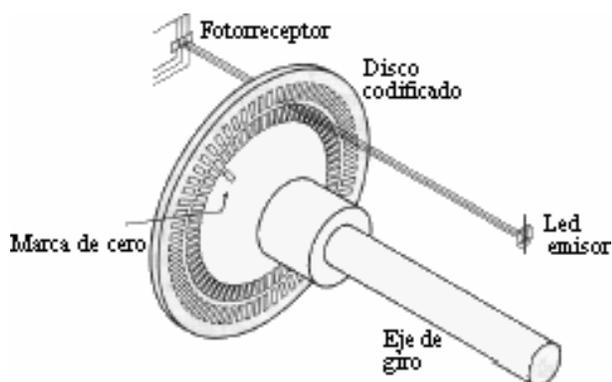


Figura 2.1.- Componentes de un sistema de odometría básico.

Las principales ventajas de la odometría son su sencillez y bajo costo, en contraparte las desventajas son la acumulación de errores y las variaciones al desplazamiento ideal del robot, tales como deslizamiento lateral (*lateral slip*), patinaje (*slipping / sliding*) y contacto superficial variable debido al desgaste de las ruedas, a la variación del peso del robot y a la variación de la compactación del terreno, Figura 2.1. [5]

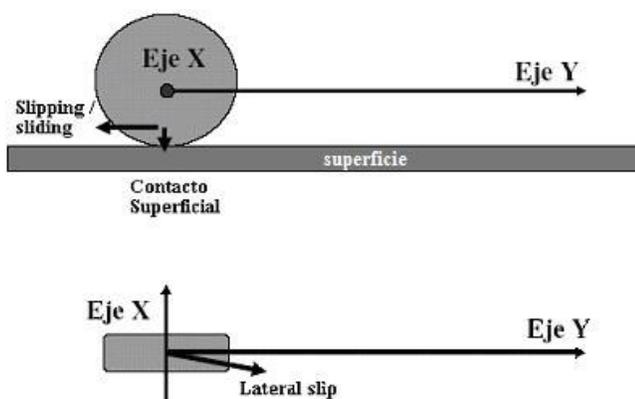


Figura 2.2.- Desventajas del sistema de odometría.

#### 2.2.1.1.2 Navegación inercial

Los métodos de navegación inercial se basan en la medición de la aceleración y rotación del robot y en base a éstas medidas determinan su posición y orientación [8]. Las medidas de aceleración son integradas, una vez para determinar la velocidad y dos veces para saber el desplazamiento.

Para realizar las mediciones de aceleración y rotación se hace uso de acelerómetros y giroscopios. Los acelerómetros, generalmente, son masas montadas sobre resortes, su aceleración se puede estimar al medir el movimiento de las masas y la relación del resorte utilizando las ecuaciones de la ley de Newton,  $F = ma$ , y la constante del resorte  $F = kx$ , de donde se obtiene:  $a = kx / m$ . Combinando 3 acelerómetros ortogonales, se obtiene un acelerómetro omnidireccional [6]. La Figura 2.3 muestra un esquema de este tipo de acelerómetro.

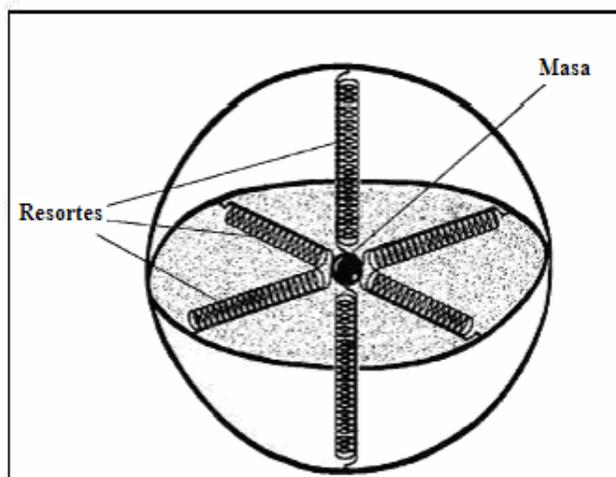


Figura 2.3. Esquema de un acelerómetro omnidireccional.

Los giroscopios son utilizados para medir ángulos de orientación, y pueden ser de tipo *mecánico* u *óptico*. Los *giroscopios mecánicos* emplean el principio de conservación de momento angular para mantener uno o más ejes internos apuntando a la misma dirección mientras el exterior del giroscopio se

traslada y rota. [7]. Los *giroscopios ópticos* consisten, básicamente, en dos rayos láser moviéndose en direcciones opuestas alrededor de una ruta de lazo cerrado. Los patrones de choque formados por las divisiones y mezclas de los dos rayos son usados para determinar el índice y dirección de la rotación del giroscopio [8]. Para la medición de los ángulos también se puede hacer uso de brújulas, pero se debe tener especial cuidado con campos magnéticos que pueden afectar su desempeño.

La navegación inercial tiene como principales ventajas ser auto-contenida y no necesita información externa del movimiento, y a diferencia de la odometría ningún error momentáneo en la orientación puede causar un crecimiento en el error [8]. Por otro lado, la principal desventaja es que la aceleración debe ser integrada dos veces y un pequeño error realizado en el momento de la medición puede causar un enorme error en el valor integrado.

#### **2.2.1.1.2 Sensores externos de posición.**

En lo que a los sensores externos se refiere, las técnicas de posicionamiento absoluto basan sus medidas en las relaciones del robot con elementos externos como marcas artificiales o satélites [9]

Los sistemas basados en sensores externos de posición constan, generalmente, de un receptor montado en el robot y varios transmisores que se encuentran en el ambiente.

#### **2.2.1.1.2.1 Sistema de posicionamiento global (GPS)**

El GPS es un sistema para navegación en exteriores desarrollado por el ministerio de Defensa de EUA, está compuesto por 24 satélites orbitando alrededor de la tierra (más 3 de reserva) que envían señales RF codificadas con información de identificación. Un receptor colocado en el robot recibe señales de los distintos satélites, la estimación de la posición se la realiza a partir de las posiciones de los satélites y midiendo el tiempo que ha tardado la señal en llegar al receptor. [10]

En el sistema GPS el gobierno de EUA ha introducido de forma intencional pequeños errores en el tiempo y posición de los satélites para prevenir que sean utilizados para dirigir armas de otras naciones. Este error intencional es denominado disponibilidad selectiva (SA) [8]

Para evitar errores de SA se utiliza un sistema de GPS diferencial (DGPS) que funciona en base a un trasmisor cercano con una posición conocida. El DGPS se basa en que un segundo receptor GPS que se encuentra próximo al primero obtiene el mismo error al usar como referencia los mismos satélites. Como el

segundo receptor se encuentra ubicado en una posición conocida este puede calcular el error introducido, corregirlo y enviar esta corrección al primero para así determinar la verdadera posición. En la actualidad existen muchas estaciones que realizan el servicio de corrección diferencial para clientes comerciales [10]

La ventaja de los Sistemas GPS es la precisión que ofrecen para grandes áreas, en su forma básica tiene una precisión del orden de 100 m y en forma diferencial hasta de 1 m. Además que es un método que ofrece una localización absoluta en áreas realmente grandes sin requerir estructuración del entorno. Las desventajas se dan por el bloqueo de señales en zonas de difícil recepción, los GPS y DGPS suelen fallar en áreas urbanas debido al efecto de valle que se produce a causa de edificios altos y árboles [11]. No es posible usar este sistema con exactitud en edificios, subsuelo y bajo el agua, tampoco en bosques densos ni entre edificios altos.

#### **2.2.1.1.2.2 Balizas**

El sistema de balizas se refiere a un conjunto de marcas artificiales ubicadas en el entorno de navegación mediante las cuales se puede determinar la ubicación y orientación del robot. Las balizas pueden ser de dos tipos: *activas* o *pasivas*. [43]

En el sistema de *balizas activas (Active Beacons)* se utilizan emisores infrarrojos, ultrasonido o de radiofrecuencia que son detectados e identificados en base a la señal y tiempo en que transmiten. Las *balizas pasivas (Passive Beacons)* se refieren a elementos pasivos como códigos de barras y símbolos, un ejemplo se muestra en [12]

Las balizas pueden ser localizadas con facilidad y rapidez, además proveen un posicionamiento exacto con un procesamiento mínimo. La principal desventaja del uso de balizas es que se requiere un tiempo para el diseño y la colocación de las marcas por lo que se incurre en un gran costo de instalación y mantenimiento. Si la colocación de las balizas no es exacta, no se podrá obtener un posicionamiento exacto. Además las balizas introducen una contaminación visual y espacial, al colocar elementos que normalmente no se encuentran en el entorno [8]

#### **2.2.1.2 Estimadores basados en la percepción del entorno**

Los estimadores basados en la percepción del entorno utilizan sensores que suministran información del entorno de navegación del robot, permitiendo determinar la localización del robot mediante comparación con otros datos o modelos conocidos. [13] [14]

### 2.2.1.2.1 Sonar

Uno de los sensores de mayor uso para la percepción del entorno del robot es el sonar. El sonar es un sensor basado en el tiempo de vuelo de una señal de ultrasonido que es transmitida continuamente en forma de pulso, la cual es captada de nuevo al chocar con algún objeto y rebotar. Mediante física elemental, la distancia al objeto es determinada multiplicando la velocidad de la onda por el tiempo que le toma en recorrer la distancia. [8]

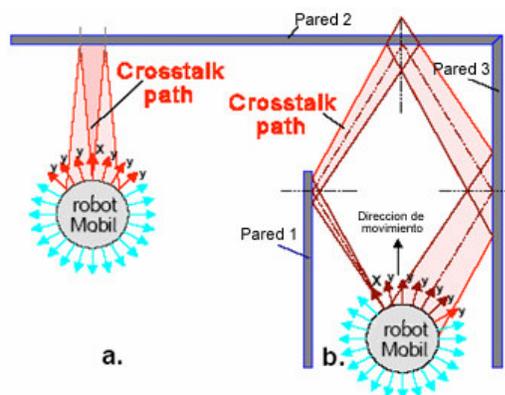


Figura 2.4. - Interferencia entre sonares (Crosstalk). Puede ser directo (a), o indirecto (b)

El uso del sonar tiene las ventajas de un reducido costo y una elevada velocidad de respuesta. Pero tiene algunas desventajas como la de obtener información muy reducida, ya que solamente adquiere la distancia sin mayor información sobre el obstáculo como por ejemplo el color o la textura. Además se producen rebotes de la señal obteniendo falsas lecturas por lo que un robot utiliza varios sonares dispuestos en forma de anillos, como por

ejemplo en [15], método que reduce la posibilidad de errores por rebotes falsos pero que introduce posibles interferencias con otros sonares (Crosstalk) como se aprecia en la Figura 2.4. [8]

#### **2.2.1.2.2 Adquisición 3D**

A partir de la información que se obtiene con un sonar sólo se pueden construir mapas en dos dimensiones. Debido a que el medio con el que interactúa el robot es tridimensional surge la necesidad de utilizar sensores más sofisticados que obtengan mayor información del entorno. En este sentido, se puede hacer uso de tecnologías de adquisición 3D para obtener un ambiente tridimensional del entorno del robot. Las dos tendencias de mayor uso con respecto a la adquisición 3D son: *escáner láser* y *visión estéreo*. [44]

##### **2.2.1.2.2.1 Escáner láser**

Los métodos de escáner láser se dividen en dos grupos: *indirectos* y *directos* [16] Los *indirectos* consisten en utilizar un patrón de luz láser y proyectarlo sobre una superficie determinada mientras es adquirida por medio de una cámara. Las coordenadas tridimensionales de un punto determinado pueden obtenerse de acuerdo a la diferencia en los patrones de luz reflejados en la superficie. En la Figura 2.5 se muestra un ejemplo para una proyección de luz

láser en forma de línea, una cámara CCD captura el reflejo por medio del cual se puede determinar la profundidad [17]

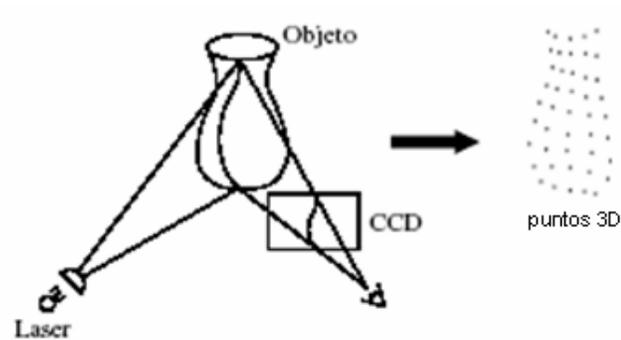


Figura 2.5. Escáner láser: método indirecto.

Los *métodos directos* consisten en emitir un pulso de luz por un láser que es reflejado al incidir sobre una superficie, en este caso la distancia al punto deseado se calcula determinando el retardo entre la emisión y la recepción de la señal.

El uso de los métodos de escáner láser permite obtener información tridimensional con una alta precisión (precisiones entre décimas y decenas de milímetros), en contraparte estos métodos presentan algunos inconvenientes, como un alto costo de equipos, alto tiempo de procesamiento, errores ante superficies reflectantes y la imposibilidad de obtener información sobre el color. [44]

#### **2.2.1.2.2.2 Adquisición 3D mediante visión**

A partir de una imagen bidimensional no se puede obtener una escena 3D sin ninguna información extra, para lograrlo se puede elegir entre 2 opciones: 1.- conocer características del medio y 2.- tomar más imágenes (visión estéreo).

En base del conocimiento de algunas propiedades geométricas conocidas de la escena, como paralelismo y ortogonalidad, se puede generar una representación 3D, un trabajo basado en esta técnica se muestra en [18], donde se requiere conocer a priori las características necesarias para la reconstrucción, lo cual debe ser realizado por una persona, imposibilitando así la automatización del proceso y su uso en la robótica.

La visión estéreo se refiere a la obtención de dos o más imágenes desde diferentes vistas de un objeto de interés, con el objetivo de encontrar sus coordenadas XYZ. Esto se obtiene en base a la configuración en el espacio físico (ubicación y orientación) de las cámaras con las que se obtienen las imágenes y a las diferencias existentes entre las imágenes adquiridas.

La adquisición mediante visión 3D permite obtener información tridimensional con una precisión bastante aceptable y con una alta velocidad de respuesta, en contraparte la principal desventaja radica en la sensibilidad a los cambios de iluminación.

### **2.2.2 Registro del movimiento. (Sistemas de localización 3D)**

Para reconstruir un modelo completo de una escena, por lo general, no basta con la información de una sola muestra realizada por el robot, sino que es necesario tener información de la escena obtenida desde varios puntos de vista. Luego es necesario realizar un registro de todos los datos obtenidos. El registro es el proceso de alinear espacialmente los datos de dos o más muestras que pertenecen a una misma escena [19].

El proceso de registro es utilizado en una gran variedad de aplicaciones, dependiendo de la aplicación varían los datos que se utilizan y el tipo de sensor con el que se obtienen dichos datos. Los datos pueden ser adquiridos por diferentes tipos de sensores al mismo tiempo, o por el mismo sensor en diferentes tiempos y pueden variar de datos 2D a datos 3D. En el caso del registro de datos 3D la meta es obtener una transformación 3D (translación y rotación) que permita unificar la información de dos puntos de vista de una escena en un solo sistema de coordenadas.

Los métodos de registro de imágenes 3D se pueden clasificar en dos grupos principales: métodos que registran un grupo de características extraídas de las imágenes, y métodos que registran los puntos presentes en las imágenes [17].

### **2.2.2.1 Registro de características**

Los métodos que registran un grupo de características extraídas de las imágenes buscan una forma de caracterizar las imágenes, para calcular la transformación de las coordenadas de las imágenes en el espacio de características (contornos, superficies, formas, etc.).

En [20] se realiza la integración de imágenes por medio de un subconjunto de superficies, para esto se divide cada imagen 3D en varias porciones representadas por modelos locales de superficie. Luego, se proyectan los modelos locales a la vista de referencia. Los modelos provistos por distintas vistas son fusionados por una función de diferencia y finalmente son triangulados, obteniendo un modelo triangulado que representa a las imágenes de distintos puntos de vistas en un solo sistema de coordenadas dado por la vista de referencia.

### **2.2.2.2 Registro de puntos**

Los métodos que registran los puntos presentes en las imágenes se basan en la minimización de una función de distancia entre los puntos presentes en imágenes de distintos puntos de vista.

Uno de los primeros trabajos en la minimización de funciones de distancia fue presentado por Chen y Medioni [21] y es considerado como el primer acercamiento al algoritmo de ICP (*Iterative Closest Point*), que fue desarrollado por Besl y McKay [22], y que en la actualidad es uno de los algoritmos más utilizados en el registro de imágenes 3D.

El algoritmo ICP realiza un proceso iterativo que, para cada punto de una imagen, realiza la búsqueda del vecino más cercano en la imagen a registrar por medio de una función de minimización de distancias, de esta forma en cada iteración acerca de manera global los dos conjuntos de datos hasta lograr que ambos converjan.

#### **2.2.2.2.1 Etapas del ICP**

Actualmente se han identificado 6 etapas en las que se desarrolla el algoritmo de ICP y que se pueden cambiar para mejorar el proceso. [23]

- 1.- Selección de los puntos (de una o ambas imágenes 3D)
- 2.- Correspondencia de puntos
- 3.- Asignación de peso de correspondencias
- 4.- Rechazo de pares
- 5.- Asignación de la métrica de error
- 6.- Minimización de la métrica error

### 2.2.2.2.1.1 Selección de puntos

La selección de puntos se refiere a muestrear un grupo de puntos que representen la información más distintiva de la imagen, entre los procesos de muestreo más utilizados se encuentran: *muestreo uniforme*, *muestreo aleatorio* y *muestreo en el espacio de normales*. [45]

**Muestreo uniforme.**- Realiza un muestreo con el objetivo de obtener una información equitativa de toda la imagen. La desventaja de este método es que se puede perder información importante para un correcto registro.

**Muestreo aleatorio.**- Realiza un muestreo de puntos al azar. Presenta un resultado similar al muestreo uniforme.

**Muestreo en el espacio de normales.**- Realiza un muestreo aleatorio de puntos de la imagen pero garantizando que habrá una distribución uniforme en las orientaciones de las normales asociadas a los puntos.

### 2.2.2.2.1.2 Correspondencia de puntos

La correspondencia de puntos se refiere a encontrar los puntos más cercanos de una imagen en otra. Una de las variaciones en esta etapa es encontrar puntos más cercanos de mayor compatibilidad que consiste en encontrar los

puntos más cercanos que cumplan con un parámetro de compatibilidad que puede ser orientación, color, distancia, etc.

Otra alternativa es simplemente usar como correspondencia el resultado de la proyección de un punto a la imagen original, lo cual reduce drásticamente el costo computacional. [23]

#### **2.2.2.2.1.3 Asignación de peso de correspondencias**

La asignación de pesos de correspondencias se refiere a asignar un valor de importancia a las correspondencias obtenidas, con el objetivo de disminuir el error por pares mal seleccionados. Los pesos de las correspondencias pueden ser asignados de forma constante, por semejanza en color, compatibilidad de normales, por distancias, entre otras. [23]

#### **2.2.2.2.1.4 Rechazo de pares**

El rechazo de pares se refiere a desechar los pares que no cumplen con un parámetro de correspondencia y que puedan afectar la etapa de minimización del error. Si el registro se lo realiza entre imágenes 3D de buena calidad esta etapa no influye mucho en el desarrollo y optimización del algoritmo. [23]

#### **2.2.2.2.1.5 Asignación y minimización de la métrica de error**

La asignación de la métrica de error se refiere a la definición de la forma en que se medirá el error entre los pares, por lo general se utiliza la suma del cuadrado de las distancias de los pares, mientras que, otra tendencia es utilizar la distancia del punto al plano que es la distancia entre un punto y el plano que es perpendicular a su normal y que pasa a través del otro punto.

Los algoritmos de minimización del error dependen de la métrica de error que se utilice. Cuando se utiliza distancia punto a punto se puede resolver el problema de minimización por descomposición de valores singulares (SVD), cuaterniones [24] o matrices ortonormales [25]. Cuando se utiliza distancia de punto a plano se debe linealizar el problema asumiendo que la variación de la rotación que minimiza el error es cercana a cero.

### **2.2.3 Integración de Imágenes**

La integración se refiere al proceso de encontrar una representación suficientemente aproximada a la escena real, utilizando la información provista en las etapas de adquisición y registro.

El proceso de integración busca eliminar la mayor cantidad de información redundante, de tal manera que cada una de las imágenes contribuya con

información nueva a la formación de una imagen única que defina de manera precisa la representación del ambiente.

La integración de imágenes se realiza principalmente por medio de dos métodos. El primero realiza aplicaciones geométricas para corregir las auto-intersecciones y agujeros en la superficie. El segundo consiste en tomar el conjunto de datos tridimensionales, extraer una o más funciones implícitas y evaluar la función para una serie de puntos en el espacio obteniendo la representación de la superficie de integración de las imágenes.

### **2.2.3.1 Métodos poligonales**

Estos métodos parten del hecho de que puntos espacialmente cercanos dentro de una imagen de rango deben ser conexos y utilizan puntos en el espacio que son triangulados previamente. [44]

La integración por los métodos poligonales se realiza generalmente en tres etapas. En la primera se determina que porciones de una imagen son visibles desde el punto de vista de otra, esto se realiza proyectando cada triángulo de una de las mallas en otra, según el punto de vista de la malla de referencia. Todos los solapes que no se encuentren en los dos puntos de vista se marcan para eliminación. En la segunda etapa se excluyen los triángulos seleccionados para eliminación y en su lugar se introducen triángulos que

hacen continua la representación de la superficie que representa la integración de las imágenes. En la última etapa se realiza un recorrido de los bordes de la malla resultante y se realiza un crecimiento en bordes cercanos cuyo recorrido se realiza en orientaciones casi inversas.

Los métodos poligonales son sensibles a saltos estadísticos, por lo que las mallas a integrar deben ser suaves.

### **2.2.3.2 Métodos basados en funciones implícitas**

El objetivo de utilizar funciones implícitas es obtener una caracterización o función que dé una idea de la localización de la superficie para todo el espacio. La función más utilizadas en este campo es la de distancia signada.

En [17] se realiza la estimación de los parámetros para la estimación de la función implícita a partir de cada imagen de rango y se extrae una función de distancia signada por cada imagen. Luego se conforma una función que resulta de la suma ponderada de las funciones anteriormente mencionadas. El factor de ponderación se calcula con base en la orientación del punto de vista desde el cual fue adquirida cada imagen de rango.

Una vez realizada la estimación de la función se limita el espacio a una caja que encierra todos los puntos. Esta caja se divide en vóxeles (elementos de volumen), y para cada vértice se obtiene el valor de la función calculada.

Los vóxeles que se tienen en cuenta son aquellos en los cuales existe cambio de signo en la función. En estos, se interpola el valor para el cual la función es igual a cero y se asume que existe conectividad entre elementos del mismo vóxel. De esta manera, se obtiene una superficie continua que representa la integración de todas las imágenes.

La principal ventaja de los métodos basados en funciones implícitas es que no son tan sensibles a saltos estadísticos y la desventaja radica en que se debe determinar con cuidado el tamaño de los vóxeles de lo contrario se puede obtener una aproximación similar a una triangulación simple en la nube de puntos con una cantidad excesiva de puntos redundantes o también se puede tener una representación muy pobre en detalles que no sea lo suficientemente próxima a la superficie real de la escena. [44]

### **2.3 ANÁLISIS DE RESULTADOS**

En esta sección se destacan las principales ventajas y desventajas del uso de las técnicas descritas anteriormente. Este análisis permite escoger la mejor

alternativa que servirá como base para la implementación de este proyecto final de carrera.

### Sistemas de adquisición

Ya que se desea realizar un modelo del ambiente, se ha decidido analizar solamente los estimadores basados en la percepción del entorno en relación al sistema de adquisición de datos del robot. Debido a que los estimadores explícitos no proporcionan información del medio que rodea al robot, imposibilitando la adquisición de datos necesarios para crear los modelos de ambientes. Por ésta razón, en la Tabla 2.1 se presenta las principales ventajas y desventajas que tienen los sonares y los sensores de adquisición 3D.

	<b>Ventajas</b>	<b>Desventajas</b>
<b>Sonar</b>	<ul style="list-style-type: none"> <li>➤ Reducido costo</li> <li>➤ Elevada velocidad de respuesta.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Información reducida.</li> <li>➤ Falsas lecturas producen poca precisión</li> <li>➤ Imposibilidad de obtener color</li> </ul>
<b>Escáner Láser</b>	<ul style="list-style-type: none"> <li>➤ Muy buena precisión.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Equipo de muy alto costo</li> <li>➤ Alto tiempo de procesamiento.</li> <li>➤ Si se desea obtener color se depende de una cámara (Sensible a la iluminación)</li> </ul>
<b>Visión 3D</b>	<ul style="list-style-type: none"> <li>➤ Buena precisión</li> <li>➤ Buena velocidad de respuesta</li> </ul>	<ul style="list-style-type: none"> <li>➤ Costos medios.</li> <li>➤ Sensible a la iluminación</li> </ul>

Tabla 2.1 Cuadro comparativo de los sistemas de adquisición

Se descarta el uso del sonar ya que sólo proporciona información 2D y en este proyecto se desea crear modelos en 3D. Para el desarrollo de este proyecto se puede tolerar el uso de un sensor que no tenga una excelente precisión o un alto tiempo de respuesta, por lo que la visión 3D y el escáner láser pueden ser una buena opción. En cuanto al costo económico de los equipos, se desea que sean los más bajos posibles. En la Tabla 2.2 se muestran tres posibles sensores y sus respectivos precios.

<b>Sensor</b>	<b>Descripción</b>	<b>Precio* (\$)</b>
DeltaSphere-3000	Escáner láser color (hw + sw)	43 690
Optix 400 series	Escáner láser color (hw + sw)	35 780
Digiclops color	Cámara estéreo color (hw + sw)	6 450
* Precio cotizado en USA		

Tabla 2.2 Precios de sensores de adquisición 3D

Como se puede apreciar en la Tabla 2.2, la diferencia de precio entre un escáner láser y una cámara estéreo es muy significativa, siendo el factor económico uno de los principales puntos a tomar en cuenta en este proyecto se ha seleccionado como sistema de adquisición a la Visión 3D.

### **Registro del movimiento**

En la Tabla 2.3 se presenta las principales ventajas y desventajas que tienen los métodos de registro.

	<b>Ventajas</b>	<b>Desventajas</b>
<b>Características</b>	➤ Alta precisión en fase de correspondencia.	➤ Alto tiempo de pre-procesamiento
<b>Puntos (ICP)</b>	➤ Alta cantidad de variantes ➤ Simpleza	➤ Impreciso para movimientos muy grandes

Tabla 2.3 Cuadro comparativo de los métodos de registro

Debido a la simpleza y a la cantidad de variantes que permiten una mayor adaptabilidad, se ha decidido realizar el registro de puntos por medio del ICP, tomando en cuenta que los desplazamientos entre cada par de adquisiciones deben ser pequeños ya que el algoritmo de ICP es impreciso para movimientos muy grandes.

### **Integración de imágenes**

En la Tabla 2.4 se presenta las principales ventajas y desventajas que tienen los métodos de integración de imágenes.

	<b>Ventajas</b>	<b>Desventajas</b>
<b>Poligonales</b>	➤ Buena eliminación de puntos redundantes	➤ Muy sensibles a saltos estadísticos
<b>Funciones implícitas</b>	➤ No tan sensibles a saltos estadísticos	➤ Dependiendo del tamaño de vóxeles puede crear mala integración y exceso de puntos redundantes

Tabla 2.4 Cuadro comparativo de los métodos de integración de imágenes

Se ha decidido prescindir de la etapa de integración para crear el modelo del ambiente de navegación del robot debido a que la navegación se puede realizar con modelos que aproximen de una forma “gruesa” a los ambientes y no necesariamente con una aproximación “fina” de alta exactitud, lo que permite realizar solamente una alineación de los datos adquiridos.

## **2.4 CONCLUSIONES**

Una vez realizado el estudio bibliográfico sobre algunas de las alternativas existentes y con la investigación inicial suficiente se recomienda tomar las siguientes acciones:

- Para el desarrollo de este proyecto se recomienda utilizar un robot con un sistema de odometría de alta precisión para poder realizar posteriores comparaciones con el sistema de localización implementado.
  
- Para el sistema de adquisición se ha decidido utilizar un sensor de visión 3D debido a su gran diferencia en costo ante el escáner láser y que, a diferencia del sonar, provee información tridimensional y a color del ambiente.

- La etapa de registro se la va a realizar a través del algoritmo de ICP debido a su simpleza y a la gran variedad de información disponible sobre el mismo.
  
- Se recomienda aprovechar las ventajas del sensor de visión haciendo uso del color en la etapa de correspondencia de puntos o de rechazo de pares durante la implementación del algoritmo de ICP.
  
- En la etapa de selección de puntos del algoritmo de ICP se recomienda el uso de un muestreo uniforme, para reducir el número de puntos a procesar y mejorar el costo de procesamiento.

## **CAPÍTULO III**

### **SISTEMA DE ADQUISICION 3D**

El sistema de adquisición 3D es una parte de vital importancia en el desarrollo de este proyecto, por lo tanto en el presente capítulo se presenta una breve introducción en la que se resume el proceso seguido en el sistema de adquisición. El sistema de adquisición tiene como principal meta el adquirir imágenes tridimensionales del ambiente de navegación del robot con la idea de crear mapas completos de ambientes mediante el sistema de registro de movimiento e integración de imágenes. Las bases del sistema de visión estéreo y de la cámara utilizada en el desarrollo del proyecto son presentados en este capítulo. Adicionalmente, se realiza un análisis del procedimiento seguido para la obtención de imágenes tridimensionales a partir de un conjunto de imágenes bidimensionales adquiridas por la cámara utilizada. Finalmente, algunos resultados experimentales que se obtuvieron en varios recorridos realizados con un robot de prueba son presentados al final del capítulo.

#### **3.1 INTRODUCCIÓN**

Para lograr que un robot móvil pueda navegar de forma autónoma el primer requisito y el de mayor importancia es que pueda obtener información del

medio que lo rodea, el cual es tridimensional, por lo que la información adquirida debe ser también tridimensional.

Para poder adquirir información tridimensional se han propuesto varios métodos mediante el uso de sensores de escáner láser y sistemas estéreo. El uso de los sensores de escáner láser permite obtener información tridimensional con una alta precisión, pero el costo de los equipos y el tiempo de procesamiento son bastante altos, tal como se explicó en el capítulo anterior. Por otro lado, los sistemas estéreo tienen un costo y el tiempo de procesamiento mucho menor, aunque no tienen una gran exactitud como los sensores de escáner láser. De esto último, sin embargo para realizar aplicaciones de navegación no se requiere un modelo exacto de los objetos existentes en el ambiente, sino que basta con adquirir imágenes con una buena aproximación que sirvan como referencia al robot.

Debido a que la navegación de robots se puede realizar con modelos que aproximen de una forma “gruesa” a los ambientes y no necesariamente con una aproximación “fina” de alta exactitud, y a motivos de costo y tiempo de procesamiento, se ha optado por el uso de un sistema de visión estéreo para realizar el sistema de adquisición 3D del robot, tal como se explicó en el capítulo anterior.

Este capítulo describe las principales características de un sistema de visión estéreo y su geometría. Además, detalla el proceso de adquisición de una imagen de profundidad y la generación de una nube de puntos que representa el ambiente. Al final del capítulo se muestran algunos resultados experimentales y las conclusiones obtenidas.

### **3.2 SISTEMA DE VISION ESTEREO**

Los sistemas de visión estéreo, trabajando sobre un ambiente, calculan la profundidad de cualquier punto sobre una superficie con respecto a un plano predefinido. Para esto, toman como base la diferencia de posición de la proyección del mismo punto en dos o más imágenes.

En el proceso de la visión estéreo destacan dos aspectos importantes, el establecer la relación de puntos de una y otra imagen denominado *correspondencia* y el calcular la profundidad de los puntos. Ambos aspectos se pueden determinar a través de la geometría estéreo o geometría epipolar [26].

En la geometría epipolar es posible fijar una línea denominada línea epipolar para cada punto de una imagen y garantizar que el equivalente de un punto en una imagen se encuentra sobre la línea epipolar de la otra imagen.

### 3.2.1 Geometría epipolar

La Figura 3.1 muestra dos cámaras, sus centros de proyección,  $C_i$ ,  $C_d$ , y sus planos imagen  $I_i$ ,  $I_d$  en coordenadas normalizadas.

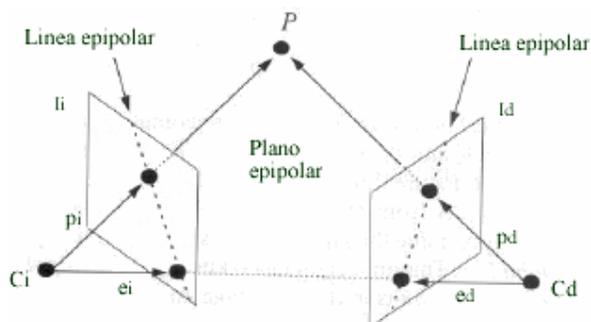


Figura 3.1.- Geometría de un sistema estéreo.

Se denominan epipolos a los puntos de los planos imagen que pertenecen a la recta que une los centros de proyección de las cámaras. En la Figura 3.1 los epipolos se denotan como  $e_i$  y  $e_d$ .

Dado cualquier punto  $P$  en el espacio, se denomina a su proyección en el plano imagen izquierda como  $p_i$  el cual se encuentra sobre la línea creada por  $P$  y  $C_i$ .

La proyección de la recta  $P$ ,  $C_i$  en el plano imagen derecha es la línea que une el epipolo derecho y la proyección de  $P$  en el plano imagen derecha,  $p_d$ , esta línea es denominada línea epipolar. La línea epipolar puede ser definida como

la intersección del plano formado por  $P$ ,  $C_i$ ,  $C_d$ , denominado plano epipolar, y el plano imagen [26].

Al ser la línea epipolar la representación en el plano imagen de la línea formada por un punto en el espacio y el centro de proyección de una cámara se puede garantizar que un punto en una imagen tiene su equivalente sobre la línea epipolar de la otra imagen. Esta implicación es de gran importancia ya que limita la búsqueda de una correspondencia en un plano a una búsqueda a lo largo de una recta. Así, en el caso más desfavorable, el número de puntos candidatos a comprobar en una imagen de tamaño  $N \times N$  es de  $\sqrt{2} N$ . [26]

### **3.2.1.1 Cámara de estereo visión Digiclops**

Para el desarrollo del presente proyecto se ha hecho uso de una cámara de estereo visión, modelo *Digiclops*, fabricada por *Point Grey Research Inc* [27]. Un módulo integrado por 3 cámaras que se encuentran distribuidas en forma de un triángulo rectángulo es usado por esta cámara, tal como se muestra en la Figura.3.2.

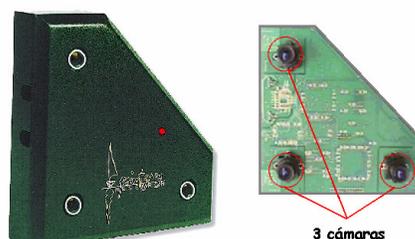


Figura 3.2.- Sistema de visión estéreo constituida por 3 cámaras:  
marca *Digiclops*.

La cámara Digiclops realiza un proceso de estéreo en dos sentidos, horizontal y vertical, por medio de una combinación de las cámaras izquierda, derecha y superior, derecha. Luego combina los resultados obteniendo un menor error.

Los planos imagen de las tres cámaras pertenecientes al módulo Digiclops son paralelos a las rectas que unen los centros de proyección, y además se encuentran alineados entre si, por lo que el epipolo de cada plano esta situado en el infinito. En este caso las líneas epipolares de dos imágenes se encuentran sobre la misma recta, tal como se aprecia en la Figura 3.3.

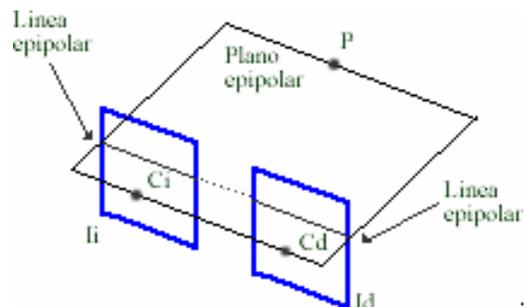




Figura 3.3.- (*Superior*) Geometría epipolar para la cámara; (*inferior*) líneas epipolares para las proyecciones del punto  $P$  –  $P_d$ ,  $P_i$  y  $P_s$  – en las 3 imágenes adquiridas con la cámara: derecha, izquierda y superior.

### 3.2.1.2 Correspondencia y profundidad de puntos

La correspondencia de imágenes tiene como objetivo establecer una relación entre los puntos en dos imágenes de una misma escena. En el presente proyecto se ha empleado la *librería Triclops SDK* [28] que utiliza un método de correlación.

Los métodos de correlación utilizan ventanas de tamaños fijos y un criterio de similitud que es una medida de semejanza entre las ventanas, la relación entre las imágenes esta dada por la ventana que maximiza el criterio de similitud dentro de una región de búsqueda

Para cada píxel de una imagen se calcula la correlación entre los valores de intensidad de una ventana centrada en dicho píxel y una ventana del mismo tamaño centrada en el píxel a analizar de la otra imagen. El problema consiste en encontrar el punto que se ajusta de manera adecuada al primero [29]

En el caso de las imágenes adquiridas con la cámara Digiclops, la búsqueda del correspondiente a un punto en la imagen derecha (referencia), se realiza a lo largo de la misma fila y de la misma columna de la imagen izquierda y superior respectivamente, siguiendo las líneas epipolares (Figura 3.3). La diferencia absoluta entre las posiciones de los píxeles de dos imágenes que corresponden a un mismo punto se denomina disparidad.

Así, por cada píxel en una imagen se toma la vecindad cuadrada de un tamaño determinado, se la compara con un número de vecindades en la otra imagen y se selecciona la mejor correspondencia. La comparación de las vecindades se la realiza mediante la suma de diferencias absolutas, por medio de la siguiente ecuación [28]:

$$d_{\min}^{d_{\max}} \sum_{i=m/2}^{m/2} \sum_{j=m/2}^{m/2} |I_{der}[x+i][y+j] - I_{izq}[x+i+d][y+j]| \quad \text{Ec. 3.1}$$

**Donde:**  $d_{\min}$  y  $d_{\max}$  son las disparidades mínima y máxima,  $m$  es el tamaño de la ventana,  $I_{der}$  y  $I_{izq}$  son las imágenes derecha e izquierda.

Los parámetros de las cámaras utilizadas, así como la configuración del sistema estéreo delimitan los posibles valores de profundidad a obtener. En la Figura 3.4 se observa como la distancia entre las cámaras, denominada línea base,  $b$ , y el ángulo de visión de las cámaras limitan el valor mínimo de profundidad y por tanto el máximo de disparidad que el sistema es capaz de proporcionar [26].

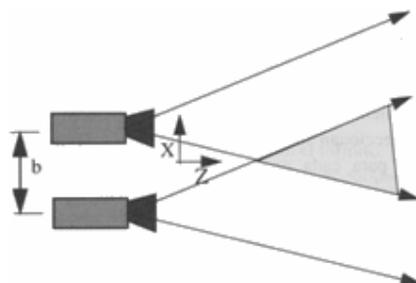


Figura 3.4.- Zona donde se puede obtener profundidad representada en color gris

Una imagen de profundidad o imagen de disparidad puede entenderse como la representación 2D de las distancias existentes entre los puntos en el espacio y un plano de referencia que, en este caso, se encuentra en la cámara. Estas imágenes expresan la distancia por medio de un valor entero por lo que puede ser visualizada como un valor de color.

La Figura.3.5 (*derecha*) muestra una imagen de profundidad donde la distancia está representada por niveles de gris y ha sido obtenida después de aplicar algunos filtros de validación como unicidad y semejanza de textura, mientras que su correspondiente imagen a color es dada en la Figura. 3.5 (*izquierda*).

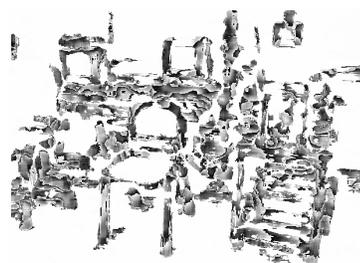


Figura 3.5.- (izquierda) Imagen a color del ambiente de trabajo; (derecha) Correspondiente imagen de disparidad.

Las ecuaciones básicas para la obtención de profundidad se basan en las relaciones de semejanzas entre triángulos. En la Figura 3.6 se muestra el sistema de coordenadas para las cámaras derecha e izquierda de un sistema de visión estéreo y se muestra un punto en el espacio de coordenadas  $(X, Y, Z)$  y sus proyecciones en las imágenes derecha e izquierda denotadas por  $(x_i, y_i)$  y  $(x_d, y_d)$  respectivamente [26].

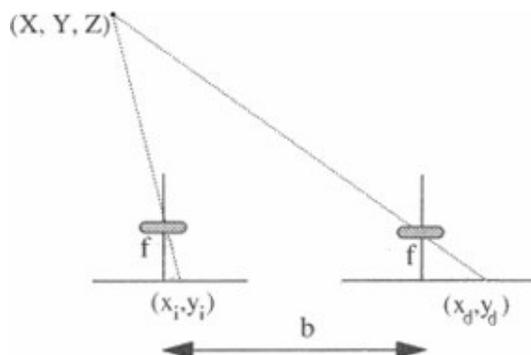


Figura 3.6.- Sistema de coordenadas para la cámara estéreo

Aplicando semejanza de triángulos con base a la distancia focal  $f$  se tienen las siguientes ecuaciones:

$$\frac{x_i}{f} = \frac{X + b}{Z} \quad \text{Ec. 3.2}$$

$$\frac{x_d}{f} = \frac{X}{Z} \quad \text{Ec. 3.3}$$

$$\frac{y}{f} = \frac{Y}{Z} \quad \text{Ec. 3.4}$$

Resolviendo las tres ecuaciones anteriores se obtiene:

$$X = \frac{b \cdot xd}{(xi - xd)} \quad \text{Ec. 3.5}$$

$$Y = \frac{b \cdot y}{(xi - xd)} \quad \text{Ec. 3.6}$$

$$Z = \frac{b \cdot f}{(xi - xd)} \quad \text{Ec. 3.7}$$

Donde  $(xi - xd)$  es el valor de disparidad  $d$ , por lo que se pueden re-escribir las relaciones anteriores como:

$$X = \frac{b \cdot xd}{d} \quad \text{Ec. 3.8}$$

$$Y = \frac{b \cdot y}{d} \quad \text{Ec. 3.9}$$

$$Z = \frac{b \cdot f}{d} \quad \text{Ec. 3.10}$$

Como se puede apreciar en las ecuaciones, con la información geométrica de las cámaras y con el valor de disparidad se puede obtener la distancia a la que se encuentra un punto determinado de la cámara. Para el presente proyecto este cálculo ha sido desarrollado por medio de la librería Triclops SDK [28], para todos los valores de disparidad de una imagen de disparidad obteniendo así una *nube de puntos*. La librería Triclops SDK ha sido utilizada ya que contiene la información de las características geométricas de la cámara estéreo utilizada.

Una nube de puntos es un conjunto de puntos tridimensionales que representa a los objetos existentes en el espacio. De esta forma, la imagen  $I$  representada por una nube de  $n$  puntos es obtenida como un listado de puntos  $P$  tal como indica la Ec. 3.11. La Figura. 3.7 muestra la nube de puntos para la imagen de la Figura 3.5 [30].

$$I = (P_1, P_2 \dots P_n); P_i = (X_i, Y_i, Z_i) \quad \text{Ec. 3.11}$$

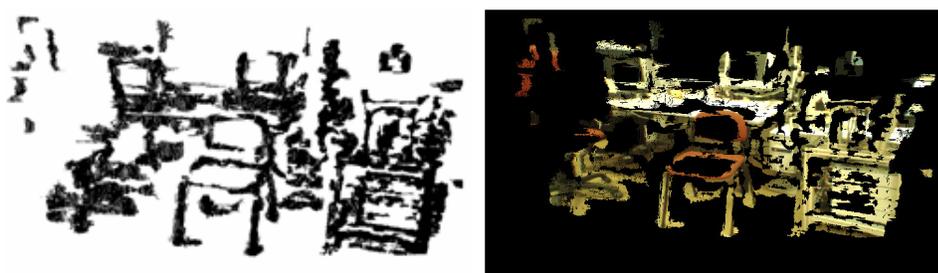


Figura 3.7.- Nube de puntos: (izquierda) imagen binaria; (derecha) imagen a color.

### 3.3 RESULTADOS

El desarrollo de las pruebas experimentales de adquisición para este proyecto se ha llevado a cabo en el *Laboratorio de Visión por Computador - Vislab* del *Instituto de Sistemas e Robótica – ISR*, perteneciente a la *Universidade Técnica de Lisboa* (Portugal), como parte de las actividades de los Proyectos de Investigación: “Programación Asistida de Robots para Tareas Industriales (PARTI)” y de la “Red Iberoamericana de Robótica (RIBERO)”, financiados por la Agencia Española de Cooperación Internacional (AECI) a través del

Programa CYTED (Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo), Subprograma VII: *Electrónica e Informática Aplicadas*, en los que el *Centro de Visión y Robótica* de la *ESPOL* participa.

El robot utilizado para realizar las pruebas experimentales fue el *TRC Labmate* de *HelpMate Robotics Inc*, el cual consiste en una plataforma móvil de control diferencial al que se le ha montado la cámara de estereo visión *Digiclops*. (Figura 3.8).

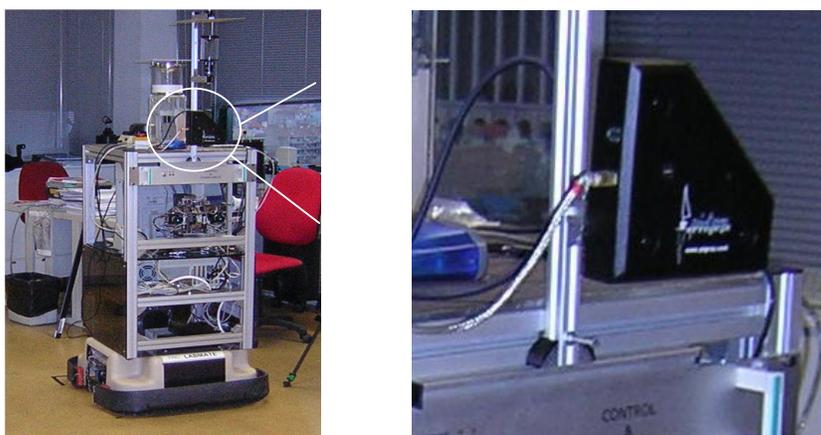


Figura 3.8.- Sistema de visión estereo montado sobre robot móvil.

Se realizaron varios recorridos con el robot, donde se adquirieron una serie de imágenes rectificadas de las cámaras izquierda, derecha y superior. Cada conjunto de imágenes fue adquirido aproximadamente cada segundo, luego se procedió a realizar un procesamiento fuera de línea para obtener las imágenes de disparidad y las respectivas nubes de puntos. La cantidad de puntos resultantes en cada nube de punto obtenida varía entre 40000 y 90000.

Para la obtención de las nubes de puntos se han realizado filtros de unicidad, detección de bordes, semejanza de textura y comprobación de superficie con el objetivo de obtener resultados de mayor robustez y en base a que se necesita obtener una referencia tridimensional de las posiciones de los objetos para crear los mapas del ambiente que permitirán una buena aproximación para futuros desarrollos de navegación y no se desea recrear un modelo exacto del ambiente. Algunos resultados son mostrados en la Figura. 3.9.

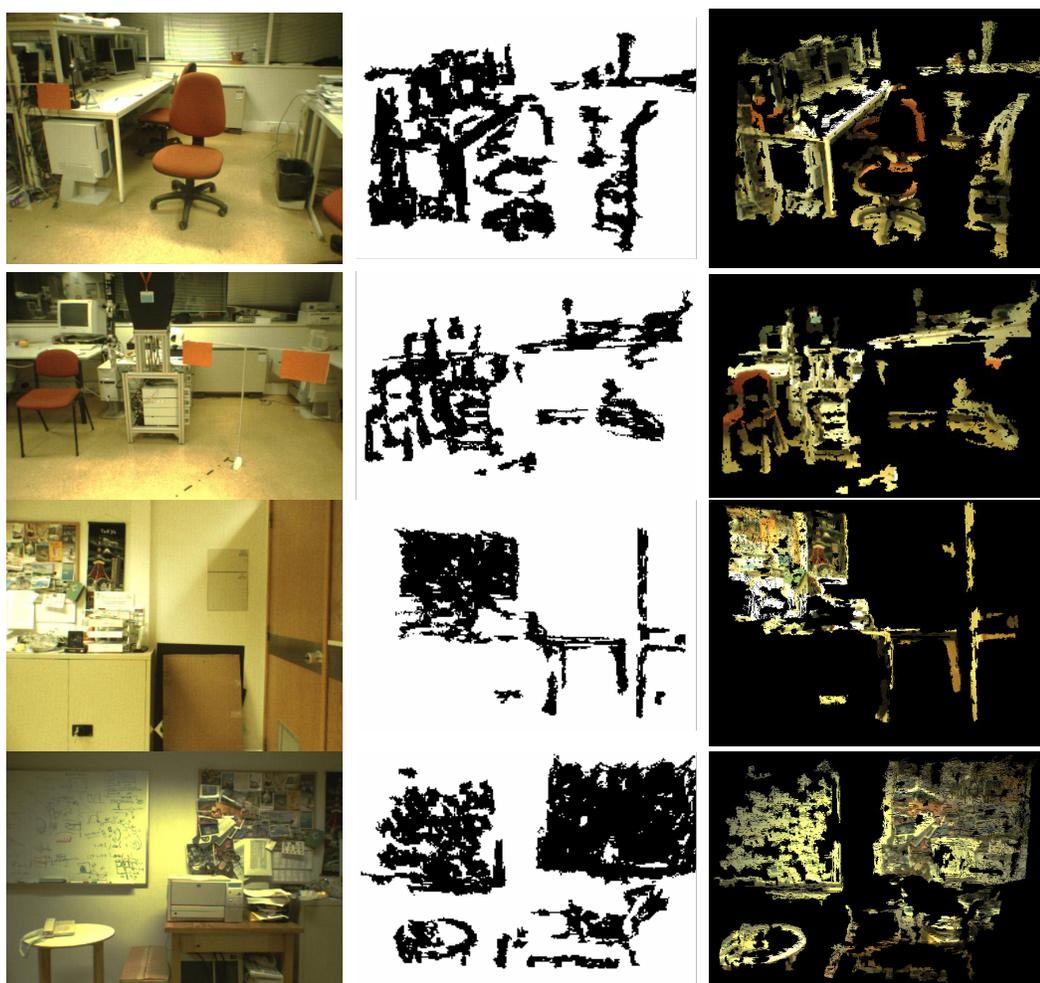


Figura 3.9.- (columna izquierda) Imágenes a color del ambiente; (columna central) Nubes de puntos B/N; (columna derecha) Nubes de puntos colores

### 3.4 CONCLUSIONES

El presente capítulo pretende solucionar el problema de la adquisición de imágenes 3D del ambiente de navegación de un robot. El objetivo es proporcionar al robot información necesaria para que pueda obtener su ubicación en el ambiente (localización) y crear mapas del mismo.

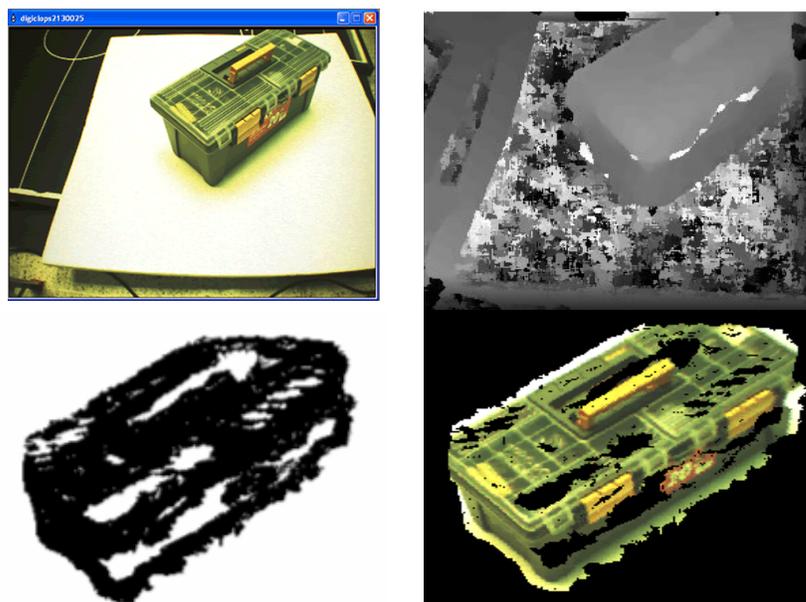
En respuesta a este problema se ha realizado la implementación de un sistema de adquisición mediante visión estéreo con el uso de una cámara de visión estéreo, modelo Digiclops, fabricada por Point Grey Research Inc [27]. Por medio de esta cámara se han adquirido una serie de imágenes mientras el robot realizaba un recorrido determinado, teniendo la ventaja de poder realizar adquisiciones de una forma rápida sin tener la restricción de detener el movimiento del robot, como sucedería en el caso de usar un escáner láser que no opere a una muy alta frecuencia. [46].

Finalmente un conjunto de pruebas experimentales al sistema de adquisición montado sobre el robot TRC Labmate fue realizado. Con estas pruebas se obtuvieron las nubes de puntos necesarias para realizar las siguientes fases en el presente proyecto.

### 3.4.1 Trabajos Paralelos

Además del desarrollo y los experimentos descritos anteriormente se realizó un sistema de modelado 3D de superficies [30], tomando como ventaja las experiencias obtenidas en el desarrollo del presente proyecto.

El sistema de modelado 3D de superficies tiene como objetivo el crear en el computador un modelo tridimensional de una superficie de un objeto real para lo cual se utiliza una malla triangular. El procedimiento para obtener la nube de puntos para el sistema de modelado 3D de superficies es el mismo que se describe en el presente capítulo y que se resume en las filas superior y central de la Figura 3.10.



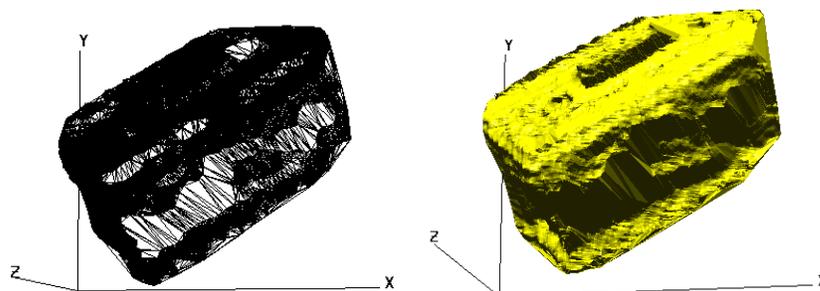


Figura 3.10.- Procedimiento para la creación del modelo 3D de una superficie: (*superior izquierda*) Imagen a color de objeto de interés; (*superior derecha*) Imagen de profundidad; (*fila central*) Nube de puntos: b/n y color; (*inferior izquierda*) Malla triangular; (*inferior derecha*) *rendered* de las mallas resultantes.

La malla triangular se genera usando un algoritmo público de Delaunay 2D [31]. La triangulación se realiza primero proyectando la nube de puntos al plano XY y luego triangulando estos en este espacio. La malla triangular resultante es mapeada al espacio 3D reasignando el valor de la coordenada Z a cada punto previamente proyectado. En la literatura este proceso es conocido como triangulación 2½D. La Figura. 3.11 ilustra este procedimiento.

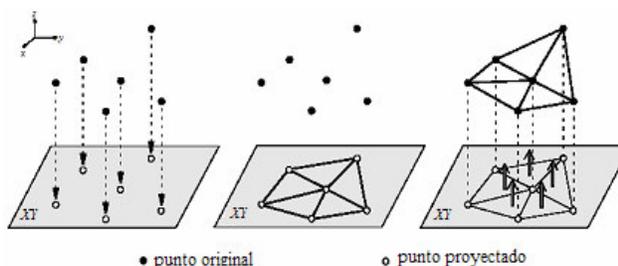


Figura 3.11.- Ilustración de un proceso de triangulación 2½D: (*izquierda*) Proyección de nube de puntos al plano XY. (*medio*) Triangulación 2D de los puntos proyectados. (*derecha*) mapeo de la triangulación 2D al espacio 3D original.

El “*rendered*” se refiere a la imagen final a mostrar (incluye características de luz, textura, color y sombra), la Figura. 3.12 muestra el *rendered* de dos

objetos diferentes obtenidos mediante el sistema de modelado 3D de superficies.[47]

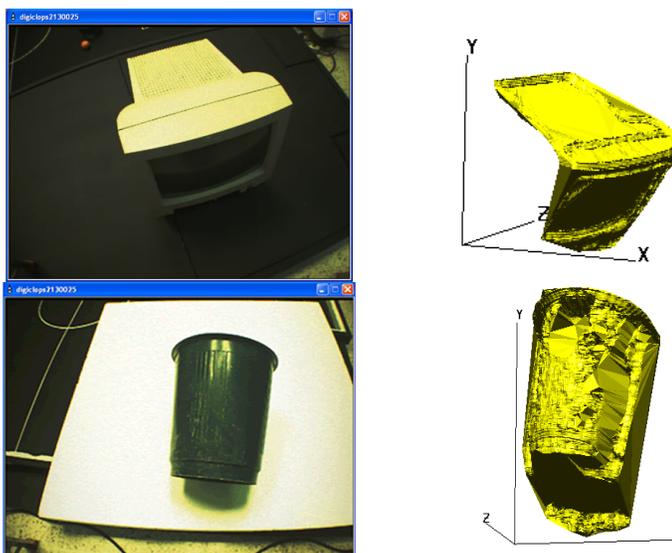


Figura 3.12.- Modelado 3D para objetos reales: (*columna izquierda*) objetos de interés; (*columna derecha*) *rendered* de las mallas resultantes.

# **CAPÍTULO IV**

## **LOCALIZACIÓN Y MAPEO 3D**

Este capítulo presenta una breve introducción en el que se resume el proceso seguido para la localización de un robot móvil en su ambiente de navegación y en la creación de mapas 3D de los ambientes navegados. La localización y mapeo (mapping) tienen como principal meta el crear un mapa tridimensional del ambiente de navegación del robot con la idea de proporcionar una herramienta eficaz para la navegación autónoma del robot. Un análisis del procedimiento seguido para la obtención de la localización del robot a partir de un conjunto de imágenes es presentado en este capítulo. Además, se muestra el proceso de creación de mapas 3D a partir de las distintas localizaciones del robot y de las imágenes 3D obtenidas. Finalmente, algunos resultados experimentales que se obtuvieron en varios recorridos realizados con un robot de prueba son presentados al final del capítulo.

### **4.1 INTRODUCCIÓN**

El proceso de construcción de una imagen 3D de un ambiente real involucra dos fases, la primera fase consiste en la adquisición del medio (escaneado 3D) y la segunda es una reconstrucción de los datos adquiridos [32].

En la fase de adquisición se obtiene información del ambiente desde varios puntos de vista por lo que es necesario realizar una reconstrucción de todos los datos adquiridos para formar un solo modelo o mapa que represente al ambiente completo. Para esto, se realiza el registro de todos los datos obtenidos, el registro es el proceso de alinear espacialmente los datos de dos o más muestras que pertenecen a una misma escena [19].

Una vez alineados todos los datos en un solo modelo, se realiza un proceso de integración de datos en el cual se eliminan puntos redundantes y se integran puntos similares que pueden pertenecer a un mismo objeto, para crear así una mejor aproximación al ambiente real y reducir errores en el modelo. Debido a que la navegación de robots se puede realizar con modelos que aproximen de una forma “gruesa” a los ambientes y no necesariamente con una aproximación “fina” de alta exactitud, se ha optado por prescindir de esta etapa y realizar solamente una alineación de todos los datos adquiridos por el robot.

En este capítulo se describen las principales características de los sistemas de registro de imágenes y alineación de imágenes detallando los pasos realizados para obtener la matriz de transformación entre un par de imágenes y la alineación de un grupo de imágenes que representan un ambiente. Al final del capítulo se muestran algunos resultados experimentales y las conclusiones obtenidas.

## 4.2 REGISTRO DE IMÁGENES 3D

Los métodos de registro de imágenes 3D se pueden clasificar en dos grupos principales: métodos que registran un grupo de características extraídas de las imágenes, y métodos que registran los puntos presentes en las imágenes [17]. El objetivo en ambos casos es encontrar la matriz de rotación y el vector de traslación que aplicados a una nube de puntos, nos permitan alinearla con la otra.

Para el presente proyecto se ha decidido realizar el registro a través del algoritmo de ICP, que registra directamente los puntos presentes en las imágenes, debido a su simpleza en relación al método de registro de un grupo de características extraídas de las imágenes.

### 4.2.1 Registro del Movimiento. (Sistemas de Localización 3D)

El algoritmo ICP realiza un proceso iterativo que, para cada punto de una imagen, realiza la búsqueda del vecino más cercano en la imagen a registrar por medio de una función de minimización de distancias. De esta forma, en cada iteración acerca de manera global los dos conjuntos de datos hasta lograr que ambos converjan.

Como se describió en el capítulo 2, existen 6 etapas en las que se desarrolla el algoritmo de ICP y que se pueden cambiar para mejorar el proceso [23]. En el presente proyecto se aplicaron 4 de los pasos mencionados siguiendo un procedimiento basado en [33], en donde se realiza un registro de puntos tridimensionales adquiridos por medio de un escáner láser, la correspondencia de los puntos se la realiza a través del vecino mas cercano y la minimización por medio de SVD, pero no se realizan los pasos de selección de puntos, rechazo de pares o asignación de peso de correspondencia ya que las nubes de puntos generadas por el escáner láser ofrecen poco ruido permitiendo obviar esas fases.

#### **4.2.1.1 Selección de puntos**

Para disminuir la cantidad de puntos a procesar se ha decidido realizar un muestreo de puntos de forma uniforme con el objetivo de obtener una información equitativa de toda la imagen. Al realizar este muestreo se espera disminuir el tiempo de procesamiento de las siguientes etapas, aunque no se minimiza el error existente en las imágenes 3D. En la Figura 4.1 (fila superior) se muestra una imagen 3D representada a través de una nube de puntos, tanto a color como en blanco y negro, mientras en la fila inferior se aprecia la misma nube de puntos después de aplicar una selección uniforme de puntos

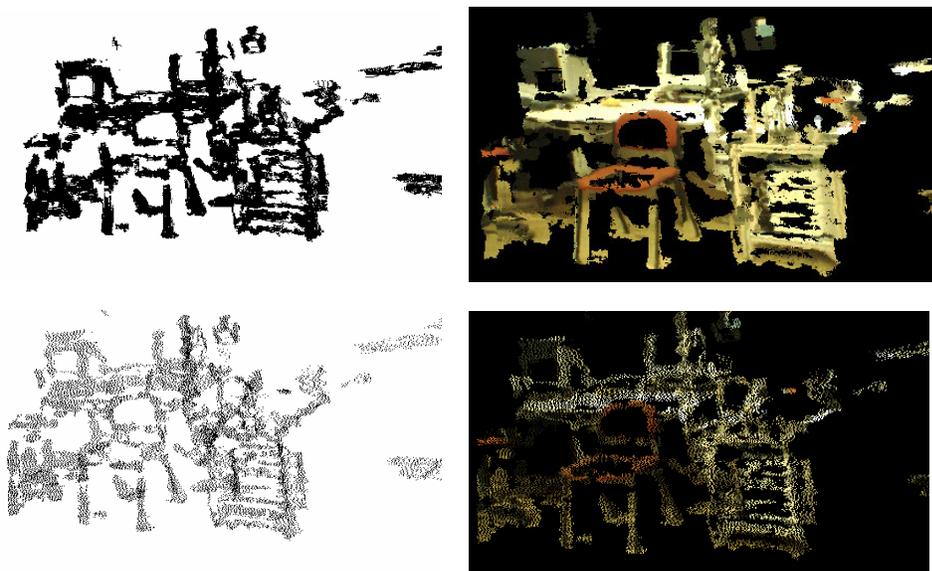


Figura 4.1.- (*fila superior*) Nube de puntos original; (*fila inferior*) Nube de puntos después de selección uniforme de puntos

#### 4.2.1.2 Correspondencia de Puntos

La correspondencia se realiza a través del vecino más cercano. Para el desarrollo del problema del vecino más cercano se tiene como entrada un grupo de puntos en el espacio de dimensión  $d$ , en el caso tridimensional  $d = 3$ . Luego, el grupo de puntos es organizado en una estructura de forma tal que dado un punto cualquiera el punto más cercano pueda ser encontrado con facilidad. Dependiendo de la aplicación, la distancia entre dos puntos puede ser definida de varias formas, como pueden ser distancia euclídea o distancia Manhattan.

Las estructuras de datos que más se utilizan en el desarrollo del problema del vecino más cercano son los árboles kd (k-dimensionales). Un árbol kd es un

árbol binario de búsqueda en el que cada nodo representa una partición del espacio de dimensión  $d$ , para crear las divisiones en el espacio en cada nodo sólo una de las  $k$  dimensiones es utilizada como discriminador. El nodo raíz representa el espacio entero y las hojas representan subespacios que contienen subconjuntos del grupo de puntos de entrada, los cuales son mutuamente excluyentes.

Se ha utilizado la librería ANN [34] para realizar la creación y búsqueda del árbol  $kd$ , el cual es creado recursivamente dividiendo el conjunto de puntos en la dimensión de mayor varianza, hasta obtener hojas con un determinado número de puntos.

El proceso de búsqueda opera recursivamente, cuando se encuentra un nodo del árbol se visita el nodo hijo que se encuentra más cerca al valor del punto consultado en la dimensión discriminante del nodo. Si el otro nodo hijo se encuentra a una distancia dentro de un porcentaje de error determinado de la distancia del vecino cercano encontrado hasta el momento, entonces el otro nodo hijo también es visitado de forma recursiva. Cuando se encuentra una hoja todos los puntos de esa hoja son evaluados en relación al punto consultado y se determina el punto más cercano.

### 4.2.1.3 Rechazo de pares

Para realizar el rechazo de pares que no representan una correspondencia válida se utiliza como filtro el color de los puntos en las imágenes 3D.

Cuando se trabaja con imágenes de color existen grandes variaciones para representar el color, las dos más utilizadas son: RGB y HSI. Por medio de la representación RGB se descompone la imagen de color en 3 imágenes correspondientes a los componentes de color rojo, verde y azul (Red, Green, Blue), es la representación de mayor uso en el mundo informático y se encuentra orientado a los equipos. El modelo RGB puede representarse como un cubo cuyos ejes son los componentes de color, tal como se muestra en la Figura 4.2. [48]

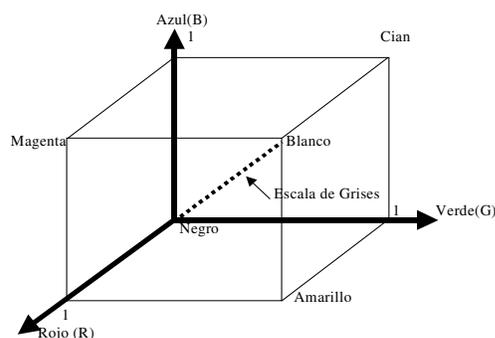


Figura 4.2.-Espacio de color RGB.

El problema que existe con la representación RGB es que los seres humanos no describimos un determinado color considerando los componentes RGB, sino mediante otras tres características que son tono, saturación y brillo. Estas

tres componentes forman la representación HSI del color, la cual se asocia a la forma en que percibimos el color. El modelo HSI puede representarse como dos conos unidos en su base tal como se aprecia en la Figura 4.3. [26]

Las tres componentes que utiliza el modelo HSI son descritas a continuación. El tono (Hue) indica el color, en la Figura 4.3 esta representado por el ángulo. La saturación (Saturation) muestra la pureza de un color determinado, lo que es igual al radio en la Figura 4.3. La intensidad o brillo (Intensity) es la iluminación del color, aparece como la altura en el cono de color de la Figura 4.3 [26].

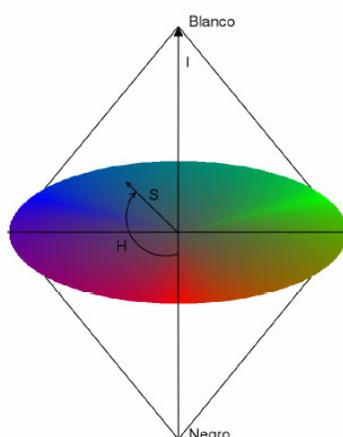


Figura 4.3.-Espacio de color HSI.

A diferencia del modelo de color RGB, el espacio HSI separa la iluminación de la información de color, esto es, el tono (H) y la saturación (S) permanecen estables bajo variaciones en la intensidad de la iluminación, [35], permitiendo realizar un filtrado más robusto al ignorar la componente I, que representa de la iluminación de la escena.

La relación existente entre RGB y HSI se muestra en las siguientes ecuaciones [26]:

$$H = a \cos \left( \frac{(1/2)((R-G) + (R-B))}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right) \text{ o } (2\pi - H) \text{ si } B > G \quad \text{Ec. 4.1}$$

$$S = 1 - \frac{3 \min(R, G, B)}{I} \quad \text{Ec. 4.2}$$

$$I = R + G + B \quad \text{Ec. 4.3}$$

La conversión de un modelo de color a otro es un proceso muy costoso computacionalmente ya que se deben realizar tantos cálculos como número de puntos existan en la imagen 3D, por lo que se ha optado por realizar una tabla de búsqueda LUT (Look Up Table), la cual se almacena en la memoria del computador con las  $2^{24}$  combinaciones de todos los valores de tono y saturación posibles para cada componente RGB utilizando 8 bits.

Una vez que se han obtenido las correspondencias entre dos imágenes 3D se comparan los valores de H y S mediante una operación de lectura de la LUT y se evalúan en relación a la media de todos los puntos de la imagen, si la diferencia de los valores excede la media se descarta el par como una correspondencia, de esa forma se realiza el registro sólo con las correspondencias más robustas. En la Figura 4.4 se aprecia una nube de puntos después de pasar por la fase de rechazo de pares a través del método de color descrito anteriormente.

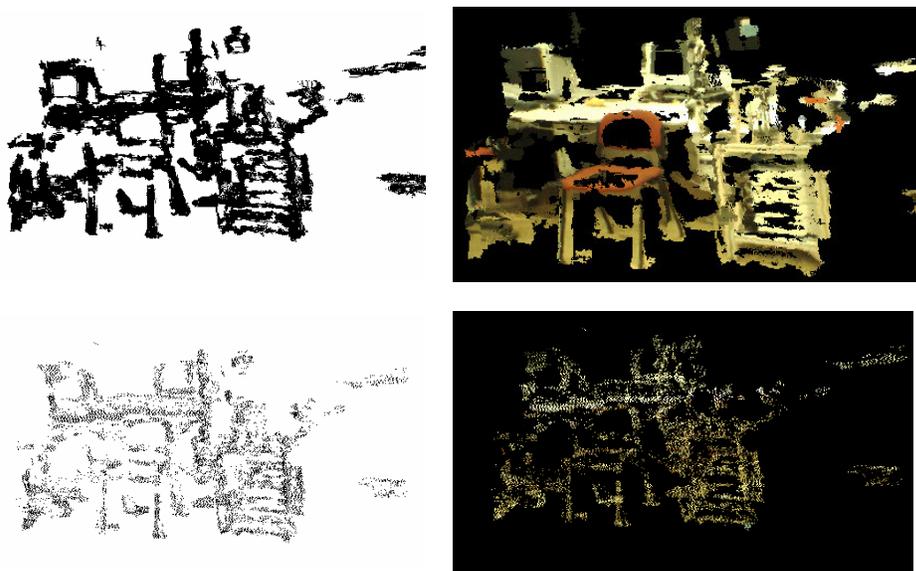


Figura 4.4.- (fila superior) Nube de puntos original; (fila inferior) Nube de puntos después del rechazo de pares.

#### 4.2.1.4 Asignación y minimización de la métrica de error

La asignación de la métrica de error se refiere a la definición de la forma en que se medirá el error entre los pares. En el presente proyecto se utiliza la suma del cuadrado de las distancias. La finalidad es encontrar la transformación, compuesta por Rotación y Translación, que minimiza la siguiente función: [49]

$$F = \sum_{i=0}^N \|p_i - (R \cdot q_i + T)\|^2 \quad \text{Ec. 4.4}$$

Donde  $\{p\}$  y  $\{q\}$  son los conjuntos de pares correspondientes entre las dos imágenes 3D y R, T son las matrices de Rotación y Translación respectivamente

Para hallar la matriz de transformación se ha empleado el método basado en la descomposición de valores singulares (SVD) propuesto en [36].

El primer paso para hallar la matriz de transformación es separar la Rotación,  $R$ , de la matriz de Translación,  $T$ , para lo cual se utilizan los centros de masas o centroides de  $\{p\}$  y  $\{q\}$  mediante las siguientes ecuaciones: [49]

$$Cp = \frac{1}{N} \sum_{i=1}^N p_i \quad \text{Ec. 4.5}$$

$$Cq = \frac{1}{N} \sum_{i=1}^N q_i \quad \text{Ec. 4.6}$$

Una vez que se han encontrado los centroides,  $Cp$  y  $Cq$ , se representan los dos grupos de puntos (imágenes 3D) en relación a los centros de masa para que el origen de cada grupo de puntos se encuentre en su respectivo centroide, tal como se expresa en Ec. 4.7 y Ec. 4.8.

$$r_i = p_i - Cp \quad \text{Ec. 4.7}$$

$$s_i = q_i - Cq \quad \text{Ec. 4.8}$$

Al realizar el cambio de coordenadas se elimina la translación de la función a minimizar consiguiendo que la transformación entre los dos grupos de puntos sea únicamente de rotación, obteniendo así una nueva función a minimizar expresada como se muestra en Ec. 4.9:

$$F = \sum_{i=0}^N \|r_i - R \cdot s_i\|^2 \quad \text{Ec. 4.9}$$

El SVD de una matriz  $H$  de  $m \times n$  (donde  $m \geq n$ ) es una descomposición de  $H$  tal que  $H = U * D * V'$  donde  $U$  es una matriz de  $m \times n$  tal que  $U' * U$  es igual a la identidad,  $D$  es una matriz diagonal de  $n \times n$  y  $V$  es una matriz ortogonal de  $n \times n$ .

En [36] se muestra que la rotación óptima puede ser obtenida por  $R = VU'$  donde  $V$  y  $U$  son derivadas por medio de la SVD de la matriz de correlación  $H$ . La matriz de correlación  $H$  es dada por Ec. 4.10

$$H = \sum_{i=1}^N s_i \cdot r_i' \quad \text{Ec. 4.10}$$

Una vez obtenida la Rotación,  $R$ , se procede a calcular la Translación,  $T$ , que esta dada por la siguiente expresión: [49]

$$T = p - R \cdot q \quad \text{Ec. 4.11}$$

#### **4.2.2 Alineación de Imágenes. (Creación de Mapas 3D)**

Se propone un algoritmo mediante el cual se van a alinear dos nubes de puntos tridimensionales, expresadas en distintas referencias y correspondientes a un mismo objeto dado. El objetivo es utilizar la matriz de transformación de cada par de imágenes 3D para representar varias nubes de puntos en un mismo

sistema de coordenadas. En la Figura 4.5 se muestran dos imágenes consecutivas expresadas cada una en su propio sistema de coordenadas.

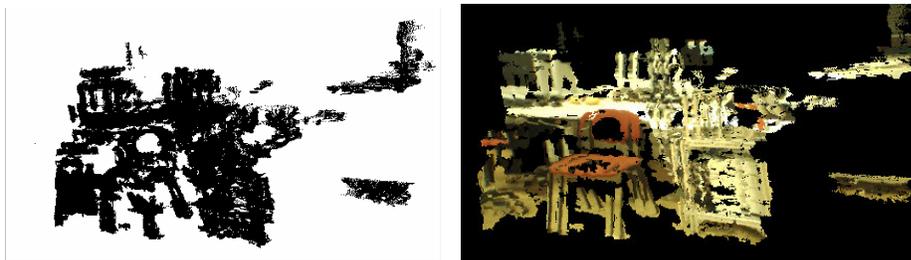


Figura 4.5.- Nubes de puntos sin alinear

Para realizar la alineación de todas las nubes de puntos se debe representar las matrices de transformación en coordenadas homogéneas. Mediante las coordenadas homogéneas se trata a todas las transformaciones con operaciones de multiplicación, para esto se utiliza una coordenada adicional para cada par  $(X, Y, Z)$ , de esa forma una coordenada homogénea se define como  $(kX, kY, kZ, k)$ , donde  $k$  es una constante arbitraria distinta de cero. [51]

#### 4.2.2.1 Translación

La translación de un punto  $(X, Y, Z)$  utilizando coordenadas homogéneas esta dado por el producto matricial: [50]

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix} \quad \text{Ec. 4.12}$$

Donde  $(T_x, T_y, T_z)$  es el vector de translación.

#### 4.2.2.2 Rotación

Para establecer la rotación de un punto en el espacio es necesario, primero definir el eje de coordenadas alrededor del cual se va a rotar. Así existen tres posibles rotaciones para un punto  $(X, Y, Z)$  utilizando coordenadas homogéneas que son: [50]

$$R(x, \alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \operatorname{sen} \alpha & 0 \\ 0 & -\operatorname{sen} \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{Ec. 4.13}$$

$$R(y, \beta) = \begin{bmatrix} \cos \beta & 0 & \operatorname{sen} \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\operatorname{sen} \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{Ec. 4.14}$$

$$R(z, \theta) = \begin{bmatrix} \cos \theta & \operatorname{sen} \theta & 0 & 0 \\ -\operatorname{sen} \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{Ec. 4.15}$$

Debido a que la variación en el orden de la rotación de un punto produce resultados distintos es necesario utilizar un orden en la composición de rotaciones tal como se muestra en Ec. 4.16.

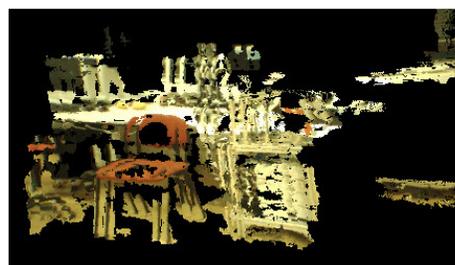
$$R = R(z, \theta)R(y, \beta)R(x, \alpha) = \begin{bmatrix} c\theta c\beta & -s\theta c\alpha + c\theta s\beta s\alpha & s\theta s\alpha + c\theta s\beta c\alpha & 0 \\ s\theta c\beta & c\theta c\alpha + s\theta s\beta s\alpha & -c\theta s\alpha + s\theta s\beta c\alpha & 0 \\ -s\beta & c\beta s\alpha & c\beta c\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{Ec. 4.16}$$

### 4.2.2.3 Alineación

Las matrices de Rotación, R, y Translación, T, son obtenidas mediante el método de SVD y luego son representadas en coordenadas homogéneas mediante una sola matriz como se muestra en Ec. 4.17 [51]

$$T(R, T) = \begin{bmatrix} c\theta c\beta & -s\theta c\alpha + c\theta s\beta s\alpha & s\theta s\alpha + c\theta s\beta c\alpha & Tx \\ s\theta c\beta & c\theta c\alpha + s\theta s\beta s\alpha & -c\theta s\alpha + s\theta s\beta c\alpha & Ty \\ -s\beta & c\beta s\alpha & c\beta c\alpha & Tz \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{Ec. 4.17}$$

Para alinear en un mismo sistema de coordenadas a varias imágenes 3D pertenecientes a un ambiente se multiplica la matriz de transformación correspondiente y se une de forma recursiva pares de imágenes. En la Figura 4.6 se muestran dos nubes de puntos representadas en su propio sistema de coordenadas (fila superior) y una vez alineadas al multiplicar una de las nubes de puntos por su respectiva matriz de transformación (fila inferior).



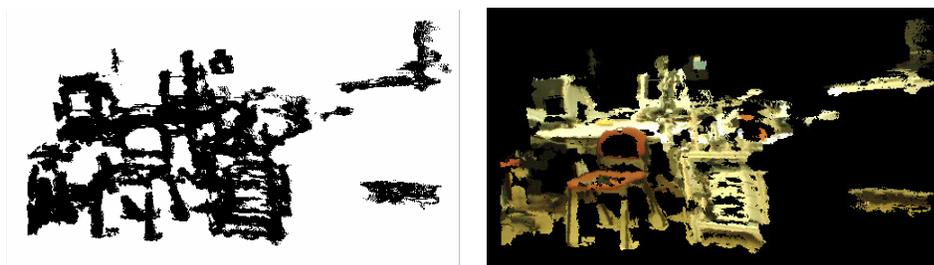


Figura 4.6.- (fila superior) Par de nubes de puntos consecutivas; (fila inferior) Nubes de puntos alineadas a partir de la matriz de transformación.

### 4.3 RESULTADOS

El robot utilizado para realizar las pruebas experimentales fue el *TRC Labmate de HelpMate Robotics Inc*, el cual consiste en una plataforma móvil de control diferencial al que se le ha montado la cámara de estéreo visión Digiclops. Este robot es del Laboratorio de Visión por Computador - Vislab del Instituto de Sistemas e Robótica – ISR, perteneciente a la Universidade Técnica de Lisboa (Portugal).

El uso de este robot es gracias a la colaboración que existe en el desarrollo de los Proyectos de Investigación: “Programación Asistida de Robots para Tareas Industriales (PARTI)” y de la “Red Iberoamericana de Robótica (RIBERO)”, financiados por la Agencia Española de Cooperación Internacional (AECI) a través del Programa CYTED (Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo), Subprograma VII: Electrónica e Informática Aplicadas, en los que el Centro de Visión y Robótica – CVR de la ESPOL

participa. Mientras que, la cámara de estéreo visión Digiclops es del CVR. La Figura 4.7 muestra el montaje de la cámara sobre el robot.



Figura 4.7.- Cámara Digiclops montada sobre el robot TRC.

Las pruebas experimentales consistieron en tres recorridos con el TRC, donde se adquirieron una serie de imágenes rectificadas de las cámaras izquierda, derecha y superior y se obtuvieron imágenes 3D a través de nubes de puntos, tal como se explica en el capítulo 3. A continuación se presentan tres pruebas realizadas, cada una con un recorrido diferente.

### **PRUEBA 1: Translación lateral**

En la primera prueba el robot realizó una trayectoria de translación lateral, teniendo un desplazamiento total de aproximadamente 2.0 metros, en el que se obtuvo 76 imágenes 3D. En este recorrido el error acumulado obtenido es aproximadamente 4 cm. en el eje de desplazamiento y 14 cm. en el eje ortogonal, tal como se muestra en la Figura 4.8 (derecha). En la Figura 4.8

(izquierda) se muestra el recorrido realizado por el robot (rojo) y la estimación realizada en base al algoritmo de SLAM (azul).

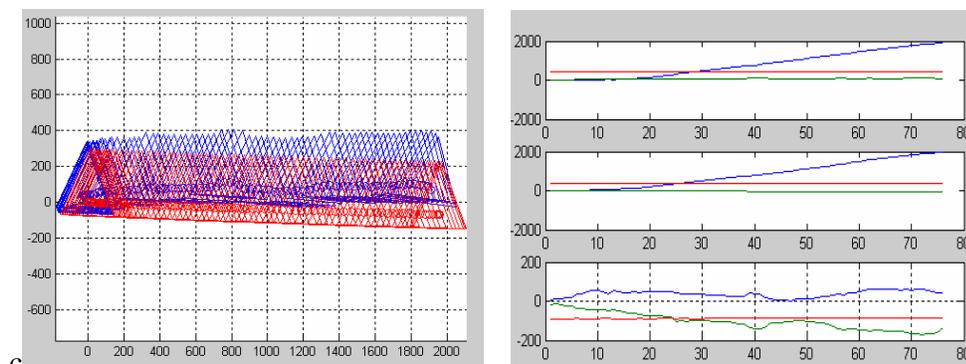


Figura 4.8.- (columna izquierda) trayecto lateral realizado por el robot ; (columna derecha) Movimiento del robot en los ejes x y y error obtenido

En la Figura 4.9 se muestran la primera y la última imagen obtenida en el recorrido del robot y sus respectivas nubes de puntos

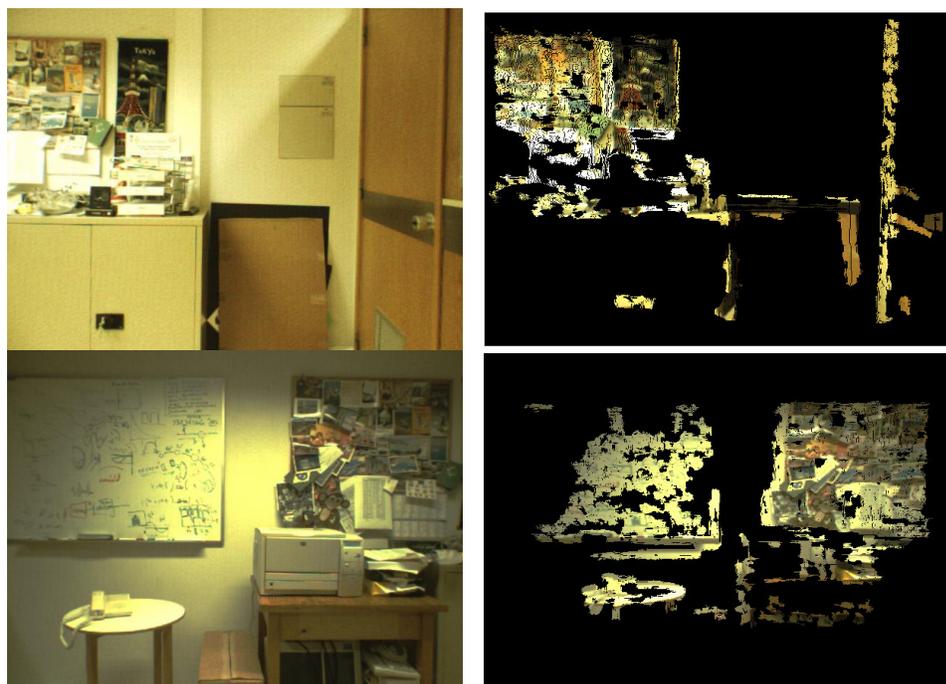


Figura 4.9.- (columna izquierda) Imagen a color del ambiente; (columna derecha) nubes de puntos resultantes.

La Figura 4.10 muestra la reconstrucción completa del ambiente (mapa 3D del ambiente) realizada a partir de las 76 imágenes obtenidas en el movimiento lateral del robot realizado en la prueba 1.

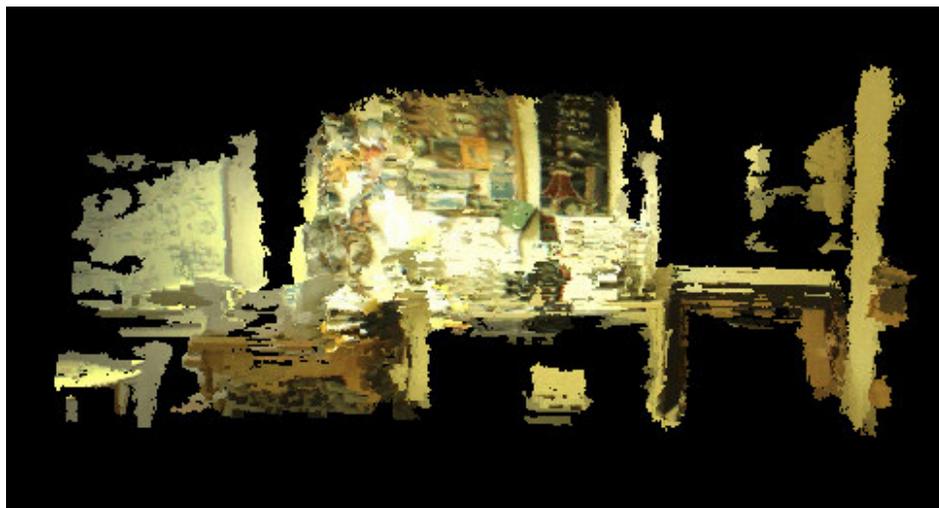


Figura 4.10.-Mapa 3D de ambiente de navegación del robot

Durante el recorrido de la prueba 1, el error total de la estimación es aproximadamente del 7%, obteniendo en cada movimiento del robot una localización bastante aproximada a la realidad y un mapa lo suficientemente bueno como para que le robot navegue autónomamente sin ningún problema.

### **PRUEBA 2: Translación frontal**

Durante el desarrollo de la segunda prueba el robot recorrió una trayectoria de translación hacia delante, desplazándose un total aproximado de 1.5 metros y obteniendo 60 imágenes 3D. El error acumulado obtenido es aproximadamente 5 cm. en el eje de desplazamiento y 12 cm. en el eje

ortogonal, tal como se muestra en la Figura 4.11 (derecha). El recorrido realizado por el robot (rojo) y la estimación realizada en base al algoritmo de SLAM (azul) son mostrados en la Figura 4.11 (izquierda).

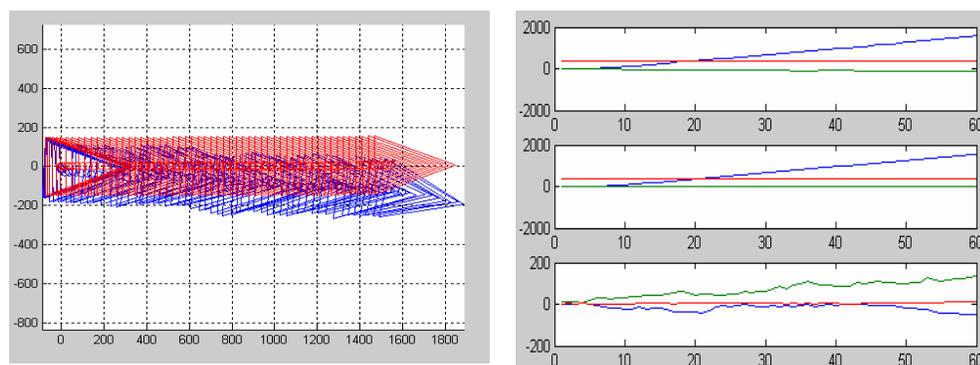


Figura 4.11.- (columna izquierda) trayecto hacia delante realizado por el robot ; (columna derecha) Movimiento del robot en los ejes x y y error obtenido

En la Figura 4.12 se muestran la primera y la última imagen obtenida en el recorrido del robot y sus respectivas nubes de punto



Figura 4.12.- (columna izquierda) Imagen a color del ambiente; (columna derecha) nubes de punto resultantes.

La Figura 4.13 muestra la reconstrucción completa del ambiente (mapa 3D del ambiente) realizada a partir de las 60 imágenes obtenidas en el movimiento de translación hacia delante del robot.



Figura 4.13.- Mapa 3D de ambiente de navegación del robot

Durante el recorrido de la prueba 2, el error total de la estimación es aproximadamente del 9%. Al igual que en la prueba 1 se obtiene una localización bastante aproximada a la realidad en cada movimiento del robot y un mapa lo suficientemente bueno como para que el robot navegue autónomamente, aunque aparecen obstáculos extras debido a la redundancia de datos al momento de realizar el registro y alineación de imágenes, pero esta información extra no influye mayormente en el robot al momento de navegar autónomamente ya que son valores pequeños.

### **PRUEBA 3: Movimiento circular**

En la tercera prueba se realizó un recorrido que mezcla rotación y translación, el recorrido elegido fue una trayectoria de dos vueltas alrededor de una circunferencia de aproximadamente 3 metros de diámetro. En la Figura 4.14 se muestra el recorrido realizado por el robot y la estimación realizada.

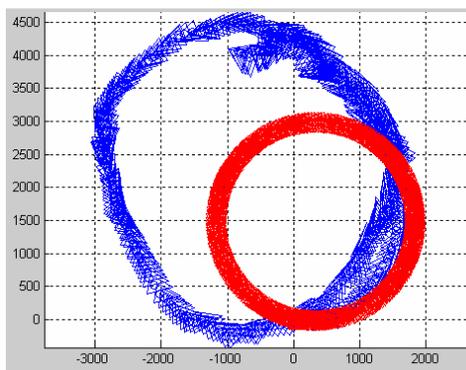


Figura 4.14.- Trayectoria circular realizada por el robot

En la Figura 4.14 se aprecia que el error de cada par de imágenes se acumula creando un error total cada vez mayor. Se trató de reducir la acumulación de error utilizando mejores estimaciones iniciales en el cálculo de las matrices de rotación y translación, para esto se tomaron las matrices obtenidas anteriormente y se obtuvo el movimiento realizado cada 5 imágenes, luego se utilizó esta información como estimación inicial en un nuevo cálculo de la matriz de movimiento, obteniendo así menores errores (Figura 4.15).

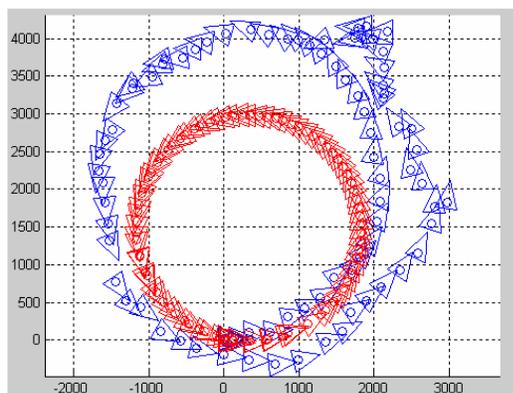


Figura 4.15.-Trayectoria calculada de forma iterativa cada 5 imágenes

En la Figura 4.16 se muestran la primera y la última imagen obtenida en el recorrido circular del robot y sus respectivas nubes de puntos.

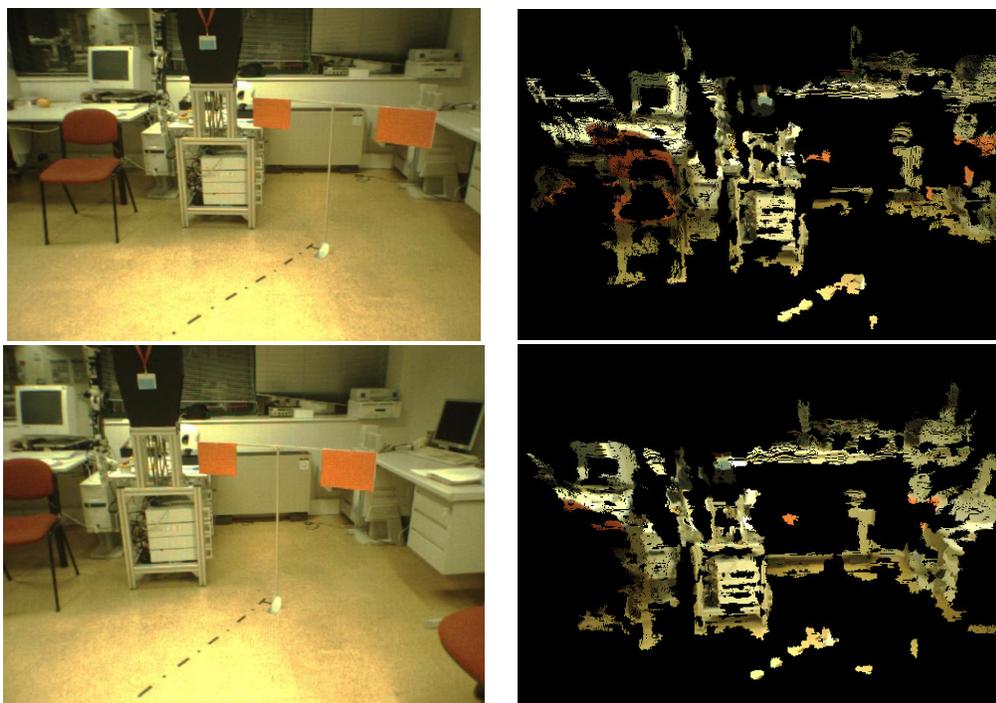


Figura 4.16.- (columna izquierda) Imagen a color del ambiente; (columna derecha) nubes de puntos resultantes.

La Figura 4.17 muestra la reconstrucción de algunos sectores del entorno de navegación del robot, estimados durante la realización de la prueba 3.

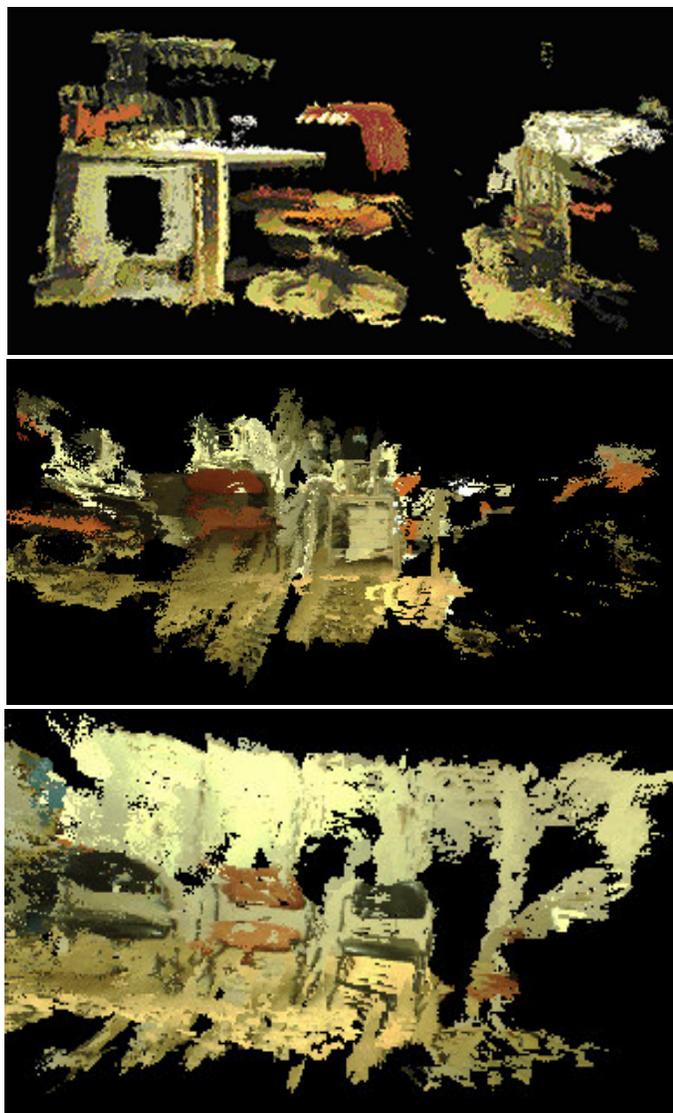


Figura 4.17.- Segmentos del mapa 3D del ambiente de navegación de la prueba 3.

Durante el desarrollo de las prueba 3 se puede apreciar que el error aumenta notablemente en relación a las pruebas 1 y 2, esto se debe principalmente a los puntos aislados (outliers) que se obtienen durante la etapa de adquisición, los

cuales mantienen el mismo patrón durante el desplazamiento, pero varían al cambiar el ángulo de adquisición de la imagen.

#### **4.4 CONCLUSIONES**

El presente capítulo pretende solucionar el problema de registro y alineación de imágenes 3D del ambiente de navegación de un robot. El objetivo es modelar el ambiente de navegación del robot para proporcionarle información necesaria para la futura navegación del robot en el ambiente.

En respuesta a este problema se ha realizado la implementación de un sistema de registro mediante el algoritmo ICP y se ha optado por prescindir de la etapa de integración y realizar solamente una alineación de todos los datos adquiridos por el robot.

Finalmente un conjunto de pruebas experimentales al sistema registro y alineación fue realizado, por medio del uso de imágenes adquiridas por un sistema de visión estéreo montado sobre el robot TRC Labmate. Con estas pruebas se obtuvieron la localización del robot en cada punto de adquisición y el modelo o mapa 3D del ambiente, con el cual se pueden desarrollar futuros proyectos de navegación de robots.

Durante el desarrollo de las pruebas experimentales se puede apreciar que el error es bajo en los recorridos que solo implican la translación del robot pero

que aumenta notablemente al introducir la rotación en el movimiento, esto se debe principalmente a los puntos aislados (outliers) que se obtienen durante la etapa de adquisición, los cuales mantienen el mismo patrón durante el desplazamiento, pero varían al cambiar el ángulo de adquisición de la imagen. Al obtener una gran cantidad de outliers en dos imágenes consecutivas se dificulta el registro, obteniendo así mayores errores.

Para mejorar el registro de imágenes se prevé utilizar un método de muestreo que permita eliminar puntos aislados (outlier) en la imagen. De esta forma, el proceso de registro de nubes de puntos trabajará sobre valores con poco ruido y se obtendrán resultados con mayor precisión. Una opción puede ser el algoritmo de RANSAC.

Además se plantea implementar un registro simultáneo de varias imágenes 3D, logrando una disminución del error global de registro al reducir o eliminar la acumulación del error para cada par de conjuntos de datos

## **CAPÍTULO V**

### **RESUMEN Y CONTRIBUCIONES**

Este capítulo presenta un resumen de los principales módulos implementados para el desarrollo de este proyecto. De igual forma se muestran las aportaciones y logros obtenidos en la realización del proyecto.

#### **5.1 INTRODUCCIÓN**

Un problema clásico en el trabajo de robots móviles autónomos que operan en entornos desconocidos es la navegación. Las actividades esenciales para llegar a la solución de este problema consisten en determinar donde se encuentra el robot dentro del medio desconocido (localización) y construir incrementalmente un mapa del entorno que ha recorrido el robot (mapeo). El mapa podrá servirle al robot para posteriormente realizar tareas como localizarse dentro de dicho entorno, planear trayectorias y otras actividades.

Un robot autónomo debe disponer de una representación interna del ambiente con el que interactúa, para poder establecer cuáles pueden ser los posibles obstáculos en el camino y poder reconocer regiones u objetos en el entorno. La representación que requiere el robot del medio (mapa) no debe ser exhaustiva (detallada), sino que debe contener la información suficiente para que el robot no colisione.

La implementación del algoritmo de Localización y Mapeo Simultáneos (SLAM) en el presente proyecto permite determinar el movimiento con 6 grados de libertad del robot TRC y, además, crear un modelo 3D del ambiente de navegación de este robot.

En el desarrollo de este capítulo se resumen los procesos necesarios para la implementación del algoritmo de SLAM. Además se describen las principales contribuciones del proyecto y se plantean recomendaciones para futuros trabajos que pueden ser realizados para la mejora del algoritmo actual y las áreas afines en la que se podrían aplicar.

## 5.2 RESUMEN

### Sistema de adquisición 3D

Una habilidad esencial para la navegación autónoma de robots es la capacidad para obtener información del medio que rodea al robot. Por lo que se ha desarrollado un sistema de adquisición 3D que utiliza una cámara de visión estéreo, modelo *Digiclops*, fabricada por *Point Grey Research Inc* [27], el cual ha sido montado sobre un robot móvil.

La cámara *Digiclops* permite la adquisición de un conjunto de imágenes al momento que el robot realizaba un recorrido determinado. Cada imagen representa una vista del ambiente del robot. El proceso de adquisición se realiza de una manera rápida sin tener la restricción de detener el movimiento del robot, como sucedería en el caso de usar un sensor de escáner láser.

En cada proceso de adquisición tres imágenes 2D son obtenidas. Estas imágenes son usadas por la librería *Triclops SDK* [28] para establecer una relación entre los puntos que conforman cada imagen. La información obtenida de esta relación son valores de profundidad (disparidad) del ambiente sensado. Luego, con la información geométrica de la cámara y con el valor de disparidad se calcula la distancia a la que se encuentran los puntos del

entorno, obteniendo así una nube de puntos 3D del ambiente de navegación del robot.

El sistema de adquisición 3D entrega como resultado varios conjuntos de puntos 3D (nubes de puntos) que representan a varias escenas del medio que ha recorrido el robot.

### **Sistema de localización**

El sistema de localización permite al robot determinar en que posición se encuentra dentro de su ambiente de navegación. El sistema de localización ha sido implementado a partir de las nubes de puntos proporcionadas por el sistema de adquisición 3D, la localización del robot se la obtiene calculando el movimiento realizado al adquirir las imágenes por medio del registro de un par de nubes de puntos consecutivas.

El registro de las nubes de puntos se lo realiza a través del algoritmo ICP por medio de un proceso iterativo que reduce una función de minimización de distancias que en cada iteración acercando de manera global los dos conjuntos de puntos hasta lograr que ambos converjan.

En el algoritmo de ICP utilizado primero se realiza un muestreo de puntos de forma uniforme, luego se encuentra una correspondencia entre el par de nubes

de puntos a través del vecino más cercano haciendo uso de árboles *kd* por medio de la librería ANN [34]. Posteriormente, se desestima los pares correspondientes que no cumplen con un filtro de color en el espacio HSI por medio de la lectura de una tabla LUT creada con anterioridad, por último la transformación que representa el movimiento de las nubes de puntos es estimada por medio del método basado en la descomposición de valores singulares (SVD), propuesto en [36], de donde se obtiene la rotación y translación realizada por la cámara. A partir del movimiento realizado por la cámara se estima el movimiento del robot y su localización, tomando como referencia su posición anterior.

### **Sistema de mapeo 3D**

El mapeo 3D se refiere a la creación de una representación del medio que rodea al robot en la cual se almacene información tridimensional del mismo. El sistema de mapeo 3D utiliza las transformaciones obtenidas por el sistema de localización y las nubes de puntos adquiridas por el sistema de adquisición 3D para crear los mapas 3D del ambiente de navegación del robot.

Por medio del sistema realizado se representan en un mismo sistema de coordenadas a varias imágenes 3D pertenecientes a un ambiente. Para realizar la alineación de todas las nubes de puntos se representan las matrices de transformación en coordenadas homogéneas, de esa forma la integración de

las imágenes se realiza multiplicando la matriz de transformación correspondiente a una nube de puntos y añadiéndola a la nube de puntos de referencia, disminuyendo el tiempo de procesamiento que se requeriría al realizar por separado el cálculo de la rotación y translación.

### **5.3 CONTRIBUCIONES**

#### **5.3.1 Filtro en el espacio de color HSI**

Por medio de la cámara utilizada se adquieren imágenes en el espacio de color RGB, el cual es un espacio muy sensible a los cambios de iluminación. A diferencia del modelo de color RGB, el espacio HSI separa la iluminación de la información de color. Por lo cual el color permanece estable bajo variaciones en la intensidad de la iluminación.

Debido a que las imágenes adquiridas presentan una variación en la iluminación correspondiente al desplazamiento realizado por el robot y a que no se controlan los parámetros de iluminación, se ha realizado un filtrado en el espacio de color HSI que permite ignorar la componente I (iluminación de la escena) para eliminar falsas correspondencias de vecinos más cercanos y poder, así, realizar un mejor registro de imágenes.

### **5.3.2 Creación de LUT (Look Up Table) para espacio HSI**

Como se ha decidido utilizar un filtro en el espacio de color HSI y las imágenes adquiridas están representadas en el espacio RGB es necesario realizar una transformación de dichos modelos de color.

La conversión de un modelo de color a otro es un proceso muy costoso computacionalmente, ya que se deben realizar una gran cantidad de cálculos, por lo que se ha optado realizar una tabla de búsqueda LUT (Look Up Table). En la LUT se almacenan los valores de tono y saturación de manera tal que el identificador de la posición de cada valor dentro de la LUT representa a la combinación de RGB correspondiente al valor de tono y saturación almacenado.

La creación de la tabla se la realiza una sola vez y se la carga en la memoria antes de realizar el registro de las imágenes. De esta manera se puede utilizar el filtro en el espacio HSI sin comprometer recurso computacional.

### **5.3.3 Modelado de superficies de objetos reales**

Se realizó un sistema de modelado 3D de superficies [30] que tiene como objetivo el crear en el computador un modelo tridimensional de una superficie de un objeto real, para lo cual se utiliza una malla triangular.

La malla triangular se genera usando un algoritmo público de Delaunay 2D [31]. La triangulación se realiza primero proyectando la nube de puntos al plano XY y luego triangulando estos en este espacio. La malla triangular resultante es mapeada al espacio 3D reasignando el valor de la coordenada Z a cada punto previamente proyectado. En la literatura este proceso es conocido como triangulación 2½D.

#### **5.3.4 Publicaciones Científicas**

Los resultados obtenidos con los trabajos de investigación de este proyecto dieron lugar a la publicación de tres artículos científicos en congresos y eventos nacionales. Las referencias a estos artículos son dadas abajo.

[1] M.A. Realpe, B.X.Vintimilla, "Modelado 3D de superficies mediante visión estereo", 1<sup>ER</sup> Congreso Nacional de Investigación, Tecnología e Innovación, Quito, Ecuador, Octubre del 2005. (versión extendida).

[2] M.A. Realpe, B.X.Vintimilla, "Modelado 3D de superficies mediante visión estereo", III Jornadas de Investigación - ESPOLCIENCIA 2004, Guayaquil, Ecuador, Noviembre - Diciembre 2004.

[3] B.X.Vintimilla, M.A. Realpe, "Sistema de reconstrucción digital 3D", Tecnociencia, Boletín Informativo, No.12, Ecuador, Noviembre 2003, 4. ISSN: 1390-1494.

### 5.3.5 Logros adicionales obtenidos con la ejecución de este proyecto

Existen logros paralelos que se han logrado en el desarrollo de este proyecto, entre los cuales se mencionan:

- Base científica para la investigación y desarrollo en nuevo proyectos de visión 3D, navegación de robots, planificación de trayectorias, representación de mapas 3D, entre otros.
- Obtención de financiamiento para los proyectos: (1) “*Sensor 3D para la adquisición de imágenes de profundidad (range images)*”, por parte de ESPOL-CICYT; y (2) “*Reconstrucción 3D de objetos sólidos a partir de múltiples imágenes de profundidad*”, por parte de FUNDACYT.
- Corta estancia de investigación en el *Laboratorio de Visión por Computador - Vislab* del *Instituto de Sistemas e Robótica – ISR*, de la *Universidade Técnica de Lisboa* (Portugal), para realizar las pruebas experimentales del proyecto implementado. El financiamiento de esta estancia se realizó con presupuesto de los proyecto: “*Programación Asistida de Robots para Tareas Industriales (PARTI)*” y de la “*Red Iberoamericana de Robótica (RIBERO)*”, a través del *Programa CYTED* (Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo), en los que el Centro de Visión y Robótica de la ESPOL participa.

## CONCLUSIONES Y RECOMENDACIONES

Al finalizar el presente trabajo se han obtenido las siguientes conclusiones:

- Se ha podido notar que la iteración inicial del algoritmo de ICP al calcular el movimiento del robot en base a un par de imágenes influye notablemente en el cálculo final de la estimación. Por lo que se sugiere utilizar el algoritmo en conjunto con un sensor de posición del robot.
- Durante la etapa de adquisición de imágenes 3D se obtienen una cantidad excesiva de puntos que se encuentran alejados de su ubicación real o puntos aislados (outliers), lo cual se sugiere mejorar con procesos de procesamiento de imágenes.
- El error al calcular los recorridos que solo implican la translación del robot es bajo, pero que aumenta notablemente al introducir la rotación en el movimiento, esto se debe principalmente a los outliers.
- Para mejorar una trayectoria estimada se puede volver a recurrir al algoritmo ICP tomando como primera iteración la trayectoria ya obtenida.
- Al obtener una gran cantidad de outliers en dos imágenes consecutivas se dificulta el registro, obteniendo así mayores errores.

Los resultados de este proyecto abren futuras líneas de trabajo que pueden ser asumidas por investigadores y estudiantes de la ESPOL o de otros centros de estudio superior, algunos trabajos y mejoras que se recomiendan son:

- Implementar un registro simultáneo de varias imágenes 3D, logrando una disminución del error global de registro al eliminar la acumulación del error para cada par de conjuntos de datos.
- Implementar la fase de integración de imágenes para obtener mapas de mayor precisión que pueden ser utilizados en proyectos donde se requiera reconocimiento de objetos y manipulación de los mismos.
- Utilizar sensores adicionales que permitan realizar una primera aproximación al movimiento del robot, que puede ser luego mejorado por medio del algoritmo de ICP.
- Utilizar el sistema de visión implementado sobre robots para la exploración de ambientes hostiles y peligrosos.
- Remover datos redundantes en el mapa final manteniendo la información necesaria para la navegación y disminuyendo la cantidad de recursos computacionales a consumir.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] JM Sáez, “Navegación 3d utilizando información estéreo”, proyecto de investigación de tercer ciclo para el reconocimiento de la suficiencia investigadora, Universidad de Alicante, Septiembre de 2001
  
- [2] Miguel Zamora Izquierdo, Luis M. Tomás Balibrea, Humberto Martínez Barberá, Antonio Gómez Skarmeta, “Navegación planificada de un robot móvil en interiores desconocidos”, I Workshop Hispano-Luso de Agentes Físicos (WAF’2000), Tarragona, 2000
  
- [3] Dinesh Manocha, “Planning for automatic capture of unknown environments/objects”, Curso de pre-grado “Motion, planning and synthesis”, Department of Computer Science Campus - The University of North Carolina at Chapel Hill, 2003
  
- [4] Miguel A. Cazorla, Otto Colomina, “Cáp. 4: Sensores”, Curso de pre-grado “Robótica (robots autónomos)”, Departamento de Ciencia de la Computación e Inteligencia Artificial – Universidad de Alicante, 2005.
  
- [5] L. Enrique Sucar, “Cáp. 2: Arquitecturas”, Curso “Robótica Móvil”, Tecnológico de Monterrey, 2004.

- [6] Robert A. Freitas Jr., “Nanomedicine, Volume I: Basic Capabilities”, <http://www.nanomedicine.com/NMI.htm>, 1999
- [7] Joseph L Jones, Anita M. Flynn, Bruce A. Seiger, “Mobile Robots, Inspiration to implementation”, Segunda edicion, Editorial A K Peters, Ltd.
- [8] J. Borenstein, H. R. Everett, L. Feng, “Where am I? Sensors and Methods for Mobile Robot Positioning”, Reporte técnico editado y compilado por J. Borenstein, Abril 1996.
- [9] P. Misra, B.P. Burke, M.M. Pratt, “GPS Performance in navigation”, Proceedings de IEEE, vol. 87, no. 1, 1999, pp. 65-85
- [10] International GNSS Service, <http://igs.cb.jpl.nasa.gov>
- [11] Chieh-Chih Wang, Charles Thorpe, Sebastian Thrun, “Online Simultaneous Localization and Mapping with Detection and Tracking of Moving Objects: Theory and Results from a Ground Vehicle in Crowded Urban Areas”, IEEE International Conference on Robotics and Automation (ICRA'03), Mayo 2003
- [12] J. García, Manuel. Mazo Quintas, Jesús Ureña, Marta Marrón Romera, “Autolocalización y posicionamiento mediante Marcas Artificiales codificadas”, TELECOM, 2002

- [13] J. Borenstein, H.R. Everet, L. Feng, D. Wehe, “Mobile Robot Positioning: Sensors and Techniques”, *Journal of Robotic Systems*, 14(4), pp. 231-249, 1997.
- [14] J. González, A. Ollero, “Estimación de la posición de un robot móvil”. *Revista de Informática y Automática*, Diciembre 1996
- [15] J.M. Pérez, J.P. Bandera, C. Urdiales, F. Sandoval, “Agente autónomo de bajo coste para la exploración de conductos teleoperado mediante realidad virtual”. I Seminario Nacional Hispabot (HISPABOT 03), Alcalá de Henares (Madrid), Abril, 2003.
- [16] F. J. Escarpa, “Introducción a los Sistemas de Medición Tridimensional con Láser” [http://www.mappinginteractivo.com/plantilla-egeo.asp?id\\_articulo=198](http://www.mappinginteractivo.com/plantilla-egeo.asp?id_articulo=198) Santiago & Cintra Ibérica S.A., Alcobendas (Madrid). Septiembre 2002
- [17] Brian Lee Curless, “New methods for surface reconstruction from range images a dissertation for the degree of doctor of philosophy”, Junio 1997
- [18] D. Robertson, R. Cipolla. “An interactive system for constraint-based modelling”, *British Machine Vision Conference*, Bristol, Septiembre 2000.
- [19] Ardy Goshtasby, *Curso de visión*, Wright State University and Image Fusion Systems Research, Junio, 2004

- [20] M. Soucy, D. Laurendeau, “A general surface approach to the integration of a set of range views”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 4, pp. 344–358, 1995
- [21] Y. Chen, G. Medioni, “Object modelling by registration of multiple range images”, *IEEE Conf. on Robotics and Automation*, 1991
- [22] P. Besl, N.D. McKay, “A method for registration of 3-d shapes”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, Febrero, 1992
- [23] S. M. Rusinkiewicz, M. Levoy. “Efficient variants of the ICP algorithm”, *Reporte tecnico, Stanford University*, 2001
- [24] B. Horn, “Closed-form solution of absolute orientation using unit quaternions”, *Journal of the Optical Society of America*, 4(4), 1987
- [25] B. Horn, B. Hilden, S. Negahdaripour, “Closed-form solution of absolute orientation using unit orthonormal matrices”, *Journal of the Optical Society of America*, 1988
- [26] Javier González Jiménez, “Visión por Computador”, *Editorial Parainfo*, 2000

- [27] Digiclops™ Stereo Vision System, “Installation Guide and Camera Control API Command Reference Version 2.4” Copyright 2001 Point Grey Research Inc.
- [28] “Triclops Stereo Vision System Manual Version 3.1” Copyright 2002 Point Grey Research Inc.
- [29] Carlos Ruiz Moreno, “Visión estereoscópica, técnicas de correspondencia”, Febrero 2005
- [30] M.A. Realpe, B.X.Vintimilla, "Modelado 3D de Superficies Mediante Visión Estereo", 1er Congreso Nacional de Investigación, Tecnología e Innovación, Quito, Ecuador, Octubre 2005. Impreso en formato CD.
- [31] J. R. Shewchuk, “Triangle: engineering a 2D quality mesh generator and Delaunay triangulator” First Workshop on Applied Computational Geometry, Philadelphia, Pennsylvania, ACM, Mayo 1996
- [32] T. Masuda, N. Yokoya, “A robust method for registration and segmentation of multiple range images”, Computer Vision and Image Understanding, pp. 295-307, vol.61, n°3, 1995.
- [33] H Surmann, A Nuchter, K Lingemann, J Hertzberg, “6D SLAM with Approximate Data Association”, ICRA, 2005

- [34] David Mount, Sunil Arya, “ANN: A Library for Approximate Nearest Neighbor Searching”
- [35] F. Perez, C. Kock, “Toward color image segmentation in analog VLSI: Algorithm and hardware”, *International Journal of Computer Vision*, vol. 12, pp.17–42, 1994.
- [36] K. S. Arun, T. S. Huang, S. D. Blostein, “Least square fitting of two 3-d point sets”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698- 700, 1987.
- [37] B. Vintimilla B., C. Monsalve A., C. Villarroel S., C. Calderón G., R. Carrillo C., V. Salamea, D. Paillacho. J. Chicala, L. Atiencia, R. Jacho, “Sistemas multiagentes aplicados al fútbol robótico: problemática existente”, *IV Jornadas Iberoamericanas de Robótica, Seminario 3, Panamá. 27-31 Enero, 2003.*
- [38] Luis Miguel Bergasa, “Sistemas robóticos para la asistencia a personas mayores”, Charla dictada en la Universidad de Alcalá, Diciembre 2003.
- [39] Melvin Andrés González Pino, José Gabriel Ojeda Rodríguez, “Robot móvil de vigilancia con sistema de omnivision y navegación autónoma con GPS”, *IV Congreso Internacional Electrónica y Tecnologías de Avanzada, Colombia.*

- [40] J. Borestein, B. Everett, L. Feng, "Navigation mobile robots: systems and techniques", Wellesley, MA: A. K. Peters, 1996
- [41] A. Pozo-Ruz, M.C. García-Alegre, A. Ribeiro, L. García, D. Guinea, F. Sandoval, "Localización de vehículos: Fusión de GPS y odometría.", Mundo Electrónico, N° 309, pp. 46-54, Mayo 2000
- [42] N. Sotomayor, "La Robótica Móvil en la Escuela Politécnica Nacional", Informativo Politécnico, Quito, Diciembre 2005.
- [43] M. Betke, L. Gurvits, "Mobile Robot Localization using landmarks", Transactions on Robotics and Automation, vol. 13, no. 2, 1997, pp. 251-263.
- [44] Branch, J.W., Gomez, J.B., Prieto, F. "Reconstrucción tridimensional a partir de imágenes de rango: Revisión de la literatura." Revista Iberoamericana de Inteligencia Artificial. España. 2004
- [45] Jairo Valencia, "Registro de Imágenes de Rango para la Construcción de Modelos Faciales", Tesis previa a título Ingeniero Electrónico, presentada a la Facultad de Ingeniería y Arquitectura de la Universidad Nacional de Colombia Sede Manizales, Enero 2004
- [46] Leica, geosystems. <http://www.leica-geosystems.com>, 2006

- [47] Wikipedia, Computer Graphics, [http://en.wikipedia.org/wiki/Rendering\\_\(computer\\_graphics\)](http://en.wikipedia.org/wiki/Rendering_(computer_graphics)), 2006
- [48] Gonzalo Pajares, Jesús M. de la Cruz, “Visión por Computador imágenes digitales y aplicaciones”, editorial RA-MA, 2001
- [49] J. Pascau, J. D. Gispert, S. Reig, R. Martínez, M. Desco, “Registro de imágenes en medicina nuclear”, Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales, Vol. 96, N° 1-2 1-2, 2002
- [50] Patricio Inostroza, “Transformaciones 3D”, Curso de pre-grado “Computación Gráfica”, Departamento de Ciencias de la Computación – Universidad de Chile, Chile, 2004.
- [51] José Sabater Navarro, “Herramientas matemáticas para la localización industrial”, Curso de pre-grado “Control de Robots y Sistemas sensoriales”, Departamento de Ingeniería de Sistemas Industriales – Universidad Miguel Hernandez, España, 2005.