

# Capítulo 1

---

# INTRODUCCIÓN

**E**l marco de esta tesis se centra en la navegación autónoma de robots. Dentro de la navegación podemos definir un conjunto de tareas que el robot se puede plantear llevar a cabo. Se nos puede plantear una única tarea, o bien combinaciones de las tres. Estas tareas son:

- Localización: dado un entorno conocido y utilizando información del entorno, el objetivo es encontrar la posición del robot dentro del entorno.
- Construcción de mapas del entorno: el robot está situado en un entorno desconocido. El robot se mueve por él y obtiene una representación geométrica, que nos puede servir para resolver las otras dos tareas.
- Planificación de trayectorias: tenemos al robot en una posición concreta, el objetivo es guiarlo hasta otra posición sin que colisione con los obstáculos.

Clásicamente, estos problemas se han venido resolviendo utilizando la información proporcionada por sensores de ultrasonidos (sonar). Estos sensores devuelven medidas de distancia que nos indican dónde se encuentran los obstáculos. Sin embargo, estas medidas, además de locales, suelen ser ruidosas y fuente de errores. El objetivo de esta tesis es desarrollar un conjunto de técnicas y métodos de extracción y agrupamiento de características en visión artificial que permitan reducir algunos de los errores cometidos. Para que estos métodos se puedan utilizar en un sistema robótico deberán tener, entre otras, las siguientes características:

- El tiempo de ejecución deberá ser lo más pequeño posible. En un sistema móvil tenemos que ser capaces de dar respuesta en el menor tiempo posible debido al movimiento del propio móvil. La mayoría de métodos propuestos para robótica móvil que utilizan visión necesitan una gran cantidad de tiempo de procesamiento.

- Los métodos de extracción deben ser robustos. Tenemos que ser capaces de poder extraer dichas características ante pequeños cambios del entorno.
- También serán genéricas. Nos planteamos extraer características que sirvan para diversas tareas.

Un ejemplo de sistema que utiliza visión de forma eficiente es el desarrollado en la Universidad de Munich [Dickmanns, 1997]. Consiguen guiar a un vehículo por una autopista en tiempo real. También controlan el vuelo de un helicóptero. El sistema que extrae características de la imagen tiene una respuesta temporal de aproximadamente 0,03s. Sin embargo, las características obtenidas sólo sirven para el entorno en el que trabajan y bajo ciertas circunstancias externas, como puede ser el clima y las condiciones de iluminación derivadas de este. Actualmente, cuando se pretende desarrollar un sistema robótico que permita trabajar en tiempo real se suelen buscar características muy específicas al problema a tratar. Sin embargo, nuestra propuesta intenta ser más genérica, teniendo una merma en cuanto al tiempo de respuesta. Es decir, buscamos desarrollar un conjunto de métodos que permitan extraer características que puedan ser utilizadas por diversas tareas. Este planteamiento es clave en aproximaciones a la percepción dirigidas por la tarea.

Algunos de los algoritmos desarrollados en esta tesis se han probado utilizando un robot Pioneer I, trabajando con el software Saphira. La Figura 1.1 muestra el robot utilizado.



**Figura 1.1:** Robot utilizado en algunos de los experimentos.

## 1.1 Visión orientada a tareas frente a visión reconstructiva

A principios de los años 80, Marr [Marr, 1982] definió uno de los primeros paradigmas sobre la visión artificial. En él establecía una división en tres niveles de la visión. El nivel más bajo se correspondía con una representación en dos dimensiones de la imagen. Para pasar al siguiente nivel aplicaríamos a la imagen una serie de operadores muy básicos que nos proporcionarían características muy locales de la imagen. Estos operadores pueden ser de detección de discontinuidades (aristas) en la imagen, detección de puntos de alta curvatura en aristas, estéreo, detección de movimiento, detección de texturas, etc. Aplicando estos operadores llegamos al segundo nivel. La representación de propiedades de las superficies visibles en un sistema centrado en el observador, tales como orientación de superficie, distancia a partir del observador, discontinuidades, etc. se denomina representación 2-1/2 D, que es una representación intermedia entre el segundo y tercer nivel. Por último llegamos al tercer nivel que se corresponde con una representación centrada en el objeto y en tres dimensiones.

Frente a este planteamiento surgió la visión orientada a tareas [Landy *et al.*, 1996] o visión activa [Yuille y Blake, 1992]. Si utilizamos un esquema de visión orientada a tareas, la tarea es la que define qué módulos u operadores utilizaremos y cómo se van a comunicar entre sí. Es la tarea la que se va a encargar de definir si utilizamos un detector de aristas en vez de uno de esquinas, si necesitamos una agrupación de aristas y en qué medida y qué representación es la más adecuada. Este nuevo esquema permite únicamente procesar aquellas zonas de la imagen útiles para la tarea, frente al esquema clásico que procesaba toda la imagen. También, el nuevo enfoque asume una limitación en los recursos del sistema. En esta tesis nos centraremos en la tarea de la localización visual en un entorno conocido. Esta tarea ha sido tratada en multitud de ocasiones con técnicas que utilizan el sonar como fuente de información. Podemos consultar [Gallardo, 1999] para una reciente propuesta utilizando información bayesiana. Sin embargo, la solución obtenida con estos métodos es sensible a los errores de odometría (medición de su propio movimiento). Los errores de odometría surgen por el ruido implícito en los sensores que nos miden el movimiento. El error más relevante es el que se comete en la detección del ángulo de giro del robot. Al margen de que la visión permite mayor rango perceptual, la extracción robusta y la utilización de características genéricas y representaciones esquemáticas permiten buenas localizaciones en tiempos aceptables. Este es el objetivo último de esta tesis.

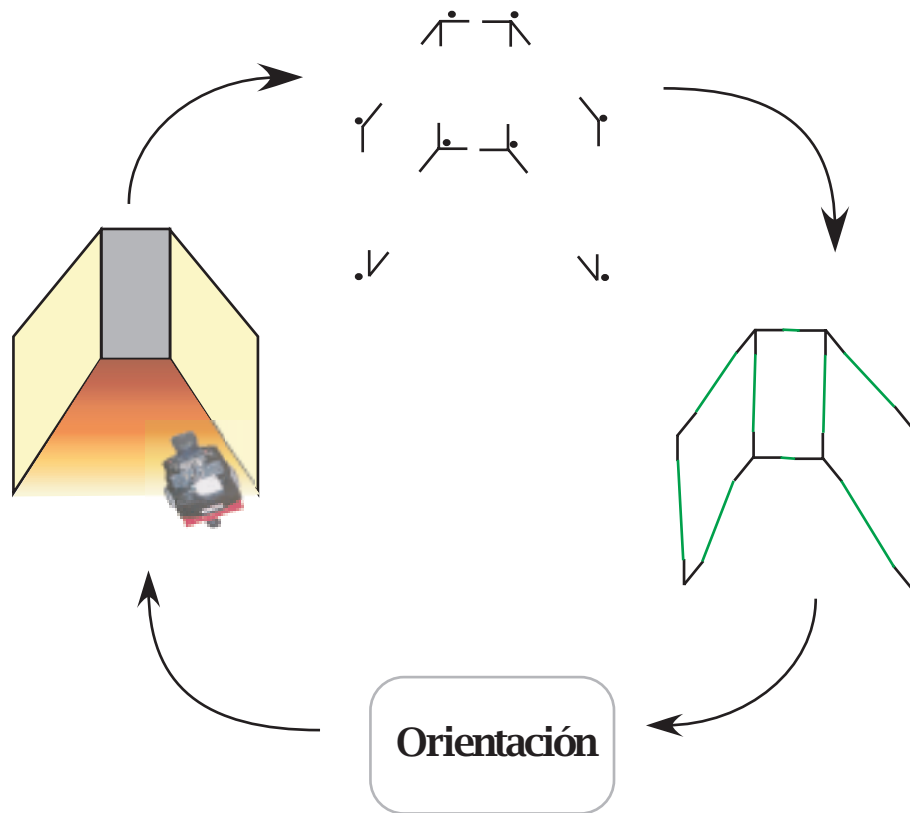


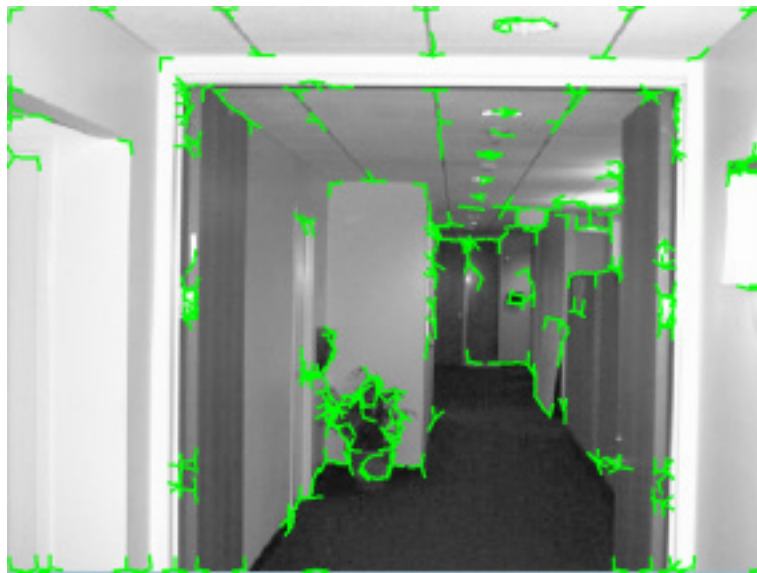
Figura 1.2: Proceso completo de aplicación de los distintos métodos propuestos.

## 1.2 Descripción del contenido de la tesis

Primero vamos a comentar cuál es el proceso de aplicación de los distintos métodos propuestos en esta tesis. Este proceso lo podemos contemplar en la Figura 1.2. Partimos de una imagen tomada por el robot. El primer paso consiste en detectar los puntos característicos de la imagen y en determinar si existe una unión en dicho punto. Al finalizar este primer paso disponemos de un conjunto de uniones, pero en algunas de ellas se van a plantear ciertos errores. Para solventar dichos errores y para obtener más características de la imagen, realizamos el segundo paso que consiste en el agrupamiento de las uniones. Disponemos ahora de un mapa con conexiones entre las uniones. Por último, utilizamos la información ob-

tenida para detectar el ángulo de orientación del robot. Este ángulo realimentará al robot, influyendo en su movimiento. A continuación comentamos la estructura general de la tesis.

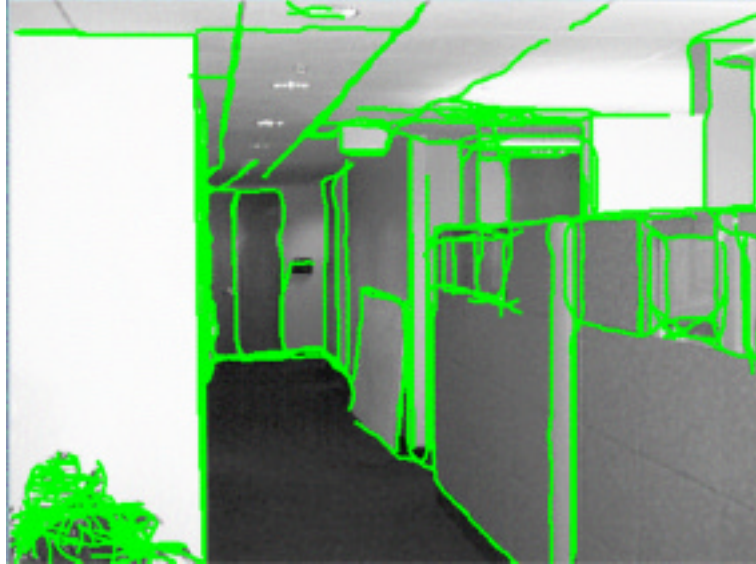
Esta tesis está estructurada de la siguiente manera: en el Capítulo 2 trataremos sobre cómo podemos extraer características de muy bajo nivel, con el menor coste computacional posible. En nuestro caso nos hemos basado en las uniones. Las uniones son puntos de la imagen donde convergen dos o más aristas. En el caso de dos aristas también podemos hablar de esquinas. Hablaremos indistintamente de uniones o esquinas. Propondremos un esquema en dos fases. La primera consistirá en detectar el centro de la unión. En la siguiente fase encontraremos los límites de las secciones angulares de la unión. Un ejemplo de obtención de uniones se muestra en la Figura 1.3.



**Figura 1.3:** Ejemplo de obtención de uniones en una imagen de gris.

En el Capítulo 3 utilizaremos las características obtenidas en el capítulo anterior como base para realizar un agrupamiento de éstas. Veremos que la detección de uniones lleva implícita un cierto error tanto de detección de límites inexistentes como de localización de dichos límites. Con la agrupación de uniones pretendemos llevar a cabo dos tareas. La primera es intentar solventar dichos errores. La segunda tarea que nos marcamos como objetivo es la de encontrar un mapa planar mínimo que mejor describa la imagen y la relación entre las uniones. En la

Figura 1.4 se muestra un ejemplo de agrupamiento de uniones.



**Figura 1.4:** Ejemplo de agrupamiento de uniones.

Por último, el Capítulo 4 resuelve el problema de encontrar el ángulo de orientación de la cámara de una manera eficiente. En este capítulo presentaremos un novedoso método de detección de dicho ángulo utilizando información bayesiana. Este método utiliza la restricción de que nos encontramos en un mundo de Manhattan, donde se cumple que la mayoría de los objetos diseñados por el hombre se alinean en dicho mundo. Por ello encontramos únicamente tres puntos de fuga. A continuación presentaremos unas ciertas mejoras al método original y cómo podemos utilizar la información de las aristas para obtener dicho ángulo en un tiempo mucho menor. La Figura 1.5 muestra un ejemplo de cálculo del ángulo de orientación y del etiquetado de los píxeles según el mundo de Manhattan.

Finalizamos con un capítulo de conclusiones y posibles ampliaciones de los métodos propuestos en esta tesis.

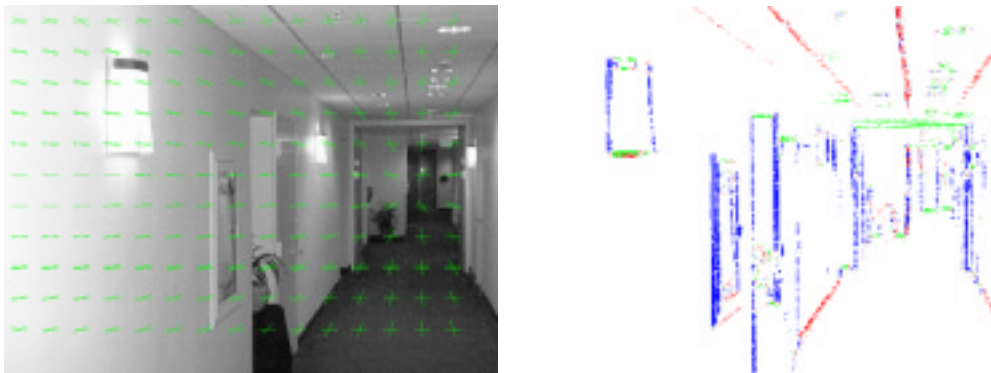


Figura 1.5: Cálculo del ángulo de orientación de la cámara.

