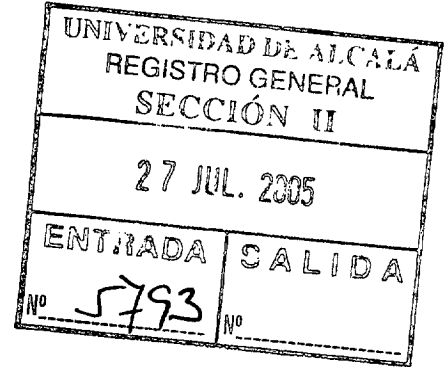


a 306557



T  
62-05  
UAH  
FER  
Sale

Universidad de Alcalá  
Escuela Politécnica Superior  
Departamento de Electrónica



# SISTEMA DE POSICIONAMIENTO ABSOLUTO DE UN ROBOT MÓVIL UTILIZANDO CÁMARAS EXTERNAS

Autor: Ignacio Fernández Lorenzo

Directores: Dr. Manuel Mazo Quintas  
Dr. José Luis Lázaro Galilea

2005

Tesis doctoral





Dr. Manuel Mazo Quintas, Catedrático de la Universidad de Alcalá, y

Dr. José Luis Lázaro Galilea, Profesor Titular de la Universidad de Alcalá

INFORMAN : Que la Tesis Doctoral titulada “Sistema de posicionamiento absoluto de un robot móvil utilizando cámaras externas”, presentada por D. Ignacio Fernández Lorenzo, y realizada bajo nuestra dirección, dentro del campo de la aplicación de los Espacios Inteligentes al guiado de robots móviles, reúne los méritos de calidad y originalidad para optar al Grado de Doctor.

Alcalá de Henares, a 18 de Julio de 2005

Fdo.: Manuel Mazo Quintas

Fdo.: José Luis Lázaro Galilea



Dr. Luis Miguel Bergasa Pascual, Director del Departamento de Electrónica de la Universidad de Alcalá,

INFORMA:

Que la Tesis Doctoral titulada “Sistema de posicionamiento absoluto de un robot móvil utilizando cámaras externas”, presentada por D. Ignacio Fernández Lorenzo, y dirigida por el Dr. Manuel Mazo Quintas y el Dr. José Luis Lázaro Galilea, cumple con todos los requisitos científicos y metodológicos para ser defendida ante un tribunal.

Alcalá de Henares, a 18 de Julio de 2005



Dr. Luis Miguel Bergasa Pascual

# Resumen

En esta Tesis Doctoral se ha realizado una nueva propuesta para el posicionamiento y guiado de robots móviles en espacios interiores, utilizando cámaras estáticas ubicadas en el entorno de movimiento y tratando de minimizar la infraestructura a bordo de los robots. La solución planteada está en la línea de lo que se conoce como “espacios inteligentes”, donde la “inteligencia” del sistema se traslada al entorno, siendo éste el encargado de captar y procesar toda la información necesaria para obtener la posición de los robots y generar las correspondientes órdenes de control, que son enviadas vía radio a los robots, en función de las trayectorias que se deseen seguir.

Los problemas más importantes que se plantean en una solución como la descrita en esta tesis son los relacionados con la localización de los robots y la obtención de las coordenadas 3D y orientación de los mismos. Para la identificación de los robots se ha propuesto la utilización de cuatro balizas circulares dispuestas en forma de “T” y construidas a partir de diodos de infrarrojos. Estas balizas ubicadas a bordo de los robots no solamente facilitan su localización, incluso en condiciones extremas de iluminación ambiental, sino que también simplifican los algoritmos de procesamiento de imágenes. La obtención de la posición y orientación de los robots se realiza a partir de los centroides de las balizas.

En la tesis se ha realizado un estudio detallado de las diferentes fuentes de error que influyen en la obtención de las coordenadas pixélicas del centroide de las balizas, y que tienen una incidencia importante en la obtención de las coordenadas 3D de los robots. Para minimizar sus efectos se han propuesto alternativas basadas en el procesamiento de imágenes.

La obtención de las coordenadas 3D de los robots, a partir de la información que en cada instante obtienen las cámaras de las balizas de a bordo, es una de las aportaciones más importantes de la tesis. En este sentido, la propuesta que se ha realizado, cuando el robot es visto por varias cámaras, se basa en minimizar un índice de comportamiento que tiene en cuenta la varianza del error en la determinación de las coordenadas de los centroides de las balizas en las diferentes cámaras. Esta solución se ha validado experimentalmente. Para el caso de que los robots sean vistos por una única cámara, se utiliza el dato de la altura del plano en el que se encuentran las balizas sobre los robots para la obtención de las coordenadas 3D. En cuanto a la orientación de los robots, si bien ésta se puede obtener a partir de su propio

movimiento, se ha propuesto una alternativa basada en la utilización de cuatro balizas dispuestas en forma de “T”.

Para la segmentación de las balizas y obtención de su centroide, se ha utilizado una estrategia completa de procesamiento de imágenes. Dado el fuerte contraste que presentan las balizas de infrarrojos(IR) con el fondo, para dicha segmentación se utilizan operadores tipo gradiente que permiten una fácil y robusta detección de bordes de las mismas. A partir de las imágenes de bordes se presenta una umbralización dinámica basada en el método de Otsu. Y teniendo presente que las balizas se proyectarán como elipses en las imágenes, para la localización de las mismas se utiliza el método de la distancia algebraica que permite obtener las elipses candidatas que más se aproximan a cada uno de los objetos segmentados. A partir de las elipses candidatas se aplican criterios de excentricidad, relación de perímetros y disposición geométrica entre las cuatro balizas, para descartar falsos candidatos. Las pruebas prácticas realizadas han demostrado la gran fiabilidad y robustez de la propuesta realizada. Una vez detectadas las balizas en la imagen, se ha propuesto una nueva alternativa para la determinación del centroide de las mismas. Dicha propuesta se basa en considerar como coordenadas de dichos centroides las de la intersección de las diagonales principales que unen los vértices del mínimo rectángulo que contiene la baliza. Esta propuesta ha demostrado en la práctica proporcionar unos resultados altamente satisfactorios, y presenta la gran ventaja de una reducida carga computacional, sobre todo si se compara con otras alternativas propuestas en la literatura y cuyos resultados desde el punto de vista de nuestra aplicación son similares.

Dada la importancia que en un sistema de este tipo tiene la calibración de las cámaras, se ha planteado un sistema de calibración multi-cámara que contempla dos fases. En una primera fase se obtienen para cada cámara los parámetros intrínsecos utilizando un patrón activo, también propuesto en esta tesis. En una segunda fase, y ya con las cámaras ubicadas en su localización definitiva dentro del entorno, se obtienen los parámetros de las matrices de rotación-traslación entre las diferentes cámaras. Con este método se evitan los errores que se cometen en la obtención de los parámetros intrínsecos si se posiciona el patrón de calibración a distancias alejadas de las cámaras. Los resultados prácticos obtenidos con esta alternativa de calibración propuesta han demostrado ser altamente satisfactorios.

Dentro de la tesis se ha desarrollado, así mismo, un sistema de guiado de robots móviles utilizando Splines cubicas para la generación de trayectorias. La posición del robot se obtiene a partir de la información del sistema de visión, añadiendo una predicción y corrección por medio del Filtro de Kalman. De esta forma se permite estimar la ubicación de las balizas a bordo del robot y en consecuencia seleccionar una ventana de interés sobre la que se aplican

los algoritmos de procesamiento de imágenes que permiten detectar los centroides de las balizas. Con ello se elimina una gran parte de ruido y al mismo tiempo se reduce el tiempo de cálculo de los algoritmos de visión.

Finalmente se ha implementado un sistema práctico de guiado, dentro del edificio de la Escuela Politécnica Superior de la Universidad de Alcalá, de un robot móvil utilizando las estrategias propuestas en esta tesis. Los resultados de estas pruebas han sido altamente satisfactorios, validando así las propuestas realizadas y demostrando que es posible realizar el guiado de robots móviles utilizando cámaras fijas ubicadas en el entorno de movimiento, con todas las ventajas que ello conlleva.



# Abstract

This Thesis proposes a new method for the positioning and guidance of indoor mobile robots, employing static cameras which are located in the surrounding environment and minimizing the equipment on the robotic platforms. The research field of the proposed solution belongs to what is known as “intelligent environments”. The external systems to the robot acquire and process the information in order to obtain the robot position and give the commands for the robot guidance.

The main problems of the given solution are robot location and obtaining the 3D coordinates and orientation. The robots are identified by means of the detection of a structure of four circular beacons arranged in “T” shape. These beacons are constructed with infrared diodes. Besides detecting the robot in poor illumination conditions also allow reducing the complexity of the algorithms of image processing. The position and orientation of the robots are extracted from the centroid of the beacons.

In the thesis a detailed study of the different error sources that have influence in the pixel coordinates obtained for the centroid of the beacons has been carried out. These errors have an important incidence obtaining the 3D coordinates of the robots. In this sense, different alternatives based on image processing have been proposed to minimize their effects.

One of the most important contributions of this thesis is the calculation of the 3D coordinates of the robots from the information obtained by the cameras capturing the beacons placed on-board the robot. In this sense, when the robot is seen for several cameras, the proposal that has been carried out it is based on minimizing a criterion that considers the error variance of the centroids coordinates of the beacons captured by the different cameras. This proposal has demonstrated to give excellent results in empirical tests. In the case that the mobile robots are captured by a single camera, the a priori data corresponding to the height of the beacons platform over the robots is used in order to obtain the 3D coordinates. The robot orientation is computed considering the four beacons placed a T shape.

The difference in the luminance level between the beacons and background permits to carry out a segmentation which uses a processing algorithm based in the image gradient. Using a dynamic threshold (Otsu's method) allows for an easy and robust detection of edges.



Afterwards, keeping in mind that the beacons were projected in the images as ellipses, an ellipsis locating method is proposed. From the ellipses candidates a criterion based on the blobs eccentricity, ratio of perimeters and geometric location among the four beacons is applied to discard false candidates. The practical tests have demonstrated a great reliability and robustness of the proposed method. After detecting the beacons in the image, a new alternative to improve the determination of their centroids has been developed. This proposal is based on considering as centroid coordinates the intersection of the main diagonals that connect the vertices of the minimum rectangle that circumscribe the beacon. This proposal provides high satisfactory results, and presents the advantage of lower computational cost than actual alternatives which have similar results from the point of view of our application.

Due to the great importance of camera calibration for a system of this type, a multi-camera calibration system has been developed considering two phases. First of all the intrinsic parameters for each camera are obtained using an active pattern, also proposed in this thesis. Second by, with the cameras located in their definitive location within the environment, the parameters of the rotation-translation matrix between the different cameras are obtained. With this method, the errors for the intrinsic parameters when the calibration pattern is positioned away from the cameras are avoided. The practical results obtained with this alternative of proposed calibration have shown to be highly satisfactory.

In this thesis a system of mobile robots guidance has also been developed using cubic Splines for path generation. The position of the robot is obtained from the information of the multi-camera system, adding a prediction and correction positioning values through a Kalman filter. The prediction of the position of the robots permits to estimate the location of the beacons on-board the robotic platform and consequently determine a region of interest to apply inside the algorithms responsible of the detection of the beacons centroids. This temporal and windowing filtering eliminates a large amount of noises and at the same time reduces the time of calculation of the algorithms.

Finally a practical guidance system of a mobile robot has been implemented, inside the building of the Escuela Politécnica Superior of Alcala University, making use of the proposed strategies in this thesis. The results of these tests have been highly satisfactory and have validated thus the developed proposals showing that it is possible to carry out the guidance of mobile robots using fixed cameras located in the surrounding environment, with all the advantages that it involves.

Las cosas se hacen  
o no se hacen.  
*Nunca se intentan.*



*A Elvira y Raquel.*



# Agradecimientos

En la vida es fundamental tener alrededor buenos amigos y compañeros con los que compartir las experiencias e inquietudes. Esta tesis no hubiera sido posible sin las muchas de esas personas que me han ayudado desde el principio, estando ahí para hablar y escuchar. Todas ellas han aportado su granito de arena para la realización de esta tesis.

Me gustaría agradecer sinceramente a mis directores de tesis, Manual Mazo Quintas y José Luis Lázaro Galilea por su continuo apoyo, ánimo, junto con un buen número de buenas ideas para que esta tesis llegase a buen puerto. Agradezco su exhaustiva revisión y crítica constructiva del texto final.

A mis amigos Enrique y José Antonio por su animo continuo y cooperación en las tareas de programación. A Felipe por sus revisiones exhaustivas y propuestas de mejoras. A Alfredo por sus consejos y aportaciones para la calibración del sistema de visión. A Elena por su inestimable ayuda en el los problemas acontecidos con el control del Peoplebot. A Miguel Ángel por sus consejos.

A esos compañeros, César, Manuel, Dani, Dori, Emilio, Raúl, Pedro, Carlos, Juan, Óscar que siempre hacen falta para seguir adelante, y no reparan en compartir su esfuerzo con los demás.

A mis compañeros de pasillo con su continuo aliento en esos momentos de desfallecimiento.

Y a todos esos compañeros que de una forma u otra han intervenido en la investigación presentada en esta tesis.

A todos aquellos alumnos que han colaborado, con el desarrollo de su Trabajo Fin de Carrera, en temas y trabajos relacionados con esta tesis.

Y a mi familia, por su eterna paciencia en algo que parece ha llegado a su fin.



# Índice

<b>Resumen</b> .....	i
<b>Abstract</b> .....	v
<b>Agradecimientos</b> .....	xi
<b>Índice</b> .....	xiii
<b>Índice de figuras</b> .....	xvii
<b>Índice de tablas</b> .....	xxvii
<b>1. Introducción.</b> .....	1
1.1. Motivación. ....	1
1.2. Objetivos planteados en la tesis. ....	2
1.3. Contexto de desarrollo de la tesis. ....	6
1.4. Estructura de la tesis. ....	6
<b>2. Revisión de trabajos previos</b> .....	9
2.1. Robots móviles. ....	10
2.1.1. Navegación de robots móviles. ....	10
2.2. Técnicas de posicionamiento de un robot móvil. ....	11
2.2.1. Métodos de posicionamiento relativos. ....	11
2.2.2. Métodos de posicionamiento absoluto. ....	12
2.3. Sistemas de posicionamiento de RM's utilizando visión artificial. ....	16
2.3.1. Sistemas que utilizan cámaras de video a bordo ( <i>on board</i> ). ....	16
Sistemas que utilizan una sola cámara. ....	17
Métodos que utilizan visión estéreo. ....	18
2.3.2. Sistemas que utilizan cámaras ubicadas en posiciones fijas (Espacios Inteligentes). ....	20
2.3.3. Guiado de robots utilizando cámaras ubicadas en puntos fijos del entorno. ....	26
2.4. Modelo de cámara. ....	35
2.4.1. Modelos sin distorsión. ....	37
2.4.2. Modelo de perspectiva con distorsión. ....	39
2.5. Calibración. ....	42
2.5.1. Patrones de calibración. ....	44
2.5.2. Calibración multicámara. ....	47
2.6. Conclusiones. ....	48



<b>3. Errores debidos al sistema de visión</b>	51
3.1. Modelo de cámara.	51
3.2. Percepción del entorno.	55
3.3. Identificación del robot.	58
3.4. Errores debidos al sistema de visión.	60
3.4.1. Errores debidos al tiempo de exposición de las cámaras.	61
3.4.2. Desplazamientos temporales debidos a la falta de sincronismo entre las cámaras y aspectos asociados al hardware de adquisición.	68
3.4.3. Errores debidos al modelo de cámara.	75
3.4.4. Errores debidos al proceso de calibración.	84
3.4.5. Errores debidos a la imprecisión en la obtención de las coordenadas pixélicas del centroide.	89
3.5. Conclusiones.	91
<b>4. Posicionamiento 3D utilizando cámaras externas</b>	93
4.1. Introducción.	93
4.2. Obtención de la posición del robot utilizando una única cámara.	94
4.2.1. Resultados prácticos.	100
4.3. Obtención de la posición utilizando dos cámaras.	102
4.3.1. Resultados prácticos.	111
4.4. Obtención de la posición del robot utilizando más de dos cámaras.	112
4.4.1. Resultados prácticos.	116
4.5. Conclusiones.	118
<b>5. Balizas de IR y procesamiento de imágenes para su identificación</b>	121
5.1. La baliza de Infrarrojos (IR).	122
5.1.1. Diseño práctico de las balizas.	124
5.2. Procesamiento de imágenes para la detección de las balizas a bordo del robot.	128
5.2.1. Segmentación.	129
5.3. Clasificación de los objetos.	138
5.3.1. Eliminación de objetos de pequeñas dimensiones	141
5.3.2. Identificación de elipses.	141
Ajuste mediante distancia algebraica.	142
Excentricidad y relación de perímetros.	144
5.3.3. Validación de candidatos.	146
5.4. Obtención del centroide de la baliza.	147
5.5. Obtención de la orientación del robot a partir de las balizas	151
5.6. Conclusiones.	153

<b>6. Calibración del sistema de visión</b> .....	155
6.1. Introducción. ....	156
6.2. Patrón de calibración. ....	157
6.2.1. Patrón de calibración propuesto. ....	157
6.2.2. Determinación automática de los puntos del patrón. ....	161
6.3. Proceso de calibración. ....	165
6.3.1. Determinación de los parámetros intrínsecos de las cámaras ....	166
6.3.2. Obtención de los parámetros extrínsecos. ....	168
Errores transmitidos de $R_{pgi}$ y $T_{pgi}$ a $R_i$ , $T_i$ . ....	173
6.3.3. Comprobación de resultados. ....	173
6.4. Conclusiones. ....	176
<b>7. Guiado</b> .....	181
7.1. Configuración general del sistema de guiado. ....	181
7.2. Modelado del entorno. ....	183
7.3. Generación de trayectorias. ....	187
7.3.1. Planificación con subgrafos de visibilidad. ....	187
7.3.2. Planificación durante la generación de trayectorias. ....	188
7.3.3. Interpolación mediante Splines. ....	189
7.4. Sistema de posicionamiento. ....	195
7.4.1. Filtro de Kalman. ....	197
7.5. Generación de consignas. ....	200
7.6. Resultados prácticos. ....	204
7.7. Conclusiones. ....	209
<b>8. Conclusiones y trabajos futuros</b> .....	211
8.1. Conclusiones y aportaciones. ....	212
8.1.1. Errores debidos al sistema de visión. ....	213
8.1.2. Posicionamiento 3D utilizando cámaras externas. ....	214
8.1.3. Balizas de IR y procesamiento de imágenes para su identificación	215
8.1.4. Calibración del sistema de visión .....	217
8.1.5. Guiado .....	218
8.2. Trabajos futuros. ....	219
8.3. Publicaciones derivadas de la tesis. ....	221
<b>Apéndice A. Calibración de cámaras con distancia focal fija</b> .....	223
A.1. Modelo <i>pin-hole</i> . ....	223
A.2. Calibración multi-imagen. ....	225
A.2.1. Detección sub-pixélica de los círculos del patrón. ....	230
<b>Bibliografía</b> .....	233



# Índice de figuras

Figura 1.1.	Estructura general del sistema de posicionamiento y guiado propuesto en esta tesis. El sistema sensor está formado por $n$ cámaras ubicadas en posiciones fijas del entorno, todas ellas conectadas a un procesador central que determina la posición del robot. Con la información proporcionada por las cámaras se generan las consignas de velocidad de las ruedas motrices, las cuales son enviadas a los robots vía radio.	3
Figura 1.2.	Estructura del sistema de control utilizado para realizar el guado de robots por un entorno a partir de la información proporcionada por un array de cámaras estática. El sistema de posicionamiento proporciona la posición 3D a partir de la información captada por el sistema sensorial (cámaras). Por su parte el módulo de guiado es el encargado de generar todas las tareas implicadas en el movimiento del robot.	4
Figura 2.1.	Sistema de visión catadióptico empleado en [Mouaddib 00]. (a) Estructura de la lente utilizada, en la que se puede observar el espejo reflector cónico que permite captar un campo de visión de 360°. (b) Montaje del sensor CCD en conjunción con la lente. (c) Ejemplo de una imagen de un entorno captada con el sistema de visión omnidireccional	17
Figura 2.2.	Sistema catadióptico estéreo utilizado en [Fiala 02]. (a) Estructura del sensor formado por una cámara CCD y dos espejos reflectantes semiesféricos. (b) Aspecto de una imagen captada por el sensor estéreo. (c) Plataforma móvil utilizada en las pruebas.	18
Figura 2.3.	Estructura del entorno Intelligent Room del Artificial Intelligence Laboratory del MIT.	22
Figura 2.4.	Entorno KidsRoom del Media Laboratory del MIT.	23
Figura 2.5.	Entorno ROR del Intelligent Cooperative Systems Laboratory de la Universidad de Tokio. (a) Elementos del entorno. (b) Vista del brazo manipulador.	25
Figura 2.6.	Sistema distribuido utilizado para los experimentos en [Soho 99].	27
Figura 2.7.	ISpace propuesto en [Lee 99]. (a) Estructura. (b) Entorno de pruebas.	27
Figura 2.8.	Tipos de marcas artificiales utilizadas para detectar a los robots en el ISpace. (a) Cuatro marcas. (b) Tres marcas.	28
Figura 2.9.	Estructura general del sistema MEPHISTO.	29
Figura 2.10.	Representación con polígonos los objetos móviles en las imágenes de las cámaras del sistema MEPHISTO.	30
Figura 2.11.	Mapa de ocupación obtenido de los trabajos realizados en [Hoover 00]. (a) Entorno de trabajo utilizado en las pruebas. (b) Mapa de ocupación realizado utilizando la información de una única cámara. (c) Mapa de ocupación obtenido con la utilización de dos cámaras. (d) Mapa de ocupación utilizando tres cámaras.	31
Figura 2.12.	Plantilla utilizada para calibrar las cámaras del sistema de visión propuesto en [Hoover 00].	32

Figura 2.13.	Sistema de calibración propuesto por Hoover y Olsen.	33
Figura 2.14.	Entorno de trabajo utilizado por Kruse y Wahl para sus trabajos.	33
Figura 2.15.	Sistema de posicionamiento propuesto [Lilienthal. 03]. (a) Robot empleado en las pruebas. (b) ubicación de las cámaras en el entorno de trabajo.	34
Figura 2.16.	Array de diodos IR utilizados en [Aitenbichler 03] para identificar el robot en una imagen.	35
Figura 2.17.	Relaciones utilizadas en [Yun 04] para determinar la posición 3D de un robot utilizando una única cámara. (a) Modelo de correspondencia para obtener la posición 3D del robot. (b) Representación de las coordenadas pixélicas del robot en una imagen.	36
Figura 2.18.	Modelo de proyección perspectiva.	37
Figura 2.19.	Proyección perspectiva con distorsión.	40
Figura 2.20.	Plantilla en forma de damero de ajedrez. (a) Estructura. (b) Detección de bordes.	45
Figura 2.21.	Plantillas 3D de calibración. (a) Con objetos a diferentes alturas. (b) Con forma de cubo.	46
Figura 2.22.	Plantilla de calibración con elementos circulares.	46
Figura 3.1.	Modelo <i>pin-hole</i> de una cámara.	52
Figura 3.2.	Relación entre las coordenadas métricas $(x,y)$ y pixélicas del plano imagen $(u,v)$ .	53
Figura 3.3.	Superficie vista por diferente número de cámaras. El eje de ordenadas representa el número de cámaras, mientras que los otros dos representan las dimensiones en metros del entorno. (a) Utilizando un total de cuatro cámaras. (b) Utilizando un total de seis cámaras.	55
Figura 3.4.	Porcentaje de la superficie del entorno de la figura 3.3 captado por un determinado número de cámaras. (a) Utilizando un total de cuatro cámaras. (b) Utilizando un total de seis cámaras.	56
Figura 3.5.	Porcentaje del entorno de la figura 3.3 captado, en función del número de cámaras (para el caso de utilizar un total de seis cámaras), para diferentes tamaños del CCD y una óptica de 7.5 mm. (a) CCD de 1/2". (b) CCD de 1/3". (c) CCD de 1/4".	57
Figura 3.6.	Porcentaje del entorno de la figura 3.3 captado, en función del número de cámaras (para un total de seis cámaras), para CCD's de 2/3" y ópticas de diferentes distancias focales. (a) Distancia focal de 3 mm. (b) Distancia focal de 10 mm. (c) Distancia focal de 20 mm.	58
Figura 3.7.	Tamaño de la escena correspondiente a un píxel en función de la distancia focal y tamaño del CCD de la cámara. (a) Para puntos separados de la cámara 4 m. (b) Para puntos separados de la cámara 8m. (c) En función de la distancia para un CCD de 2/3".	59
Figura 3.8.	Baliza de IR utilizada a bordo del robot que facilita la identificación de su posición y orientación. (a) Distribución de los emisores de infrarrojos $(A, B, C, D)$ . (b) Ejemplo de detección de las balizas a bordo de un prototipo de robot.	60
Figura 3.9.	Ejemplo de detección de la posición del robot utilizando una única baliza de IR a bordo.	60

Figura 3.10.	Imágenes de un objeto en movimiento tomadas con cámaras con diferentes tiempos de exposición.	62
Figura 3.11.	Deformación de la proyección sobre el plano imagen de una baliza circular de radio 20mm, como consecuencia del tiempo de exposición de la cámara y movimiento de la baliza.. (a) $\ V\ =0$ . (b) $\ V\ =500$ mm/s y $t_{exp}=1/100$ s. (c) $\ V\ =2000$ mm/s y $t_{exp}=1/100$ s. (d) $\ V\ =2000$ mm/s y $t_{exp}=1/1000$ s.	64
Figura 3.12.	Error en la determinación de las coordenadas pixélicas de una baliza que se mueve a una velocidad con componentes $V_x=V_y=2$ m/s y $V_z=0$ , para dos tiempos de exposición de la cámara. (a) Con un tiempo de exposición de 1/100 s. (b) Con un tiempo de exposición de 1/1000 s. (c) Error en unidades métricas, teniendo en cuenta la resolución, para la figura (a). (d) Error en unidades métricas, teniendo en cuenta la resolución, para la figura (b).	65
Figura 3.13.	Errores de las coordenadas del centroide de la baliza en movimiento en función del tiempo de exposición de la cámara, para dos posiciones ( $P_1$ y $P_2$ ) de la baliza. (a) Coordenadas pixélicas. (b) Coordenadas métricas.	66
Figura 3.14.	Error (en píxeles) de la coordenada $u$ del centroide de la baliza con ésta en movimiento, para diferentes tiempos de exposición de la cámara, posiciones de la baliza y tamaño de sensor CCD. (a) CCD con tamaño 1/4". (b) CCD con tamaño 1/3". (c) CCD con tamaño 1/2".	66
Figura 3.15.	Error en unidades métricas correspondientes a los errores pixélicos de las figuras 3.14. (a) Figura 3.14.a (b) Figura 3.14.b (c) Figura 3.14.c.	67
Figura 3.16.	Error (en píxeles) de la coordenada $u$ del centroide de la baliza en función del tiempo de exposición de la cámara, para diferentes distancias focales, posiciones de la baliza y velocidades del robot. (a) Distancia focal de 3 mm. (b) Distancia focal de 10 mm. (c) Distancia focal de 20 mm.	68
Figura 3.17.	Error en mm para las gráficas de la figura 3.16, teniendo en cuenta la resolución de la cámara. (a) Figura 3.16.a. (b) Figura 3.16.b. (c) Figura 3.16.c.	69
Figura 3.18.	Error en las coordenadas del centroide de la baliza entre dos cámaras, para diferentes valores de velocidad del robot, posición de la baliza y para diferentes valores de desplazamiento temporal entre las mismas. (a) Error en coordenadas pixélicas. (b) Error en coordenadas métricas.	71
Figura 3.19.	Error en las coordenadas pixélicas del centroide de la baliza entre dos cámaras, para diferentes tamaños de CCD, diferentes valores de velocidad del robot y diferentes valores de desplazamiento temporal entre las cámaras. (a) CCD con un tamaño de 1/4". (b) CCD con un tamaño de 1/3". (c) CCD con un tamaño de 1/2".	71
Figura 3.20.	Error en las coordenadas pixélicas del centroide de la baliza entre dos cámaras, para ópticas de diferentes distancias focales, diferentes velocidades del robot y diferentes valores de desplazamiento temporal entre las cámaras.. (a) Distancia focal 3mm. (b) Distancia focal 10 mm. (c) Distancia focal 20 mm.	72
Figura 3.21.	Errores en mm para las gráficas de las figura 3.19. (a) Figura 3.19.a. (b) Figura 3.19.b. (c) Figura 3.19.c.	73
Figura 3.22.	Errores en coordenadas métricas para las gráficas de la figura 3.20. (a) Figura 3.20.a. (b) Figura 3.20.b. (c) Figura 3.20.c.	74

Figura 3.23.	Errores en las coordenadas píxelicas del centroide de la baliza debidos a la digitalización secuencial para un sistema formado por 16 cámaras y con tiempos de adquisición de cada una de ellas, $t_a = 40\text{ms}$ , para distintas posiciones de la baliza. (a) Error en coordenadas píxelicas. (b) Error en coordenadas métricas.	74
Figura 3.24.	Error de las coordenadas píxelicas del centroide de una baliza para un sistema de 15 cámaras monocromo utilizando tarjetas digitalizadoras con entrada RGB. (a) Error en coordenadas píxelicas. (b) Error en coordenadas métricas.	75
Figura 3.25.	Efecto de las componentes de distorsión introducida por la óptica: componentes radial y tangencial.	76
Figura 3.26.	Efecto sobre las imágenes de la distorsión de la óptica de la cámara. (a) Utilizando una óptica con una fuerte distorsión. (b) Utilizando una óptica con baja distorsión.	76
Figura 3.27.	Representación del desplazamiento sufrido por las coordenadas píxelicas de una imagen debido a la distorsión, en verde se representa su posición ideal y en rojo las reales teniendo en cuenta la distorsión (la línea azul refleja el emparejamiento entre ellos). (a) Óptica con alta distorsión (figura 3.26.a). (b) Óptica con baja distorsión (figura 3.26.b).	78
Figura 3.28.	Valor del desplazamiento en las coordenadas $(u, v)$ de los píxeles de una imagen debido a la distorsión introducida por la óptica. (a) Para una óptica similar a la utilizada en la captación de la imagen de la figura 3.26.a. (b) Para una óptica similar a la utilizada en la captación de la imagen de la figura 3.26.b.	78
Figura 3.29.	Proceso iterativo dado por la ecuación (3.23) empleado para corregir la distorsión de un píxel.	79
Figura 3.30.	Errores debidos a la distorsión de cámara, en función del número de iteraciones realizadas para corregir la distorsión utilizado en esta tesis. (a) error máximo. (b) error medio.	80
Figura 3.31.	Efecto de la corrección de la distorsión de la óptica. (a) para el ejemplo de la imagen de la figura 3.27.a. (b) para el ejemplo de la imagen de la figura 3.27.b.	81
Figura 3.32.	Valor del desplazamiento en las coordenadas $(u, v)$ de los píxeles tras corregir la distorsión. (a) Para la óptica utilizada en la imagen de la figura 3.26.a. (b) Para la óptica utilizada en la imagen de la figura 3.26. c.	81
Figura 3.33.	Desplazamiento de las coordenadas del centroide de la baliza circular debido a la distorsión. (a) Campo de visión de la cámara. (b) Variación de la coordenada $u$ del centroide. (c) Variación de la coordenada $v$ del centroide.	82
Figura 3.34.	Desplazamiento de las coordenadas del centroide de la baliza una vez realizada la corrección de la distorsión. (a) Coordenada $u$ realizando la corrección sobre todos los píxeles correspondientes a la proyección de la baliza. (b) Coordenada $v$ realizando la corrección sobre todos los píxeles correspondientes a la proyección de la baliza. (c) Coordenada $u$ realizando la corrección sólo del centroide de la proyección de la baliza. (d) Coordenada $v$ realizando la corrección sólo del centroide de la proyección de la baliza.	83

Figura 3.35.	Valores experimentales reales del error que se comete en la determinación de la posición de una baliza circular para un sistema formado por tres cámaras, con y sin corrección de la distorsión. (a) Distribución de los puntos de prueba y ubicación de las cámaras. (b) Error utilizando la información de una única cámara para obtener la posición de la baliza. (c) Error utilizando dos cámaras. (d) Error utilizando tres cámaras.	84
Figura 3.36.	Error cometido en la obtención de las coordenadas pixélicas del centroide de la baliza, en función de la posición de la baliza respecto a la cámara, para $e_x = e_y = 1$ píxel (a) Error en la coordenada $u$ . (b) Error en la coordenada $v$ .	86
Figura 3.37.	Error en las coordenadas pixélicas del centroide de la baliza en función del error en la obtención de los parámetros intrínsecos $f_x$ y $f_y$ ( $e_{f_x}$ y $e_{f_y}$ )	86
Figura 3.38.	Error cometido en las coordenadas pixélicas del centroide de la baliza para un error en el cálculo del ángulo $\alpha$ de $0.4^\circ$ . (a) Error en la coordenada $u$ . (b) Error en la coordenada $v$ .	88
Figura 3.39.	Error máximo de las coordenadas pixélicas del centroide de la baliza en función de las variaciones de los ángulos de Euler. (a) Suponiendo errores en el ángulo $\alpha$ . (b) Suponiendo errores en el ángulo $\beta$ . (c) Suponiendo errores en el ángulo $\gamma$ .	88
Figura 3.40.	Errores en las coordenadas pixélicas del centroide de la baliza en función de los errores en los coeficientes de la matriz de traslación. (a) Para el coeficiente $T_x$ . (b) Para el coeficiente $T_y$ . (c) Para el coeficiente $T_z$ .	89
Figura 3.41.	Propuesta para la obtención sub-pixélica de las coordenadas del centroide de las balizas. (a) caso de cuatro balizas a bordo del robot. (b) caso de una baliza.	90
Figura 3.42.	Comparación entre los métodos propuesto en esta tesis y el propuesto en [Lavest 99] para la obtención de las coordenadas pixélicas del centroide de la baliza. (a) Coordenada $u$ . (b) Coordenada $v$ . En ambos casos se representa la diferencia entre los resultados obtenidos con las dos propuestas.	90
Figura 4.1.	Error de posicionamiento utilizando una única cámara, para dos valores de $\sigma_u$ y $\sigma_v$ . (a) $\sigma_u = \sigma_v = 0.2$ píxeles. (b) $\sigma_u = \sigma_v = 2$ píxeles	96
Figura 4.2.	Representación de los estadísticos del error de posicionamiento para el caso de utilizar una única cámara, en función del error en la obtención de las coordenadas pixélicas del centroide de la baliza. La baliza se ha ubicado en el punto del campo de visión de la cámara para el cual el error de posicionamiento es máximo (a) Valor máximo. (b) Valor medio. (c) Desviación típica.	96
Figura 4.3.	Error de posicionamiento de la baliza para $\sigma_u = \sigma_v = 2$ píxeles en función de la altura a la que se ubica la cámara. (a) altura de 2.5 m. (b) altura de 4 m.	97
Figura 4.4.	Error de posicionamiento utilizando una única cámara, para diferentes valores de $\sigma_u$ y $\sigma_v$ , suponiendo pequeñas irregularidades en altura en superficie por la que se mueve el robot. Se ha considerado una variación en altura de media nula y desviación típica ( $\sigma_z$ ) de 10mm. (a) $\sigma_u = \sigma_v = 0.2$ píxeles (b) $\sigma_u = \sigma_v = 2$ píxeles.	98
Figura 4.5.	Estadísticos del error de posicionamiento en función del error en $Z_0$ , para diferentes valores de $\sigma_u$ y $\sigma_v$ . (a) $\sigma_u = \sigma_v = 0.2$ píxeles (b) $\sigma_u = \sigma_v = 2$ píxeles.	99



Figura 4.6.	Representación del error de posicionamiento, utilizando un única cámara, en un entorno de 8x8 m sobre el que se han ubicado cuatro cámaras equiespaciadas y tomando la más cercana a la baliza para obtener su posición, considerando $\sigma_u = \sigma_v = 2$ píxeles. (a) incertidumbre de la posición. (b) error de posicionamiento.	99
Figura 4.7.	Valor del error de posicionamiento para las condiciones de trabajo de la figura 4.6, al que se le añade un error gaussiano en la coordenada $Z_0$ de media nula y desviación típica 10 mm.	100
Figura 4.8.	Error prácticos de posición del robot utilizando una única cámara, para un entorno de 6x4 m sobre el que se han ubicado cuatro cámaras. Se reflejan los valores sin corregir y corrigiendo la distorsión. (a) Utilizando la cámara más cercana al robot (b) Utilizando la cámara más alejada del robot (c) Ubicación física de los puntos de test y de las cámaras.	101
Figura 4.9.	Concepto de triangulación, a partir de dos cámaras, para determinar las coordenadas 3D de un punto $P$ .	103
Figura 4.10.	Ilustración gráfica del punto de cruce de las rectas de proyección de dos cámaras	104
Figura 4.11.	Ejemplos simulados del error de posicionamiento utilizando dos cámaras. (a) Campo de visión de las cámaras. (b) Caso de contaminar las medidas con un ruido de media nula y desviación típica de 0.2 píxeles. (c) Ruido con desviación típica de 2 píxeles.	110
Figura 4.12.	Comparación entre el método propuesto en esta tesis y el propuesto por [Aitenbichler 03][Hartley 03]. (a) Para el caso de un ruido gaussiano de media cero y desviación típica de 0.2 píxeles. (b) ídem, para una desviación típica de 2 píxeles.	111
Figura 4.13.	Error de posicionamiento de un sistema formado por seis cámaras, utilizando, para la obtención de la posición del robot, las dos cámaras más próximas a éste. (a) Con un error en las coordenadas píxelicas de desviación típica 0.2 píxeles. (b) Con un error en las coordenadas píxelicas de desviación típica 2 píxeles.	111
Figura 4.14.	Valores reales del error de posicionamiento para un entorno de 6x4 m sobre el que se han dispuesto cuatro cámaras, utilizando las dos más cercanas al robot para determinar su posición. Se reflejan los valores sin y con corrección de distorsión.	112
Figura 4.15.	Representación del punto óptimo $P$ que minimiza la distancia a las rectas de proyección de $n$ cámaras, ejemplo para $n=5$ , obtenido a partir de los posibles puntos obtenidos por el cruce entre dos rectas.	113
Figura 4.16.	Representación del error de posicionamiento de un entorno rectangular de 8x8m en el que se han ubicado seis cámaras. La posición del robot se obtiene utilizando todas las cámaras que en cada punto lo captan. (a) Desviación típica de 0.2 píxeles. (b) Desviación típica de 2 píxeles.	115
Figura 4.17.	Ejemplo de estudio comparativo del error de posicionamiento, en función del número de cámara utilizadas para determinar la posición del robot.	115
Figura 4.18.	Comparación entre los resultados de posicionamiento obtenidos utilizando la propuesta realizada en esta tesis y la realizada en [Bouguet 99]. Se representan los valores del error medio de posicionamiento para el caso de utilizar dos, tres, cuatro, cinco y seis cámaras. (a) Para una varianza de 0.2 píxeles. (b) Para una varianza de 2 píxeles.	116
Figura 4.19.	Valores prácticos del error de posicionamiento para el caso de utilizar más de dos cámaras. (a) tres cámaras. (b) cuatro cámaras.	116

Figura 4.20.	Resultados de los errores de posicionamiento de un robot en función del número de cámaras utilizadas para calcular su posición. (a) Sin realizar la corrección de la distorsión. (b) Corrigiendo la distorsión de las ópticas.	117
Figura 5.1.	Balizas utilizadas a bordo del robot para facilitar su identificación. (a) Caso de utilización de una única baliza. (b) Caso de utilización cuatro balizas, que facilita la identificación del robot y su orientación.	123
Figura 5.2.	Respuestas espectrales de: cámara CCD comercial (línea de color azul), filtro fotóptico ubicado entre el CCD y la óptica (línea de color verde) y silicio (línea de color rojo). Está última alternativa es la utilizada en esta tesis.	123
Figura 5.3.	Patrón de radiación de dos diodos IR. (a) Diodo muy directivo. (b) Diodo con una radiación casi plana.	124
Figura 5.4.	Vista detalla de las cuatro balizas de diodos IR.	125
Figura 5.5.	Ejemplo de captación de la baliza de IR a bordo del robot para dos situaciones de distancia entre robot y cámara. (a) Robot situado a 4 m de la cámara. (b) Niveles de gris de los píxeles de la imagen captada. (c) Robot situado a 10 m de la cámara. (d) Niveles de gris de los píxeles de la imagen correspondiente.	125
Figura 5.6.	Ejemplos de imágenes captadas del robot ante diferentes condiciones de iluminación del entorno. (a) Iluminación artificial. (b) Niveles de grises correspondientes a la figura (a). (c) Presencia de reflejos debidos al sol. (d) Niveles de grises para la imagen (c). (e) Ausencia de iluminación. (f) Niveles de grises para la imagen (e).	126
Figura 5.7.	Reducción del efecto de los brillos utilizando polarizadores. (a) Enmascaramiento de la baliza por un brillo. (b) Cámara, óptica y polarizador utilizados en nuestra aplicación. (c) Imagen del la figura (a) tomada con un cámara cuya óptica incorpora un polarizador lineal.	127
Figura 5.8.	Proceso general de localización de las balizas.	128
Figura 5.9.	Ventana de interés sobre la que se realiza el proceso de segmentación.	130
Figura 5.10.	Etapas incluidas en la propuesta de segmentación de las balizas de IR.	130
Figura 5.11.	Detección de bordes aplicando diferentes máscaras para emular el operador gradiente. (a) Imagen de entrada. (b) Mascara $H_v = [-1, 1]$ . (c) Mascara de Sobel. (d) Mascara de Prewitt. (e) Mascara de Frei-Chen. (f) Máscaras de Roberts.	132
Figura 5.12.	Resultado de aplicar la máscara de Prewitt a las imágenes de la figura 5.6, reflejando la ventana de interés de la imagen. (a) Imagen de la Figura 5.6.a. (b) Imagen de la figura 5.6.c (c) Imagen de la figura 5.6.e.	133
Figura 5.13.	Histogramas de las imágenes de la figura 5.12. (a) Imagen de la figura 5.12.a. (b) Imagen de la figura 5.12.b (c) Imagen de la figura 5.12.c.	134
Figura 5.14.	Resultado de realizar la umbralización de las imágenes de la figura 5.12. (a) Imagen figura 5.12.a, $T_k=48$ . (b) Imagen figura 5.12.b, $T_k=41$ . (c) Imagen figura 5.12.c, $T_k=45$ .	136
Figura 5.15.	Ejemplos de histogramas de las imágenes de la figura 5.12 tras la realización de la entropía. (a) Figura 5.12.a. (b) Figura 5.12.b. (c) Figura 5.12.c.	137

Figura 5.16.	Resultado de umbralización utilizando la entropía y $T_k=254$ para las imágenes de la figura 5.12. (a) Imagen figura 5.12.a. ( b) Imagen figura 5.12.b. (c) Imagen figura 5.12.c.	138
Figura 5.17.	Resultado de realizar la erosión y dilatación sobre las imágenes de la figura 5.14. (a) Imagen figura 5.14.a. (b) Imagen figura 5.14.b. (c). Imagen figura 5.14.c.	138
Figura 5.18.	Organigrama de clasificación seguido para identificar las balizas IR en una imagen.	140
Figura 5.19.	Ejemplos de “elipses estimadas” utilizando la distancia algebraica. (a) Caso de una estimación correcta (color rojo) y tres falsas (color azul). (b) Caso de cuatro balizas estimadas correctamente y una falsa. (c) Ejemplo de “elipse estimada” falsa que puede ser eliminada tras la comparación de perímetros. (d) Ejemplo de “elipse estimada” falsa que puede ser descartada por excentricidad.	141
Figura 5.20.	Representación de los parámetros $d_w$ , $E$ y $R_p$ para cada una de las “elipses estimadas” asociadas a cada uno de los objetos de la imagen. (a) Imagen con cuatro balizas. (b) Imagen con una baliza.	146
Figura 5.21.	Estrategia utilizada en la validación de candidatos a balizas a partir de su disposición geométrica.	147
Figura 5.22.	Obtención de las coordenadas del centroide de la baliza como intersección de las diagonales del rectángulo mínimo que la contiene.	148
Figura 5.23.	Representación de la luminancia de la imagen de la figura 5.11.a (en color azul) y su aproximación parabólica (color rojo). (a) Luminancia horizontal. (b) Luminancia vertical.	149
Figura 5.24.	Representación de las coordenadas de los centroides de la baliza para las imágenes de la figura (a) 5.6.a, (b) 5.6.b y (c) 5.6.c, obtenidas utilizando diferentes alternativas de cálculo.	150
Figura 5.25.	Distribución de las cuatro balizas situadas a bordo del robot. (a) Distribución y nomenclatura utilizada. (b) Ejemplo de identificación.	152
Figura 5.26.	Identificación de los centroides correspondientes a las cuatro balizas de la figura 5.24.b.	152
Figura 6.1.	Patrón de calibración utilizado en [Hoover 00].	159
Figura 6.2.	Patrón de calibración activo formado por diodos IR.	160
Figura 6.3.	Vista del patrón de calibración propuesto y comparación con un patrón formado por círculos negros.	161
Figura 6.4.	Identificación de los puntos de los patrones de calibración de la figura 6.3.	161
Figura 6.5.	Organigrama empleado en la obtención del centroide de cada uno de los objetos del patrón de calibración.	162
Figura 6.6.	Proyección de los puntos de un patrón de calibración sobre el plano imagen de una cámara. (a) Diodos IR del patrón propuesto. (b) Círculos reflectantes de un patrón estándar.	162
Figura 6.7.	Distribución de los objetos de calibración en el patrón.	163
Figura 6.8.	Organigrama seguido para la ordenación de los objetos del patrón de calibración.	164

Figura 6.9.	Distribución aleatoria de los índices de los puntos del patrón de calibración tras la segmentación e identificación de objetos.	164
Figura 6.10.	Representación gráfica de los pasos seguidos para reordenar los puntos del patrón de calibración.	166
Figura 6.11.	Representación de la figura 6.8 con los puntos reordenados de acuerdo a la numeración de la figura 6.7.	167
Figura 6.12.	Imágenes utilizadas para obtener los parámetros intrínsecos de una cámara. (a) Situada en una plataforma externa. (b) Situada en su posición final.	169
Figura 6.13.	Convergencia de los valores tomados por los parámetros intrínsecos de una cámara sometida al proceso iterativo de calibración en dos fases y utilizando la plantilla propuesta. (a) Valores de $f_x$ y $f_y$ . (b) Valores de $u_0$ y $v_0$ . (c) Valores de los coeficientes de la distorsión.	170
Figura 6.14.	Relación entre los sistemas de referencia global, de la cámara y del patrón.	172
Figura 6.15.	Representación del sistema de referencia del patrón de calibración.	172
Figura 6.16.	Medida de la posición del robot utilizando una única cámara.	173
Figura 6.17.	Relación entre los sistemas de referencia para el caso de $n$ cámaras.	174
Figura 6.18.	Proceso de medida de la posición del robot con $n$ cámaras.	175
Figura 6.19.	Influencia del error en la determinación de los ángulos de Euler de la matriz $R_{pgi}$ en los ángulos de Euler de la cámara. (a) Error en $\alpha_{pgi}$ . (b) Error en $\beta_{pgi}$ . (c) Error en $\gamma_{pgi}$ .	176
Figura 6.20.	Variación de los parámetros del vector de traslación de una cámara debida al error en la posición del patrón de calibración. (a) Error en $T_{Xpgi}$ . (b) Error en $T_{Ypgi}$ . (c) Error en $T_{Zpgi}$ . (d) Error en $\alpha_{pgi}$ . (e) Error en $\beta_{pgi}$ . (f) Error en $\gamma_{pgi}$ .	177
Figura 6.21.	Patrones de calibración utilizados para la obtención de los parámetros extrínsecos de una cámara. (a) Plantilla propuesta. (b) Plantilla utilizada con la herramienta de calibración de Matlab <sup>®</sup> .	178
Figura 6.22.	Calculo del error de posicionamiento utilizando el método de calibración propuesto en esta tesis (en color azul) y la herramienta de calibración de Matlab <sup>®</sup> (en color rojo).	179
Figura 7.1.	Estructura del sistema de control utilizado para realizar el guiado de robots por un entorno a partir de la información proporcionada por un array de cámaras estáticas y utilizando la odometría como sistema auxiliar.	182
Figura 7.2.	Ejemplo de mapa geométrico de un entorno.	184
Figura 7.3.	Definición de los elementos de la configuración o posición de un robot móvil en el plano XY.	184
Figura 7.4.	Obtención del espacio libre de obstáculos para el caso de un robot circular de radio $\rho$ . (a) Estructura del entorno. (b) Espacio libre tras la dilatación de los obstáculos y contracción del polígono límite.	186
Figura 7.5.	Obtención del espacio libre de obstáculos. (a) Entorno original. (b) Expansión de obstáculo y contracción del entorno. (c) Representación (en color azul) del espacio libre.	186

Figura 7.6.	Generación de una ruta entre un punto de partida $x_a$ y uno de llegada $x_f$ utilizando subgrafos de visibilidad.	188
Figura 7.7.	Generación de trayectorias. (a) Método de planificación durante la generación de trayectorias. (b) Método de planificación con subgrafos de visibilidad.	189
Figura 7.8.	de una Spline cúbica parametrizada con las funciones $fX(w)$ y $fY(w)$ para una trayectoria con cinco puntos de paso.	191
Figura 7.9.	Generación de una Spline para diferentes valores del factor de suavidad, $\mu$ .	194
Figura 7.10.	Trazado de un Spline con puntos de paso próximo. (a) Construyéndola sin dividirla por tramos. (b) Construyéndola dividiéndola en tramos.	195
Figura 7.11.	Robot Peoplebot de Active Media utilizado en las pruebas prácticas. (a) Vista del robot. (b) Definición de los parámetros proporcionados por la odometría.	196
Figura 7.12.	Estructura del sistema de posicionado.	197
Figura 7.13.	Estructura del Filtro de Kalman.	197
Figura 7.14.	Organigrama para el cálculo del filtro de Kalman	199
Figura 7.15.	Corrección de la posición del robot utilizando un Filtro de Kalman. (a) Trayectoria recta. (b) Trayectoria sinuosa.	200
Figura 7.16.	Trazado de trayectorias para cada posición del robot durante un proceso de navegación.	201
Figura 7.17.	Representación del cambio de posición del robot durante un periodo de muestreo.	202
Figura 7.18.	Linealización de la trayectoria que permite obtener las consignas de guiado del robot móvil.	204
Figura 7.19.	Distribución de las cámaras en entorno de trabajo empleado en las pruebas práctica.	205
Figura 7.20.	Ejemplos de seguimiento de trayectorias rectas. (a) Horizontal. (b) Oblicua .	205
Figura 7.21.	Ejemplos de seguimiento de trayectorias complejas con tres puntos intermedios de paso.	206
Figura 7.22.	Ejemplo de seguimiento de una trayectoria cerrada.	207
Figura 7.23.	Ejemplo de una trayectoria controlada por cuatro cámaras. (a) Entorno de trabajo. (b) Trayectoria seguida por el robot.	208
Figura 7.24.	Ejemplo de un experimento práctico realizado sobre un entorno de dimensiones grandes utilizando cuatro cámaras. (a) Dimensiones (en m) del entorno. (b) Trayectoria realizada. (c) Dispositivos (cámaras y odometría) utilizados para obtener la posición del robot dentro de los diferentes tramos.	209
Figura A.1.	Modelo geométrico de la cámara <i>pin-hole</i> .	224
Figura A.2.	Sistemas de referencia utilizados en la calibración de la cámara.	226
Figura A.3.	Modelo de luminancia de un punto circular.	231

# Índice de tablas

Tabla 3.1.	Valores de los parámetros extrínsecos de las cámaras utilizadas para obtener la gráfica de la figura 3.3.a.	55
Tabla 3.2.	Valores de los parámetros extrínsecos de las cámaras utilizadas para obtener la gráfica de la figura 3.3.b.	56
Tabla 4.1.	Estadísticos del errores prácticos de posicionamiento utilizando la cámara más cercana a la baliza.	101
Tabla 4.2.	Estadísticos de los errores prácticos de posicionamiento utilizando la cámara más alejada de la baliza.	102
Tabla 4.3.	Errores prácticos de posicionamiento utilizando dos cámaras.	112
Tabla 4.4.	Errores de posicionamiento prácticos para un sistema de visión formado por cuatro cámaras, realizando la corrección de la distorsión.	117
Tabla 5.1.	Valores de las coordenadas del centroide de la baliza para las imágenes de la figura 5.6 y para diferentes métodos de cálculo.	151
Tabla 6.1.	Valor medio de la desviación típica del centroide de los puntos de un patrón de calibración frente a diferentes condiciones de iluminación.	163
Tabla 6.2.	Comparación de los valores de los parámetros intrínsecos obtenidos con el método de calibración in situ y de laboratorio.	169
Tabla 6.3.	Comparación de los valores de los parámetros extrínsecos obtenidos con el método propuesto en esta tesis y con la herramienta de calibración de Matlab <sup>®</sup> .	178
Tabla 7.1.	Parámetros extrínsecos de las cámaras de la figura 7.17	205
Tabla 7.2.	Coordenadas de los puntos de paso de las trayectorias de la figura 7.21.	206
Tabla 7.3.	Coordenadas de los puntos de paso de la trayectoria de la figura 7.22.	206
Tabla 7.4.	Parámetros extrínsecos de las cámaras C3 y C4 de la figura 7.23.	207
Tabla 7.5.	Parámetros extrínsecos de las cámaras C3 y C4 de la figura 7.24.a.	208
Tabla 7.6.	Coordenadas de los puntos de paso de la trayectoria de la figura 7.24.b.	208



# Capítulo 1

---

## Introducción

---

Entre los diversos retos planteados en el campo de la robótica móvil, uno de los que presenta mayor interés es el de dotar a los robots de movilidad y autonomía. El fin perseguido, es que los robots sean capaces de moverse y manipular objetos y todo ello en entornos total o parcialmente conocidos y desconocidos. De entre los diversos aspectos que confluyen en los robots autónomos hay dos de especial interés, uno su capacidad para moverse y el otro su capacidad para reconocer y aprender el entorno. Ello no quiere decir que los aspectos de control de bajo y alto nivel, estructura mecánica, sistemas sensoriales, generación de trayectorias, etc., no sean aspectos también importantes, aunque de alguna forma están implícitos en los anteriores.

### 1.1. Motivación.

El guiado autónomo de robots es un campo multidisciplinar en el que existen aún muchos problemas por resolver. Dentro de las múltiples tareas que se plantean en este campo, una de las que está demandando mayor atención es la que se refiere a la localización. Esto es, conocer con precisión suficiente la posición y orientación del robot dentro del entorno.



Una amplia variedad de métodos de localización han sido propuestos y se han desarrollado, con éxito, diferentes prototipos de laboratorio. Los diferentes métodos de localización se pueden clasificar en dos grupos: aquellos que requieren modificar el entorno y los que no. La primera alternativa presenta ventajas importantes ya que, en muchos casos, permite reducir notablemente la complejidad de los robots, facilita la navegación de múltiples robots al mismo tiempo y en el mismo entorno, permite conocer lo que está sucediendo en el entorno y tomar decisiones en función de ello (es lo que se conoce como “espacios inteligentes” [Lee 01] [Steinhaus 04]). Estos “espacios inteligentes” se caracterizan por la utilización de un array de cámaras ubicadas en posiciones fijas del entorno y distribuidas de forma estratégica para abarcar la totalidad del campo de movimiento de los robots. La información proporcionada por todas las cámaras debe permitir determinar tanto la posición de los robots como de cualquier otro objeto que se encuentre en el entorno. Por su parte, los métodos que no necesitan modificar el entorno requieren que los robots lleven a bordo sistemas muy avanzados, complejos y voluminosos para el procesamiento de la información que obtienen del entorno. Además, la desventaja más importante que presentan estos métodos (paramétricos y basados en rejilla) es que necesitan modelos explícitos del entorno.

Los “espacios inteligentes” presentan importantes ventajas en aplicaciones para espacios interiores, ya que por una parte la información del entorno de movimiento del robot es más completa, los sistemas a bordo de los robots se simplifican notablemente, la infraestructura a incorporar al entorno es única (e independiente del número de robots que se muevan por el entorno), los costes se reducen notablemente sobre todo cuando en el entorno se mueven varios robots, se puede tener de forma sencilla información sobre la ubicación de todos los robots, facilitando la cooperación entre ellos, etc.

Hasta la fecha son escasos los trabajos publicados [Hoover 00][Lilienthal 03][Yun 04], relacionados con el guiado de robots móviles a partir de la información obtenida por un array de cámaras ubicadas en puntos fijos del entorno. En ellos, los experimentos prácticos se han realizado en entornos muy estructurados en los que no se producen cambios de iluminación, y con un fuerte contraste entre el robot y la superficie por la que se mueve.

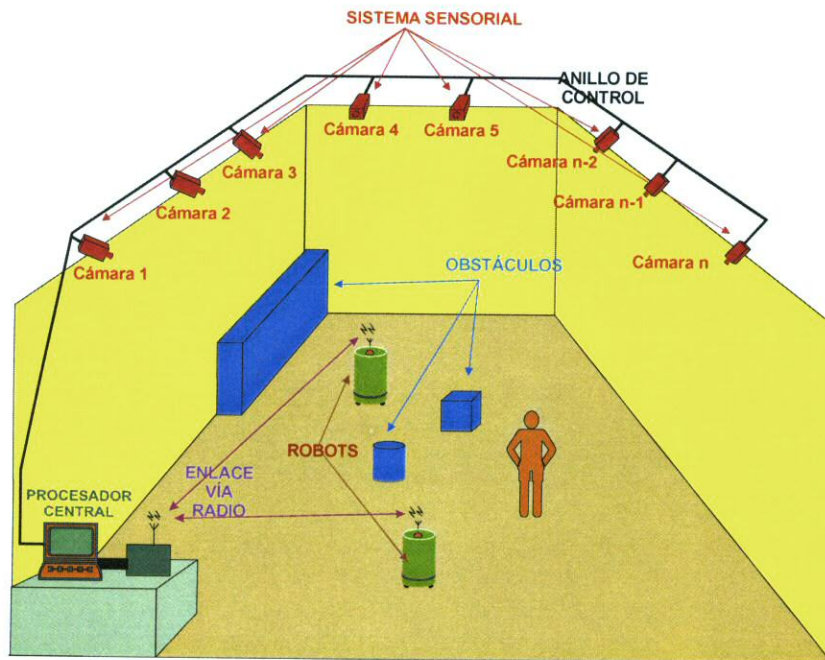
## **1.2. Objetivos planteados en la tesis.**

En esta tesis se plantea como objetivo principal el desarrollo de un sistema de posicionamiento y guiado de robots móviles en espacios interiores, utilizando cámaras estáticas ubicadas en el

entorno de movimiento y con un mínimo de infraestructura a bordo de los mismos. El entorno, a diferencia de los trabajos antes comentados, puede ser parcialmente estructurado, y se admiten cambios en la iluminación del mismo.

Las cámaras se ubican en posiciones conocidas del entorno, con una distribución y características ópticas tales que garanticen la completa visión de las trayectorias por las que deben navegar los robots. En general para realizar un correcto posicionamiento de los robots es necesario que todas las zonas por las que éstos deben navegar sean captadas por al menos dos cámaras, si bien cuando se trate de superficies planas este número se puede reducir a una sola. En caso de haber pequeñas zonas que no sean captadas por ninguna cámara se utiliza la odometría para obtener la posición del robot.

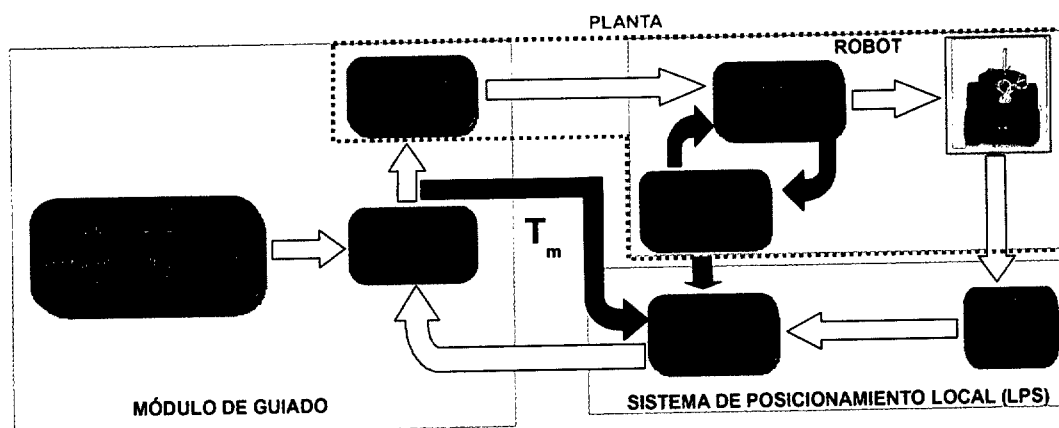
En la figura 1.1 se muestra, de forma simplificada, el sistema que se propone en esta tesis. En ella se puede observar que a partir de la información captada por las cámaras se obtienen las posiciones de los robots. A partir de éstas se generan (en un procesador central) las correspondientes consignas de velocidad de las ruedas motrices, que son enviadas a los robots vía radio.



**Figura 1.1** Estructura general del sistema de posicionamiento y guiado propuesto en esta tesis. El sistema sensorial está formado por  $n$  cámaras ubicadas en posiciones fijas del entorno, todas ellas conectadas a un procesador central que determina la posición del robot. Con la información proporcionada por las cámaras se generan las consignas de velocidad de las ruedas motrices, las cuales son enviadas a los robots vía radio.

Las ventajas más destacables que presenta este nuevo enfoque de posicionamiento y guiado, se pueden resumir en que se obtiene una posición absoluta de los robots, el hardware y el software pueden estar centralizados en un punto fijo (con las ventajas que ello supone de poder incrementar la capacidad de cálculo), se reduce el peso y volumen de los sistemas electrónicos a bordo del robot y el consumo energético (lo que facilita, entre otras, la portabilidad de los robots) y un mismo sistema hardware puede dar respuesta al guiado de varios robots.

En la figura 1.2 se muestra la estructura general del sistema de control propuesto en esta tesis. En él se han identificado dos grandes bloques, el “módulo de guiado” y el “sistema de posicionamiento local (LPS)”. El primero de ellos es el encargado de generar la consigna de velocidad de cada ruedas motriz, a partir de la información 3D de la posición real del robot, proporcionada por el “sistema de posicionamiento local”, y de la trayectoria que se desea seguir. El “sistema de posicionamiento local” es el encargado de proporcionar las coordenadas 3D de los robots, a partir de la información captada por las cámaras. El lazo de control principal, como se puede observar en la figura 1.2, lo constituyen el “sistema de posicionamiento local” y el “módulo de guiado”. El periodo de muestreo,  $T_m$ , de este lazo de control viene impuesto, fundamentalmente por el tiempo requerido para realizar el procesamiento de las imágenes.



**Figura 1.2** Estructura del sistema de control utilizado para realizar el guiado de robots por un entorno a partir de la información proporcionada por un array de cámaras estática. El sistema de posicionamiento proporciona la posición 3D a partir de la información captada por el sistema sensorial (cámaras). Por su parte el módulo de guiado es el encargado de generar todas las tareas implicadas en el movimiento del robot.

Dentro del objetivo general de la tesis, de realizar el guiado de los robots en espacios interiores, a partir de la información obtenida por las diversas cámaras ubicadas en el entorno, se plantea como objetivo más destacado la obtención de la posición 3D de los robots. Además,

esta posición 3D debe realizarse con la precisión suficiente como para permitir realizar un guiado fiable en las condiciones más exigentes que se presentan en espacios interiores (guiado por pasillos, habitaciones, etc.). Por otra parte, el sistema debe ser robusto a las distintas condiciones de iluminación que son frecuentes en espacios interiores. Finalmente, el sistema propuesto debe permitir realizar el guiado a velocidades máximas del orden de 1m/s.

Para alcanzar estos objetivos generales se plantean los siguientes objetivos concretos:

- Definición de una arquitectura global y requerimientos de los subsistemas de posicionamiento y guiado mediante cámaras.
- Definición de las características de las marcas a bordo del robot que faciliten su identificación en condiciones cambiantes de iluminación.
- Análisis de las características generales de las cámaras a utilizar (intentando que sean cámaras de bajo coste) examinando su influencia en los errores de posicionamiento.
- Diseñar un sistema de calibración que permita obtener de forma sencilla los parámetros intrínsecos y extrínsecos de múltiples cámaras ubicadas en espacios de grandes dimensiones.
- Desarrollar algoritmos y estrategias de segmentación de imágenes que permitan identificar a los robots de una forma fiable y rápida, para diferentes condiciones de iluminación, tanto artificial como natural (y con presencia de posibles brillos).
- Desarrollar algoritmos robustos que permitan obtener la posición del robot a partir de las imágenes de una, dos o más cámaras.
- Proponer soluciones que minimicen el error de medida de la posición 3D de los robots móviles, corrigiendo las diferentes fuentes de error.
- Realizar algoritmos para la generación y actualización de la posición del robot y control de su seguimiento, para guiarlo desde la ubicación actual hasta la de destino, evitando colisiones con los posibles obstáculos del entorno.

- Realizar pruebas prácticas que demuestren la viabilidad de la propuesta realizada, destacando sus virtudes y deficiencias.

### 1.3. Contexto de desarrollo de la tesis.

Esta tesis se enmarca dentro de las actividades de investigación que se vienen realizando durante los últimos años en el Departamento de Electrónica de la Universidad de Alcalá, en el campo de los robots móviles. En concreto, esta tesis se ha desarrollado dentro de los proyectos “Procesamiento de Imágenes orientado al Reconocimiento de Objetos 3D y al Guiado Autónomo de móviles Utilizando Cámaras Externas” (PIROGAUCE), financiado por la Consejería de Educación de la Comunidad de Madrid (Ref.: 07T/0057/2000), y “Entornos Inteligentes para el Guiado de Móviles de modo Autónomo”(ENIGMA), financiado por la Consejería de Educación de la Comunidad de Madrid (Ref.: 07T/0051/2003). En el primero de los proyectos se planteaba, como objetivo general, la obtención de la posición 3D de un robot móvil mediante cámaras fijas externas, ubicadas en el entorno de movimiento del robot. En cuanto al proyecto ENIGMA, se puede considerar una continuación del proyecto PIROGAUCE, incorporando, además de visión, sensores de ultrasonidos e infrarrojos.

### 1.4. Estructura de la tesis.

La organización de la memoria de esta tesis es el resultado de la exposición sistemática de los estudios teóricos, y de la experimentación efectuada sobre cada uno de los temas que se han propuesto para su desarrollo. La memoria está estructurada en ocho capítulos y un apéndice.

En el **capítulo primero** se realiza una breve introducción a los “espacios inteligentes”, se plantean los objetivos perseguidos con la realización de esta tesis, contexto de desarrollo y estructura de la memoria.

En el **capítulo segundo** se expone una síntesis del estado del arte en sistemas de posicionamiento y guiado de robots móviles en interiores, prestando especial atención a aquellos que utilizan cámaras estáticas (“espacios inteligentes”).

El **tercer capítulo** se dedica al estudio de las fuentes de error del sistema de visión y su influencia en la precisión del cálculo de la posición 3D del robot.

En el **capítulo cuarto** se describen los algoritmos utilizados para obtener la posición 3D del robot, a partir de su posición en el plano imagen de cada una de las cámaras del sistema de visión.

En el **capítulo quinto** se describe la estructura de las marcas a ubicar a bordo del robot, las cuales son utilizadas para facilitar su identificación por el sistema de visión, justificando su estructura y componentes. Así mismo se desarrollan los algoritmos de visión utilizados para su detección.

En el **capítulo sexto** se describe el método de calibración del sistema de visión propuesto para alcanzar los objetivos de precisión fijados en el capítulo tercero.

El **capítulo séptimo** se dedica a la descripción del sistema de guiado propuesto, mostrándose los resultados obtenidos, los cuáles validan las soluciones propuestas.

En el **capítulo octavo** se presentan las conclusiones finales junto a un resumen de las aportaciones realizadas en esta tesis. Así mismo, se indican futuras líneas de investigación relacionadas con el trabajo desarrollado, y publicaciones asociadas.

En el **apéndice A** se presentan los algoritmos de calibración para una cámara de distancia focal fija. Se escoge el método multi-imagen por ser la base de todos los cálculos posteriores, sirviendo además para estudiar la problemática asociada al patrón de calibración utilizado.

Finalmente se incluyen la relación de las referencias bibliográficas citadas en esta tesis.



## Capítulo 2

---

### Revisión de trabajos previos

---

La característica principal de los robots móviles (RM's) es la capacidad de desplazarse libremente en un espacio, sin la dependencia de un sistema rígido de guiado (raíles cables, etc.). Para conseguir este objetivo es necesario resolver aspectos diversos, como son los relacionados con la localización, generación de trayectorias, guiado, etc., siendo la localización uno de los más importantes.

En este capítulo se analizan las técnicas más importantes de posicionamiento de RM's que han sido propuestas por diversos grupos de investigación. De todas ellas, se ha prestado especial atención a aquellas que guardan relación con esta tesis. Esto es, sistemas de posicionamiento en entornos interiores, y en particular aquellos que utilizan la visión como sensor para determinar la posición de los robots. La ventaja que presenta la visión es la gran cantidad de información que proporciona, aunque su procesamiento para la extracción de características útiles para el control del RM requiere de complejos algoritmos con un coste computacional generalmente alto. Relacionado con el proceso de visión está la elección del modelo de cámara a utilizar, razón por la cual en este capítulo se analizarán diferentes modelos con el fin de elegir el que más se adecue a la aplicación propuesta en esta tesis. También se presentan algunas propuestas para obtener los parámetros intrínsecos, que describen el modelo de la cámara, y los extrínsecos de un sistema multi-cámara. Finalmente se expondrán las conclusiones más importantes sobre todos estos aspectos.



## 2.1. Robots móviles.

Los robots móviles autónomos se caracterizan por una unión “inteligente” entre las operaciones de percepción y actuación controlada, que definen su comportamiento y les permite llevar a cabo los objetivos programados. El grado de autonomía dependerá, en gran medida, de la facultad del robot para abstraer el entorno y convertir la información obtenida en órdenes, de tal modo que aplicadas sobre los actuadores del sistema de locomoción garanticen una acción eficaz. Las dos grandes características que diferencian a los robots móviles de cualquier otro vehículo son:

- *La percepción.* Determina la relación del robot con su entorno de trabajo mediante el uso de sensores.
- *La actuación controlada.* Determina las acciones que se han de realizar en cada momento, según el estado del robot y del entorno, para alcanzar las metas perseguidas.

### 2.1.1. Navegación de robots móviles.

La característica principal de los robots móviles es, como ya se ha mencionado, su capacidad de movimiento sin estar ligado físicamente a ningún tipo de infraestructura fija (raíles, cables, etc.). Para ello los robots móviles requieren de un sistema de posicionamiento que les permita conocer, en todo momento, su posición dentro del entorno. Esta información permite planificar la trayectoria o camino a seguir para poder alcanzar su objetivo, evitando aquellos obstáculos que se puedan presentar en el camino. El robot debe poseer un sistema sensorial que le permita percibir y representar el mundo que le rodea. En la navegación se conjugan tres tareas:

1.- *Localización.* Agrupa al conjunto de técnicas que permiten determinar, en todo momento, la posición del robot en el entorno a partir de la información suministrada por el sistema sensorial. La ubicación de un robot en el espacio de trabajo viene determinada por el vector de coordenadas cartesianas  $(X, Y, Z)$  y la orientación  $(\theta_x, \theta_y, \theta_z)$ . En el caso de movimiento sobre superficies planas la ubicación viene dada por el vector  $(X, Y)$  y el ángulo  $\theta$  en el plano.

2.- *Planificación y seguimiento de trayectorias.* Implica el conjunto de técnicas utilizadas para definir un camino entre un punto de origen y otro destino, dentro del espacio de movimiento. Esta tarea debe asegurar que los caminos propuestos sean: seguros y estén libres de obstáculos, lo más cortos posibles y dinámicos; pudiendo ser modificados ante una serie de eventos como

puede ser un error en la posición del robot o la presencia de un obstáculo no previsto. Así como, el seguimiento de la trayectoria minimizando los errores lateral y longitudinal.

3.- *Técnicas de detección y evitación de obstáculos no previstos.* Es el conjunto de técnicas que permiten la detección de obstáculos en el recorrido planificado del robot, así como las acciones a realizar para evitarlos.

De las tres tareas implicadas en el proceso de navegación de los RM's, la primera (localización) plantea una dificultad importante, siendo uno de los problemas fundamentales de la robótica móvil. En este sentido, se han propuesto un gran número de modelos, enfoques y técnicas para resolver este problema [Taeseok 03][Nguyen 03][Esteves 03][Valin 03][Gijeong 03]. La mayoría de ellas sólo son útiles en situaciones muy específicas.

## 2.2. Técnicas de posicionamiento de un robot móvil.

Para navegar en un espacio, el robot necesita conocer, con mayor o menor precisión, la posición en la que se encuentra en cada momento o, cuanto menos, su posición relativa con respecto al objetivo. El problema de determinar la posición de un robot móvil es uno de los aspectos más importantes en el campo de la robótica móvil y son muchas las soluciones parciales que existen para resolverlo [Borenstein 96a]. Se pueden clasificar básicamente en dos grupos:

- *Métodos de posicionamiento relativo:* determinan la posición y orientación actual del robot respecto de una posición inicial dada. Para ello se utilizan las informaciones proporcionadas por sensores internos (encoders, giróscopos, acelerómetros, etc.).
- *Métodos de posicionamiento absoluto:* determinan la ubicación del robot respecto a un sistema de referencia externo, utilizando sensores que permiten realizar medidas absolutas respecto a marcas, naturales o artificiales, dispuestas en puntos conocidos del entorno por el que se debe mover el robot. También se pueden usar sensores de posicionamiento global (GPS, DGPS).

### 2.2.1. Métodos de posicionamiento relativos.

Estos métodos determinan la posición actual y orientación respecto de una referencia inicial,

utilizando sensores internos del robot. En este grupo se incluye la odometría, la navegación inercial basada en giroscopios y acelerómetros [Borenstein96b][Duckett 98] [Antonelli 03][Bostan 03] que permiten determinar la velocidad angular y la aceleración del robot sobre el que van montados. Los errores de estos sistemas son acumulativos [García 93][Nakju 03], aumentando a medida que transcurre el tiempo de navegación. Estos errores se pueden dividir en dos tipos:

- *Errores sistemáticos.* Se producen por el cambio en algún parámetro de la geometría del robot (radio de las ruedas, distancia entre las ruedas motrices, deslizamientos, etc). Estos errores son impredecibles.
- *Errores de medida.* Son los originados por el propio sensor de medida y las operaciones que en él se realizan para obtenerla. Por ejemplo, la utilización de encoders lleva inherente un proceso de cuantificación con su error asociado. Para su compensación se utilizan modelos matemáticos estocásticos.

Los errores asociados a los sistemas de posicionamiento relativo y su crecimiento con el tiempo hacen que éstos solamente sean válidos para distancias cortas, siendo necesario cada cierto tiempo obtener medidas de la posición absoluta que permitan actualizar las medidas y con ello eliminar o reducir los errores acumulados por el sistema de posicionamiento relativo.

### **2.2.2. Métodos de posicionamiento absoluto.**

Estos métodos permiten determinar la posición del robot a partir de la información generada por sensores y marcas, artificiales o naturales, externas al robot. Dentro de estos métodos se enmarca el GPS (*Global Positioning System*) desarrollado por el departamento de defensa de EE.UU en los años 70, para su utilización en espacios exteriores. Su uso en espacios interiores no es posible debido a la oclusión de las señales proporcionadas por los satélites geostacionarios. El uso de GPS *pseudolites* (acrónimo de pseudo satélite)[Stone 98][Wang 01][LeMaster 02] permite aplicar las técnicas de navegación GPS en entornos interiores. Un *pseudolite* es un emisor de señales similares a las del GPS, el cual se ubica en el entorno de trabajo. En la recepción se utilizan receptores GPS estándares. Con estos sistemas se obtienen unos errores de posicionamiento de dos a tres centímetros [Kee 00], según indican sus autores.

También se han realizado trabajos basados en el servicio E911 de la telefonía móvil [Caffery 98] para obtener la posición de un robot móvil, aunque los errores cometidos (entorno a un metro) no los hacen recomendables para entornos interiores.

Dado que una gran parte de las aplicaciones de los RM's son en espacios interiores (oficinas, fábricas, hospitales, etc.), es en este campo donde se han desarrollado una gran cantidad de sistema de posicionamiento. En lo que sigue, y por tratarse de sistemas de posicionamiento locales, nos vamos a referir a ellos como LPS (*Local Positioning Systems*). Dentro de este grupo se incluyen los métodos basados en: balizas activas, marcas artificiales, marcas naturales y correspondencia con mapas.

a) *Balizas activas (Active Beacons)*. Éstas son elementos activos que emiten algún tipo de energía, normalmente electromagnética (luz, sonido o radiofrecuencia), estando dispuestas en el entorno por el que se mueve el robot, en posiciones conocidas. Conocida la posición de éstas y el ángulo de incidencia de la señal emitida por varias de ellas y/o la distancia a ellas, es posible determinar la posición del robot.

Las balizas activas se pueden clasificar atendiendo a la naturaleza de la señal emitida. Así se puede hablar de balizas de: ultrasonidos, infrarrojos, radiofrecuencia y campos magnéticos, entre las más importantes.

- *Ultrasonidos*. Estos sistemas utilizan transmisores de ultrasonidos ubicados estratégicamente en el entorno de trabajo, de forma que realicen una emisión sincrona [Hazas 03] o simultanea codificada [Villadangos 04], mientras que sobre el robot se coloca el receptor. Midiendo los tiempos de vuelo es posible determinar la posición del robot. Estos sistema proporcionan precisiones del orden de 3 cm. Otra alternativa consiste en la existencia de un único emisor (ubicado en el robot) y varios receptores (ubicados en posiciones fijas). Esta última alternativa presenta inconvenientes cuando se encuentran más de un robot móvil en la misma área de trabajo, pudiendo producirse interferencias entre sus sistemas sensoriales.
- *Infrarrojos*. Presentan una estructura similar a la anterior ubicando múltiples emisores IR [Eren 97][Arai 03] o láser [Hernández 03]. La posición del robot se determina a partir de los ángulos de incidencia de cada baliza. Con estos sistemas se obtienen precisiones entre 2 cm y 4 cm.
- *Radiofrecuencia*. Se basan en una emisión RF, de la que se utilizan los tiempos de llegada para medir la separación entre emisor y receptor y así calcular la posición del robot [Kudoh 00]. Estos sistemas proporcionan precisiones entre 1m y 3m, degradándose con la existencia de rebotes en los objetos y en personas que se mueven por el entorno. Otros autores [Youssef 02][Ladd 02] proponen la

utilización de redes WiFi como elementos de posicionamiento. Para ello se mide la energía recibida por un receptor (ubicado en el robot) correspondiente a las emisiones de diferentes puntos de acceso. Al utilizarse frecuencia de 2 GHz (microondas) se producen fuertes atenuaciones en la energía de la señal recibida debido a elementos con un alto contenido en agua (personas, animales, etc); además son muy sensibles a las reflexiones (multicaminos). Las precisiones que se obtienen son del orden de 40 cm [Ocaña 04].

- *Magnéticos*. Utilizan campos magnéticos [Prigge 02] codificados para determinar la posición. La precisión obtenida es de 3 cm a 4 cm, la cual se disminuye a medida que aumenta el número de sensores utilizados. Estos sistemas se utilizan en espacios reducidos y que por tanto requieren de un número bajo de sensores.

Un inconveniente que presentan todos estos sistemas que utilizan balizas activas es que la precisión en la determinación de la posición del robot es función de la exactitud en el conocimiento de la posición de las balizas [Skrzypczynski 98]. Otros inconvenientes son que requieren un mantenimiento regular para comprobar su funcionamiento, un coste alto y son sistemas altamente invasivos. A esto, hay que añadir los errores debidos a las reflexiones de las señales emitidas.

*b) Marcas artificiales (Artificial Landmarks)*. Las marcas artificiales son elementos fácilmente distinguibles que se ubican en posiciones fijas del entorno del robot, haciendo las veces de balizas pasivas. Estas marcas se diseñan para que sean fácilmente detectadas, aún en condiciones adversas. A partir de la información extraída de ellas y su ubicación en el entorno es posible determinar la posición y orientación del robot. La ventaja principal de las marcas artificiales es que es un método económico, por la sencillez de su fabricación, son flexibles ya que no necesitan una costosa infraestructura ni modificaciones grandes del entorno y requieren un mínimo o nulo mantenimiento. La forma de la marca viene impuesta, en gran medida, por la información que se quiera obtener de ella, pudiendo ser una información métrica relativa (permite sólo obtener la posición y orientación relativas) [Romero 00], o global. En este último caso, a las marcas se les añade un código identificador que permite diferenciar unas de otras, dando lugar a sistemas muy versátiles. La información codificada puede ser muy variada, utilizándose códigos de barras [García 00], puntos de colores [Adán 02], etc. Para detectar las marcas, la mayor parte de los trabajos utilizan cámaras CCD, si bien otros autores [Everet94] proponen la utilización de escáneres láser.

La precisión obtenida con los métodos que utilizan marcas artificiales depende de la precisión

con la que se obtienen los parámetros geométricos de las mismas, téngase presente que la forma captada de la marca depende de la posición relativa entre el robot y ella. Esta precisión, en general, empeora a medida que aumenta la separación entre marcas y robots. Por otra parte, la trayectoria del robot está limitada al recorrido en el que se han dispuesto las marcas.

*c) Marcas naturales (Natural Landmarks).* Dentro de las marcas naturales se incluyen los objetos estáticos, como pueden ser puertas, paredes, esquinas, etc., que se encuentran en el entorno del robot y que hacen las veces de balizas pasivas. En consecuencia, no es necesario realizar modificaciones del entorno, pero éste debe ser conocido a priori. Una vez identificados estos elementos y conocida su ubicación en un mapa del entorno es posible calcular la posición y orientación del robot. Debido a la dificultad de detección de las marcas y su emparejamiento con un elemento del mapa a priori, hace que la obtención de la ubicación sea más compleja que el que utiliza marcas artificiales. Los sensores utilizados para captar dichas marcas pueden ser cámaras de video [Murray 00] [Fiala 02] [Marrón 02], escáneres láser (*laser rangefinder*) [Santiso 03], ultrasonidos [Althaus 01], emisores de infrarrojos [Maaref 02] o una combinación de ellos. Algunos autores refuerzan el sistema añadiendo la detección de marcas artificiales [Mata 01][Rupp 00].

*d) Correspondencia de modelos o mapas (Map Matching).* Se utilizan en conjunción con balizas activas, marcas naturales o artificiales. La información captada por los sensores de a bordo (*on board*) es comparada con un modelo del entorno (mapa local), que contiene datos sobre los elementos que hay en él. Los resultados obtenidos dependen fuertemente de la exactitud del mapa utilizado y, sobre todo, del tipo de escenario.

Dado un sistema de posicionamiento, no siempre es fácil determinar en qué grupo de los anteriores se puede clasificar. La diferencia se encuentra en la finalidad de la información procesada por los sensores y en los mecanismos utilizados para deducir la posición del robot. La clasificación anterior también se podría haber realizado en función de los sensores externos utilizados, así se tendría: ultrasonidos, infrarrojos, láser, visión, etc.

En esta tesis se propone un sistema de posicionamiento local (*indoor*) que utiliza visión artificial en conjunción con una baliza de infrarrojos. A diferencia de la mayor parte de las propuestas, la baliza se encuentra sobre el propio robot mientras que las cámaras se ubican en posiciones fijas del entorno. A continuación se presenta una visión general de aquellos trabajos que utilizan la visión para obtener la posición del robot, para, seguidamente, realizar un estudio pormenorizado de los que plantean el uso de cámaras ubicadas en posiciones fijas, en línea con la propuesta realizada en esta tesis.

## 2.3. Sistemas de posicionamiento de RM's utilizando visión artificial.

Al igual que sucede en el ser humano, la capacidad de dotar de visión a un robot móvil le permite responder al entorno de una manera flexible e inteligente, en comparación con otros mecanismos. La visión es considerada como uno de los más importantes y sofisticados sistemas de percepción. Pero como era de esperar, los sensores, las técnicas y los procesos asociados a la visión son considerablemente más complejos que los de otras alternativas de percepción.

En los sistemas prácticos, la mayor parte de los trabajos de investigación demandan al sistema sensorial una gran cantidad de información para proporcionar unas mejores capacidades de interacción robot-entorno. La percepción visual puede proporcionar a un robot toda esta información. Sin embargo, debido a la gran cantidad de información que suministra, la extracción de características para determinar la posición del robot móvil es compleja, quedando aún muchas tareas por realizar. A pesar de lo anteriormente comentado, el notable descenso en el coste de las cámaras, la gran oferta existente en el mercado, las mejoras en sus prestaciones y el precio de los sistemas de procesamiento (hardware) hacen de este tipo de sistemas los de mayores posibilidades potenciales en la robótica móvil. Los sistemas que utilizan visión para realizar el guiado de robots se pueden clasificar, desde el punto de vista de la ubicación de las cámaras, en dos grandes bloques: los que utilizan cámaras montadas sobre el propio robot (*on board*) y los que ubican las cámaras en posiciones fijas de su entorno (*off board*).

### 2.3.1. Sistemas que utilizan cámaras de video embarcadas (*on board*).

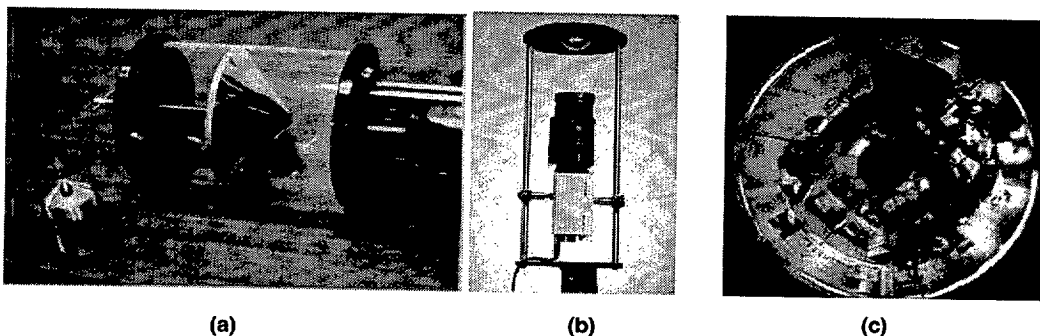
En el problema del guiado de robots móviles utilizando visión, la mayoría de los trabajos de investigación se han desarrollado en la línea de embarcar la visión a bordo (*on board*). Esta alternativa presenta la desventaja de la pobre perspectiva que el robot tiene del entorno, estando limitada a su campo de visión. A ello hay que añadir las dificultades de obtener información del entorno con el robot en movimiento y los errores de medida que se cometen como consecuencia de ese movimiento. Además, hay que añadir que los costes se incrementan a medida que aumenta el número de robots que se mueven por un entorno.

Dentro de los trabajos en robótica móvil que utilizan la visión como elemento de percepción del entorno, se puede hacer una clasificación entre aquellos que utilizan una o varias cámaras.

## Sistemas que utilizan una sola cámara.

Pocos son los trabajos que utilizan un sistema monocular para determinar la ubicación del robot móvil. En [Paz 97] se describe un sistema que integra la información odométrica y la correlación de la información visual de marcas naturales (vértices de esquinas y columnas) con un mapa CAD (plano global). La odometría se utiliza como apoyo durante el tiempo de procesamiento de las imágenes. En este trabajo los autores realizan el guiado de un robot a una velocidad de 1m/s y con una precisión de  $\pm 2$  cm y  $\pm 0.2^\circ$ , pero el sistema requiere de un hardware específico compuesto por una red de cuatro *transputers*. Dentro de esta misma línea de trabajo, pero utilizando marcas artificiales, está el trabajo desarrollado en [García 03]. En él los autores son capaces de obtener un error de posición de 4 cm para una velocidad de 1m/s.

Un problema común a los sistemas anteriores, y en general a los que utilizan una sola cámara a bordo del robot, es el limitado campo visual que tiene el robot del entorno que le rodea, por las limitaciones propias del campo de visión de la cámara. Para evitar este inconveniente, recientes trabajos [Mouaddib 00] plantean el uso de sensores omnidireccionales que proporcionan un campo de visión de  $360^\circ$ . Estos sensores (figura 2.1.a y b) están formados por una cámara CCD, la correspondiente óptica y un espejo esférico o cónico, que permite captar información más amplia del entorno (figura 2.1.c). Utilizando este tipo de sistemas y marcas naturales (líneas verticales y horizontales) junto con un mapa del entorno, los autores consiguen una precisión de 5 cm en el posicionamiento del robot.

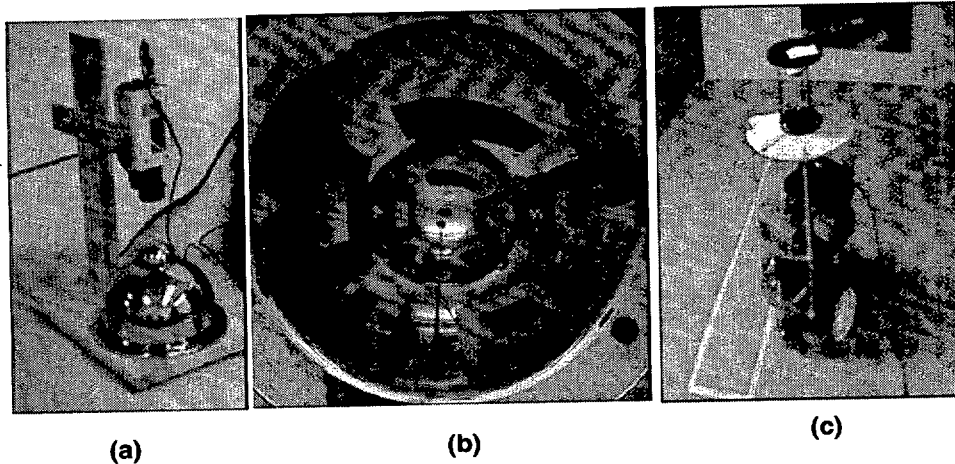


**Figura 2.1.** Sistema de visión catadióptrico empleado en [Mouaddib 00]. (a) Estructura de la lente utilizada, en la que se puede observar un espejo reflector cónico que permite captar un campo de visión de  $360^\circ$ . (b) Montaje del sensor CCD en conjunción con la lente. (c) Ejemplo de una imagen de un entorno captada con el sistema de visión omnidireccional.

Otro trabajo con sensores omnidireccionales es el desarrollado por Fiala [Fiala 02]. En éste se propone el uso de un sensor estéreo (figura 2.2. a) con el fin de obtener dos imágenes panorámicas del entorno (figura 2.2.b). Sobre una imagen panorámica se detectan líneas



verticales y horizontales (marcas naturales) utilizando la transformada panorámica de Hough (PHT). Estas líneas son comparadas con un mapa del entorno para localizar la posición del robot, en este caso el error máximo de posicionamiento conseguido es de 3 cm. El sistema propuesto por Fiala requiere de un tiempo de procesamiento alto, lo que hace que la velocidad máxima de proceso sea de una imagen por segundo.



**Figura 2.2.** Sistema catadióptrico estéreo utilizado en [Fiala 02]. (a) Estructura del sensor formado por una cámara CCD y dos espejos reflectantes semiesféricos. (b) Aspecto de una imagen captada por el sensor estéreo. (c) Plataforma móvil utilizada en las pruebas.

### Métodos que utilizan visión estéreo.

Teniendo en cuenta que la información proporcionada por una sola cámara no permite obtener información 3D del espacio y que la información 3D puede resultar de gran interés para posicionar un RM en el entorno de movimiento, la utilización de sistemas de visión multicámara adquiere una especial importancia. La solución más frecuente a este problema es utilizar dos cámaras, si bien se podría obtener información tridimensional utilizando una sola cámara y diferentes perspectivas o técnicas basadas en la variación de la distancia focal.

Son numerosos los trabajos [Weckesser 95][Burschka 97][Ponweiser 01] [Okada 03][Saeedi 03][Jae 03][Fujiyoshi 03] que utilizan dos cámaras, formando un par estereoscópico para determinar la posición de un robot. Generalmente en estos casos se utilizan marcas naturales (columnas, vértices de paredes o puertas, etc.) en conjunción con un mapa a priori. Las precisiones obtenidas con estas alternativas están en el orden de los milímetros (entorno a 10 mm) y las velocidades de procesamiento están en el orden de los 200 ms. Para obtener la correspondencia entre puntos en las dos imágenes se utiliza la correlación entre ellas, proceso que requiere tiempos de cómputo altos.

Una variante del sistema estéreo, la forman los sistema trinoculares, en los que al sistema binocular se le añade una tercera cámara para simplificar el problema de la correspondencia entre imágenes. En este caso, se simplifican los algoritmos de correspondencias y con ello se reducen notablemente los tiempos de computación. Las precisiones que se consiguen están próximas a los 8 mm, obteniéndose unos tiempos de procesado de 150 ms. Trabajos en esta línea se pueden encontrar en [Martínez 03][Saeedi 00][Murray 00].

Los sistemas estéreo buscan en las imágenes elementos (marcas) que tienen una determinada característica (líneas rectas, esquinas, etc.). Esto presenta una limitación importante a la hora de su implementación en entornos cerrados formados por paredes y suelos, en los que puede haber pocas de esas características. Por ejemplo, utilizando una técnica estéreo se pueden detectar los bordes de una pared, pero no la propia pared. Por otra parte, hay situaciones en las que no se puede determinar la posición correcta del robot, sobre todo cuando tiene que distinguir entre dos posiciones que poseen entornos similares, como por ejemplo un pasillo en el que existen gran cantidad de puertas iguales. La solución que proponen la mayoría de autores se basa en la interpolación de las marcas detectadas con las anteriores posiciones del robot, con lo que la precisión depende de la cantidad de datos almacenados en la memoria del robot móvil. Otra solución fue propuesta por Fraundorfer [Fraundorfer02]; en ella se realiza de forma simultánea la localización del robot y la construcción de un mapa (SLAM *Simultaneous Localization and Map building*), mapa construido con una definición mayor para ser utilizado en futuros guiados. El autor indica que este método proporciona una buena identificación de la posición del RM, si bien no indica cuál es el coste computacional necesario para realizar dicho posicionamiento.

Los métodos anteriores que utilizan visión estéreo, a partir de dos o tres cámaras, tienen un coste computacional alto (sobre todo por los problemas de correspondencia), lo que limita la velocidad del robot móvil. Otras alternativas se proponen en [Kidono 00] y [Ohya 01] en las que en una primera fase se realiza el aprendizaje del camino a seguir mediante un control remoto, generando un mapa, y posteriormente el robot navega de forma autónoma a partir del conocimiento adquirido y la información captada. Durante la fase de guiado manual, el robot observa el entorno con un sistema de visión estéreo, generando un mapa de ocupación. La integración de múltiples observaciones permite determinar la zona libre de obstáculos por la cual se puede mover. A continuación el robot pasa a trabajar en modo autónomo de forma que calcula el camino más corto entre los puntos origen y destino, evitando los obstáculos. Con ello los autores indican que consiguen una reducción del 20% en el tiempo de procesado. Esta alternativa supone que el entorno por el que se mueve el robot es estático, no habiendo

cambiado tras la fase de aprendizaje. Estos métodos consiguen una desviación de 4 cm en la trayectoria seguida por el robot respecto a la de la fase de aprendizaje.

Los sistemas de guiado de robots móviles que utilizan cámaras a bordo, presentan una importante limitación: por una parte la perspectiva que el robot tiene del entorno es pobre, estando limitada a su campo de visión (una pequeña parte del entorno). A ello hay que añadir: las dificultades de obtener información del entorno con los robots en movimiento, los errores de medida que se cometen como consecuencia del movimiento y el volumen que ocupa el hardware necesario para procesar las imágenes. Además cada robot debe incorporar gran cantidad de electrónica.

Frente a esta alternativa están surgiendo, en los últimos años, propuestas donde el sistema de visión se ubica en posiciones fijas (*off board*) del entorno de desplazamiento del robot. Este nuevo enfoque presenta importantes ventajas en los sistemas LPS. Entre estas ventajas se puede destacar que: la información del entorno de movimiento del robot es más completa y por tanto la navegación puede ser más segura, el sistema de percepción es fijo e independiente del número de robots lo que redundará en una reducción importante del hardware a bordo, y ello permite minimizar el tamaño, costes y utilizar múltiples robots de idénticas características.

Al disponer de mayor información se facilita el guiado de los robots y al mismo tiempo permite que varios, con independencia de sus características, puedan moverse de forma autónoma dentro de un mismo entorno sin necesidad de sistemas complejos de comunicación con los robots (solamente se envían al robot, desde un puesto central, los comandos de guiado). Es en esta línea donde se enmarca esta tesis y por ello se van a presentar los trabajos más importantes que se han desarrollado en este campo hasta la fecha. Conviene también indicar que éste es un campo novedoso de investigación, lo que hace que no sean muchos los trabajos previos publicados directamente relacionados con este tema. Sin embargo, cada día son más los grupos de investigación interesados en esta idea, que es uno de los aspectos incluidos dentro de lo que se conoce como “espacios inteligentes”.

### **2.3.2. Sistemas que utilizan cámaras ubicadas en posiciones fijas (Espacios Inteligentes).**

En los últimos años está surgiendo dentro de la comunidad científica que trabaja en el ámbito de las interfaces de usuario (HMI, Human Computer Interfaces) una nueva área de trabajo que está recibiendo una gran atención y que son los llamados “Entornos Activos” (*Active Environments*), o “Entornos/Espacios Inteligentes” (*Intelligent Environments/Spaces, Smart Environments, Smart Rooms*). En este nuevo concepto se busca que los entornos interactúen

de forma natural con el individuo y le ayuden, de manera no intrusiva, en la realización de las tareas cotidianas. Los ordenadores dentro de un “Entorno Activo” son transparentes a los usuarios, de tal forma que éstos obtienen servicios del “entorno” mediante una interacción en lenguaje natural y sensible al contexto. De este modo se consigue que las habitaciones u oficinas tengan una entidad propia y tomen la iniciativa en la interacción con el objetivo de facilitar a los usuarios sus actividades y en definitiva la calidad de vida. Estos entornos deben permitir a los ordenadores participar en actividades que nunca habían implicado computación y a los usuarios interactuar con sistemas computacionales de la misma forma que lo harían con otros usuarios. Las personas en los entornos inteligentes podrán, entre otras, visualizar información multimedia y experimentar ambientes virtuales compartidos sin necesidad de usar el teclado o sensores especiales.

Dentro de un entorno inteligente se pueden encontrar tecnologías diversas; desde dispositivos hardware como sensores, interruptores, electrodomésticos, webcams, etc., hasta aplicaciones software tales como reconocedores de voz, servidores multimedia, agentes de correo electrónico, etc. Estas entidades se tienen que integrar y controlar utilizando la misma interfaz de usuario. Por ejemplo, un usuario tiene que poder realizar una conexión a internet tan fácilmente como apagar las luces.

Los entornos inteligentes se basan en el concepto de computación ubicua, definido originalmente por Weiser [Weiser 91]. Según Weiser, la ubicuidad y transparencia del sistema son las características más sobresalientes de la computación ubicua. Los espacios inteligentes se pueden caracterizar por sus elementos básicos: ubicuidad, conocimiento, inteligencia e interacción natural. La ubicuidad se refiere a situaciones en las cuales estamos rodeados por múltiples sistemas embebidos, interconectados entre ellos. El conocimiento se refiere a la capacidad del sistema para localizar y reconocer objetos y personas y sus intenciones. La inteligencia hace referencia al hecho de que el entorno digital pueda analizar el contexto, autoadaptarse a las personas y objetos que viven en él, aprender su conducta y finalmente mostrar emociones. La interacción natural se refiere a aspectos tales como lenguaje natural, reconocimiento de gestos, reconocimiento de voz, los cuales facilitan la comunicación entre los usuarios y el entorno digital. Algunos autores definen también los espacios inteligentes como “un espacio donde se puede interactuar en forma simple con ordenadores y robots, y nos permite obtener servicios útiles de ellos”. Los sistemas de computación ubicua proporcionan acceso a servicios computacionales pero ocultan los dispositivos a los usuarios. Los usuarios no tienen que encontrar las interfaces, sino que el sistema tiene la responsabilidad de servirles.

Uno de los precursores de estos sistemas ubicuos es el proyecto Xerox PARCs's *Ubiquitous Computing* (UbiComp) [Weiser 91] desarrollado a finales de los 80 y principios de los 90, el cual, obtenía información de las personas que interactúan con él. En este proyecto no se utiliza la visión artificial para obtener información de las personas que se encuentran en el entorno, sino que se utilizan otros sensores más sencillos, debido a las limitaciones técnicas de los sistemas electrónicos existentes en esas fechas. Por ejemplo, se podía seguir un persona por un edificio mediante una insignia de infrarrojos. De forma similar, se podía detectar que una persona estaba sentada en una silla por medio de un sensor de presión colocado en un cojín. Un sistema más avanzado, que ya utilizaba la visión como elemento de detección es el Active Badge [Want 92], cuya primera aplicación fue la localización de personas en un campus.

En la actualidad varios centros de investigación están realizando importantes aportaciones en el campo de la computación ubicua, algunos de los más notorios se indican a continuación:

- *Intelligent Room* [Coen 98]. Desarrollado por el grupo de investigación del Artificial Intelligence Laboratory del MIT, es uno de los más desarrollados en la actualidad. Se trata de una habitación que posee cámaras, micrófonos y otros sensores (figura 2.3) que intentan interpretar lo que los usuarios están haciendo, con el fin de poder ayudarles del mismo modo que si fueran mayordomos invisibles. Los usuarios pueden interactuar con el entorno mediante la voz, gestos y contexto.

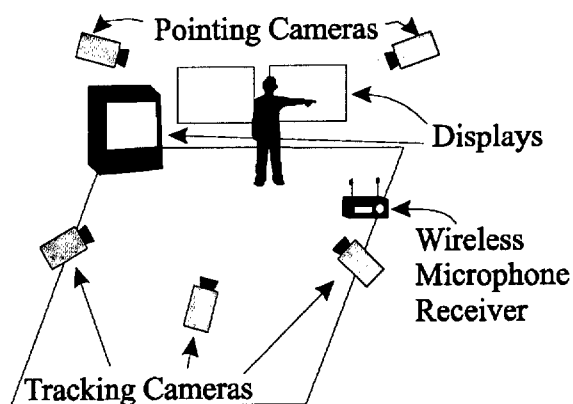


Figura 2.3. Estructura del entorno *Intelligent Room* del Artificial Intelligence Laboratory del MIT.

- *Smart Room* [Pentland 97]. Desarrollado por el Media Laboratory del MIT, utilizando cámaras, micrófonos y otros sensores, pretenden desarrollar una especie de mayordomo como un observador pasivo que constantemente está viendo la oportunidad de ayudar, de forma similar al proyecto de "*Intelligent Room*". Una vez que una persona es localizada el paso

siguiente es identificarla, utilizando su cara o su voz. Sus autores recuperan la geometría del cuerpo para realizar diferentes tareas de reconocimiento de gestos, incluyen también, un lector de tiempo real del lenguaje de señas y un sistema que reconoce los movimientos de T'ai chi y entrena a lo usuarios para que los ejecuten correctamente. Su trabajo actual se centra en un sistema que aprende palabras con interacciones naturales; el usuario le enseña al sistema palabras simplemente señalando el objeto y nombrándolo.

- *Smart Desks*, del mismo departamento que “*Intelligent Room*”, es un tipo de *Smart Room*, pero especializado en el entorno de negocios. El objetivo de este proyecto es desarrollar un escritorio que actúe como si fuese un ayudante de oficina. Así, el escritorio deberá conocer los hábitos, preferencias, y sensaciones del usuario, o simplemente recordar dónde poner las cosas.

- *The KidsRoom* [Bobick 97], también del mismo departamento que el anterior, es un espacio de juego totalmente automatizado e interactivo para niños (figura 2.4). Dos de las paredes de la habitación son tabiques reales, mientras que las otras dos son pantallas de proyección usadas para transformar la apariencia del ambiente de la habitación. Se emplea visión artificial para el reconocimiento de acciones, de modo que estas informaciones puedan ser utilizadas por el ordenador como datos de entrada para el sistema de control de la narración. El reconocimiento perceptual hace posible la interacción con las “criaturas virtuales” proyectadas en las pantallas, las cuales responden en función de las acciones de los niños. A través de la aventura, el sistema de cómputo detecta la posición de los niños y responde a eventos tales como: ¿están todos en la cama?, ¿están los niños en un grupo?, ¿están siguiendo los niños la ruta?, etc. La música y la narrativa de la historia cambia dependiendo de qué están haciendo los niños.

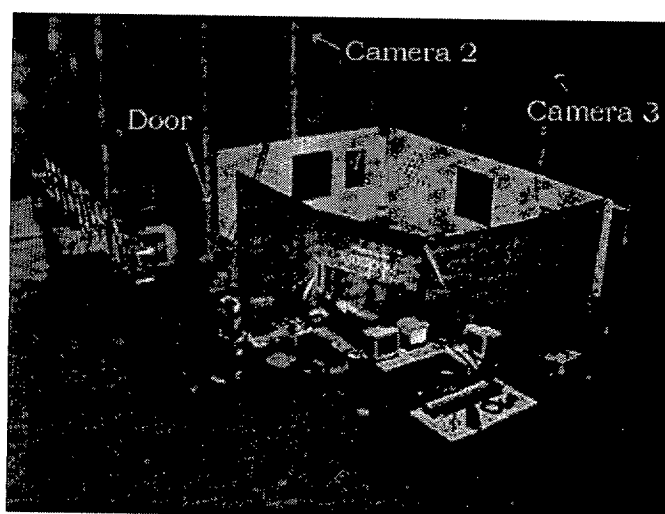
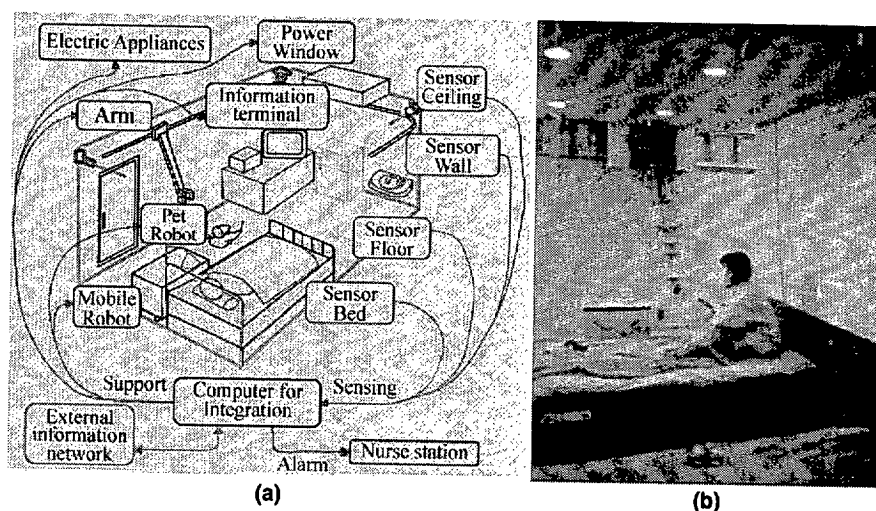


Figura 2.4. Entorno KidsRoom del Media Laboratory del MIT.

- *I.A.E. (Intelligent and Aware Environments)* también conocido como Smart Spaces [Essa 98], que consiste en espacios que han sido transformados en áreas de trabajo inteligentes donde cámaras, pantallas, micrófonos y otros sensores se fusionan mostrando un entorno integrado.
- *EasyLiving [Shafer 98]*, se trata del esfuerzo investigador de Microsoft para desarrollar “entornos activos” completos, de modo que a través de computación ubicua ayude a las personas en multitud de tareas cotidianas. Uno de los mayores retos de este grupo es que la interacción se realice en lenguaje natural. Este entorno inteligente permite realizar el seguimiento e “identificación” de personas en una habitación. El concepto de identificación está acotado en su significado, indicando que puede distinguir a una persona dentro del grupo que se encuentra en la habitación. El sistema propuesto está formado por dos módulos binoculares en color, conectados a sendos PC's, encargados de determinar la posición de cada persona en el entorno. Estos dos PC's se comunican con un tercero que es el encargado de realizar las tareas de seguimiento y control. Con la información proporcionada por el sistema de visión pueden realizar diferentes tareas, como por ejemplo encender la televisión cuando una persona se sienta frente a ella o apagarla cuando se levanta, etc.
- *ODISEA [Haya 04]*, desarrollado por el Departamento de Ingeniería de la Universidad Autónoma de Madrid, proponen desarrollar un entorno inteligente en el cual se pueda interactuar en lenguaje natural. Está implementando un entorno domótico y un entorno ofimático reales, en los que se podrán realizar las tareas habituales en dichos lugares contando con el soporte proactivo del propio entorno. La habitación inteligente tendrá constancia de los ocupantes presentes en cada momento, de sus preferencias y de las tareas que en ese momento están realizando. A partir de esta información del contexto, la habitación inteligente realizará, de manera automática, acciones que ayuden a los ocupantes e interactuará con ellos mediante diálogos en lenguaje natural. Desde el punto de vista de la tecnología del habla, se propone la viabilidad de entablar diálogos en lenguaje natural apoyados en una información de contexto muy rica, obtenida a partir de otros modos de interacción (visión artificial y otros tipos de sensores). Se propone diseñar diálogos muy específicos para cada una de las posibles necesidades del usuario, así como un gestor de diálogos que decida cuál de estos se emplea en cada momento, cuándo se cambia de un diálogo a otro, así como qué información se intercambia entre ellos.

Los Entornos Inteligentes descritos hasta aquí se centran en la interacción usuario-entorno, pero no contemplan la presencia de robots como sistemas capaces de realizar diferentes tareas. Sin embargo también existen proyectos que incluyen robots móviles como elementos fundamentales del “espacio inteligente”. En este sentido, uno de los primeros proyectos que

incluye robots, con el objetivo de facilitar las tareas al usuario, es el “*Robotic Office Room*” (ROR) desarrollado por el Intelligent Cooperative System Laboratory de la Universidad de Tokio [Mizoguchi 96](figura 2.5), pensado para espacios interiores.



**Figura 2.5.** Entorno ROR del Intelligent Cooperative Systems Laboratory de la Universidad de Tokio. (a) Elementos del entorno. (b) Vista del brazo manipulador.

Pero los “entornos inteligentes” no sólo se limitan a espacios interiores, sino a todos los entornos de la vida cotidiana del hombre. Así, el proyecto *Smatr Cars* del Media Laboratory del MIT pretende crear un sistema de control inteligente de un coche, capaz de estimar las acciones a realizar en función de como se mueva el volante o se pise el acelerador.

Aunque las aportaciones en “entornos inteligentes” se centran fundamentalmente en la interacción entre el entorno y las personas, tal como ya se ha comentado, también están adquiriendo protagonismo los proyectos que incluyen robots móviles. Esta nueva concepción presenta importantes ventajas en aplicaciones para espacios interiores (*indoor*), ya que la información que se puede tener del entorno de movimiento del robot es más completa. Además, al disponer de diferentes puntos de vista se reduce la incertidumbre de la escena evitándose las oclusiones. Por otra parte, y este es un aspecto muy importante, los costes no aumentan por el hecho de introducir un nuevo robot en el entorno. A ello hay que añadir que la complejidad de los robots se reduce notablemente. Todas estas razones son las que están haciendo que la creación de “entornos inteligentes” para el control y guiado de robots móviles sea una nueva e interesante línea de investigación. Seguidamente se van a presentar algunos de los trabajos previos más importantes que tienen como objetivo fundamental el guiado de robots móviles utilizando cámaras situadas en puntos fijos del entorno por el que se deben mover los robots.



### 2.3.3. Guiado de robots utilizando cámaras ubicadas en puntos fijos del entorno.

Como ya se ha comentado, este nuevo enfoque en el guiado de robots móviles se caracteriza por la utilización de un array de cámaras ubicadas en puntos fijos del entorno y distribuidas de forma estratégica para abarcar la totalidad del campo de movimiento de los robots. La información proporcionada por todas las cámaras debe permitir determinar tanto la posición y orientación de los robots como cualquier otro objeto que se encuentre en el entorno.

Hasta la fecha son escasos los trabajos publicados, relacionados con el guiado de robots móviles a partir de la información obtenida por un array de cámaras ubicadas en puntos fijos del entorno. Pese a todo, dentro de los trabajos existentes, se puede hablar de dos tendencias, en función de ubicación de los elementos de proceso de imágenes. Así, se puede hablar de sistemas donde cada cámara lleva un procesador de imagen asociado (sistemas distribuidos) y aquellos que utilizan un sistema central de procesado. En el primer caso, cada cámara dispone del hardware necesario para procesar las imágenes, la información obtenida de ellas es comunicada a los robots, encargándose éstos de realizar una selección de esa información y actuar en consecuencia. En el segundo caso, existe una unidad de procesamiento central que es la encargada de procesar toda la información y generar las consignas de control.

Dentro de la primera línea (sistemas distribuidos) se pueden destacar los trabajos realizados por T. Sogo [Sogo99], que propone la utilización de 16 cámaras fijas (llamadas *Vision Agents*, VA) para realizar el guiado de dos robots móviles sobre la maqueta de una ciudad (figura 2.6). La particularidad de este trabajo está en la utilización de cámaras de color omnidireccionales no calibradas. El camino a seguir por los robots no se especifica como punto origen y punto final, sino que necesitan que un operador les “enseñe” el camino a seguir, camino que es memorizado por los diferentes VA’s para un posterior guiado. Durante la fase de guiado autónomo, cada VA compara la posición que ocupan los robots en la imagen con la de referencia, obtenida en la fase de aprendizaje, generando y enviando a cada robot las correspondientes consignas de guiado. La posición de los robots en una imagen se determina mediante la resta de esta imagen con la de referencia que contienen el fondo de la escena. Todos los VA’s envían sus consignas de guiado a cada robot, siendo éstos los encargados de elegir la más adecuada. Dicha consigna se corresponde con la proporcionada por el VA que se encuentra más cercana al correspondiente robot. Para evitar colisiones, se realiza la integración de las imágenes de cada una de las cámaras en una imagen general de todo el entorno. El sistema propuesto tiene una velocidad de captación de cinco imágenes por segundo, si bien los autores no especifican la precisión conseguida en el guiado. Este sistema presenta algunos inconvenientes, por una parte no se tiene información métrica de la posición

de los robots, y por otra la determinación de la posición de los robots en una imagen es poco robusta, por ser muy sensible a los cambios de iluminación.

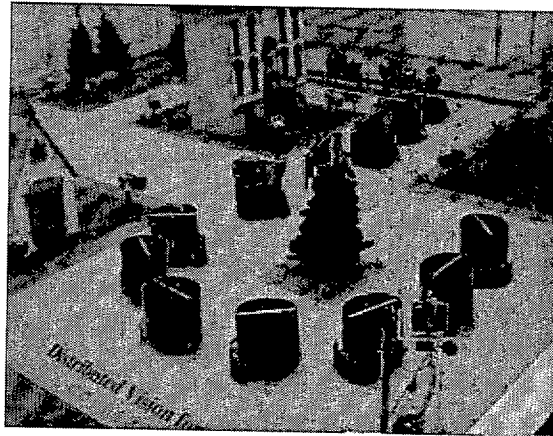


Figura 2.6. Sistema distribuido utilizado para los experimentos en [Soho 99].

Un trabajo más avanzado en esta misma línea es el propuesto por Lee [Lee 01][Lee 04]; en este caso se plantean la creación de un “espacio inteligente” (ISpace) [Lee99] (figura 2.7), por el que se pueden mover tanto robots como personas. Este ISpace es una zona (habitación, pasillo, calle, etc.) que está equipada con un conjunto de sensores (cámaras CCD, micrófonos, etc.), actuadores (altavoces, displays, etc.), una base de datos, un sistema de procesamiento y una red de comunicaciones entre todos ellos.

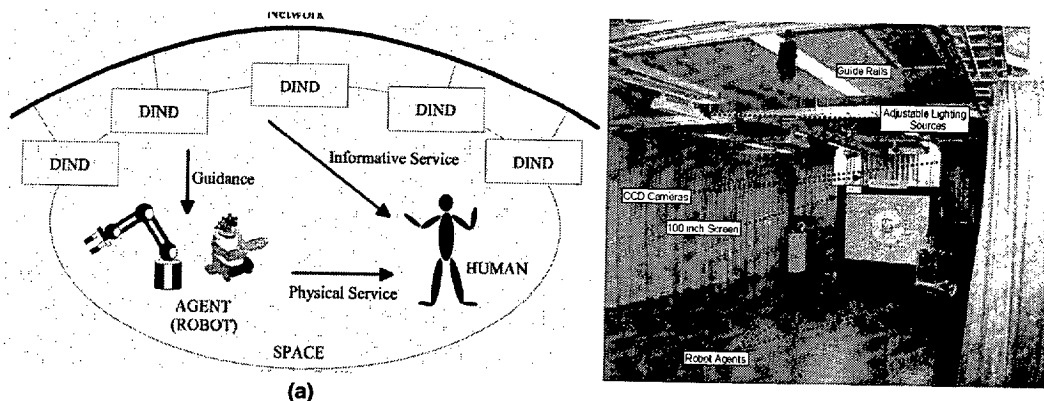
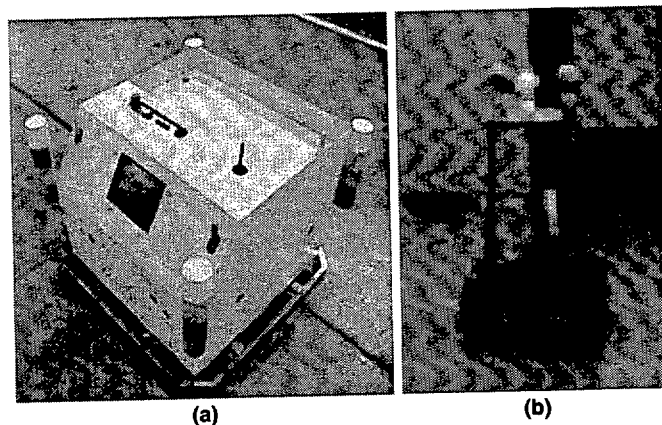


Figura 2.7. ISpace propuesto en [Lee 99]. (a) Estructura. (b) Entorno de pruebas.

Cada elemento del sistema de visión, denominado DIND (*Distributed Intelligent Networked Devices*) está formado por una cámara CCD de color, con capacidad de procesamiento, conectada a la red de comunicaciones. Cada uno de estos DIND capta el entorno identificando la presencia de una persona o un robot. Para ello realiza la resta de la imagen captada con la

que contiene sólo el fondo de la habitación. Esta información, que es comunicada a los otros DIND's, y a los robots, es la utilizada para realizar el guiado. Para localizar un robot en la imagen de una cámara, así como para distinguir un robot de otro, los autores proponen la utilización de marcas de colores ubicadas sobre los robots. En sus diferentes trabajos el número y forma de dichas marcas es distinto (figura 2.8). Para determinar la posición de los robots, los autores proponen la utilización de una única cámara, para ello suponen que la altura de los robots es conocida y que se mueven sobre una superficie plana. Para comprobar la eficacia de su propuesta han utilizado un entorno de prueba rectangular de 6 x 4 m, en el que han ubicado cinco DIND's. El error máximo cometido en la determinación de la posición de los robots es de 17 cm, la duración del ciclo de control es de 300 ms y la velocidad máxima de los robots es de 30cm/s.



**Figura 2.8.** Tipos de marcas artificiales utilizadas para detectar a los robots en el ISpace. (a) Cuatro marcas. (b) Tres marcas.

La detección, en una imagen, de una persona la realizan reconociendo el color de la piel de la cara y manos. En este caso, se propone la utilización de dos cámaras para determinar la posición de una persona.

Este espacio inteligente (ISpace) posibilita tanto la interacción hombre-robot (los autores muestran el seguimiento de una persona por parte de un robot [Morioka02]), como el guiado de un robot por dicho entorno, evitando los posibles obstáculos que se pueda encontrar.

Al utilizar una única cámara para determinar la posición del robot se pueden cometer errores muy elevados en aquellos puntos en los que la distancia entre cámara y robot es grande. A su vez, al realizar la localización de objetos móviles en la escena como una resta entre la imagen actual y la de referencia, hace al sistema sensible a los cambios de iluminación, igual que sucede en el trabajo de Sogo [Sogo99]. Otro inconveniente que comentan los autores es la baja

velocidad (30 cm/s) de desplazamiento de los robots que se consigue con este sistema. La ventaja de ISpace viene derivada de la independencia de los DIND's, de forma que el añadir más no requiere modificaciones en el mismo.

Otra propuesta que utiliza un sistema distribuido es la realizada en el proyecto MEPHISTO [Steinhaus 04] del Institute for Computer Design and Fault Tolerance de Alemania (figura 2.9). En él, las LAPU's (*Local Area Processing Units*) disponen de cámaras CCD dotadas de capacidad de procesamiento, encargadas de examinar el entorno y detectar objetos móviles. También se encargan de planificar y controlar las trayectorias de los robots que se encuentran en su área de influencia. Cada LAPU está interconectada a otras LAPU's que examinan zonas contiguas. Cada LAPU se conecta a una RCU (*Robot Control Unit*) encargada de comunicarse con los robots para enviar las consignas de control. Las RCU's se encargan de planificar las trayectorias globales concernientes a sus zonas de influencia. Al conjunto de una RCU y sus LAPU's lo denominan GAPU (*Global Area Processing Unit*), siendo la encargada de planificar la trayectoria global, que es dividida en secciones y comunicada a cada RCU.

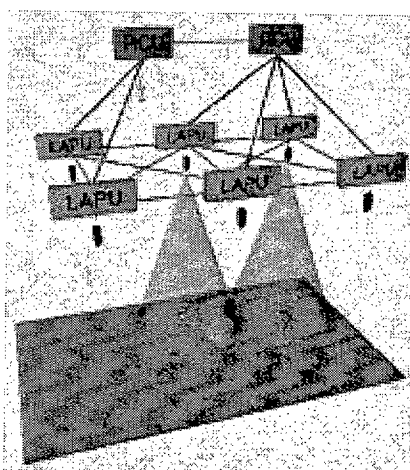
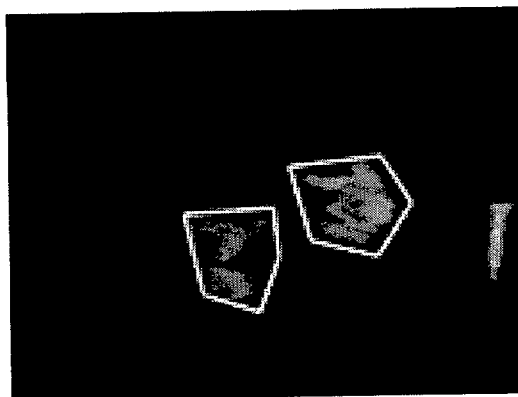


Figura 2.9. Estructura general del sistema MEPHISTO.

Para la detección de objetos móviles los autores proponen la resta de la imagen actual con una de referencia que contiene el fondo del entorno. Con el fin de que el sistema sea insensible a los cambios de iluminación, dicha imagen de referencia se va actualizando de forma continua. Un objeto móvil en la imagen se representa por un polígono (figura 2.10), cuya coordenadas píxelicas son convertidas a coordenadas métricas (globales) mediante una "look-up table", obtenida en el proceso de calibración. La intersección de los polígonos correspondiente a un mismo objeto captado por varias cámaras, permiten obtener la posición 3D del mismo.

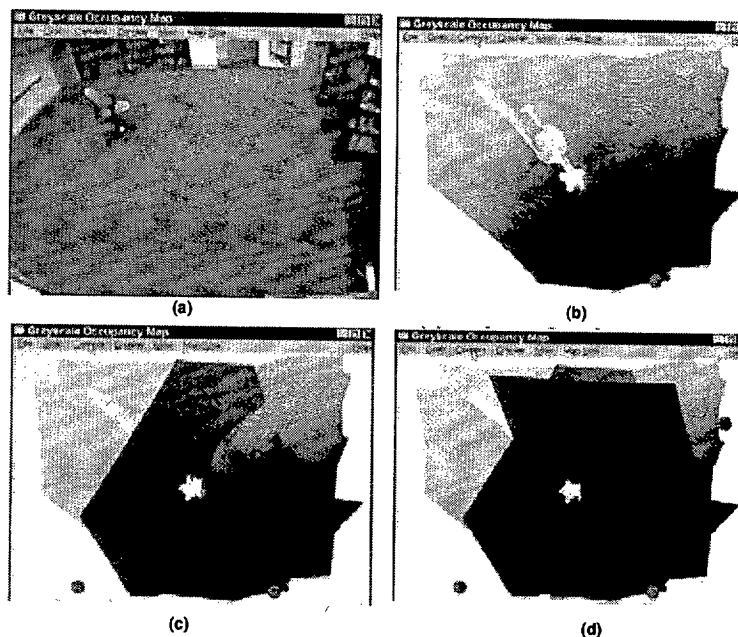


**Figura 2.10.** Representación con polígonos los objetos móviles en las imágenes de las cámaras del sistema MEPHISTO.

MEPHISTO es un sistema flexible que no precisa cambios significativos por el hecho de añadir nuevos sensores al sistema de visión. Sobre la precisión alcanzada por este sistema no se pueden dar datos ya que los autores en ninguno de sus trabajos aportan información al respecto.

Es evidente que los sistemas distribuidos presentan grandes posibilidades, pero tienen un coste elevado que en algunas aplicaciones resulta prohibitivo. A esto hay que añadir la dificultad en la sincronización de los diferentes sensores ubicados en el entorno (por ejemplo las cámaras). Como alternativa a esta solución están los sistemas centralizados, con los cuales se pueden reducir costes, aunque sus prestaciones pueden ser inferiores a los sistemas distribuidos.

Dentro de los sistemas centralizados se pueden encuadrar los trabajos desarrollados por Hoover y Olsen [Hoover 99][Hoover 00], en los que los autores utilizan cuatro cámaras ubicadas en posiciones fijas de una habitación para realizar el guiado de un robot. Las cámaras se encuentran conectadas a un PC, encargado de realizar el procesado de las imágenes y la elaboración de un mapa de ocupación que determina el entorno libre de obstáculos. Este mapa es transmitido a una *workstation* encargada de realizar la planificación del camino a seguir por el robot, generando las consignas de control, las que son transmitidas vía radio. El tiempo empleado para actualizar el mapa de ocupación es de 800 ms, consiguiendo mover al robot por una habitación de 6.5x9m con una velocidad de 20cm/s, mientras que las consignas se generan con una frecuencia de 1 Hz. En la figura 2.11.a se muestra el entorno de trabajo utilizado por los autores de este trabajo, y en las figuras 2.11.b, c, d se refleja el mapa de ocupación utilizando la información procedente de una, dos y tres cámaras, respectivamente.

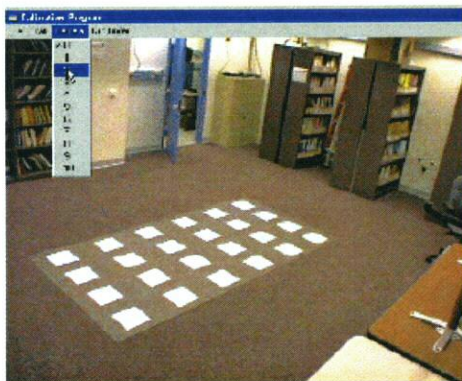


**Figura 2.11.** Mapa de ocupación obtenido de los trabajos realizados en [Hoover 00]. (a) Entorno de trabajo utilizado en las pruebas. (b) Mapa de ocupación realizado utilizando la información de una única cámara. (c) Mapa de ocupación obtenido con la utilización de dos cámaras. (d) Mapa de ocupación utilizando tres cámaras.

El sistema propuesto por Hoover y Olsen requiere de una primera fase de inicialización en la que se capta el entorno libre de obstáculos. Durante esta fase, en la escena no debe aparecer ningún objeto móvil ya que sería considerado como un obstáculo. Los autores utilizan un suelo cuyo color presenta un alto contraste con otros objetos de la escena (sillas, mesas, etc.). La zona libre de obstáculos estáticos determina el espacio por el que se pueden mover los robots. Durante esta fase de inicialización, también crea un patrón de referencias para cada una de las cuatro cámaras. Cada vez que una cámara capta una nueva imagen, ésta se compara con su patrón de referencia, píxel a píxel. Si la diferencia es inferior a un valor umbral, se considera el píxel vacío, en caso contrario se considera que estará ocupado. La ventaja de este método está en que las operaciones se realizan a nivel de píxeles, evitando el uso de complejos algoritmos de segmentación.

Para obtener las coordenadas 3D de un píxel los autores proponen la utilización de tablas de correspondencia o “*look-up tables*” (una por cada cámara), las cuales emparejan cada píxel con su correspondiente coordenada 3D de la escena. Estas “*look-up tables*” (LUT’s) se obtienen durante el proceso de calibración del sistema, operación que se realiza off-line bajo la supervisión de una persona. En primer lugar se obtienen los parámetros intrínsecos y extrínsecos de cada una de las cámaras, utilizando el algoritmo de Tsai para puntos coplanarios [Tsai 87] y una plantilla de calibración, como la mostrada en la figura 2.12, sobre

la que se indica el número de la cámara a calibrar, el tamaño de los elementos de la rejilla y las coordenadas 3D de un punto de referencia. Una vez obtenidos los parámetros intrínsecos y extrínsecos de cada una de las cámaras, obtienen sus LUT's haciendo uso de las ecuaciones de proyección. Con este sistema los autores obtienen errores de posicionamiento en torno a 4.3 mm de media y 2.7 mm de varianza.



**Figura 2.12.** Plantilla utilizada para calibrar las cámaras del sistema de visión propuesto en [Hoover 00].

El sistema propuesto por Hoover y Olsen presenta varios inconvenientes. El primero de ellos es que si en el entorno aparece un objeto con un color que difiere del fondo en un valor menor del umbral, no será detectado. Así, el funcionamiento del sistema es sensible al valor del umbral seleccionado, debiendo éste estar comprendido entre 15 y 50 para que el sistema funcione satisfactoriamente. Además, al utilizar una única imagen para realizar la calibración de las cámaras la precisión obtenida es baja. Por otra parte, los parámetros extrínsecos de las cámaras deben estar referidos a un mismo sistema global y en consecuencia todas las LUT's. Esto obliga a colocar la plantilla de calibración en posiciones conocidas respecto de una referencia global; para ello los autores proponen rellenar todo el entorno con plantillas de calibración tal y como se muestra en la figura 2.13.

Un sistema similar al anterior es el presentado por Kruse y Wahl [Kruse98] (figura 2.14) en el proyecto MONAMOVE (MONitoring and NAVigation for MOBILE VEHICLES) para realizar el guiado de un robot móvil en un entorno industrial. En este trabajo los autores proponen utilizar un sistema sensorial formado por cámaras estáticas, encargadas de percibir el entorno de movimiento del robot, en conjunción con otros sensores (ultrasonidos, sensores de contacto, etc.) montados sobre el robot, los cuales son utilizados para realizar paradas de emergencia y realizar el posicionamiento en movimientos precisos.

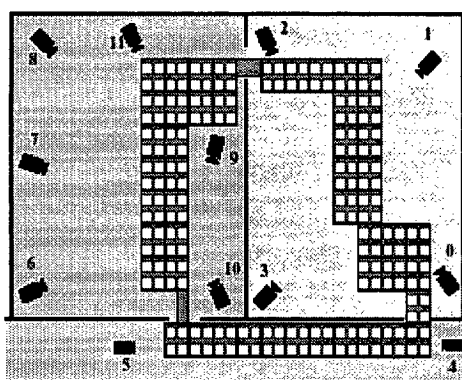


Figura 2.13. Sistema de calibración propuesto por Hoover y Olsen.

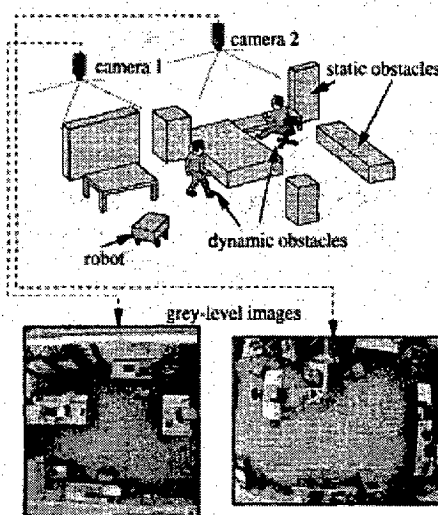


Figura 2.14. Entorno de trabajo utilizado por Kruse y Wahl para sus trabajos.

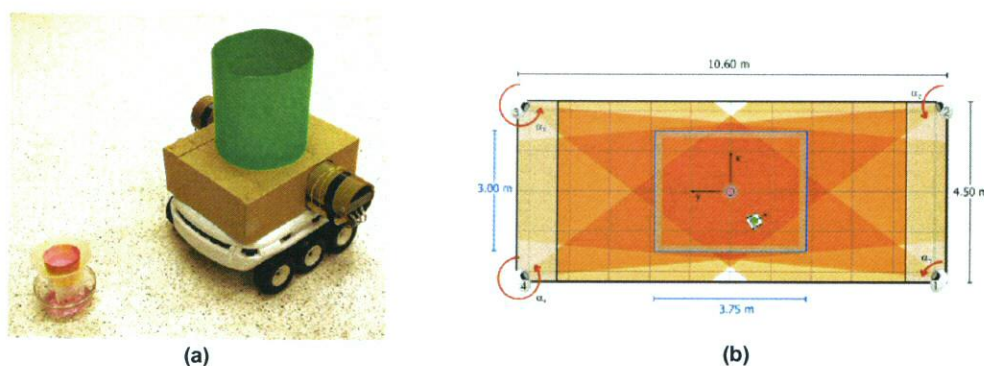
Al igual que Hoover y Olsen, Kruse y Wahl proponen la utilización de tablas de correspondencia para obtener las coordenadas 3D de un píxel en cada una de las cámaras y con esta información realizar un mapa de ocupación. Para la obtención de estas tablas se utilizan plantillas con marcas blancas colocadas sobre todo el suelo del entorno de trabajo, en posiciones conocidas. Los obstáculos estáticos son marcados de forma manual sobre las imágenes de las diferentes cámaras, determinándose el espacio libre por el cual se puede mover el robot. Por otra parte, la detección de obstáculos móviles se realiza en base a un análisis diferencial de las imágenes, de forma que cada nueva imagen es comparada con las de referencia tomadas sobre el entorno estático, al igual que el método propuesto por Hoover y Olsen. Existe una pequeña diferencia entre estos métodos, ya que Kruse y Wahl utilizan dos imágenes de referencia, en vez de una sola. Estas dos imágenes de referencia son tomadas con dos niveles de iluminación, uno alto y otro bajo; estas imágenes son sumadas ponderándolas



con un factor de escala. Según los autores, con esta técnica el sistema propuesto es insensible a los cambios de iluminación. Kruse y Wahl proponen la colocación de las cámaras de forma que los campos de visión de las mismas se solapen parcialmente y así conseguir que la mayor parte de las zonas sólo sean captadas por una única cámara. En este trabajo, los autores utilizan la odometría como sistema auxiliar de guiado, actualizando de forma periódica su posición absoluta utilizando el sistema de visión.

Otro trabajo que se puede enmarcar dentro de los sistemas centralizados es el realizado por J. Santos [Santos01] para el guiado de un pequeño robot utilizando una única cámara. Para calcular la posición 3D el autor supone que el robot se mueve por una superficie plana con coordenada Z nula ( $Z=0$ ). La cámara previamente es calibrada obteniéndose los 9 coeficientes de la matriz de proyección (la matriz de proyección de una cámara tiene un tamaño de  $3 \times 4$ , pero si se considera  $Z=0$  se reduce a  $3 \times 3$ ). El sistema propuesto fue probado en un entorno libre de obstáculos, proporcionando unos resultados aceptables, pero el autor no indica el valor de los errores cometidos durante el guiado. Obviamente, al utilizar una única cámara los errores de posicionamiento van a ser mayores, tanto más cuanto más se aleje el robot de la misma.

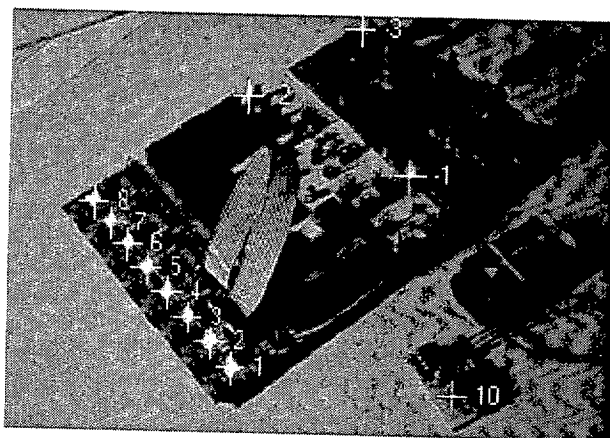
Otro sistema de posicionamiento global es el presentado en [Lilienthal 03]. En él, los autores utilizan un cilindro de color verde (figura 2.15.a) situado sobre el robot, como marca para facilitar la detección. La pruebas las han realizado en un entorno de  $10 \times 3$  m sobre el cual se han colocado cuatro cámaras RGB (figura 2.15.b).



**Figura 2.15.** Sistema de posicionamiento propuesto [Lilienthal. 03]. (a) Robot empleado en las pruebas. (b) ubicación de las cámaras en el entorno de trabajo.

Un sistema de posicionamiento, que se asemeja al propuesto en esta tesis, es el realizado por Aitenbichler y Mühlhäuser [Aitenbichler 03]. Los autores utilizan un sistema binocular, previamente calibrado, ubicado en una posición fija del entorno de trabajo, para determinar

la posición del robot. Para la identificación del robot, a partir de la información proporcionada por el sistema binocular, los autores colocan sobre él un conjunto de LED's infrarrojos (figura 2.15). Para comprobar la precisión de su sistema los autores han realizado medidas en diferentes posiciones (un total de 138) obteniendo un error medio de 16.67 cm. Para dichas pruebas utilizaron un entorno muy idealizado, ya que la placa que contiene los diodos IR se coloca en frente del sistema binocular a diferentes distancias (comprendidas entre 1m y 3 m), en una sala sin ventanas y con las luces apagadas. Además, las medidas se realizan con el array de LED's en régimen estático.



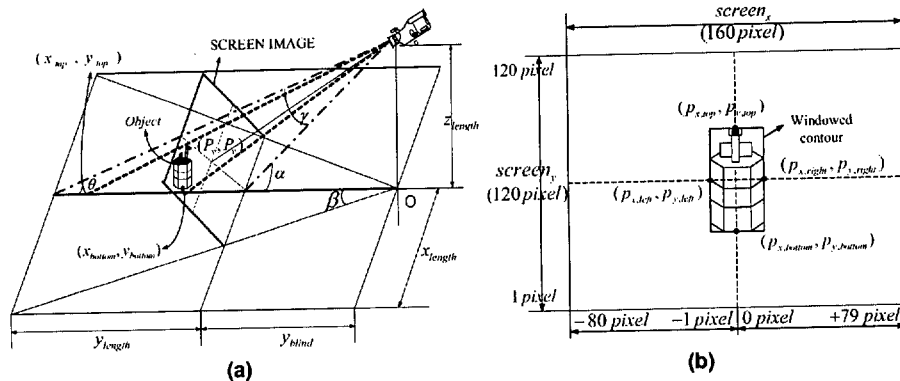
**Figura 2.16.** Array de diodos IR utilizados en [Aitenbichler 03] para identificar el robot en una imagen.

Un trabajo singular, por la ausencia de marcas colocadas sobre los robots, es el desarrollado en [Yun 04]. En él, los autores obtienen la posición 3D del robot utilizando una única cámara. Para la identificación de los objetos móviles, al igual que otros autores, utilizan la resta de la imagen actual con una de referencia. Una vez identificado el robot en una imagen y obtenidas las coordenadas píxelicas correspondientes a los bordes (figura 2.17.a) es posible determinar las coordenadas 3D (figura 2.17.b) si se conoce las dimensiones del robot. Con esta propuesta, el error de posicionamiento cometido es de 2.6 cm cuando el robot se encuentra a una distancia de la cámara de 3m, mientras que el error es de 4.8 cm si dicha distancia es de 4m.

## 2.4. Modelo de cámara.

Un aspecto importante en todos los sistemas que utilizan cámaras, y de forma especial cuando éstas tienen que obtener información de puntos alejados, es tener un conocimiento preciso de las relaciones que ligan las coordenadas en el espacio 3D con las coordenadas en el plano imagen (coordenadas 2D). Estas relaciones dependen, por una parte, de la geometría de cámara

(posición y orientación en el espacio relativa a algún sistema de coordenadas de referencia - sistema de coordenadas del mundo) y, por otra, del modelo de cámara (geometría de cámara), que describe las características internas de la misma (distancia focal, distorsión de las lentes, etc.). Las características internas se denominan generalmente *parámetros intrínsecos*, mientras que los parámetros que describen la geometría se denominan *parámetros extrínsecos*.



**Figura 2.17.** Relaciones utilizadas en [Yun 04] para determinar la posición 3D de un robot utilizando una única cámara. (a) Modelo de correspondencia para obtener la posición 3D del robot. (b) Representación de las coordenadas píxelicas del robot en una imagen.

Dada la importancia que el modelo de cámara tiene en la aplicación propuesta en esta tesis, en este apartado se va a hacer una breve revisión de los trabajos más importantes sobre modelos de cámaras. En este sentido, hay que decir que se han propuesto diferentes modelos [Ito 91] y que la elección de uno u otro depende, en gran medida, del tipo de escenas con las que se está trabajando y de la precisión que se desee obtener en las medidas. Evidentemente, esta precisión estará también estrechamente relacionada con la exactitud en la obtención de los parámetros del modelo de la cámara utilizado. El proceso que permite obtener los distintos parámetros que intervienen en la transformación 3D-2D se conoce como calibración de cámara.

Seguidamente se va a presentar un breve resumen sobre algunos de los modelos más significativos de cámara propuestos y que permitirán realizar la elección más adecuada a la propuesta de aplicación de esta tesis. Así mismo, se presentan algunas de las alternativas propuestas sobre calibración, tanto de una sola cámara como de múltiples cámaras, siendo este último caso el de mayor interés desde el punto de nuestra aplicación.

### 2.4.1. Modelos sin distorsión.

En este punto se van a analizar los modelos matemáticos que se aproximan al proceso de formación de imagen, considerando que las lentes utilizadas en la óptica son ideales (no presenta distorsiones ni aberraciones ópticas).

**Modelo de proyección perspectiva o cónica:** Es el modelo de cámara más empleado debido fundamentalmente a su sencillez y a sus buenas prestaciones en aplicaciones prácticas. En la literatura tradicional este modelo es referido como modelo *pin-hole* o modelo de lente delgada [Faugeras 93] (figura 2.18).

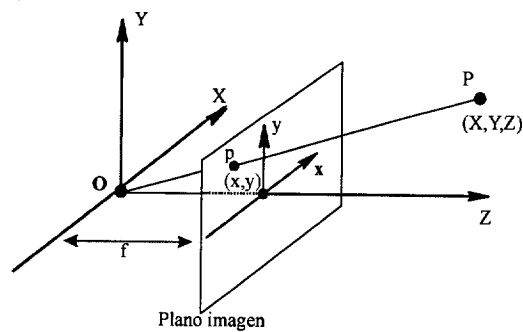


Figura 2.18. Modelo de proyección perspectiva.

En este modelo un punto  $P$  en el espacio 3D con coordenadas  $(X, Y, Z)$  referidas a un sistema de coordenadas ubicado en la propia cámara (tal como se muestra en la figura 2.18) se proyecta en el plano imagen en un punto  $p$ , con coordenadas  $(x, y)$  (coordenadas de la imagen), a través de las ecuaciones (2.1).

$$\begin{aligned} x &= f \cdot \frac{X}{Z} \\ y &= f \cdot \frac{Y}{Z} \end{aligned} \quad (2.1)$$

o en coordenadas homogéneas:

$$\begin{bmatrix} \lambda \cdot x \\ \lambda \cdot y \\ \lambda \end{bmatrix} = \begin{bmatrix} f \cdot X \\ f \cdot Y \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.2)$$

**Modelo de proyección ortográfica:** Este modelo es una simplificación del modelo de perspectiva que linealiza las ecuaciones de proyección. La proyección ortográfica ideal se da en el caso de considerar que la distancia a la cual está el objeto, respecto de la cámara, es infinita, motivo por el cual los rayos procedentes del objeto son paralelos al eje óptico, esto es, perpendiculares al plano imagen [Horaud 94]. Este tipo de proyección se da en varias circunstancias de la vida real, como por ejemplo en imágenes aéreas tomadas por un avión o satélites. Bajo este modelo se ignora completamente la información de profundidad. Las ecuaciones de proyección de modelo de proyección ortográfica son:

$$\begin{aligned}x &= X \\y &= Y\end{aligned}\tag{2.3}$$

**Modelo de proyección débil:** Este modelo, también denominado de proyección ortográfica escalada, es otra simplificación del modelo de perspectiva. En este caso se supone que todos los puntos se encuentran en un plano (plano objeto) paralelo al plano imagen y situado a una distancia fija de este último ( $Z=Z_0$ ). Para que esta aproximación sea válida, se supone que el objeto es suficientemente pequeño con respecto de la distancia a la cámara [DeMenthon 95]. Esto es razonable por ejemplo en imágenes aéreas. Las ecuaciones de proyección son:

$$\begin{aligned}x &= f \cdot \frac{X}{Z_0} \\y &= f \cdot \frac{Y}{Z_0}\end{aligned}\tag{2.4}$$

**Modelo de proyección paraperspectiva:** En la proyección de perspectiva débil, un punto de la escena se proyecta primero sobre el plano objeto, usando rectas paralelas al eje óptico. Esto provoca un error de aproximación que es función de la distancia de un objeto al eje óptico, siendo mayor cuanto más grande sea esta distancia. La proyección paraperspectiva minimiza este error, haciendo que las rectas de proyección sean paralelas a la denominada *recta de proyección central*, recta que une el centroide del objeto de la escena (de coordenadas  $[X_G, Y_G, Z_G]$ ) con el centro óptico. De forma similar al modelo de proyección débil, la proyección paraperspectiva emplea un plano objeto. Las rectas de proyección de todos los puntos se proyectan paralelas a la recta de proyección central hasta este plano. Posteriormente los puntos se proyectarán del plano de profundidad media al plano de imagen de forma perspectiva. Las ecuaciones de proyección de este modelo son:

$$\begin{aligned}x &= f \cdot \frac{X}{Z_G} - \frac{X_G}{Z_G^2} Z + \frac{X_G}{Z_G} \\y &= f \cdot \frac{Y}{Z_G} - \frac{Y_G}{Z_G^2} Z + \frac{Y_G}{Z_G}\end{aligned}\tag{2.5}$$

**Modelo de proyección afín:** En la transformación afín se preservan las distancias y los ángulos existentes en la imagen original, además de las transformaciones lineales como son las imágenes en espejo, escaladas o bien con cierto desplazamiento de la imagen original. De esta manera la única condición a cumplir es que las líneas paralelas siguen siendo paralelas después de la transformación. La diferencia entre el modelos afín y el proyectivo está en la identificación del plano en el infinito, cuya definición y uso se expone en [Quan, 96] [Quan 97].

En [Xu 96] se realiza un interesante estudio de los distintos modelos expuestos, cuantificando el error que se comete en cada uno de ellos. Así, la proyección débil puede verse como una aproximación de orden cero al modelo perspectivo, mientras el modelo paraperspectivo es una aproximación de orden uno.

#### 2.4.2. Modelo de perspectiva con distorsión.

Como es sabido, las ópticas introducen aberraciones que distorsionan la proyección en el plano imagen de los puntos en el espacio 3D. Las aberraciones son causadas por la naturaleza del medio óptico, ya que este puede presentar distinto comportamiento a los fenómenos de refracción (cambio de la dirección de un rayo al pasar de un medio óptico a otro) y dispersión (el ángulo de refracción difiere para distintas longitudes de onda) en función de la longitud de onda. En consecuencia, la distorsión va a introducir nuevos parámetros intrínsecos para la descripción del modelo interno de la cámara.

La distorsión modifica la posición de los puntos en la imagen, de tal modo que éstos realmente no se encuentran donde señalan las ecuaciones de proyección (ecuaciones 2.1), sino desplazados. Así, un punto  $P$  que se debería proyectar en el punto  $p$  si no existiera distorsión, realmente lo hace en  $p'$  (figura 2.19). La magnitud de este desplazamiento es función de la posición en la que se encuentre el punto  $p$  en el plano de imagen.

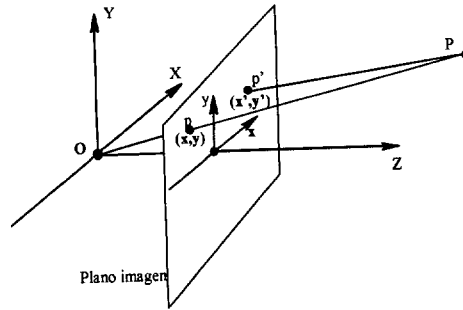


Figura 2.19. Proyección perspectiva con distorsión.

Identificando por  $(x,y)$  y  $(x',y')$  las coordenadas en la imagen de los puntos  $p$  y  $p'$ , respectivamente, se tiene:

$$\begin{aligned} x' &= x + \delta_x(x,y) = x + \delta_{xr}(x,y) + \delta_{xt}(x,y) \\ y' &= y + \delta_y(x,y) = y + \delta_{yr}(x,y) + \delta_{yt}(x,y) \end{aligned} \quad (2.6)$$

Los términos  $\delta_x$  y  $\delta_y$  resumen la distorsión de forma genérica, la cual puede modelarse de distintas formas según el tipo de distorsión predominante. Los valores de la componente radial ( $\delta_{xr}$ ,  $\delta_{yr}$ ) y tangencial ( $\delta_{xt}$ ,  $\delta_{yt}$ ) para las coordenadas horizontal ( $x$ ) y vertical ( $y$ ) se modelan como una serie de infinitos términos. Sin embargo, para la mayoría de las aplicaciones es suficiente con la utilización de dos términos para cada una de ellas. Así, la distorsión se puede modelar como [Lázaro 00]:

$$\begin{aligned} \delta_{xr} &= x(k_1 r^2 + k_2 r^4) \\ \delta_{yr} &= y(k_1 r^2 + k_2 r^4) \\ \delta_{xt} &= 2p_1 xy + p_2(r^2 + 2x^2) \\ \delta_{yt} &= p_1(r^2 + 2y^2) + 2p_2 xy \\ r &= \sqrt{(x^2 + y^2)} \end{aligned} \quad (2.7)$$

En consecuencia, la ecuación (2.6) se puede expresar como:

$$\begin{aligned} x' &= x + x(k_1 r^2 + k_2 r^4) + 2p_1 xy + p_2(r^2 + 2x^2) \\ y' &= y + y(k_1 r^2 + k_2 r^4) + p_1(r^2 + 2y^2) + 2p_2 xy \end{aligned} \quad (2.8)$$

donde  $k_1$  y  $k_2$  son los coeficientes que modelan la distorsión radial y  $p_1$  y  $p_2$  la tangencial. La ecuación (2.8) se puede expresar de forma matricial como:

$$p' = p + \delta(p) \quad (2.9)$$

donde,  $\delta(p)$  refleja la distorsión del punto  $p$ .

**Corrección de la distorsión:** La distorsión afecta de forma significativa en los sistemas que utilizan la visión para obtener una información métrica precisa. Estos errores debidos a la distorsión han sido analizados en [Lázaro 00a] y [Lázaro 00b] con un sistema de ecuaciones que resuelve, de forma general, las diferentes aberraciones que puedan presentarse en una lente, considerando que dichas deformaciones se producen como consecuencia de la variación de forma no-lineal del módulo y la fase de los píxeles de la imagen a partir del punto principal.

Para poder utilizar los modelos de cámara descritos en la sección 2.4.1 es necesario corregir la distorsión introducida por la óptica de la cámara. Esto es, obtener las coordenadas del punto  $p$  a partir de las coordenadas del punto  $p'$ . De ahí, la necesidad de tener un método rápido para pasar de una imagen con distorsión a una imagen libre de distorsiones, si bien esta operación no tiene una solución analítica. Varias han sido las propuestas realizadas para obtener una aproximación. A modo de ejemplo, Toscani y Faugeras [Toscani 87] proponen la utilización de una transformación bilineal en pequeñas regiones de la imagen.

Otra propuesta más simple, y no por ello menos precisa, fue la realizada por Melen [Melen 94] en la que propone el uso del siguiente modelo:

$$p \approx p' - \delta(p') \quad (2.10)$$

Esta ecuación se puede convertir en el modelo recursivo:

$$\begin{aligned} p &= p' - \delta(p) \\ &= p' - \delta(p' - \delta(p)) \\ &= p' - \delta(p' - \delta(p' - \delta(p))) \\ &= p' - \delta(p' - \delta(p' - \delta(p' - \delta(p)))) \\ &= \dots \end{aligned} \quad (2.11)$$

En la práctica, 2 o 3 iteraciones son suficientes para corregir la distorsión, incluso en casos extremos.

Heikkilä [Heikkilä 00] propone otro modelo alternativo al modelo recursivo de Melen, reduciendo su tiempo de cómputo. Para ello utiliza el desarrollo en series de Taylor de la función  $\delta(p)$  (ecuación 2.9) al rededor de  $p$  obteniendo:



$$p \approx p' - \frac{\delta(p')}{4k_1r^2 + 6k_2r^2 + 8p_1y + 8p_2y + 1} \quad (2.12)$$

Un modelo más elaborado, y en consecuencia con un coste computacional mayor, es el presentado en [Tamaki 02]. Este autor propone un algoritmo que resuelve la corrección de la distorsión entre la imagen original y la corregida como el proceso inverso de la imagen corregida a la original, unificando ambos caminos.

## 2.5. Calibración.

La calibración es el proceso mediante el cual se determinan los parámetros internos o intrínsecos (matriz de proyección) y externos o extrínsecos (posición de la cámara respecto de un sistema de referencia global). Por tanto, tal como ya se comentó, el proceso de calibración permite obtener todos los parámetros que intervienen en las ecuaciones que relacionan las coordenadas de los puntos en el mundo 3D y las coordenadas 2D de una imagen. Dependiendo de la precisión requerida, los modelos de proyección utilizados son los de perspectiva y ortográfico [Heikkilä 00]. El modelo ortográfico se utiliza en aquellas aplicaciones en las que no se requiere una precisión elevada y representa una solución con una carga computacional pequeña. Por su parte, el modelo de proyección de perspectiva aporta una precisión mayor, que se puede aumentar añadiendo el modelo de distorsión de la lente.

Los métodos estándar de calibración de cámaras utilizan un patrón cuya geometría es conocida, sobre el que se disponen  $N$  puntos de interés con coordenadas  $(P_i = [X_i, Y_i, Z_i]^T, \text{ para } i = 1 \dots N)$ . El patrón es captado por la cámara y sus puntos de interés son vistos como puntos 2D con coordenadas píxelicas  $(u_i, v_i)$ . Partiendo del modelo de proyección es posible obtener una estimación teórica de los puntos 3D y a partir de ella determinar los parámetros de la cámara. Este proceso se divide en dos pasos: en primer lugar se emparejan las coordenadas 2D en la imagen de cada uno de los puntos del patrón de calibración con sus coordenadas 3D. A continuación, se genera un conjunto de ecuaciones que, utilizando las correspondencias anteriores, permita obtener el modelo de la cámara. Obviamente, es en este último paso donde se centra la literatura existente.

Toscani y Faugeras [Toscani 86] proponen el método de la Transformación Lineal Directa (DLT). En este caso, en primer lugar se calculan los elementos de la matriz de proyección  $H$ , para a partir de ellos obtener los parámetros intrínsecos y extrínsecos. En este método se realiza una aproximación lineal de la calibración considerando que la matriz de rotación  $R$  está

formada por 9 parámetros  $r_{ij}$  independientes, cuando en realidad  $R$  sólo tiene 3. La matriz  $R$  es una matriz ortogonal de  $3 \times 3$  que viene definida por los ángulos de Euler ( $\alpha$ ,  $\beta$ ,  $\gamma$ ), los cuales reflejan la rotación que tiene el sistema de coordenadas de la cámara respecto al sistema de referencia global. Por ello, en la solución final obtenida existirán errores debidos a que el sistema de ecuaciones es no lineal, en contra de lo supuesto por este método. Otra desventaja que presentan es que no permiten obtener el modelo de distorsión de la lente, lo que genera una pobre precisión. La ventaja que presenta es la sencillez de cálculo.

Un método más preciso es la calibración geométrica, que tiene en cuenta otros factores que intervienen en el proceso de formación de las imágenes como la distorsión introducida por la lente. El método seguido es no lineal [Salvi 97]. Los parámetros de las cámaras se obtienen a través de iteraciones hasta minimizar la distancia entre los puntos de la imagen y el modelo de proyección obtenido con las sucesivas iteraciones. Este método aporta una mayor precisión a costa de un mayor coste computacional. Además, es necesario proporcionar unas buenas condiciones iniciales para que el sistema converja.

Otras técnicas realizan la calibración en dos pasos (*Two-Step*) [Tsai 87] [Weng 92], las cuales, en un primer paso calculan algunos parámetros utilizando una optimización lineal y en un segundo paso el resto de parámetros, todo ello de forma iterativa. Los resultados son una calibración rápida y precisa, y además se garantiza la convergencia.

Estos métodos de calibración están basados en el procesamiento de una única imagen. La captura del patrón debe incluir numerosos puntos de los cuales se conoce su posición 3D. Hay que dejar patente que dichos métodos introducen errores en la medida de las coordenadas de los puntos posteriormente utilizados en el algoritmo de optimización. Además, presentan problemas cuando la distancia focal es muy pequeña (caso de objetivos de gran angular) no corrigiendo las distorsiones que presentan dicho tipo de objetivos. Estos métodos suponen conocidas las coordenadas reales de los elementos del patrón, formando parte del proceso de optimización, por lo que los errores e incertidumbres de las medidas serán arrastrados en todos los cálculos posteriores. Las técnicas multi-imagen complementan dichos algoritmos e introducen como incógnitas las coordenadas de los puntos del patrón, mejorando el proceso de optimización del modelo de formación de las imágenes. Así, los métodos tradicionales que hacían uso de una única imagen han dejado paso a métodos donde se considera una secuencia de imágenes en las que se realiza un seguimiento de una serie de puntos característicos fácilmente reconocibles.

En [Lavest 92] se muestra un método de calibración multi-imagen que utiliza una secuencia de imágenes y ecuaciones fotogramétricas, resolviendo el problema de la corrección de la distorsión a la vez que se obtienen los parámetros intrínsecos del sistema y del patrón utilizado para la calibración. La principal característica de dicho trabajo es la utilización de varias imágenes captando diferentes posiciones de un mismo patrón de calibración. Con ello no se tiene la necesidad de realizar medidas de coplanaridad ni medidas precisas sobre los elementos y marcas que conforman el patrón, ya que se introducen las posiciones de los puntos del patrón como incógnitas adicionales a optimizar. Esto proporciona una calibración más precisa. En el apéndice A se analizan más en profundidad las mejoras y los avances introducidos en las técnicas de calibración de cámara multi-imagen. Trabajos similares pueden verse en [Zhang 99] y [Heikkilä 00].

Distintos investigadores están buscando soluciones que permitan una calibración automática, haciendo uso de un patrón de puntos generados virtualmente con la ayuda bien de una pantalla TFT de grandes dimensiones o de un video proyector [Gardel 03]. Actualmente el problema de dichas aproximaciones es que la precisión en la generación que se requiere al reproducir el patrón de forma virtual debe ser muy preciso, dado que de ella depende la exactitud de los parámetros de calibración obtenidos posteriormente.

Otros autores [Cipolla 98] [Harley 93] proponen la realización de la calibración utilizando la información contenida en las propias imágenes (auto-calibración). En estos métodos se toman varias imágenes de una misma escena y a partir de la correspondencia entre puntos de distintas imágenes se pueden encontrar los mejores parámetros del modelo que puedan otorgar estas correspondencias. La reconstrucción 3D realizada con este modelo, estará afectada por un factor de escalado, ya que no se puede conocer cuál es el tamaño real de los objetos.

### **2.5.1. Patrones de calibración.**

Como se ha puesto de manifiesto, para calibrar la cámara se deben captar puntos de la escena correspondientes a unas marcas artificiales pertenecientes a un patrón. A continuación se muestran una serie de patrones ampliamente utilizados en la calibración de las cámaras. Se analizan las ventajas y los problemas que posee cada uno de ellos.

La precisión de la calibración depende de la exactitud en la determinación de las coordenadas  $(u, v)$  de los puntos de interés (centro círculos, intersección de líneas, etc.), por lo que deben estar perfectamente definidos en una imagen. Así, por ejemplo, en el caso de utilizar un patrón con círculos es necesario que el contorno sea perfecto para que los algoritmos de detección

subpíxelica puedan determinar el centroide con precisión, además, el tamaño de los puntos debe ser el adecuado (en torno a unos 20 píxeles de diámetro) [Lavest 98].

### Damero de ajedrez.

Este patrón plano (figura 2.20) es muy utilizado dado que la detección de los puntos de intersección tras pasar un filtro de bordes resulta bastante sencilla. Por el contrario la precisión subpíxelica del mismo no es muy elevada. Existen algoritmos [Bouguet 99] de detección que introduciendo directamente el número de recuadros horizontales y verticales del damero y obtiene los puntos de cruce que luego se utilizan en el proceso de calibración.

Conviene hacer notar que el simple hecho de obtener la línea con un filtro detector de bordes hace que la precisión posterior no sea elevada. Además, si sólo se utiliza una imagen del mismo se pierde el factor profundidad. Si se utiliza una técnica multi-imagen se puede hacer uso de un patrón plano en diferentes localizaciones, lo que proporciona las medidas de profundidad necesarias para realizar una calibración fiable. En este apartado se ha de reseñar la aplicación desarrollada en [Zhang 99], la cual hace uso de este patrón de calibración, comúnmente denominado como patrón de Tsai.

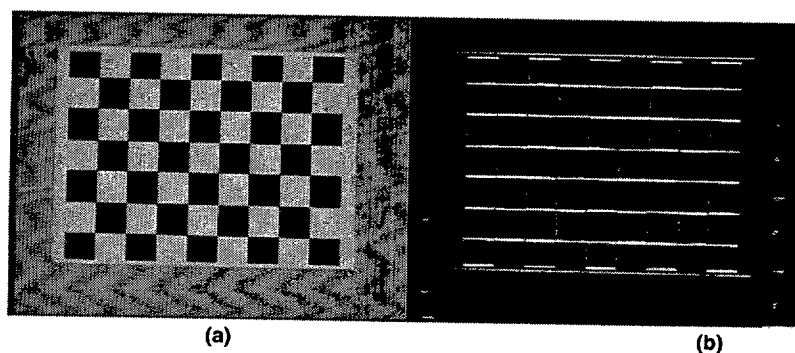
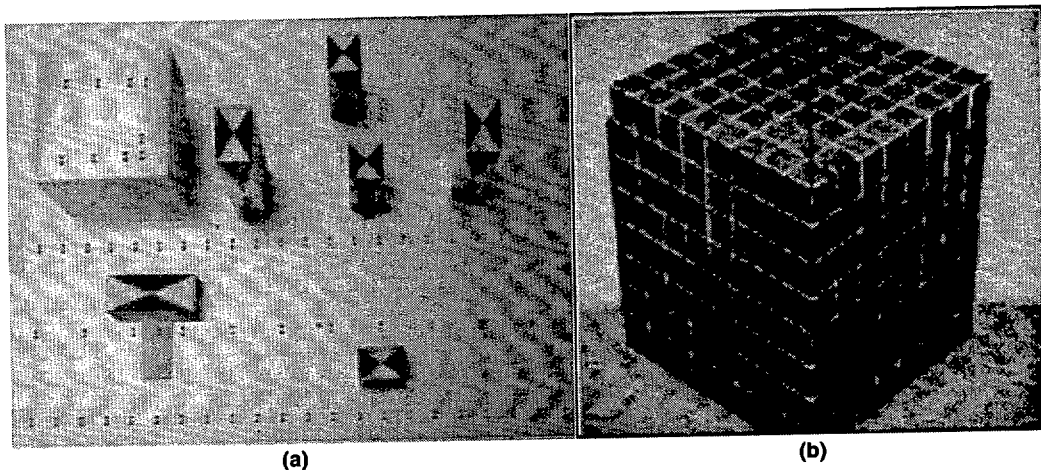


Figura 2.20. Plantilla en forma de damero de ajedrez. (a) Estructura. (b) Detección de bordes.

### Plantilla 3D.

Para captar en una sola imagen puntos a diferentes alturas (figura 2.21.a), algunos investigadores han realizado objetos de calibración dispuestos a varias alturas o profundidades.

La ventaja de estas plantilla radica en que en una sola imagen se obtiene puntos suficientes para calibrar el sistema sin los problemas del factor profundidad que conlleva disponer de una sola imagen de un patrón plano.

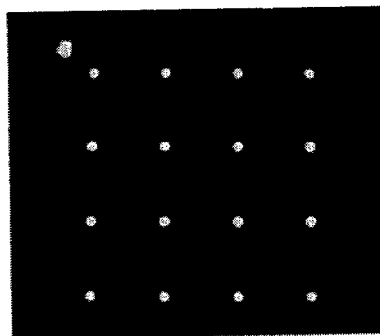


**Figura 2.21.** Plantillas 3D de calibración. (a) Con objetos a diferentes alturas. (b) Con forma de cubo.

En otras ocasiones la utilización de cubos tridimensionales (figura 2.21.b) permite la optimización y obtención de los parámetros de calibración al incorporar diferentes orientaciones de puntos con respecto a la cámara [Lavest 92]. Con una única imagen y conociendo las posiciones reales de los puntos extraídos se puede calibrar la cámara.

### **Plantilla plana con círculos.**

Otro patrón muy utilizado es aquel que consta de círculos (figura 2.22). Su ventaja radica en una mejor detección del centro del círculo en cualquier perspectiva. En este caso se puede utilizar un modelo basado en la luminancia [Heikkilä 00] para obtener, con una elevada precisión, su centro.



**Figura 2.22.** Plantilla de calibración con elementos circulares.

La utilización de patrones planos resulta muy adecuada en los algoritmos multi-imagen donde la precisión de la calibración no recae en la precisión con la que se haya fabricado el patrón sino en la detección exacta del mismo punto en todas las imágenes de la secuencia [Dhôme

89]. Nótese que al captar el patrón desde diferentes localizaciones y ángulos ya se está considerando el factor profundidad.

### **Plantilla esférica.**

Una idea para poder calibrar con la misma escena no sólo una cámara sino un conjunto de ellas es utilizar esferas como patrones de calibración, pudiendo ser captadas simultáneamente por todas las cámaras [Agrawal 03]. Dicha forma tridimensional hace que sea posible captar el mismo punto 3D, el centro de la esfera, desde múltiples puntos de vista. Es sin duda un buen método para adquirir rápidamente los datos de calibración para diferentes cámaras además de añadir nuevas restricciones a los parámetros de calibración teniendo en cuenta todo el conjunto de cámaras [Faugeras 01].

### **2.5.2. Calibración multicámara.**

La calibración de un sistema multicámara es mucho más complejo que la calibración de una única cámara, siendo necesario encontrar la relación geométrica existente entre los sistemas de referencia de todas las cámaras respecto de uno global. Dependiendo del tamaño del entorno, la forma de realizar la calibración es diferente. Así, para entornos pequeños en los que existe una zona que es captada simultáneamente por todas las cámaras, se utilizan las técnicas de calibración de una cámara. Trabajos en esta línea pueden consultarse en [Wiles 99][Svoboda 03][Ueshiba 03]

En el caso de trabajar con entornos de dimensiones grandes, cada una de las cámaras capta sólo una pequeña parte del mismo y en consecuencia no existe una zona común a todas las cámaras. En estos casos se hace necesario desplazar la plantilla de calibración, por todo el entorno, para calibrar de forma individual cada una de ellas. Dentro de esta línea está el trabajo desarrollado por Rander y Kanade [Rander 98] que ha sido utilizado para calibrar un sistema formado por 50 cámaras. Aunque el sistema proporcione buenos resultados, la precisión depende de la correcta ubicación de la plantilla de calibración, amén de tener un coste computacional alto. Estos sistemas no son aplicables en entornos de grandes dimensiones. Otro trabajo en esta línea se puede consultar en [Hoover01].

Una propuesta más ambiciosa es la realiza por Barreto y Daniilidis [Barreto 04] en la que proponen la calibración de un sistema formado por 52 cámaras, de las que previamente se han obtenido sus parámetros intrínsecos. Además, de cuatro de ellas se conocen sus parámetros extrínsecos respecto del sistema de referencia global. En un entorno sin iluminación, los

autores utilizan un LED que se mueve para generar los puntos de calibración de las 48 cámaras restantes. De los diferentes puntos generados, para calibrar una cámara sólo son utilizados aquellos que tienen correspondencia en, al menos, dos de las cámaras calibradas, para poder conocer su posición 3D. En [Chen 00] proponen la utilización de un objeto móvil como punto de referencia cuya posición 3D se obtiene utilizando un Filtro de Kalman extendido. Otro trabajo en la misma línea se pueden ver en [Baker 03]

## 2.6. Conclusiones.

En este capítulo se han mostrado diferentes alternativas de guiado de robots móviles utilizando información visual. Así mismo se han presentado diferentes modelos de cámaras y técnicas de calibración. En ambos casos se ha prestado especial atención a aquellas alternativas más próximas a los objetivos planteados en esta tesis. Esto es, sistemas que utilizan múltiples cámaras y que éstas estén ubicadas en el entorno de movimiento de los robots. En lo que a la ubicación de las cámaras se refiere, se ha realizado una breve revisión de la alternativa más ampliamente utilizada, que es la de embarcarlas a bordo de los robots (*on board*). Frente a esta alternativa, se está planteando en los últimos años la realización del guiado de robots a partir de cámaras ubicadas en posiciones fijas y conocidas del entorno (*off-board*). Esta segunda alternativa está dentro de la línea de los denominados "espacios inteligentes" y presenta diversas ventajas frente a la solución de las cámaras embarcadas a bordo. Entre estas ventajas se pueden citar la de que estos sistemas permiten tener una visión más amplia y completa del entorno por el que se desplazan los robots, reducen la complejidad de los mismos y facilitan la interacción hombre-robot. Esta nueva tendencia, constituye una novedosa línea de investigación en la que, aunque escasos, ya existen grupos de investigación en diversos países con resultados interesantes; los trabajos más destacados en esta línea han sido presentados dentro de este capítulo. Dentro de la alternativa de cámaras externas, se puede establecer una clasificación atendiendo a distribución de la "inteligencia". Así se puede hablar de sistemas distribuidos y centralizados. En los primeros, cada cámara dispone del hardware necesario para procesar las imágenes y conectarse con otras, con el robot o con el sistema de control general. En la segunda alternativa se disponen de una unidad de procesamiento central que es la encargada de procesar toda la información de las cámaras y generar las consignas de control a los robots.

Como conclusión general de la alternativa de guiado de robots utilizando cámaras externas, cabe decir que son escasos los trabajos que se han desarrollado hasta ahora en esta línea. El hecho de que se trate de un tema novedoso, hace que existan muchos aspectos por resolver.

Así por ejemplo, del estudio realizado se ha podido constatar que en ninguna de las soluciones propuestas se aborda la sincronización entre cámaras ni la corrección de errores debidos a la distorsión, aspectos éstos que tienen una gran incidencia en los errores de posicionamiento.

Otro problema asociado con la precisión de los sistemas multi-cámara se encuentra en el método de calibración propuesto. En este sentido, la gran mayoría de los trabajos analizados utilizan entornos de trabajos de pequeñas dimensiones, en los que la calibración se puede realizar utilizando los procesos tradicionales para una sola cámara. Sin embargo no existen trabajos en los que se planteen soluciones precisas para espacios de grandes dimensiones.





## Capítulo 3

---

### Errores debidos al sistema de visión

---

La precisión en la obtención de la posición de los robots móviles utilizando cámaras externas, no sólo viene impuesta por la excelencia de los algoritmos utilizados en la obtención de las coordenadas de los puntos significativos sobre el robot, sino también por la precisión en la obtención de las coordenadas en los planos imagen de esos puntos para cada una de las cámaras del sistema de visión.

En este capítulo se analizarán cuáles son las fuentes de errores existentes en el proceso de visión, analizando su influencia en la determinación de la posición del robot en cada una de las imágenes, así como las posibles soluciones que los minimicen.

#### 3.1. Modelo de cámara.

Las técnicas de posicionamiento de robots móviles requieren del conocimiento y modelado del sensor empleado. En el caso de los sistemas que utilizan visión artificial, el sensor es la cámara y el fenómeno a modelar es el proceso de formación de la imagen sobre el sensor CCD. Así, en la precisión obtenida en la determinación de la posición del robot influirá de forma

significativa el modelo de cámara utilizado. Cuantos más parámetros se utilicen en el modelo de cámara, mayor será la similitud con el dispositivo real, si bien ello supondrá una mayor carga computacional.

El modelo de proyección de perspectiva *pin-hole* es uno de los modelos más utilizados por su sencillez y por las buenas prestaciones que proporciona en un amplio campo de aplicaciones reales de formación de imágenes. Dicho modelo mejora si se añade el modelo de distorsión de las lentes. En la figura 3.1 se muestra este modelo proyectivo, donde  $O$  representa el centro óptico de la cámara,  $\Pi$  es el plano imagen, paralelo al plano formado por los ejes  $x$  e  $y$ , mientras que  $z$  representa el eje óptico.

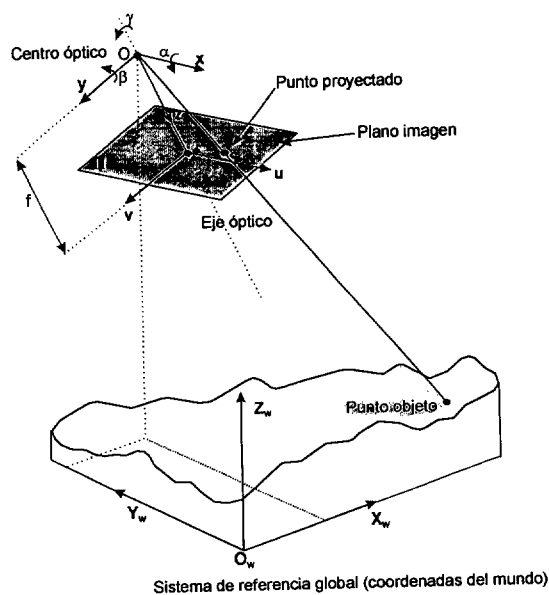


Figura 3.1. Modelo *pin-hole* de una cámara.

En el modelo *pin-hole* la relación entre un punto de la escena (punto objeto) de coordenadas  $[X', Y', Z']^T$  respecto del sistema de referencia global y su proyección en el plano imagen de coordenadas  $[x, y, z]^T$ , viene dada por la relación:

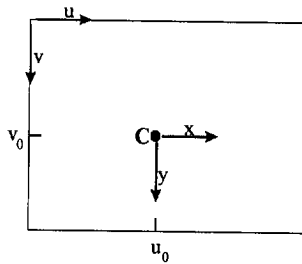
$$z \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot X' \\ f \cdot Y' \\ Z' \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} \quad (3.1)$$

Si las coordenadas de la imagen se expresan en píxeles, un punto de coordenadas  $(x, y)$  en el plano imagen se convierten en un píxel de coordenadas  $(u, v)$  según indica la expresión (3.2).

$$\begin{aligned} u &= \frac{1}{l_x} x \\ v &= \frac{1}{l_y} y \end{aligned} \tag{3.2}$$

donde  $l_x$  y  $l_y$  representan el tamaño horizontal y vertical, respectivamente, de los píxeles en el plano imagen. En la obtención de la ecuación (3.2) se ha supuesto que el origen del sistema de referencia  $(x, y)$  coincide con el de  $(u, v)$ , lo cual no tiene por qué suceder, como se muestra en la figura 3.2. Así, si  $u_0$  y  $v_0$  representan las coordenadas del punto principal del plano imagen (centro de formación de la imagen), la ecuación (3.2) se puede escribir como:

$$\begin{aligned} (u - u_0) &= \frac{x}{l_x} \\ (v - v_0) &= \frac{y}{l_y} \end{aligned} \tag{3.3}$$



**Figura 3.2.** Relación entre las coordenadas métricas  $(x,y)$  y píxelicas del plano imagen  $(u,v)$ .

Teniendo en cuenta (3.3), la ecuación (3.1) se puede expresar como (3.4).

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{l_x} f & 0 & u_0 & 0 \\ 0 & \frac{1}{l_y} f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \mathbf{K} \cdot \mathbf{P}' \tag{3.4}$$

Aquí  $\mathbf{K}$  es la matriz de proyección, y  $f_x, f_y, u_0$  y  $v_0$  son los llamados parámetros intrínsecos de la cámara.

Para la obtención de la ecuación (3.4) se parte del supuesto de que el sistema de referencia global coincide con el de la cámara. Por lo general, una cámara se ubica en una posición arbitraria, desplazada y girada respecto al sistema de referencia global, de ejes  $(X_w, Y_w, Z_w)$  y origen  $O_w$  (figura 3.1). Si un punto  $P'$  tiene de coordenadas  $[X', Y', Z']^T$  respecto del sistema de referencia de la cámara y  $[X, Y, Z]^T$  respecto al sistema de referencia global, la relación entre estas coordenadas será:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + \mathbf{T} \quad (3.5)$$

o bien:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \mathbf{R}^{-1} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \mathbf{R}^{-1} \cdot \mathbf{T} \quad (3.6)$$

De las ecuaciones (3.4) y (3.5) se obtiene:

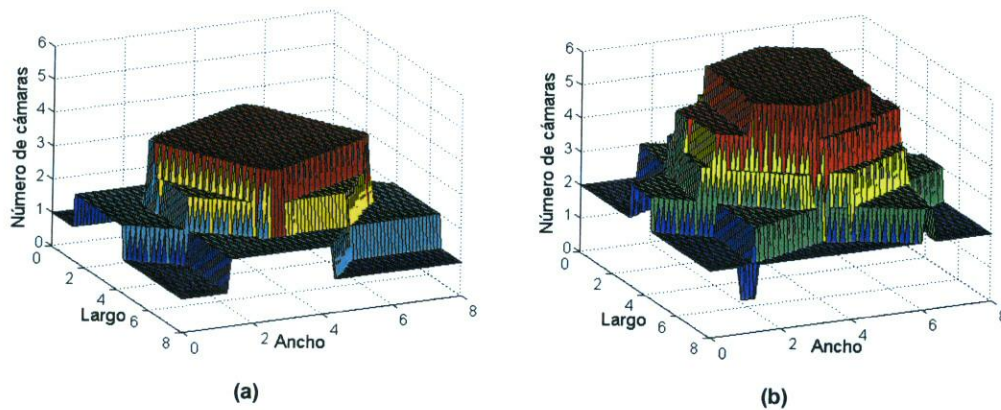
$$\lambda \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.7)$$

La ecuación (3.7) relaciona las coordenadas píxelicas  $(u, v)$  correspondientes a la proyección de un punto de coordenadas  $[X, Y, Z]^T$  respecto al sistema de coordenadas global. El vector  $\mathbf{T} = [T_x, T_y, T_z]^T$  representa la traslación entre los dos sistemas de referencia, mientras que  $\mathbf{R}$  es una matriz ortogonal de 3x3 que viene definida por los ángulos de Euler  $(\alpha, \beta, \gamma)$  (figura 3.1). Los parámetros  $T_x, T_y, T_z, \alpha, \beta$  y  $\gamma$  son los denominados parámetros extrínsecos de la cámara. Tanto el valor de los parámetros intrínsecos como el de los extrínsecos se determinan en el proceso de calibración. De forma compacta la ecuación (3.7) se puede escribir:

$$\lambda \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{M} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.8)$$

### 3.2. Percepción del entorno.

En su recorrido por el entorno, el robot es captado por un número diferente de cámaras, según su posición. Las cámaras se distribuyen por el entorno en posiciones estratégicas y con ángulos de giro,  $\alpha$  (*pan*) respecto al eje  $X_w$  y  $\beta$  (*tilt*) respecto al eje  $Y_w$ , tratando de garantizar que todos los puntos del entorno por el que navegan los robots sean vistos por, al menos, una cámara. Con estas premisas se han obtenido los gráficos de las figura 3.3.a y b, donde se representa el número de cámaras que captan al robot según su posición, para el caso de trabajar con cuatro cámaras (figura 3.3.a) ó seis cámaras (figura 3.3.b). En la obtención de las gráficas de la figura 3.3 se utilizaron cámaras con un tamaño de CCD de 2/3", dotadas de ópticas con distancia focal de 7.5 mm. En lo referente a los parámetros extrínsecos de las cámaras, respecto a un sistema centrado en  $[0, 0, 0]^T$ , se muestran sus valores en las tablas 3.1 y 3.2, para el caso de cuatro y seis cámaras, respectivamente.



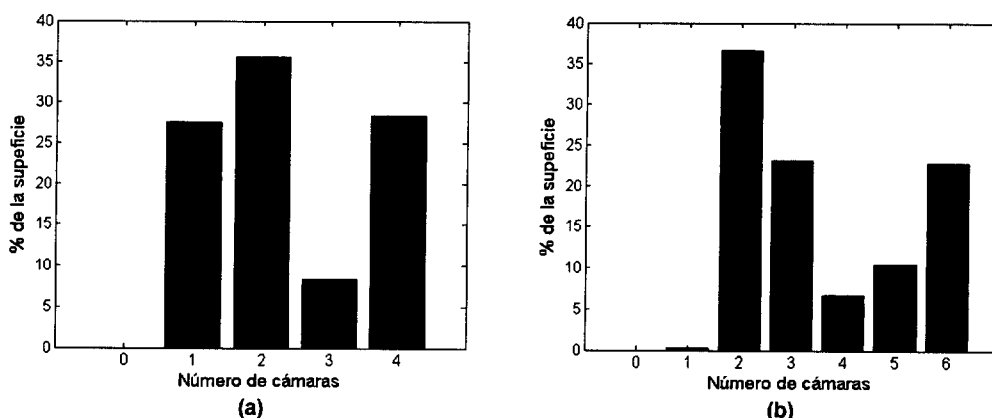
**Figura 3.3.** Superficie vista por diferente número de cámaras. El eje de ordenadas representa el número de cámaras, mientras que los otros dos representan las dimensiones en metros del entorno. (a) Utilizando un total de cuatro cámaras. (b) Utilizando un total de seis cámaras.

TABLA 3.1 Valores de los parámetros extrínsecos de las cámaras utilizadas para obtener la gráfica de la figura 3.3.a.						
Cámara	$T_x$ (m)	$T_y$ (m)	$T_z$ (m)	$\alpha$ (°)	$\beta$ (°)	$\gamma$ (°)
1	2	0	3	30	26	0
2	0	6	3	30	116	0
3	6	8	3	30	206	0
4	8	2	3	30	296	0

**TABLA 3.2** Valores de los parámetros extrínsecos de las cámaras utilizadas para obtener la gráfica de la figura 3.3.b.

Cámara	$T_x$ (m)	$T_y$ (m)	$T_z$ (m)	$\alpha$ (°)	$\beta$ (°)	$\kappa$ (°)
1	2	0	3	30	26	0
2	0	3.8	3	30	86	0
3	1.4	8	3	30	146	0
4	6	8	3	30	206	0
5	8	4.2	3	30	266	0
6	6.6	0	3	30	326	0

En las gráficas de la figura 3.4. se muestran los porcentajes de superficie del entorno que es captado por diferente número de cámaras. Como se puede apreciar, en el caso de utilizar cuatro cámaras, existe un 28% del entorno que es captado por una única cámara. Este porcentaje se reduce a un 0.2% para el caso de utilizar seis cámaras. Como posteriormente se demostrará, cuanto mayor sea el número de cámaras que se puedan utilizar para determinar la posición del robot, menor será el error cometido en su posicionamiento.



**Figura 3.4.** Porcentaje de la superficie del entorno de la figura 3.3 captado por un determinado número de cámaras. (a) Utilizando un total de cuatro cámaras. (b) Utilizando un total de seis cámaras.

Gráficas similares a las de la figura 3.4 se muestran en la figura 3.5, realizadas, en este caso, para diferentes tamaños del CCD y para el caso de utilizar seis cámaras. Como se puede observar, a medida que disminuye el tamaño del CCD aumenta la superficie del entorno que es percibido por una cámara o por ninguna. Evidentemente al disminuir el tamaño del CCD también lo hace el campo de visión de la cámara. Hay que hacer notar que en el caso de utilizar un CCD de 1/4", más de un 44% del entorno no es captado por ninguna cámara.

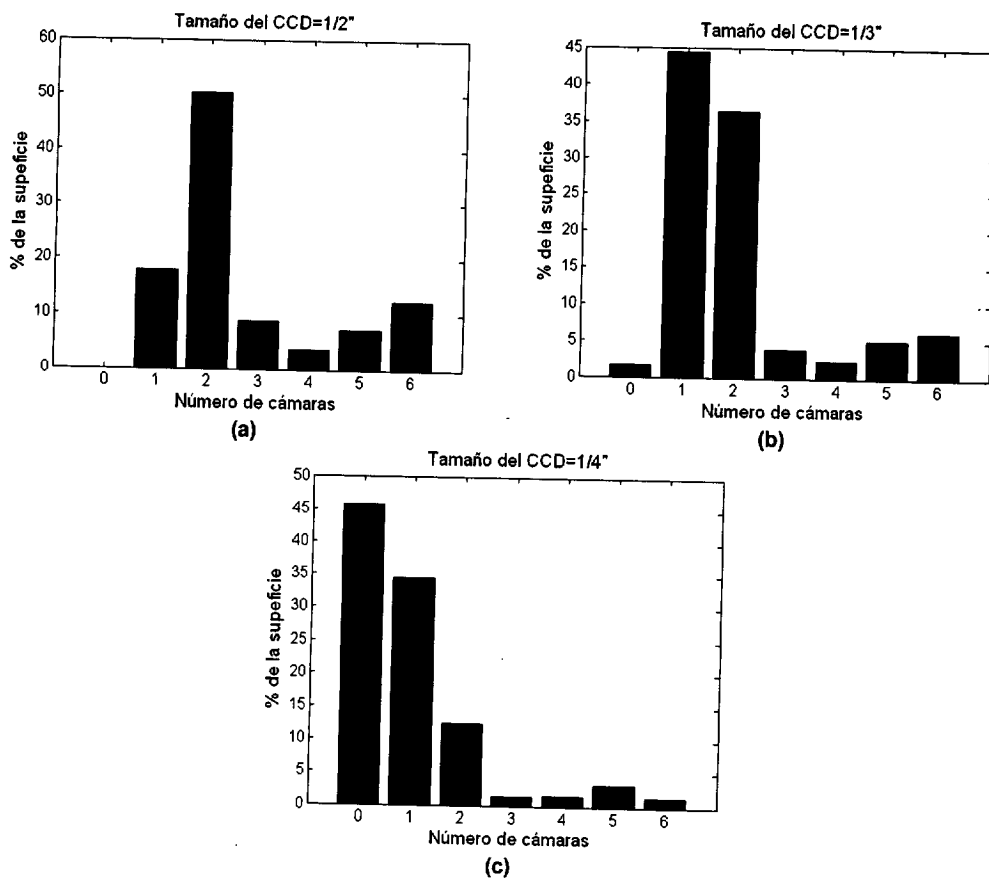
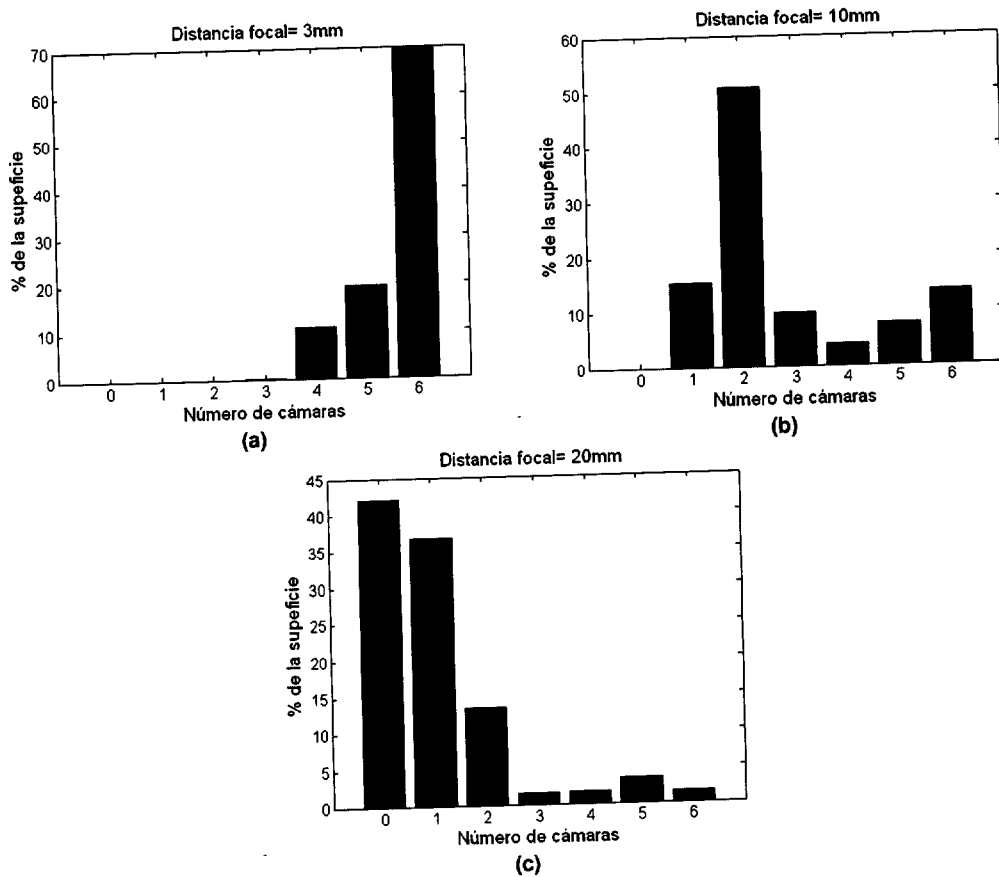


Figura 3.5. Porcentaje del entorno de la figura 3.3 captado, en función del número de cámaras (para el caso de utilizar un total de seis cámaras), para diferentes tamaños del CCD y una óptica de 7.5 mm. (a) CCD de 1/2". (b) CCD de 1/3". (c) CCD de 1/4".

Similar análisis al de la figura 3.5 se puede hacer teniendo en cuenta la distancia focal de la óptica montada en las cámaras, obteniéndose así las gráficas de la figura 3.6. En este caso, los mejores resultados, considerando coordenadas píxelicas, se obtienen con ópticas de distancias focales pequeñas que proporcionan campos de visión grandes. Ahora bien, estas ópticas presentan una elevada distorsión y aunque el sistema capta un campo mayor, su resolución empeora y por tanto también la precisión en la medida. En la figura 3.7 se muestra el tamaño de la escena abarcado por un píxel, cuando la separación entre ésta y la cámara están separados una distancia de 4 m (figura 3.7.a) y 8 m (figura 3.7.b), en función de la distancia focal de la óptica y el tamaño del CCD.





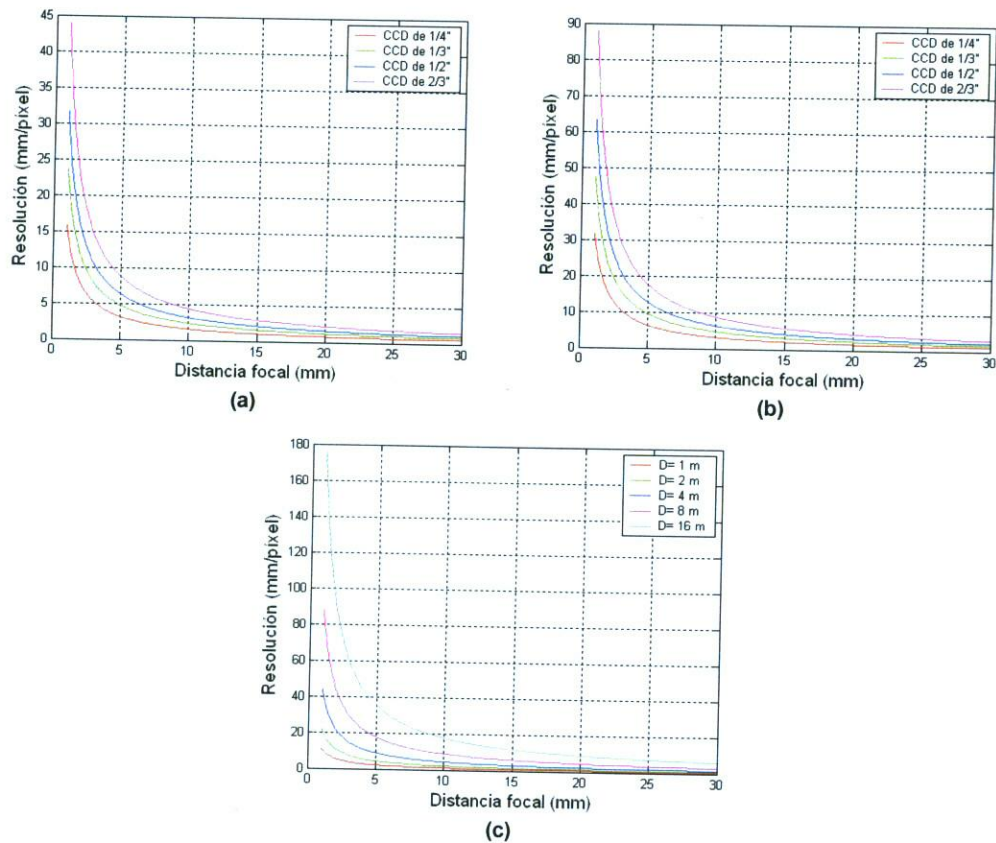
**Figura 3.6.** Porcentaje del entorno de la figura 3.3 captado, en función del número de cámaras (para un total de seis cámaras), para CCD's de 2/3" y ópticas de diferentes distancias focales. (a) Distancia focal de 3 mm. (b) Distancia focal de 10 mm. (c) Distancia focal de 20 mm.

Para el cálculo de la resolución se ha considerado que se podrá detectar la posición del centroide de un objeto con precisión de un píxel, por lo que el error máximo cometido debido a la cuantificación de la imagen en la cámara será de  $\pm 0.5$  píxeles. Nótese, que el campo correspondiente a un píxel aumenta a medida que nos vamos alejando de la cámara y en consecuencia, también lo hará el error en unidades métricas correspondiente a un píxel (figura 3.7.c). En esta tesis, para realizar las pruebas prácticas, se han utilizado cámaras con un CCD de 2/3" y ópticas de 7.5 mm.

### 3.3. Identificación del robot.

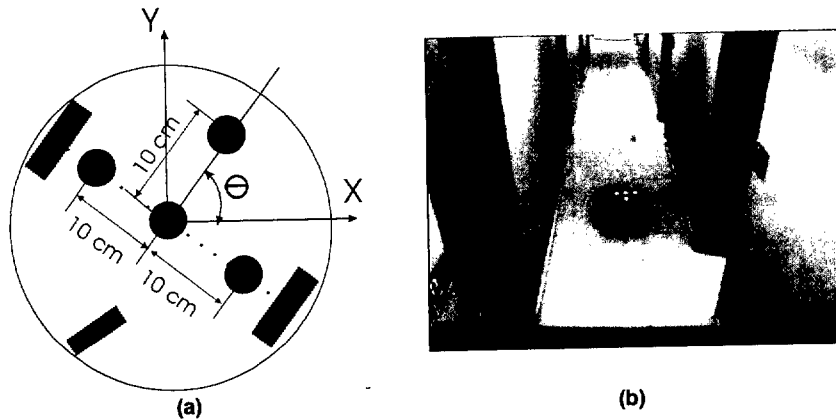
Con el fin de facilitar la identificación del robot en las imágenes, se ha ubicado a bordo del mismo (figura 3.8.b) un array de emisores de infrarrojos (IR) (balizas de infrarrojos) con una distribución como la mostrada en la figura 3.8.a. En esta figura se pueden apreciar cuatro

balizas denominadas *A*, *B*, *C* y *D*, todas ellas del mismo tamaño. La baliza *C* se utiliza para obtener la posición del robot, su posición se identificará con la posición de éste. Con la disposición física de las balizas mostrada en la figura 3.8.a se puede obtener también la dirección u orientación del robot. En los apartados y capítulos siguientes, cuando se haga referencia a la posición del robot se estará haciendo referencia a la posición del centro de la baliza *C*, hablando indistintamente de posición del robot o de la baliza.

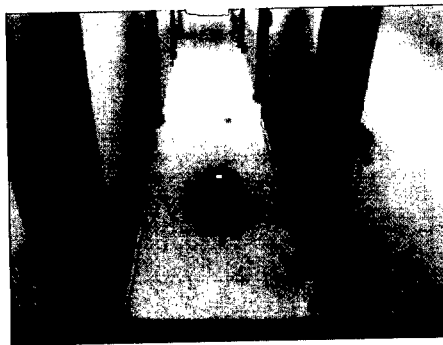


**Figura 3.7.** Tamaño de la escena correspondiente a un píxel en función de la distancia focal y tamaño del CCD de la cámara. (a) Para puntos separados de la cámara 4 m. (b) Para puntos separados de la cámara 8m. (c) En función de la distancia para un CCD de 2/3".

Si bien la configuración de balizas de infrarrojos propuesta (*A*, *B*, *C*, *D*) facilita la obtención precisa de la orientación del robot, a medida que éste se aleja de las cámaras la separación entre las balizas en el plano imagen va disminuyendo, llegando a solaparse. Esta situación se produce para distancias próximas a los 7 m. En este caso, al no poder identificar las cuatro balizas no se puede obtener la orientación del robot y su posición no será precisa (ya que no se podrá identificar la baliza *C*). En esta tesis también se han realizados pruebas utilizando una única baliza de IR (figura 3.9) del mismo tamaño que las anteriores, detectando la baliza hasta posiciones alejadas unos 15 m de la cámara.



**Figura 3.8.** Baliza de IR utilizada a bordo del robot que facilita la identificación de su posición y orientación. (a) Distribución de los emisores de infrarrojos (*A, B, C, D*). (b) Ejemplo de detección de las balizas a bordo de un prototipo de robot.



**Figura 3.9.** Ejemplo de detección de la posición del robot utilizando una única baliza de IR a bordo.

En el capítulo cinco se abordan aspectos relacionados con el diseño práctico de la estructura de balizas (número de balizas, diámetro de cada una de ellas, longitud de onda del infrarrojo, disposición geométrica, etc.)

### 3.4. Errores debidos al sistema de visión.

Una vez identificado el robot y obtenidas las coordenadas en el plano imagen de los centroides de las balizas de a bordo, se puede calcular la posición 3D (coordenadas  $[X, Y, Z]^T$ ) y orientación del robot. Evidentemente uno de los factores que influye en la precisión con la que se obtienen las coordenadas en el plano imagen de los centroides de las balizas, son los errores

propios del sistema de visión. Hay que tener presente que las imágenes de los robots se captan con éstos en movimiento. Los errores más importantes debidos al sistema de visión son:

- Tiempo de exposición (*shutter*) de las cámaras.
- Desplazamientos temporales debidos a la falta de sincronismo entre las cámaras y aspectos asociados al hardware de adquisición.
- Modelo de cámara utilizado.
- Error en el proceso de calibración de las cámaras.
- Obtención de las coordenadas pixélicas de las balizas en el plano imagen.

Cada una de estas fuentes de error provoca una variación de las coordenadas pixélicas ( $u, v$ ) de los centroides de las proyecciones de las balizas en cada una de las cámaras. Teniendo en cuenta que estos errores están incorrelados, el error final será igual a la suma de todos ellos. A continuación se analiza la influencia de cada uno de estos errores.

### 3.4.1. Errores debidos al tiempo de exposición de las cámaras.

Además de los ajustes de enfoque e iris, la mayoría de las cámaras tienen un ajuste de la velocidad del obturador (*shutter*). Para el caso de cámaras CCD, esta velocidad representa el tiempo que el CCD recoge señal luminosa. A este tiempo se le conoce con el nombre de tiempo de exposición de la cámara. Dado que el robot se está moviendo, el tiempo de exposición provocará una deformación de las balizas en el plano imagen. Con tiempos de exposiciones bajos, por ejemplo 1/1000 s, cualquier movimiento puede ser "congelado" sin pérdida de nitidez o rasgos de movimiento (figura 3.10). Sin embargo, la disminución del tiempo de exposición provoca una pérdida de iluminación en la imagen, razón por la cual estos tiempos sólo pueden ser pequeños cuando los niveles de iluminación de la escena son altos. Para nuestro objetivo, un tiempo de exposición grande significaría no poder determinar la posición de las balizas en la imagen, observando únicamente una estela de luz.

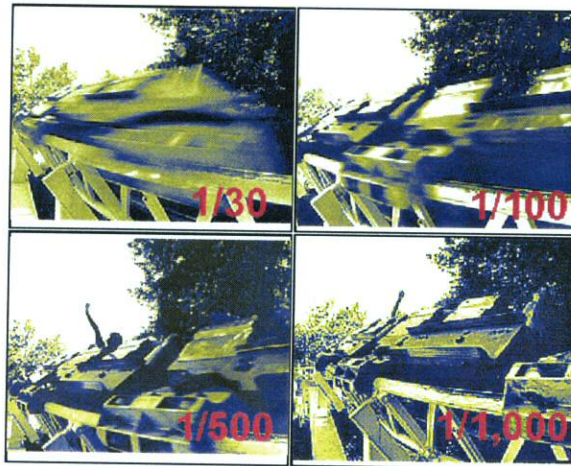
Si un punto  $P$  del espacio 3D se mueve con una velocidad  $V$ , partiendo de una posición  $[X_p, Y_p, Z_p]^T$ , transcurrido un tiempo igual al de exposición de la cámara ( $t_{exp}$ ) las coordenadas 3D de ese mismo punto  $[X_{pe}, Y_{pe}, Z_{pe}]^T$  serán:

$$\begin{bmatrix} X_{pe} \\ Y_{pe} \\ Z_{pe} \end{bmatrix} = \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} \quad (3.9)$$

donde:

$$\begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} = \mathbf{V} \cdot t_{\text{exp}} = t_{\text{exp}} \cdot \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} \quad (3.10)$$

siendo  $V_x$ ,  $V_y$  y  $V_z$  las componentes del vector velocidad.



**Figura 3.10.** Imágenes de un objeto en movimiento tomadas con cámaras con diferentes tiempos de exposición.

A partir de (3.10) y (3.4) se pueden obtener las diferencias entre las coordenadas en el plano imagen  $(u, v)$  supuesto el punto  $P$  con velocidad nula y las correspondientes  $(u_d, v_d)$  a la posición que ocupa ese mismo punto tras el tiempo de exposición, supuesto que este se mueve con una velocidad  $\mathbf{V}$  en el espacio 3D. Estas diferencias de coordenadas en el plano imagen vienen dadas por las ecuaciones (3.11), siendo  $[X', Y', Z']^T$  las coordenadas del punto  $P$  referidas al sistema de coordenadas de la cámara. Si en lugar de utilizar la ecuación (3.4) se utiliza la (3.8) se obtienen las ecuaciones (3.12).

$$\begin{aligned} u_d - u &= f_x \left( \frac{(X' + \Delta X')}{Z' + \Delta Z'} - \frac{X'}{Z'} \right) \\ v_d - v &= f_y \left( \frac{(Y' + \Delta Y')}{Z' + \Delta Z'} - \frac{Y'}{Z'} \right) \end{aligned} \quad (3.11)$$

$$\begin{aligned}
u_d - u = f_x \cdot \frac{r_{11} \cdot (X + \Delta X) + r_{12} \cdot (Y + \Delta Y) + r_{13} \cdot (Z + \Delta Z) + T_x}{r_{31} \cdot (X + \Delta X) + r_{32} \cdot (Y + \Delta Y) + r_{33} \cdot (Z + \Delta Z) + T_z} \\
- f_x \cdot \frac{r_{11} \cdot X + r_{12} \cdot Y + r_{13} \cdot Z + T_x}{r_{31} \cdot X + r_{32} \cdot Y + r_{33} \cdot Z + T_z}
\end{aligned}
\tag{3.12}$$

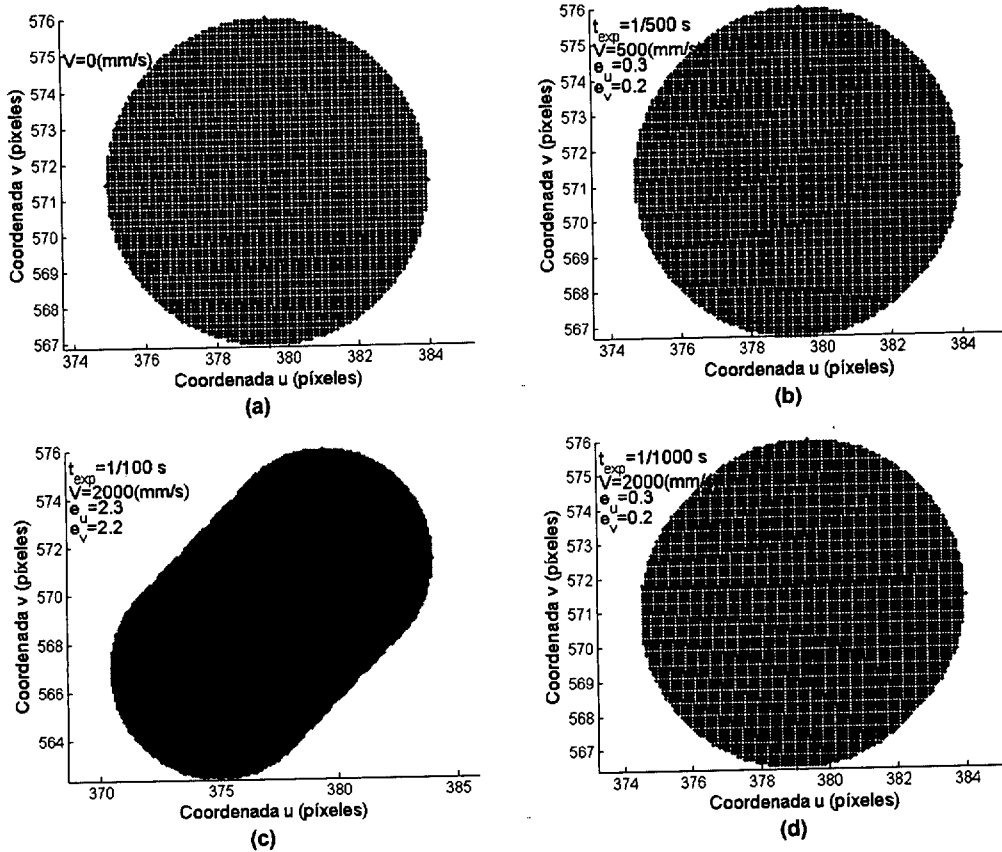
$$\begin{aligned}
v_d - v = f_y \cdot \frac{r_{21} \cdot (X + \Delta X) + r_{22} \cdot (Y + \Delta Y) + r_{23} \cdot (Z + \Delta Z) + T_y}{r_{31} \cdot (X + \Delta X) + r_{32} \cdot (Y + \Delta Y) + r_{33} \cdot (Z + \Delta Z) + T_z} \\
- f_y \cdot \frac{r_{21} \cdot X + r_{22} \cdot Y + r_{23} \cdot Z + T_y}{r_{31} \cdot X + r_{32} \cdot Y + r_{33} \cdot Z + T_z}
\end{aligned}$$

Evidentemente las diferencias de coordenadas en el plano imagen dadas por la ecuación (3.11) y (3.12) son función de las componentes de velocidad en el espacio 3D y del tiempo de exposición. A modo de ejemplo, en la figura 3.11 se muestran las proyecciones en el plano imagen de una baliza de radio 20 mm (similar a las utilizadas en las pruebas prácticas) ubicada en las coordenadas  $X_p=4$  m,  $Y_p=1.2$  m,  $Z_p=0.2$  m, con la cámara ubicada en:  $X=4$  m,  $Y=0$  m,  $Z=3$  m,  $\alpha=45^\circ$ ,  $\beta=\gamma=0$ , y para distintos tiempos de exposición y velocidades de la baliza. Como se puede observar, debido al movimiento de la baliza y al tiempo de exposición de las cámaras, la imagen obtenida de la baliza se verá deformada, siendo esta deformación tanto mayor cuanto mayor sean su velocidad y el tiempo de exposición de la cámara.

La deformación en la proyección de la baliza, sobre todo cuando ésta es grande, puede introducir errores apreciables en el cálculo de la posición y orientación del robot. Evidentemente el error que se comete en la determinación de las coordenadas en el plano imagen debido al tiempo de exposición y velocidad del robot, es la diferencia entre las coordenadas de los centroides de las proyecciones de las balizas con y sin movimiento. Identificando por  $e_u$  y  $e_v$  estas diferencias de coordenadas, se tiene:

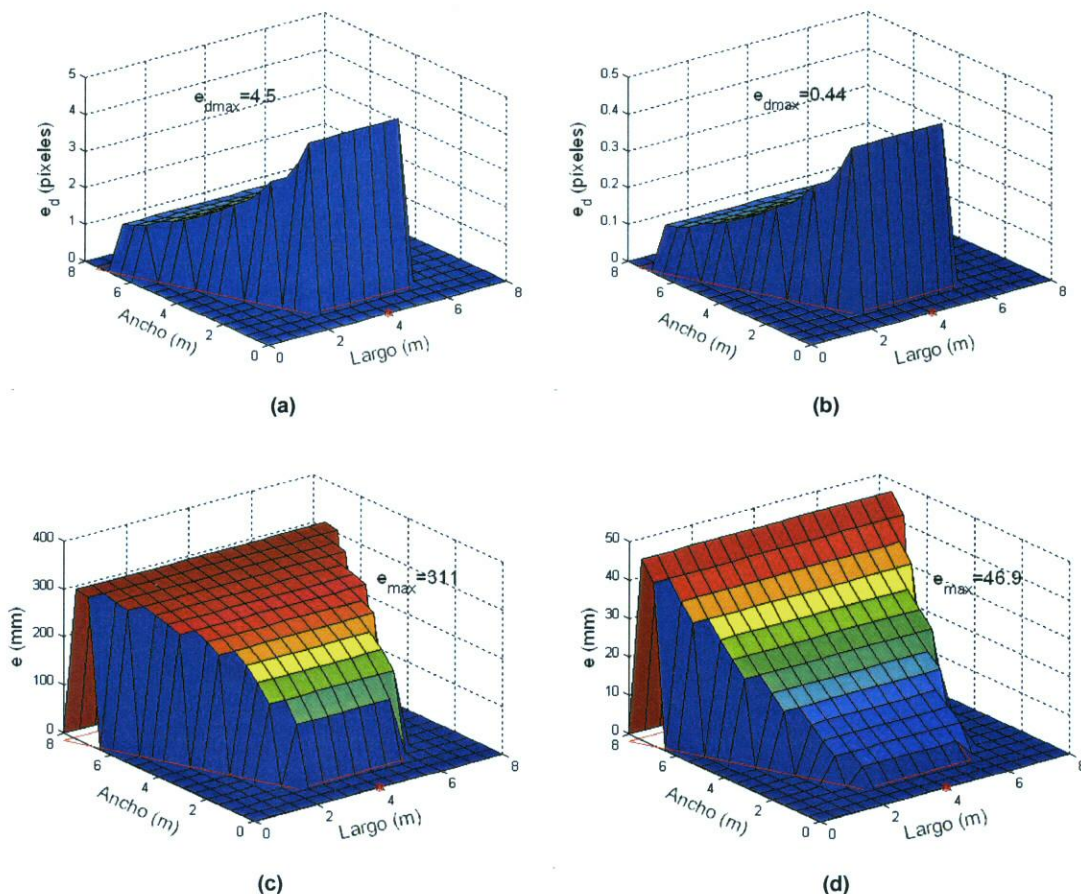
$$\begin{aligned}
e_u &= |u_V - u_{V=0}| \\
e_v &= |v_V - v_{V=0}|
\end{aligned}
\tag{3.13}$$

donde  $u_{V=0}$  y  $v_{V=0}$  representan las coordenadas del centroide de la proyección de la baliza con velocidad nula, mientras que  $u_V$  y  $v_V$  representan ese mismo centroide cuando la baliza se mueve con una velocidad  $V$ .



**Figura 3.11.** Deformación de la proyección sobre el plano imagen de una baliza circular de radio 20mm, como consecuencia del tiempo de exposición de la cámara y movimiento de la baliza.. (a)  $\|V\|=0$ . (b)  $\|V\|=500$  mm/s y  $t_{exp} = 1/100$  s. (c)  $\|V\|=2000$  mm/s y  $t_{exp}=1/100$  s. (d)  $\|V\|=2000$ mm/s y  $t_{exp}=1/1000$  s.

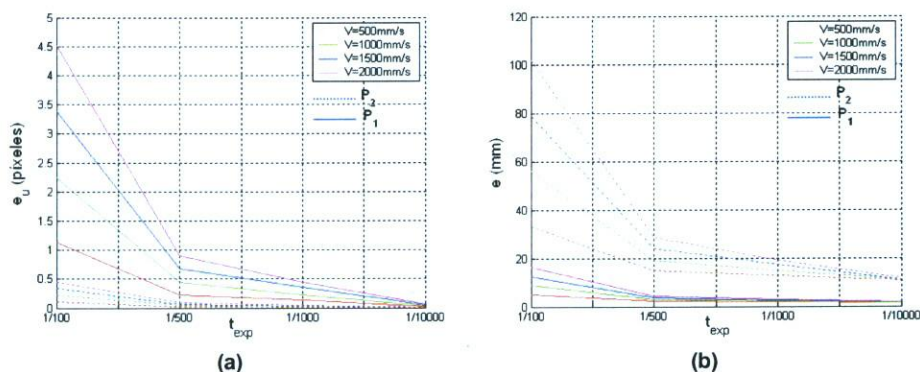
Para realizar una estimación del error que se comete en la determinación de las coordenadas píxelicas, debido al movimiento de la baliza y al tiempo de exposición de la cámara, se han realizado pruebas para una velocidad de la baliza con componentes  $V_x=V_y=2$  m/s,  $V_z=0$ , dos tiempos de exposición 1/100 s. (figura 3.12.a) y 1/1000 s. (figura 3.12.b), con la cámara ubicada en la misma posición que para la obtención de la figura 3.11, y para ubicaciones de la baliza dentro del campo de visión de la cámara. Indicar, que aunque el error en píxeles disminuye a medida que nos alejamos de la cámara, el error en unidades métricas, teniendo en cuenta la resolución, aumenta. Este efecto puede observarse en las figura 3.12.c y d, para el caso de errores en píxeles como los mostrados en las figura 3.12.a y b, respectivamente.



**Figura 3.12.** Error en la determinación de las coordenadas pixélicas de una baliza que se mueve a una velocidad con componentes  $V_x=V_y=2\text{m/s}$  y  $V_z=0$ , para dos tiempos de exposición de la cámara. (a) Con un tiempo de exposición de  $1/100\text{ s}$ . (b) Con un tiempo de exposición de  $1/1000\text{ s}$ . (c) Error en unidades métricas, teniendo en cuenta la resolución, para la figura (a). (d) Error en unidades métricas, teniendo en cuenta la resolución, para la figura (b).

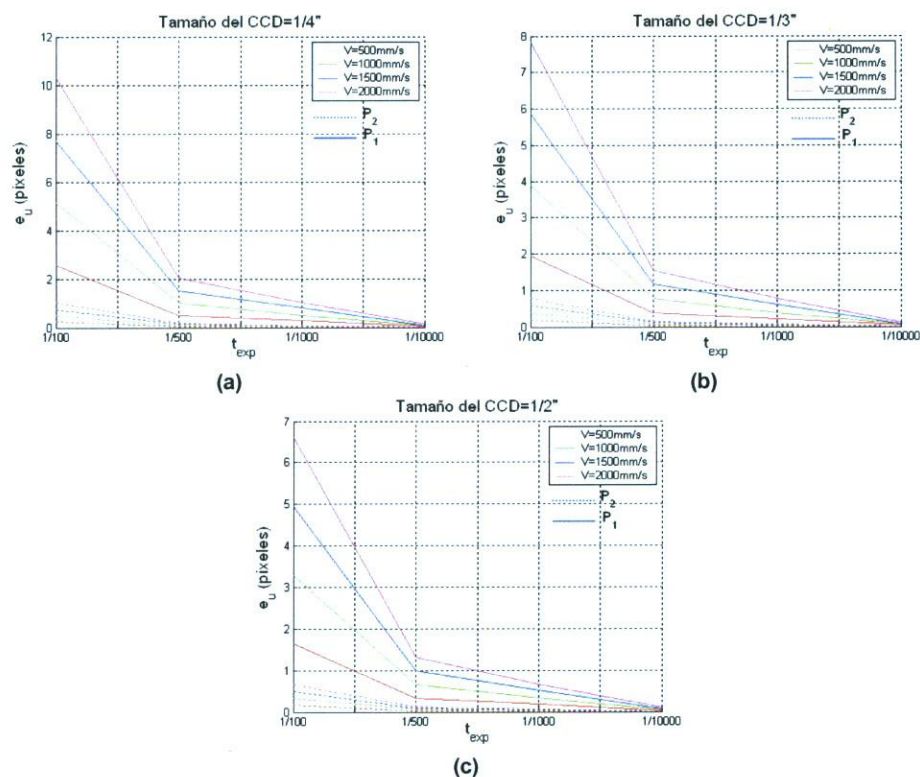
En las figura 3.13.b se muestra el error en las coordenadas métricas, debidas al tiempo de exposición y velocidad del robot, para diferentes valores de estos dos parámetros (suponiendo la cámara ubicada en la misma posición que para la figura 3.11) y considerando la baliza situada en las posiciones  $P_1$  y  $P_2$ , situados a un distancia de  $2.2\text{m}$  ( $P_1$ ) y  $7.3\text{m}$  ( $P_2$ ) de la cámara, respectivamente. Hay que indicar que, al no ser cuadrados los píxeles en el sensor CCD, la variación del error en la coordenada  $u$  es siempre mayor que en la coordenada  $v$ . Examinando la figura 3.13.a se puede observar cómo para tiempos de exposición inferiores a  $1/1000\text{ s}$ , el error de las coordenadas del centroide de la baliza es inferior a  $0.5$  píxeles, para las diferentes velocidades y posiciones de la baliza analizadas.





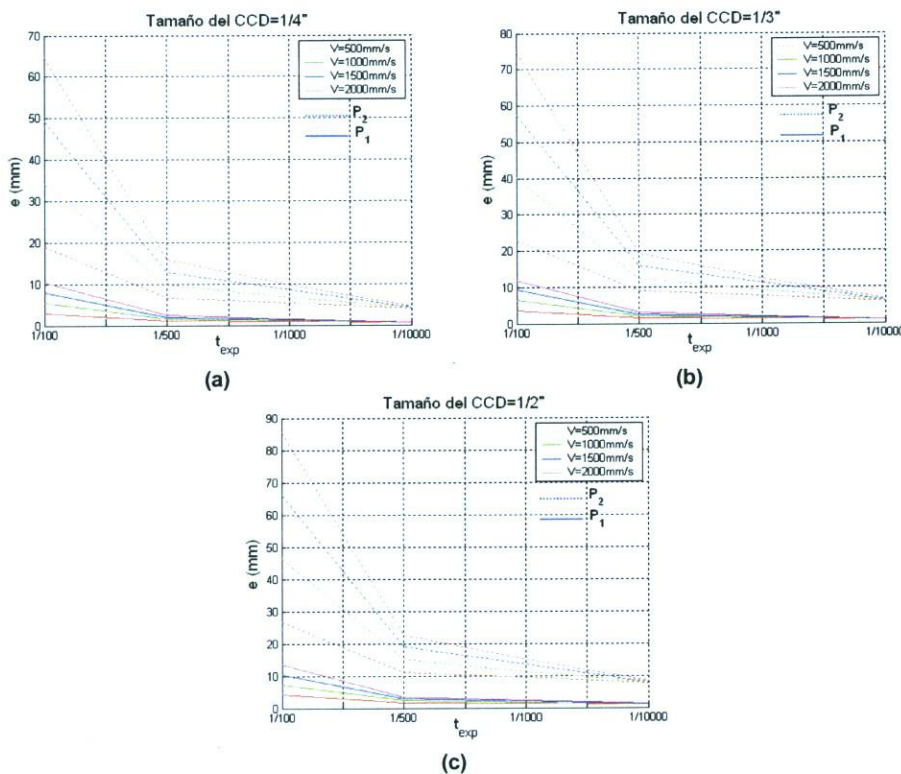
**Figura 3.13.** Errores de las coordenadas del centroide de la baliza en movimiento en función del tiempo de exposición de la cámara, para dos posiciones ( $P_1$  y  $P_2$ ) de la baliza. (a) Coordenadas pixélicas. (b) Coordenadas métricas.

Las gráficas de la figura 3.13 se han obtenido para una cámara con un tamaño de CCD de 2/3" con una óptica de 7.5 mm de focal. En la figura 3.14 se muestra el error de la coordenada  $u$  del centroide de la baliza para diferentes tamaños de CCD.



**Figura 3.14.** Error (en píxeles) de la coordenada  $u$  del centroide de la baliza con ésta en movimiento, para diferentes tiempos de exposición de la cámara, posiciones de la baliza y tamaño de sensor CCD. (a) CCD con tamaño 1/4". (b) CCD con tamaño 1/3". (c) CCD con tamaño 1/2".

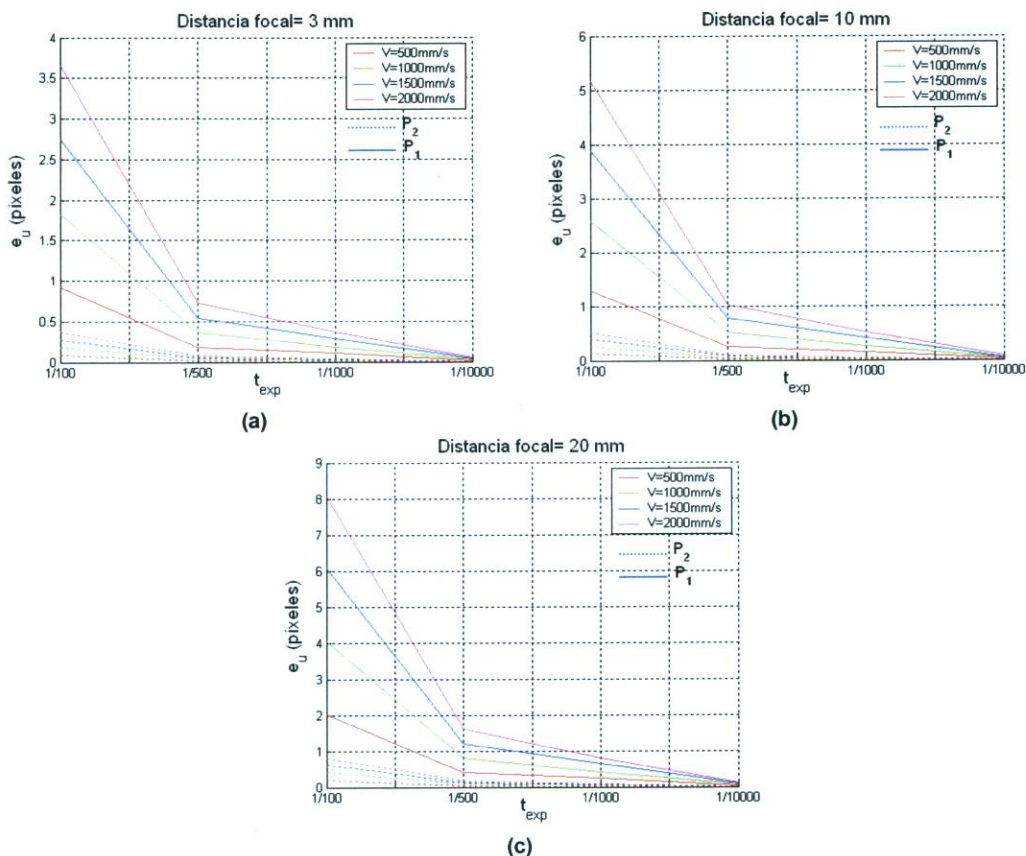
En la figura 3.14 el error en píxeles aumenta a medida que disminuye el tamaño del CCD. Ahora bien, como ya se indicó anteriormente, si se tiene en cuenta la resolución de la cámara, el error (en coordenadas métricas) empeora a medida que el objeto se aleja de la misma, y en consecuencia el resultado global indica que el error aumenta con la distancia. Este error en coordenadas métricas se muestra en la figura 3.15.



**Figura 3.15.** Error en unidades métricas correspondientes a los errores pixélicos de las figuras 3.14. (a) Figura 3.14.a (b) Figura 3.14.b (c) Figura 3.14.c.

Otro parámetro que interviene en el proceso de formación de imágenes, es la distancia focal ( $f$ ) de la óptica. En la figura 3.16 se ha representado la influencia de la distancia focal en los errores debidos al tiempo de exposición y movimiento del robot. Únicamente se muestra cómo varía el error en la coordenada  $u$  del centroide de la baliza para diferentes distancias focales y velocidades del robot (utilizando una cámara con un CCD de 2/3").

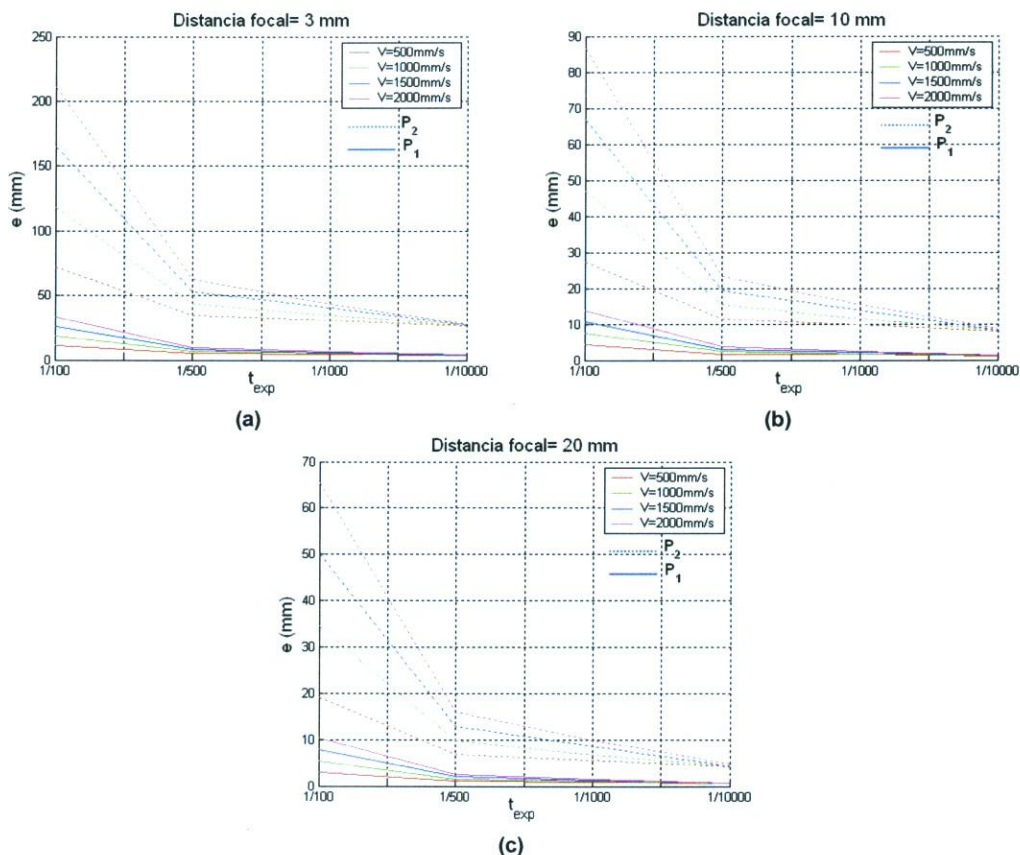
Realizando sobre las gráficas de la figura 3.16 el mismo razonamiento que el empleado en la figura 3.15 para obtener el error en coordenadas métricas debido a la posición de la baliza, para diferentes distancias focales, se obtiene la figura 3.17. Como se puede observar, este error, para una misma posición, aumenta a medida que disminuye la distancia focal de la óptica.



**Figura 3.16.** Error (en píxeles) de la coordenada  $u$  del centroide de la baliza en función del tiempo de exposición de la cámara, para diferentes distancias focales, posiciones de la baliza y velocidades del robot. (a) Distancia focal de 3 mm. (b) Distancia focal de 10 mm. (c) Distancia focal de 20 mm.

### 3.4.2. Desplazamientos temporales debidos a la falta de sincronismo entre las cámaras y aspectos asociados al hardware de adquisición.

Dado que en nuestro sistema se utiliza la información procedente de diferentes cámaras para obtener la posición (centroide de la baliza de a bordo) del robot en movimiento, es importante conocer los instantes de tiempo en que captan las imágenes las diferentes cámaras. Evidentemente, si las cámaras no captan en el mismo instante la posición del robot, se cometerá un error al obtener las coordenadas 3D. Por ello, en este punto se van a analizar, por una parte, las causas que hacen que diferentes cámaras no capturen al mismo tiempo la información del entorno, y, por otra, los errores a que esto puede dar lugar.



**Figura 3.17.** Error en mm para las gráficas de la figura 3.16, teniendo en cuenta la resolución de la cámara. (a) Figura 3.16.a. (b) Figura 3.16.b. (c) Figura 3.16.c.

Las causas que pueden hacer que un sistema multi-cámaras no capten simultáneamente, son dos: a) falta de un sincronismo común a todas ellas, b) hardware de adquisición con exploración secuencial de las cámaras. Como se pondrá de manifiesto más adelante, es necesario evitar estas dos fuentes de errores (utilizando una señal de sincronismo común y un hardware que permita la adquisición simultánea de varias cámaras) ya que de lo contrario se cometería un error apreciable en la determinación de las coordenadas 3D del robot, sobre todo cuando éste se mueva a velocidades altas.

En primer lugar, se abordan los errores de posicionamiento debidos a la falta de sincronismo entre cámaras. Como es sabido, un sensor CCD transforma la energía de la luz incidente en energía eléctrica. La carga eléctrica almacenada en cada una de sus celdas es posteriormente transferida, utilizando un registro de desplazamiento, para conformar una señal de vídeo estándar, utilizando para ello señales de sincronismo (horizontal y vertical). En la mayoría de casos, es la propia cámara la que genera dichas señales de sincronismo. En un sistema de visión multi-cámara es importante la perfecta sincronización entre todas ellas, de forma que

todas captan el entorno en el mismo instante de tiempo.

Si identificamos por  $t_i$  la diferencia de tiempo de captación entre una cámara de referencia (que podemos identificar por  $C_0$ ) y otra cámara  $C_i$ , y suponemos un robot desplazándose a una velocidad  $V$  ( $[V_x, V_y, V_z]^T$ ), es evidente que si la cámara de referencia capta al robot en una posición  $P_0 = [X_0, Y_0, Z_0]^T$ , la cámara  $C_i$  lo captará en otra posición  $P_i = [X_i, Y_i, Z_i]^T$ , cumpliéndose:

$$P_i = \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} X_0 + t_i \cdot V_x \\ Y_0 + t_i \cdot V_y \\ Z_0 + t_i \cdot V_z \end{bmatrix} \quad (3.14)$$

Por tanto, si la obtención de la posición del robot se realiza a partir de varias cámaras, y éstas no están perfectamente sincronizadas se cometerá un error en la determinación de su posición. Este error será tanto mayor cuanto mayor sea el desplazamiento temporal existente entre las cámaras.

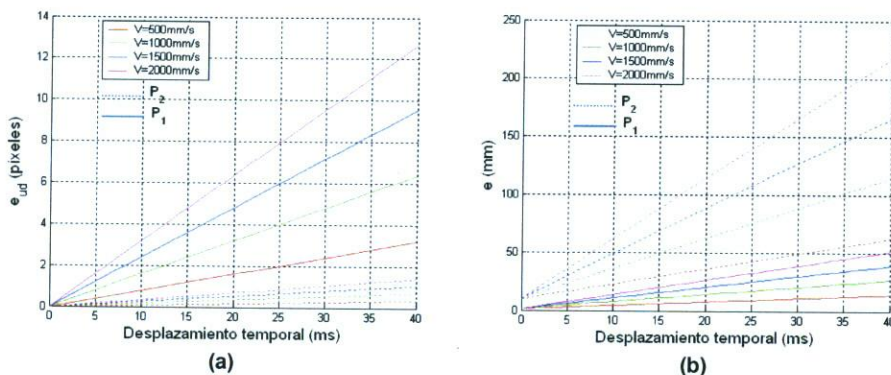
Para evaluar el error que se produce en las coordenadas del plano imagen de la cámara  $C_i$ , con respecto a la de referencia (debido a la falta de sincronismo), basta con sustituir en la ecuación (3.8) los valores de  $P_0$  y  $P_i$ . Así, identificando por  $(u, v)$  las coordenadas de la proyección de  $P_0$ , por  $(u_i, v_i)$  las de  $P_i$ , y por  $e_{ud}$  y  $e_{vd}$  los errores en las coordenadas  $u$  y  $v$  debidas al desplazamiento temporal entre la cámara de referencia y la  $C_i$ , se pueden escribir:

$$e_{ud} = |u - u_i| = \left| \frac{m_{11} \cdot X_0 + m_{12} \cdot Y_0 + m_{13} \cdot Z_0 + m_{14}}{m_{31} \cdot X_0 + m_{32} \cdot Y_0 + m_{33} \cdot Z_0 + m_{34}} - \frac{m_{11} \cdot X_i + m_{12} \cdot Y_i + m_{13} \cdot Z_i + m_{14}}{m_{31} \cdot X_i + m_{32} \cdot Y_i + m_{33} \cdot Z_i + m_{34}} \right| \quad (3.15)$$

$$e_{vd} = |v - v_i| = \left| \frac{m_{21} \cdot X_0 + m_{22} \cdot Y_0 + m_{23} \cdot Z_0 + m_{24}}{m_{31} \cdot X_0 + m_{32} \cdot Y_0 + m_{33} \cdot Z_0 + m_{34}} - \frac{m_{21} \cdot X_i + m_{22} \cdot Y_i + m_{23} \cdot Z_i + m_{24}}{m_{31} \cdot X_i + m_{32} \cdot Y_i + m_{33} \cdot Z_i + m_{34}} \right|$$

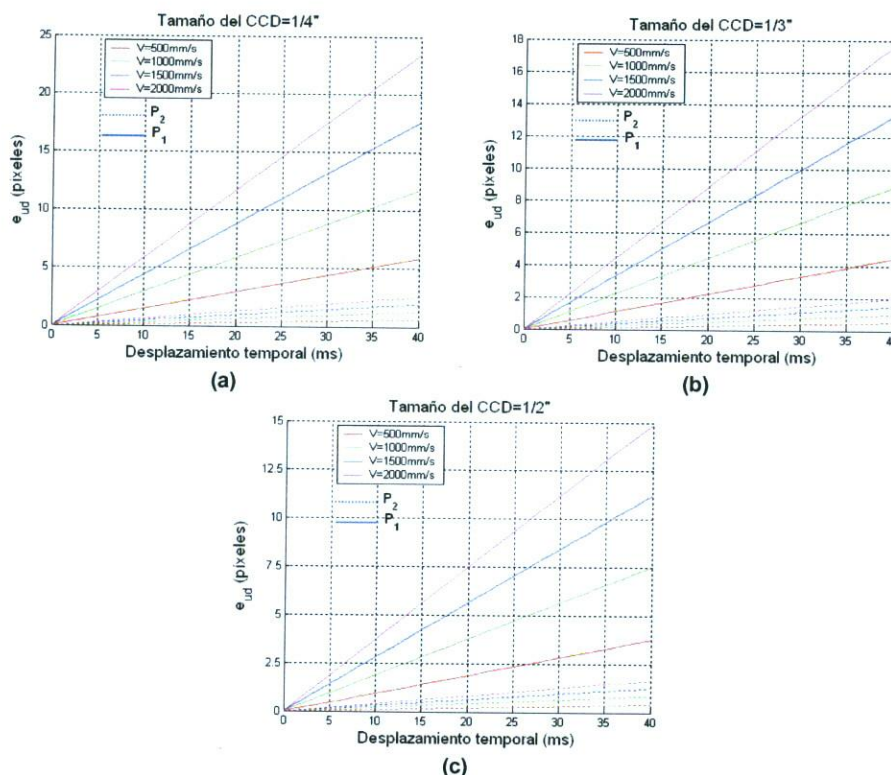
Al igual que sucedía con el tiempo de exposición de las cámaras, la magnitud del error de la ecuación (3.15) es función de la posición del robot respecto a las cámaras.

En la figura 3.18.a y b se muestra el error máximo de la proyección de una baliza circular en función del desplazamiento de tiempo existente entre dos cámaras. En este caso se ha supuesto que la baliza se mueve sobre el plano  $Z=Z_0$ , para diferentes valores de velocidad pero con componentes de velocidad horizontal y vertical iguales ( $V=V_x=V_y$ ) y situada en las posiciones  $P_1$  y  $P_2$ .

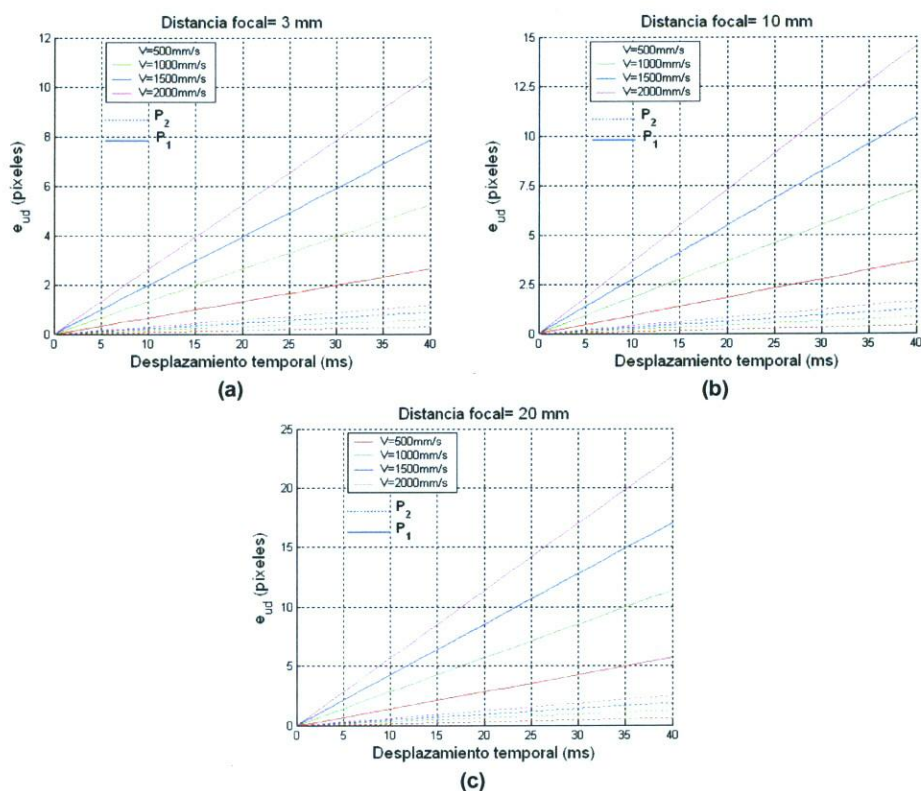


**Figura 3.18.** Error en las coordenadas del centroide de la baliza entre dos cámaras, para diferentes valores de velocidad del robot, posición de la baliza y para diferentes valores de desplazamiento temporal entre las mismas. (a) Error en coordenadas pixélicas. (b) Error en coordenadas métricas

Las gráficas de la figura 3.18 se han obtenido para una cámara con un CCD de 2/3" y una óptica de 7.5 mm. En la figura 3.19 y 3.20 se representa el error, en coordenadas pixélicas, debido al desplazamiento temporal para cámaras con diferente tamaños de CCD y distancia focal.



**Figura 3.19.** Error en las coordenadas pixélicas del centroide de la baliza entre dos cámaras, para diferentes tamaños de CCD, diferentes valores de velocidad del robot y diferentes valores de desplazamiento temporal entre las cámaras. (a) CCD con un tamaño de 1/4". (b) CCD con un tamaño de 1/3". (c) CCD con un tamaño de 1/2".



**Figura 3.20.** Error en las coordenadas píxelicas del centroide de la baliza entre dos cámaras, para ópticas de diferentes distancias focales, diferentes velocidades del robot y diferentes valores de desplazamiento temporal entre las cámaras.. (a) Distancia focal 3mm. (b) Distancia focal 10 mm. (c) Distancia focal 20 mm.

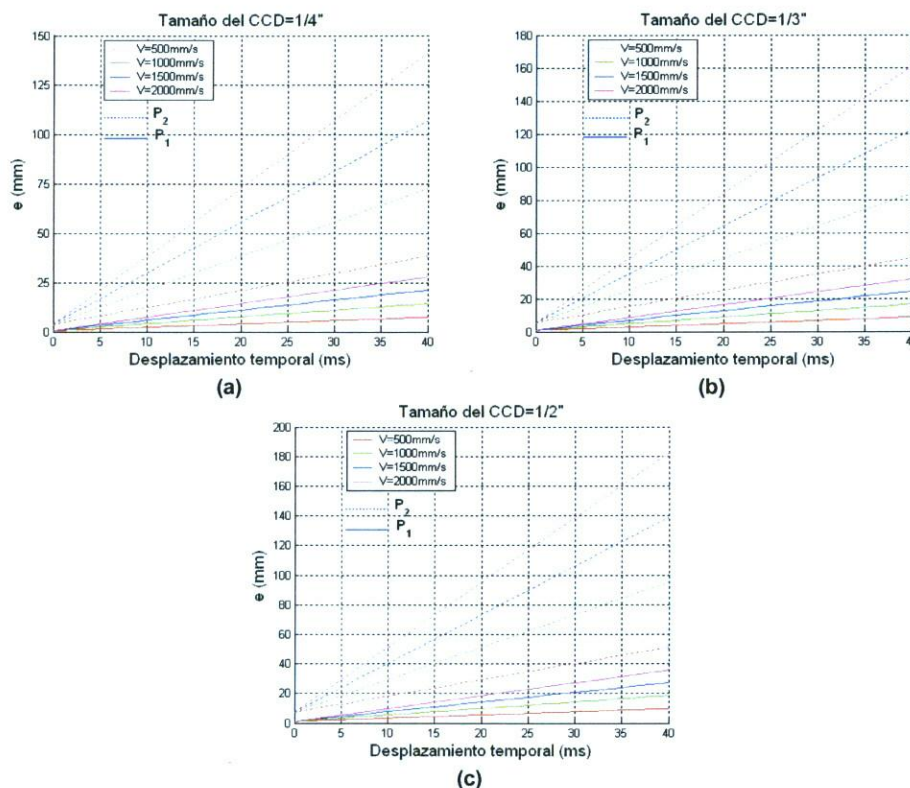
Si se tiene en cuenta la resolución de la cámara, los errores en píxeles de las figura 3.18, 3.19 y 3.20 se traducen a unidades métricas, tal y cómo se muestra en las figura 3.21 y 3.22. En esta última, se puede comprobar cómo este error, al igual que sucedía con el error debido al tiempo de exposición de la cámara, aumenta a medida que disminuye la distancia focal de la óptica.

Finalmente, para analizar el efecto que tiene la exploración secuencial de las diferentes cámaras vamos a suponer un sistema formado por  $n$  cámaras, un tiempo de adquisición de cada cámara que identificaremos por  $t_a$ , y una secuencia de adquisición  $C_1, C_2, \dots, C_n$ . Dado que sólo después de una secuencia completa de adquisición se analizan las imágenes, la diferencia de tiempo entre dos cámaras  $C_i, C_j$  ( $i < j < n$ ) se puede escribir:

$$t_{ij} = (j - i) \cdot t_a \tag{3.16}$$

Llevando este tiempo a la ecuación (3.14) se puede obtener el desplazamiento de las coordenadas del centroide de la proyección de la baliza en una cámara debido a la

digitalización secuencial de las señales provenientes de cada una de ellas. En la figura 3.23. se muestran los valores de estos errores para un sistema con  $n=16$  y para  $t_a = 40$ ms.



**Figura 3.21.** Errores en mm para las gráficas de las figura 3.19. (a) Figura 3.19.a. (b) Figura 3.19.b. (c) Figura 3.19.c.

Como ya se comentó, los errores debidos a la digitalización secuencial son prohibitivos y por ello es necesario utilizar en la práctica el hardware que garantice que todas las cámaras que “ven” una misma zona del entorno capten de forma simultánea. Una posible solución, cuando se trabaja con cámaras monocromas, consiste en la utilización tarjetas digitalizadoras que permitan trabajar en modo RGB y sincronizarse con otras. Esto nos permite trabajar con tres cámaras simultáneamente. A modo de ejemplo, utilizando esta estrategia, en la figura 3.24 se puede observar la reducción de este error.

Evidentemente este tipo de error se puede anular utilizando un sistema de visión distribuido, en el que cada cámara disponga de un hardware adicional que permite digitalizar y procesar sus imágenes, con sincronización entre los diferentes subsistemas.



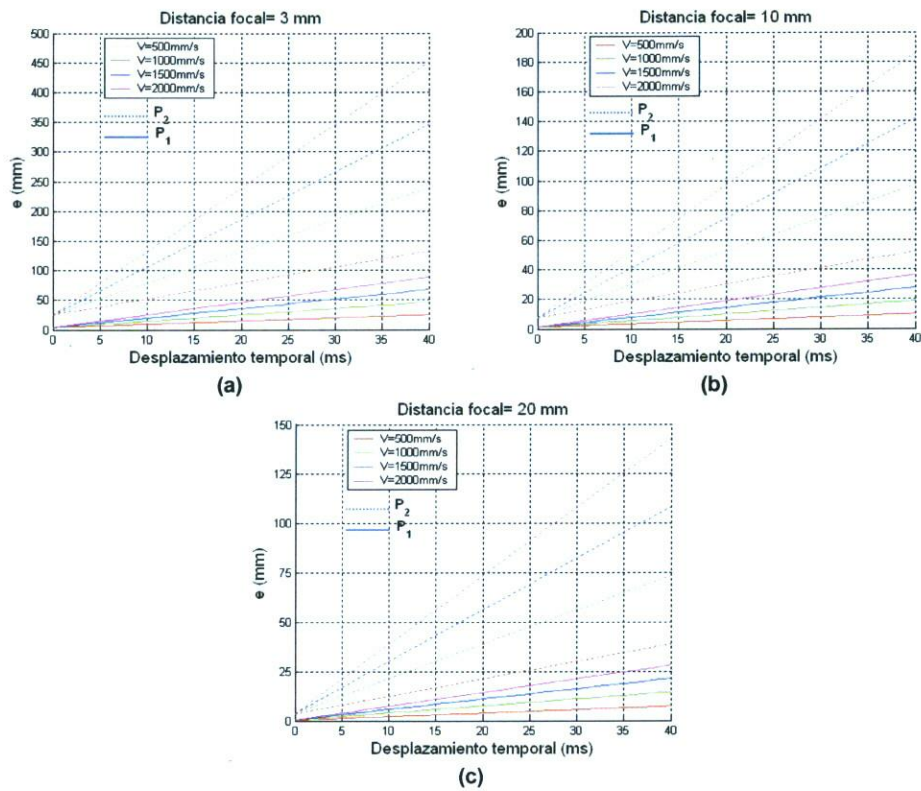


Figura 3.22. Errores en coordenadas métricas para las gráficas de la figura 3.20. (a) Figura 3.20.a. (b) Figura 3.20.b. (c) Figura 3.20.c.

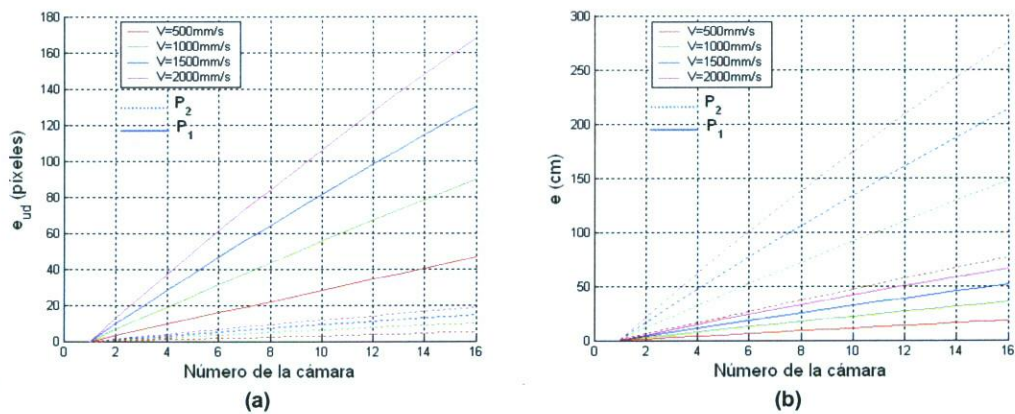
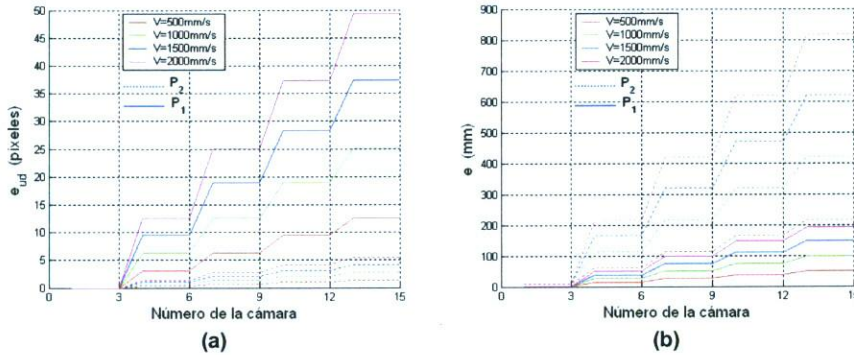


Figura 3.23. Errores en las coordenadas píxelicas del centroide de la baliza debidos a la digitalización secuencial para un sistema formado por 16 cámaras y con tiempos de adquisición de cada una de ellas,  $t_a = 40ms$ , para distintas posiciones de la baliza. (a) Error en coordenadas píxelicas. (b) Error en coordenadas métricas.



**Figura 3.24.** Error de las coordenadas pixélicas del centroide de una baliza para un sistema de 15 cámaras monocromo utilizando tarjetas digitalizadoras con entrada RGB. (a) Error en coordenadas pixélicas. (b) Error en coordenadas métricas.

### 3.4.3. Errores debidos al modelo de cámara.

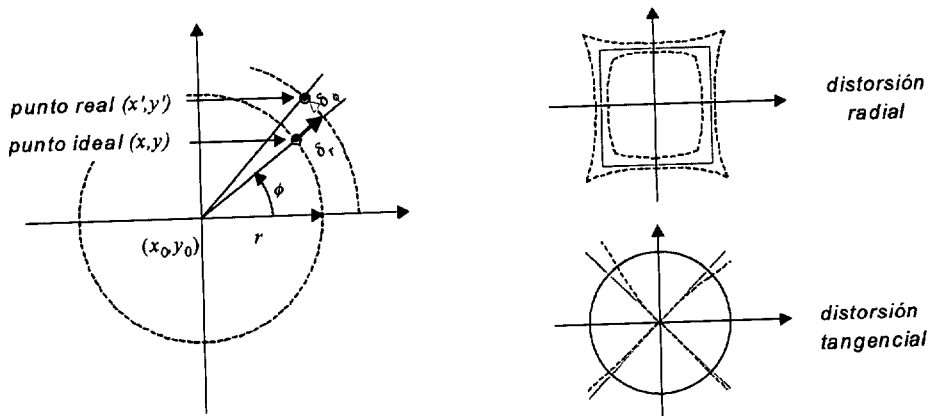
En todos los estudios de errores presentados hasta ahora se ha utilizado el modelo de cámara ideal *pin-hole*. Sin embargo, en la práctica las lentes introducen deformaciones (distorsión). Dicha distorsión se produce cuando los rayos de luz que pasan a través de la óptica son desviados, modificando sus direcciones e interceptando, en el plano imagen, en posiciones alejadas de lo que sucedería en el modelo ideal. Este desplazamiento (distorsión) es poco significativo en la zona central de la imagen, pero su efecto es importante en los bordes de la misma. Generalmente, la distorsión se modela mediante dos componentes, una radial ( $\delta_r$ ) y otra tangencial ( $\delta_\phi$ ) (figura 3.25). La componente radial de la lente hace que los puntos de la imagen se distorsionen a lo largo de líneas radiales a partir del punto principal de la misma ( $x_\phi, y_\phi$ ), mientras que la componente tangencial se produce en líneas perpendiculares a las radiales a partir del punto principal.

La distorsión introducida por la óptica hace que un punto  $P$  de la escena, que idealmente se debería proyectar en el punto  $p$  de coordenadas  $(x,y)$  del plano imagen, se proyecte en otro punto  $p'$  de coordenadas  $(x',y')$ . La relación entre las coordenadas de ambos puntos (a las que ya se hizo referencia en el capítulo 2) son, según la teoría de la fotogrametría, las que se muestran a continuación.

$$x' = x + x(k_1r^2 + k_2r^4) + 2p_1xy + p_2(r^2 + 2x^2) \tag{3.17}$$

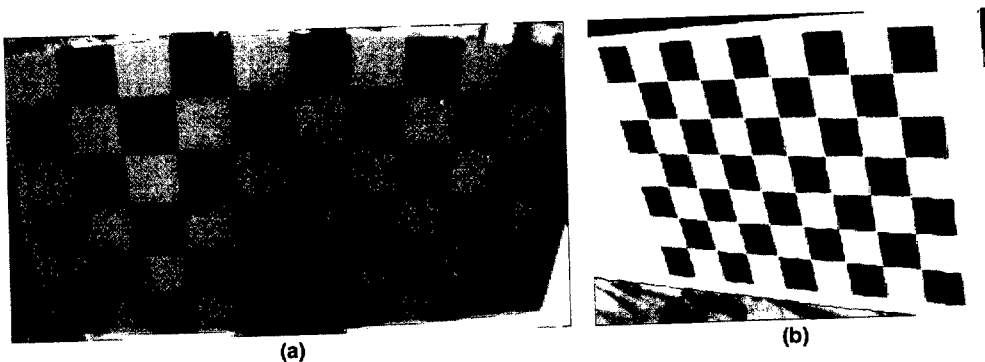
$$y' = y + y(k_1r^2 + k_2r^4) + p_1(r^2 + 2y^2) + 2p_2xy$$

$$p' = p + \delta(p) \tag{3.18}$$



**Figura 3.25.** Efecto de las componentes de distorsión introducida por la óptica: componentes radial y tangencial.

En este caso se han limitado a dos los parámetros que modelan los efectos de la distorsión radial, y también a dos los que modelan la tangencial. Toda óptica distorsiona la imagen, en mayor o menor medida, pero su grado de distorsión será tanto mayor cuanto mayor lo sean los coeficientes  $k_1$ ,  $k_2$ ,  $p_1$  y  $p_2$ . En la figura 3.26 se muestran dos imágenes con distorsión. Dichas imágenes han sido captadas utilizando dos ópticas diferentes, una con coeficientes de distorsión :  $k_1 = -0.3444$   $k_2 = 0.1518$ ,  $p_1 = -0.0001$  y  $p_2 = -0.0025$ , para la imagen de la figura 3.26.a, y otra con :  $k_1 = -0.1564$   $k_2 = 0.1717$ ,  $p_1 = -0.0037$  y  $p_2 = 0.0057$ , para la de la figura 3.26.b. Es evidente que la distorsión en la figura 3.26.a es más significativa que en la figura 3.26.b, lo cual era de esperar a la vista de los valores de los coeficientes de distorsión. Los parámetros de distorsión de las ópticas utilizadas en las pruebas prácticas de esta tesis tienen valores similares a los de la óptica utilizada para captar la imagen de la figura 3.26.b.



**Figura 3.26.** Efecto sobre las imágenes de la distorsión de la óptica de la cámara. (a) Utilizando una óptica con una fuerte distorsión. (b) Utilizando una óptica con baja distorsión.

Un píxel de la imagen con coordenadas  $(u, v)$ , según (3.4) tendrá unas coordenadas métricas  $(x, y)$  dadas por:

$$\begin{aligned}x &= \frac{u - u_0}{f_x} \\y &= \frac{v - v_0}{f_y}\end{aligned}\tag{3.19}$$

pero si se tiene en cuenta el efecto de la distorsión (ecuación (3.17)) se convierten en las coordenadas  $(x', y')$  dadas por:

$$\begin{aligned}x' &= \frac{u' - u_0}{f_x} \\y' &= \frac{v' - v_0}{f_y}\end{aligned}\tag{3.20}$$

Los errores en las componentes horizontal y vertical de las coordenadas pixélicas debidos a la distorsión serán:

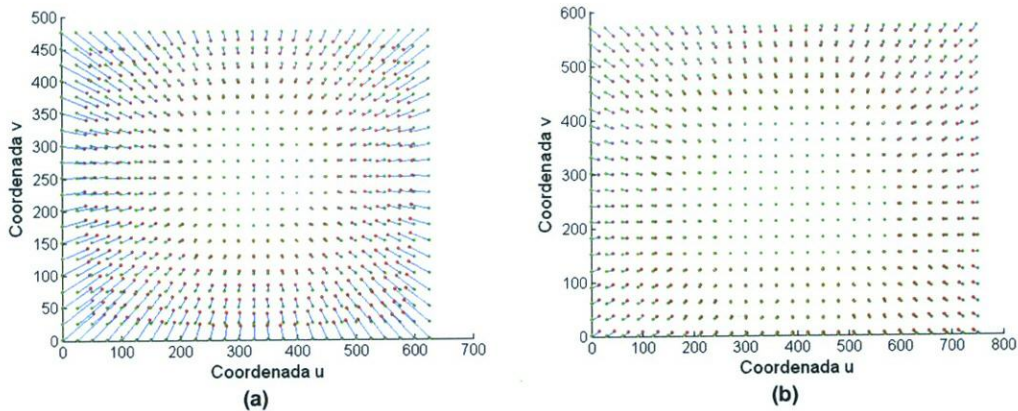
$$\begin{aligned}e_u &= |u - u'| \\e_v &= |v - v'|\end{aligned}\tag{3.21}$$

En la figura 3.27 se puede observar como los puntos de una imagen utilizando un modelo de cámara sin distorsión (puntos reflejados en color verde) se desplazan a otras posiciones debido al efecto de la distorsión (puntos indicados en color rojo). Como se puede observar, el desplazamiento es poco significativo en la parte central del plano imagen, aumentando de forma apreciable en los bordes.

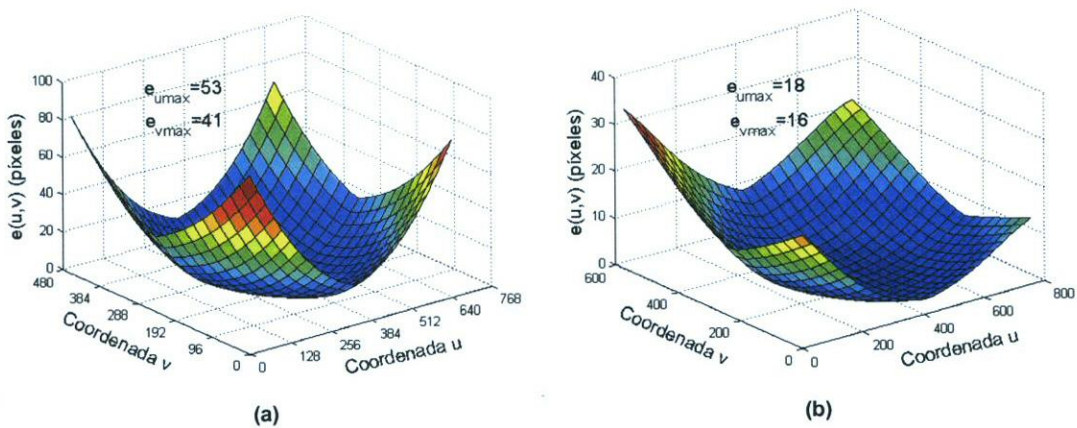
En la figura 3.28 se muestran los errores de las coordenadas pixélicas debidos a los efectos de distorsión de la óptica utilizada. Como queda de manifiesto, los valores máximos de estos errores (desplazamiento) son considerables, lo que obliga en la práctica a buscar alternativas para reducirlos (corrección de la distorsión de la óptica).

El proceso de corrección consiste en obtener las coordenadas  $(x, y)$  del punto  $p$  sin distorsión a partir de las coordenadas  $(x', y')$  del punto distorsionado  $p'$ . La utilización de la ecuación (3.17) para corregir la distorsión no es inmediata, ya que se trata de una ecuación no lineal. Evidentemente lo que se busca con el proceso de corrección de distorsión es permitir el uso del modelo *pin-hole* como modelo de proyección, sin tener en cuenta las no linealidades del sistema óptico. En este sentido son varias las propuestas [Heikkilä 00][Melen 94][Tamaki 02]

que se han realizado para corregir la distorsión de las ópticas. En esta tesis se utiliza la propuesta de Melen [Melen 94] por los excelentes resultados proporcionados y el bajo coste computacional que requiere.



**Figura 3.27.** Representación del desplazamiento sufrido por las coordenadas píxelicas de una imagen debido a la distorsión, en verde se representa su posición ideal y en rojo las reales teniendo en cuenta la distorsión (la línea azul refleja el emparejamiento entre ellos). (a) Óptica con alta distorsión (figura 3.26.a). (b) Óptica con baja distorsión (figura 3.26.b).



**Figura 3.28.** Valor del desplazamiento en las coordenadas  $(u,v)$  de los píxeles de una imagen debido a la distorsión introducida por la óptica. (a) Para una óptica similar a la utilizada en la captación de la imagen de la figura 3.26.a. (b) Para una óptica similar a la utilizada en la captación de la imagen de la figura 3.26.b.

Así, despejando  $p$  de la ecuación (3.18), se obtiene:

$$p = p' - \delta(p) \tag{3.22}$$

Dado que la magnitud del desplazamiento al que se ven sometidos los píxeles debido a la distorsión depende del valor del vector  $\delta(p) = p' - p$ , si en el término  $\delta(p)$  de la expresión (3.22)

se sustituye el valor de  $p$  dada por la propia ecuación se obtiene el proceso iterativo descrito en (3.23):

$$\begin{aligned}
 p &= p' - \delta(p) \\
 &= p' - \delta(p' - \delta(p)) \\
 &= p' - \delta(p' - \delta(p' - \delta(p))) \\
 &= p' - \delta(p' - \delta(p' - \delta(p' - \delta(p)))) \\
 &= \dots
 \end{aligned}
 \tag{3.23}$$

Según (3.23), el proceso a seguir para obtener las coordenadas corregidas, que identificaremos por  $(u_c, v_c)$ , del píxel distorsionado  $(u', v')$  es el mostrado en la figura 3.29.

```

x'=(u'-u0)/fx
y'=(v'-v0)/fy
xi=x'
yi=y'
for j=1:1:k
  r = (xi2 + yi2)
  cr = 1 + k1 · r + k2 · r2
  dx=2 · p1 · xi · yi + p2 · (r + 2 · xi2)
  dy=p1 · (r + 2 · yi2) + 2 · p2 · xi · yi
  dxr=(x' - dx)/cr
  dyr=(y' - dy)/cr
  xi=dxr
  yi=dyr
end
x=xi
y=yi
uc=fx · x + u0
vc=fy · y + v0

```

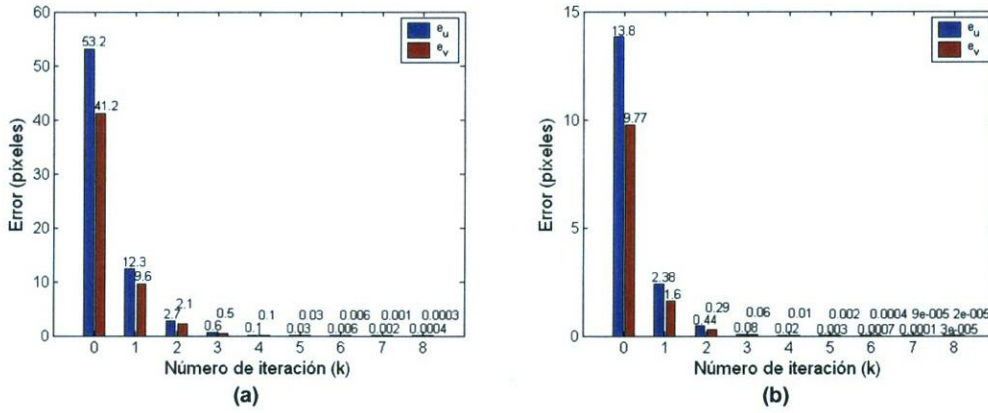
**Figura 3.29.** Proceso iterativo dado por la ecuación (3.23) empleado para corregir la distorsión de un píxel.

Ahora, el valor de la desviación que sufren los píxeles debido a la distorsión vendrá dada por:

$$\begin{aligned}
 e_u &= |u - u_c| \\
 e_v &= |v - v_c|
 \end{aligned}
 \tag{3.24}$$

La figura 3.30 refleja los valores máximos (figura 3.30.a) y medios (figura 3.30.b) del desplazamiento que sufren las coordenadas píxelicas de una imagen debido a la distorsión de

la óptica, en función del número de iteraciones ( $k$ ) del proceso de la figura 3.29. Los valores para  $k=0$  se corresponden con los puntos con distorsión, sin realizar ninguna corrección. Los valores representados en la figura 3.30 corresponden a los puntos en que la distorsión es máxima.



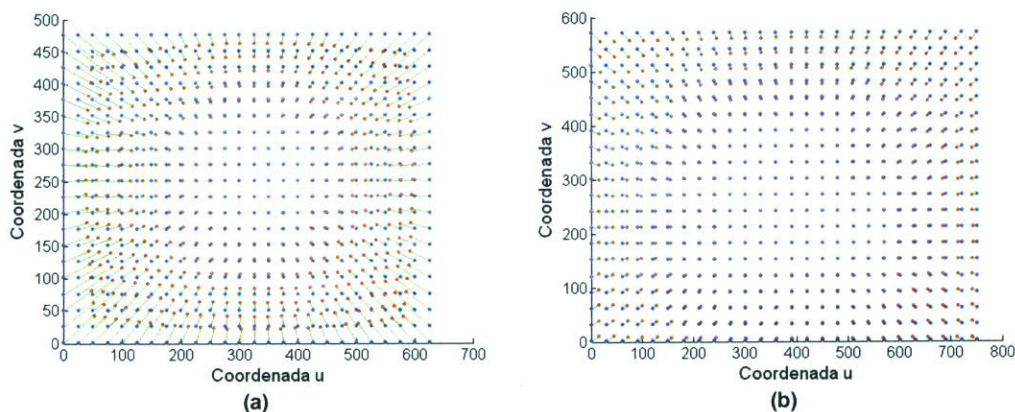
**Figura 3.30.** Errores debidos a la distorsión de cámara, en función del número de iteraciones realizadas para corregir la distorsión utilizado en esta tesis. (a) error máximo. (b) error medio.

De la figura 3.30 se puede concluir que tres iteraciones ( $k=3$ ) son suficientes para que la serie converja. Así, realizando tres iteraciones en (3.23), se tiene que el modelo de corrección de la distorsión propuesto viene dado por:

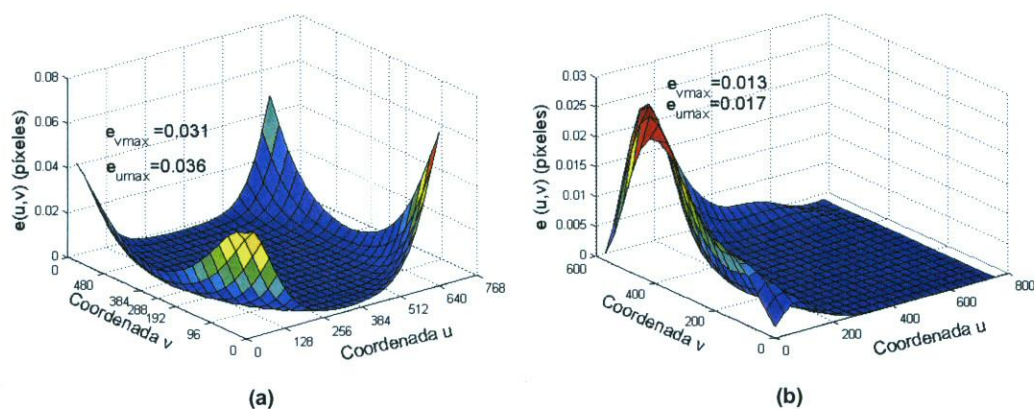
$$p = p' - \delta(p' - \delta(p' - \delta(p'))) \tag{3.25}$$

Aplicando la ecuación (3.25) a cada uno de los píxeles de las imágenes de la figura 3.27 se obtiene la figura 3.31. En ella se muestra cómo los píxeles distorsionados (en color rojo) son desplazados a su posición corregida (en color azul), la cual se asemeja más a la real. La figura 3.31.a se corresponde con la corrección de la distorsión de la figura 3.27.a, y la figura 3.31.b con la figura 3.27.b.

Realizando la corrección de distorsión (ecuación (3.25)) sobre las coordenadas de los píxeles de una imagen y obteniendo la magnitud de su desplazamiento respecto a su valor ideal (ecuación (3.24)) se obtienen las gráficas de la figura 3.32. Como se puede observar, ahora el valor de este desplazamiento es muy pequeño, pudiendo considerarse prácticamente depreciable.



**Figura 3.31.** Efecto de la corrección de la distorsión de la óptica. (a) para el ejemplo de la imagen de la figura 3.27.a. (b) para el ejemplo de la imagen de la figura 3.27.b.



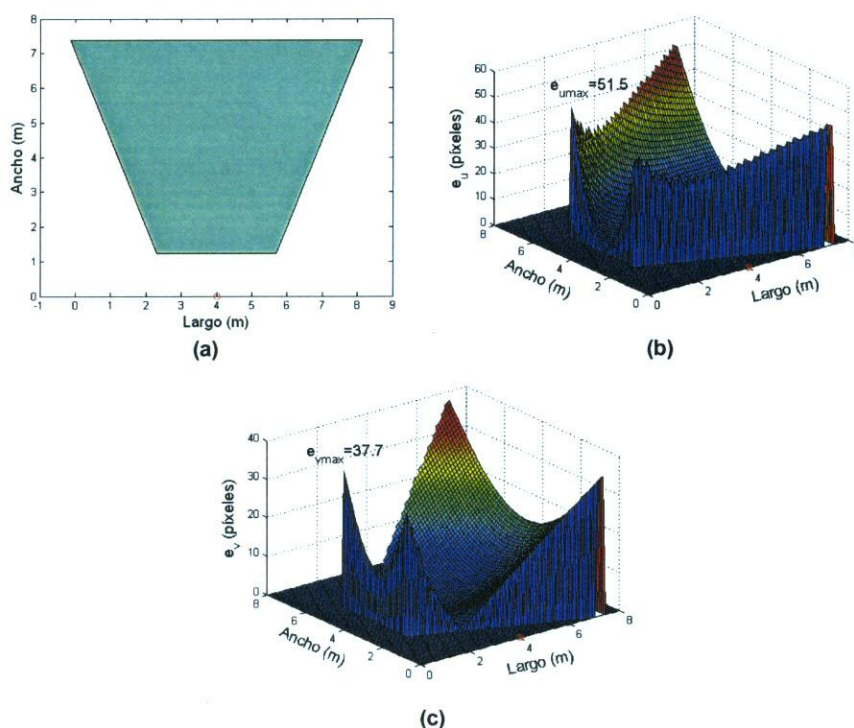
**Figura 3.32.** Valor del desplazamiento en las coordenadas  $(u, v)$  de los píxeles tras corregir la distorsión. (a) Para la óptica utilizada en la imagen de la figura 3.26.a. (b) Para la óptica utilizada en la imagen de la figura 3.26. c.

Para obtener las gráficas de la figura 3.28 se ha analizado el efecto de la distorsión sobre cada uno de los píxeles de la imagen. Evidentemente, lo que interesa para la aplicación de esta tesis es conocer cómo afecta la distorsión al cálculo del valor de las coordenadas del centroide de la proyección de la baliza a bordo del robot. Para estimar este efecto se han realizado pruebas sintéticas utilizando una cámara con un campo de visión como el que se muestra en la figura 3.33.a y una baliza circular de 2 cm de radio. La baliza se ubicó en distintas posiciones del campo de visión de la cámara. En ausencia de distorsión el centroide de la baliza se corresponde con las coordenadas  $(u, v)$ . Aplicando las ecuaciones (3.19), (3.17) y (3.20) se obtiene la imagen distorsionada de la baliza cuyo centroide será ahora  $(u', v')$ , siendo el valor de su desplazamiento respecto al ideal igual a:



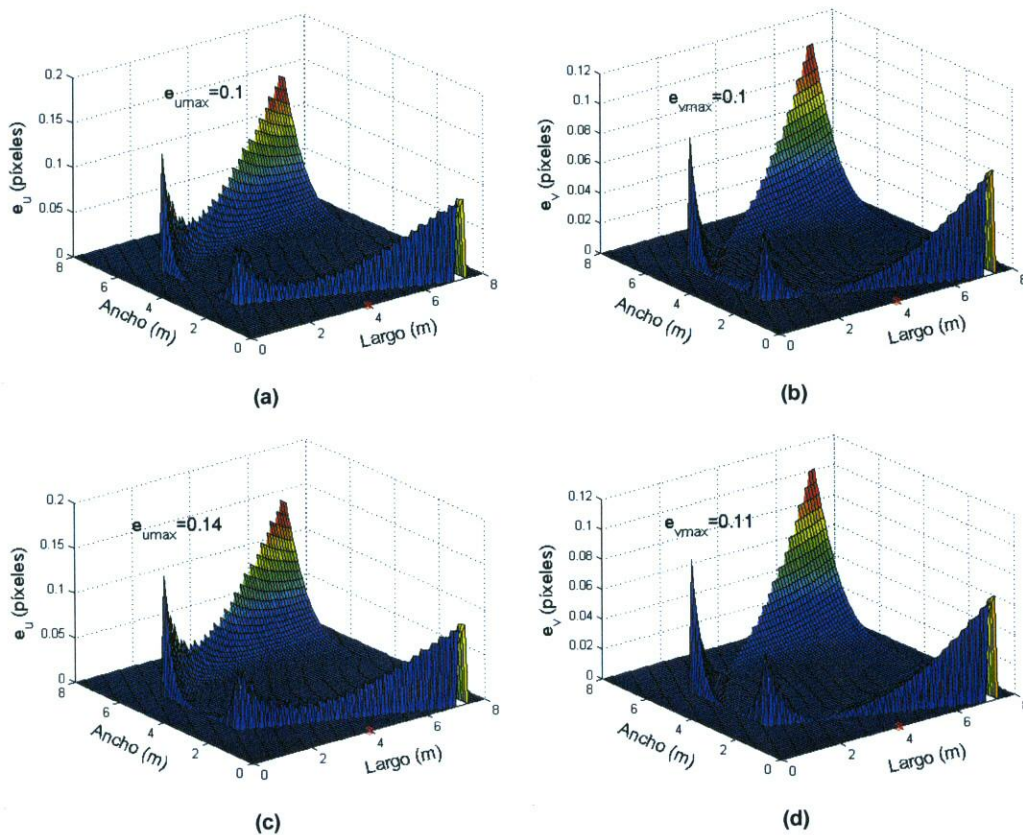
$$\begin{aligned}
 e_u &= |u - u'| \\
 e_v &= |v - v'|
 \end{aligned}
 \tag{3.26}$$

En la figura 3.33 se muestra la variación de las coordenadas del centroide de la baliza debido a la distorsión, para unos coeficientes de distorsión:  $k_1=-0.3444$ ,  $k_2=0.1518$ ,  $p_1=-0.0001$  y  $p_2=-0.0025$ .



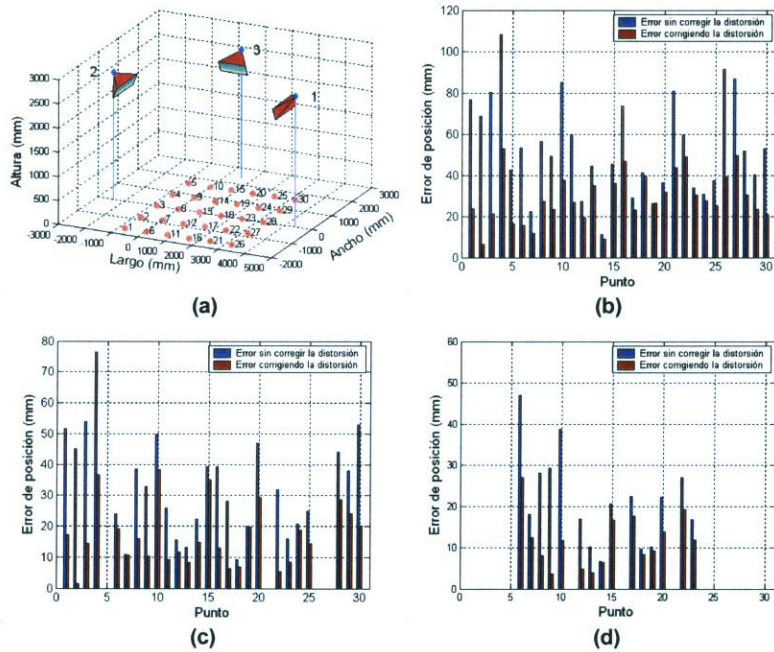
**Figura 3.33.** Desplazamiento de las coordenadas del centroide de la baliza circular debido a la distorsión. (a) Campo de visión de la cámara. (b) Variación de la coordenada  $u$  del centroide. (c) Variación de la coordenada  $v$  del centroide.

Evidentemente, la proyección de la baliza en la imagen ocupa un número variable de píxeles que será función del tamaño de la baliza y de la distancia a la cámara. Así, a la hora de corregir la distorsión, aplicando la ecuación (3.25), caben dos propuestas: una aplicando la corrección a cada uno de los píxeles de su proyección (figura 3.34.a y b) y otra aplicándola sólo al punto central de su proyección (figura 3.34.c y d). Como se puede apreciar, en ambos casos, los resultados son similares, pero la segunda alternativa presenta la ventaja de un menor coste computacional.



**Figura 3.34.** Desplazamiento de las coordenadas del centroide de la baliza una vez realizada la corrección de la distorsión. (a) Coordenada  $u$  realizando la corrección sobre todos los píxeles correspondientes a la proyección de la baliza. (b) Coordenada  $v$  realizando la corrección sobre todos los píxeles correspondientes a la proyección de la baliza. (c) Coordenada  $u$  realizando la corrección sólo del centroide de la proyección de la baliza. (d) Coordenada  $v$  realizando la corrección sólo del centroide de la proyección de la baliza.

En la figura 3.35 se muestran los valores reales del error de posicionamiento, con y sin corrección de la distorsión, para el caso de una baliza circular y un sistema formado por tres cámaras. En la figura 3.35.a se muestra la posición en la que se han ubicado las cámaras, así como las posiciones en las que se ha colocado la baliza. Los errores se muestran para los casos en que se utilice la información de una (figura 3.35.b), dos (figura 3.35.c) o tres cámaras (figura 3.35.d) para obtener la posición de la baliza. Como se puede observar, al realizar la corrección de la distorsión se obtiene una reducción importante en el error de posicionamiento.



**Figura 3.35.** Valores experimentales reales del error que se comete en la determinación de la posición de una baliza circular para un sistema formado por tres cámaras, con y sin corrección de la distorsión. (a) Distribución de los puntos de prueba y ubicación de las cámaras. (b) Error utilizando la información de una única cámara para obtener la posición de la baliza. (c) Error utilizando dos cámaras. (d) Error utilizando tres cámaras.

### 3.4.4. Errores debidos al proceso de calibración.

Todo proceso de calibración para la obtención de los parámetros intrínsecos y extrínsecos de una cámara está sujeto a imprecisiones. Imprecisiones que se van a traducir en un error a la hora de obtener la posición de las balizas. Para evaluar el efecto que sobre la obtención de las coordenadas del robot tienen los errores cometidos en la obtención de los parámetros de la cámara, basta con analizar el efecto de los errores de calibración sobre los valores de las coordenadas pixélicas del centroide de la baliza a bordo. Para ello se puede hacer un estudio similar al que se ha realizado en apartados anteriores. Así, partiendo de la ecuación (3.7) se pueden obtener las coordenadas  $(u, v)$  de dicho centroide:

$$\begin{aligned}
 u &= f_x \cdot \frac{r_{11} \cdot X + r_{12} \cdot Y + r_{13} \cdot Z + T_X}{r_{31} \cdot X + r_{32} \cdot Y + r_{33} \cdot Z + T_Z} + u_0 = f_x \frac{N_u}{D_u} + u_0 \\
 v &= f_y \cdot \frac{r_{21} \cdot X + r_{22} \cdot Y + r_{23} \cdot Z + T_Y}{r_{31} \cdot X + r_{32} \cdot Y + r_{33} \cdot Z + T_Z} + v_0 = f_y \frac{N_v}{D_v} + v_0
 \end{aligned}
 \tag{3.27}$$

donde:

$$\begin{aligned}
 r_{11} &= \cos(\gamma) \cos(\beta); & r_{12} &= \cos(\gamma) \text{sen}(\beta) \text{sen}(\alpha) - \text{sen}(\gamma) \cos(\alpha); \\
 r_{13} &= \cos(\gamma) \text{sen}(\beta) \cos(\alpha) + \text{sen}(\gamma) \text{sen}(\alpha); & r_{21} &= \text{sen}(\gamma) \cos(\beta); \\
 r_{22} &= \text{sen}(\gamma) \text{sen}(\beta) \text{sen}(\alpha) + \cos(\gamma) \cos(\alpha); & r_{23} &= \text{sen}(\gamma) \text{sen}(\beta) \cos(\alpha) - \cos(\gamma) \text{sen}(\alpha); \\
 r_{31} &= -\text{sen}(\beta); & r_{32} &= \cos(\beta) \text{sen}(\alpha); & r_{33} &= \cos(\beta) \cos(\alpha)
 \end{aligned}
 \tag{3.28}$$

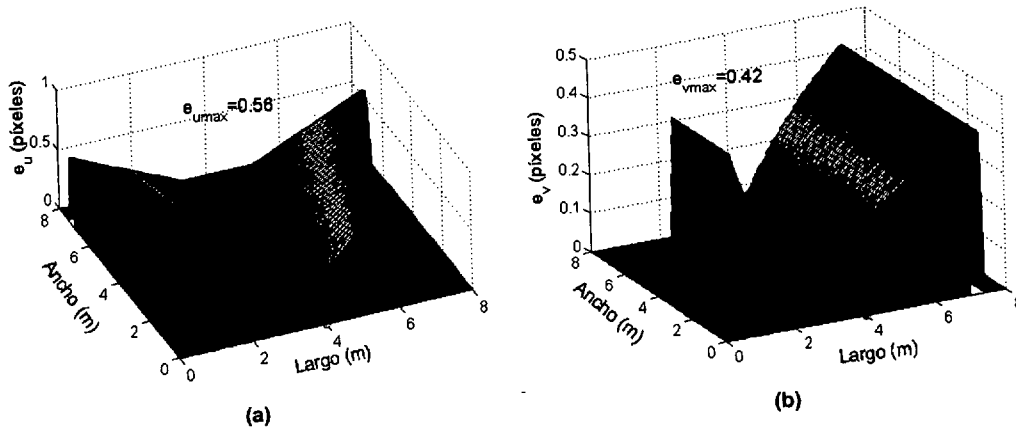
En primer lugar se va analizar la influencia de la variación de los parámetro intrínsecos ( $f_x, f_y, u_0, v_0$ ) sobre la variación de  $u$  y  $v$ . Derivando (3.27) respecto a cada uno de estos parámetros se tiene:

$$\begin{aligned}
 \frac{\partial u}{\partial f_x} &= \frac{N_u}{D_v} ; & \frac{\partial u}{\partial f_y} &= 0 \\
 \frac{\partial v}{\partial f_y} &= \frac{N_v}{D_v} ; & \frac{\partial v}{\partial f_x} &= 0 \\
 \frac{\partial u}{\partial u_0} &= 1 ; & \frac{\partial u}{\partial v_0} &= 0 \\
 \frac{\partial v}{\partial v_0} &= 1 ; & \frac{\partial v}{\partial u_0} &= 0
 \end{aligned}
 \tag{3.29}$$

De las ecuaciones (3.29) se puede concluir que los errores cometidos en la obtención de los parámetros  $u_0$  y  $v_0$  se trasladan directamente a las coordenadas  $u$  y  $v$  del centroide, mientras que el efecto en  $u$  y  $v$  de los errores cometidos en la obtención de  $f_x$  y  $f_y$  depende de las coordenadas 3D del punto proyectado. Así, para una baliza ubicada en las coordenadas  $[X, Y, Z]^T$ , la relación entre los errores cometidos en la obtención de  $f_x$  ( $e_{f_x}$ ) y  $f_y$  ( $e_{f_y}$ ) y los cometidos en la obtención de las coordenadas pixélicas de un punto ( $e_u, e_v$ ) viene dada por:

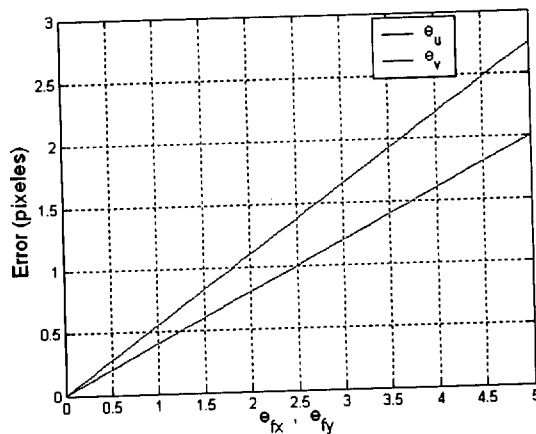
$$\begin{aligned}
 \frac{e_u}{e_{f_x}} &= \frac{r_{11} \cdot X + r_{12} \cdot Y + r_{13} \cdot Z + T_X}{r_{31} \cdot X + r_{32} \cdot Y + r_{33} \cdot Z + T_Z} \\
 \frac{e_v}{e_{f_y}} &= \frac{r_{21} \cdot X + r_{22} \cdot Y + r_{23} \cdot Z + T_Y}{r_{31} \cdot X + r_{32} \cdot Y + r_{33} \cdot Z + T_Z}
 \end{aligned}
 \tag{3.30}$$

En la figura 3.36 se muestran los errores en las coordenadas pixélicas del centroide de una baliza en función de la posición, dentro del campo de visión reflejado en la figura 3.33.a, suponiendo  $e_{f_x} = e_{f_y} = 1$  píxel. Como se puede observar el error aumenta a medida que la baliza se aleja de la cámara.



**Figura 3.36.** Error cometido en la obtención de las coordenadas pixélicas del centroide de la baliza, en función de la posición de la baliza respecto a la cámara, para  $e_{fx} = e_{fy} = 1$  píxel (a) Error en la coordenada  $u$ . (b) Error en la coordenada  $v$ .

Repitiendo el proceso seguido para la obtención de la figura 3.36 pero para diferentes  $e_{fx}$  y  $e_{fy}$ , y tomando el valor máximo del error de las coordenadas pixélicas del centroide, se obtiene la gráfica de la figura 3.37.



**Figura 3.37.** Error en las coordenadas pixélicas del centroide de la baliza en función del error en la obtención de los parámetros intrínsecos  $f_x$  y  $f_y$  ( $e_{fx}$  y  $e_{fy}$ )

En el caso de los parámetros extrínsecos, se puede seguir un procedimiento similar al de los intrínsecos. En las ecuaciones (3.31) y (3.32.) se muestran las variaciones que sufren las coordenadas pixélicas de un punto en función de las variaciones en los parámetros extrínsecos de la cámara.

$$\begin{aligned}
 \frac{\partial u}{\partial \alpha} &= f_x \frac{D_u (r_{11}^\alpha X + r_{12}^\alpha Y + r_{13}^\alpha Z) + N_u (r_{31}^\alpha X + r_{32}^\alpha Y + r_{331}^\alpha Z)}{D_u^2} \\
 \frac{\partial u}{\partial \beta} &= f_x \frac{D_u (r_{11}^\beta X + r_{12}^\beta Y + r_{13}^\beta Z) + N_u (r_{31}^\beta X + r_{32}^\beta Y + r_{331}^\beta Z)}{D_u^2} \\
 \frac{\partial u}{\partial \gamma} &= f_x \frac{D_u (r_{11}^\gamma X + r_{12}^\gamma Y + r_{13}^\gamma Z) + N_u (r_{31}^\gamma X + r_{32}^\gamma Y + r_{331}^\gamma Z)}{D_u^2} \\
 \frac{\partial v}{\partial \alpha} &= f_y \frac{D_v (r_{21}^\alpha X + r_{22}^\alpha Y + r_{23}^\alpha Z) + N_v (r_{31}^\alpha X + r_{32}^\alpha Y + r_{331}^\alpha Z)}{D_v^2} \\
 \frac{\partial v}{\partial \beta} &= f_y \frac{D_v (r_{21}^\beta X + r_{22}^\beta Y + r_{23}^\beta Z) + N_v (r_{31}^\beta X + r_{32}^\beta Y + r_{331}^\beta Z)}{D_v^2} \\
 \frac{\partial v}{\partial \gamma} &= f_y \frac{D_v (r_{21}^\gamma X + r_{22}^\gamma Y + r_{23}^\gamma Z) + N_v (r_{31}^\gamma X + r_{32}^\gamma Y + r_{331}^\gamma Z)}{D_v^2}
 \end{aligned} \tag{3.31}$$

$$\begin{aligned}
 \frac{\partial u}{\partial T_x} &= f_x \frac{1}{D_u}; & \frac{\partial u}{\partial T_y} &= 0; & \frac{\partial u}{\partial T_z} &= f_x \frac{N_u}{D_u^2} \\
 \frac{\partial v}{\partial T_x} &= 0; & \frac{\partial v}{\partial T_y} &= f_y \frac{1}{D_v}; & \frac{\partial v}{\partial T_z} &= f_y \frac{N_v}{D_v^2}
 \end{aligned} \tag{3.32}$$

donde  $r_{ij}^\xi$  representa la derivada del coeficiente  $r_{ij}$  respecto al ángulo de Euler  $\xi$  ( $\xi = \alpha, \beta, \gamma$ ).

Utilizando la ecuación (3.31) para los puntos de la baliza circular y para diferentes ubicaciones de ésta dentro del campo de visión de la cámara (figura 3.33.a), se obtienen las variaciones en las coordenadas del centroide debido a los errores cometidos en la determinación de los ángulos de Euler ( $\alpha, \beta, \gamma$ ). Así, en la figura 3.38 se representan estas variaciones para un error en el ángulo  $\alpha$  de  $0.4^\circ$ .

Repetiendo el proceso anterior para diferentes valores del error en la obtención de los ángulos de Euler se obtienen las gráficas de la figura 3.39.

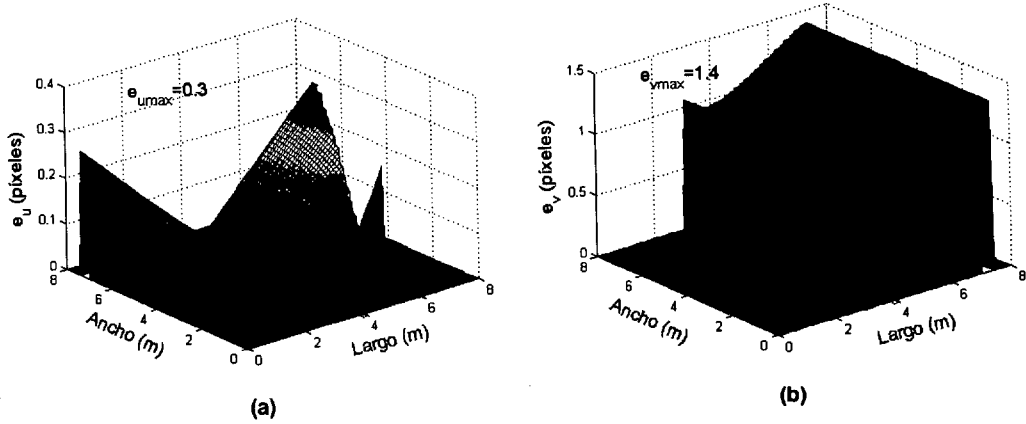


Figura 3.38. Error cometido en las coordenadas píxelicas del centroide de la baliza para un error en el cálculo del ángulo  $\alpha$  de  $0.4^\circ$ . (a) Error en la coordenada  $u$ . (b) Error en la coordenada  $v$ .

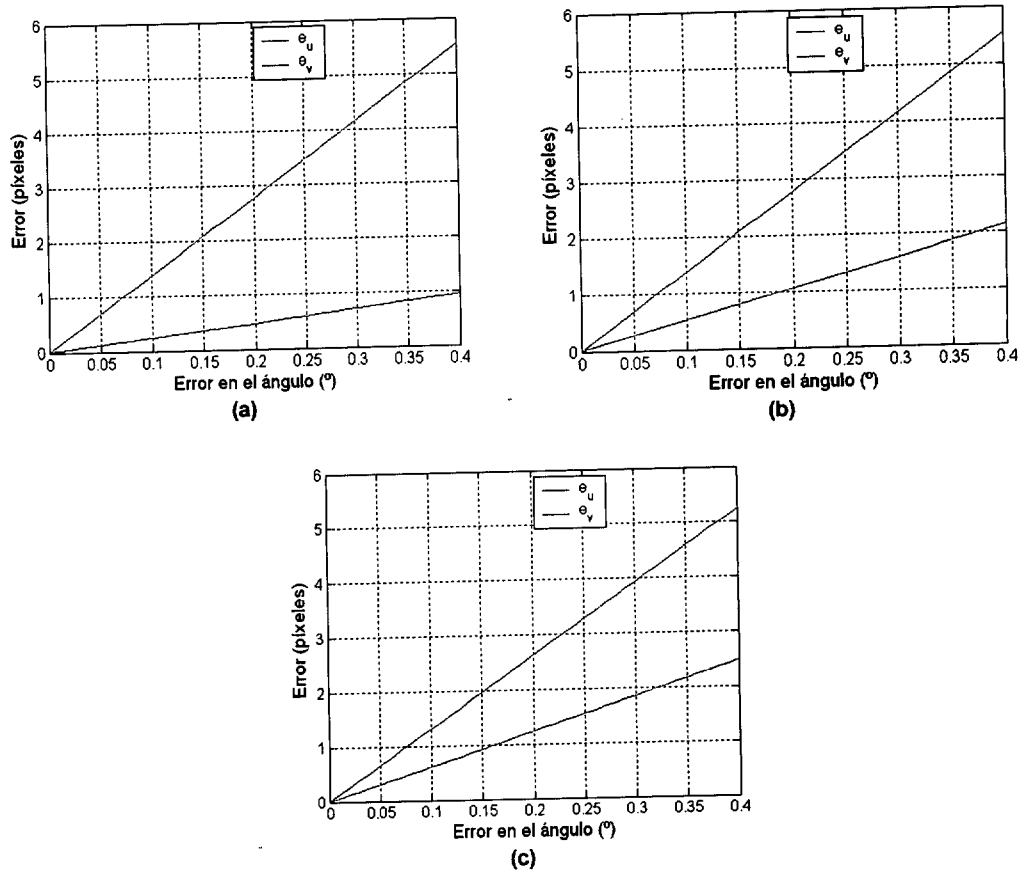


Figura 3.39. Error máximo de las coordenadas píxelicas del centroide de la baliza en función de las variaciones de los ángulos de Euler. (a) Suponiendo errores en el ángulo  $\alpha$ . (b) Suponiendo errores en el ángulo  $\beta$ . (c) Suponiendo errores en el ángulo  $\gamma$ .

Por otra parte, haciendo uso de la ecuación (3.32), se puede analizar la influencia de los errores cometidos en la obtención de los de los parámetros de traslación ( $T_x$ ,  $T_y$  y  $T_z$ ). Este efecto se muestra en las gráficas de la figura 3.40.

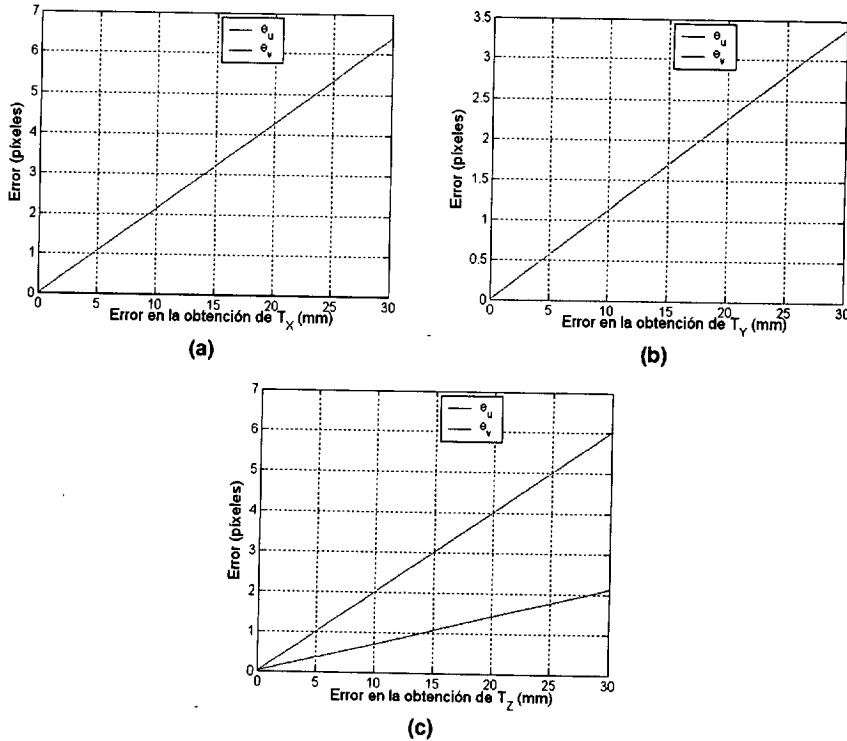
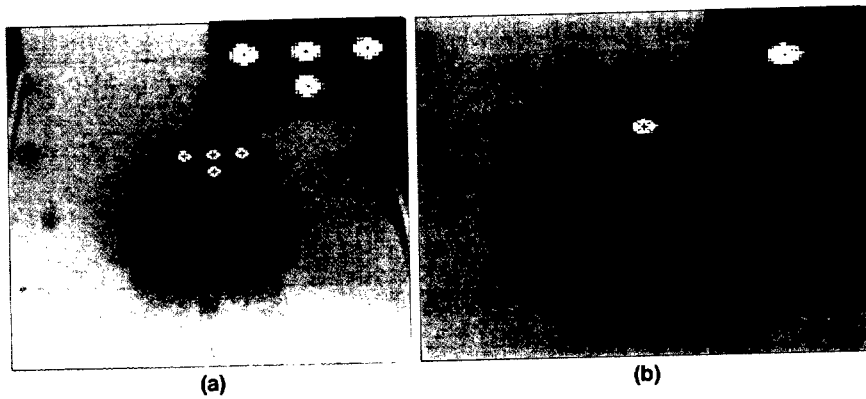


Figura 3.40. Errores en las coordenadas pixélicas del centroide de la baliza en función de los errores en los coeficientes de la matriz de traslación.(a) Para el coeficiente  $T_x$ .(b) Para el coeficiente  $T_y$ . (c) Para el coeficiente  $T_z$ .

### 3.4.5. Errores debidos a la imprecisión en la obtención de las coordenadas pixélicas del centroide.

Una vez identificadas las balizas en una imagen, el paso siguiente es la obtención de las coordenadas pixélicas de los correspondientes centroides. Evidentemente para minimizar el error en la obtención de las coordenadas 3D de las balizas, habrá que minimizar el error en la determinación de las coordenadas pixélicas de dicho centroide. En este punto se propone para este fin una solución con precisión sub-pixélica. La propuesta realizada en esta tesis consta de dos fases: en una primera se obtiene, en el plano imagen, el rectángulo mínimo que circunscribe todos los puntos de la proyección de la baliza (figura 3.41), y a continuación se obtiene la intersección de las diagonales de dicho rectángulo. Las coordenadas de la intersección de esas dos diagonales son precisamente las tomadas como coordenadas pixélicas del centroide de la baliza.

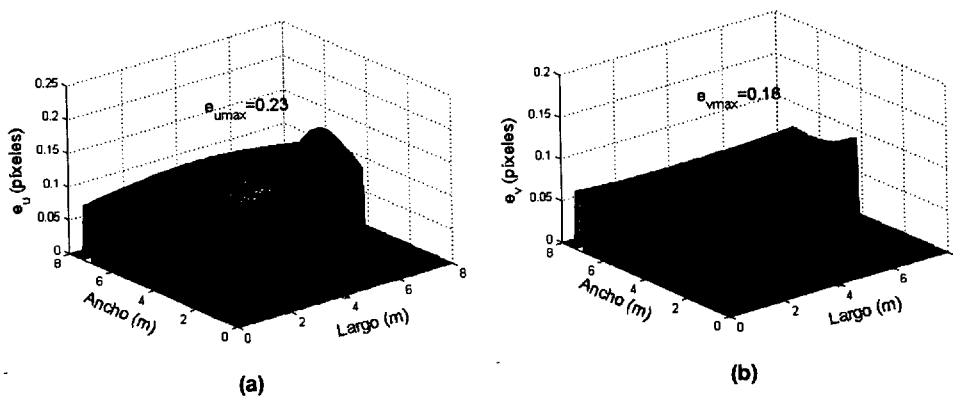




**Figura 3.41.** Propuesta para la obtención sub-pixélica de las coordenadas del centroide de las balizas..(a) caso de cuatro balizas a bordo del robot.(b) caso de una baliza.

Esta alternativa es computacionalmente muy sencilla y los resultados son muy similares a los obtenidos en la propuesta realizada en [Lavest, 99] y que se basa en el análisis de luminancia y el ajuste a un modelo local paramétrico, capaz de simular la respuesta del CCD y la cámara gracias a un proceso de optimización. Esta última alternativa ha sido contrastada como una excelente solución, aunque requiere una elevada carga computacional.

A modo de ejemplo en la figura 3.42 se muestra la comparación entre los resultados obtenidos en [Lavest 99] y los obtenidos con nuestra propuesta, para el caso de una baliza de 2 cm de radio y una cámara con un campo de visión similar al mostrado en la figura 3.33.a. Como se puede observar la diferencia entre ambas propuestas es mínima.



**Figura 3.42.** Comparación entre los métodos propuesto en esta tesis y el propuesto en [Lavest 99] para la obtención de las coordenadas pixélicas del centroide de la baliza. (a) Coordenada  $u$ . (b) Coordenada  $v$ . En ambos casos se representa la diferencia entre los resultados obtenidos con las dos propuestas.

### 3.5. Conclusiones.

En este capítulo se ha realizado un estudio detallado de las diferentes fuentes de error que influyen en la obtención de las coordenadas pixélicas del centroide de las balizas (tras definir brevemente la estructura de balizas a ubicar a bordo del robot para facilita su identificación y obtener información sobre la posición y orientación de éste) y, en consecuencia, la influencia en la obtención de las coordenadas 3D del robot. Se ha puesto de manifiesto que las diferentes fuentes de error son acumulativas y afectan en diferente grado a la determinación de las coordenadas en el plano imagen del centroide de la baliza.

Las fuentes de error analizadas han sido el tiempo de exposición (*shutter*) de las cámaras, los desplazamientos temporales debidos a la falta de sincronismo entre las cámaras y aspectos asociados al hardware de adquisición, el modelo de cámara, el proceso de calibración de las cámaras, y el proceso de obtención de las coordenadas pixélicas de las balizas en el plano imagen. Como conclusiones del estudio realizado se puede decir:

- a) Que el tiempo de exposición tiene un incidencia importante en el cálculo de las coordenadas pixélicas de la baliza, y que su efecto se puede minimizar, además de tratando de reducir el tiempo de exposición y la velocidad del móvil, utilizando cámaras de grandes distancias focales y tamaño de CCD.
- b) Los desplazamientos temporales debidos a la falta de sincronismo entre las cámaras y el hardware de adquisición pueden constituir una de las mayores fuentes de error, lo cual obliga a que todas las cámaras tengan una señal de sincronismo común y que el hardware de adquisición permita que ésta se realice de forma simultánea para todas las cámaras.
- c) Las distorsiones introducidas por la óptica también son una fuente de error importante, sobre todo para los puntos proyectados en los bordes del plano imagen; por ello en esta tesis se ha propuesto una técnica para reducir el efecto tanto de la distorsión radial como tangencial.
- d) En cuanto al proceso de calibración, también se ha realizado un estudio sobre el efecto que tienen en nuestra aplicación los errores cometidos en el proceso de calibración, tanto de los parámetros intrínsecos como extrínsecos; esto obliga a prestar una especial atención al proceso de calibración.

e) En el proceso de obtención de las coordenadas píxelicas del centroide de la baliza, se ha realizado una propuesta de obtención con precisión subpíxelica que ha demostrado dar excelentes resultados, y además con baja carga computacional.

Como aportaciones más destacadas, dentro de este capítulo están las relacionadas con las fuentes de error debidas al sistema de visión. En este sentido se puede decir que se ha realizado un detallado estudio sobre las diferentes fuentes de errores, se ha evaluado su efecto (partiendo de un estudio matemático y realizando numerosas simulaciones), y se han propuesto soluciones para reducir este efecto en la obtención de la posición del robot. En relación con estos últimos aspectos, es de destacar la propuesta realizada para minimizar los efectos de la distorsión de cámara y la de obtención de las coordenadas píxelicas del centroide de la baliza.

## Capítulo 4

---

### Posicionamiento 3D utilizando cámaras externas

---

Uno de los aspectos más importantes en el guiado de robots móvil es el de la localización dentro del entorno donde se deben mover éstos. Una localización rápida y fiable favorece sin duda todos los aspectos posteriores relacionados con la navegación. Todo sistema de localización o posicionamiento debe permitir que los robots se desplacen de forma continua sin necesidad de tener que detenerse cada vez que requieran determinar su posición.

En este capítulo se analizan los algoritmos de posicionamiento propuestos en esta tesis basados en la información proporcionada por las diferentes cámaras ubicadas en el entorno, y suponiendo que a bordo de los robots se disponen balizas de infrarrojos (IR) que facilitan la localización dentro de cada una de las imágenes que conforman el sistema de visión.

#### 4.1. Introducción.

En el capítulo anterior se analizó la influencia de las diferentes fuentes de error en el proceso de formación de las imágenes y que sin duda influyen, en mayor o menor grado, en la determinación de la localización de los robots. En este capítulo se presenta la propuesta de

localización 3D de los robots en función del número de cámaras que “ven” simultáneamente a cada uno de ellos. Para ello vamos a partir de un sistema de visión compuesto por  $n$  cámaras ubicadas en posiciones fijas del entorno. Para la propuesta de localización que se realiza en esta tesis se van a diferenciar tres casos: que el robot sea visto por una sola cámara, por dos cámaras, y por más de dos cámaras, si bien el caso más frecuente será el de dos.

A continuación se estudian los algoritmos propuestos en cada uno de los tres casos anteriores, donde se analizan su robustez y fiabilidad frente a los errores del sistema de visión. Los algoritmos desarrollados son independientes de si se utiliza una o cuatro balizas embarcadas en el robot, ya que lo que se pretende obtener es la posición 3D de una baliza a partir del centroide de su proyección en las diferentes cámaras que forman el sistema de visión. Por tanto, a partir de la determinación de la posición de una o de las cuatro balizas se obtiene la posición del robot.

## 4.2. Obtención de la posición del robot utilizando una única cámara.

Identificando por  $(u, v)$  las coordenadas que representan el centroide de la proyección de una baliza en una cámara, y por  $[X, Y, Z]^T$  las coordenadas 3D de su punto central, referidas al sistema de referencia global, la relación entre ellas viene dada por la ecuación (3.7), la cual, a modo de introducción, se muestra a continuación.

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.1)$$

La ecuación (4.1) se puede representar también como:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda \cdot \mathbf{R}^{-1} \begin{bmatrix} u - u_0 \\ \frac{f_x}{f_y} \frac{v - v_0}{1} \end{bmatrix} - \mathbf{R}^{-1} \cdot \mathbf{T} \quad (4.2)$$

La ecuación (4.2) representa un sistema de tres ecuaciones con 4 incógnitas, por tanto conociendo la altura a la que se encuentra la baliza sobre el robot ( $Z=Z_0$ ), de (4.2) se obtiene un sistema de tres ecuaciones con tres incógnitas. Evidentemente la suposición de  $Z=Z_0=cte$  es válida para muchos entornos interiores, donde los robots se mueven en gran medida sobre superficies planas (plano XY).

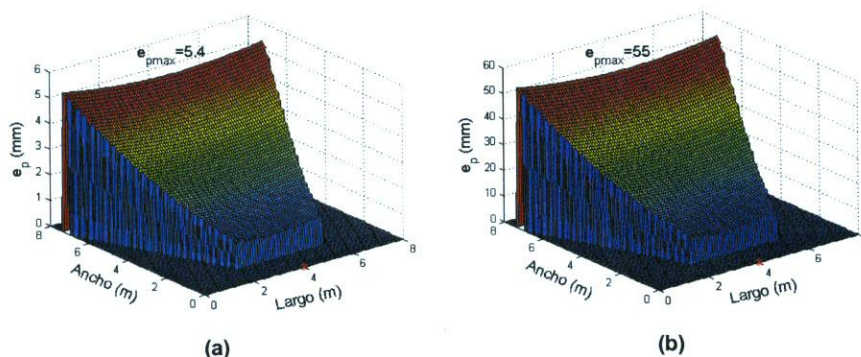
Debido al error producido en la obtención de las coordenadas del centroide de la proyección de la baliza se van a producir errores en la determinación de la localización 3D (error de posicionamiento). Así, si la baliza se encuentra en la posición  $[X, Y, Z]^T$ , a la que le corresponden, según (4.1), unas coordenadas  $(u, v)$  para el centroide de su proyección, y éstas se encuentran desplazadas una cantidad  $e_u$  y  $e_v$ , para las coordenadas  $u$  y  $v$  respectivamente, la posición asignada a la baliza (suponiendo  $Z=Z_0$ ) será:

$$\begin{bmatrix} X_e \\ Y_e \\ Z_0 \end{bmatrix} = \lambda_e \cdot \mathbf{R}^{-1} \begin{bmatrix} \frac{u + e_u - u_0}{f_x} \\ \frac{v + e_v - v_0}{f_y} \\ 1 \end{bmatrix} - \mathbf{R}^{-1} \cdot \mathbf{T} \quad (4.3)$$

de forma que el error de posicionamiento ( $e_p$ ) vendrá dado por:

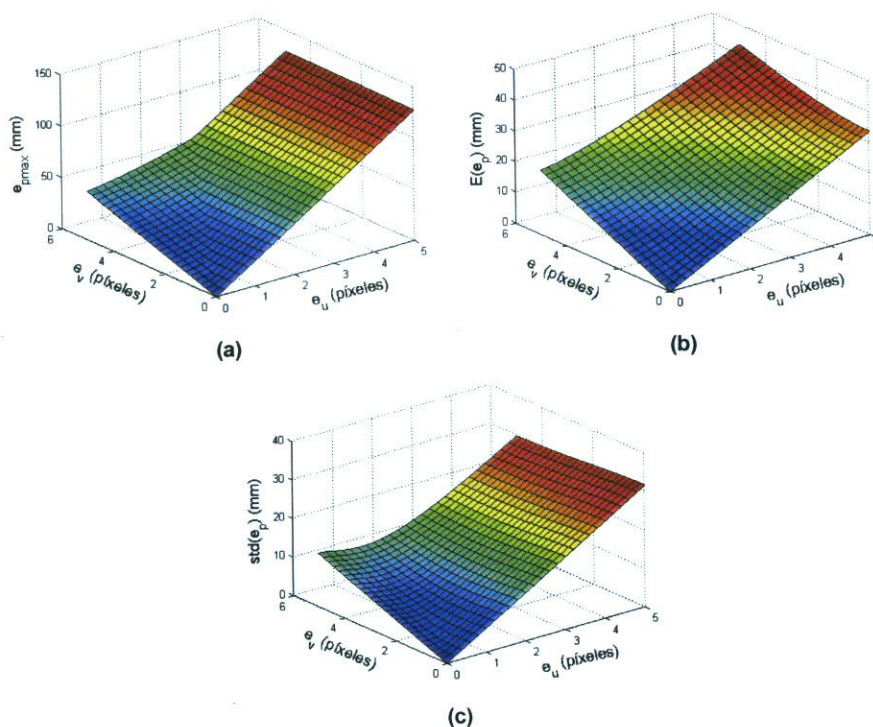
$$e_p = \left\| \begin{bmatrix} X \\ Y \end{bmatrix} - \begin{bmatrix} X_e \\ Y_e \end{bmatrix} \right\| \quad (4.4)$$

Para ver la precisión en la obtención de la posición de una baliza utilizando una sola cámara, se han realizado pruebas prácticas con un campo de visión como el reflejado en la figura 3.33.a, y con ésta ubicada en la posición:  $X=4$  m,  $Y=0$  m,  $Z=3$  m,  $\alpha=45^\circ$ ,  $\beta=\gamma=0^\circ$ . La baliza se ha situado en diferentes puntos del campo de visión de la cámara y las coordenadas ideales de su centroide se contaminan con unos error ( $e_u$  y  $e_v$ ). Haciendo uso de las ecuaciones (4.3) y (4.4), se han obtenido las gráficas de la figura 4.1. En la figura 4.1 se ha considerado que las coordenadas píxelicas del centroide están contaminadas con un error de media nula y desviación típica  $\sigma_u$  y  $\sigma_v$ . En la figura 4.1.a se representa el error de posicionamiento para el caso  $\sigma_u = \sigma_v = 0.2$  píxeles, mientras que para la figura 4.1.b se ha utilizado  $\sigma_u = \sigma_v = 2$  píxeles. Como se puede observar, el error de posicionamiento aumenta a medida que la baliza se aleja de la cámara, siendo apreciable para los puntos extremos del campo de visión.



**Figura 4.1.** Error de posicionamiento utilizando una única cámara, para dos valores de  $\sigma_u$  y  $\sigma_v$ . (a)  $\sigma_u = \sigma_v = 0.2$  píxeles. (b)  $\sigma_u = \sigma_v = 2$  píxeles

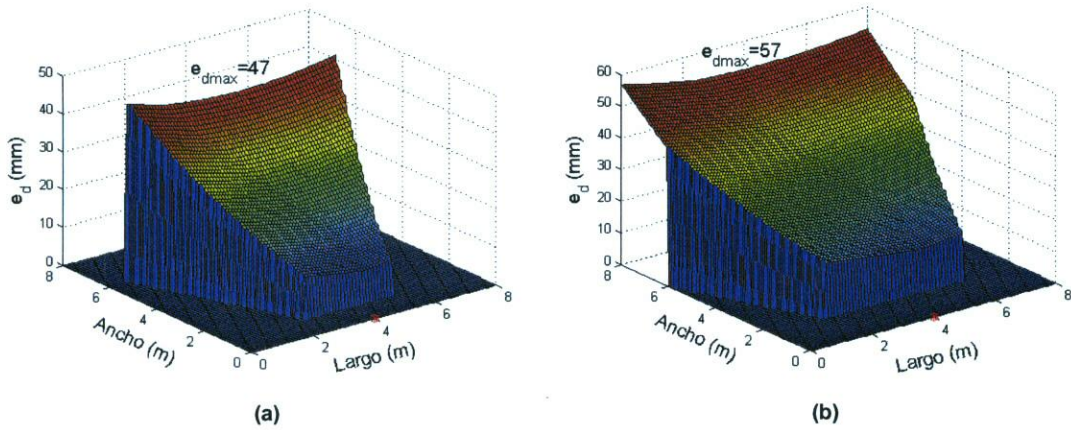
Si bien para la obtención de las gráficas de la figura 4.1 se ha supuesto que  $\sigma_u = \sigma_v$ , en la práctica, evidentemente, pueden tomar valores muy diferentes. Por ello, en la figura 4.2 se muestra el error de posicionamiento de la baliza para diferentes ubicaciones de ésta dentro del campo de visión de la cámara y tomando como parámetros  $\sigma_u$  y  $\sigma_v$ .



**Figura 4.2.** Representación de los estadísticos del error de posicionamiento para el caso de utilizar una única cámara, en función del error en la obtención de las coordenadas píxelicas del centroide de la baliza. La baliza se ha ubicado en el punto del campo de visión de la cámara para el cual el error de posicionamiento es máximo (a) Valor máximo. (b) Valor medio. (c) Desviación típica.

En la figura 4.2.a se representa el valor máximo del error de posicionamiento, mientras que las figuras 4.2.b y 4.2.c reflejan el valor medio y la desviación típica de este error. Hay que destacar que para valores de  $\sigma_u$  y  $\sigma_v$  superiores a 4 píxeles el error de posicionamiento es notable tomando valores del orden de los 10 cm.

Como se ha puesto de manifiesto en la figura 4.1, los errores en la obtención de la posición 3D de la baliza aumentan a medida que ésta se aleja de la cámara. En este sentido cabe decir que los resultados mostrados en las figuras 4.1 y 4.2 se han obtenido con las cámaras situadas a una altura de 3 m; evidentemente si esta altura se disminuye el error de posicionamiento también se reduce. A modo de ejemplo, en la figura 4.3 se representa este error para el caso en que las cámaras se ubiquen a una altura de 2.5 m (figura 4.3.a) y 4 m (figura 4.3.b), tomando en ambos casos  $\sigma_u = \sigma_v = 2$  píxeles.



**Figura 4.3.** Error de posicionamiento de la baliza para  $\sigma_u = \sigma_v = 2$  píxeles en función de la altura a la que se ubica la cámara. (a) altura de 2.5 m. (b) altura de 4 m.

En todo lo presentado hasta ahora se ha supuesto que el robot se mueve por una superficie perfectamente plana. En la práctica sin embargo las superficies pueden presentar pequeños desniveles. Para ver la influencia que estos desniveles pueden producir en la obtención de las coordenadas 3D del robot se han realizado simulaciones incluyendo los efectos de estos cambios de altura en la superficie. Para reflejar esta circunstancia la ecuación (4.3) se puede escribir como:

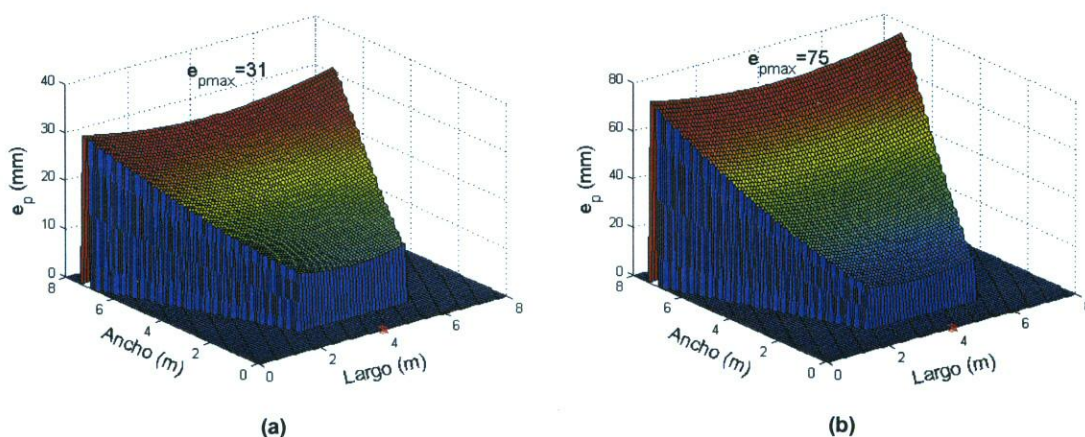
$$\begin{bmatrix} X_e \\ Y_e \\ Z_e \end{bmatrix} = \begin{bmatrix} X_e \\ Y_e \\ Z_0 + e_{z0} \end{bmatrix} = \lambda_e \cdot \mathbf{R}^{-1} \begin{bmatrix} \frac{u + e_u - u_0}{f_x} \\ \frac{v + e_v - v_0}{f_y} \\ 1 \end{bmatrix} - \mathbf{R}^{-1} \cdot \mathbf{T} \tag{4.5}$$



siendo ahora el valor del error de posicionamiento :

$$e_p = \left\| \begin{bmatrix} X \\ Y \\ Z_0 \end{bmatrix} - \begin{bmatrix} X_e \\ Y_e \\ Z_e \end{bmatrix} \right\| \tag{4.6}$$

Realizando el planteamiento seguido en la obtención de la figura 4.1 y contaminando la coordenada  $Z=Z_0$  con un error de media nula y una desviación típica ( $\sigma_{z_0}$ ) de 10 mm, se obtienen las gráficas de la figura 4.4. Ahora el error máximo de posicionamiento es del orden de los 31 mm y 75 mm para desviaciones típicas de ruido en las coordenadas pixélicas de 0.2 y 2 píxeles, respectivamente. Evidentemente estos errores son superiores a los 5.4 mm y 55 mm que se obtienen si se considera la superficie sin desnivel.

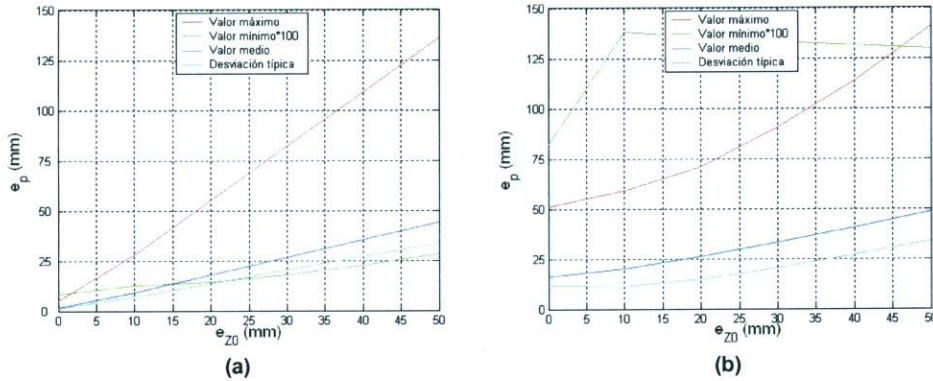


**Figura 4.4.** Error de posicionamiento utilizando una única cámara, para diferentes valores de  $\sigma_u$  y  $\sigma_v$ , suponiendo pequeñas irregularidades en altura en superficie por la que se mueve el robot. Se ha considerado una variación en altura de media nula y desviación típica ( $\sigma_{z_0}$ ) de 10mm. (a)  $\sigma_u = \sigma_v = 0.2$  píxeles (b)  $\sigma_u = \sigma_v = 2$  píxeles.

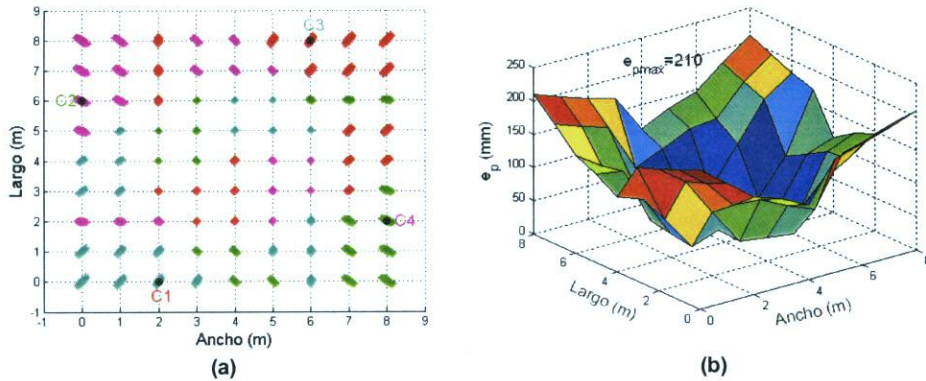
En la figura 4.5 se muestran los estadísticos del error de posicionamiento en el caso se realizar el procedimiento anterior con diferentes valores de la desviación típica del error en  $Z_0$  ( $\sigma_{z_0}$ ).

Dado que el error de posición está estrechamente relacionado con la distancia entre cámara y robot, es evidente que cuando se disponga de varias cámaras, si solamente fuera posible utilizar una de ellas habrá que utilizar la más próxima. A modo de ejemplo, para un entorno rectangular de 8x8 m sobre el que se ubican cuatro cámaras equiespaciadas a una altura de 3m, suponiendo unos parámetros intrínsecos como los indicados en la tabla 3.1, y para el caso de  $\sigma_u = \sigma_v = 2$  píxeles, se obtienen los errores de posicionamiento mostrados en la figura 4.6 . Los

colores utilizados en las representaciones se corresponden con la cámara utilizada (obsérvese que cada cámara se identifica por una letra de color diferente). Como se puede apreciar el error de posicionamiento es máximo en los bordes del entorno, ya que en esos puntos la distancia entre cámaras y robot es máxima. En el caso de los puntos centrales del entorno, el error se encuentra entre los 30 mm y 40 mm.

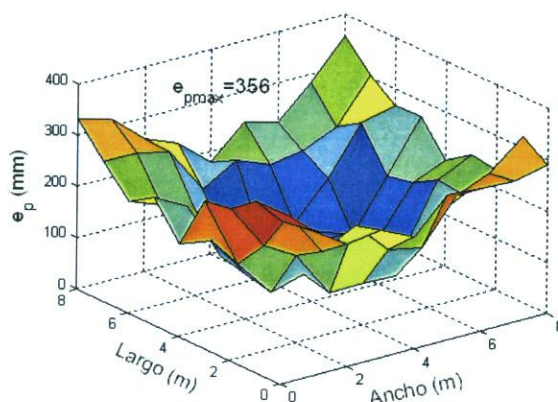


**Figura 4.5.** Estadísticos del error de posicionamiento en función del error en  $Z_0$ , para diferentes valores de  $\sigma_u$  y  $\sigma_v$ . (a)  $\sigma_u = \sigma_v = 0.2$  píxeles (b)  $\sigma_u = \sigma_v = 2$  píxeles.



**Figura 4.6.** Representación del error de posicionamiento, utilizando un única cámara, en un entorno de 8x8 m sobre el que se han ubicado cuatro cámaras equiespaciadas y tomando la más cercana a la baliza para obtener su posición, considerando  $\sigma_u = \sigma_v = 2$  píxeles. (a) incertidumbre de la posición. (b) error de posicionamiento.

Para la obtención de las gráficas de la figura 4.6 se ha supuesto que la altura a la cual se encuentra la baliza ( $Z=Z_0$ ) es constante. Si se supone un cambio de altura tipo ruido gaussiano de media nula y desviación típica de  $\sigma_{z0} = 10$  mm se obtienen los errores mostrados en la figura 4.7. En este caso el valor del error en los puntos centrales del entorno toman valores comprendidos entre 70 mm y 90 mm.

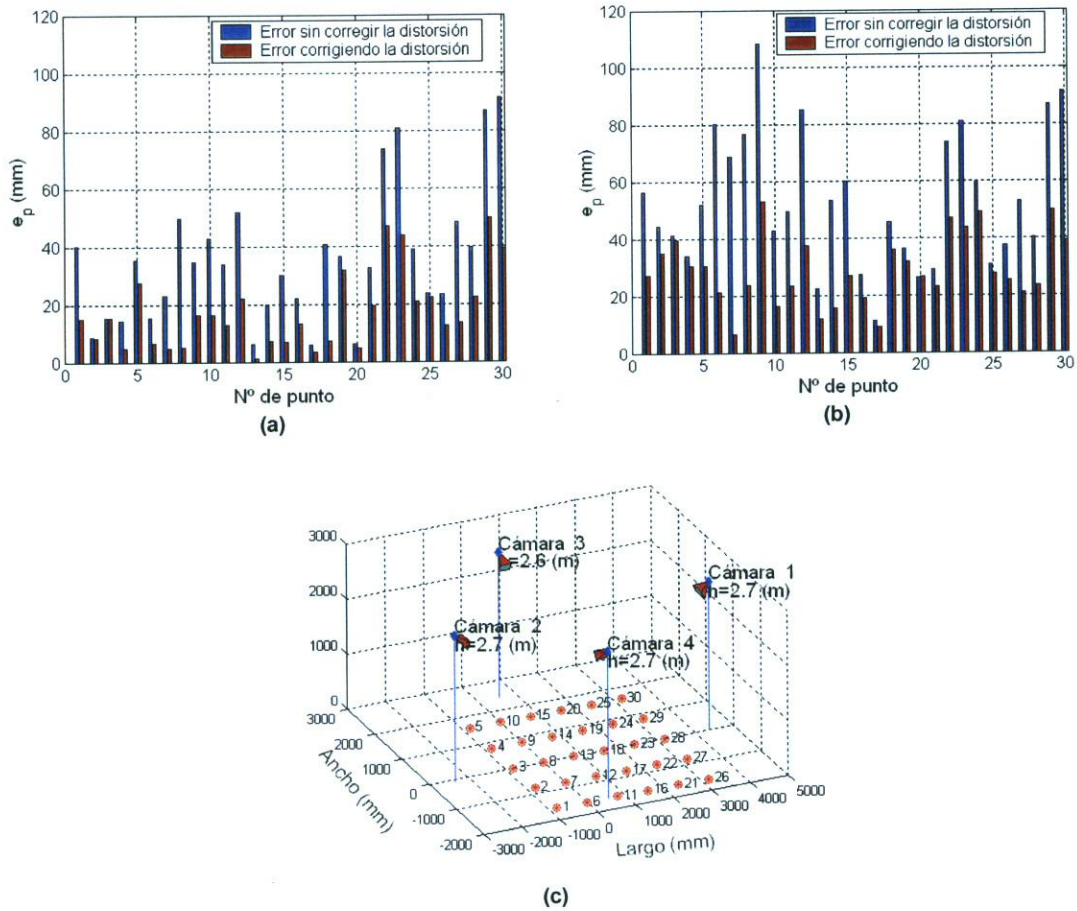


**Figura 4.7.** Valor del error de posicionamiento para las condiciones de trabajo de la figura 4.6, al que se le añade un error gaussiano en la coordenada  $Z_0$  de media nula y desviación típica 10 mm.

#### 4.2.1. Resultados prácticos.

Para validar el sistema de posicionamiento utilizando una única cámara (y suponiendo, por tanto, superficies planas), se han realizado pruebas en un entorno rectangular de 6x4m en el que se han ubicado cuatro cámaras con unas ópticas que cubren la totalidad del espacio por el que se mueve el robot. Se ha ubicado el robot en diferentes puntos del entorno y en cada uno de ellos se realizaron un total de mil medidas, obteniéndose su valor medio. Los resultados obtenidos en estas pruebas se muestran en la figura 4.8, tanto para el caso de no realizar corrección de la distorsión como de realizarla utilizando la ecuación (3.25). En el caso de la figura 4.8.a se utilizó la cámara más cercana al robot, mientras que para la figura 4.8.b se utilizó la más alejada. En la figura 4.8.c se representa las posiciones de los puntos de ensayo en los que se ubicó el robot y la de las cámaras.

En las tablas 4.1 y 4.2 se muestran los valores de los estadísticos del error de posición cometido al utilizar una única cámara ( cámara más cercana y más alejada al robot). Las conclusiones que se puede sacar de la tabla 4.1 son, por una parte, la eficacia del algoritmo de corrección de distorsión propuesto en esta tesis, y por otra, los buenos resultados obtenidos en aquellos puntos próximos a las cámaras. Hay que indicar que los errores de posicionamiento obtenidos con esta propuesta son en general menores que los obtenidos en las propuestas de otros autores [Lee 03][Yun 04].



**Figura 4.8.** Errores prácticos de posición del robot utilizando una única cámara, para un entorno de 6x4 m sobre el que se han ubicado cuatro cámaras. Se reflejan los valores sin corregir y corrigiendo la distorsión. (a) Utilizando la cámara más cercana al robot (b) Utilizando la cámara más alejada del robot (c) Ubicación física de los puntos de test y de las cámaras.

**TABLA 4.1.** Estadísticos del errores prácticos de posicionamiento utilizando la cámara más cercana a la baliza.

	Error mínimo (mm)	Error máximo (mm)	Error medio (mm)	Desviación típica (mm)
<b>Sin corregir la distorsión</b>	5.7	91.4	35.7	23.1
<b>Corrigiendo la distorsión</b>	1.2	49.8	17.3	13.3

<b>TABLA 4.2.</b> Estadísticos de los errores prácticos de posicionamiento utilizando la cámara más alejada de la baliza.				
	Error mínimo (mm)	Error máximo (mm)	Error medio (mm)	Desviación típica (mm)
<b>Sin corregir la distorsión</b>	10.9	108.1	53.4	23.6
<b>Corrigiendo la distorsión</b>	6.5	53.4	28.1	12.1

### 4.3. Obtención de la posición utilizando dos cámaras.

Dado que el uso de una sola cámara está restringido a espacios donde la componente vertical (altura) se mantiene constante, es necesario pensar en el uso simultáneo de varias cámaras para determinar la posición de los robots, con independencia del tipo de superficie por la que se muevan. Por ello, en este punto se aborda el tema de la obtención de las coordenadas 3D de la baliza a bordo del robot, a partir de la información proporcionada por dos cámaras. Es lo que se conoce popularmente como triangulación [Hartley 93]. La triangulación consiste en inferir las coordenadas 3D de un punto  $P$  de la escena a partir de las rectas (rectas de proyección) que van desde los centros ópticos ( $O_i, O_j$ ) de cada cámara hasta dicho punto, pasando por su proyección en los planos imagen ( $p_i, p_j$ ). En consecuencia, las coordenadas 3D del punto de la escena se corresponden con el punto de intersección de estas rectas (líneas de color rojo en la figura 4.9). Desafortunadamente, debido a los errores cometidos tanto en el proceso de detección, como en el de calibración (determinación de los parámetros intrínsecos y extrínsecos de las cámaras), las rectas de proyección no se cortan, sino que se cruzan, dando lugar a una zona de incertidumbre, reflejado en la figura 4.9 en color azul.

En la figura 4.9 los puntos  $p_i$  y  $p_j$  representa la proyección del punto  $P$  sobre los planos imagen de las cámaras  $i$  y  $j$ . Las coordenadas  $[X, Y, Z]^T$  del punto  $P$ , referidas al sistema de referencia global, a partir de las coordenadas píxelicas de sus proyecciones en cada uno de los planos imágenes ( $u_i, v_i$  y  $u_j, v_j$ ), según (4.2), vienen dadas por:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda_i \cdot R_i^{-1} \cdot \begin{bmatrix} u_i - u_{oi} \\ f_{xi} \\ v_i - v_{oi} \\ f_{yi} \\ 1 \end{bmatrix} - R_i^{-1} \cdot T_i = \lambda_i \cdot R_i^{-1} \cdot r_i - R_i^{-1} \cdot T_i = \lambda_i \cdot R_i^{-1} \cdot r_i + T_i' \quad (4.7)$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda_j \cdot \mathbf{R}_j^{-1} \cdot \begin{bmatrix} \frac{u_j - u_{oj}}{f_{xj}} \\ \frac{v_j - v_{oj}}{f_{yj}} \\ 1 \end{bmatrix} - \mathbf{R}_j^{-1} \cdot \mathbf{T}_j = \lambda_j \cdot \mathbf{R}_j^{-1} \cdot \mathbf{r}_j - \mathbf{R}_j^{-1} \cdot \mathbf{T}_j = \lambda_j \cdot \mathbf{R}_j' \cdot \mathbf{r}_j + \mathbf{T}_j' \quad (4.8)$$

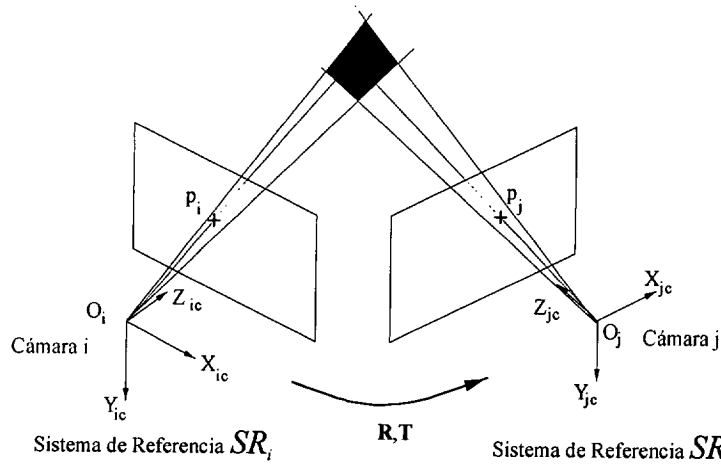


Figura 4.9. Concepto de triangulación, a partir de dos cámaras, para determinar las coordenadas 3D de un punto  $P$ .

donde  $\mathbf{R}'_k = \mathbf{R}^{-1}_k$  y  $\mathbf{T}'_k = -\mathbf{R}^{-1}_k \cdot \mathbf{T}_k$ , siendo  $\mathbf{R}_k$  y  $\mathbf{T}_k$  las matrices de rotación y traslación de la cámaras  $k$  ( $k=i, j$ ) referidas al sistema de referencia global. Igualando las expresiones (4.7) y (4.8) se obtiene:

$$\lambda_i \cdot \mathbf{r}_i = \lambda_j \cdot \mathbf{R} \cdot \mathbf{r}_j + \mathbf{T} \quad (4.9)$$

donde

$$\mathbf{R} = \mathbf{R}_{ij} = \mathbf{R}_i \cdot \mathbf{R}_j^{-1} \quad \text{y} \quad \mathbf{T} = \mathbf{T}_{ij} = \mathbf{T}_i - \mathbf{R}_i \cdot \mathbf{R}_j^{-1} \cdot \mathbf{T}_j \quad (4.10)$$

Las matrices  $\mathbf{R}$  y  $\mathbf{T}$  se pueden considerar como los parámetros de la transformación rígida (matriz de rotación y vector de traslación, respectivamente) entre los dos sistemas de referencia ( $SR_i$  y  $S_j$ ). La ecuación (4.9) conduce al sistema lineal:

$$[-\mathbf{R} \cdot \mathbf{r}_j, \mathbf{r}_i] \cdot \begin{bmatrix} \lambda_j \\ \lambda_i \end{bmatrix} = \mathbf{T} \quad (4.11)$$

Haciendo  $A = [-R_i r_i, r_i]$  y  $B = [\lambda_j, \lambda_i]^T$ , la ecuación (4.11) se puede escribir como:

$$A \cdot B - T = 0 \tag{4.12}$$

La solución de (4.12) viene dada por:

$$B = \begin{bmatrix} \lambda_j \\ \lambda_i \end{bmatrix} = (A^T \cdot A)^{-1} \cdot A^T \cdot T \tag{4.13}$$

Introduciendo los valores de  $\lambda_j, \lambda_i$  proporcionados por (4.13) en las ecuaciones (4.7) y (4.8) se puede obtener las coordenadas del punto  $P$ . En ausencia de error en las coordenadas pixélicas ( $u_i, v_i$  y  $u_j, v_j$ ), las soluciones obtenidas a partir de (4.7) y (4.8) deben coincidir, y por tanto se corresponde con el punto de intersección de las rectas de proyección. En caso contrario, las coordenadas proporcionadas por (4.7) y (4.8) corresponden a los puntos, pertenecientes a cada una de las rectas de proyección, en los que la distancia entre ellas es mínima (figura 4.10). Hay autores [Aitenbichler 03][Hartley 03] que asocian al punto  $P$  el valor del punto medio ( $P_m$ ) del segmento  $h$ , donde:

$$P_i = \lambda_i \cdot R'_i \cdot r_i + T'_i \quad \text{y} \quad P_j = \lambda_j \cdot R'_j \cdot r_j + T'_j \tag{4.14}$$

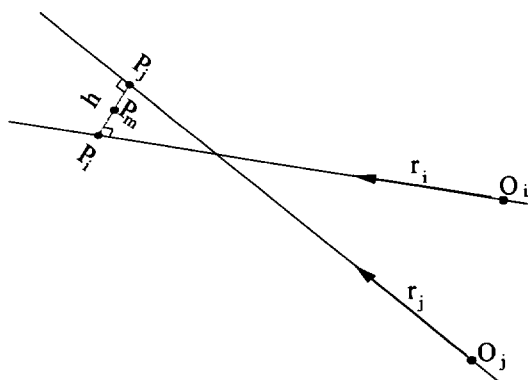


Figura 4.10. Ilustración gráfica del punto de cruce de las rectas de proyección de dos cámaras

En la asignación del punto medio ( $P_m$ ) del segmento  $h$  como valor de  $P$  se está presuponiendo que los errores cometidos en la obtención de los centroides de la baliza son idénticos para ambas cámaras. Sin embargo en la práctica esto no es así, debido a diversas razones prácticas propias del proceso de calibración, y además porque estos errores dependen de la distancia

entre cámaras y robot. Por tanto, sería más correcto suponer que el punto  $P$  se encuentra más próximo a la recta de proyección de la cámara en la que los errores son menores. En consecuencia, una propuesta más realista es aquella que asigna a  $P$  un punto, dentro del segmento  $h$ , más próximo a  $P_i$  o a  $P_j$ , en función de las varianzas de las coordenadas de dichos puntos.

La obtención de las coordenadas del punto  $P$  a partir del cruce de las dos rectas se puede abordar utilizando diferentes estrategias. Evidentemente, la mayor probabilidad de ubicación del punto  $P$  será sobre el segmento que determina la mínima distancia entre las dos rectas. Por otra parte, si los errores cometidos en la determinación de las coordenadas de los puntos  $P_i$  y  $P_j$ , fueren idénticos, también se podría pensar, como ya se ha comentado anteriormente, que el punto  $P$  debería ser equidistante a los puntos  $P_i$  y  $P_j$ . Sin embargo en la práctica estos errores no serán iguales y por tanto parece lógico premiar la proximidad de  $P$  al punto de menor error. Teniendo presentes estas consideraciones, en esta tesis se propone para la obtención de  $P$ , la minimización de una función de coste definida por la ecuación (4.15), donde  $C_k$  ( $k=j, i$ ) representa la matriz de covarianza asociada a las coordenadas de  $P_k$ . ( $k=j, i$ ). Como se puede observar, cada uno de los sumandos tiene un gran parecido con la distancia de Mahalanobis, la única diferencia radica en que aquí se utiliza  $C$  en lugar de  $C^{-1}$ ; ello se debe a que aquí se trata de que el punto  $P$  se aproxime más al punto de mayor certidumbre. En cualquier caso, los resultados obtenidos utilizando la distancia de Mahalanobis y las correspondiente operaciones posteriores que hagan que  $P$  se aproxime más al punto de menor certidumbre son idénticos. Por tanto, la función de coste propuesta viene dada por:

$$E = \sum_{k=i,j} (P - P_k)^T C_k (P - P_k) \quad (4.15)$$

Como lo que se busca es el valor de  $P$  que minimiza esta función, entonces:

$$\frac{\partial E}{\partial P} = 0 \Rightarrow \frac{\partial E}{\partial P} = \sum_{k=i,j} (C_k + C_k^T) \cdot (P - P_k) = 0 \quad (4.16)$$

Teniendo en cuenta que  $C_k$  es simétrica ( $C_k = C_k^T$ ) la ecuación (4.16) se convierte en:

$$\frac{\partial E}{\partial P} = \sum_{k=i,j} C_k \cdot (P - P_k) = 0 \quad (4.17)$$

Las matrices  $C_k$  ( $k=i,j$ ) vienen dadas por:



$$C_k = \begin{bmatrix} \sigma_{X(k)}^2 & \rho_{XY(k)} \cdot \sigma_{X(k)} \cdot \sigma_{Y(k)} & \rho_{XZ(k)} \cdot \sigma_{X(k)} \cdot \sigma_{Z(k)} \\ \rho_{XY(k)} \cdot \sigma_{Y(k)} \cdot \sigma_{X(k)} & \sigma_{Y(k)}^2 & \rho_{YZ(k)} \cdot \sigma_{Y(k)} \cdot \sigma_{Z(k)} \\ \rho_{XZ(k)} \cdot \sigma_{Z(k)} \cdot \sigma_{X(k)} & \rho_{YZ(k)} \cdot \sigma_{Y(k)} \cdot \sigma_{Z(k)} & \sigma_{Z(k)}^2 \end{bmatrix} \quad (4.18)$$

donde  $\sigma_{X(k)}$ ,  $\sigma_{Y(k)}$  y  $\sigma_{Z(k)}$  representan las varianzas de las coordenadas  $X$ ,  $Y$  y  $Z$ , respectivamente, del punto  $P_k$  ( $k=i, j$ ), y  $\rho_{XY(k)}$ ,  $\rho_{XZ(k)}$  y  $\rho_{YZ(k)}$  los coeficiente de correlación entre dichas coordenadas. En lo que sigue, los términos de las matrices  $C_k$  se van a identificar como:

$$C_k = \begin{bmatrix} c_{11(k)} & c_{12(k)} & c_{13(k)} \\ c_{21(k)} & c_{22(k)} & c_{23(k)} \\ c_{31(k)} & c_{32(k)} & c_{33(k)} \end{bmatrix} \quad (4.19)$$

En consecuencia, la ecuación (4.17) se puede expresar como:

$$\sum_{k=i,j} \begin{bmatrix} c_{11(k)} & c_{12(k)} & c_{13(k)} \\ c_{21(k)} & c_{22(k)} & c_{23(k)} \\ c_{31(k)} & c_{32(k)} & c_{33(k)} \end{bmatrix} \begin{bmatrix} X - X_k \\ Y - Y_k \\ Z - Z_k \end{bmatrix} = 0 \quad (4.20)$$

y desarrollando el producto de matrices de (4.20):

$$\sum_{k=i,j} \begin{bmatrix} c_{11(k)} \cdot X & c_{12(k)} \cdot Y & c_{13(k)} \cdot Z \\ c_{21(k)} \cdot X & c_{22(k)} \cdot Y & c_{23(k)} \cdot Z \\ c_{31(k)} \cdot X & c_{32(k)} \cdot Y & c_{33(k)} \cdot Z \end{bmatrix} = \sum_{k=i,j} \begin{bmatrix} c_{11(k)} \cdot X_k & c_{12(k)} \cdot Y_k & c_{13(k)} \cdot Z_k \\ c_{21(k)} \cdot X_k & c_{22(k)} \cdot Y_k & c_{23(k)} \cdot Z_k \\ c_{31(k)} \cdot X_k & c_{32(k)} \cdot Y_k & c_{33(k)} \cdot Z_k \end{bmatrix} \quad (4.21)$$

Si se suponen variables incorreladas, la matriz  $C_k$  se reduce a:

$$C_k = \begin{bmatrix} \sigma_{X(k)}^2 & 0 & 0 \\ 0 & \sigma_{Y(k)}^2 & 0 \\ 0 & 0 & \sigma_{Z(k)}^2 \end{bmatrix} \quad (4.22)$$

y en consecuencia la ecuación (4.21) se convierte en

$$\sum_{k=i,j} \begin{bmatrix} \sigma_{X(k)}^2 \cdot X \\ \sigma_{Y(k)}^2 \cdot Y \\ \sigma_{Z(k)}^2 \cdot Z \end{bmatrix} = \sum_{k=i,j} \begin{bmatrix} \sigma_{X(k)}^2 \cdot X_k \\ \sigma_{Y(k)}^2 \cdot Y_k \\ \sigma_{Z(k)}^2 \cdot Z_k \end{bmatrix} \quad (4.23)$$

De (4.23) se obtienen las coordenadas del punto  $P$ , que vienen dadas por:

$$\begin{aligned} X &= \frac{\sigma_{X_i}^2 \cdot X_j + \sigma_{X_j}^2 \cdot X_i}{\sigma_{X_i}^2 + \sigma_{X_j}^2} \\ Y &= \frac{\sigma_{Y_i}^2 \cdot Y_j + \sigma_{Y_j}^2 \cdot Y_i}{\sigma_{Y_i}^2 + \sigma_{Y_j}^2} \\ Z &= \frac{\sigma_{Z_i}^2 \cdot Z_j + \sigma_{Z_j}^2 \cdot Z_i}{\sigma_{Z_i}^2 + \sigma_{Z_j}^2} \end{aligned} \quad (4.24)$$

Como se puede observar en (4.24), el valor de las coordenadas del punto  $P$  es función de los puntos de cruce de las rectas de proyección y del valor de sus varianzas  $\sigma_{X_i}^2$ ,  $\sigma_{Y_i}^2$ ,  $\sigma_{Z_i}^2$ ,  $\sigma_{X_j}^2$ ,  $\sigma_{Y_j}^2$  y  $\sigma_{Z_j}^2$ . En el caso de que éstas fueran iguales, las coordenadas del punto  $P$  se corresponden con el punto medio del segmento  $h$ .

Una vez obtenidos los valores de las coordenadas del punto  $P$ , en función de las varianzas de las coordenadas 3D, solamente queda expresar éstas en función de las varianzas de las coordenadas pixélicas. Se trata, en definitiva, de expresar las varianzas  $\sigma_{X(k)}^2$ ,  $\sigma_{Y(k)}^2$  y  $\sigma_{Z(k)}^2$  en función de las varianzas de las coordenadas  $(u_i, v_i)$  y  $(u_j, v_j)$ ,  $\sigma_{u_i}^2$ ,  $\sigma_{v_i}^2$ ,  $\sigma_{u_j}^2$  y  $\sigma_{v_j}^2$ . Para ello, generalizando las ecuaciones (4.7) y (4.8) se tiene:

$$P_k = \begin{bmatrix} X_k \\ Y_k \\ Z_k \end{bmatrix} = \lambda_{(k)} \cdot \begin{bmatrix} r_{11(k)} & r_{21(k)} & r_{31(k)} \\ r_{12(k)} & r_{22(k)} & r_{32(k)} \\ r_{13(k)} & r_{23(k)} & r_{33(k)} \end{bmatrix} \begin{bmatrix} \frac{u_k - u_{0(k)}}{f_{x(k)}} \\ \frac{v_k - v_{0(k)}}{f_{y(k)}} \\ 1 \end{bmatrix} + \begin{bmatrix} T'_{X(k)} \\ T'_{Y(k)} \\ T'_{Z(k)} \end{bmatrix} \quad (4.25)$$

y derivando en (4.25) respecto a  $\lambda_{(k)}$ ,  $u_k$  y  $v_k$  se obtiene:

$$\begin{aligned} \sigma_{X(k)}^2 &= J_{X(k)} \cdot \begin{bmatrix} \sigma_{\lambda(k)}^2 & 0 & 0 \\ 0 & \sigma_{u(k)}^2 & 0 \\ 0 & 0 & \sigma_{v(k)}^2 \end{bmatrix} \cdot J_{X(k)}^T \\ \sigma_{Y(k)}^2 &= J_{Y(k)} \cdot \begin{bmatrix} \sigma_{\lambda(k)}^2 & 0 & 0 \\ 0 & \sigma_{u(k)}^2 & 0 \\ 0 & 0 & \sigma_{v(k)}^2 \end{bmatrix} \cdot J_{Y(k)}^T \\ \sigma_{Z(k)}^2 &= J_{Z(k)} \cdot \begin{bmatrix} \sigma_{\lambda(k)}^2 & 0 & 0 \\ 0 & \sigma_{u(k)}^2 & 0 \\ 0 & 0 & \sigma_{v(k)}^2 \end{bmatrix} \cdot J_{Z(k)}^T \end{aligned} \quad (4.26)$$

donde  $\mathbf{J}_{X(k)}$ ,  $\mathbf{J}_{Y(k)}$ ,  $\mathbf{J}_{Z(k)}$  representan los Jacobianos de las coordenadas  $X_k$ ,  $Y_k$  y  $Z_k$ , que vienen dados por:

$$\begin{aligned} \mathbf{J}_{X(k)} &= \left[ -r_{11(k)} \frac{u_k - u_{0(k)}}{f_{x(k)}} - r_{22(k)} \frac{v_k - v_{0(k)}}{f_{y(k)}} - r_{31(k)} \quad , \quad -\frac{r_{11(k)}}{f_{x(k)}} \quad , \quad -\frac{r_{21(k)}}{f_{y(k)}} \right] \\ \mathbf{J}_{Y(k)} &= \left[ -r_{12(k)} \frac{u_k - u_{0(k)}}{f_{x(k)}} - r_{22(k)} \frac{v_k - v_{0(k)}}{f_{y(k)}} - r_{32(k)} \quad , \quad -\frac{r_{12(k)}}{f_{x(k)}} \quad , \quad -\frac{r_{22(k)}}{f_{y(k)}} \right] \\ \mathbf{J}_{Z(k)} &= \left[ -r_{13(k)} \frac{u_k - u_{0(k)}}{f_{x(k)}} - r_{23(k)} \frac{v_k - v_{0(k)}}{f_{y(k)}} - r_{33(k)} \quad , \quad -\frac{r_{13(k)}}{f_{x(k)}} \quad , \quad -\frac{r_{23(k)}}{f_{y(k)}} \right] \end{aligned} \quad (4.27)$$

A partir de las ecuaciones (4.26) y (4.27) se obtienen las varianzas de las coordenadas de los puntos  $P_i$ ,  $P_j$  en función de las varianzas de las coordenadas pixélicas ( $\sigma_{u_i}^2$ ,  $\sigma_{v_i}^2$ ,  $\sigma_{u_j}^2$  y  $\sigma_{v_j}^2$ ) y de las varianzas de  $\sigma_{\lambda_i}^2$  y  $\sigma_{\lambda_j}^2$ . Finalmente, para obtener los valores de  $\sigma_{\lambda_i}^2$  y  $\sigma_{\lambda_j}^2$  se deriva (4.9) respecto a  $u_i$ ,  $v_i$  y  $u_j$ ,  $v_j$ ,  $\lambda_i$  y  $\lambda_j$ , obteniéndose:

$$\begin{bmatrix} \sigma_{\lambda_i}^2 & \rho_{\lambda_i \lambda_j} \sigma_{\lambda_i} \sigma_{\lambda_j} \\ \rho_{\lambda_i \lambda_j} \sigma_{\lambda_i} \sigma_{\lambda_j} & \sigma_{\lambda_j}^2 \end{bmatrix} = \mathbf{J}_1^+ \cdot \mathbf{J}_2 \cdot \mathbf{C}_e \cdot \mathbf{J}_2^T \cdot \mathbf{J}_1^+ \quad (4.28)$$

donde:

$$\mathbf{J}_1 = \begin{bmatrix} \frac{u_i - u_{0i}}{f_{xi}} & - \left( r_{11} \frac{u_i - u_{0i}}{f_{xi}} + r_{12} \frac{v_i - v_{0i}}{f_{yi}} + r_{13} \right) \\ \frac{v_i - v_{0i}}{f_{yi}} & - \left( r_{21} \frac{u_i - u_{0i}}{f_{xi}} + r_{22} \frac{v_i - v_{0i}}{f_{yi}} + r_{23} \right) \\ 1 & - \left( r_{31} \frac{u_i - u_{0i}}{f_{xi}} + r_{32} \frac{v_i - v_{0i}}{f_{yi}} + r_{33} \right) \end{bmatrix} \quad (4.29)$$

$$\mathbf{J}_2 = \begin{bmatrix} \frac{\lambda_j}{f_{xi}} & -\lambda_j \frac{r_{11}}{f_{xj}} & -\lambda_j \frac{r_{12}}{f_{yj}} \\ 0 & -\lambda_j \frac{r_{21}}{f_{xj}} & -\lambda_j \frac{r_{22}}{f_{yj}} \\ 0 & -\lambda_j \frac{r_{31}}{f_{xj}} & -\lambda_j \frac{r_{32}}{f_{yj}} \end{bmatrix}$$

siendo  $r_{ij}$  los coeficientes de la matriz  $\mathbf{R}$  dada en (4.10). Por otra parte, la matriz  $\mathbf{J}_1^+$  representa la pseudoinversa de  $\mathbf{J}_1$ , cuyo valor viene dado por:

$$\mathbf{J}_1^+ = \mathbf{J}_1^T (\mathbf{J}_1 \cdot \mathbf{J}_1^T)^{-1} \quad (4.30)$$

Al ser  $\mathbf{J}_1 \cdot \mathbf{J}_1^T$  singular no es posible aplicar la expresión anterior para obtener la pseudoinversa de  $\mathbf{J}_1$ . Una posible solución a este problema consiste en descomponer en valores singulares (SVD) de la forma:  $\mathbf{J}_1 = \mathbf{Q} \cdot \mathbf{S} \cdot \mathbf{P}^T$ . Con esta descomposición el valor de la pseudoinversa viene dado por:

$$\mathbf{J}_1^+ = \mathbf{P} \cdot \mathbf{S}^+ \cdot \mathbf{Q}^T \quad (4.31)$$

donde  $\mathbf{S}^+$  es una matriz diagonal de 2x3 de la forma:

$$\mathbf{S}^+ = \begin{bmatrix} \frac{1}{\sqrt{a_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{a_2}} & 0 \end{bmatrix} \quad (4.32)$$

donde  $a_1$  y  $a_2$  son los autovalores de la matriz  $\mathbf{J}_1^T \cdot \mathbf{J}_1$ .

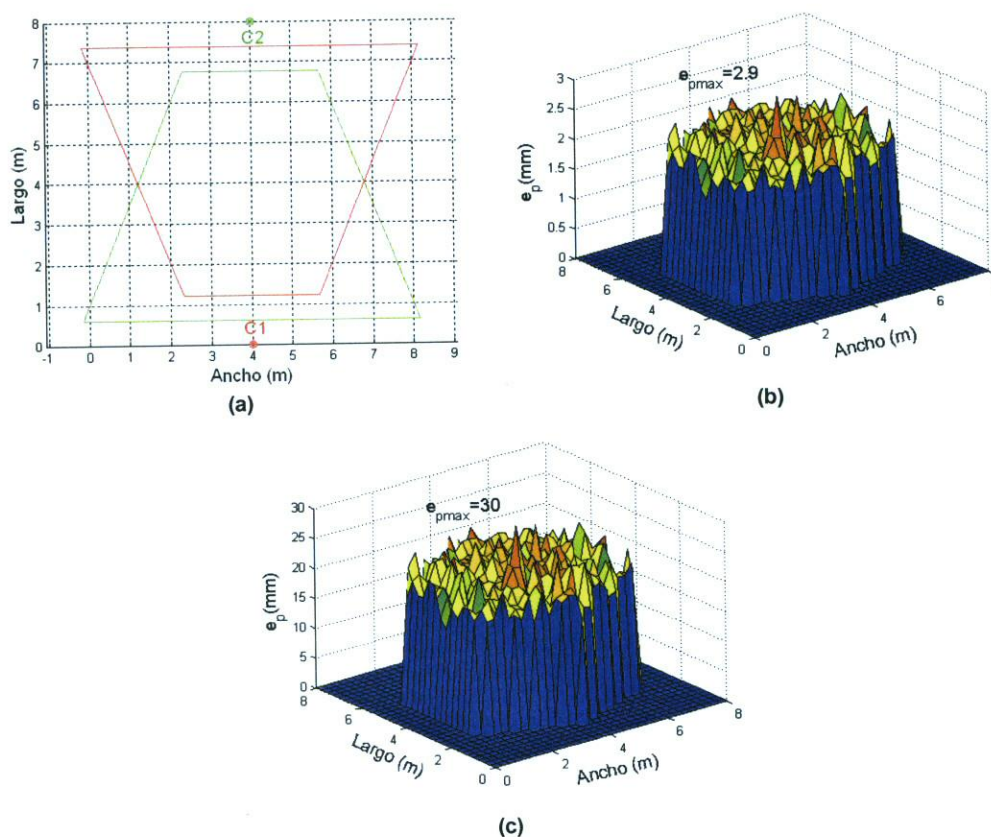
A partir de las ecuaciones (4.24), (4.26) y (4.28) se obtiene el valor de las coordenadas  $[X, Y, Z]^T$  del punto  $\mathbf{P}$ .

Para comprobar la eficacia del método propuesto se han realizado diversas simulaciones y pruebas reales del error de posicionamiento utilizando dos cámaras. Para ello, la cámara  $C_1$  se ubicó en la posición  $X=4$  m,  $Y=0$  m,  $Z=3$ ,  $\alpha=45^\circ$ ,  $\beta=0=\gamma$ , y la cámara  $C_2$  en  $X=4$  m,  $Y=0$  m,  $Z=3$ ,  $\alpha=45^\circ$ ,  $\beta=0$  y  $\gamma=180^\circ$ . En estas condiciones el campo de visión se muestra en la figura 4.11.a. Las coordenadas del centroide de la baliza se contaminaron con errores gaussianos de media nula y desviación típica  $\sigma_{u_i}$ ,  $\sigma_{v_i}$ ,  $\sigma_{u_j}$  y  $\sigma_{v_j}$ . En estas condiciones se obtuvieron las gráficas mostradas en las figuras 4.11.b y c, para un ruido con desviación típica 0.2 píxeles (figura 4.11.b), y 2 píxeles (figura 4.11.c).

Comparando las gráficas de la figura 4.11 con las de la figura 4.1, se puede observar que el error en la obtención de la posición de la baliza (robot) utilizando dos cámaras se reduce, aproximadamente, en un 70 % respecto a la utilización de una sola.

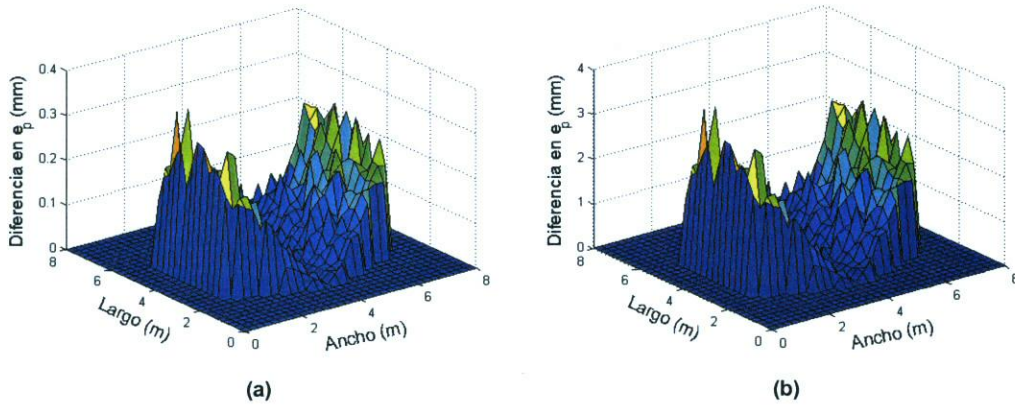
Los resultados obtenidos a partir de la propuesta realizada en esta tesis se han comparado con otras propuestas realizadas por otros investigadores, en concreto con las realizadas por [Aitenbichler 03] y [Hartley 03]. Los resultados de esta comparación se muestran en la figura 4.12, donde se observa la diferencia de error de posicionamiento entre la propuesta de estos

investigadores y la realizada en esta tesis; esta diferencia siempre es positiva y la mejora es del orden del 15 %.

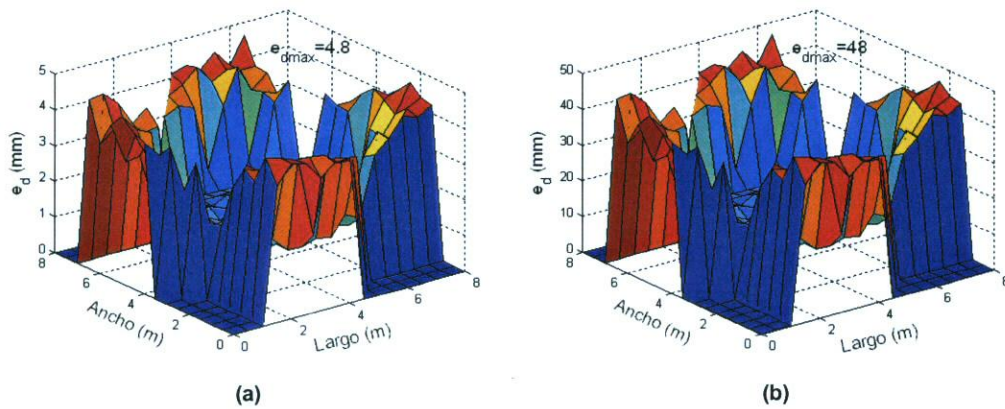


**Figura 4.11.** Ejemplos simulados del error de posicionamiento utilizando dos cámaras.. (a) Campo de visión de las cámaras. (b) Caso de contaminar las medidas con un ruido de media nula y desviación típica de 0.2 píxeles. (c) Ruido con desviación típica de 2 píxeles.

Para ver el efecto que tiene la distancia entre cámaras y robot se han realizado pruebas, utilizando el mismo sistema de cámaras que el empleado en la obtención de la figura 4.6, pero ahora eligiendo las dos cámaras más próximas al robot. En estas condiciones los resultados obtenidos se muestran en la figura 4.13, suponiendo que las coordenadas píxelicas se contaminan con un ruido gaussiano de media nula y desviación típica 0.2 píxeles (figura 4.13.a) y 2 píxeles (figura 4.13.b). Como se puede observar, para el primer caso, el error máximo es de 4.8 mm, mientras que para el segundo es de 48 mm. Al igual que sucedía en la figura 4.6, en los bordes del espacio de observación el valor del error de posicionamiento es elevado debido al hecho de que las cámaras empleadas para obtener la posición están muy alejadas de esos puntos.



**Figura 4.12.** Comparación entre el método propuesto en esta tesis y el propuesto por [Aitenbichler 03][Hartley 03]. (a) Para el caso de un ruido gaussiano de media cero y desviación típica de 0.2 píxeles. (b) ídem, para una desviación típica de 2 píxeles.



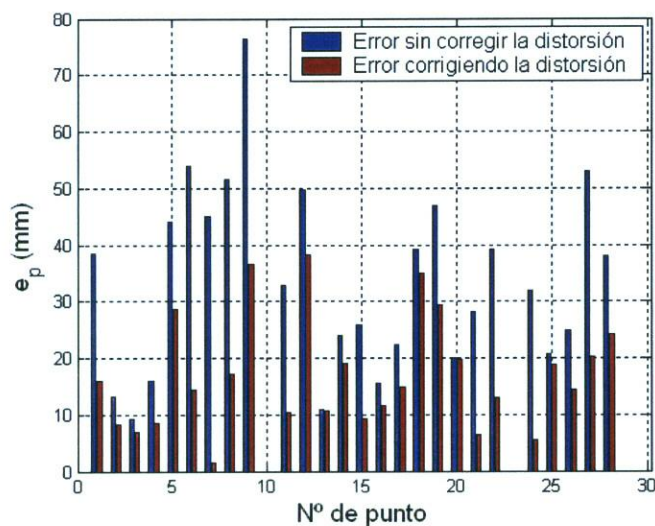
**Figura 4.13.** Error de posicionamiento de un sistema formado por seis cámaras, utilizando, para la obtención de la posición del robot, las dos cámaras más próximas a éste. (a) Con un error en las coordenadas píxelicas de desviación típica 0.2 píxeles. (b) Con un error en las coordenadas píxelicas de desviación típica 2 píxeles.

### 4.3.1. Resultados prácticos.

Para comprobar el algoritmo de posicionamiento utilizando dos cámaras, se hizo uso del sistema de la figura 4.8.c. En este caso para obtener la posición del robot se han utilizado las dos cámaras más próximas al él. En la figura 4.14 se muestran los resultados obtenidos. En la tabla 4.3. se muestran los valores de los estadísticos del error de posición para el ejemplo de la figura 4.14.

<b>TABLA 4.3.</b> Errores prácticos de posicionamiento utilizando dos cámaras.				
	Error mínimo (mm)	Error máximo (mm)	Error medio (mm)	Desviación típica (mm)
<b>Sin corrección de distorsión</b>	9	76	29	19
<b>Con corrección de distorsión</b>	1	38	14	10

Comparando los resultados de las tablas 4.1 y 4.2. con la tabla 4.3. se puede comprobar que al utilizar dos cámaras el error de posicionamiento se reduce prácticamente a la mitad.



**Figura 4.14.** Valores reales del error de posicionamiento para un entorno de 6x4 m sobre el que se han dispuesto cuatro cámaras, utilizando las dos más cercanas al robot para determinar su posición. Se reflejan los valores sin y con corrección de distorsión.

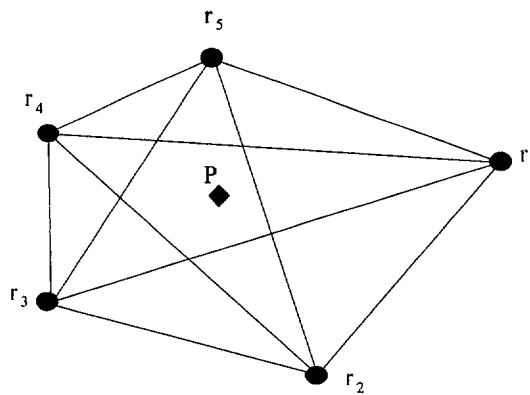
#### 4.4. Obtención de la posición del robot utilizando más de dos cámaras.

Una vez vistas las alternativas de obtención de la posición del robot utilizando una y dos cámaras (situación está última más frecuente en aplicaciones prácticas), en este apartado se presenta una generalización de la propuesta realizada en esta tesis para el caso de utilizar  $n$  cámaras en la determinación de la posición del robot. En el caso de que el robot sea visto

simultáneamente por  $n$  cámaras ( $n > 2$ ), existirán  $M$  posibles cruces entre las rectas de proyección de las diferentes cámaras, donde  $M$  viene dada por:

$$M = \binom{n}{2} = \frac{n(n-1)}{2} \tag{4.33}$$

Lo que se busca es el punto  $P$  (figura 4.15) cuya suma de distancias a cada una de las rectas de proyección sea mínima, y además tenga en cuenta las varianzas de las coordenadas pixélicas del centroide de la baliza para aproximar  $P$  al punto de mayor certidumbre.



**Figura 4.15.** Representación del punto óptimo  $P$  que minimiza la distancia a las rectas de proyección de  $n$  cámaras, ejemplo para  $n= 5$ , conseguido a partir de los posibles puntos obtenidos por el cruce entre dos rectas.

Generalizando para  $n$  cámaras el caso de dos, la función de coste a minimizar (ecuación 4.15) se convierte en:

$$E = \sum_{i=1}^{2M} (\mathbf{P} - \mathbf{P}_i)^T C_i (\mathbf{P} - \mathbf{P}_i) \tag{4.34}$$

y en consecuencia:

$$\frac{\partial E}{\partial \mathbf{P}} = \sum_{i=1}^{2M} C_i \cdot (\mathbf{P} - \mathbf{P}_i) = 0 \tag{4.35}$$

Teniendo en cuenta (4.18) y que las coordenadas  $[X_p, Y_p, Z_p]^T$  correspondientes al punto  $P_i$  están incorreladas, la matriz  $C_i$  es:



$$C_i = \begin{bmatrix} \sigma_{X(i)}^2 & 0 & 0 \\ 0 & \sigma_{Y(i)}^2 & 0 \\ 0 & 0 & \sigma_{Z(i)}^2 \end{bmatrix} \quad (4.36)$$

Llevando (4.36) a (4.35) se obtiene:

$$\sum_{i=1}^{2M} \begin{bmatrix} \sigma_{X(i)}^2 \cdot X \\ \sigma_{Y(i)}^2 \cdot Y \\ \sigma_{Z(i)}^2 \cdot Z \end{bmatrix} = \sum_{i=1}^{2M} \begin{bmatrix} \sigma_{X(i)}^2 \cdot X_i \\ \sigma_{Y(i)}^2 \cdot Y_i \\ \sigma_{Z(i)}^2 \cdot Z_i \end{bmatrix} \quad (4.37)$$

donde  $[X_p, Y_p, Z_p]^T$  representan las coordenadas de los  $2M$  puntos, obtenidas a partir de la ecuación (4.13). De (4.37) se determinan las coordenadas 3D del punto  $P$ , las cuales vienen dadas por:

$$X = \frac{\sum_{i=1}^{2M} X_i \cdot \sigma_{Xi}^2}{\sum_{i=1}^{2M} \sigma_{Xi}^2}; \quad Y = \frac{\sum_{i=1}^{2M} Y_i \cdot \sigma_{Yi}^2}{\sum_{i=1}^{2M} \sigma_{Yi}^2}; \quad Z = \frac{\sum_{i=1}^{2M} Z_i \cdot \sigma_{Zi}^2}{\sum_{i=1}^{2M} \sigma_{Zi}^2} \quad (4.38)$$

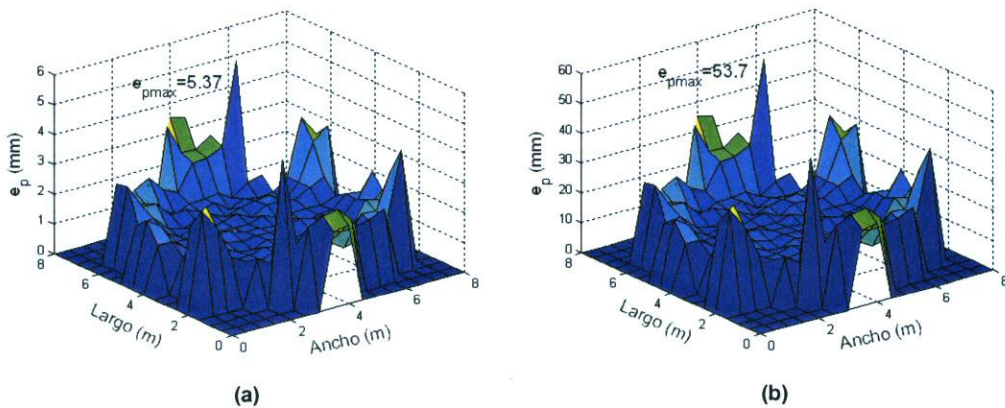
Si en la ecuación (4.38) todas las varianzas  $\sigma_{Xi}^2$ ,  $\sigma_{Yi}^2$  y  $\sigma_{Zi}^2$  se consideran iguales, el punto  $P$  correspondería al centroide de todos los punto de cruce de las rectas de proyección:

$$X = \frac{\sum_{i=1}^{2M} X_i}{2M}; \quad Y = \frac{\sum_{i=1}^{2M} Y_i}{2M}; \quad Z = \frac{\sum_{i=1}^{2M} Z_i}{2M} \quad (4.39)$$

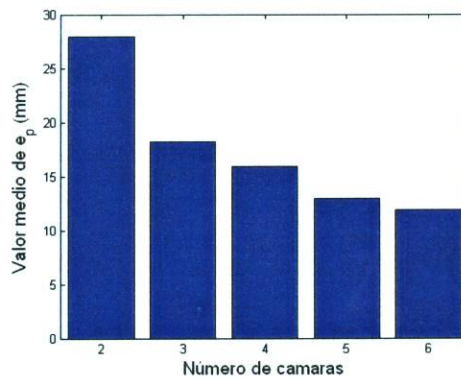
En la figura 4.16 se muestra el valor del error de posicionamiento cometido en un entorno rectangular de 8x8m, sobre el que se han ubicado seis cámaras equiespaciadas por su periferia. En cada zona del entorno se han utilizado, para la obtención de la posición del robot, todas las cámaras que captan dicha zona. Al igual que en apartados anteriores, las gráficas se han realizado para dos valores de desviación típica en las coordenadas pixélicas: 0.2 píxeles (figura 4.16.a) y 2 píxeles (figura 4.16.b).

De la figura 4.16 se puede concluir que cuanto mayor sea del número de cámaras utilizadas para determinar la posición del robot, menor es el error en la determinación de su posición. Para ver la influencia que tiene en el error de posición el número de cámaras que se utilizan,

se han realizado experimentos ubicando el robot (baliza) en diferentes puntos del entorno y suponiendo diferentes valores de errores en la determinación de las coordenadas pixélicas. A modo de ejemplo, en la figura 4.17 se muestra la dependencia entre el número de cámaras que se utilizan y el error que se comete en la determinación de la posición del robot, suponiendo una desviación típica de las coordenadas pixélicas de 2 píxeles. Como se puede observar en la figura 4.17 hay una mejora sustancial (aproximadamente 1cm) cuando se utilizan tres cámaras frente al uso de dos. También se puede observar que para un número de cámaras superior a tres las mejoras no son tan significativas.



**Figura 4.16.** Representación del error de posicionamiento de un entorno rectangular de 8x8m en el que se han ubicado seis cámaras. La posición del robot se obtiene utilizando todas las cámaras que en cada punto lo captan. (a) Desviación típica de 0.2 píxeles. (b) Desviación típica de 2 píxeles.

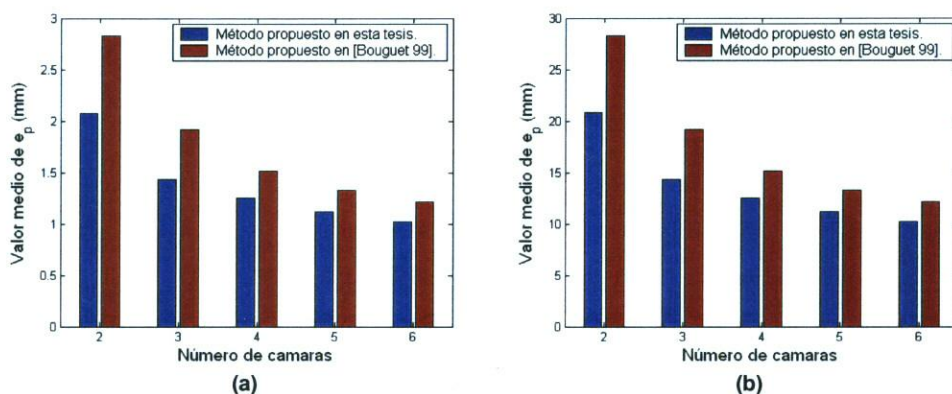


**Figura 4.17.** Ejemplo de estudio comparativo del error de posicionamiento, en función del número de cámara utilizadas para determinar la posición del robot.

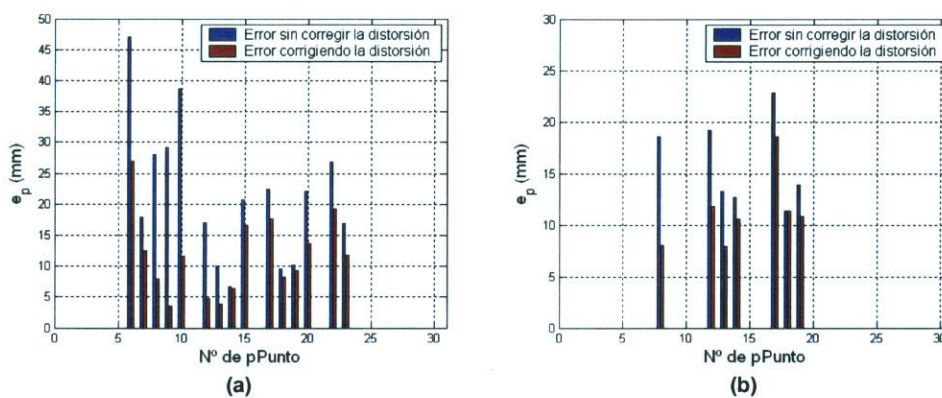
Para comprobar la eficacia del algoritmo de posicionamiento propuesto en esta tesis, se han realizado también comparaciones con las propuestas realizadas por otros autores, en concreto con la propuesta realizada en [Bouguet 99]. En la figura 4.18 se muestra esta comparativa. Como se puede comprobar la propuesta realizada en esta tesis mejora en un 10% la propuesta realizada por [Bouguet 99].

### 4.4.1. Resultados prácticos.

En la figura 4.19 se muestran los resultados reales del error de posicionamiento obtenido utilizando la propuesta realizada en esta tesis. Los resultados reflejados corresponden a los casos de utilización de tres (figura 4.19.a) y cuatro cámaras (figura 4.19.b).



**Figura 4.18.** Comparación entre los resultados de posicionamiento obtenidos utilizando la propuesta realizada en esta tesis y la realizada en [Bouguet 99]. Se representan los valores del error medio de posicionamiento para el caso de utilizar dos, tres, cuatro, cinco y seis cámaras. (a) Para una varianza de 0.2 píxeles. (b) Para una varianza de 2 píxeles.

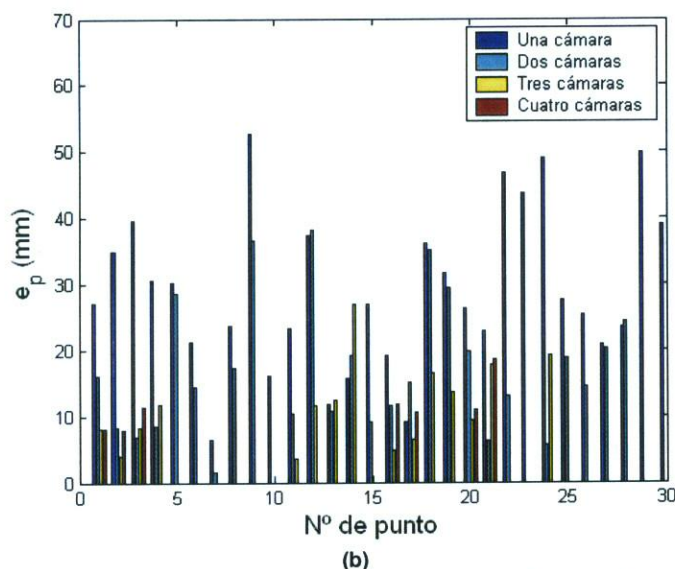
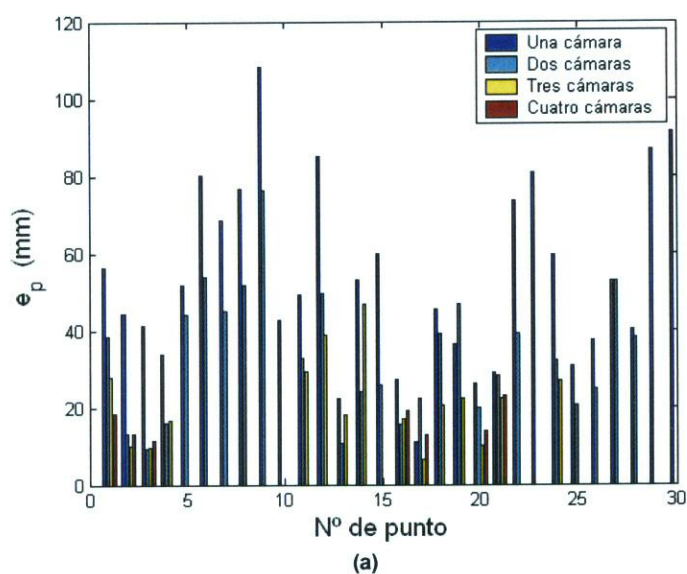


**Figura 4.19.** Valores prácticos del error de posicionamiento para el caso de utilizar más de dos cámaras. (a) tres cámaras. (b) cuatro cámaras.

Como resumen final, en las gráficas de la figura 4.20 se muestra el error de posicionamiento dependiendo del número de cámaras utilizadas, tanto para el caso de no realizar corrección de distorsión (figura 4.20.a), como para el caso de sí realizar corrección (figura 4.20.b). En la tabla 4.4 se muestran los valores numéricos de los errores de la figura 4.20.b

**TABLA 4.4.** Errores de posicionamiento prácticos para un sistema de visión formado por cuatro cámaras, realizando la corrección de la distorsión.

	Número de cámaras			
	1	2	3	4
<b>Error medio (mm)</b>	28.9	16.8	11.6	11.3
<b>Desviación típica (mm)</b>	12.1	9.9	6.4	3.5



**Figura 4.20.** Resultados de los errores de posicionamiento de un robot en función del número de cámaras utilizadas para calcular su posición. (a) Sin realizar la corrección de la distorsión. (b) Corrigiendo la distorsión de las ópticas.

## 4.5. Conclusiones.

En este capítulo se ha propuesto un sistema de posicionamiento de un robot móvil utilizando  $n$  cámaras distribuidas por su entorno de movimiento. El algoritmo propuesto permite obtener la posición del robot a partir de los centroides de las balizas de a bordo y teniendo presente los errores cometidos en la determinación de las coordenadas de dichos centroides. La propuesta realizada permite obtener dicha posición a partir de una única cámara (suponiendo que el robot se mueva por un plano y la altura del plano en el que se encuentra la baliza es conocida), dos o más cámaras. Demostrando la precisión obtenida en la posición en función del número de cámaras que en cada instante perciben información del robot.

Para el caso más frecuente de que el robot sea captado en cada instante por más de una cámara, se ha propuesto un algoritmo que minimiza el error de posicionamiento y que ha demostrado proporcionar mejores resultados que las propuestas realizadas por otros autores. El algoritmo propuesto se basa en minimizar una función de coste, teniendo presente que las coordenadas del punto en el espacio 3D se encontrará más próximo a la recta de proyección de la cámara en la que los errores son menores. Por ello, la función de coste a minimizar tiene un gran parecido con la distancia de Mahalanobis, ya que tienen en cuenta las varianzas de las medidas obtenidas por las diferentes cámaras, asignado mayor protagonismo a las medidas más fiables.

Las simulaciones y pruebas prácticas realizadas ponen de manifiesto la viabilidad de la propuesta realizada en esta tesis, siendo los errores de posicionamiento perfectamente asumibles para la navegación de robots en espacios interiores, utilizando una distribución de cámaras cuyo coste es aceptable en la mayor parte de las aplicaciones prácticas

En el caso de que el robot sea captado por una única cámara, suponiendo que el robot se mueve sobre una superficie plana y que la altura a la que se encuentra la baliza ( $Z = Z_0$ ) es conocida, los errores de posicionamiento obtenidos son aceptables dentro de un amplio rango de distancia entre cámara y robot. Este error es del orden de los 30 mm para distancias robot-cámara del orden de los 6 m, aumentando a medida que esta distancia aumenta. En consecuencia, para un sistema de  $n$  cámaras en el que solamente una de ellas capta en cada instante al robot, será necesario garantizar que la máxima distancia entre cámara y robot no supere los límites de error prefijados. Otro aspecto a tener presente en este caso es la sensibilidad del error de posicionamiento a la variación en la altura de la baliza ( $Z_0$ ).

En este capítulo se ha puesto de manifiesto que una de las mejores soluciones para obtener la

posición del robot consiste en utilizar en todo momento la información proporcionada por dos cámaras. En este caso se ha demostrado que se pueden obtener las tres coordenadas ( $X, Y, Z$ ) de la posición del robot con independencia de la superficie por la que se mueva, siendo la precisión obtenida muy superior al caso de utilizar una única cámara. Evidentemente, si se aumenta el número de cámaras utilizadas en cada instante para obtener la posición, también se reduce el error de posicionamiento, si bien esta mejora no justifica, en aplicaciones prácticas, el coste añadido que supone la utilización simultánea de más de dos cámaras. Por tanto, se puede concluir que empleando dos cámaras para determinar la posición del robot se consigue una precisión alta y a un coste reducido; en consecuencia, la propuesta realizada consiste en distribuir las cámaras por el entorno de forma que cada punto del mismo sea captado por, al menos, dos.

Como aportación más destacada de este capítulo se pueden citar los algoritmos de posicionamiento utilizando dos o más cámaras, en función de la desviación típica del error en la determinación de las coordenadas del centroide de la baliza en las diferentes cámaras del sistema de visión. En este sentido se realiza una comparación con las propuestas de otros autores, reflejando en todos los casos, que la aquí propuesta mejora notablemente los resultados de posicionamiento obtenidos. Todo ello se ha avalado con resultados simulados y pruebas reales en diferentes condiciones de entornos reales.



## Capítulo 5

---

### Balizas de IR y procesamiento de imágenes para su identificación

---

Uno de los inconveniente que presenta la utilización de la visión en el posicionamiento de robots móviles es la gran cantidad de información que, en general, se debe procesar para obtener los datos de interés. Dependiendo de los objetivos que se persigan, estas dificultades pueden ser mayores o menores, llegando en algunos casos a la necesidad de tener que utilizar un hardware muy específico [Paz 97] para poder trabajar en tiempo real. Esta dificultad se ve incrementada a medida que se incrementa el número de cámaras utilizadas en el proceso; así, por ejemplo, en el caso de utilizar un array de cámaras fijas la información a procesar aumenta de forma considerable. En aras de reducir la complejidad de la detección del robot y el proceso de obtención de la información precisa que permita obtener, a partir de ella, su posición 3D, se pueden utilizar diferentes estrategias. En esta tesis se propone, como ya se ha comentado anteriormente, la utilización de unas balizas de infrarrojos a bordo del robot. La utilización de dichas balizas facilita, en gran medida, todas las tareas de procesamiento de imágenes y reduce la carga computacional. Es evidente que para la detección del robot se pueden utilizar diferentes alternativas y que ello constituye una interesante línea de investigación futura; en este sentido se puede plantear la detección de marcas naturales de la propia estructura de robot (esquinas, por



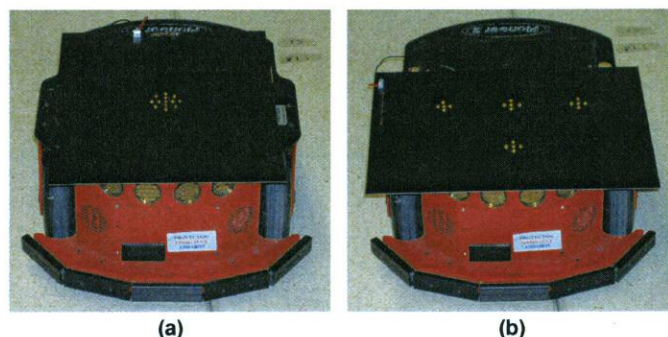
ejemplo). Sin embargo, y dado que el objetivo fundamental de esta tesis es el de demostrar la viabilidad de realizar el guiado a partir de la detección de algunos puntos, se ha simplificado la detección de éstos al máximo, con ayuda de las balizas de IR a bordo. En este capítulo se van a presentar, por tanto, las características de las balizas de IR y las técnicas de procesamiento utilizadas para su detección por las diferentes cámaras ubicadas en el entorno de movimiento del robot.

## 5.1. La baliza de Infrarrojos (IR).

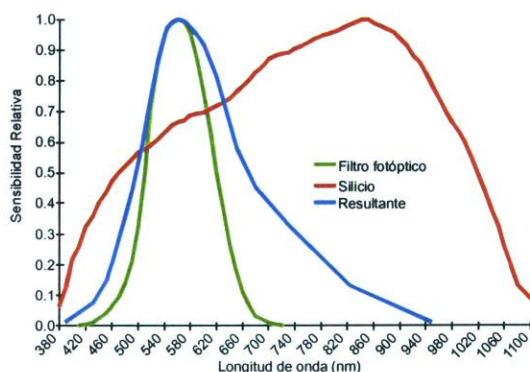
Para detectar e identificar los robots utilizando visión computacional, los investigadores que trabajan en este campo han propuesto diferentes alternativas de elementos identificadores (balizas) ubicados sobre los robots. El tipo y la forma de estas balizas pueden ser diferentes: código de barras de colores [Lee 01], círculos en blanco y negro [McNamee 03], etc. En todos estos casos, los sistemas propuestos son sencillos pero al mismo tiempo son muy sensibles a las condiciones de iluminación del entorno. Y si bien en entornos interiores estas condiciones no tiene porque ser muy cambiantes, pueden presentarse situaciones en las que, por diferentes razones, esta condición no se cumpla, como por ejemplo las originadas por los cambios de la iluminación artificial, o los cambios debidos a la propia iluminación natural cuando el robot se mueva por entornos con ventanas, cristaleras, etc. Por otra parte, la identificación de las balizas se ve dificultada a medida que aumenta su distancia con la cámara. En muchas aplicaciones, sobre todo para entornos de grandes dimensiones, las restricciones impuestas por la distancia entre balizas y cámaras son más importantes que las impuestas por las condiciones de iluminación. Por todo ello, en esta tesis se propone la utilización de una baliza formada por diodos infrarrojos (IR) de media potencia, que presenta la ventaja de ser fácilmente detectada, incluso en condiciones ambientales de cambios importantes de iluminación y facilita en gran medida los algoritmos de procesamiento de imágenes a utilizar. Además, esta alternativa es muy sencilla de realizar en la práctica y con costes de implementación y mantenimiento muy bajos. En lo que se refiere a la configuración geométrica y número de balizas a utilizar, como ya se comentó en el capítulo 3, el número de balizas ubicadas sobre el robot puede ser una (figura 5.1.a) o un número superior, por ejemplo cuatro (figura 5.1.b), si lo que se pretende, además de detectar al robot, es conocer su orientación a partir directamente de dichas balizas.

Otro aspecto importante de las balizas de IR es la longitud de onda de emisión, que como es lógico estará condicionado por las características de respuesta espectral de las cámaras a utilizar. En este sentido el silicio y en consecuencia los CCD's tienen una respuesta espectral mucho más

amplia que la del ojo humano (curva roja de la figura 5.2.), y presenta su mayor sensibilidad en el infrarrojo cercano (longitudes de onda comprendidas entre 700 nm y 5  $\mu\text{m}$ ). En general su respuesta máxima se encuentra en el margen de 820-900 nm. Sin embargo, si se pretenden utilizar cámaras comerciales hay que tener presente que los fabricantes tratan de adaptar la respuesta de la cámara a la curva de sensibilidad del ojo humano y para ello utilizan diferentes estrategias, como la de incorporar un filtro fotóptico (curva de color verde de la figura 5.2). Este filtro fotóptico, ubicado entre la óptica y el CCD, hace que la respuesta espectral de las cámaras sea más estrecha (curva azul en la figura 5.2). Al trabajar con balizas de emisores IR es necesario eliminar dicho filtro, y de esta forma la baliza se va a percibir en las imágenes como un objeto con un nivel de intensidad muy alto (próximo al nivel de 255, para una imagen en blanco y negro) y con una baja dependencia de la distancia entre cámaras y balizas.

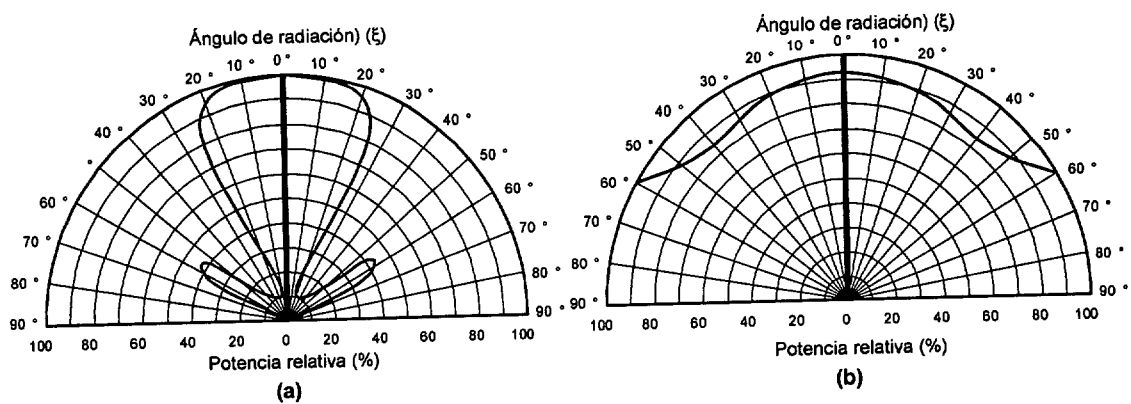


**Figura 5.1.** Balizas utilizadas a bordo del robot para facilitar su identificación. (a) Caso de utilización de una única baliza. (b) Caso de utilización cuatro balizas, que facilita la identificación del robot y su orientación.



**Figura 5.2.** Respuestas espectrales de: cámara CCD comercial (línea de color azul), filtro fotóptico ubicado entre el CCD y la óptica (línea de color verde) y silicio (línea de color rojo). Esta última alternativa es la utilizada en esta tesis.

Otra de las características importantes a tener en cuenta en los diodos emisores de infrarrojos utilizados para diseñar las balizas, es que tengan un patrón de emisión amplio, tal como el mostrado en la figura 5.3.b, que facilite la visualización de las balizas para diferentes ángulos de visión entre cámaras y robot. Evidentemente diodos con un ángulo de emisión más estrecho (figura 5.3.a) serán más fáciles de detectar pero sólo dentro de un ángulo de visión más pequeño. En este último caso, para que una cámara pueda ver la baliza en cualquier posición de su campo de visión, el eje de radiación debe ser, en todo momento, prácticamente paralelo al eje óptico de la cámara.



**Figura 5.3.** Patrón de radiación de dos diodos IR. (a) Diodo muy directivo. (b) Diodo con una radiación casi plana.

Eligiendo unos diodos IR cuya respuesta espectral se encuentre centrada en la zona de máxima respuesta del CCD y con un patrón de radiación lo más plano posible, se pueden diseñar balizas cuya segmentación se vea facilitada en gran medida, ya que dichas balizas se perciben en las diferentes cámaras como un punto luminoso de alta energía, siendo su valor, para una imagen en blanco y negro, próximo a 255, y todo ello con independencia, en gran medida, de las posiciones relativas de robots y cámaras.

### 5.1.1 Diseño práctico de las balizas.

Para las pruebas prácticas realizadas en esta tesis se han construido balizas utilizando diodos de IR cuyo patrón de emisión se encuentra centrado en 890 nm y se han utilizado un total de cinco diodos por baliza distribuidos dentro de un círculo de radio 2 cm (figura 5.4).

A modo de ejemplo, en la figura 5.5 se muestra uno de los robots utilizados en nuestros experimentos con una de estas balizas a bordo y para diferentes distancias entre robot y una de las cámaras: la figura 5.5.a se corresponde con una distancia de 4 m, y figura 5.5.c, para una

distancia de 10 m. En las figuras 5.5.b y 5.5.d se muestran los niveles de gris de la imagen del robot captada por la cámara para las dos situaciones de las figuras 5.5.a y 5.5.c. Como se puede observar los puntos correspondiente a la baliza presenta un nivel gris próximo a 255, en ambas situaciones.

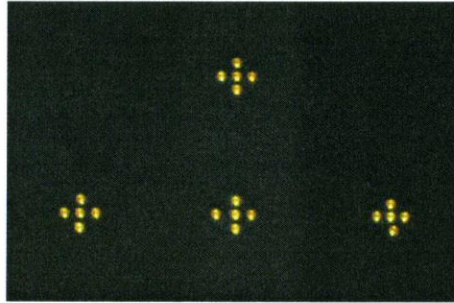


Figura 5.4. Vista detalla de las cuatro balizas de diodos IR.

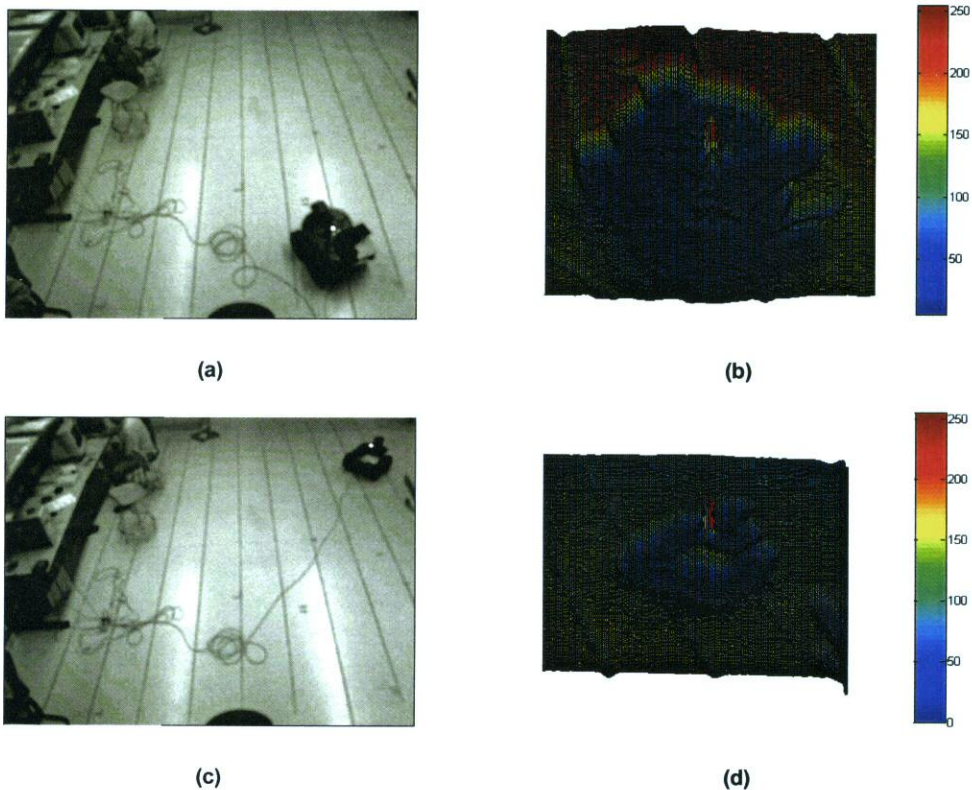
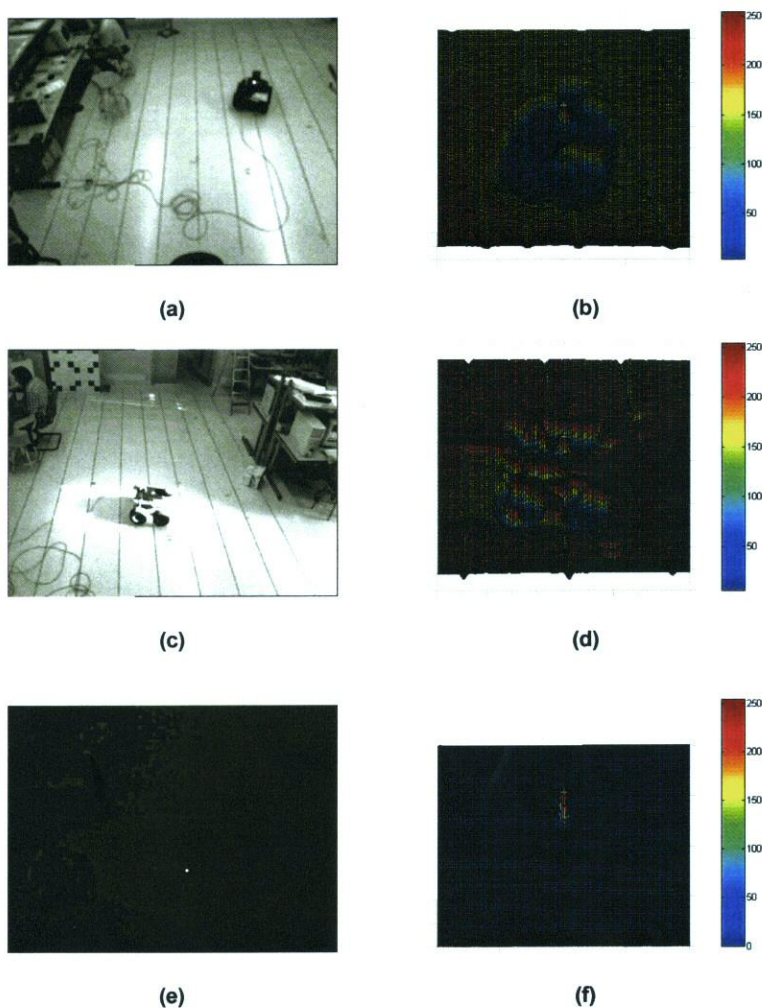


Figura 5.5. Ejemplo de captación de la baliza de IR a bordo del robot para dos situaciones de distancia entre robot y cámara. (a) Robot situado a 4 m de la cámara. (b) Niveles de gris de los píxeles de la imagen captada. (c) Robot situado a 10 m de la cámara. (d) Niveles de gris de los píxeles de la imagen correspondiente.

Para ver la robustez de este tipo de baliza y su detección frente a cambios de iluminación, en las imágenes de la figura 5.6 se muestra al robot ante tres condiciones de iluminación muy dispares: en condiciones normales de iluminación artificial (figura 5.6.a), con incidencia directa del sol (figura 5.6.c) sobre la baliza, y en ausencia de iluminación (figura 5.6.e). En las figuras 5.6.b, 5.6.d y 5.6.f se muestran los niveles de grises correspondientes a cada una de las imágenes anteriores. Como se puede observar la situación más crítica es la de la figura 5.6.c, y a pesar de ello la baliza no se ve enmascarada por la iluminación del sol. Evidentemente la situación más favorable es la de la figura 5.6.e, donde la identificación de la baliza se realiza de forma muy sencilla.



**Figura 5.6.** Ejemplos de imágenes captadas del robot ante diferentes condiciones de iluminación del entorno. (a) Iluminación artificial. (b) Niveles de grises correspondientes a la figura (a). (c) Presencia de reflejos debidos al sol. (d) Niveles de grises para la imagen (c). (e) Ausencia de iluminación. (f) Niveles de grises para la imagen (e).

Otro problema importante en la detección de las balizas es la presencia de brillos. Como es sabido, un rayo de luz consta de un haz formado por infinitos planos radiales (radiación vibrando con la misma probabilidad en todas las direcciones perpendiculares a la dirección de propagación), de forma que cuando se refleja sobre una superficie que no sea metálica, produce una radiación (brillo) parcialmente polarizada, la cual vibra en una sola dirección. El plano en el que vibra la luz se le conoce como plano de polarización. El efecto de estos brillos en nuestra aplicación, puede provocar un enmascaramiento de la baliza (figura 5.7.a). Para reducir el efecto de los brillos se pueden utilizar diferentes estrategias (estrategias software por medio de filtros homomórficos, polarizadores, etc), siendo una de las más efectivas la utilización de polarizadores lineales ubicados sobre la óptica (figura 5.7.b). Un polarizador actúa como una rejilla que permite eliminar la luz que vibra en un plano de polarización perpendicular a él, eliminando el brillo casi por completo (figura 5.7.c).



(a)



(b)



(c)

**Figura 5.7.** Reducción del efecto de los brillos utilizando polarizadores. (a) Enmascaramiento de la baliza por un brillo. (b) Cámara, óptica y polarizador utilizados en nuestra aplicación. (c) Imagen de la figura (a) tomada con un cámara cuya óptica incorpora un polarizador lineal.

## 5.2. Procesamiento de imágenes para la detección de las balizas a bordo del robot.

En este punto se van a presentar las diferentes estrategias de procesamiento de imágenes utilizadas para realizar una detección fiable y robusta de las balizas de IR a bordo del robot. Como ya se ha comentado, el hecho de utilizar balizas de IR facilita en gran medida los algoritmos de procesamiento de imágenes y reduce drásticamente el tiempo de procesamiento, eliminando de esta forma el “cuello de botella” que en muchas aplicaciones suponen el procesamiento de imágenes y máxime si éstos deben aplicarse sobre un array de cámaras. En nuestro caso, el algoritmo propuesto para la localización de las balizas incluye tres fases que de forma resumida se ilustran en la figura 5.8:

- Selección de la ventana de interés en función de la predicción de la posición del robot.
- Proceso de segmentación, que permite aislar las balizas del resto de la imagen.
- Proceso de extracción de características (perímetro, área, circularidad, etc), a partir de las cuales se pueden descartar algunos de los falsos candidatos a balizas (zonas de la imagen con niveles de intensidad similares al de las balizas).
- Proceso de clasificación, que permite identificar las balizas.

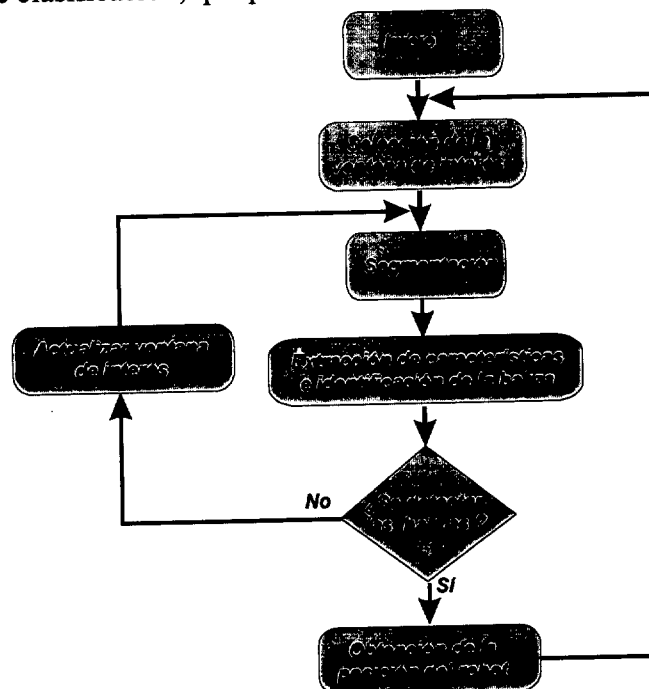


Figura 5.8. Proceso general de localización de las balizas.

Como se puede observar en la figura 5.8 la estructura general del algoritmo de procesamiento de imágenes es muy similar al de otras aplicaciones de visión por computador. Hay que indicar que en nuestro caso se han eliminado aspectos relacionados con el preprocesamiento, ya que tras diversas pruebas se ha podido comprobar que aspectos tales como filtrado, ecualización, etc., no introducían mejoras significativas.

Evidentemente, las tareas de segmentación, extracción de características y clasificación se ven facilitadas dado que entre dos imágenes consecutivas se puede hacer una predicción de la posición que va a ocupar la baliza, y centrar el algoritmo de procesamiento en una ventana (ventana de interés) entorno a la posición predicha en el espacio 3D ( en el capítulo 7 se describe el proceso de predicción utilizando un Filtro de Kalman).

Obviamente, cuando se inicia por primera vez el sistema la ventana de interés abarcará toda la imagen, pero tras este proceso inicial el tamaño de la misma se va variando en función de la predicción de posición realizada. En nuestro caso el tamaño mínimo de la ventana de interés viene determinada por el rectángulo que contiene todas las balizas, tal como se muestra en la figura 5.9, donde  $(\hat{u}_A, \hat{v}_A)$ ,  $(\hat{u}_B, \hat{v}_B)$ ,  $(\hat{u}_C, \hat{v}_C)$  y  $(\hat{u}_D, \hat{v}_D)$  son los valores estimados de las coordenadas en el plano imagen de los centroides de las balizas. Estas coordenadas se obtienen a partir de las estimadas en el espacio 3D:

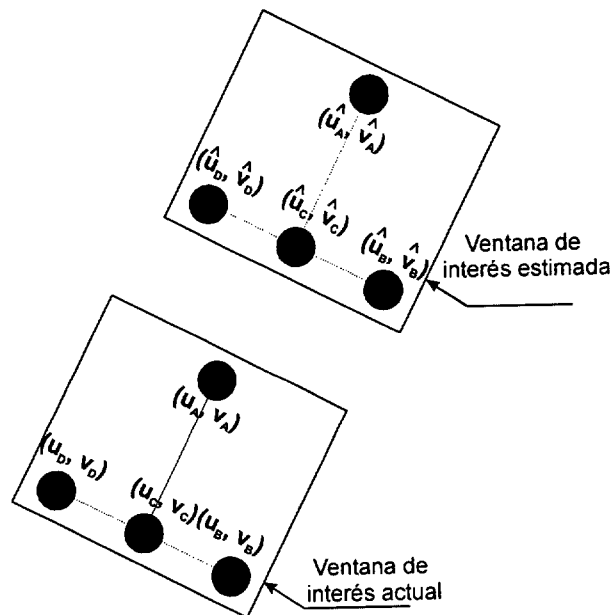
$$\begin{aligned}\hat{u}_i &= \frac{m_{11} \cdot \hat{X}_i + m_{12} \cdot \hat{Y}_i + m_{13} \cdot \hat{Z}_i + m_{14}}{m_{31} \cdot \hat{X}_i + m_{32} \cdot \hat{Y}_i + m_{33} \cdot \hat{Z}_i + m_{34}} \\ \hat{v}_i &= \frac{m_{21} \cdot \hat{X}_i + m_{22} \cdot \hat{Y}_i + m_{23} \cdot \hat{Z}_i + m_{24}}{m_{31} \cdot \hat{X}_i + m_{32} \cdot \hat{Y}_i + m_{33} \cdot \hat{Z}_i + m_{34}}\end{aligned}\quad (5.1)$$

Si utilizando la ventana de interés de mínima dimensión no se detectan las balizas, se procede a incrementar las dimensiones de dicha ventana. En nuestro caso se propone duplicar el tamaño de los lados en cada iteración, hasta alcanzar las dimensiones máximas de la imagen.

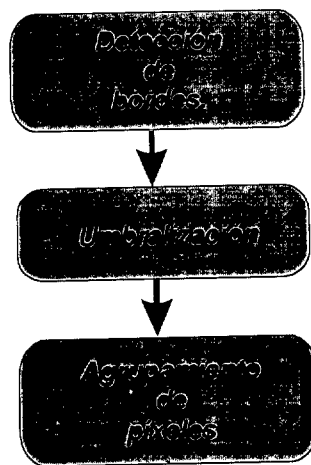
### 5.2.1. Segmentación.

Por segmentación se entiende el proceso mediante el cual una imagen es dividida en regiones inconexas con el propósito de separar las zonas de interés del resto de la imagen. En nuestro caso, las zonas de interés serán las balizas de IR a bordo del robot y el proceso de segmentación propuesto para este fin se muestra en la figura 5.10.





**Figura 5.9.** Ventana de interés sobre la que se realiza el proceso de segmentación.



**Figura 5.10.** Etapas incluidas en la propuesta de segmentación de las balizas de IR.

Como se puede observar, en nuestra propuesta el proceso de segmentación se inicia con la detección de bordes de los objetos presentes en la ventana de interés de la imagen procesada. Esta propuesta se justifica por el hecho de que las balizas de IR presentan un nivel de gris muy elevado, comparado con el resto de la imagen, y por tanto se pretende resaltar aún más dichos bordes con el objeto de magnificarlos, lo que facilita su posterior segmentación. Como la

presencia de la baliza en la imagen supone la presencia de unos bordes bastante definidos, para realzarlos aún más basta con realizar un filtrado paso alto, y para ello se han probado diferentes máscaras (figura 5.11) basadas en el operador gradiente:

$$|\nabla I(u, v)| = |G_u| + |G_v| \quad (5.2)$$

con:

$$\begin{aligned} G_u &= I(u, v) \otimes H_u(u, v) \\ G_v &= I(u, v) \otimes H_v(u, v) \end{aligned} \quad (5.3)$$

siendo  $\otimes$  el operador convolución.

Resultando que la solución más efectiva es utilizar máscaras de 3x3 tipo Prewitt:

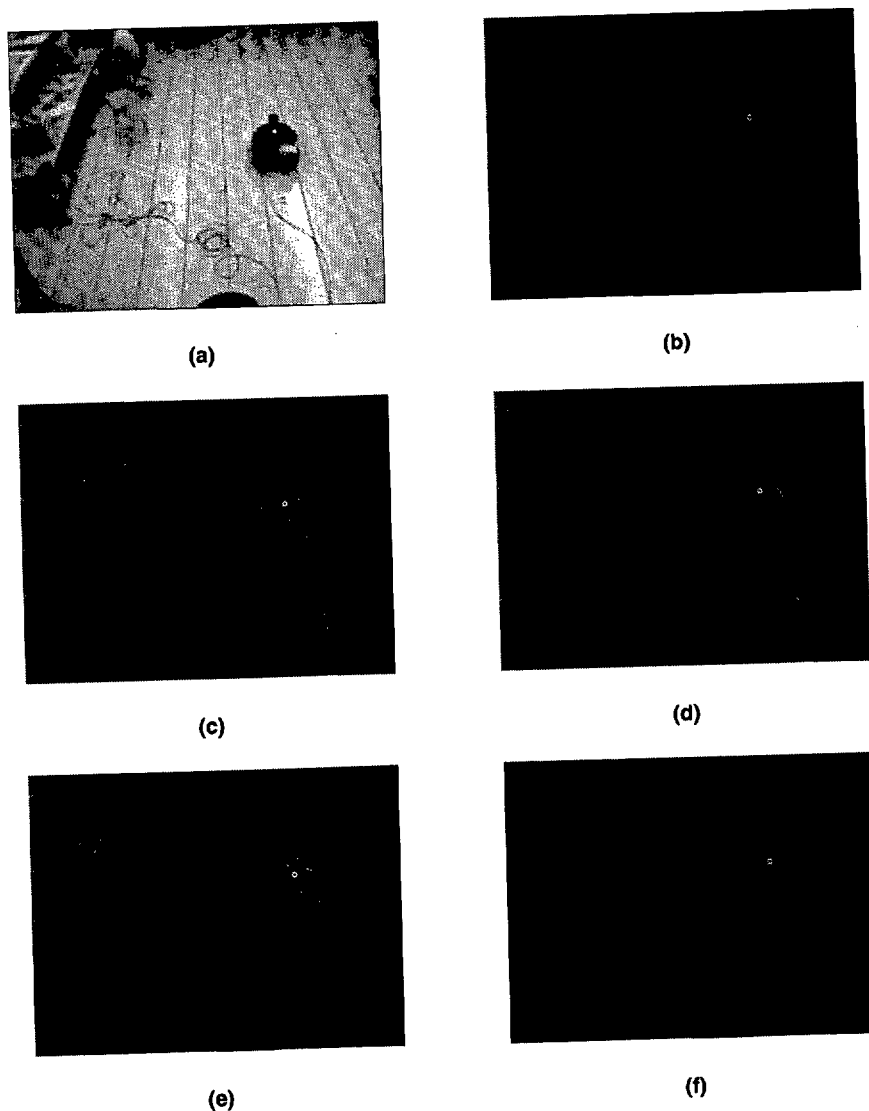
$$H_u = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}; \quad H_v = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (5.4)$$

En la figura 5.12 se muestra el resultado de aplicar el operador de Prewitt a las imágenes de la figura 5.6. El nivel de gris correspondiente al borde de un objeto, será tanto mayor cuanto mayor sea el contraste entre él y el fondo. Así, para la figura 5.6.a la baliza es el objeto que presenta mayor contraste en la imagen, en consecuencia, su borde tendrá un nivel de gris superior al borde de otros objetos. No sucede lo mismo para el caso de la figura 5.6.c, aquí el contraste entre el robot y el fondo es el mismo que tiene la baliza, teniendo, por tanto, sus borde un nivel de gris similar.

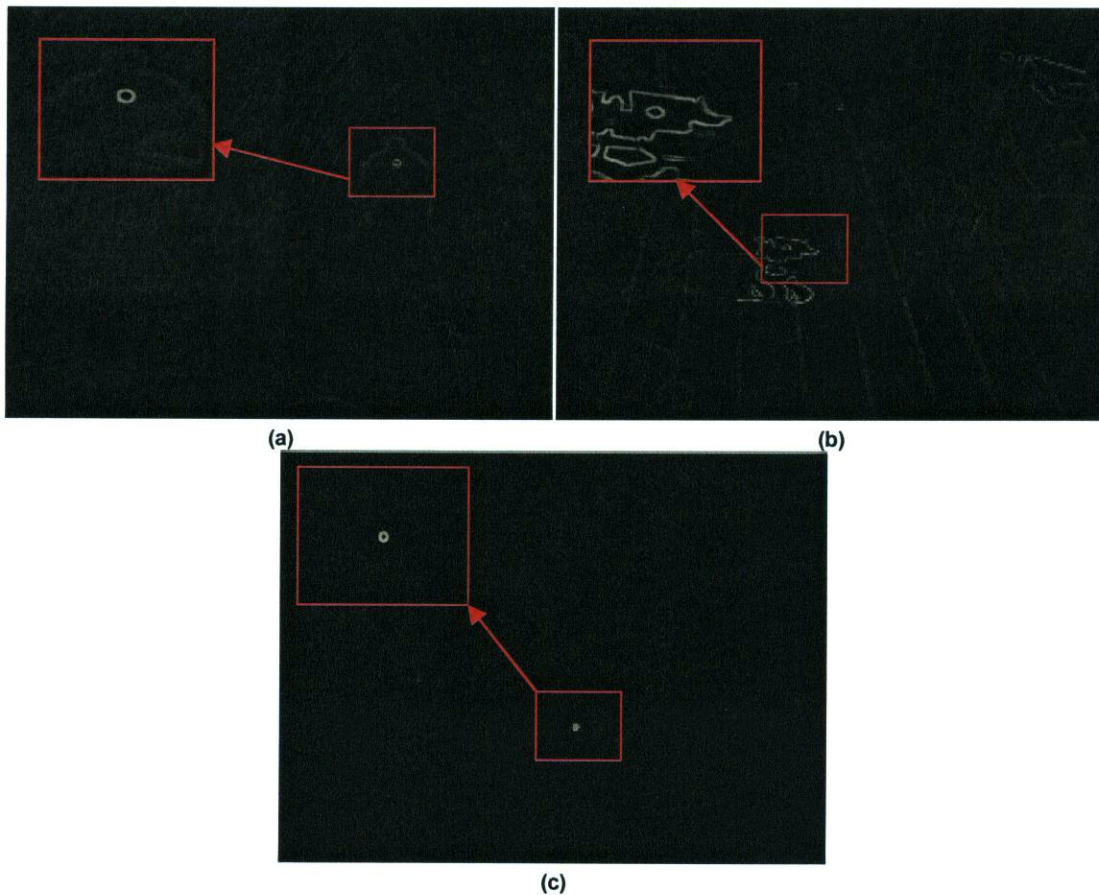
Dado que tras aplicar el operador gradiente se resaltan los bordes, es evidente que los correspondientes a las balizas tomarán niveles de gris altos si se comparan con el resto de la imagen procesada. Por tanto el histograma de las imágenes que resultan de aplicar el detector de bordes presentará dos zonas perfectamente definidas, una correspondiente a los bordes y otro al resto de la imagen (figura 5.13).

Una vez aplicado el operador gradiente, el siguiente paso es realizar un proceso de umbralizado sobre la imagen resultante. Evidentemente, el problema de la segmentación es la elección del umbral. Teniendo en cuenta que en el entorno por donde se mueve el robot pueden existir

diferentes niveles de iluminación y que la distancia entre el robot y cada una de las cámaras también es variable, es de suponer que no se puede utilizar un valor fijo del umbral  $T_k$ . En nuestro caso se propone utilizar un umbral adaptativo, de tal forma que en cada instante  $k$  el valor de  $T_k$  se obtiene a partir del histograma de la ventana de interés. En cualquier caso hay que decir que en las pruebas prácticas realizadas se ha demostrado que el valor del umbral no es crítico, aún en condiciones variables de iluminación y para diferentes distancias entre robot y cámaras, ya que los niveles de gris correspondientes a los bordes de la baliza son significativamente superiores al resto de la imagen.



**Figura 5.11.** Detección de bordes aplicando diferentes máscaras para emular el operador gradiente. (a) Imagen de entrada. (b) Máscara  $H_v = [-1, 1]$ . (c) Máscara de Sobel. (d) Máscara de Prewitt. (e) Máscara de Frei-Chen. (f) Máscaras de Roberts.

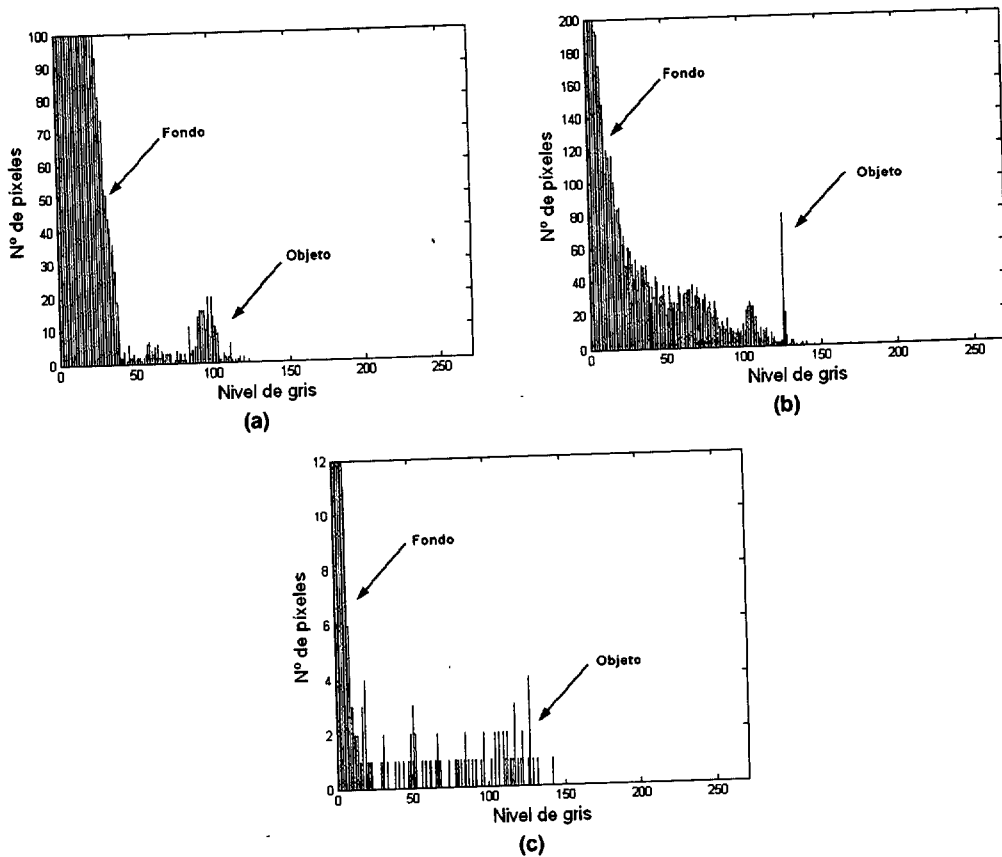


**Figura 5.12.** Resultado de aplicar la máscara de Prewitt a las imágenes de la figura 5.6, reflejando la ventana de interés de la imagen. (a) Imagen de la Figura 5.6.a. (b) Imagen de la figura 5.6.c (c) Imagen de la figura 5.6.e.

En nuestro caso se ha utilizado un umbral adaptativo  $T_k$  basado en el método de Otsu [Otsu 79][Liao 99] dentro de la zona o ventana de interés de la imagen (entorno de predicción de ubicación de la baliza). Por tanto, la imagen de salida tras la umbralización,  $U_k(u, v)$ , será:

$$U_k(u, v) = \begin{cases} 255 & \text{si } I_k(u, v) > T_k \\ 0 & \text{si } I_k(u, v) \leq T_k \end{cases} \quad (5.5)$$

donde  $I_k(u, v)$  se corresponde con el nivel de gris del píxel  $(u, v)$  de la imagen obtenida tras aplicar la detección de borde y  $T_k$  es el umbral. En consecuencia, un píxel de  $U_k(u, v)$  cuyo nivel de gris sea 255 será un candidato a pertenecer al borde de la baliza o a otro borde de características similares.



**Figura 5.13.** Histogramas de las imágenes de la figura 5.12. (a) Imagen de la figura 5.12.a. (b) Imagen de la figura 5.12.b (c) Imagen de la figura 5.12.c.

El método de Otsu se basa en que la medida de la homogeneidad de una región es la varianza (regiones con alta homogeneidad tendrán baja varianza). Además se supone que el histograma de la imagen es la suma de  $n$  gaussianas (tantas como clases existan) con desviaciones típicas iguales o diferentes. El método de Otsu selecciona los umbrales tratando de minimizar la varianza interna de las clases. Si se trata de imágenes con  $n$  clases/regiones, Otsu permite discriminar los píxeles en varias clases o regiones,  $C_1 = \{0, \dots, t_1\}$ ,  $C_2 = \{t_1 + 1, \dots, t_2\}$  ...  $C_n = \{t_{n-1} + 1, \dots, L-1\}$ , donde  $L$  el número de niveles de gris de un píxel, y donde cada clase agrupa todos los píxeles hasta un umbral detectado por Otsu, siendo  $T = \{t_1, t_2, \dots, t_n\}$  el conjunto de umbrales detectados. El modo en que Otsu selecciona los umbrales  $T$ , para crear el conjunto de clases/regiones se fundamenta en la minimización de la varianza interna de las clases o lo que es lo mismo la maximización de varianza entre clases/regiones

En nuestro caso solamente existen dos clases, una  $C_b$  (*background*) con niveles de grises  $\{0, 2, \dots, T_k\}$  y otra  $C_o$  (*objeto*) con niveles de grises  $\{T_k + 1, \dots, L-1\}$ . Considerando que tenemos una

imagen con  $L$  niveles de grises y su histograma normalizado (esto es, para cada nivel de gris  $i$ ,  $p(i)$  es la frecuencia normalizada de  $i$ ), asumiendo que existe un umbral  $T_k$ , la fracción de píxeles que serán clasificados como “objeto” y “fondo” será:

$$q_b(T_k) = \sum_{i=0}^{T_k} p(i); \quad q_o(T_k) = \sum_{i=T_k+1}^{L-1} p(i); \quad \text{de forma que } (q_b(T_k) + q_o(T_k) = 1) \quad (5.6)$$

Las medias para cada clase son:

$$\begin{aligned} \mu_b(T_k) &= \frac{\sum_{i=0}^{T_k} i \cdot p(i)}{\sum_{i=0}^{T_k} p(i)} = \frac{1}{q_b(T_k)} \sum_{i=0}^{T_k} i \cdot p(i) \\ \mu_o(T_k) &= \frac{\sum_{i=T_k+1}^{L-1} i \cdot p(i)}{\sum_{i=T_k+1}^{L-1} p(i)} = \frac{1}{q_o(T_k)} \sum_{i=T_k+1}^{L-1} i \cdot p(i). \end{aligned} \quad (5.7)$$

y la de la imagen total será

$$\mu(T_k) = \frac{\sum_{i=0}^{L-1} i \cdot p(i)}{\sum_{i=0}^{L-1} p(i)} = \sum_{i=0}^{L-1} i \cdot p(i) \quad (5.8)$$

Por otra parte, las varianzas del objeto y fondo serán

$$\begin{aligned} \sigma_b^2(T_k) &= \frac{\sum_{i=0}^{T_k} (1 - \mu_b)^2 p(i)}{\sum_{i=0}^{T_k} p(i)} = \frac{1}{q_b(T_k)} \sum_{i=0}^{T_k} (1 - \mu_b)^2 p(i) \\ \sigma_o^2(T_k) &= \frac{\sum_{i=T_k+1}^{L-1} (1 - \mu_o)^2 p(i)}{\sum_{i=T_k+1}^{L-1} p(i)} = \frac{1}{q_o(T_k)} \sum_{i=T_k+1}^{L-1} (1 - \mu_o)^2 p(i) \end{aligned} \quad (5.9)$$

y la varianza de toda la imagen será:

$$\sigma^2 = \sum_{i=1}^L (1 - \mu)^2 p(i) \quad (5.10)$$

Se puede demostrar que la varianza de la imagen  $\sigma^2$  se puede escribir:

$$\begin{aligned}\sigma^2 &= q_b(T_k)\sigma_b^2(T_k) + q_o(T_k)\sigma_o^2(T_k) + q_b(T_k)(\mu_b(T_k) - \mu)^2 + q_o(T_k)(\mu_o(T_k) - \mu)^2 = \\ &= \sigma_w^2(T_k) + \sigma_B^2(T_k)\end{aligned}\quad (5.11)$$

donde  $\sigma_w^2$  representan las varianzas interna (*within*) de las clases y  $\sigma_B^2$  la varianza entre clases (*between*).

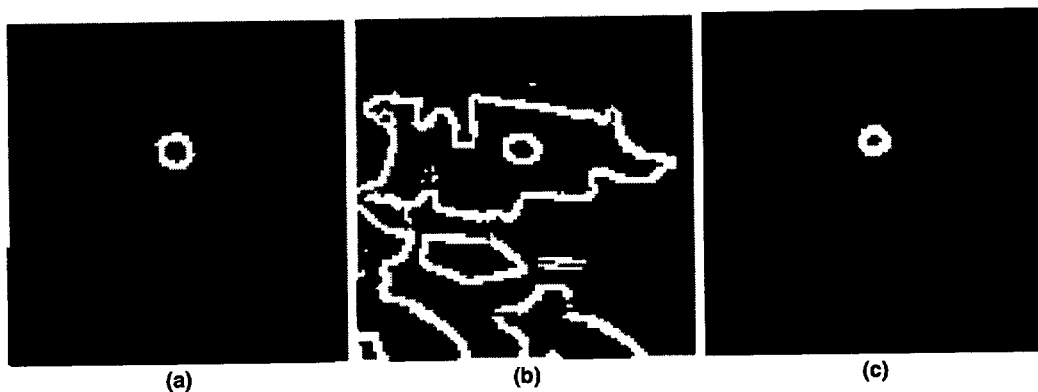
Dado que la varianza total no depende de  $T_k$ , la  $T_k$  que minimiza  $\sigma_w^2$  también maximiza  $\sigma_B^2$ .

Si lo que se busca es maximizar la varianza entre clases, se puede escribir

$$\sigma_B^2 = \frac{[\mu(T_k) - \mu q_b(T_k)]^2}{q_b(T_k)q_o(T_k)}\quad (5.12)$$

Por tanto, el proceso consiste en recorrer todos los niveles de gris, calculando el valor anterior y eligiendo el umbral que maximice esta varianza.

En la figura 5.14 se muestran los resultados de la umbralización de las imágenes de la figura 5.12 con la estrategia propuesta



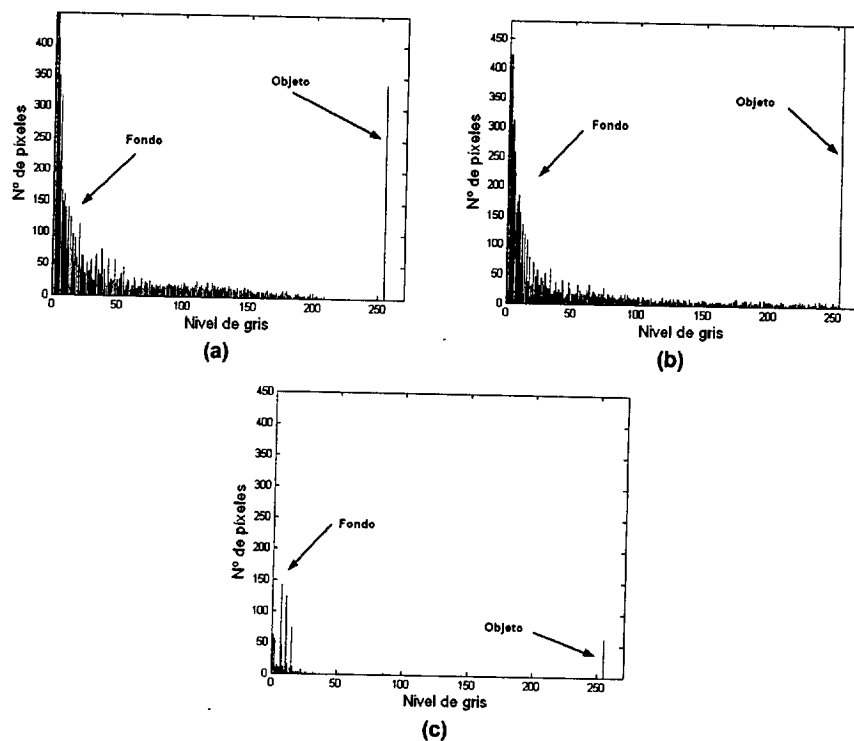
**Figura 5.14.** Resultado de realizar la umbralización de las imágenes de la figura 5.12. (a) Imagen figura 5.12.a,  $T_k=48$ . (b) Imagen figura 5.12.b,  $T_k=41$ . (c) Imagen figura 5.12.c,  $T_k=45$ .

También se han hecho pruebas realizando algunas transformaciones sobre las imágenes obtenidas tras aplicar el operador gradiente, previa a la umbralización, con el fin de obtener un histograma donde objeto y fondo queden más separados (lo que facilita la elección del umbral). Una de las

transformaciones que mejores resultados proporcionó fue la entropía, definida como:

$$I'_k(u, v) = I_k(u, v) \cdot \ln(I_k(u, v)) \quad (5.13)$$

donde  $I'_k(u, v)$  es la imagen de salida e  $I_k(u, v)$  es la imagen de entrada. A modo de ejemplo, en la figura 5.15 se muestran los histogramas obtenidos tras la realización de la entropía a las imágenes de la figura 5.12. En la figura 5.16 se muestran los resultados tras umbralizar la imagen de la figura 5.12 siguiendo esta estrategia. En cualquier caso los resultados obtenidos son muy similares aplicando la entropía y sin ella.

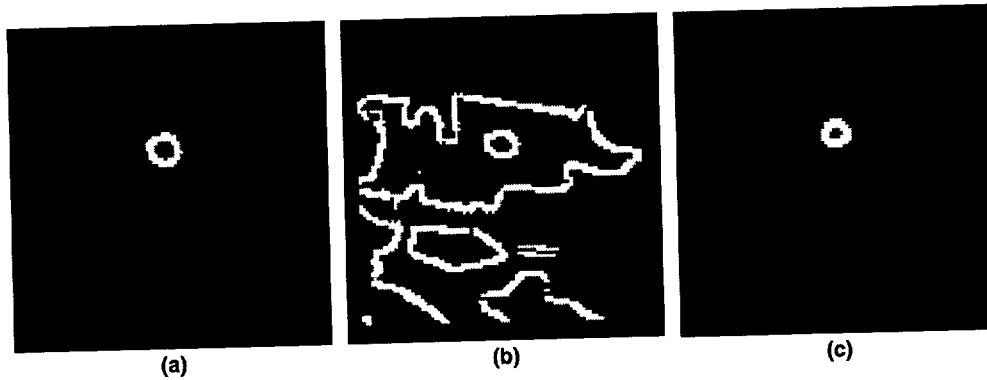


**Figura 5.15.** Ejemplos de histogramas de las imágenes de la figura 5.12 tras la realización de la entropía. (a) Figura 5.12.a. (b) Figura 5.12.b. (c) Figura 5.12.c.

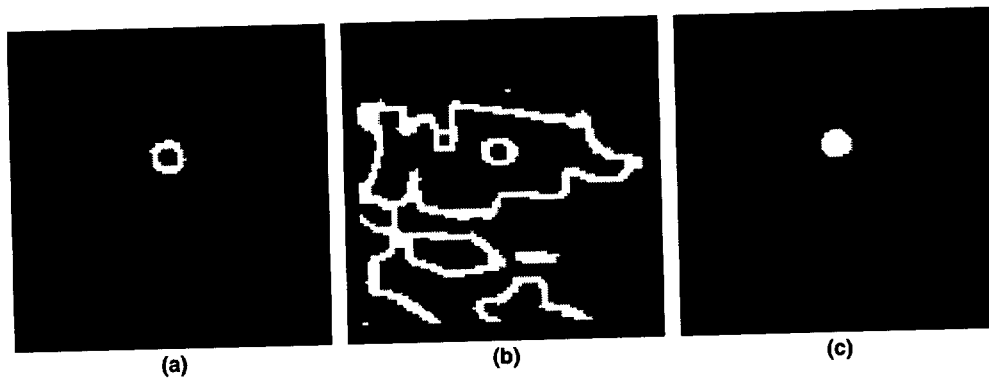
Evidentemente tras la umbralización los bordes de las balizas pueden presentar pequeñas discontinuidades en su contorno. Para eliminar estas discontinuidades se realiza un procesamiento morfológico; en nuestro caso se ha optado por una dilatación seguida de un erosionado. La dilatación de una imagen produce su expansión de los objetos, mientras que el erosionado ocasiona una reducción. El resultado obtenido es una imagen en la que se ha suavizado el contorno de los objetos, pero se unen discontinuidades y se eliminan pequeños



elementos aislados. En la figura 5.17 se muestra el resultado de aplicar la erosión y dilatación en las imágenes de la figura 5.14.



**Figura 5.16.** Resultado de umbralización utilizando la entropía y  $T_k=254$  para las imágenes de la figura 5.12. (a) Imagen figura 5.12.a. (b) Imagen figura 5.12.b. (c) Imagen figura 5.12.c.



**Figura 5.17.** Resultado de realizar la erosión y dilatación sobre las imágenes de la figura 5.14. (a) Imagen figura 5.14.a. (b) Imagen figura 5.14.b. (c) Imagen figura 5.14.c.

Para el caso de las figuras 3.17.a y c, tras la umbralización se ha conseguido aislar la baliza del resto de objetos de la ventana de interés de la imagen, pero no sucede lo mismo en la figura 5.17.b, siendo necesaria una fase de clasificación.

### 5.3. Clasificación de los objetos.

Una vez finalizado el proceso de segmentación dentro de la ventana de interés, y como paso previo a la obtención del centroide de las balizas, se procede a la obtención de las características que permitan diferenciarlas de otros posibles objetos que pudieran haberse segmentado (debido a brillos, ruido, etc) dentro de la ventana de interés.

En éste y en capítulos anteriores se han mostrado dos estructuras de balizas utilizadas para facilitar la identificación del robot : una formada por una única baliza de IR y otra formada por cuatro distribuidas en forma de “T”. A continuación se detalla el proceso de clasificación seguido para identificar la estructura formada por cuatro balizas.

Una vez segmentados los objetos dentro de las ventanas de interés, el proceso general que se propone en esta tesis para la detección de las balizas es el que se resume en la figura 5.18. En una primera fase se eliminan aquellos objetos de área reducida (área inferior a 10 píxeles) teniendo presente el tamaño de las balizas, las características de la óptica utilizada y la distancia máxima a la que pueden ser captadas. En el segundo, y dado que las balizas utilizadas son circulares, es evidente que al proyectarse en el plano imagen en general aparecerán como elipses. Por tanto, una alternativa para encontrar, dentro de los objetos segmentados, los candidatos a elipses consiste en encontrar para cada uno de ellos la elipse que mejor se ajusta, que denominaremos “elipse estimada”. En esta segunda fase se van a eliminar aquellas “elipses estimadas” cuya distancia algebraica no supere un determinado valor (en este caso se ha fijado en 0,5 píxeles). En una tercera fase se calcula la excentricidad y la relación entre los perímetros de las “elipses estimadas” y los de los correspondientes objetos. La excentricidad permite eliminar aquellos objetos de características similares a los mostrados en la figura 5.19.d, los cuales tienen una distancia algebraica pequeña y sin embargo no pueden representar a las balizas, ya que éstas siempre presentarán excentricidades próximas a la unidad; tras esta fase se seleccionan las cuatro balizas de menor excentricidad, que identificaremos por  $BE_1$ ,  $BE_2$ ,  $BE_3$  y  $BE_4$ . En cuanto a la relación entre perímetros permite eliminar objetos del tipo mostrados en la figura 5.19.c, donde se puede observar que la “elipse estimada” se ajusta perfectamente a los puntos de borde utilizados en su generación, pero sin embargo el perímetro de la elipse estimada es muy superior a los puntos de borde utilizados en su generación; de nuevo, tras esta fase se seleccionan las cuatro balizas de menor relación de perímetros, y que identificaremos por  $BP_1$ ,  $BP_2$ ,  $BP_3$  y  $BP_4$ . En la fase cuarta se busca la coincidencia entre las balizas  $BE_1$ ,  $BE_2$ ,  $BE_3$ ,  $BE_4$  con  $BP_1$ ,  $BP_2$ ,  $BP_3$ ,  $BP_4$ :  $\{BE_1, BE_2, BE_3, BE_4\} = \{BP_1, BP_2, BP_3, BP_4\}$ . Dicha coincidencia se puede realizar comparando las coordenadas de los centroides de los objetos o con una etiqueta identificativa que se puede asociar a cada objeto (*labeling*). Si existe coincidencia entre los dos grupos, el paso siguiente es comprobar que las cuatro balizas seleccionadas cumplen las relaciones geométricas propias de la ubicación física relativa entre ellas. En caso afirmativo se da por finalizado el algoritmo de clasificación. Tanto para el caso de no haber coincidencia entre los grupos de cuatro balizas,  $\{BE_1, BE_2, BE_3, BE_4\} \neq \{BP_1, BP_2, BP_3, BP_4\}$ , como para el caso de no cumplirse las relaciones geométricas entre ellas, se procederá a ampliar la ventana de búsqueda y repetir el proceso, hasta que dicha ventana abarque toda la imagen. Si analizando toda la imagen tampoco

se detectan las cuatro balizas, se procede a captar y analizar otra nueva imagen.

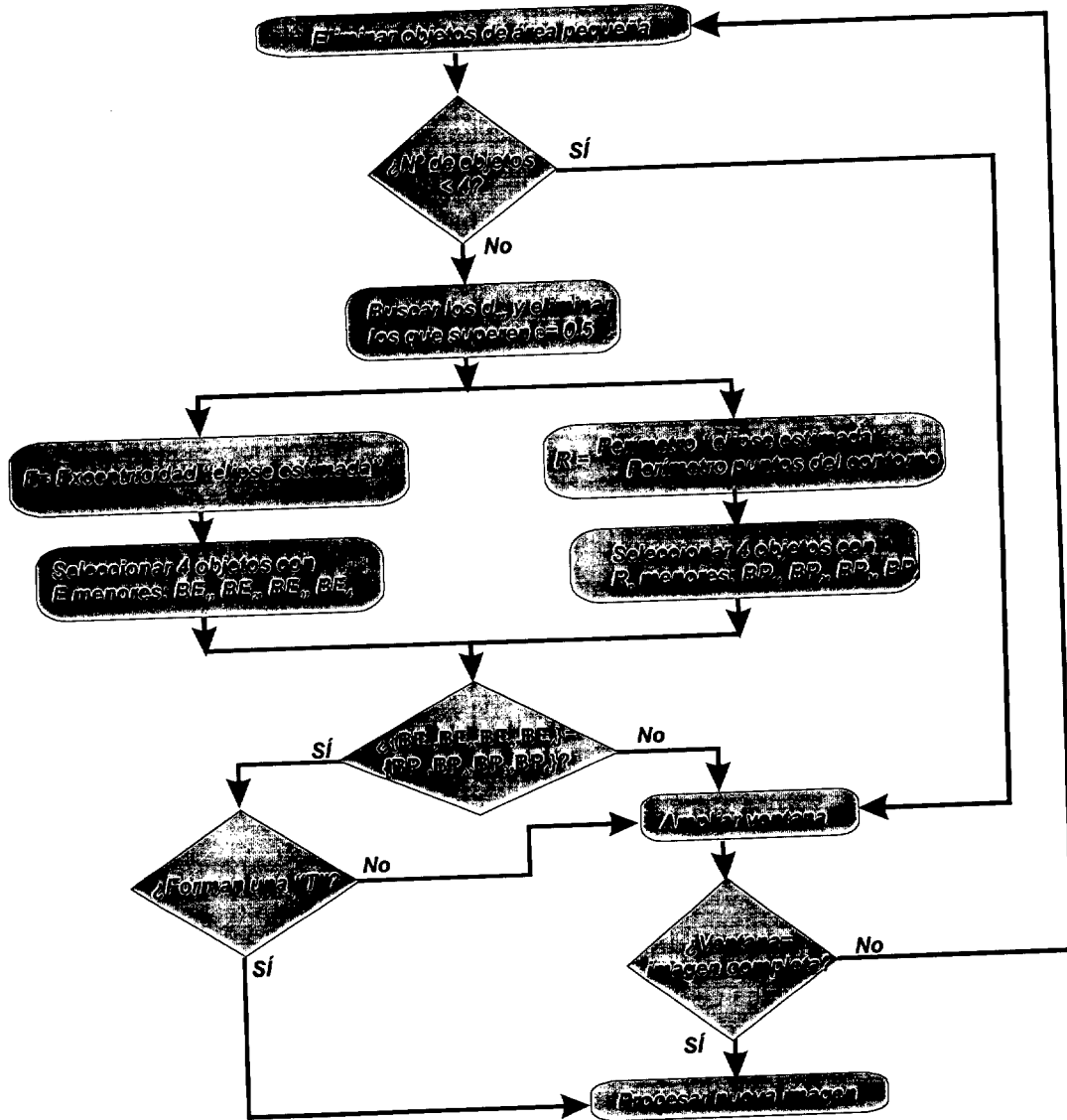
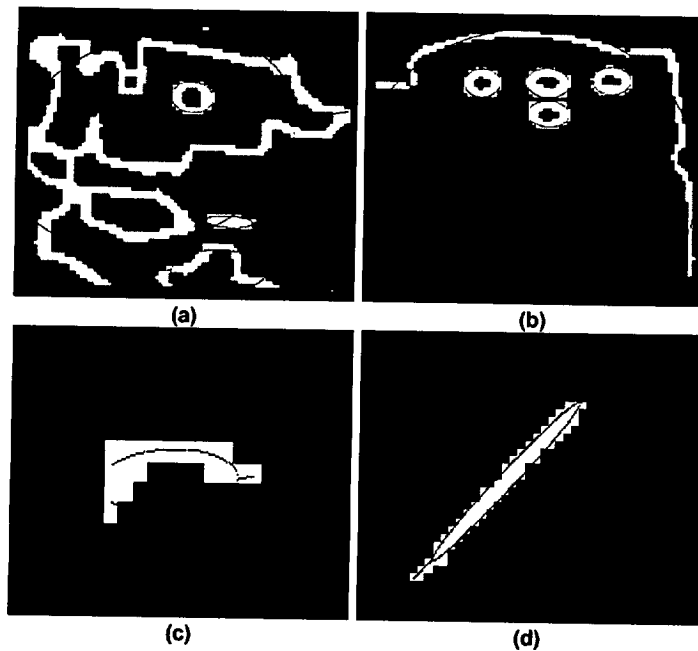


Figura 5.18. Organigrama de clasificación seguido para identificar las balizas IR en una imagen.

Seguidamente se describen los algoritmos utilizados en cada una de las fases descritas en la figura 5.18.



**Figura 5.19.** Ejemplos de “elipses estimadas” utilizando la distancia algebraica. (a) Caso de una estimación correcta (color rojo) y tres falsas (color azul). (b) Caso de cuatro balizas estimadas correctamente y una falsa. (c) Ejemplo de “elipse estimada” falsa que puede ser eliminada tras la comparación de perímetros. (d) Ejemplo de “elipse estimada” falsa que puede ser descartada por excentricidad.

### 5.3.1. Eliminación de objetos de pequeñas dimensiones.

Es evidente que en el proceso de segmentación, por razones diversas, se generan objetos de pequeñas dimensiones que nada tienen que ver con las balizas. Considerando el tamaño de las balizas y las condiciones más desfavorables desde el punto de vista de las dimensiones de su proyección en el plano imagen, se pueden descartar todos aquellos objetos cuya área sea inferior a un determinado umbral. En nuestro caso se descartan todos aquellos objetos cuya área,  $A$ , no supere 10 píxeles. Por tanto el conjunto de candidatos a balizas tras esta primera fase estará formada por un conjunto  $\beta = \{B_1, B_2, B_3, \dots\} \ni A \geq 10 \text{ píxeles}$ . El proceso continúa siempre y cuando los objetos del conjunto  $\beta$  sea igual o superior a cuatro, en caso contrario se debe proceder a ampliar la ventana de interés.

### 5.3.2. Identificación de elipses.

Para obtener la elipse (“elipse estimada”) que mejor se ajusta a cada uno de los objetos segmentados se pueden utilizar diferentes alternativas. En esta tesis se han propuesto: ajuste

mediante distancia algebraica, utilizando técnicas de Análisis de Componentes Principales (PCA) y utilizando ángulos de Feret, si bien se ha podido comprobar en las pruebas prácticas que la alternativa que mejores resultados proporciona es la de ajuste mediante distancia algebraica.

### Ajuste mediante distancia algebraica.

La distancia algebraica de un punto  $p$  a una curva  $g(p, w) = 0$  es simplemente  $|g(p, w)|$ . Por tanto, si lo que se trata es de buscar la elipse, que denominaremos "elipse estimada", que mejor se adapta a un objeto con  $N$  puntos de borde con coordenadas píxeles  $p_i = [u_i, v_i]$ ,  $i = 1, 2, \dots, N$ , el objetivo es determinar el vector de parámetros  $w = [a, b, c, d, e, f]^T$  que caracteriza a dicha elipse. Para ello se utiliza la distancia algebraica definida en (5.15):

$$d_w = \min_w \sum_{i=1}^N |x_i^T \cdot w| \quad (5.15)$$

siendo

$$g(p, w) = x^T \cdot w = a \cdot u^2 + b \cdot u \cdot v + c \cdot v^2 + d \cdot u + e \cdot v + f = 0 \quad (5.16)$$

la ecuación implícita de la cónica genérica que para  $4ac - b^2 > 0$  se convierte en la ecuación de la elipse y donde  $x_i^T$  viene dada por:

$$x_i^T = [u_i^2, u_i \cdot v_i, v_i^2, u_i, v_i, 1] \quad (5.17)$$

Por tanto, el objetivo es buscar el vector de parámetros de la elipse que minimice (5.15) a partir de los  $N$  puntos de borde de cada uno de los objetos segmentados. En la minimización de (5.15), para evitar la solución trivial  $w=0$ , se debe imponer una restricción en  $w$ . En este sentido, de las múltiples soluciones posibles se elige aquella que fuerce la solución de (5.16) para que sea una elipse. Esta condición se puede escribir como:

$$b^2 - 4ac = w^T \cdot \begin{bmatrix} 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot w = w^T \cdot B \cdot w = -1 \quad (5.18)$$

Para encontrar una solución a este problema se puede escribir (5.15) como:

$$d_w = \min_w \|w^T X^T X w\| = \min_w \|w^T S w\| \quad (5.19)$$

donde

$$X = \begin{bmatrix} u_1^2 & u_1 \cdot v_1 & v_1^2 & u_1 & v_1 & 1 \\ u_2^2 & u_2 \cdot v_2 & v_2^2 & u_2 & v_2 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ u_N^2 & u_N \cdot v_N & v_N^2 & u_N & v_N & 1 \end{bmatrix} \quad (5.20)$$

En la terminología de mínimos cuadrados  $X$  se conoce como la matriz de diseño,  $S = X^T X$  es la "scatter matrix" y  $B$  la matriz de restricción.

Para minimizar (5.19), imponemos la restricción de  $w$  mediante el multiplicador de Lagrange ( $\lambda$ ), y definimos la función (5.21):

$$\chi = w^T S w - \lambda (w^T B w + 1) \quad (5.21)$$

Derivando (5.21) con respecto a  $w$ , se obtiene:

$$\frac{\partial \chi}{\partial w} = 2S w - 2\lambda B w = 0 \rightarrow S w = \lambda B w \quad (5.22)$$

Por tanto, la solución de (5.15) se puede escribir:

$$S w = \lambda B w \quad (5.23)$$

que se conoce como el problema de *autovalores generalizado*.

Introduciendo el valor dado por (5.23) en (5.21) se obtiene:

$$\chi_{\min} = w^T \lambda B w - \lambda w^T B w - \lambda = -\lambda \quad (5.24)$$

Como el valor mínimo de (5.21) es positivo, se concluye que el autovector asociado al único autovalor negativo se corresponde con el vector de parámetros asociado a la elipse que mejor se ajusta a los  $N$  puntos  $p_i = [u_i, v_i]$ ,  $i = 1, 2, \dots, N$ .

Una vez finalizado este proceso, si en la imagen hay  $r$  objetos segmentados, ordenando las distancias algebraicas de menor a mayor:  $d_{w1} < d_{w2} < d_{w3} < d_{w4} < \dots < d_{wr}$ , el conjunto,  $\Gamma$ , de los posibles candidatos a balizas, estará formado por aquellos cuya  $d_{wi}$  sea inferior a 0.5 ( $\Gamma = \{\text{Objetos}\} \ni d_{wi} \leq 0.5 \text{ píxeles}$ ).

### Excentricidad y relación de perímetros.

Como se puede observar en el organigrama de la figura 5.18, para cada elemento del conjunto  $\Gamma$  de elipses estimadas, se calcula su excentricidad y relación de perímetros, con el fin de descartar aquellos potenciales candidatos que en realidad son falsos.

*Excentricidad (E).* La excentricidad de una elipse se define como:

$$E = \sqrt{1 - \left(\frac{e_{jm}}{e_{jM}}\right)^2} \quad (5.26)$$

siendo  $e_{jm}$  y  $e_{jM}$  el tamaño de los semiejes menor y mayor, respectivamente, de la elipse. Si en (5.16)  $b \neq 0$  los ejes de la elipse no son paralelos a los ejes de la imagen, es necesario realizar una rotación para obtener la elipse equivalente con los ejes paralelos a los de la imagen, siendo la nueva ecuación de la elipse:

$$\lambda_1 \cdot u^2 + \lambda_2 \cdot v^2 + [d \ e] \cdot \mathbf{D} \cdot \begin{bmatrix} u \\ v \end{bmatrix} + f = 0 \quad (5.27)$$

donde  $\mathbf{D}$  son los autovectores asociados a los autovalores  $\lambda_1$  y  $\lambda_2$  de la matriz:

$$\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \quad (5.28)$$

Según (5.26) y (5.27) el valor de la excentricidad vendrá dada por:

$$E = \sqrt{1 - \frac{\left(\frac{1}{2} \cdot (\min(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}))\right)}{\left(\frac{1}{2} \cdot (\max(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}))\right)}} = \sqrt{1 - \left(\frac{\lambda_b}{\lambda_a}\right)^2} \quad (5.29)$$

donde  $\lambda_a$  representa la longitud del semieje mayor y  $\lambda_b$  la del menor.

Es evidente que cualquiera que sea la posición de las balizas dentro del entorno captado por una determinada cámara, su proyección en el plano imagen será una elipse, pero su relación entre los ejes mayor y menor no puede ser muy dispar. Por ello se seleccionan como “elipses estimadas” las cuatro que tengan menor excentricidad. Por tanto, ordenando las elipses de menor a mayor excentricidad:  $E_1 < E_2 < E_3 < E_4 \dots$ , los candidatos a balizas será el conjunto  $\Pi = \{BE_1, BE_2, BE_3, BE_4\}$  asociados a las cuatro excentricidades menores. Con ello se eliminan objetos tales como los mostrados en la figura 5.19.d.

*Relación de perímetros.* La relación de perímetros,  $R_p$ , se define como el cociente entre los perímetros de la “elipse estimada”, que identificaremos por  $P_e$ , y el de los puntos de contorno del objeto utilizados en la obtención de la correspondiente elipse, que identificaremos por  $P_c$ :

$$R_p = \frac{\text{Per. "elipse estimada"}}{\text{Per. puntos de contorno}} = \frac{P_e}{P_c} \quad (5.30)$$

Donde, el perímetro de la “elipse estimada” se aproxima por:

$$P_e \approx \pi \cdot (\lambda_a + \lambda_b) \cdot \left( 1 + \frac{3 \cdot \left( \frac{\lambda_a - \lambda_b}{\lambda_a + \lambda_b} \right)^2}{10 + \sqrt{4 - 3 \cdot \left( \frac{\lambda_a - \lambda_b}{\lambda_a + \lambda_b} \right)^2}} \right) \quad (5.31)$$

siendo el error relativo ( $e_{rel}$ ) de la aproximación (5.31)

$$e_{rel} = \frac{3 \cdot e^{20}}{2^{36}} \quad (5.32)$$

Ahora, ordenando de mayor a menor los valores de  $R_p$ :  $R_{p1} > R_{p2} > R_{p3} > R_{p4}$ , los candidatos a balizas será el conjunto,  $\eta = \{BP_1, BP_2, BP_3 \text{ y } BP_4\}$  asociados a las cuatro relaciones de perímetros mayores. Esto permite eliminar candidatos de las características mostradas en la figura 5.19.c.

En la figura 5.20 se muestran los valores de los parámetros  $d_w$ ,  $E$  y  $R_p$  para cada una de las “elipses estimadas” de dos imágenes.



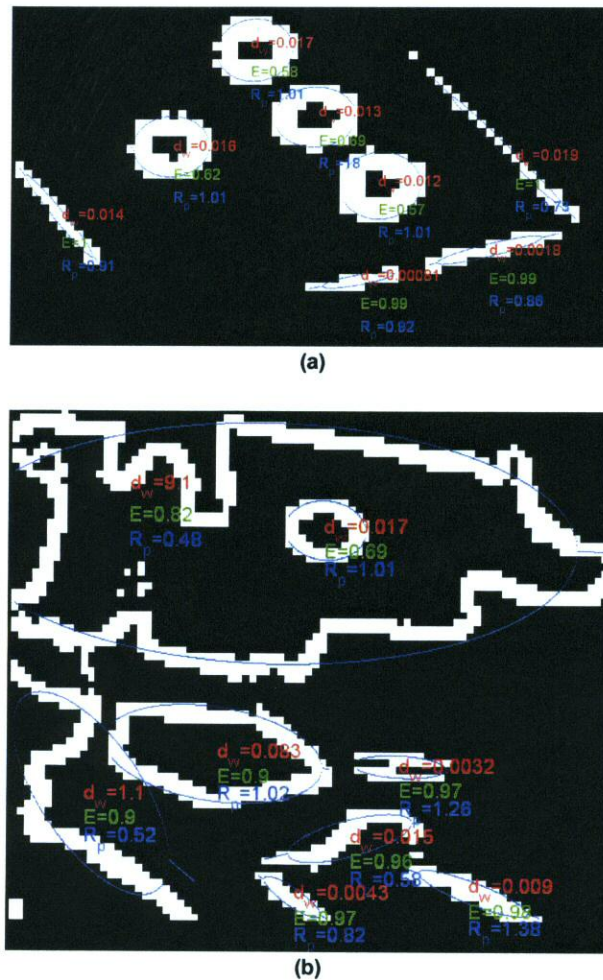


Figura 5.20. Representación de los parámetros  $d_w$ ,  $E$  y  $R_p$  para cada una de las “elipses estimadas” asociadas a cada uno de los objetos de la imagen. (a) Imagen con cuatro balizas. (b) Imagen con una baliza.

### 5.3.3. Validación de candidatos.

Para el caso de utilizar cuatro balizas sobre el robot, se puede utilizar su distribución para discriminar las balizas de otros objetos. Para ello, una vez obtenidos los conjuntos  $\mathbf{II} = \{BE_1, BE_2, BE_3, BE_4\}$  y  $\boldsymbol{\eta} = \{BP_1, BP_2, BP_3, BP_4\}$  las cuatro candidatas a balizas se darán, en principio, como válidos si ambos conjuntos contienen los mismos elementos.

Para el caso de la plantilla formada por cuatro balizas, cuando se cumpla que  $\{BE_1, BE_2, BE_3, BE_4\} = \{BP_1, BP_2, BP_3, BP_4\}$ , se procede a comprobar si las relaciones geométricas entre las

coordenadas  $(u, v)$  de los centroides de los objetos que cumplan la relación anterior corresponden a las de la "T" formada por las balizas. Para ello se comprueba si entre dichas coordenadas existen tres alineadas. El método utilizado para calcular el centroide de un objeto de la imagen se discutirá en el apartado 5.4.

En efecto, sean  $p_1, p_2, \dots, p_4$  los centroides de los objetos detectados en una imagen que cumplen con las restricciones impuestas sobre los parámetros  $d_w, E$  y  $R_p$  y con coordenadas  $(u_1, v_1), (u_2, v_2), \dots, (u_4, v_4)$ , respectivamente. El proceso utilizado para comprobar que cuatro de ellos se corresponden con las balizas, consiste en obtener las ecuaciones de las posibles rectas determinadas por cada dos puntos. Si identificamos por  $r_{ij}$  la ecuación de la recta que pasa por los puntos  $p_i$  y  $p_j$ ,  $i, j = 1, 2, 3, 4$ , se calculan las distancias de los otros puntos a cada una de esas rectas. Identificando por  $d(r_{ij}, p_k)$ ,  $i \neq j \neq k$ , las distancias del punto  $p_k$  a la recta  $r_{ij}$ , si una de estas distancias es inferior a un determinado umbral (en nuestro caso se ha fijado en 5 píxeles) será indicativo de que los puntos que cumplen esta condición ( $p_i, p_j$  y  $p_k$ ) se corresponden con las tres balizas que son candidatas a formar la parte horizontal de la "T" del sistema de balizas (balizas B,C,D) (ver figura 5.21) y la cuarta baliza con la A.

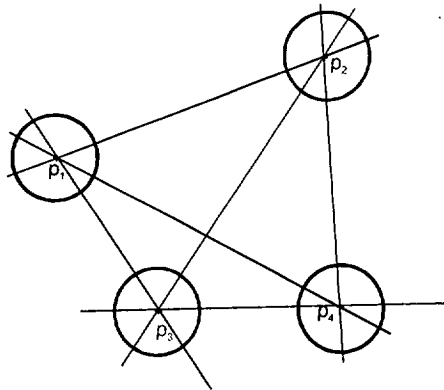


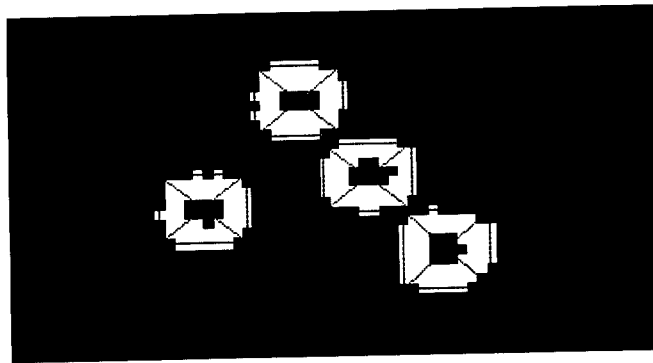
Figura 5.21. Estrategia utilizada en la validación de candidatos a balizas a partir de su disposición geométrica.

## 5.4. Obtención del centroide de la baliza.

Tras identificar y segmentar las balizas, el paso siguiente consiste en obtener las coordenadas  $(u, v)$  de cada uno de sus centroides. Este es un aspecto de gran importancia ya que la precisión con la que se obtenga tiene una influencia directa en la obtención de la posición del robot (tal como se demostró en 3.4.5). Con carácter general se puede decir que existen dos grandes alternativas

para la obtención del centroide, una es trabajar con luminancias y la otra con imágenes binarizadas. Trabajar con luminancias proporciona una mayor precisión en la detección del centroide, tal y como ha quedado de manifiesto en los trabajos desarrollados en [Lavest, 99][Lázaro 00b][Heikkilä 00], sin embargo su coste computacional es muy alto, lo que lo hace prohibitivo en nuestra aplicación. Por otra parte, y si bien interesa calcular el centroide con precisión, la diferencia entre las dos alternativas se ha demostrado que no tiene una incidencia notable en el posicionamiento del robot, por lo que finalmente se ha optado por realizar el cálculo a partir de imágenes binarizadas, lo que justifica todo el estudio realizado en los apartados anteriores.

En concreto, la propuesta realizada en esta tesis para obtener las coordenadas del centroide ( $u_c$ ,  $v_c$ ) es considerar que sus coordenadas coinciden con las del punto de intersección de las diagonales principales que unen los vértices del mínimo rectángulo que contiene la baliza (figura 5.22).



**Figura 5.22.** Obtención de las coordenadas del centroide de la baliza como intersección de las diagonales del rectángulo mínimo que la contiene.

Para comprobar la eficacia del algoritmo propuesto, en comparación con los métodos basados en luminancia, se han realizado varias pruebas. Para ello, una vez identificada una baliza en la imagen binarizada, se define una región o ventana de trabajo (que contiene la baliza) sobre la imagen original, y sobre ella se realiza una aproximación parabólica de la luminancia, tanto en el eje horizontal (coordenadas  $u$ ) como en el vertical (coordenadas  $v$ ), asignándole al centroide las coordenadas en las que las parábolas anteriores son máximas. La ecuación general de estas parábolas es:

$$\begin{aligned} f(u) &= a_u u^2 + b_u u + c_u \\ g(v) &= a_v v^2 + b_v v + c_v \end{aligned} \quad (5.33)$$

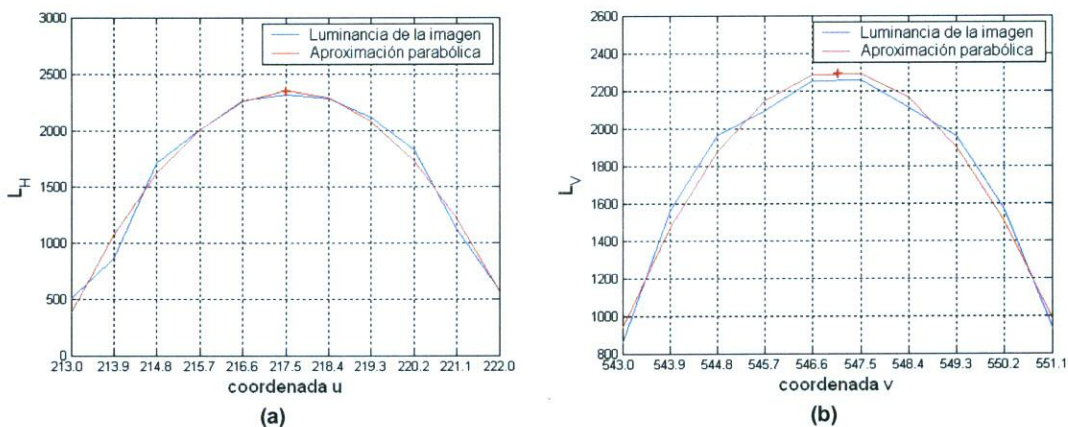
donde las coordenadas del centroide son las de los puntos donde las curvas toman su valor máximo:

$$\begin{aligned} u_p &= -\frac{b_u}{2a_u} \\ v_p &= -\frac{b_v}{2a_v} \end{aligned} \quad (5.34)$$

El valor de la luminancia de un objeto  $\mathbf{O}$  se obtiene sumando el nivel de gris de todos los píxeles que lo forman, realizando un barrido horizontal (luminancia horizontal,  $L_H$ ) y vertical (luminancia vertical,  $L_V$ ) según:

$$\begin{aligned} L_H(u) &= \sum_i I(u, v_i) \quad \forall v_i \in \mathbf{O} \\ L_V(v) &= \sum_j I(u_j, v) \quad \forall u_j \in \mathbf{O} \end{aligned} \quad (5.35)$$

En la figura 5.23 se muestran en color azul las luminancias horizontal (figura 5.23.a) y vertical (figura 5.23.b) correspondientes a la baliza de la imagen de la figura 5.11.a. Sobre las gráficas de luminancia se ha representado, en color rojo, la aproximación parabólica dada por (5.33). Así mismo, sobre las gráficas de la figura 5.23 se ha marcado con una cruz roja el punto máximo (ecuación (5.34)) de la parábola.



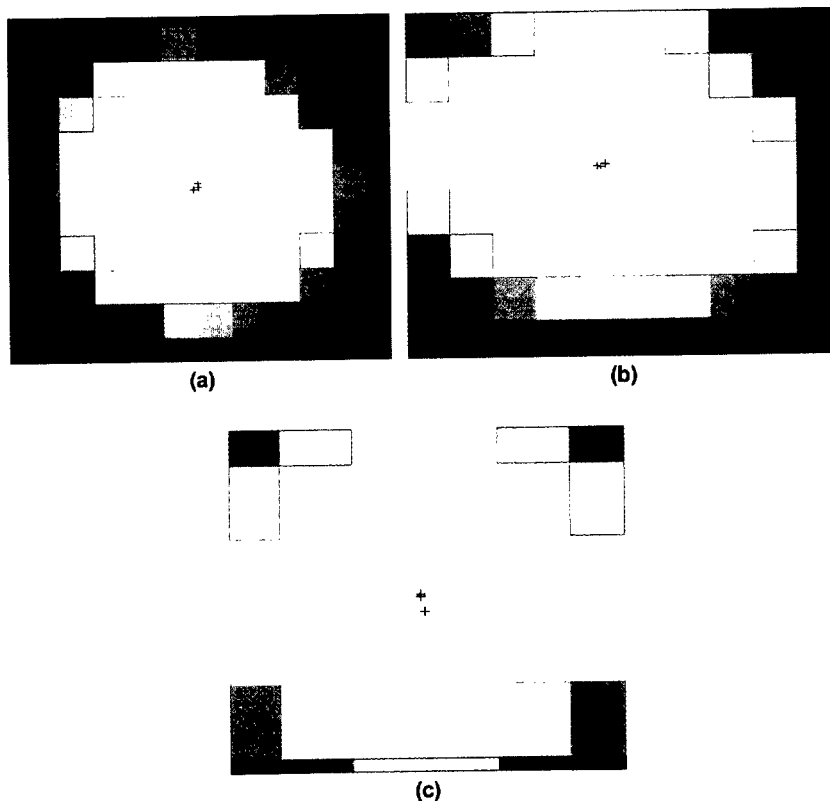
**Figura 5.23.** Representación de la luminancia de la imagen de la figura 5.11.a (en color azul) y su aproximación parabólica (color rojo). (a) Luminancia horizontal. (b) Luminancia vertical.

Otra alternativa posible para la obtención del centroide puede ser la de considerar como

coordenadas de éste las del centro de gravedad  $(u_g, v_g)$  del objeto en la imagen binarizada (figura 5.22), las cuales vienen dadas por:

$$u_g = \frac{\sum_i u_i}{Area}, \quad v_g = \frac{\sum_i v_i}{Area} \quad \forall \quad u_i, v_i \in \mathbf{O} \quad (5.36)$$

En la figura 5.24 se muestran las coordenadas del centroide de la baliza de la figura 5.6, calculado utilizando los tres métodos anteriores. En las figura 5.24 se indica con una cruz azul el centroide obtenido con el método propuesto en esta tesis, en rojo el obtenido con la aproximación parabólica y en verde el del centro de gravedad de la imagen binarizada.



**Figura 5.24.** Representación de las coordenadas de los centroides de la baliza para las imágenes de la figura (a) 5.6.a, (b) 5.6.b y (c) 5.6.c, obtenidas utilizando diferentes alternativas de cálculo.

En la tabla 5.1 se muestran los valores numéricos de las coordenadas de los centroides reflejados en la figura 5.24 y para las tres alternativas indicadas.

**TABLA 5.1.** Valores de las coordenadas del centroide de la baliza para las imágenes de la figura 5.6 y para diferentes métodos de cálculo.

Método	Figura 5.6.a		Figura 5.6.c		Figura 5.6.e	
	$u_c$	$v_c$	$u_p$	$v_p$	$u_g$	$v_g$
Intersección de diagonales principales	218.19	547.49	334.93	324.11	385.30	417.44
Aproximación Parabólica	218.12	547.53	334.98	323.91	385.27	417.43
Centro de Gravedad	218.36	547.7	335.25	324.12	385.56	417.52

Comparando las coordenadas del centroide obtenidas mediante el método propuesto en esta tesis con la aproximación parabólica de la luminancia, se observa una pequeña diferencia que apenas tiene efecto en el resultado final de posicionamiento del robot.

## 5.5. Obtención de la orientación del robot a partir de las balizas

Como ya se ha comentado anteriormente, nuestro interés en la utilización de balizas de IR es la de determinar la posición del robot, pudiéndose obtener la orientación a partir de los propios movimientos de éste. Sin embargo, también puede resultar de interés, sobre todo en la fase de inicio, poder obtener la orientación a partir de la información proporcionada por las balizas. Por tanto, una vez obtenidos los centroides de las cuatro balizas es necesario identificarlas a fin de poder saber la orientación del robot. En definitiva, nombrando las balizas con las letras *A*, *B*, *C* y *D* (figura 5.25.a), se trata de identificarlas con los objetos segmentados. De la validación de candidatos realizada, en 5.3 se obtiene una primera identificación de las balizas (figura 5.25.b), determinando los objetos *i*, *j*, *k* que forman la parte horizontal de la "T", mientras que el cuarto se corresponde con la baliza *A*.

Ahora solamente queda identificar las balizas *B*, *C* y *D*; para ello se calculan las distancias entre los puntos *i*, *j* y *k*:

$$\begin{aligned}
 d_{ij} &= \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \\
 d_{ik} &= \sqrt{(u_i - u_k)^2 + (v_i - v_k)^2} \\
 d_{jk} &= \sqrt{(u_j - u_k)^2 + (v_j - v_k)^2}
 \end{aligned}
 \tag{5.37}$$

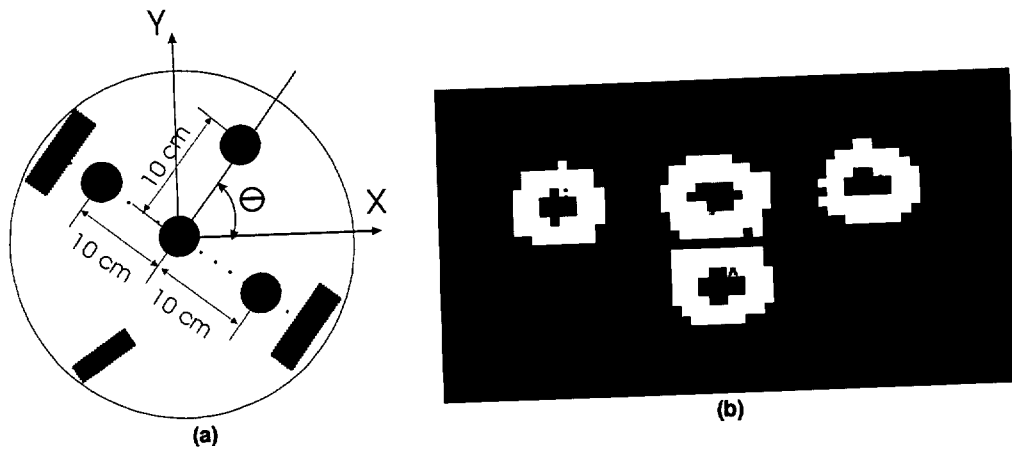


Figura 5.25. Distribución de las cuatro balizas a bordo del robot. (a) Distribución y nomenclatura utilizada. (b) Ejemplo de identificación.

Las coordenadas de la baliza  $C$ , se corresponderán con las del punto cuyas distancias a los otros dos sean las más pequeñas. Finalmente, para identificar las balizas  $B$  y  $D$  basta con obtener el producto vectorial de los vectores que van desde el punto  $A$  al  $C$  y del  $A$  a uno de los otros dos ( $D$  o  $B$ ), que identificaremos por  $k$ :

$$(u_A - u_C)(v_A - v_{Ak}) - (v_A - v_C)(u_A - u_{Ak}) \quad (5.38)$$

Si el resultado de (5.38) es positivo indica que el punto  $k$  se corresponde con la baliza  $B$ , en caso contrario será la baliza  $D$ . En la figura 5.26 se muestra la reordenación de los centroides de la figura 5.25.b según el proceso descrito con anterioridad.

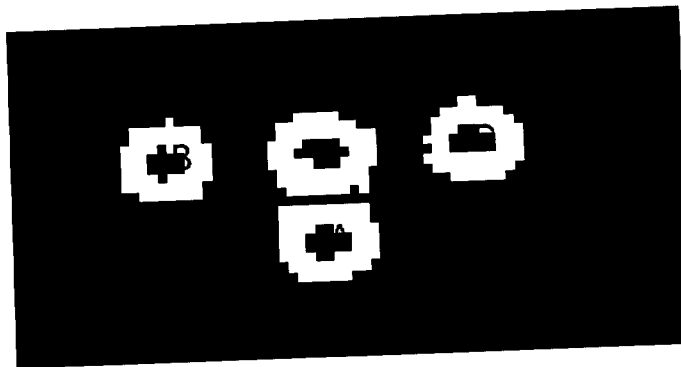


Figura 5.26. Identificación de los centroides correspondientes a las cuatro balizas de la figura 5.24.b.

Finalmente, a partir de los centroides de las balizas, obtenidos en las diferentes cámaras que en un mismo instante “ven” el robot, se obtienen las coordenadas 3D de las balizas ( $[X_A, Y_A, Z_A]^T$ ,  $[X_B, Y_B, Z_B]^T$ ,  $[X_C, Y_C, Z_C]^T$  y  $[X_D, Y_D, Z_D]^T$ ), según las propuestas realizadas en el capítulo 4. La posición del robot se asigna al centroide de la baliza  $C$ , mientras que su orientación se corresponde con la del vector que va desde el punto  $C$  al  $A$ .

## 5.6. Conclusiones.

En este capítulo se han abordado los aspectos más importantes que aparecen en la identificación de los robots, a partir de las imágenes captadas por las diferentes cámaras ubicadas en el entorno de movimiento. En nuestro caso se ha propuesto la utilización de balizas de IR ubicadas a bordo de los robots, con el objetivo de facilitar la localización de éstos incluso en condiciones extremas de iluminación ambiental y al mismo tiempo simplificar los algoritmos de procesamiento de imágenes. Por una parte se propone la utilización de una configuración de cuatro balizas circulares dispuestas físicamente en forma de “T”, una en cada extremo de la “T” y una cuarta en el punto de intersección de las líneas horizontal y vertical. Por otra parte, se han propuesto diferentes algoritmos de procesamiento de imágenes que facilitan la detección de las balizas y de los puntos de interés (centroides), a partir de los cuales se obtienen la posición y orientación de los robots, de forma fiable, robusta y con tiempos de proceso muy reducidos.

Como aportaciones a destacar dentro de este capítulo se pueden citar:

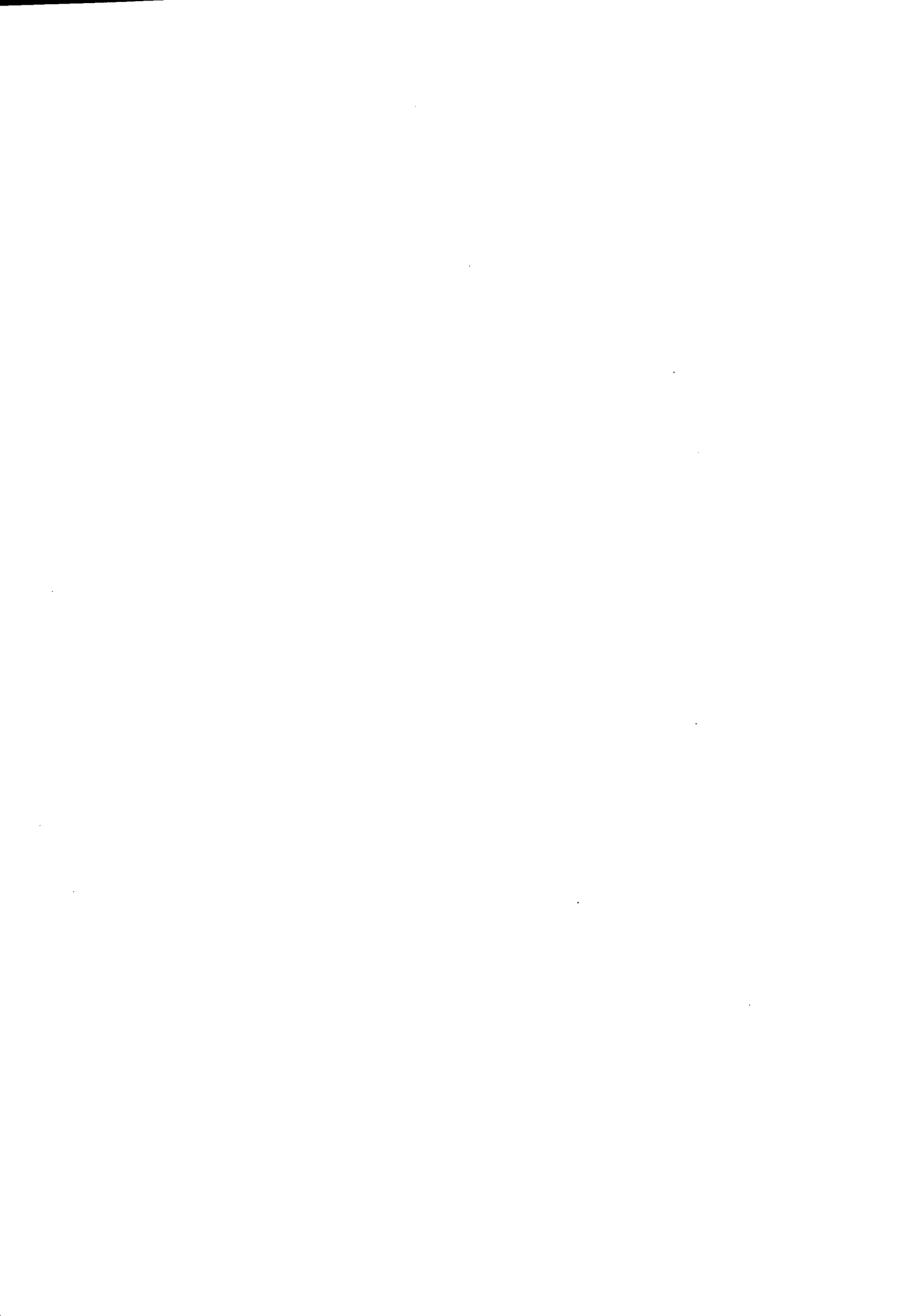
- Se ha propuesto una baliza de IR cuya dimensión y características de emisión (longitud de onda, radiación), dimensiones, potencia de emisión, etc, permiten una fácil identificación, incluso en condiciones cambiantes y extremas de iluminación ambiental, utilizando cámaras CCD estándares y ampliamente comercializadas. Además se ha propuesto una configuración de cuatro balizas dispuestas en forma de “T”, que facilita la identificación de las mismas (aprovechando su disposición geométrica) y permiten obtener información sobre la orientación del robot.
- Dentro de la fase de procesamiento de imágenes se ha propuesto, para cada una de las cámaras que en cada instante “ven” al robot, una completa estrategia que facilita la identificación de las balizas y permite obtener de forma precisa sus centroides, información ésta que es utilizada para la obtención de la posición y orientación del robot.



- La propuesta de procesamiento de imágenes incluye la utilización de una “ventana de interés” dinámica entorno a la posición estimada del robot. Esta propuesta permite centrar el análisis de imágenes sólo dentro de ella, lo cual reduce el efecto del ruido y al mismo tiempo la carga computacional. Para la segmentación de las balizas se ha propuesto la utilización de operadores tipo gradiente para la detección de los bordes de las balizas y a partir de las imágenes de bordes se ha propuesto una umbralización dinámica basada en el método de Otsu. Todo ello ha dado como resultado un eficiente y robusto sistema de segmentación de las balizas, que ha sido probado en entornos reales y que ha demostrado su eficacia incluso en condiciones extremas de iluminación, distancia entre cámaras y balizas, etc.
- Se ha realizado una propuesta de validación de balizas para los diferentes “objetos” que se obtienen tras la segmentación de imágenes. En nuestro caso esta validación incluye diferentes fases: eliminación de objetos de pequeñas dimensiones (teniendo presente el tamaño de las balizas, las características de la óptica utilizada y la distancia máxima a la que pueden ser captadas); identificación de elipses (dado que las balizas son circulares) utilizando la distancia algebraica (obtención de “elipses estimadas”); obtención de la excentricidad y la relación entre los perímetros de las “elipses estimadas” y los de los correspondientes objetos, lo que permite eliminar falsos candidatos. Finalmente se comprueban las relaciones geométricas propias de la ubicación física relativa entre las cuatro balizas. Esta propuesta ha sido probada en diferentes condiciones ambientales y para diferentes posiciones relativas cámaras balizas y los resultados obtenidos han sido muy satisfactorios.
- Así mismo, se ha realizado un estudio de las alternativas más importantes para la obtención de las coordenadas del centroide de las balizas, que es el punto utilizado para determinar la posición de los robots. En este sentido, se ha realizado una propuesta que consiste en considerar como coordenadas del centroide las de intersección de las diagonales principales que unen los vértices del mínimo rectángulo que contiene la baliza. Esta propuesta ha demostrado en la práctica proporcionar unos resultados altamente satisfactorios (comparables a otras alternativas propuestas en la literatura) y con la ventaja de requerir una baja carga computacional.

En conclusión, en este capítulo se ha presentado una propuesta de balizas activas (a bordo de los robots), y una estrategia de procesamiento de imágenes que permite la detección, segmentación

y localización de las balizas, de forma fiable, robusta, precisa y con tiempos de procesamiento reducidos. Además muchos de estos algoritmos de procesamiento podrían ser implementados utilizando hardware de propósito específico, por ejemplo FPGAs, lo que facilitaría su implementación práctica.



## Capítulo 6

---

### Calibración del sistema de visión

---

El objetivo principal del sistema de visión artificial utilizado en esta tesis es poder realizar mediciones en el entorno de trabajo, determinando la posición y orientación del robot, tal y como se describió en el capítulo 4. Para poder realizar medidas con una determinada precisión, mediante una o varias cámaras, es necesario conocer el modelo de comportamiento de éstas y los parámetros de dicho modelo que caracterizan a las mismas. Cada cámara, por tanto, se sustituirá por un modelo matemático, cuyos parámetros es necesario obtener. El proceso por el que se obtiene dichos parámetros es conocido como calibración. De la precisión en la obtención de los mismos dependerá la exactitud de las medidas futuras.

En este capítulo se aborda la descripción del método y patrón utilizados para calibrar el conjunto de cámaras de que consta el sistema de visión, de manera que se obtengan los parámetros intrínsecos y extrínsecos de las mismas, permitiendo obtener la posición 3D del robot.

## 6.1. Introducción.

La calibración de un sistema multicámara consiste en, además de obtener los parámetros intrínsecos correspondientes al modelo de cada cámara, determinar sus posiciones y orientaciones relativas con respecto a un sistema de referencia fijo o global (parámetros extrínsecos). Para llevar a cabo la calibración se deberán proponer, programar y ejecutar los algoritmos que permitan obtener todos los parámetros del modelo empleado para el proceso de formación de una imagen.

En el proceso de calibración es común utilizar patrones y plantillas con formas conocidas (círculos, ajedrezados, etc.), de manera que sus imágenes en los planos sensores de cada cámara sean las entradas para los algoritmos de dicho proceso. En algunos casos, la geometría de las plantillas se asume perfectamente conocida, pero actualmente son mucho más utilizados algoritmos que determinan su geometría en el propio proceso de calibración, eliminando así, los errores introducidos en la determinación de las dimensiones de las plantillas o patrones. Un patrón cuya utilización está muy extendida es el formado por  $N$  círculos con sus centros en las coordenadas 3D  $[X_i, Y_i, Z_i]^T$  ( $i = 1 \dots N$ ) los cuales se proyectan en las coordenadas  $(u, v)$  de una imagen. Captando las imágenes del patrón situado en diferentes posiciones de la escena se pueden utilizar las coordenadas de los puntos de las imágenes (centroides de los círculos en las imágenes), junto con el modelo de proyección, para poder estimar los parámetros del modelo de cámara y las coordenadas 3D de los puntos en el espacio.

Existen múltiples técnicas [Zhang 99][Heikkilä 00][Lázaro 00a][Chen 01][Gardel 04] que permiten obtener la calibración de una cámara con lente fija. Algunas de éstas usan varias imágenes del mismo patrón en diferentes posiciones para calibrar, por lo que son conocidas con el nombre de calibración multi-imagen. Los algoritmos más clásicos fueron propuestos por Brown en los setenta [Brown 71] y Tsai a finales de los ochenta [Tsai 87], en éstos el modelo de cámara utilizado es el modelo *pin-hole* modificado con una serie de términos correctores de la distorsión radial y tangencial.

Es importante mencionar también dos de los problemas más relevantes en la descripción de un proceso de calibración [Salvi 02]: la falta de una notación estandarizada que facilite la interpretación de los sistemas de coordenadas en las diferentes técnicas de calibración y la dificultad de evaluar la precisión de la técnica utilizada, ya que los métodos utilizados en cada uno de los casos no utilizan los mismos criterios.

## 6.2. Patrón de calibración.

Una vez que se dispone de un algoritmo de calibración eficiente hay que optar por seleccionar un patrón de calibración apropiado [Gardel 04] que permita obtener los menores errores en dicho proceso. En la revisión realizada a este respecto en el estado del arte se han analizado diferentes propuestas, viendo las ventajas y mejoras que introducen en los resultados de calibración cada una de ellas, en las aplicaciones para las que han sido desarrolladas. Queda demostrado que tienen una importancia capital: la forma geométrica, las distribución, las dimensiones, la iluminación, la ausencia de vibraciones, etc., de los objetos del patrón de calibración, a la hora de obtener los parámetros del modelo de cámara con determinada precisión.

### 6.2.1. Patrón de calibración propuesto.

Como se comentó en el apartado 2.5.1, la importancia del patrón elegido tiene suma importancia a la hora de llevar a cabo la calibración del sistema [Lavest 98]. La precisión de la calibración depende de la exactitud en la determinación de las coordenadas  $(u_p, v_p)$  de los puntos de interés (centro círculos, intersección de líneas, etc.), por lo que deben estar perfectamente definidos en la imagen. Así, por ejemplo, en el caso de utilizar un patrón con círculos es necesario que el contorno sea perfecto para que los algoritmos de detección subpíxelica puedan determinar el centroide con precisión, además de que el tamaño de los puntos debe ser el adecuado (en torno a unos 20 píxeles de diámetro) [Lavest 98]. Esto es difícil de conseguir en un sistema con un conjunto de cámaras fijas, colocadas en el techo del entorno, si no se quiere trabajar con plantillas de dimensiones excesivamente grandes (figura 6.1).

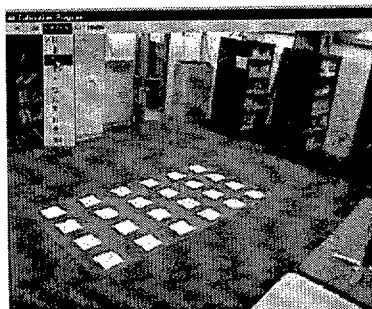


Figura 6.1. Patrón de calibración utilizado en [Hoover 00].

Otro aspecto importante a la hora de determinar el contorno de los elementos del patrón es el contraste que debe existir entre éstos y el fondo de dicho patrón. Basándonos en este hecho, ampliamente estudiado en la literatura, en lugar de utilizar un patrón pasivo construido con un material altamente reflectante con reflexión difusa casi ideal (sobre el que se proyecta una luz, también difusa), en esta tesis se propone el uso de un patrón activo, construido con diodos IR (figura 6.2), para calibrar el sistema de visión.

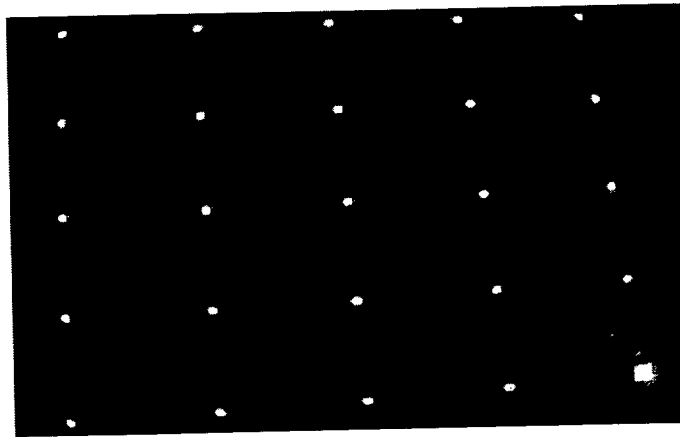
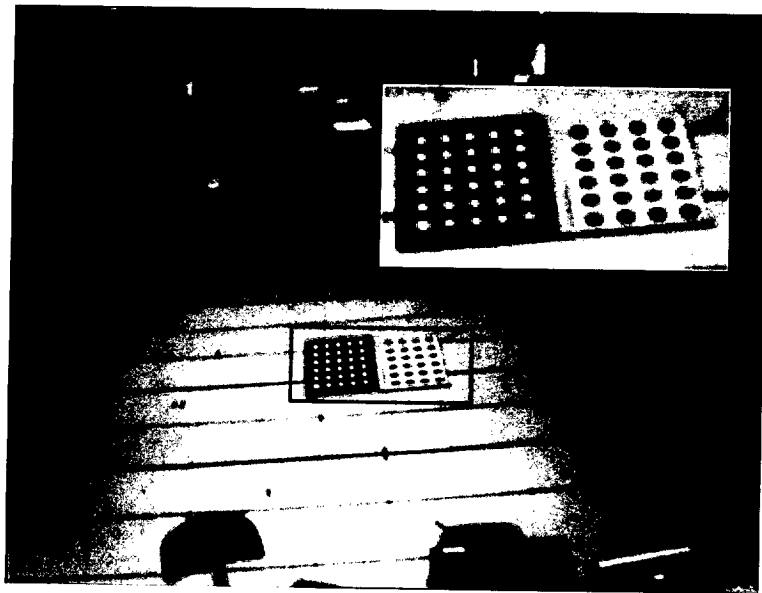


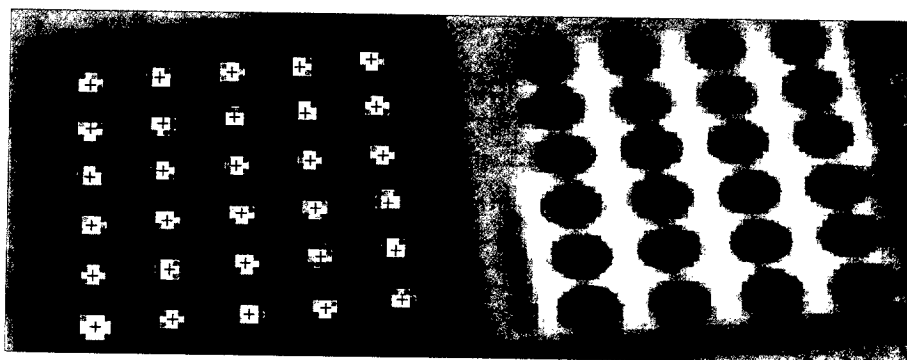
Figura 6.2. Patrón de calibración activo formado por diodos IR.

La plantilla está formada por 30 objetos luminosos (elementos de calibración), que se proyectan en las imágenes de las cámaras como puntos circulares. Éstos se distribuyen en forma matricial de cinco filas y seis columnas. Cada uno de los elementos de calibración de la plantilla está constituido por un único diodo, a excepción de uno de ellos que está construido por cuatro diodos y que es utilizado como referencia para facilitar la identificación de cada uno de los puntos, la localización y orientación del patrón, de manera que tras captar  $n$  imágenes el proceso pueda llevar a cabo la calibración sin la intervención del usuario. Se ha constatado, prácticamente, que es posible identificar todos los puntos de la baliza, situada ésta a una distancia de seis metros de la cámara (figura 6.3).

En la figura 6.3 se puede apreciar la calidad en la percepción de los puntos de calibración del patrón propuesto, en comparación con otra de uso extendido y de dimensiones similares, cuando ambas se sitúan en posiciones alejadas de una cámara. Así mismo, en la figura 6.4 se muestra la correcta identificación de los objetos de calibración del patrón propuesto, para la imagen de la figura 6.3. Observese, cómo en el otro patrón no es posible detectar todos los puntos.



**Figura 6.3.** Vista del patrón de calibración propuesto y comparación con un patrón formado por círculos negros.

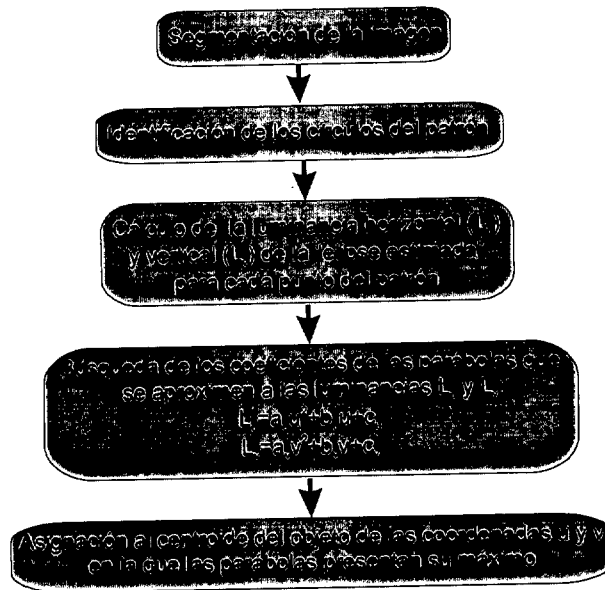


**Figura 6.4.** Identificación de los puntos de los patrones de calibraciones de la figura 6.3.

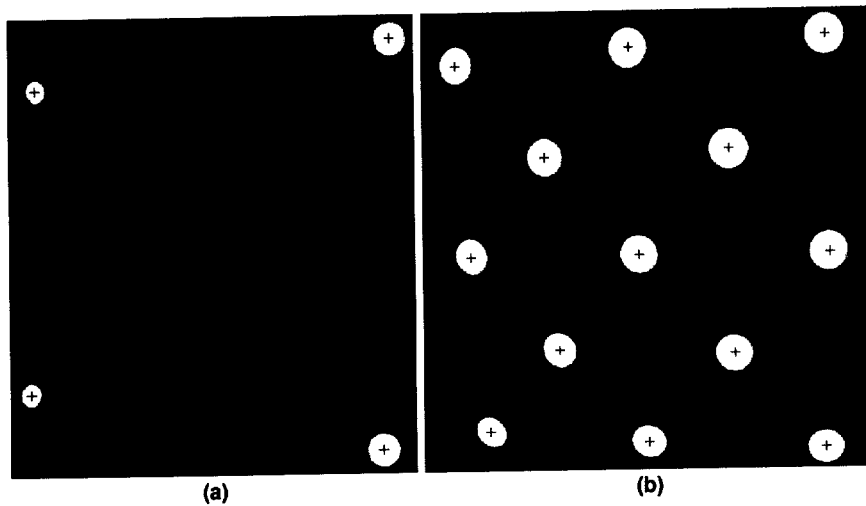
Para la obtención del centroide (figura 6.5) de cada uno de los objetos del patrón en una imagen, se utiliza el mismo procedimiento empleado en la detección de las balizas a bordo del robot (capítulo 5). Además, para obtener una mayor precisión se utiliza el método de la aproximación parabólica dada por las ecuaciones (5.33) y (5.34).

El hecho de utilizar diodos IR como objetos de calibración incrementa la calidad de proyección en las imágenes de las cámaras (figura 6.6). Se proyectan como objetos elípticos que, unido a la naturaleza del proceso de identificación utilizado (“elipses estimadas”) corrige pequeñas deformaciones, mejorando las características de las imágenes de calibración.





**Figura 6.5.** Organigrama empleado en la obtención del centroide de cada uno de los objetos del patrón de calibración.



**Figura 6.6.** Proyección de los puntos de un patrón de calibración sobre el plano imagen de una cámara. (a) Diodos IR del patrón propuesto. (b) Círculos reflectantes de un patrón estándar.

Como se comentó con anterioridad, para obtener unos buenos resultados en la calibración es preciso disponer de unas condiciones de iluminación del patrón idóneas. Así, para comprobar la eficacia del patrón propuesto se han tomado 25 imágenes de los dos patrones de la figura 6.6, ubicados siempre en la misma posición y con pequeñas variaciones de las condiciones de iluminación. De dichas imágenes se obtienen los centroides de los puntos de calibración y se analiza su desviación típica. Los resultados se muestran en la tabla 6.1. Como se puede observar en dicha tabla, los resultados obtenidos con la plantilla propuesta presentan una

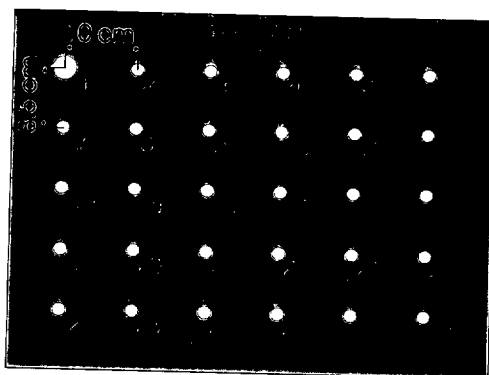
desviación típica de los centroides inferior, y por tanto, la calidad y precisión del proceso son superiores a las obtenidas utilizando una plantilla reflectante.

**TABLA 6.1.** Valor medio de la desviación típica del centroide de los puntos de un patrón de calibración frente a diferentes condiciones de iluminación.

Coordenada	Patrón construido con diodos IR	Patrón construido con círculos reflectantes
u	0.0065	1.6013
v	0.0024	0.5084

### 6.2.2. Determinación automática de los puntos del patrón.

Los algoritmos de calibración parten de una estructura de baliza conocida de la que se sabe, aproximadamente, la separación entre los objetos de calibración. En la figura 6.7 se muestra la distribución e identificación de los objetos de calibración del patrón propuesto.



**Figura 6.7.** Distribución de los objetos de calibración en el patrón

Una vez captadas las imágenes del patrón necesarias, los puntos deben ser identificados antes de proceder a la calibración del sistema, de manera que los algoritmos sepan cuál es cada punto en cada imagen. En el algoritmo propuesto en esta tesis para obtener una determinación automática de los elementos, el índice asignado en la fase de segmentación a cada punto de la baliza es aleatorio entre 1 y 30 (figura 6.8), razón por la cual es necesario reordenarlos para adaptarlos a la numeración de la figura 6.7, antes de poder proceder a la calibración.

El proceso de ordenación se basa en que los objetos de calibración del patrón se encuentran alineados, tanto en el espacio 3D como en las imágenes. En la figura 6.9 se muestra el organigrama utilizado para identificar cada uno de los objetos del patrón de calibración en una imagen.

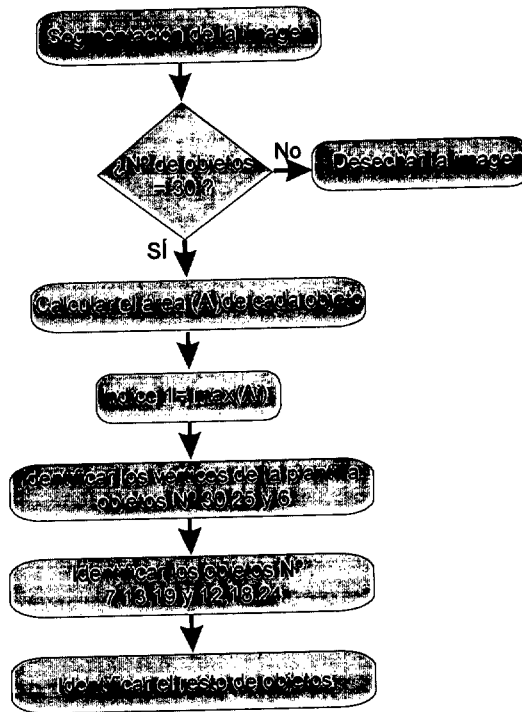


Figura 6.8. Organigrama seguido para la ordenación de los objetos del patrón de calibración.

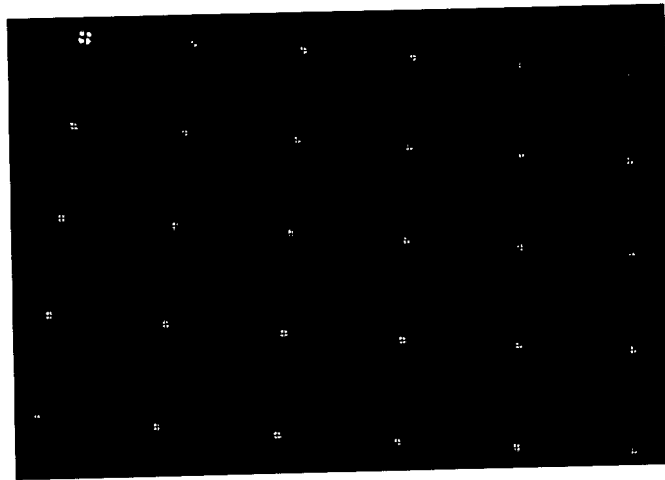


Figura 6.9. Distribución aleatoria de los índices de los puntos del patrón de calibración tras la segmentación e identificación de objetos.

Una vez segmentados los objetos de una imagen y comprobado que su número es igual a 30, se procede a su reordenación utilizando las coordenadas de sus centroides  $(u_i, v_i)$ . Para ello se realizan los pasos que se describen a continuación y que se muestran de forma gráfica en la figura 6.10:

1°. El punto de referencia  $l$ , llamado punto principal del patrón, se asigna a aquél de los 30 puntos que tenga un área mayor. En este caso se procede al renombrando de sus coordenadas  $(u_i, v_i)$  por  $(u_l, v_l)$ .

2°. Se obtiene el centro de gravedad de la baliza que utilizaremos para la clasificación de los puntos del patrón. Su valor viene dado por:

$$u_g = \frac{\sum_{i=1}^{30} u_i}{30} \quad y \quad v_g = \frac{\sum_{i=1}^{30} v_i}{30} \quad (6.1)$$

3°. Se obtiene la distancia de todos los puntos a la recta  $r_{lg}$  que pasa por  $(u_l, v_l)$  y  $(u_g, v_g)$ , siendo esta distancia:

$$d_i = \frac{(u_1 - u_g, v_1 - v_g) \wedge (u_i - u_g, v_i - v_g)}{\|(u_1 - u_g, v_1 - v_g)\|} \quad (6.2)$$

4°. Se calcula la distancia de cada punto al punto de índice  $l$  ( $d_{il}$ ):

$$d_{il} = \|(u_1, v_1) - (u_i, v_i)\| \quad (6.3)$$

El punto de índice 30 es aquél en el que  $d_{il}$  es máxima y  $d_i$  es mínima. Por tanto, se renombran sus coordenadas por  $(u_{30}, v_{30})$ .

5°. Uno de los objetos que se encuentran en las esquinas restantes del patrón será aquel que se encuentre más alejado ( $\max(d_i)$ ) de la recta  $r_{lg}$ , identificándolo sus coordenadas por  $(u_{vrt1}, v_{vrt1})$ .

6°. La esquina restante se elige entre aquellos cuya distancia a la recta que pasa por  $(u_{vrt1}, v_{vrt1})$  ( $u_g, v_g$ ) sea mínima y, a su vez, la distancia al punto principal sea máxima. Sus coordenadas se identifican por  $(u_{vrt2}, v_{vrt2})$ .

Notar que este punto no puede calcularse de manera similar al paso quinto debido a que en función de la perspectiva con la que se capte la imagen, en ésta pueden existir puntos más alejados de la recta  $r_{lg}$  que él.

7°. Para identificar los dos vértices anteriores se realiza el producto vectorial del

vector formado por los punto 1 y 30 con el vector formado por el punto 1 y uno de los vértices, por ejemplo  $(u_{vrt1}, v_{vrt1})$ , y se examina el signo:

$$signo\left((u_1 - u_{30})(v_1 - v_{vrt1}) - (v_1 - v_{30})(u_1 - u_{vrt1})\right) \tag{6.4}$$

si éste es positivo, al vértice de coordenadas  $(u_{vrt1}, v_{vrt1})$  se le asigna el índice 6, en caso contrario el 25. Por su parte el índice del punto  $(u_{vrt2}, v_{vrt2})$  tendrá el valor 25 ó 6.

8°. Una vez obtenidas las referencias de los objetos que se corresponden con los vértices del patrón de calibración, el paso siguiente consiste en la identificación de los puntos de la primera columna (7, 13 y 19) y de la última (12, 18 y 24). Para los primeros se obtienen los tres puntos más cercanos a la recta que une los puntos 1 y 25, reordenándolos en función de su distancia al punto principal (1). Para los puntos de la última columna se procede de forma análoga pero a partir de los elemntos 6 y 30.

9°. El último paso consiste en identificar los puntos de cada fila, los candidatos serán aquellos que se encuentran más cerca de la recta delimitada por los puntos primero y último de cada fila, reordenándolos en función de su distancia al primer punto de la fila.

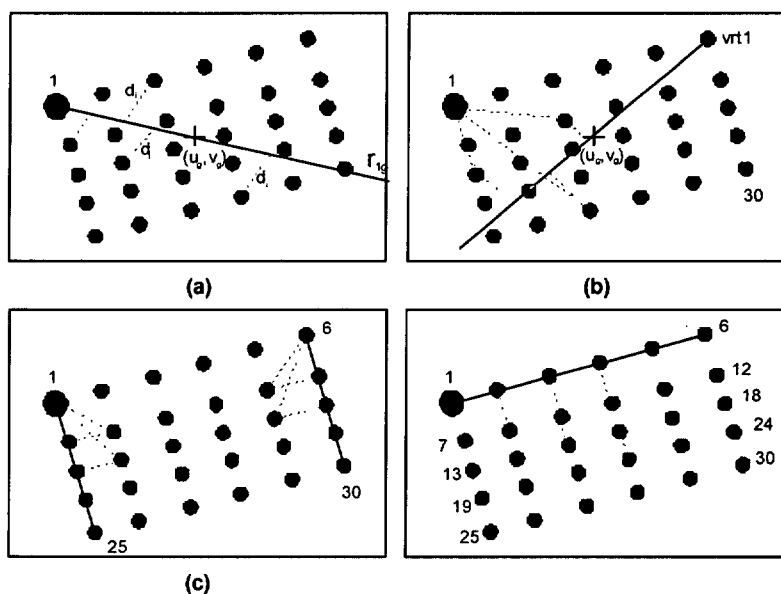
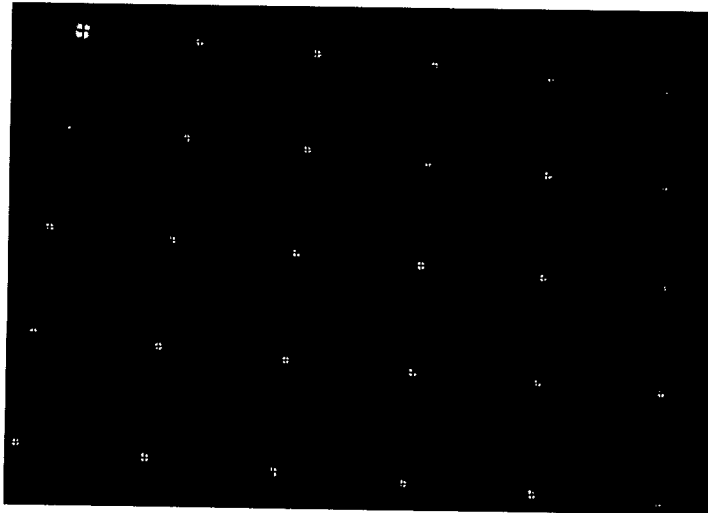


Figura 6.10. Representación gráfica de los pasos seguidos para reordenar los puntos del patrón de calibración.

En la figura 6.11 se muestra los puntos de la imagen de la figura 6.8 reordenados, de forma acorde con la distribución de la figura 6.7.



**Figura 6.11.** Representación de la figura 6.8 con los puntos reordenados de acuerdo a la numeración de la figura 6.7.

Hay que hacer notar que muchas de las herramientas [Bouguet 99] de calibración necesitan la supervisión del usuario para indicar, sobre las imágenes, el lugar donde se encuentran los elementos del patrón de calibración, dando lugar a una labor tediosa y con riesgo de equívoco que puede provocar una calibración errónea, cosa que no sucede con el proceso propuesto.

### 6.3. Proceso de calibración.

En el capítulo 4 se ha desarrollado el método de determinación de la posición 3D del robot, indicando: cómo es el error en la obtención de las coordenadas en función del número de cámaras que lo “ven”, las restricciones que se deben imponer en determinadas ocasiones, etc. El punto de partida para poder aplicar la metodología y algoritmos propuestos allí, es conocer perfectamente: *a)* los parámetros intrínsecos de cada cámara que permitan determinar la geometría de las rectas de proyección que comprenden los centros ópticos  $O_k$  de cada cámara y que pasan por los puntos  $(p_k)$  en los planos imagen y *b)* la matriz de rotación y vector de traslación de cada una de las cámaras,  $R_k$  y  $T_k$ , respecto a un sistema de referencia global (y por tanto las matrices  $R_{ij}$  y  $T_{ij}$  de parámetros de la transformación rígida entre los sistemas de referencia  $SR_i$  y  $SR_j$  de cada par de cámaras). Por tanto, para completar el proceso, se deben obtener los parámetros intrínsecos de cada cámara, y los extrínsecos entre las cámaras y el

sistema de referencia.

Tal y como se ha descrito en el capítulo 2, existen múltiples procedimientos y algoritmos de calibración para la obtención de dichos parámetros. En el apéndice A se detalla el método de calibración multi-imagen a utilizar en la calibración de nuestro sistema, mostrando la matemática y problemática asociada a la calibración de una cámara con una óptica de distancia focal fija.

Hay que tener presente que este método de calibración se basa en la localización de los elementos de un patrón y que, para calibrar los parámetros intrínsecos, calcula también la posición  $R_j$  y  $T_j$  del patrón (calibración multi-imagen) respecto del sistema de coordenadas de la cámara. Este hecho lo podemos utilizar para poder determinar con este mismo algoritmo, una vez calibrada la cámara, la posición relativa de un plano (patrón) respecto de la cámara o viceversa. Para obtener los valores iniciales de localización del patrón con respecto a la cámara, se hace uso del algoritmo de DeMenthon [DeMenthon 95], utilizándolos posteriormente como valores iniciales del algoritmo iterativo de minimización no-lineal (algoritmo de optimización de Levenberg-Marquadt) del modelo *pin-hole*.

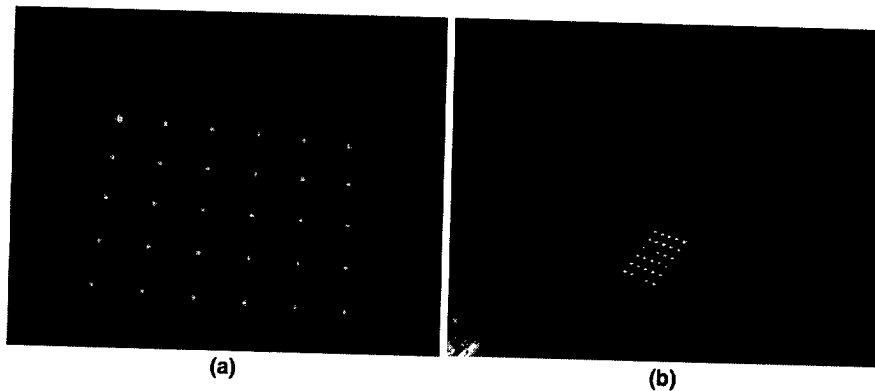
### 6.3.1. Determinación de los parámetros intrínsecos de las cámaras

Para el caso de espacios de grandes dimensiones, como es el caso que nos ocupa en esta tesis, si se desea captar el patrón desde varias cámaras para calibrarlas y obtener en el mismo proceso los parámetros de la transformación rígida ( $R_{ij}$ ,  $T_{ij}$ ) entre ellas, las posiciones en las que se debe colocar el patrón estarán muy alejadas de cualquiera de las cámaras y, como consecuencia, la precisión en la obtención de los parámetros intrínsecos no será óptima, trasladándose las imprecisiones, posteriormente, a las medidas realizadas.

En esta tesis se propone la realización de la calibración del sistema completo en dos pasos. En un primer paso se obtienen los parámetros intrínsecos con las cámaras fuera de su emplazamiento definitivo (realizado en un entorno de calibración de laboratorio), lo que permite obtenerlas con una alta precisión. En un segundo paso se calcula la posición relativa entre cámaras ( $R_{ij}$  y  $T_{ij}$ ), con éstas en su posición definitiva y utilizando los parámetros intrínsecos obtenidos en el primer paso. En ambos casos se aplican los algoritmos descritos en el apéndice A.

Para poder evaluar los beneficios de calibrar previamente los parámetros intrínsecos de cada cámara se han calculado los parámetros intrínsecos de las mismas, primero con la cámara

montada sobre una plataforma de calibración de laboratorio (una de las imágenes utilizadas se muestra en la figura 6.12.a) y segundo con la cámara localizada en su emplazamiento definitivo y el patrón en una posición donde pudieran verlo varias cámaras (una de las imágenes utilizadas se muestra en la figura 6.12.b).



**Figura 6.12.** Imágenes utilizadas para obtener los parámetros intrínsecos de una cámara. (a) Situada en una plataforma externa. (b) Situada en su posición final.

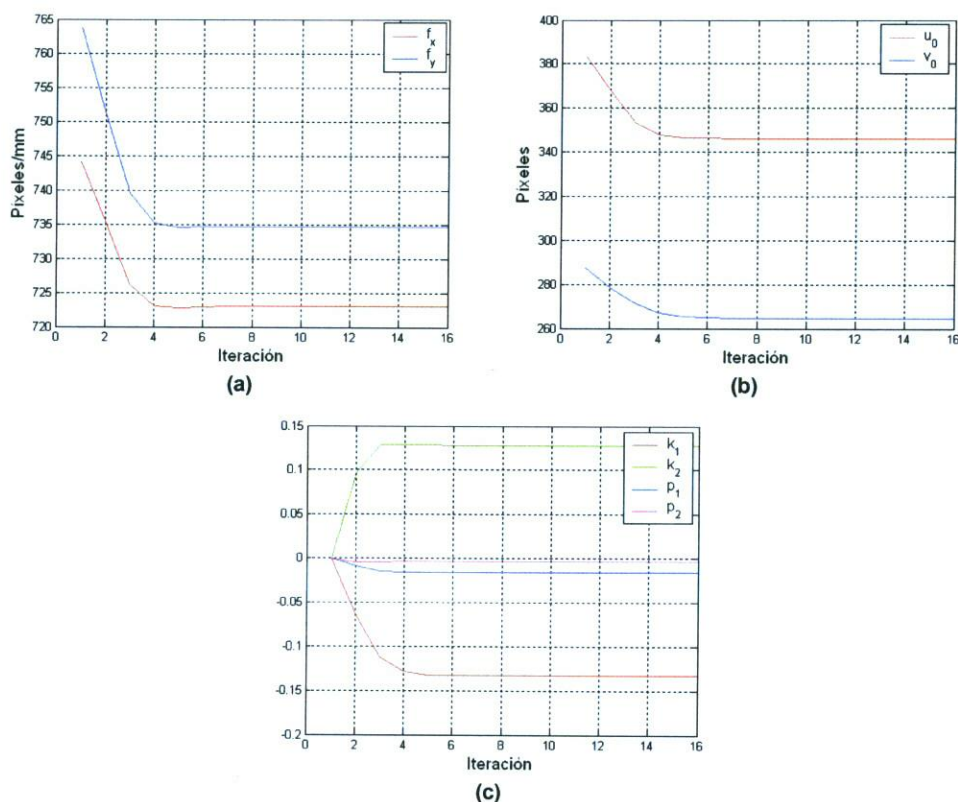
En la tabla 6.2 se muestran los resultados obtenidos en las calibraciones mencionadas en el párrafo anterior. Hay que destacar que al algoritmo iterativo de calibración se le requirió alcanzar una precisión (residuos) de 0.02 píxeles en la determinación de los parámetros. En el primer caso, el algoritmo consiguió el objetivo con muy pocas iteraciones y en el segundo caso, después de haber agotado el límite fijado para las iteraciones (200), el algoritmo se detuvo indicando una precisión de 0.3 píxeles.

**TABLA 6.2.** Comparación de los valores de los parámetros intrínsecos obtenidos con el método de calibración in situ y de laboratorio.

Parámetro	Calibración de laboratorio		Calibración in situ	
	Media	Desv. típica	Media	Desv. típica
$u_0$	346.35	1.123	346.67	2.282
$v_0$	264.85	1.009	260.72	2.733
$f_x$	723.69	0.875	709.78	2.772
$f_y$	723.11	0.856	709.56	2.762
$k_1$	-0.1321	0.003	-0.1259	0.002
$k_2$	0.1281	0.004	0.1178	0.04
$p_1$	-0.0152	0.0006	-0.0045	0.007
$p_2$	-0.0028	0.0005	-0.0067	0.006



En la figura 6.13 se muestra la convergencia de los valores de los parámetros intrínsecos utilizando el método de la obtención en laboratorio. En ella puede observarse que los valores convergen rápidamente y que a partir de la quinta iteración alcanzan su valor final.



**Figura 6.13.** Convergencia de los valores tomados por los parámetros intrínsecos de una cámara sometida al proceso iterativo de calibración en dos fases y utilizando la plantilla propuesta. (a) Valores de  $f_x$  y  $f_y$ . (b) Valores de  $u_0$  y  $v_0$ . (c) Valores de los coeficientes de la distorsión.

### 6.3.2. Obtención de los parámetros extrínsecos.

Una vez obtenidos los parámetros intrínsecos de cada cámara y con las mismas ya ubicadas en su localización definitiva, se procede a calcular la matriz de rotación y el vector de traslación entre cada cámara y el sistema de referencia global  $R_i$  y  $T_i$ . Las situaciones que podemos encontrar en los diferentes entornos de trabajo se pueden resumir en dos. En primer lugar cuando existan espacios donde el robot sea visto por una única cámara (pasillos estrechos con cámaras en sus extremos, por ejemplo) y en segundo lugar entornos donde el robot sea visto por  $n$  cámaras. En el capítulo 4 se realizó un análisis de los errores que se cometían y las restricciones que se debían imponer en el caso de detectar el robot con una y con  $n$  cámaras.

Al igual que allí, en el caso de tener que detectar el robot con una sola cámara, se supondrá que éste se desplaza sobre una superficie plana no inclinada, por lo que la coordenada Z se considerará constante,  $Z=cte$ .

En todos los casos, para obtener los parámetros  $R_i$  y  $T_i$  de cada cámara se utilizará el mismo patrón que para la obtención de los parámetros intrínsecos y el mismo algoritmo que para la calibración de cámaras que, como se comentó, es capaz de proporcionar con una muy alta precisión la posición de la cámara ( $R_i$  y  $T_i$ ) con respecto a un plano (patrón de calibración).

El patrón de calibración, para calibrar una cámara  $i$ , se sitúa en una posición arbitraria (figura 6.14) respecto del sistema de referencia global, asumiendo, además, que se conoce la transformación rígida existente (rotación y traslación) entre referencia y patrón y que vendrá dada por la matriz de rotación  $R_{pgi}$  y el vector de traslación  $T_{pgi}$ . Para calcular  $R_{pgi}$  y  $T_{pgi}$ , se hace uso de cualquiera de los métodos empleados tradicionalmente, como son teodolitos o con la ayuda de una cámara externa a nuestro sistema, previamente calibrada, que capte el patrón y la localización donde se quiere emplazar el sistema de referencia. Cualquiera de los métodos aportará una solución de  $R_{pgi}$  y  $T_{pgi}$  suficientemente precisa. Sus valores vienen dados por:

$$R_{pgi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha_{pgi}) & \text{sen}(\alpha_{pgi}) \\ 0 & -\text{sen}(\alpha_{pgi}) & \cos(\alpha_{pgi}) \end{bmatrix} \begin{bmatrix} \cos(\beta_{pgi}) & 0 & -\text{sen}(\beta_{pgi}) \\ 0 & 1 & 0 \\ \text{sen}(\beta_{pgi}) & 0 & \cos(\beta_{pgi}) \end{bmatrix} \begin{bmatrix} \cos(\gamma_{pgi}) & \text{sen}(\gamma_{pgi}) & 0 \\ -\text{sen}(\gamma_{pgi}) & \cos(\gamma_{pgi}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.5)$$

$$T_{pgi} = \begin{bmatrix} X_{pgi} \\ Y_{pgi} \\ Z_{pgi} \end{bmatrix}$$

donde  $X_{pgi}$ ,  $Y_{pgi}$  y  $Z_{pgi}$  representan la posición del punto principal respecto del sistema de referencia y  $\alpha_{pgi}$ ,  $\beta_{pgi}$  y  $\gamma_{pgi}$  son los ángulos de Euler del sistema de referencia del patrón respecto al global.

Para realizar este proceso únicamente es necesaria una imagen que nos permitirá calcular los ángulos de euler y los componentes del vector de traslación del plano patrón respecto del sistema de coordenadas de la cámara. Ahora bien, para optimizar estos parámetros se capta la imagen del patrón en diferentes posiciones y con el sistema sobredimensionado se obtienen los valores finales de los mismos. Las pruebas empíricas demuestran que, dada la precisión en la determinación de la posición del patrón, se obtiene una muy buena solución con una sola imagen.

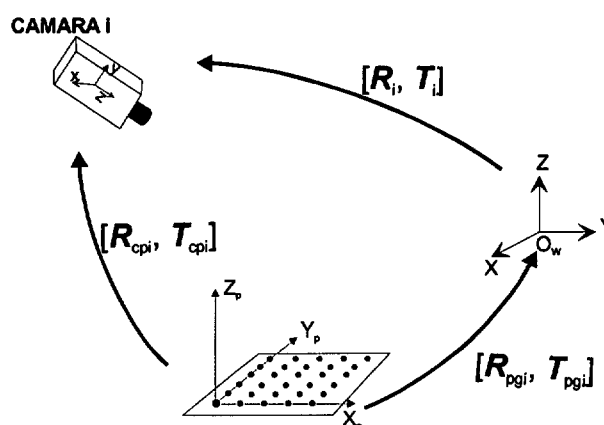


Figura 6.14. Relación entre los sistemas de referencia global, de la cámara y del patrón.

El sistema de referencia del patrón de calibración (figura 6.15) se encuentra centrado en el punto principal (de índice  $l$ ). El eje  $X$  se corresponde con la fila formada por cinco diodos IR, mientras que el eje  $Y$  se encuentra alineado con la primera columna constituida por seis diodos, como se refleja en la figura 6.7. En la misma, se puede comprobar que a pesar de la distancia del patrón a la cámara, la nitidez y contraste en la detección de sus elementos hacen del patrón activo propuesto un elemento fundamental para la calibración con suficiente calidad.

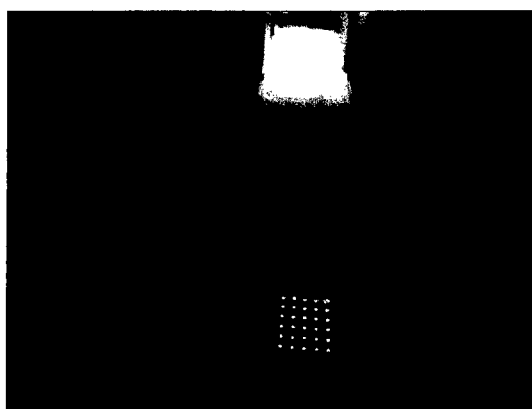


Figura 6.15. Representación del sistema de referencia del patrón de calibración.

Teniendo en cuenta lo descrito hasta este punto, se concluye que el valor de las matrices de rotación y traslación de una cámara  $i$ , respecto al sistema de referencia global, vienen dadas por:

$$\begin{aligned} R_i &= R_{cpi} \cdot R_{pgi}^{-1} \\ T_i &= T_{cpi} - T_{pgi} \cdot R_{cpi} \cdot R_{pgi}^{-1} \end{aligned} \quad (6.6)$$

donde  $R_{cpi}$  y  $T_{cpi}$  representan la matriz de rotación y el vector de traslación, respectivamente, de la cámara  $i$  respecto al sistema de referencia del patrón de calibración (parámetros extrínsecos de la cámara), obtenidos en el segundo paso de la calibración.

En la figura 6.16 se ilustra el proceso de obtención de  $R_i$  y  $T_i$  para la situación de medida con una única cámara.

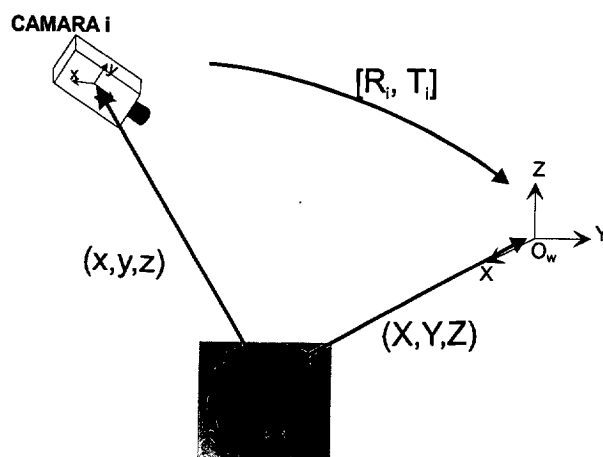


Figura 6.16. Medida de la posición del robot utilizando una única cámara.

Una vez determinados para la cámara  $i$  los parámetros de  $R_i$  y  $T_i$ , cuando dicha cámara capte la imagen de la baliza en forma de "T" utilizada por el robot (asumiendo la restricción para una sola cámara de  $Z=Z_0=Cte.$ ), el sistema podrá determinar las coordenadas 3D de su punto central  $[X, Y, Z]^T$  respecto de la cámara y la posición del mismo respecto de la referencia global (figura 6.16).

En este caso, el procedimiento y la matemática son los mismos que los descritos en el capítulo 4, siendo la ecuación de la posición de los puntos:

$$\begin{bmatrix} X \\ Y \\ Z_0 \end{bmatrix} = R^{-1} \begin{bmatrix} x \\ y \\ z \end{bmatrix} - R^{-1} \cdot T = \lambda \cdot R^{-1} \begin{bmatrix} \frac{u - u_0}{f_x} \\ \frac{v - v_0}{f_y} \\ 1 \end{bmatrix} - R^{-1} \cdot T \quad (6.7)$$

Para la situación en la que el robot sea captado por varias cámaras, el proceso de obtención de

las diferentes  $R_i$  y  $T_i$  correspondientes a cada una, es similar al descrito. En este caso, no se asume ninguna restricción y se obtienen las coordenadas 3D del robot  $[X, Y, Z]^T$  respecto de la referencia global, a partir de las imágenes de las diferentes cámaras.

En la figura 6.17 se muestra el proceso de obtención de las diferentes  $R_i$  y  $T_i$  entre cada cámara y el sistema de referencia global, así como de las diferentes  $R_{ij}$  y  $T_{ij}$  que representan la rotación y traslación relativa entre los sistemas de referencia locales de las cámara  $i$  y  $j$ .

$$R_{ij} = R_i \cdot R_j^{-1} \quad y \quad T_{ij} = T_i - R_i \cdot R_j^{-1} \cdot T_j \quad (6.8)$$

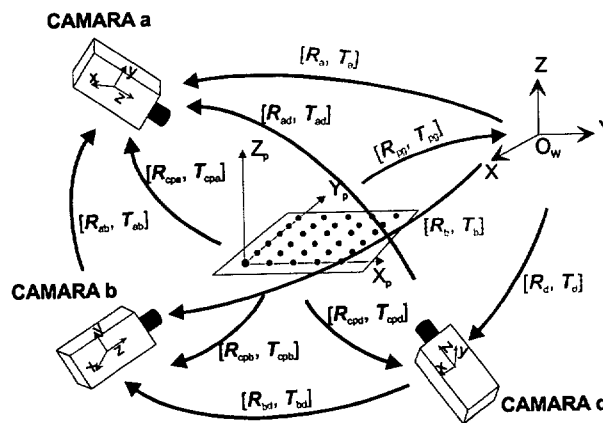


Figura 6.17. Relación entre los sistemas de referencia para el caso de  $n$  cámaras.

Determinados los parámetros de  $R_i$ ,  $T_i$  y  $R_{ij}$ ,  $T_{ij}$ , cuando las cámaras captan la imagen de la baliza en forma de "T", el sistema podrá determinar las coordenadas 3D de su punto central  $[X, Y, Z]^T$  respecto de la cámara y la posición del mismo respecto de la referencia global (figura 6.18), según el procedimiento descrito en el capítulo 4. En este caso, el procedimiento y la matemática son los mismos que los descritos allí. A continuación se detalla la matemática para la determinación de la posición del robot:

$$X = \frac{\sum_{i=1}^{2M} X_i \cdot \sigma_{Xi}^2}{\sum_{i=1}^{2M} \sigma_{Xi}^2}, \quad Y = \frac{\sum_{i=1}^{2M} Y_i \cdot \sigma_{Yi}^2}{\sum_{i=1}^{2M} \sigma_{Yi}^2}, \quad Z = \frac{\sum_{i=1}^{2M} Z_i \cdot \sigma_{Zi}^2}{\sum_{i=1}^{2M} \sigma_{Zi}^2} \quad (6.9)$$

siendo  $\sigma_{Xi}^2$ ,  $\sigma_{Yi}^2$ ,  $\sigma_{Zi}^2$  las varianzas de los puntos de cruce de las rectas de proyección tomando

dos a dos las cámaras que “ven” al robot y  $X_i$ ,  $Y_i$  y  $Z_i$  sus coordenadas.

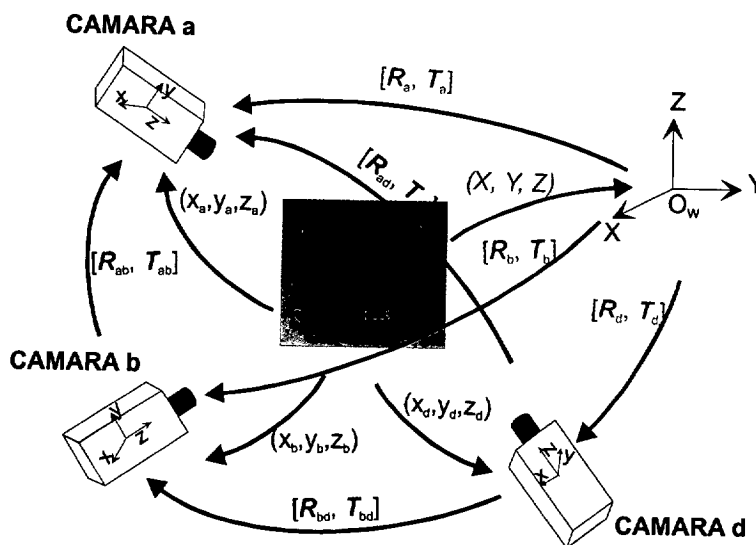


Figura 6.18. Proceso de medida de la posición del robot con  $n$  cámaras.

### Errores transmitidos de $R_{pgi}$ y $T_{pgi}$ a $R_i$ , $T_i$ .

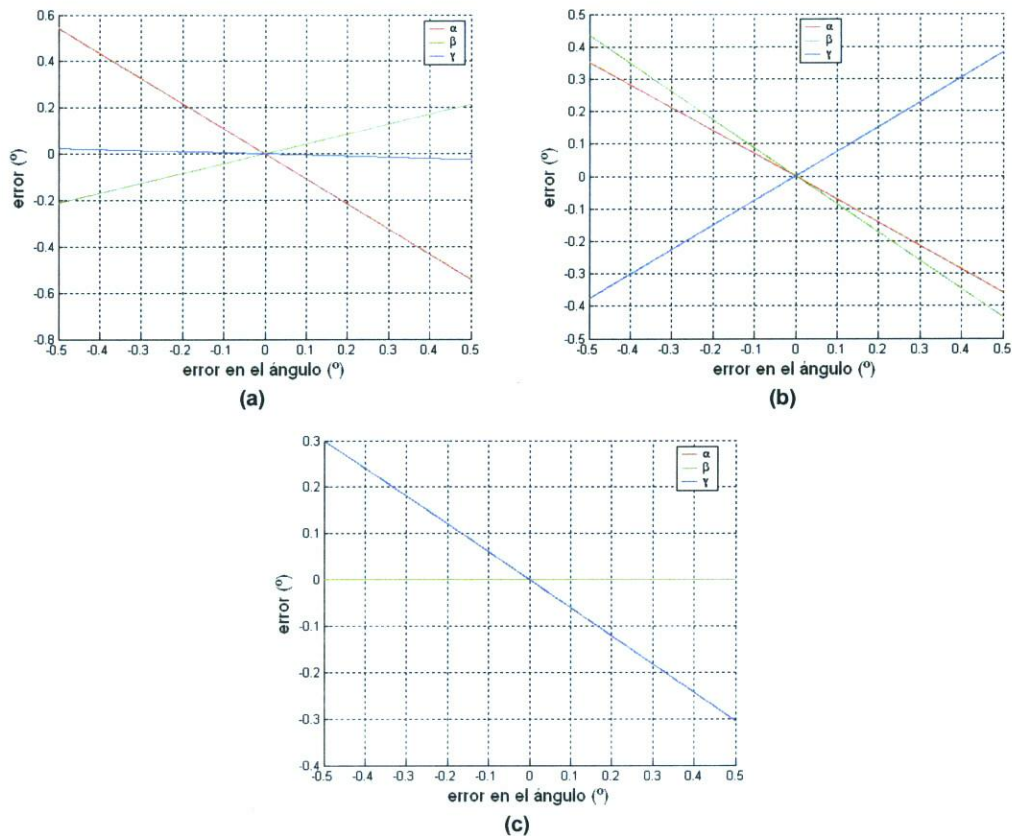
En los procesos de obtención de los parámetros extrínsecos del sistema detallado hasta este momento, la determinación de  $R_{pgi}$  y  $T_{pgi}$  puede llevar implícito un error que se transmitirá hasta los parámetros de  $R_i$ ,  $T_i$  de las diferentes cámaras. En este punto se va a analizar cuál es la influencia de los errores cometidos en la determinación de la posición del patrón en la obtención de los parámetros extrínsecos ( $T_x$ ,  $T_y$ ,  $T_z$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ) de cada cámara.

Si consideramos las ecuaciones (6.5) y (6.6) donde,  $X_{pgi}$ ,  $Y_{pgi}$ ,  $Z_{pgi}$ ,  $\alpha_{pgi}$ ,  $\beta_{pgi}$  y  $\gamma_{pgi}$  son, respectivamente, las coordenadas del punto principal de la baliza y la rotación de los ejes de la misma respecto el sistema de referencia global, si en alguno de estos parámetros hubiese un error, repercutiría en el valor de los parámetros ( $T_x$ ,  $T_y$ ,  $T_z$ ,  $\alpha$ ,  $\beta$  y  $\gamma$ ) de las cámaras con una influencia como la que se muestra en las figuras 6.19 y 6.20.

### 6.3.3. Comprobación de resultados.

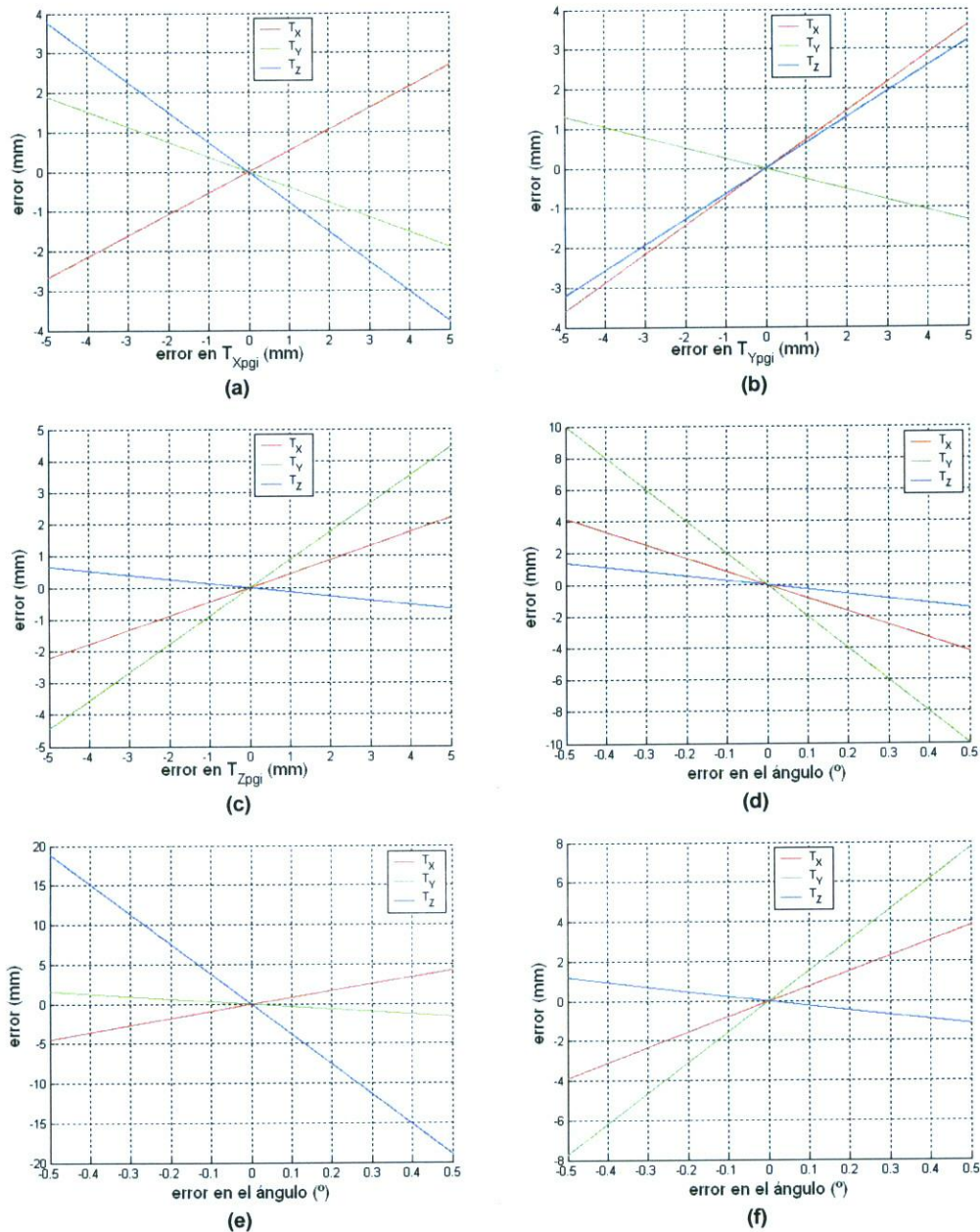
Para comprobar la eficacia, tanto del método como de la plantilla de calibración propuestos se han comparado los resultados con los obtenidos utilizando las herramientas de calibración de Matlab® [Bouquet 99]. Estas herramientas utilizan patrones del tipo damero en el cual se utilizan los vértices de los cuadrados blancos como elementos de calibración. Estas

herramientas proporcionan además unos buenos resultados cuando la plantilla de calibración abarca la totalidad de la imagen. Esto obliga a trabajar con patrones muy grandes y difíciles de manejar. Así, la utilizada en esta comprobación, para el cálculo de los parámetros extrínsecos, posee ocho puntos (figura 6.21.b) separados 400 mm. En la figura 6.21.a se muestra la posición del patrón de infrarrojos para obtener los parámetros extrínsecos de una cámara y la del patrón utilizado con herramienta de calibración de Matlab<sup>®</sup> (figura 6.21.b).



**Figura 6.19.** Influencia del error en las determinación de los ángulos de Euler de la matriz  $R_{pgi}$  en los ángulos de Euler la cámara. (a) Error en  $\alpha_{pgi}$ . (b) Error en  $\beta_{pgi}$ . (c) Error en  $\gamma_{pgi}$

En la tabla 6.3 se muestran los parámetros extrínsecos de la calibración de una cámara, utilizando el método propuesto en esta tesis y las herramientas de calibración de Matlab<sup>®</sup>, ampliamente utilizadas por muchos investigadores [Aitenbichler 03]. En sí esta tabla representa únicamente valores de la determinación de  $R_{cpi}$  y  $T_{cpi}$  con ambos métodos, sin poder detallar cuál es mejor. Sin embargo, si a partir de estos parámetros obtenidos se realiza el proceso inverso para determinar la posición de un punto de la plantilla, se puede deducir cuál comete un error mayor.



**Figura 6.20.** Variación de los parámetros del vector de traslación de una cámara debida al error en la posición del patrón de calibración. (a) Error en  $T_{xpgi}$ . (b) Error en  $T_{ypgi}$ . (c) Error en  $T_{zpgi}$ . (d) Error en  $\alpha_{pgi}$ . (e) Error en  $\beta_{pgi}$ . (f) Error en  $\gamma_{pgi}$ .

Así, con los parámetros de la tabla 6.3 se procede al cálculo del error de posicionamiento ( $e_p$ ) cometido al determinar la posición de cada uno de los puntos de la plantilla de calibración en posiciones diferentes a la utilizada para la calibración. Los resultados se muestran en la figura 6.22. Hacer notar que la precisión obtenida con el método propuesto en esta tesis y con el



patrón activo (en color azul) es muy superior a la obtenida con la plantilla de cuadrados pasiva.

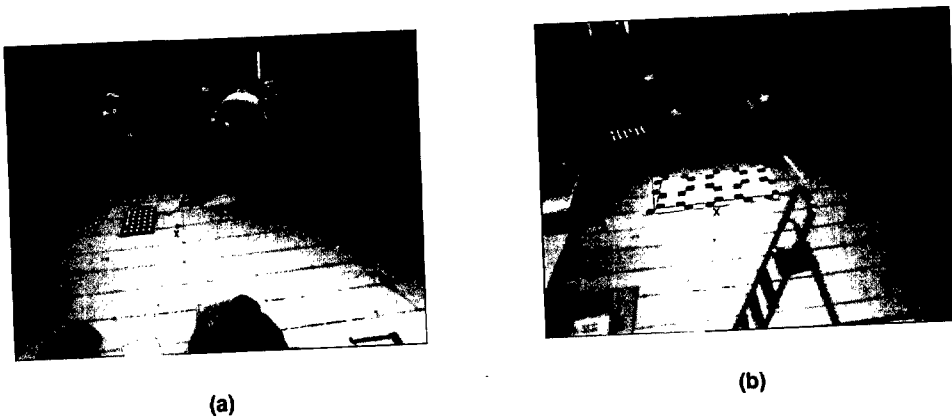


Figura 6.21. Patrones de calibración utilizados para la obtención de los parámetros extrínsecos de una cámara. (a) Plantilla propuesta. (b) Plantilla utilizada con la herramienta de calibración de Matlab®.

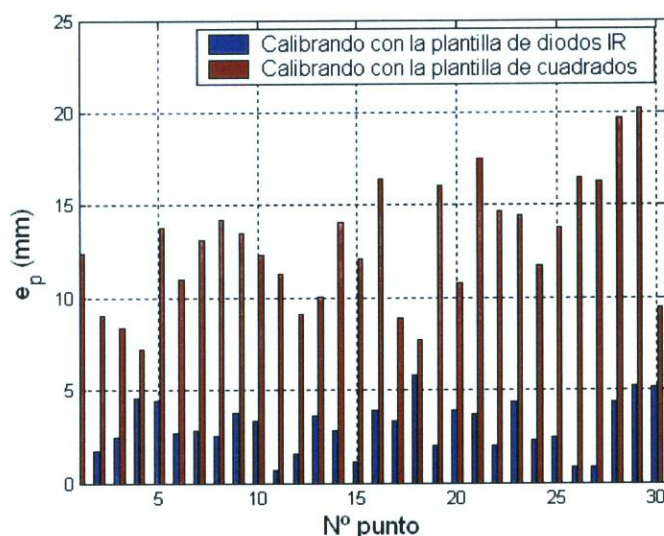
**TABLA 6.3.** Comparación de los valores de los parámetros extrínsecos obtenidos con el método propuesto en esta tesis y con la herramienta de calibración de Matlab®.

Parámetro	Método propuesto en esta tesis.	Herramienta de calibración de Matlab®
	Valor	Valor
$\alpha$ (°)	-29.42	-30.78
$\beta$ (°)	15.68	15.49
$\gamma$ (°)	-129.31	-129.07
$T_x$ (mm)	-1895	-1851
$T_y$ (mm)	1099	1113
$T_z$ (mm)	2692	2678

## 6.4. Conclusiones.

En este capítulo se ha analizado el proceso de calibración de los parámetros intrínsecos para cada una de las cámaras que conforman el sistema de visión y la calibración del sistema global para poder obtener la posición del robot en el espacio de observación del sistema. Así mismo, se han realizado propuestas tanto de algoritmos como de diseño y utilización de modelos y patrones que aportan un valor añadido a los procesos de calibración, fundamentalmente en

sistemas a utilizar en espacios de grandes dimensiones.



**Figura 6.22.** Cálculo del error de posicionamiento utilizando el método de calibración propuesto en esta tesis (en color azul) y la herramienta de calibración de Matlab<sup>®</sup> (en color rojo).

Se propone la utilización de una patrón de calibración novedoso que, por una parte simplifica el proceso de calibración de cada cámara y por otra, permite obtener con una buena precisión los parámetros extrínsecos de posicionamiento relativo entre cámaras y respecto del sistema de referencia global. Este patrón, frente a los tradicionales, introduce la variante de ser activo (realizado con diodos IR). Como puede deducirse de los análisis y resultados de las pruebas y calibraciones realizadas con el mismo, queda patente que este tipo de patrón es mucho más insensible a posibles variaciones de iluminación, tamaño de captura del mismo, perspectiva, etc., que los pasivos, permitiendo obtener un buen contraste entre los objetos de calibración y el fondo y unas características muy estáticas, por lo que cabe concluir que para poder calibrar en situaciones no idealizadas suministra resultados de mayor precisión. Otra de las conclusiones obtenidas es que la desviación típica en la determinación de los centroides de los elementos de calibración es muy inferior al caso de los patrones tradicionales. Igualmente, cuando se utiliza este patrón para poder calibrar los parámetros extrínsecos del sistema global, donde el patrón debe estar muy alejado (dada la posición de las cámaras), se demuestra que se puede obtener la imagen de los puntos de calibración con gran nitidez y contraste, conservando en un alto grado la proporcionalidad de su forma. Esto permite a los algoritmos de detección localizarlos y obtener su centroide con calidad suficiente para determinar su posición y orientación con precisión.

En relación con la identificación automática, en las imágenes de calibración, de la orientación del patrón, de su posición y las de los elementos que lo componen, se deduce y concluye que sí es posible llevarla a cabo y poder hacer que todo el proceso sea autónomo, utilizando el algoritmo propuesto al efecto en esta tesis. Si el patrón utilizado tiene una distribución regular, el algoritmo propuesto permite, de una forma sencilla y con baja carga computacional que todo el proceso se realice automáticamente sin la intervención del usuario, con las consiguientes ventajas de rapidez, comodidad, precisión, etc.

Analizada y estudiada la posibilidad de realizar la calibración global del sistema en una sola fase, se concluye que este procedimiento llevaría a obtener los parámetros intrínsecos de las cámaras con una baja precisión debido a la lejanía en que se debería ubicar el patrón de calibración para ser captado por varias cámaras, dadas las dimensiones del entorno a cubrir y la separación y altura de colocación de dichas cámaras. Por ello, se ha propuesto en esta tesis un proceso de calibración en dos fases, calibrando primero los parámetros intrínsecos de las cámaras, independientemente, fuera de su emplazamiento definitivo y posteriormente los parámetros extrínsecos del sistema global. A tenor de los resultados obtenidos y la comprobación de resultados, cabe concluir que el proceso propuesto y la secuencia de fases son correctos e idóneos para calibrar el sistema visión en este tipo de entornos para poder llevar a cabo la calibración del sistema global de medida y que suministran resultados óptimos.

## Capítulo 7

---

### Guiado

---

El guiado es la tarea que indica a un robot móvil el camino o trayectoria a seguir para llegar desde la posición actual a un destino, teniendo en cuenta el entorno que le rodea. Esta tarea debe asegurar que los caminos propuestos sean: seguros (estén libres de obstáculos), lo más cortos posibles, y dinámicos; pudiendo ser modificados ante una serie de eventos como puede ser un error en la posición del robot o la presencia de un obstáculo no previsto. En este capítulo se abordarán todos los procesos implicados en la tarea de guiado utilizando como información fundamentalmente la proporcionada por el sistema de visión formado por las cámaras ubicadas en el entorno de movimiento del robot (“espacio inteligente”).

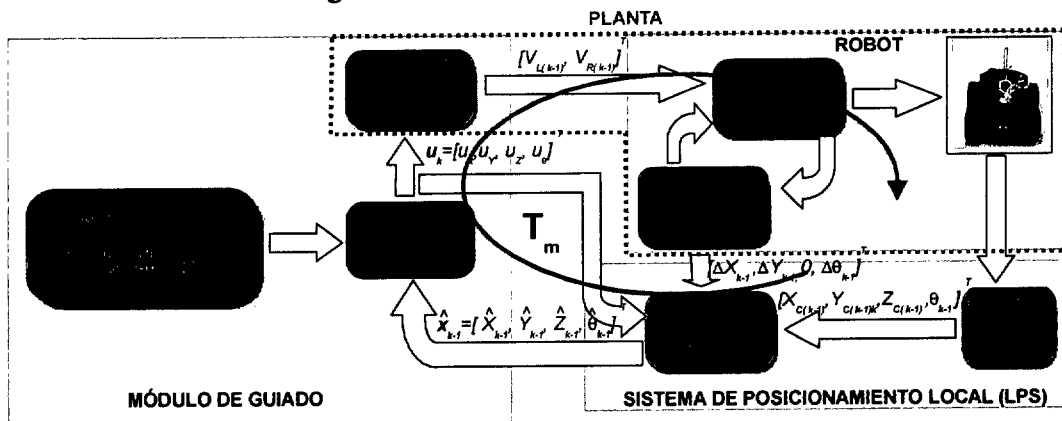
#### 7.1. Configuración general del sistema de guiado.

Como es sabido, el guiado de un robot móvil tiene como objetivo conseguir que el robot alcance un punto destino a partir de un punto de inicio, pasando por determinados puntos intermedios. En definitiva se trata de conseguir que el robot ejecute una determinada trayectoria entre los puntos de partida y final. Para conseguir este objetivo es necesario realizar diferentes tareas, las cuales se pueden clasificar en: posicionamiento, planificación de rutas, generación de trayectorias, y seguimiento de trayectorias, cumpliendo cada una las funciones resumidas

a continuación:

- *Posicionamiento*: Es sin duda una de las más importantes, y tiene como objetivo conocer las coordenadas espaciales  $([X, Y, Z]^T)$  y orientación del robot  $(\theta)$ .
- *Planificación de rutas*: Esta tarea tiene como función obtener los objetivos intermedios que deben ser alcanzados por el robot dependiendo de su posición actual y final.
- *Generación de trayectorias*: Su misión es obtener una función continua con cambios de dirección suave que representa el camino que debe seguir el robot.
- *Seguimiento de trayectorias*: Constituye la última fase y es la encargada de generar las órdenes al robot (velocidades de las ruedas activas) para que éste efectúa los desplazamientos para seguir de la forma más fiable posible la trayectoria deseada.

En la figura 7.1 se muestra la estructura general del sistema de guiado propuesto para esta aplicación, y donde se reflejan los módulos más importantes encargados que realizar las diferentes tareas asociadas al guiado del robot.



**Figura 7.1.** Estructura del sistema de control utilizado para realizar el guiado de robots por un entorno a partir de la información proporcionada por un array de cámaras estáticas y utilizando la odometría como sistema auxiliar.

Como se puede observar en la figura 7.1, se han identificado dos grandes bloques: módulo de guiado y sistema de posicionamiento local (LPS). El primero se encarga de generar las consignas de velocidad a enviar al robot, y el segundo es el encargado de obtener una posición fiable del robot a partir del sistema sensorial.

Para la implementación del módulo de guiado se parte de un mapa del entorno, de las especificaciones de la tarea de navegación (punto de partida y punto de destino) y la posición actual del robot. Con estos datos, en cada instante, se calcula la trayectoria o camino que debe

seguir el robot, de forma que se eviten los obstáculos presentes. A partir de la trayectoria, el generador de consignas obtiene las velocidades de las ruedas activas del robot que permitan alcanzar un objetivo deseado; un punto de la trayectoria.

En cuanto al sistema de posicionamiento local (LPS), en nuestro caso se propone para su implementación utilizar un Filtro de Kalman, que es el encargado de proporcionar la estimación de la posición del robot, a partir de la información de entrada proporcionada por el sistema de visión y el sistema odométrico. La necesidad de utilizar odometría, se justifica por el hecho de que desde que el sistema de visión realiza una captura del entorno hasta que se estima la posición del robot transcurre un tiempo  $T_m$ , durante el cual, el robot continua en movimiento. Así, a la posición proporcionada por el sistema de visión hay que añadirle el desplazamiento ocurrido durante  $T_m$ , el cual es proporcionado por el sistema de odometría. En lo que sigue se describen en detalle las propuestas realizadas en esta tesis para la implementación de cada uno de los diferentes subsistemas que componen los módulos principales del sistema de guiado.

## 7.2. Modelado del entorno.

En todo sistema de navegación de robots móviles es necesario conocer total o parcialmente el entorno por el cual se desplaza el robot. Este conocimiento puede ser previo (utilización de mapas CAD) [Santiso 03] o iterativo (a medida que se desplaza el robot se va construyendo el mapa del entorno) [Fraundorfer 02].

En nuestra propuesta se parte de que el entorno es conocido mediante la descripción 2D proporcionada por un mapa geométrico (figura 7.2). Este mapa contiene las dimensiones físicas del entorno así como la posición, forma y dimensiones de los obstáculos fijos (columnas, paredes, muebles, etc.) que forman parte del mismo. Estos obstáculos normalmente vendrán especificados con formas poligonales definidos por sus vértices. Además, en dicho entorno se deberá especificar cuáles serán los puntos de partida y destino que deberá alcanzar el robot.

Por otra parte, es necesario conocer las características geométricas y capacidad de movimiento del robot, que junto con el mapa geométrico, nos permita generar un modelo del entorno del que se puedan sacar conclusiones acerca de las posibles trayectorias a seguir. En este sentido, todo robot puede ser representado en el espacio de las configuraciones [Latombe 91], también

llamadas posiciones, entendiendo por configuración o posición, un vector  $x$  cuyas componentes proporcionan información completa sobre las coordenadas 3D ( $[X, Y, Z]^T$ ) y la orientación  $\theta$  del mismo:

$$x = [X, Y, Z, \theta_{XY}, \theta_{XZ}, \theta_{YZ}]^T \quad (7.1)$$

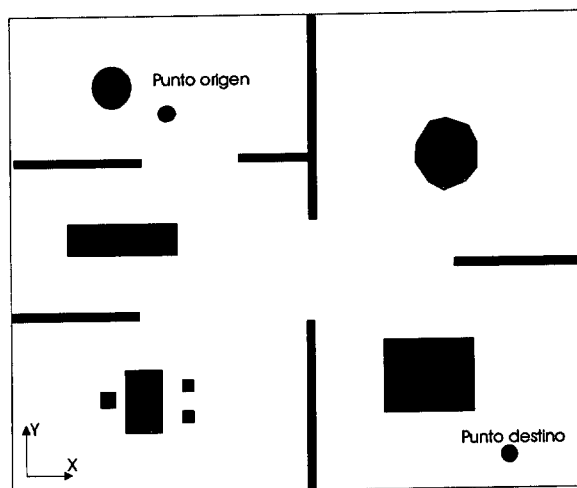


Figura 7.2. Ejemplo de mapa geométrico de un entorno

Para el caso de que el robot se mueva sobre el plano  $XY$  (figura 7.3) la ecuación (7.1) se convierte en:

$$x = [X, Y, \theta]^T \quad (7.2)$$

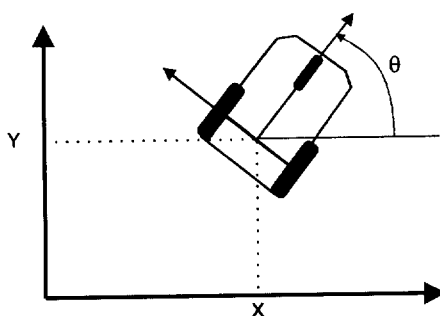


Figura 7.3. Definición de los elementos de la configuración o posición de un robot móvil en el plano  $XY$ .

Así mismo, se denomina espacio de configuraciones,  $C$ , de un robot,  $R$ , a todas las configuraciones  $x$  que pueda tomar el robot en su entorno de trabajo. El subconjunto de  $C$  ocupado por el robot  $R$ , cuando éste se encuentra en  $x$ , se denota por  $R(x)$ . Si el robot se

modela de forma circular con radio  $\rho$ ,  $R(x)$  se define como:

$$R(x) = \{x_i \in C / \|x, x_i\| \leq \rho\} \quad (7.3)$$

En el caso de un robot puntual ( $\rho=0$ ), la ecuación (7.3) se convierte en:

$$R(x) = \{x_i\} \quad (7.4)$$

Suponiendo que por el espacio de configuración  $C$  se encuentran distribuidos una serie de obstáculos definidos como polígonos  $\Psi$ :

$$\Psi = \{b_1, b_2, b_3 \dots b_n\} \quad (7.5)$$

donde  $b_i(x)$  representa al conjunto de configuraciones del espacio  $C$  ocupadas por un obstáculo, el espacio libre de obstáculos vendrá dado por:

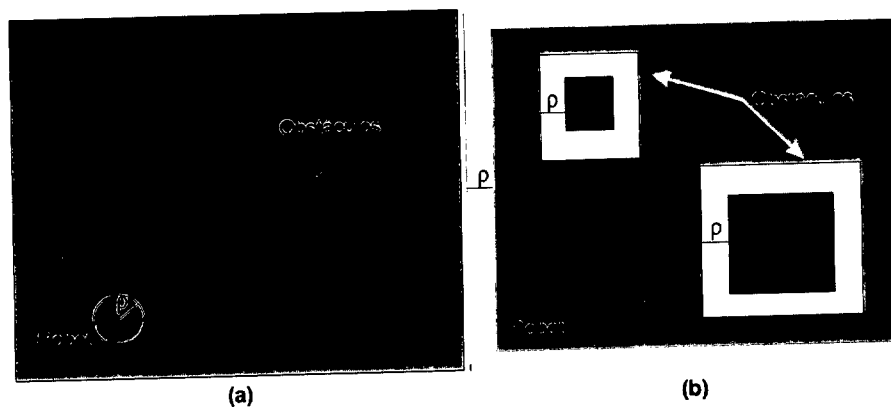
$$C_i = \{x \in C / R(x) \cap (\bigcup_{i=1}^n b_i(x_i)) = \emptyset\} \quad (7.6)$$

La construcción física de los robots móviles (formas y tamaño), junto con la distribución de los obstáculos del entorno, pueden invalidar o prohibir ciertas configuraciones. Por ejemplo, un robot con forma rectangular en un pasillo estrecho, en algunos casos, sólo puede asumir ciertas orientaciones, aquellas que siguen la dirección del pasillo. En consecuencia, el mayor problema asociado con la ubicación del robot en el espacio de las configuraciones es la complejidad de su representación y la búsqueda del espacio libre. Conocidas las dimensiones de los obstáculos, éstas se "ensanchan" en una proporción equivalente a las dimensiones del polígono que representa el robot móvil, más un margen de seguridad que garantice el movimiento libre de colisiones [Lozano 83][Foux 93][Lázaro 98]. Así, se puede asumir que el robot puede ser representado como un punto y es capaz de realizar movimientos omnidireccionales. Para el caso de que los robots sean de forma circular, un mecanismo de seguridad para el guiado consiste en dilatar todos los obstáculos del entorno en una cantidad igual al radio del robot (figura 7.4).

A la hora de realizar la expansión del entorno, es preciso considerar que su modelo se compone de dos clases de cuerpos, un primer grupo formado por los obstáculos definidos según el conjunto  $\Psi$  y un segundo que es el polígono  $p_L$  (polígono límite) que define los

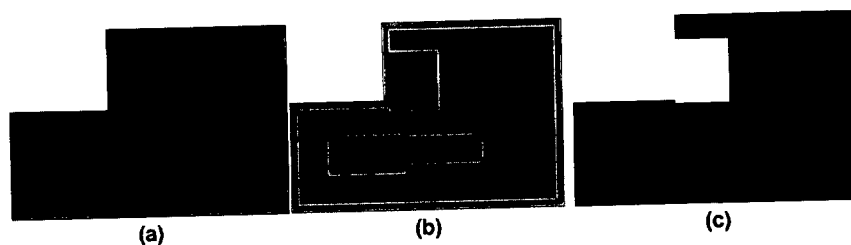


límites físicos del espacio de trabajo del robot, el cual, especifica los confines del espacio de configuraciones  $C$ . El tratamiento que se debe efectuar en cada uno de estos elementos es diferente, ya que mientras que los obstáculos se expanden hacia fuera, el polígono límite  $p_L$  sufre una contracción. Al considerar las dos clases de componentes que posee el modelo del entorno, el algoritmo de expansión consiste en aplicar el tratamiento adecuado a cada uno de las partes integrantes, para, de forma posterior, unirlos en la versión expandida. Esta acción se realiza en tres fases: expansión de obstáculos y contracción del polígono límite, cálculo del nuevo polígono límite, cálculo del nuevo conjunto de obstáculos.



**Figura 7.4.** Obtención del espacio libre de obstáculos para el caso de un robot circular de radio  $\rho$ . (a) Estructura del entorno. (b) Espacio libre tras la dilatación de los obstáculos y contracción del polígono límite.

De este modo, se verifica que el nuevo entorno consta de un número de obstáculos inferior o igual al original, lo que supone cierta ventaja a la hora de la construcción de la trayectoria del robot móvil. En la figura 7.5 se muestra gráficamente el proceso para realizar la expansión de obstáculos y la contracción del polígono límite, para obtener el espacio libre por el que se puede mover el robot.



**Figura 7.5.** Obtención del espacio libre de obstáculos. (a) Entorno original. (b) Expansión de obstáculo y contracción del entorno. (c) Representación (en color azul) del espacio libre.

### 7.3. Generación de trayectorias.

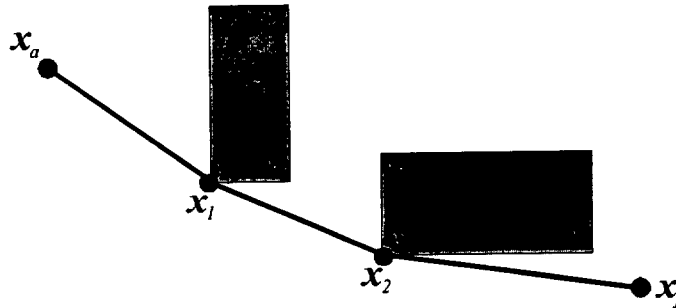
Se puede considerar el proceso de generación de trayectorias como la búsqueda de las configuraciones ( $x$ ) que debe tomar un robot móvil para llegar desde una posición de partida hasta otra final, a través del entorno de trabajo, evitando cualquier contacto con los obstáculos. Partiendo del supuesto de que el entorno es conocido en su totalidad, resulta factible la utilización de un algoritmo de búsqueda en grafos, que emplee cierta función de coste para la obtención de la trayectoria. Sin embargo, la aplicación directa de esta metodología implica efectuar ciertas consideraciones acerca del robot móvil. Para ello, se supone un robot puntual, omnidireccional, y que recorre los caminos consignados de forma perfecta. Las suposiciones anteriores son limitaciones que deberán ser eliminadas por los sucesivos módulos del sistema de navegación, como puede ser la dilatación de obstáculos o la generación de trayectorias mediante curvas.

En primer lugar se deben obtener los puntos por los que debe pasar el robot, el conjunto de estos puntos es lo que se conoce como ruta. A la hora de generar una ruta segura en un entorno totalmente conocido, se pueden utilizar diferentes métodos de planificación, como son los grafos [Nilsson 69] y subgrafos de visibilidad [Rao 88], los diagramas de Voronoi [Rombaut 91], o la descomposición en celdas [Thorpe 84].

#### 7.3.1. Planificación con subgrafos de visibilidad.

Este método se caracteriza por no generar el grafo de visibilidad completo. En vez de generar todas las rutas posibles entre la posición inicial ( $x_a$ ) y final ( $x_f$ ) y elegir la más corta, se va generando una ruta en la que los segmentos de la mismas sean los más cortos posibles. Partiendo de que el robot se encuentra en su posición inicial, se elige entre todos los vértices de los obstáculos visibles, desde dicha posición, aquél cuya distancia a la posición final del robot sea menor. El punto de partida y el vértice elegido se marcan como puntos de paso, y se introducen en la ruta. El proceso se repite tomando dicho vértice como nuevo punto de observación y así sucesivamente hasta alcanzar la posición final. En el caso de que desde un punto no se tenga visibilidad de vértices y no sea el punto destino, se retrocede al vértice anterior y se selecciona, entre los visibles, el más cercano al punto destino que no esté marcado como punto de paso. En la figura 7.6 se muestra cómo se calcula la función ruta desde la posición de partida  $x_a$  hasta la final  $x_f$ . Partiendo de  $x_a$ , se tiene que el conjunto de vértices visibles  $v_2$  y  $v_3$ , que coinciden con los vértices que pueden ser unidos en un grafo de visibilidad con  $x_a$ . Si  $v_3$  es el vértice más cercano a  $x_f$ , se tomará como punto de paso de la ruta, a la vez

que se convierte en el nuevo vértice que debe expandirse según la metodología anterior, obteniéndose  $v_6$  como el siguiente punto de paso y desde éste ya se alcanza  $x_f$ . Así, la ruta estará formada por las configuraciones  $\{x_a, x_1, x_2, x_f\}$ .



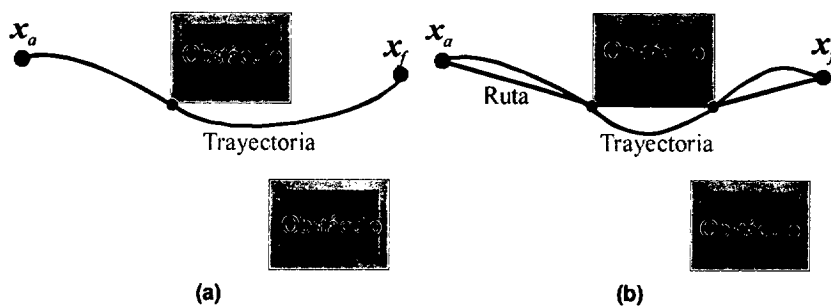
**Figura 7.6.** Generación de una ruta entre un punto de partida  $x_a$  y uno de llegada  $x_f$  utilizando subgrafos de visibilidad.

### 7.3.2. Planificación durante la generación de trayectorias.

Este método tiene bastantes similitudes con los subgrafos de visibilidad, desde el punto de vista que en una determinada fase del algoritmo es necesario determinar qué vértices se encuentran visibles desde un punto de paso.

Partiendo del conocimiento del entorno, de la definición de los obstáculos mediante polígonos y de los puntos de partida y de destino, se van trazando directamente caminos válidos mediante el uso de Splines cúbicas. En cuanto se detecta el cruce con cualquiera de los obstáculos del conjunto  $\mathcal{P}$ , se procede al estudio de los posibles puntos de paso (vértices) para evitarlo. El punto de paso elegido será un vértice del obstáculo, que además debe ser visible desde el punto de paso anterior a él y que de entre todos los vértices posibles del obstáculo sea el que posea la menor distancia hasta el punto final. Así, el punto que resulte elegido formará parte de los puntos a interpolar mediante la función Spline; estos pasos se repetirán hasta alcanzar el punto final.

La ventaja de este método, con respecto al de subgrafos de visibilidad, estriba en que aprovechando las curvaturas propias de la Spline se puede evitar la intersección con obstáculos, mientras que utilizando el método de los subgrafos de visibilidad posiblemente se tendrían que considerar, lo que permite acortar la longitud de camino y tiempo de cálculo. En la figura 7.6 se muestra una comparativa entre este método (figura 7.a) y el de subgrafos de visibilidad (figura 7.7.b).



**Figura 7.7.** Generación de trayectorias. (a) Método de planificación durante la generación de trayectorias. (b) Método de planificación con subgrafos de visibilidad.

En esta tesis se propone utilizar la alternativa de planificación durante la generación de la trayectoria para la definición del camino a seguir por el robot. Este sistema presenta, respecto a los subgrafos de visibilidad, una reducción en el número de puntos intermedios de paso, con la consiguiente disminución del tiempo de computo y además se consiguen trayectorias más suaves.

### 7.3.3. Interpolación mediante Splines.

La Spline es una curva utilizada para interpolación. Básicamente, se trata de un método para unir varios puntos  $(X, Y)$  entre sí mediante una curva suave que asegura la continuidad de ciertos parámetros en dichos puntos de unión. Aunque existen varios tipos de Splines, en la generación de la trayectoria se va a utilizar la Spline cúbica [Lázaro 98], que enlaza cada punto con el siguiente mediante un polinomio de tercer grado, asegurando la continuidad hasta la derivada de orden dos en los puntos de unión. Esto, aplicado al guiado de robots móviles, supone una forma sencilla de generar trayectorias que pasen por varios puntos y que asegure en todo momento la continuidad en dirección (primera derivada) y en la curvatura (segunda derivada). Otros autores [Muñoz 97][Romero 00] proponen la utilización de las curva  $\beta$ -Spline, que presenta la particularidad de que los puntos de la ruta correspondientes a los vértices de los obstáculos no son puntos de paso de la trayectoria. Una alternativa a la utilización de las Splines son las clotoides [Kanayama 86], pero estas curvas sólo garantizan continuidad hasta la primera derivada, con lo que las variaciones angulares son más bruscas y dan lugar a trayectorias más largas. Otro inconveniente que presentan las clotoides son los altos tiempos de computo necesarios para obtener la trayectoria.

El problema esencial que se presenta cuando se trabaja con funciones Splines cúbicas está relacionado con su condición de función matemática real, lo que implica que no puede

representar dos ordenadas diferentes para una misma abscisa. Para salvar esta limitación se puede trabajar con varios sistemas de referencia [Lázaro98], realizando cambios de ejes de referencia, o hacer uso de las curvas parametrizadas, solución tomada en esta tesis, por su sencillez y facilidad de cálculo.

Las curvas parametrizadas representan una curva  $Q(w)$  en un plano bidimensional  $(X, Y)$  en función de un parámetro de avance denominado 'w':

$$Q(w) = (f_X(w), f_Y(w)) \quad (7.7)$$

donde  $f_X(w)$  y  $f_Y(w)$  representa la variación de las abscisas y ordenadas, respectivamente, en función del parámetro "w". Para implementar las funciones  $f_X(w)$  y  $f_Y(w)$  se usan polinomios de tercer grado. Teniendo en cuenta que es posible definir en el camino tantos puntos intermedios como sean necesarios, la curva se divide en tramos (uno por cada dos puntos consecutivos) estando definido cada uno de ellos por un polinomio. En consecuencia, si se tienen  $n$  puntos de paso  $P_i$ , de coordenadas  $[X_i, Y_i]^T$  ( $i=0, 1, \dots, n-1$ ), la Spline tendrá un total de  $n-1$  tramos ( $Q_0, Q_1, \dots, Q_{n-2}$ ). Así, el tramo  $i$  estará definido por dos polinomios del tipo:

$$\begin{aligned} f_{X_i}(w) &= a_{iX} + b_{iX} \cdot w + c_{iX} \cdot w^2 + d_{iX} \cdot w^3 \quad \text{con } 0 \leq w \leq 1 \\ f_{Y_i}(w) &= a_{iY} + b_{iY} \cdot w + c_{iY} \cdot w^2 + d_{iY} \cdot w^3 \quad \text{con } 0 \leq w \leq 1 \end{aligned} \quad (7.8)$$

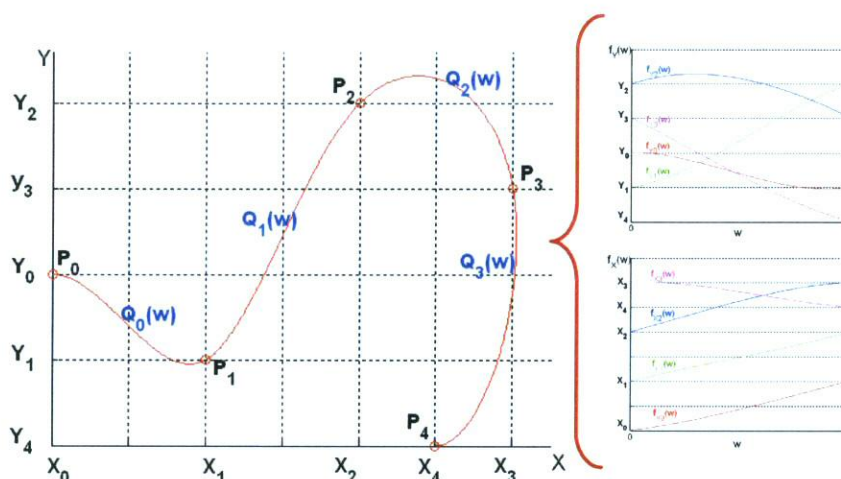
siendo  $(a_{iX}, b_{iX}, c_{iX}, d_{iX})$  y  $(a_{iY}, b_{iY}, c_{iY}, d_{iY})$  los coeficientes de la curva de abscisas y ordenadas, respectivamente, para cada tramo de la trayectoria. De forma gráfica, en la figura 7.8 se muestra una curva Spline así como la representación de las curvas  $f_X(w)$  y  $f_Y(w)$ .

Para realizar el cálculo de los coeficientes de la ecuación (7.8) en cada tramo, se imponen las condiciones de continuidad en la primera y segunda derivada en todos los puntos intermedios de la Spline; asegurando además una orientación (derivada primera) de la curva para sus puntos extremos (inicio y destino).

A continuación se expone el desarrollo del algoritmo necesario para calcular los cuatro coeficientes de la curva  $f_{Y_i}(w)$  para el tramo  $i$ , siendo idéntico para la  $f_{X_i}(w)$ .

El polinomio de la curva  $f_{Y_i}(w)$  para el tramo  $i$  tiene la expresión:

$$f_{Y_i}(w) = a_{iY} + b_{iY} \cdot w + c_{iY} \cdot w^2 + d_{iY} \cdot w^3 \quad \text{Con } 0 \leq w \leq 1 \quad (7.9)$$



**Figura 7.8.** Representación de una Spline cúbica parametrizada con las funciones  $f_X(w)$  y  $f_Y(w)$  para una trayectoria con cinco puntos de paso.

Sobre la ecuación (7.9) se aplican las condiciones que debe cumplir la curva de interpolación:

1º. La curva debe pasar por los puntos intermedios:

$$\begin{aligned} f_{Y_i}(0) &= Y_i = a_{iY} \\ f_{Y_i}(1) &= Y_{i+1} = a_{iY} + b_{iY} + c_{iY} + d_{iY} \end{aligned} \tag{7.10}$$

2º. Debe existir continuidad en la derivada primera.

$$\begin{aligned} \left. \frac{df_{Y_i}}{dw} \right|_{w=0} &= D_i = b_{iY} \\ \left. \frac{df_{Y_i}}{dw} \right|_{w=1} &= D_{i+1} = d_{iY} + 2c_{iY} + 3d_{iY} \end{aligned} \tag{7.11}$$

3º. Debe existir continuidad en la derivada segunda.

$$\left. \frac{d^2 f_{Y_{(i-1)}}}{d^2 w} \right|_{w=1} = \left. \frac{d^2 f_{Y_i}}{d^2 w} \right|_{w=0} = 2c_{(i-1)Y} + 6d_{(i-1)Y} = 2c_{iY} \tag{7.12}$$

A partir de (7.10), (7.11) y (7.12) es posible obtener los cuatro coeficientes de  $f_{Y_i}(w)$ :

$$\begin{aligned}
 a_{iY} &= Y_i \\
 b_{iY} &= D_i \\
 c_{iY} &= 3(Y_{i+1} - Y_i) - 2D_i - D_{i+1} \\
 d_{iY} &= 2(Y_i - Y_{i+1}) + D_i + D_{i+1}
 \end{aligned} \tag{7.13}$$

Sustituyendo (7.13) en (7.12) se obtiene:

$$D_{i-1} + 4D_i + D_{i+1} = 3.(Y_{i+1} - Y_{i-1}) \tag{7.14}$$

Sobre los puntos extremos de la trayectoria se impone la restricción de que debe existir continuidad en la orientación. Estas orientaciones inicial y final deben ser traducidas a pendientes (o derivadas) inicial y final de los polinomios  $X_0(w)$ ,  $Y_0(w)$  y  $X_{n-1}(w)$ ,  $Y_{n-1}(w)$ . Así, en el primer punto la pendiente de la curva  $Q_0(w)$  deberá ser la tangente de la orientación inicial  $\theta_0$ :

$$\tan(\theta_0) = \left. \frac{dY}{dX} \right|_{(X_0, Y_0)} = \left. \frac{df_{Y_0}/dw}{df_{X_0}/dw} \right|_{(w=0)} = \frac{\mu \cdot \sin(\theta_0)}{\mu \cdot \cos(\theta_0)} \tag{7.15}$$

Por tanto, al sistema de ecuaciones para determinar los coeficientes del polinomio  $f_{Y_0}(w)$  de la Spline, hay que añadir el valor de la primera derivada en el punto  $[X_0, Y_0]^T$ :

$$D_0 = \mu \cdot \sin(\theta_0) \tag{7.16}$$

De la misma forma, para el último punto  $[X_{n-1}, Y_{n-1}]^T$  la pendiente de la curva  $Q_{n-2}(w)$  debe ser la tangente de la orientación final  $\theta_{n-1}$ :

$$\tan(\theta_{n-1}) = \left. \frac{dY}{dX} \right|_{(X_{n-1}, Y_{n-1})} = \left. \frac{df_{Y(n-2)}/dw}{df_{X(n-2)}/dw} \right|_{(w=1)} = \frac{\mu \cdot \sin(\theta_{n-1})}{\mu \cdot \cos(\theta_{n-1})} \tag{7.17}$$

Al sistema de ecuaciones para determinar los coeficientes del polinomio  $f_{Y(n-2)}(w)$  de la Spline hay que añadir el valor de la primera derivada en el último punto:

$$D_{n-1} = \mu \cdot \sin(\theta_{n-1}) \tag{7.18}$$

Las ecuaciones de (7.14), (7.16) y (7.18) se pueden poner de forma matricial como:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \\ \vdots \\ D_{n-2} \\ D_{n-1} \end{bmatrix} = \begin{bmatrix} \mu \cdot \sin(\theta_0) \\ 3 \cdot (Y_2 - Y_0) \\ 3 \cdot (Y_3 - Y_1) \\ 3 \cdot (Y_4 - Y_2) \\ \vdots \\ 3 \cdot (Y_{n-1} - Y_{n-3}) \\ \mu \cdot \sin(\theta_{n-1}) \end{bmatrix} \tag{7.19}$$

Triangulando (7.19) se obtiene:

$$\begin{bmatrix} 1 & \gamma_0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \gamma_1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & \gamma_2 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_3 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & \gamma_{n-2} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \gamma_{n-1} \end{bmatrix} \cdot \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ D_3 \\ \vdots \\ D_{n-2} \\ D_{n-1} \end{bmatrix} = \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_{n-2} \\ \lambda_{n-1} \end{bmatrix} \tag{7.20}$$

donde:

$$\gamma_i = \begin{cases} 0 & \text{para } i = 0; \\ \frac{1}{4 - \gamma_{i-1}} & \text{para } 1 \leq i \leq n-1; \end{cases} \tag{7.21}$$

$$\lambda_i = \begin{cases} \mu \cdot \sin(\theta_0) & \text{para } i = 0; \\ \gamma_i \cdot (3 \cdot (Y_{i+1} - Y_{i-1}) - \lambda_{i-1}) & \text{para } 1 \leq i \leq n-2; \\ \mu \cdot \sin(\theta_{n-1}) & \text{para } i = n-1; \end{cases}$$

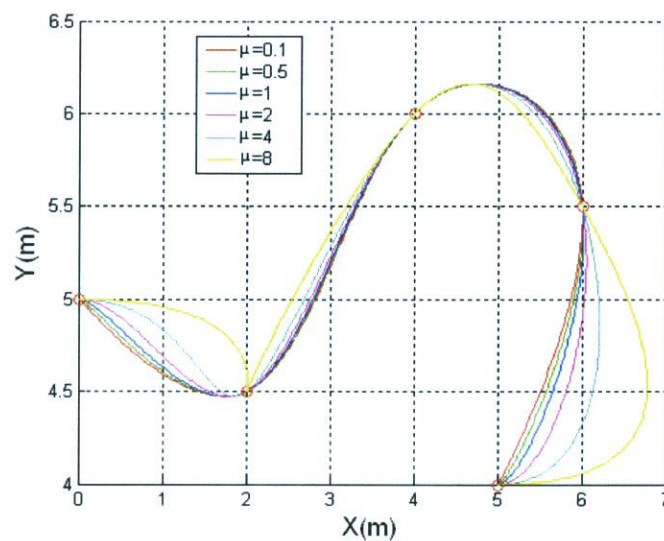
La solución de (7.20) viene dada por:

$$D_i = \begin{cases} \lambda_i & \text{para } i = n-1 \\ \lambda_i - \gamma_i \cdot D_{i+1} & \text{para } 0 \leq i \leq n-2 \end{cases} \tag{7.22}$$



Sustituyendo el valor de  $D_i$  obtenido en (7.22) en (7.13) se obtienen los coeficientes  $a_{iY}$ ,  $b_{iY}$ ,  $c_{iY}$  y  $d_{iY}$  de cada tramo. Para obtener los coeficientes  $a_{iX}$ ,  $b_{iX}$ ,  $c_{iX}$  y  $d_{iX}$  del polinomio  $f_X(w)$  se procede de forma similar.

El parámetro  $\mu$  es un factor de suavidad en la variación de la orientación, respecto a las orientaciones iniciales y finales. Cuanto mayor sea el valor de dicho parámetro el cambio de orientación se hará de una forma más suave y prolongada (figura 7.9), si bien valores altos dan lugar a trayectorias más largas.

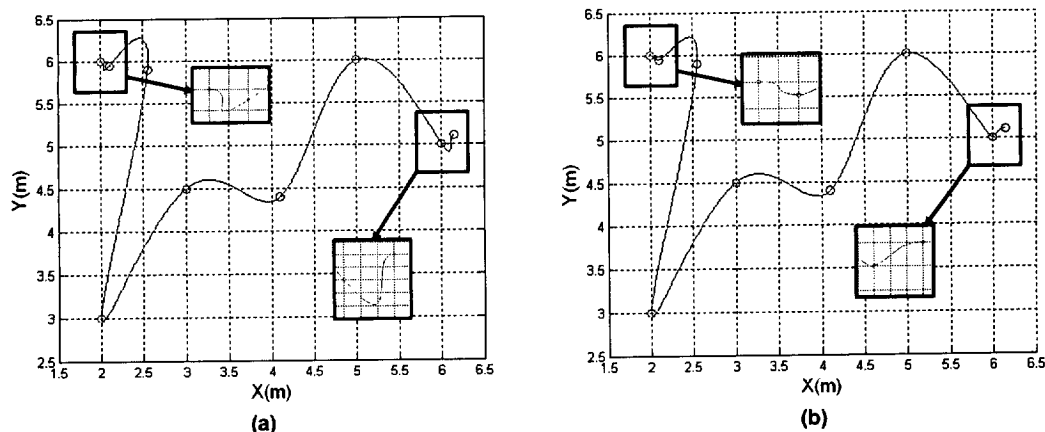


**Figura 7.9.** Generación de una Spline para diferentes valores del factor de suavidad,  $\mu$ .

La utilización de las curvas parametrizadas de las Splines presentan un problema cuando las separaciones entre dos puntos de paso es pequeña (inferior a 20 cm), tal y como puede apreciarse en la figura 7.10.a, dando lugar a tramos sinuosos y abruptos, siendo más acuciante cuanto mayor sea el valor de parámetro  $\mu$ . Este error se produce por la condición de continuidad en la orientación impuesta para los diferentes tramos, siendo independientemente de la posición que ocupen los puntos en la curva (primero, segundo,..., último).

Para dar solución al problema de los puntos próximos de la Spline se propone descomponerla en varias secciones. Así, dado un conjunto de  $n$  puntos ( $P_0, P_1, \dots, P_{n-1}$ ), de los cuales los puntos  $P_k$  y  $P_{k+1}$  se encuentran próximos, se van a generar tres Spline: una utilizando los puntos  $P_l, P_2, \dots, P_k$ , otra con los puntos  $P_k$  y  $P_{k+1}$  y una tercera con los puntos  $P_{k+1}, P_{k+2}, \dots, P_n$ . En el caso

de existir más de dos puntos próximos se procede de forma similar. En la figura 7.10.b se muestra el resultado de aplicar este método a la curva de la figura 7.10.a.

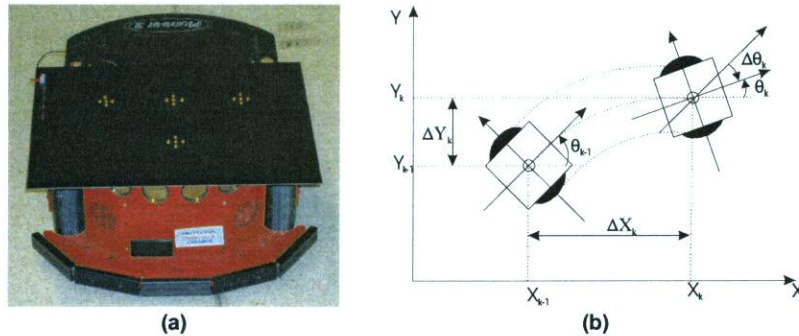


**Figura 7.10.** Trazado de un Spline con puntos de paso próximo. (a) Construyéndola sin dividirla por tramos. (b) Construyéndola dividiéndola en tramos.

## 7.4. Sistema de posicionamiento.

Para poder generar las consignas de control que permitan al robot seguir una trayectoria determinada, es necesario conocer en cada instante de tiempo cuál es su posición. Para obtener dicha posición cabe pensar en utilizar sólo la información proporcionada con el sistema de visión. Ahora bien, debido a la gran cantidad de información que se debe procesar, el tiempo invertido ( $T_m$ ) por el sistema de visión en determinar la posición del robot, hace necesario utilizar también la información proporcionada por el sistema odométrico. Evidentemente el tiempo  $T_m$  se puede reducir dotando a cada cámara de un sistema de procesamiento dedicado. En cualquier caso, el tiempo  $T_m$  es lo suficientemente pequeño como para que a las velocidades de desplazamiento normales de los robots, en entornos interiores, el error cometido por los sistemas odométricos sea despreciable. Por ejemplo, para el caso de un sistema de visión formado por cuatro cámaras, y un único procesador de imágenes,  $T_m$  es 0.3 sg. Considerando que el robot se mueve con una velocidad máxima de 1m/sg, durante el tiempo  $T_m$  éste habrá recorrido una distancia de 30 cm. Por tanto la posición real del robot será igual a la proporcionada por el sistema de visión, más el desplazamiento medido por los encoders durante el tiempo  $T_m$ . Para las pruebas prácticas se ha utilizado un robot Peoplebot de Active Media (figura 7.11.a), el cual tiene una estructura tipo triciclo, con dos ruedas

delanteras motrices y una trasera libre (loca). Esta estructura favorece la maniobrabilidad, además de ser sencillo de controlar.



**Figura 7.11.** Robot Peoplebot de Active Media utilizado en las pruebas prácticas. (a) Vista del robot. (b) Definición de los parámetros proporcionados por la odometría.

El sistema odométrico del robot proporciona la variación  $(\Delta x_k, \Delta y_k, \Delta \theta_k)$  de las coordenadas del punto central del eje que une las ruedas motrices. Así, la salida  $\mathbf{z}_k$  del sistema de medida (figura 7.12) será:

$$\mathbf{z}_k = \begin{bmatrix} X_k \\ Y_k \\ Z_k \\ \boldsymbol{\theta}_k \end{bmatrix} = \begin{bmatrix} X_{Ck} + \Delta X_k \\ Y_{Ck} + \Delta Y_k \\ Z_{Ck} + 0 \\ \boldsymbol{\theta}_k + \Delta \boldsymbol{\theta}_k \end{bmatrix} \quad (7.23)$$

siendo  $[X_{Ck}, Y_{Ck}, Z_{Ck}]^T$  las coordenadas de la baliza C asignadas a la posición del robot, obtenidas con el sistema de visión, el cual también proporciona la orientación de este último  $\boldsymbol{\theta}_k = [\theta_{Xk}, \theta_{Yk}, \theta_{Zk}]^T$ .

Teniendo presente que las medidas proporcionadas, tanto por el sistema de visión como por la odometría, están contaminada con ruido, para minimizar su efecto se propone en esta tesis la utilización de un Filtro de Kalman, el cual, también se utiliza para estimar en un instante  $t$  la posición del robot en el instante  $t+T_m$  y con ello la posición de la ventana de interés, sobre la que se centran los algoritmos de visión.

En la figura 7.12 se muestra la estructura del sistema de posicionamiento, donde  $\mathbf{u}_k$  representa las consignas  $[X_{fk}, Y_{fk}, Z_{fk}]^T$  y  $\hat{\mathbf{x}}_k$  la estimación de la posición del robot  $[\hat{X}_k, \hat{Y}_k, \hat{Z}_k]^T$ .

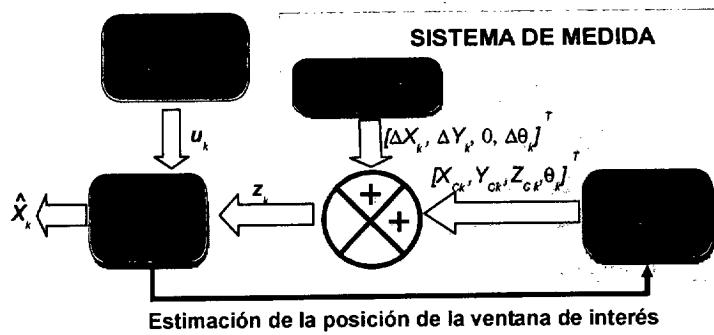


Figura 7.12. Estructura del sistema de posicionado.

### 7.4.1. Filtro de Kalman.

El Filtro de Kalman [Welch 01] es un potente algoritmo de estimación, basado en observaciones ruidosas y en un modelo de la incertidumbre de la dinámica del sistema (figura 7.13), que permite, entre otros resolver el problema general de estimación del estado de un sistema expresado en tiempo discreto. A continuación, se muestra las ecuaciones generales, para después particularizarlas a la aplicación de la tesis.

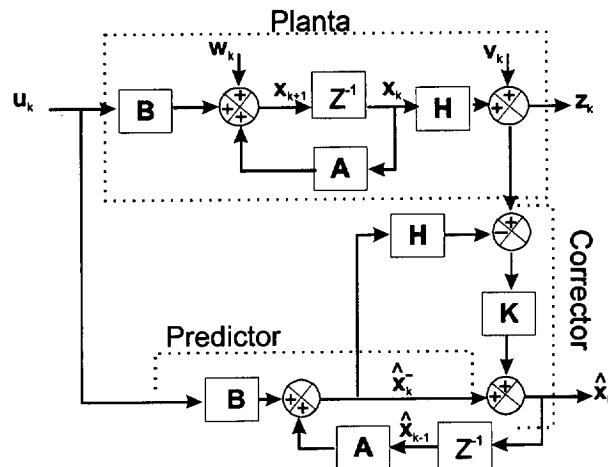


Figura 7.13. Estructura del Filtro de Kalman.

El Filtro de Kalman se basa en el modelo de la planta, incluyendo la evaluación del vector de estado, y el proceso de medida. La ecuación de estado (7.24), describe cómo evoluciona la posición o configuración del robot ( $x$ ) en respuesta a la señal de control ( $u$ ), junto con los errores del sistema ( $w$ ).

$$x_k = A \cdot x_{k-1} + B \cdot u_k + w_{k-1} \tag{7.24}$$

El proceso de medida, ecuación (7.25), expresa la relación entre el vector de estado y la medida obtenida con el sistema de visión y la odometría, junto con sus errores ( $v_k$ ).

$$z_k = H \cdot x_k + v_k \quad (7.25)$$

Los errores del sistema y de medida se asume que tienen una distribución gaussiana con media nula y matriz de covarianza  $Q$  y  $R$ , respectivamente. Por tanto, su distribución de probabilidad vendrá dada por la distribución normal:

$$\begin{aligned} p(w) &= N(0, Q) \\ p(v) &= N(0, R) \end{aligned} \quad (7.26)$$

En la práctica las matrices de covarianza  $Q$  y  $R$  podrían cambiar con el tiempo, pero en general se asumen como constantes.

El proceso de cálculo característico del Filtro de Kalman se puede dividir en dos etapas: predicción y corrección. La primera predice el valor del vector de estado en un instante  $k$ , tomando como referencia el estado en el instante  $k-1$ . La segunda incorpora nueva información a la predicción anterior y obtienen una corrección del estado.

El Filtro de Kalman predice el próximo estado  $\hat{x}_k^-$  del sistema y la matriz de covarianza esperada  $P_k^-$ , basándose en la información disponible del modelo:

$$\begin{aligned} \hat{x}_k^- &= A \cdot \hat{x}_{k-1} + B \cdot u_k \\ P_k^- &= A_k \cdot P_{k-1} \cdot A_k^T + Q_{k-1} \end{aligned} \quad (7.27)$$

La ganancia del Filtro de Kalman viene dada por:

$$K_k = P_k^- \cdot H_k^T \cdot (H_k \cdot P_k^- \cdot H_k^T + R_k)^{-1} \quad (7.28)$$

Una vez que las medidas  $z_k$  están disponible se pueden incorporar para corregir el estado estimado  $\hat{x}_k$ :

$$\hat{x}_k = \hat{x}_k^- + K_k \cdot (z_k - H \cdot \hat{x}_k^-) \quad (7.29)$$

El nuevo valor de la matriz de covarianza viene dado por:

$$P_k = (I - K_k \cdot H_k) P_k^- \tag{7.30}$$

En la figura 7.14 se muestra el diagrama de flujo que permite implementar el Filtro de Kalman.

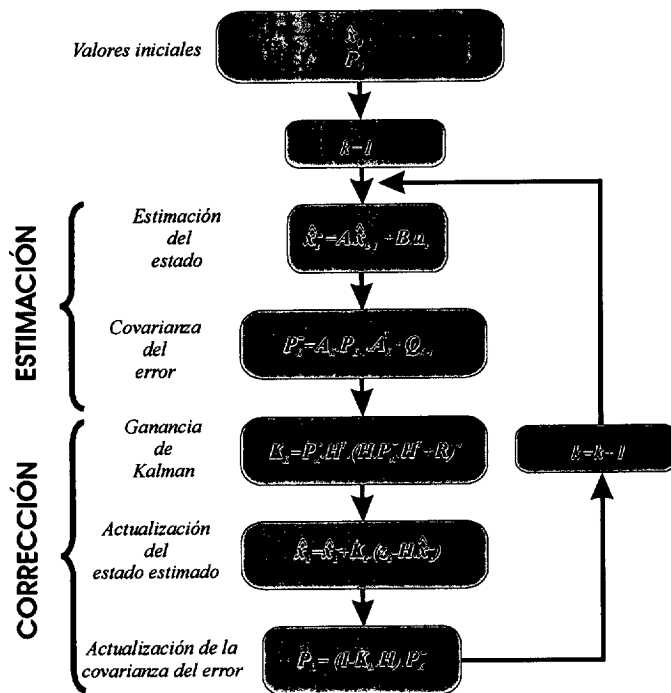


Figura 7.14. Organigrama para el cálculo del filtro de Kalman

En el guado del robot se realiza un control de posición, esto es, si en un instante de tiempo  $k-1$  se encuentra en la posición  $x_{k-1}$ , se generan las consignas necesarias para que en el instante  $k$  se encuentre en  $x_k$ , siendo esta última un punto de la trayectoria. Así, la ecuación del estado viene dada por:

$$\begin{bmatrix} X_k \\ Y_k \\ Z_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} X_{k-1} \\ Y_{k-1} \\ Z_{k-1} \\ \theta_{k-1} \end{bmatrix} + \begin{bmatrix} X_f \\ Y_f \\ Z_f \\ \theta_f \end{bmatrix} \tag{7.31}$$

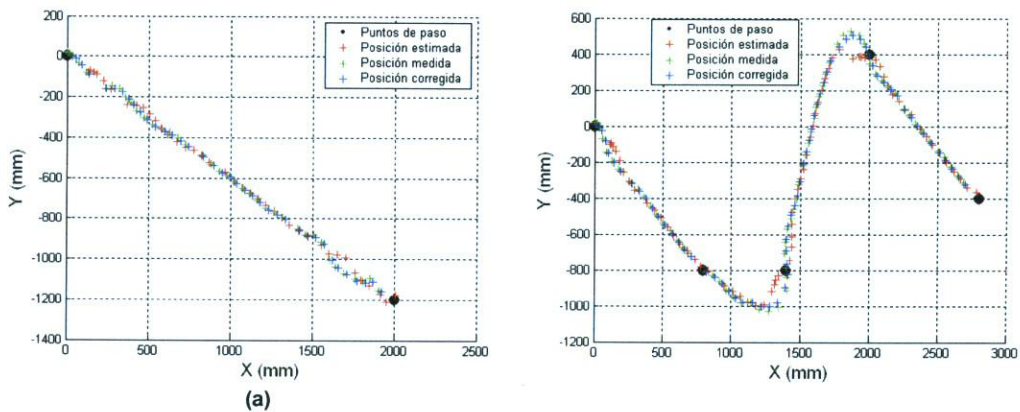
Identificando las ecuaciones (7.24) y (7.31) se obtiene que las matrices  $A$  y  $B$  son sendas matrices identidad.

Por su parte la ecuación de medida tiene la forma:

$$\begin{bmatrix} X_k \\ Y_k \\ Z_k \\ \theta_k \end{bmatrix} = \mathbf{H} \cdot \begin{bmatrix} X_k \\ Y_k \\ Z_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} X_k \\ Y_k \\ Z_k \\ \theta_k \end{bmatrix} \quad (7.32)$$

con lo que la matriz  $\mathbf{H}$  (ecuación 7.25) también es igual a la matriz identidad.

En la figura 7.15 se muestra los resultados de utilizar el Filtro de Kalman para corregir la posición del robot para dos trayectorias diferentes.



**Figura 7.15.** Corrección de la posición del robot utilizando un Filtro de Kalman. (a) Trayectoria recta. (b) Trayectoria sinuosa.

## 7.5. Generación de consignas.

Para realizar el guiado del robot se utiliza el sistema de control mostrado en la figura 7.1. Inicialmente se determinan los puntos de pasos para generar la trayectoria del robot. A continuación y durante cada ciclo de control, una vez obtenida y corregida la posición del robot, se realiza un nuevo cálculo de la trayectoria (figura 3.16).

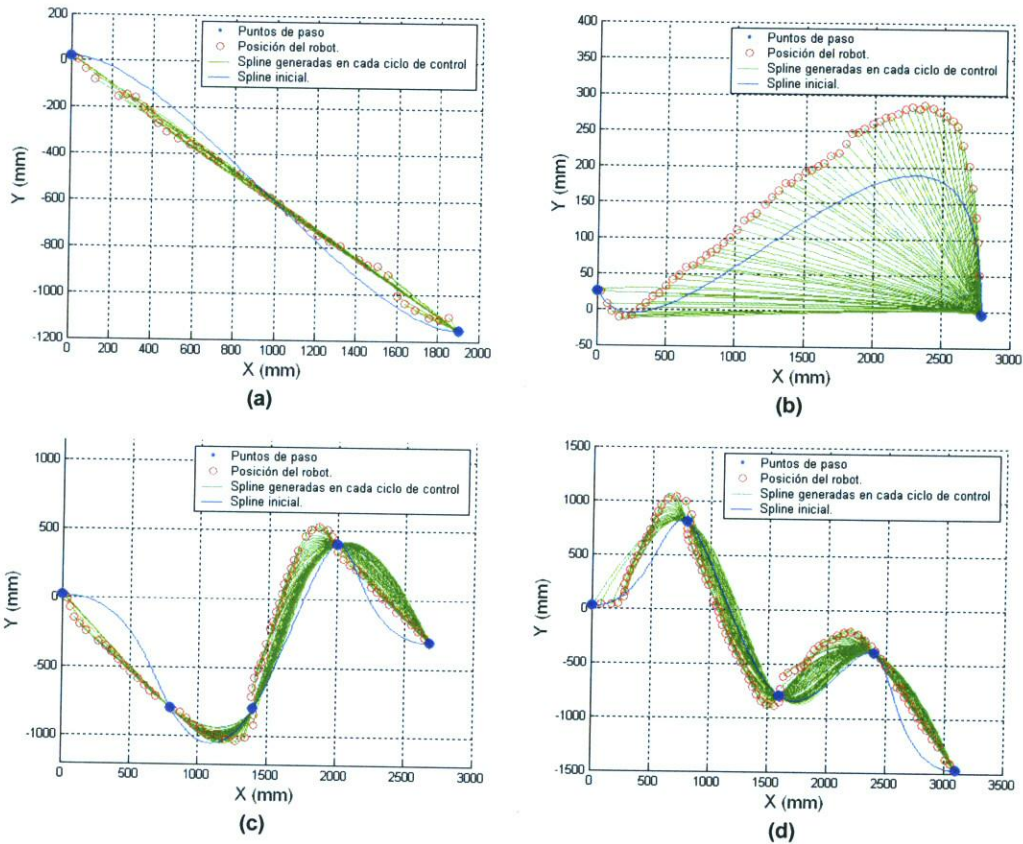


Figura 7.16. Trazado de trayectorias para cada posición del robot durante un proceso de navegación.

Considerando que el robot se desplaza en el plano  $XY$  del sistema de referencia global, su posición viene dada por el vector de estado  $\mathbf{x}=[X,Y,\theta]^T$ , siendo  $X$  e  $Y$  las coordenadas del centro del eje que une las ruedas motrices y  $\theta$  su orientación (véase figura 7.3). La relación existente entre las velocidades lineales de cada una de las ruedas motrices ( $V_R, V_L$ ) con la velocidad lineal ( $V$ ) y angular ( $\Omega$ ) del robot viene dada por:

$$\begin{aligned} \dot{X} &= \frac{(V_R + V_L)}{2} \cdot \cos(\theta) = V \cdot \cos(\theta) \\ \dot{Y} &= \frac{(V_R + V_L)}{2} \cdot \text{sen}(\theta) = V \cdot \text{sen}(\theta) \\ \dot{\theta} &= \frac{(V_R - V_L)}{B} = \Omega \end{aligned} \tag{7.33}$$

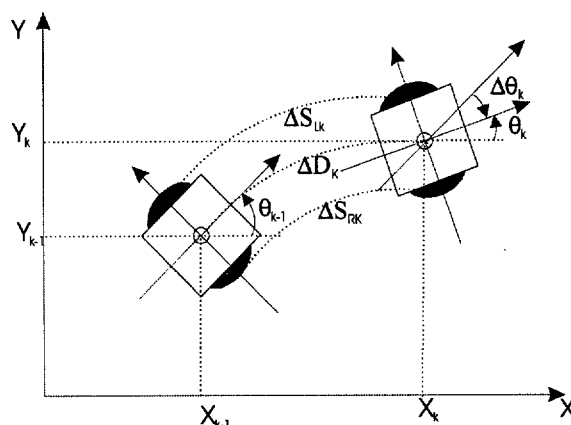
donde  $B$  representa la separación entre las ruedas motrices.



Con el robot en movimiento, la variación de su posición (figura 7.17) durante cada periodo de muestreo ( $T_m$ ) se compone de un desplazamiento ( $\Delta D$ ) y una rotación ( $\Delta\theta$ ) dados por:

$$\begin{aligned}\Delta D &= \frac{(\Delta S_R + \Delta S_L)}{2} \\ \Delta\theta &= \frac{(\Delta S_R - \Delta S_L)}{B}\end{aligned}\quad (7.34)$$

donde  $\Delta S_R$  y  $\Delta S_L$  representa los desplazamientos lineales sufridos por las ruedas derecha e izquierda, respectivamente.



**Figura 7.17.** Representación del cambio de posición del robot durante un periodo de muestreo.

Integrando (7.33) por el método de Euler [Paz 97], se puede escribir:

$$\begin{aligned}X_k &= X_{k-1} + \Delta D_k \cdot \cos(\theta_k) \\ Y_k &= Y_{k-1} + \Delta D_k \cdot \text{sen}(\theta_k) \\ \theta_k &= \theta_{k-1} + \Delta\theta_k\end{aligned}\quad (7.35)$$

La ecuación (7.35) equivale a aproximar el movimiento del robot por segmentos rectilíneos. Una mejor aproximación [Wang 88] se consigue considerando que la trayectoria está formada por arcos de circunferencia y tomando como argumento de las funciones seno y coseno el ángulo medio en el periodo de muestreo:

$$\overline{\theta}_k = \theta_k + \frac{\Delta\theta_k}{2}\quad (7.36)$$

así, la ecuación (7.35) se puede escribir como:

$$\begin{aligned} X_k &= X_{k-1} + \Delta D_k \cdot \cos\left(\theta_k + \frac{\Delta\theta_k}{2}\right) \\ Y_k &= Y_{k-1} + \Delta D_k \cdot \text{sen}\left(\theta_k + \frac{\Delta\theta_k}{2}\right) \\ \theta_k &= \theta_{k-1} + \Delta\theta_k \end{aligned} \quad (7.37)$$

Las aproximaciones (7.37) requieren que la curvatura de la trayectoria, definida por  $\Delta\theta_k/\Delta D_k$ , sea constante durante el periodo de muestro y pequeña respecto a las aceleraciones lineal y angular.

Teniendo en cuenta que el robot se desplaza a una velocidad media  $V$ , durante el tiempo de muestreo  $T_m$  recorrerá una distancia  $e$ :

$$e = V \cdot T_m = \sqrt{(X_k - X_{k-1})^2 + (Y_k - Y_{k-1})^2} \quad (7.38)$$

siendo  $[X_k, Y_k]^T$  y  $[X_{k-1}, Y_{k-1}]^T$  las coordenadas de las posiciones del robot en un instante  $k$  y en  $k-1$  (figura 7.18), respectivamente. El valor de las coordenadas  $[X_k, Y_k]^T$  se obtiene de la ecuación de la Spline (ecuación (7.8)) imponiendo la condición (7.38). A partir de estas coordenadas y con (7.37) se obtienen las consignas de control del robot, las cuales van a ser las velocidades lineales de las ruedas motrices ( $V_L, V_R$ ) que vienen dadas por (7.39).

$$\begin{aligned} V_L &= \frac{1}{T_m} \left( \frac{X_k - X_{k-1}}{\cos\varphi} - \frac{B \cdot (\theta_k - \theta_{k-1})}{2} \right) \\ V_R &= \frac{1}{T_m} \left( \frac{X_k - X_{k-1}}{\cos\varphi} + \frac{B \cdot (\theta_k - \theta_{k-1})}{2} \right) \end{aligned} \quad (7.39)$$

con

$$\begin{aligned} \varphi &= \frac{\theta_k + \theta_{k-1}}{2} \\ \theta_k &= \arctan\left(\frac{\Delta Y_k}{\Delta X_k}\right) \end{aligned} \quad (7.40)$$

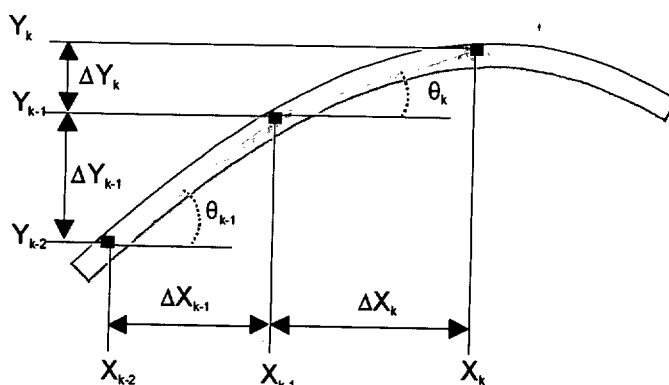


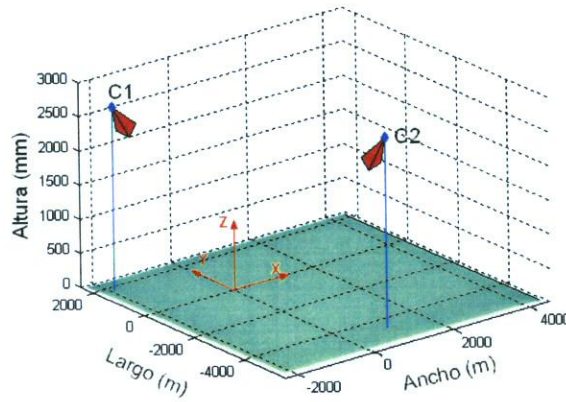
Figura 7.18. Linealización de la trayectoria que permite obtener las consignas de guiado del robot móvil.

## 7.6. Resultados prácticos.

A continuación se muestran los resultados prácticos obtenidos con el sistema de guiado propuesto, para diferentes distribuciones de las cámaras. En todas las pruebas, el robot parte de una posición inicial y se especifica la posición final  $([X_f, Y_f, \theta_f]^T)$  y las coordenadas de los puntos intermedios por los cuales tiene que pasar  $([X_p, Y_p]^T)$ . A partir de estos puntos (inicial, final e intermedios) se obtiene una primera trayectoria, la cual se va recalculando a medida que el robot se va moviendo.

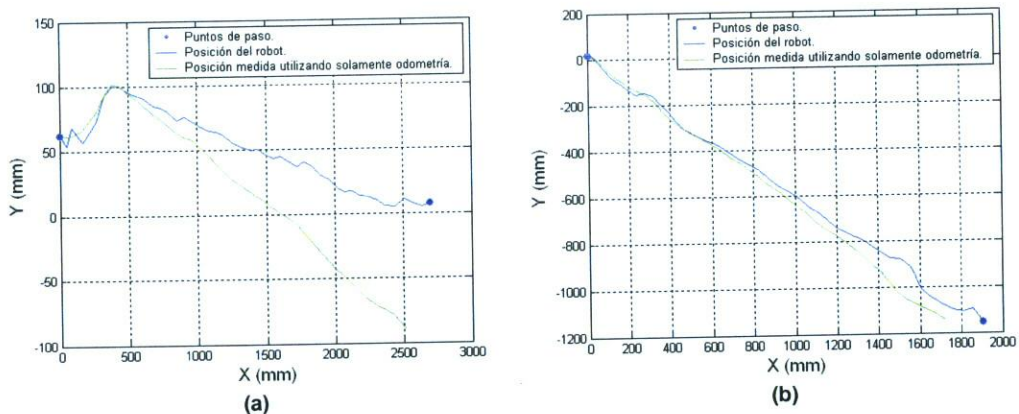
Un primer grupo de pruebas se han realizado en el entorno de la figura 7.19, sobre el que se han dispuesto dos cámaras. Los parámetros extrínsecos de las cámaras se muestran en la tabla 7.1.

TABLA 7.1. Parámetros extrínsecos de las cámaras de la figura 7.17		
	C1	C2
$T_x$ (mm)	-1895	1099
$T_y$ (mm)	1803	-4404
$T_z$ (mm)	2692	2.768
$\alpha$ (°)	-37.6	-26.7
$\beta$ (°)	13.7	5.3
$\gamma$ (°)	-58	98.2



**Figura 7.19.** Distribución de las cámaras en entorno de trabajo empleado en las pruebas práctica.

Un primer grupo de experimentos prácticos consistió en realizar trayectorias rectas, indicando la posición final a alcanzar  $([X_f \ Y_f \ \theta_f]^T)$ . En el primero se tomó  $X_f=2700$  mm,  $Y_f=0$  y  $\theta_f=0^\circ$ , siendo los resultados los mostrados en la figura 7.20.a. En color azul se muestra la posición del robot obtenida con el sistema de visión propuesto en esta tesis, mientras que en color verde se refleja la posición del robot utilizando solamente odometría. La posición alcanzada por el robot fue  $X_{fR}=2701$  mm,  $Y_{fR}=6.9$  mm y  $\theta_{fR}=1.7^\circ$ . Para una segunda recta se tomó  $X_f=1900$  mm,  $Y_f=-1100$  mm y  $\theta_f=0^\circ$ , siendo los resultados los mostrados en la figura 7.20.b, donde la posición final del robot fue  $X_{fR}=1905$  mm,  $Y_{fR}=-1156$  mm y  $\theta_{fR}=3^\circ$ .



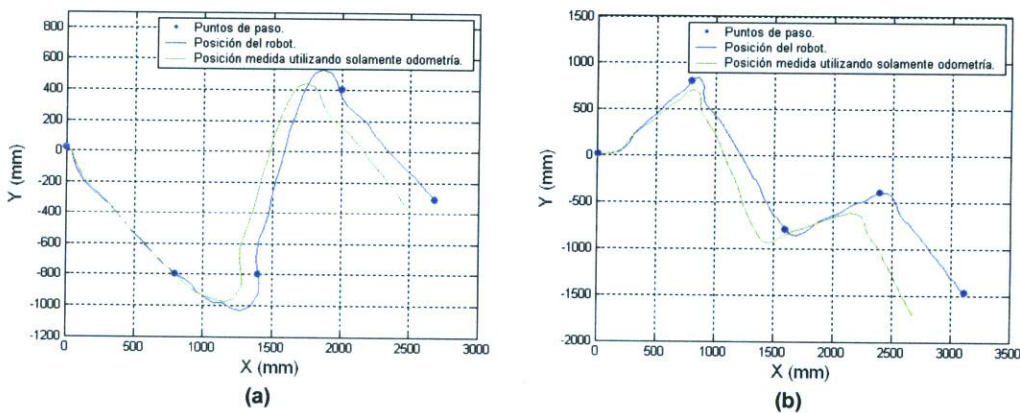
**Figura 7.20.** Ejemplos de seguimiento de trayectorias rectas. (a) Horizontal. (b) Oblicua.

La segunda prueba consistió en la realización de caminos más elaborados (figura 7.21) con diferentes cambios de orientación, donde las coordenadas de los puntos de paso para cada

trayectoria se muestran en la tabla 7.2. Para la figura 7.21.a se fijó una posición final  $X_f=2800$  mm,  $Y_f=-300$  mm y  $\theta_f=0^\circ$ , alcanzándose la posición  $X_{fR}=2787$ mm,  $Y_{fR}=-313$ mm y  $\theta_{fR}=-4^\circ$ . Por su parte, para la figura 7.21.b se fijó como posición destino  $X_f=3200$  mm,  $Y_f=-1600$  mm y  $\theta_f=0^\circ$ , alcanzándose la posición  $X_{fR}=3115$ mm,  $Y_{fR}=-1576$ mm y  $\theta_{fR}=-3^\circ$ .

**TABLA 7.2.** Coordenadas de los puntos de paso de las trayectorias de la figura 7.21.

Figura 7.21.a		Figura 7.21.b	
X (mm)	Y (mm)	X (mm)	Y (mm)
800	-800	800	800
1400	-800	1600	-800
2000	400	2400	-400

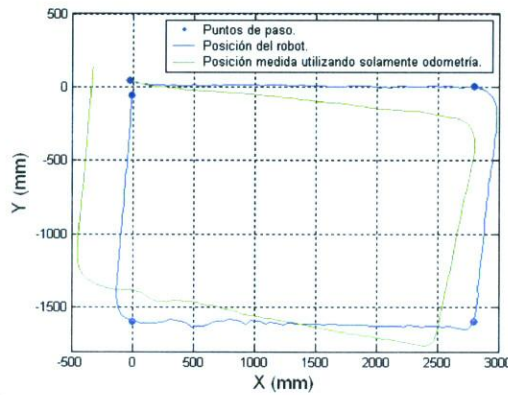


**Figura 7.21.** Ejemplos de seguimiento de trayectorias complejas con tres puntos intermedios de paso.

Se realizó un tercer experimento de seguimiento de una trayectoria cerrada en forma de cuadrado (figura 7.22), fijándose la posición final del robot en  $X_f=0$  mm,  $Y_f=0$  mm y  $\theta_f=90^\circ$ , mientras que las coordenadas de los puntos de paso fueron las mostradas en la tabla 7.3. En este caso la posición alcanzada por el robot fue:  $X_{fR}=-6$  mm,  $Y_{fR}=-59$ mm y  $\theta_{fR}=-84^\circ$ .

**TABLA 7.3.** Coordenadas de los puntos de paso de la trayectoria de la figura 7.22.

X (mm)	Y (mm)
2800	0
2800	-1600
0	-1600



**Figura 7.22.** Ejemplo de seguimiento de una trayectoria cerrada.

Se realizó también un cuarto experimento utilizando cuatro cámaras para la obtención de la posición del robot. El entorno utilizado en este experimento fue el de la figura 7.19 y los parámetros extrínsecos de las cámaras fueron los mostrados en la tabla 7.4. En este caso se generó una trayectoria (figura 7.23.b) con posición final  $X_f= 3000$  mm,  $Y_f= 100$  mm y  $\theta_f= 0^\circ$  y puntos de paso (1000, -50) y (2800, 650). La posición final alcanzada fue:  $X_{fR}= 2993$  mm,  $Y_{fR}= 106$  mm y  $\theta_{fR}=-2^\circ$ .

TABLA 7.4. Parámetros extrínsecos de las cámaras C3 y C4 de la figura 7.23.		
	C3	C4
$T_x$ (mm)	1113	4385
$T_y$ (mm)	2398	1612
$T_z$ (mm)	2567	2650
$\alpha$ (°)	-40.2	-35.7
$\beta$ (°)	8.7	-3.9
$\gamma$ (°)	67.8	160.4

Un último grupo de experimentos se realizó sobre un entorno de dimensiones mayores a los anteriores (figura 7.24.a), datado de cuatro cámaras: las cámaras C1 y C2 ocupan las mismas posiciones que en los entornos anteriores, mientras que C3 y C4 tienen los parámetro extrínseco mostrados en la tabla 7.5. En este caso se contemplaron todas las posibilidades para el posicionado: que sea visto por dos, una o ninguna de las cámaras (figura 7.24.c).

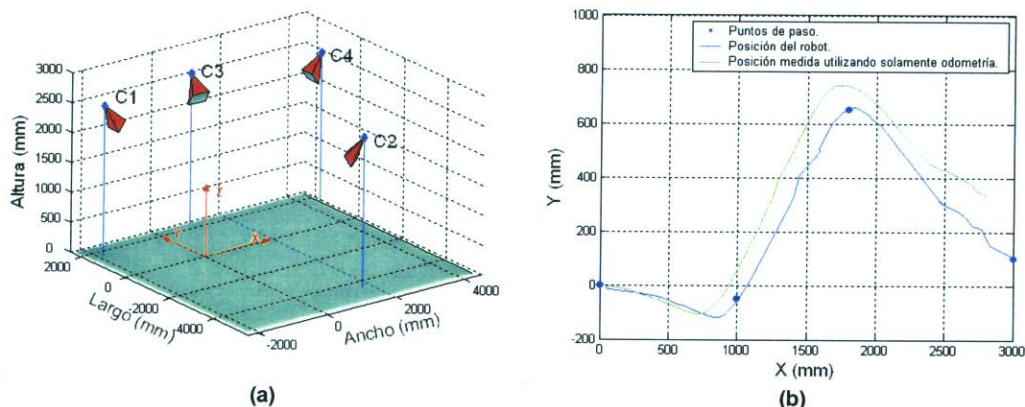


Figura 7.23. Ejemplo de una trayectoria controlada por cuatro cámaras. (a) Entorno de trabajo. (b) Trayectoria seguida por el robot.

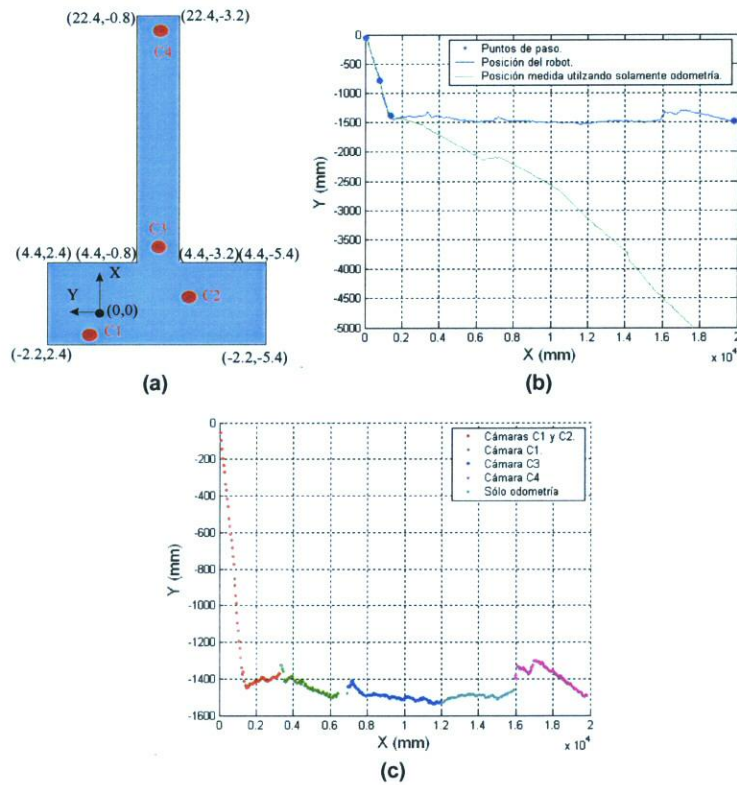
**TABLA 7.5.** Parámetros extrínsecos de las cámaras C3 y C4 de la figura 7.24.a.

	C3	C4
$T_x$ (mm)	4787	22135
$T_y$ (mm)	-1276	-1625
$T_z$ (mm)	2540	2600
$\alpha$ (°)	-60.2	-58.7
$\beta$ (°)	3.7	-2.6
$\gamma$ (°)	167.8	16.4

Para la trayectoria de la figura 7.24 se fijó una posición final  $X_f= 20$  m,  $Y_f= -1.5$  m y  $\theta_f= 0^\circ$ , siendo la posición alcanzada por el robot:  $X_{fR}= 19.932$  m,  $Y_{fR}= -1.507$  m y  $\theta_{fR}=-2^\circ$ . Por otra parte las coordenadas de los puntos de pasos son las mostradas en la tabla 7.6.

**TABLA 7.6.** Coordenadas de los puntos de paso de la trayectoria de la figura 7.24.b.

X (mm)	Y (mm)
800	-800
1400	-1400



**Figura 7.24.** Ejemplo de un experimento práctico realizado sobre un entorno de dimensiones grandes utilizando cuatro cámaras. (a) Dimensiones (en m) del entorno. (b) Trayectoria realizada. (c) Dispositivos (cámaras y odometría) utilizados para obtener la posición del robot dentro de los diferentes tramos.

En las pruebas anteriores el tiempo empleado en el lazo de control: obtención de la posición, cálculo de la trayectoria y generación de consignas (velocidades lineales de las ruedas) fue de 0.33 sg. Para las pruebas se ha utilizado un ordenador personal con un procesador Pentium II a 200 Mhz.

Finalmente indicar que en todas las pruebas realizadas los errores de posicionamiento entre las trayectorias reales y las ideales no superó: 5 cm para los tramos de las trayectorias vistos por una sola cámara, de 2 cm para los de dos cámaras, y 1.8 cm para los de cuatro.

## 7.7. Conclusiones.

En este capítulo se ha desarrollado el sistema de guiado del robot a partir de la información del sistema de visión, añadiendo una predicción y corrección por medio del Filtro de Kalman.



La utilización de las Splines cúbicas para la generación de la trayectoria del robot, constituye una alternativa atractiva para conseguir trayectorias continuas con variaciones suaves, pasando por los puntos intermedios prefijados.

La propuesta de utilizar el Filtro de Kalman permite realizar una corrección efectiva de la posición del robot, además de una estimación fiable de dicha posición, que es utilizada por la etapa de procesamiento de imágenes para determinar la ventana de interés.

La validez de la propuesta de guiado realizada en esta tesis, ha sido constatada en las pruebas prácticas de guiado realizadas, como lo ponen de manifiesto los resultados de las trayectorias obtenidas y que han sido mostradas dentro de este capítulo.

Como aportaciones de este capítulo se puede indicar, por una parte, la propuesta para obtener un camino libre de obstáculos que pueda seguir el robot para alcanzar un punto destino, utilizando, por ejemplo, subgrafos de visibilidad (que permiten además una reducción en los tiempos de ejecución). Otra aportación es la de realizar el desdoblamiento en el cálculo de las Splines para puntos próximos (en los cuales estas curvas producen trayectorias muy sinuosas), situación común cuando el robot se acerca a los puntos de paso y final.

Otra aportación es la utilización del Filtro de Kalman para realizar la corrección de la posición y la estimación de la “ventana de interés” sobre la que se aplican los algoritmos de procesamiento de imágenes para la segmentación de las balizas.

## Capítulo 8

---

### Conclusiones y trabajos futuros

---

En esta tesis se ha propuesto y desarrollado un nuevo sistema de guiado de robots móviles a partir de la información proporcionada por un array de cámaras ubicadas en el entorno de movimiento de los mismos. Este nuevo enfoque, en el que el sistema de percepción es externo a los robots, está en línea con los denominados “espacios inteligentes”.

En esta nueva concepción la “inteligencia” se traslada al entorno y es éste el encargado de percibir, procesar y generar las consignas de control a los robots.

Se trata en definitiva, de evitar tener que incorporar en los robots una compleja estructura sensorial y de procesadores, trasladando la parte más compleja de dicha estructura sensorial y de computación hacia el entorno, simplificándose así los robots. Incorporando los sensores de posicionamiento y procesado al entorno, la tarea de identificación de objetos, estáticos o dinámicos se simplifica, pues las restricciones geométricas de la estructura de sensores fijos ayudan a interpretar más fiablemente dicha información, minimizando, por una parte, la electrónica a bordo de los robots (con las ventajas que ello conlleva), y, por otro lado, facilitando que múltiples robots puedan moverse simultáneamente en el mismo entorno. La estrategia de localización propuesta en esta tesis es mediante cámaras de video ubicadas en posiciones fijas del entorno. La comunicación entre el sistema central de procesamiento y los

robots, para realizar el control y guiado, se realiza mediante enlaces vía radio.

Para alcanzar el objetivo final de guiado de los robots móviles a partir del array de cámaras externo, se han tenido que abordar diferentes tareas, que van desde la propuesta de las balizas a ubicar a bordo del robot y que facilitan su identificación, hasta la generación de trayectorias para el guiado, pasando por la calibración de múltiples cámaras, algoritmos para la obtención de las coordenadas 3D y orientación de los robots, procesamiento de imágenes para obtener de forma precisa los centroides de las balizas, estudio de las fuentes de errores y sus efectos en la obtención de la posición de los robots, y aplicación del Filtro de Kalman para realizar una estimación más precisa de la posición y acotar la ventana de interés de dimensiones reducidas que facilita los procesos de segmentación de las balizas. En todos estos aspectos se han realizado aportaciones de mayor o menor relevancia, las cuales se han ido presentando en los diferentes capítulos. A todo ello hay que añadir la puesta en práctica de la propuesta realizada, demostrando su viabilidad y prestaciones. Así, se ha creado un entorno de movimiento de robots (que incluye pasillos y habitaciones del edificio del Edificio Politécnico de la Universidad de Alcalá), donde se ha ubicado un array de cuatro cámaras y donde se han realizado diferentes tipos de pruebas, entre las que se incluyen diferentes situaciones de visibilidad del robot por las cámaras (robot visto por dos, una o ninguna cámara), diferentes condiciones de iluminación, etc.

En este último capítulo se realiza un breve resumen de las principales aportaciones realizadas en la tesis, así como las conclusiones finales derivadas de su desarrollo. Finalmente se plantean algunas de las líneas futuras de investigación que pueden ser de interés dentro de esta nueva propuesta de guiado de robots móviles (algunas de estas líneas de investigación ya son objeto de desarrollo por otros miembros del grupo de investigación)

## **8.1. Conclusiones y aportaciones.**

Como conclusión general de esta tesis se puede decir que se ha propuesto un novedoso sistema de guiado de robots móviles donde la parte más compleja de la estructura sensorial y de procesamiento se traslada al entorno. Esta nueva propuesta simplifica notablemente la configuración de los robots y presenta numerosas ventajas, frente a las soluciones que embarcan toda la “inteligencia” a bordo del robot. Esta propuesta se puede considerar como muy novedosa ya que hasta el 2002 apenas existían trabajos en esta línea, si bien en los últimos años están adquiriendo un mayor protagonismo. En cualquier caso, el trabajo

desarrollado en esta tesis se puede considerar muy novedoso.

Las aportaciones más destacables, así como las conclusiones que se derivan de las soluciones propuestas se pueden resumir en cinco grandes bloques (que se corresponden con los capítulos 3 a 7 de la tesis): errores debidos al sistema de visión, posicionamiento 3D utilizando cámaras externas, balizas de IR y procesamiento de imágenes para su identificación, calibración del sistema de visión, y guiado. A continuación se realiza un breve resumen para cada una de ellas.

### **8.1.1. Errores debidos al sistema de visión.**

Un tema importante dentro del sistema propuesto, una vez definidos los aspectos generales sobre el tipo de baliza y la ubicación de las cámaras, es el estudio de las diferentes fuentes de error y su incidencia en la obtención de la posición de los robots. Dado que la posición del robot se obtiene a partir de las balizas a bordo, se realizó un estudio detallado de los errores que influyen en la obtención de las coordenadas pixélicas del centroide de las balizas (y en consecuencia en la obtención de las coordenadas 3D del robot). Dentro de las fuentes de error, se concluye que las más importantes son las asociadas al tiempo de exposición, desplazamientos temporales debidos a la falta de sincronismo entre las cámaras, modelo de cámara, proceso de calibración, y obtención de las coordenadas pixélicas de las balizas en el plano imagen. Por ello es necesario buscar soluciones que minimicen su efecto, por ejemplo tratando de reducir el tiempo de exposición, hacer que todas las cámaras tengan una señal de sincronismo común y que el hardware de adquisición permita que ésta se realice de forma simultánea para todas las cámaras; utilizar técnicas para reducir el efecto tanto de la distorsión radial como tangencial, realizar una calibración precisa y obtener con precisión subpixélica las coordenadas del centroide de las balizas.

Dentro de las aportaciones realizadas en este tema, se pueden destacar:

- Propuesta de una estructura de baliza fácilmente identificable, lo que simplifica los procesos posteriores de procesamiento de imágenes y da robustez al sistema de detección, aún en condiciones extremas de cambios de iluminación.
- Estudio detallado sobre las diferentes fuentes de errores, evaluando su efecto (partiendo de un estudio matemático y realizando numerosas simulaciones).

- Propuesta de soluciones para reducir el efecto de las fuentes de error en la obtención de la posición del robot. En relación con esto, es de destacar la propuesta realizada para minimizar los efectos de la distorsión de cámara y la de obtención de las coordenadas pixélicas del centroide de la baliza.

### **8.1.2. Posicionamiento 3D utilizando cámaras externas.**

En relación con la obtención de la posición del robot a partir de la información captada por las cámaras de las balizas a bordo, en esta tesis se ha realizado una propuesta que permite obtener dicha posición a partir de una única cámara (suponiendo que el robot se mueva por un plano y la altura del plano en el que se encuentra la baliza es conocida), dos o más cámaras.

La conclusión final de este estudio es que la mejor alternativa para realizar el posicionamiento es que en todo instante el robot sea captado por dos cámaras, ya que esta alternativa permite obtener las coordenadas 3D, con independencia del tipo de superficie por la que se muevan, y los errores de posicionamiento son aceptables para muchos entornos prácticos. Se ha demostrado también que la mejora obtenida utilizando simultáneamente un número de cámaras superior a dos no justifica el coste añadido que, en general, ello conlleva. Las pruebas de simulación y prácticas llevadas a cabo dentro del estudio realizado en este capítulo avalan esta conclusión.

Como aportaciones más importante se pueden destacar:

- Propuesta de un algoritmo basado en la minimización de una función de coste, para la obtención de las coordenadas 3D a partir de la información obtenida por dos o más cámaras. Dicho algoritmo tiene presente los errores cometidos por las cámaras implicadas, dando mayor protagonismo a las que cometen menor error (se asigna mayor protagonismo a las cámaras con medidas de menor varianza).
- Evaluación de los errores cometidos en el posicionamiento en función de la disposición geométrica entre cámaras y robot, para las diferentes alternativas del número de cámaras que en cada instante se utilizan para la obtención de la posición 3D.

### **8.1.3. Balizas de IR y procesamiento de imágenes para su identificación.**

Otro de los aspectos más importantes en relación con los objetivos planteados es la capacidad de obtener los centroides de las balizas a bordo del robot, ya que a partir de esta información se obtienen sus coordenadas 3D y orientación. Además, interesa que esta detección no requiera una gran carga computacional y sea robusta ante diferentes cambios y condiciones de iluminación. Para dar respuesta a estos objetivos se ha seleccionado una estructura de baliza formada por cuatro balizas de infrarrojos con la longitud de onda adecuada (teniendo presente, entre otros, la respuesta espectral de las cámaras) dispuestas en forma de "T". Así mismo, se han propuesto diferentes estrategias y algoritmos de procesamiento de imágenes que facilitan la detección de las balizas y de los puntos de interés (centroides). Y todo ello garantizando resultados fiables y robustos en diferentes condiciones ambientales y con tiempos de proceso reducidos.

Como conclusiones generales, en primer lugar se puede decir que la utilización de balizas de infrarrojos de longitud de onda y potencia adecuadas (teniendo presentes, entre otras, las características de las cámaras y las condiciones del entorno) constituye una alternativa muy atractiva para la detección de los robots, ya que con ello no solamente se facilitan los algoritmos de procesamiento de imágenes sino que también se dota al sistema de una mayor robustez e insensibilidad a las condiciones cambiantes de iluminación del entorno. Por otra parte, una disposición adecuada de las balizas facilita la obtención de la orientación del robot.

Otra conclusión importante es que, a pesar de utilizar balizas de infrarrojos, es necesario prestar especial atención a la fase de procesamiento de imágenes ya que esta fase debe garantizar una detección fiable y robusta de las balizas y la obtención precisa de las coordenadas de sus centroides. De ello depende, en gran medida, la precisión con la que se puede obtener la posición del robot. En este sentido, las propuestas realizadas en esta tesis han sido probadas en condiciones reales y ponen de manifiesto su viabilidad incluso en condiciones extremas de iluminación, distancia entre cámaras y balizas, etc.

En resumen, se puede decir que la propuesta realizada en esta tesis, permite la detección, segmentación y localización de las balizas, de forma fiable, robusta, precisa y con tiempos de procesamiento reducidos.

Como aportaciones a destacar se pueden citar:

- Propuesta de las características (dimensión, longitud de onda, potencia, etc.) de las balizas de IR para facilitar su identificación, incluso en condiciones cambiantes y extremas de iluminación ambiental, utilizando cámaras CCD estándares.
- Propuesta de una configuración de cuatro balizas dispuestas en forma de “T”, que facilita la identificación de las mismas (aprovechando su disposición geométrica) y permiten obtener información sobre la orientación del robot.
- Estrategia de procesamiento de imágenes que facilita la identificación de las balizas y permite obtener de forma precisa las coordenadas de sus centroides (a partir de las cuales se obtiene la posición y orientación del robot). La propuesta de procesamiento de imágenes incluye: utilización de una “ventana de interés” dinámica entorno a la posición estimada del robot (que permite centrar el análisis de imágenes sólo dentro de ella, lo cual reduce el efecto del ruido y al mismo tiempo la carga computacional); utilización de operadores tipo gradiente para la detección de los bordes de las balizas; umbralización dinámica basada en el método de Otsu.
- Propuesta de validación de balizas basada en: eliminación de objetos de pequeñas dimensiones; identificación de elipses utilizando la distancia algebraica; eliminación de falsos candidatos a partir de las características de excentricidad y la relación entre perímetros de las “elipses estimadas” y los potenciales candidatos; relaciones geométricas propias de la ubicación física relativa entre las cuatro balizas.
- Estudio de las alternativas más importantes para la obtención de las coordenadas del centroide de las balizas y realización de una propuesta en la que las coordenadas del centroide se obtienen a partir de la intersección de las diagonales principales que unen los vértices del mínimo rectángulo que las contiene.

### 8.1.4. Calibración del sistema de visión

Un factor determinante para poder realizar las medidas de posición 3D con cámaras externas es la calibración del sistema de visión. En este proceso se debe obtener una calibración con suficiente calidad y precisión y que no requiera de: alta carga computacional, sofisticados mecanismos ni logística compleja para llevarla a cabo, a pesar de realizarse *off-line*. En el resultado del proceso de calibración van a influir los elementos utilizados, la estructura del proceso y la algoritmia. Para dar respuesta a estos requerimientos se han realizado propuestas tanto de algoritmia como de diseño y utilización de modelos y patrones que aportan un valor añadido a los procesos de calibración. Se ha utilizado un patrón de calibración activo (realizado con diodos IR) que simplifica el proceso de calibración de cada cámara y permite, además, obtener con una buena precisión los parámetros extrínsecos de posicionamiento de las cámaras en el sistema global. Desde el punto de vista de la estructura del proceso y la algoritmia, se ha analizado el proceso de calibración y se ha propuesto una estructura de proceso en dos fases que optimiza los resultados de obtención de parámetros y que hace que el proceso sea robusto. Fundamentalmente se fortalece la parte más sensible, que es la calibración de los parámetros intrínsecos de las cámaras.

Como conclusiones se puede comentar que la utilización del patrón de calibración propuesto simplifica el proceso de calibración de cada cámara y permite obtener con una buena precisión los parámetros extrínsecos de posicionamiento relativo en el sistema de referencia global. Queda patente a través de los resultados, que este tipo de patrón es mucho más insensible a posibles variaciones de los múltiples factores que afectan a la calibración, por lo que aun no disponiendo de condiciones ideales de calibración suministra resultados de gran precisión. Otra de las conclusiones es que con patrones del tipo propuesto, la desviación típica en la determinación de los centroides de los elementos de calibración es muy inferior al caso de los patrones tradicionales. Inclusive se puede utilizar para calibrar los parámetros extrínsecos del sistema global, donde el patrón debe estar muy alejado de la posición de las cámaras, obteniendo su centroide con calidad suficiente para determinar su posición y orientación con precisión.

Respecto de la calibración global del sistema, una conclusión importante es que la calibración en una sola fase de todo los parámetros intrínsecos y extrínsecos, llevaría a obtener los parámetros intrínsecos de las cámaras con una baja precisión debido a la lejanía en que se debería ubicar el patrón de calibración para ser captado por varias cámaras. Por ello, el proceso de calibración del sistema visión en dos fases propuesto, a tenor de la constatación de las



hipótesis y de los resultados obtenidos, es idóneo para este tipo de entornos y los resultados son óptimos.

En resumen, se puede decir que la propuesta de patrón y de estructura de calibración, permite obtener los parámetros intrínsecos de cada una de las cámaras y los extrínsecos de su localización en el sistema global, de forma rápida, robusta, precisa y sin requerimientos de infraestructura sofisticada.

Como aportaciones a destacar se pueden citar:

- Propuesta de un patrón de calibración novedoso que simplifica el proceso de calibración de cada cámara y permite obtener con una buena precisión los parámetros extrínsecos de posicionamiento relativo entre cámaras y respecto del sistema de referencia global.
- Estrategia de detección de centroides, a partir del patrón propuesto, con desviación típica en la determinación de los mismos inferior a la de los sistemas tradicionales existentes, con patrones tanto cercanos como lejanos.
- Estrategia de identificación automática de los elementos que componen los patrones de calibración, que posibilita realizar el proceso de forma autónoma, de una forma sencilla y con baja carga computacional y sin riesgos de calibraciones defectuosas debidas a errores por la intervención del usuario.
- Estudio de la influencia de un proceso de calibración en una única fase, en la determinación de parámetros de calibración y propuesta de una estructura del proceso en dos fases, calibrando primero los parámetros intrínsecos de las cámaras y posteriormente los parámetros extrínsecos del sistema global.

### **8.1.5. Guiado**

Finalmente, en el apartado de guiado de robots se ha propuesto una completa estrategia para conseguir realizar éste, siguiendo unas determinadas trayectorias. La propuesta realizada en esta tesis incluye la utilización de un Filtro de Kalman con un doble objetivo, por una parte permite obtener la posición del robot de forma más precisa, y por otra permite realizar una estimación de la posición del robot y con ello de una “ventana de interés” que incluye las

balizas, con el objetivo de centrar en ella el procesamiento de imágenes. Para la generación de trayectorias se propone la utilización de Splines cubicas, por sus excelentes prestaciones para conseguir trayectorias continuas con variaciones suaves, y con la posibilidad de forzar el paso por puntos intermedios prefijados.

Los resultados obtenidos tanto en las simulaciones como en las pruebas prácticas han sido altamente satisfactorios, confirmando la validez de la propuesta global de guiado realizada.

Como aportaciones destacables en el tema de guiado se pueden citar:

- Utilización de un Filtro de Kalman para realizar la corrección de la posición del robot independientemente del ruido del proceso y de medida, y la estimación de una “ventana de interés” que facilita la fase de procesamiento de imágenes.
- Propuesta para obtener un camino libre de obstáculos utilizando subgrafos de visibilidad y desdoblamiento en el cálculo de las Splines para puntos próximo (en los cuales estas curvan producen trayectorias sinuosas).
- Validación práctica, en un entorno de movimiento dentro del Edificio Politécnico, de la propuesta global de guiado a partir, fundamentalmente, de la información proporcionada por las cámaras ubicadas en el entorno. Conviene indicar que únicamente en aquellos tramos (en general muy cortos y que se pueden evitar añadiendo más cámaras) donde el robot no es visto por ninguna cámara se puede utilizar odometría para realizar el guiado.

## 8.2. Trabajos futuros.

Como ya se ha comentado anteriormente, la propuesta de guiado de robots móviles realizada en esta tesis es muy novedosa. De hecho hasta hace muy pocos años apenas existían trabajos en esta línea, si bien más recientemente está despertando el interés de varios grupos de investigación. Este aspecto novedoso hace que el trabajo realizado en esta tesis sea solamente un primer paso, y en consecuencia son muchos los aspectos que quedan por resolver hasta conseguir los ambiciosos objetivos que, dentro del concepto de “espacios inteligentes”, se plantean para conseguir espacios donde se puede inter-actuar en forma simple con ordenadores

y robots, y nos permite obtener servicios útiles de ellos. Por tanto, son numerosos los temas de investigación que se pueden abordar en este campo. En este sentido, y centrándonos en el guiado de robots móviles utilizando cámaras ubicadas en posiciones fijas del entorno, se pueden enumerar las siguientes líneas futuras de investigación (sobre algunas de las cuales ya se han iniciado trabajos dentro del Departamento de Electrónica):

- Diseño de un sistema de localización que permita al entorno reconocer robots móviles en ausencia de marcas artificiales en su estructura. Cualquier robot en el entorno deberá ser reconocido por el mismo mediante la propiedad de controlabilidad que el entorno tiene sobre éste. Esta nueva propuesta, donde no se tiene conocimiento previo del robot, obliga a la definición de un conjunto de marcas naturales pertenecientes a su estructura. Dichas marcas deberán ser identificadas de manera robusta en las sucesivas imágenes generadas por el movimiento del robot. Además, si se desea realizar un posicionamiento métrico en el espacio tridimensional se deberá relacionar la posición de dichas marcas en el espacio bidimensional de la imagen con la posición del robot. Por otra parte, este enfoque tiene como consecuencia el cálculo de la posición 3D del conjunto de marcas naturales detectadas y con ello un incremento en el conocimiento acerca de la estructura del robot. Por tanto, el sistema utilizando marcas naturales se convierte en uno de localización simultánea a la reconstrucción, puesto que se obtiene información acerca de la estructura al mismo tiempo que se localiza. Dar respuesta a todas estas cuestiones supone abordar diferentes tipos de problemas, entre los que se pueden destacar: definir un proceso de detección e identificación de marcas naturales robusto que resuelvan el problema de correspondencia entre marcas ya conocidas y las nuevas que surgen del movimiento; generalizar el proceso de localización para una o varias cámaras; solucionar el problema de la inicialización cuando no se tiene información previa de la estructura.
- Calibración de Múltiples cámaras ubicadas en el entorno utilizando un patrón activo. Este es un aspecto importante ya que los resultados de posicionamiento están estrechamente ligados a la calibración realizada. Para ello se puede abordar el diseño de un método de calibración de cámaras utilizando un panel LCD como patrón de calibración. Esto permitirá modificar la posición de los puntos fiduciales a voluntad, facilitando el proceso de calibración y mejorando el proceso de detección de los mismos, generando un sistema completamente autónomo. Todo ello conlleva: abordar aspectos relacionados con la definición

de las características a cumplir por el patrón de calibración a generar sobre el LCD (pantalla de ordenador portátil, por ejemplo), en función de los parámetros generales de las cámaras y de las distancias máximas cámara-patrón; desarrollar algoritmos de detección de puntos fiduciales sobre LCD con precisión subpíxelica; desarrollar un proceso de calibración basado en Homografías, etc.

- Desarrollo de hardware específico (reconfigurable y sobre FPGA's) para el procesamiento de imágenes, con el objetivo de descentralizar y aumentar la velocidad del procesamiento en cada nodo de visión del espacio inteligente.
- Generalización de algoritmos para permitir el guiado, e incluso la colaboración de múltiples robots en el entorno, generando múltiples trayectorias y evitando posibles colisiones entre ellos y con los obstáculos.
- Abordar líneas de investigación encaminadas a dotar al entorno, utilizando la información visual, de la capacidad para localizar y reconocer robots y personas que permitan la interacción natural entre los usuarios y los robots a partir, por ejemplo del reconocimiento de gestos.

### 8.3. Publicaciones derivadas de la tesis.

Parte de los trabajos desarrollado en esta tesis se han reflejado en las siguientes publicaciones:

- I. Fernández. J.L. Lázaro. M.Mazo. S. García. "*A non linear multicamera system calibration method for tele-operated robot*". International Conference TELECOM'02. Universidad de Alcalá (España). Vol.1.pp. 15-24. July 2002.

En este congresos se presento parte del sistema de calibración utilizado en esta tesis.

- I. Fernández. M. Mazo. J. L. Lázaro. P. Martín. S. García. "*Local Positioning System (LPS) for Indoor Environments Using a Camera Array*". The 11th International Conference on Advanced Robotics. ICAR'2003. Vol.2. pp. 613-618. Coimbra (Portugal) 2003.

En la comunicación presentada a esta conferencia se describen los métodos propuestos para obtener la posición 3D del robot.

- I. Fernández. M. Mazo. J. L. Lázaro. P. Martín. S. García. "*Posicionado y seguimiento de un robot móvil a partir de detección multicámara*". Seminario Anual de Automática Electrónica Industrial E Instrumentación. SAAEI'2002 Vol. 2. pp. 321-324. Alcalá de Henares 2002.

En esta comunicación se da una visión global del trabajo realizado en esta tesis.

- I. Fernández. M. Mazo. J. L. Lázaro. P. Martín. S. García. "*Sistema de Posicionamiento Local (LPS) en Interiores Utilizando un Array de Cámaras*". Seminario Anual De Automática Electrónica Industrial E Instrumentación. SAAEI'2003. Vol. 1. pp. 165-170. Vigo 2003.

En esta ponencia se presentaron fundamentalmente aspectos relacionados con la identificación de las balizas y el posicionamiento del robot.

## Apéndice A

---

### Calibración de cámaras con distancia focal fija

---

En este apéndice se muestra la problemática asociada a la calibración de una cámara con una óptica de distancia focal fija.

#### A.1. Modelo *pin-hole*.

El modelo de cámara que tradicionalmente se utiliza para pasar de coordenadas reales 3D a coordenadas 2D pertenecientes a la imagen captada, suele ser el modelo de proyección perspectiva denominado modelo *pin-hole* (figura A.1), siendo el utilizado en esta tesis.

Tal como se examinó en el Tema 3, las ecuaciones de este modelo son:

$$\lambda \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (\text{A.1})$$

$$\lambda \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{M} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (\text{A.2})$$

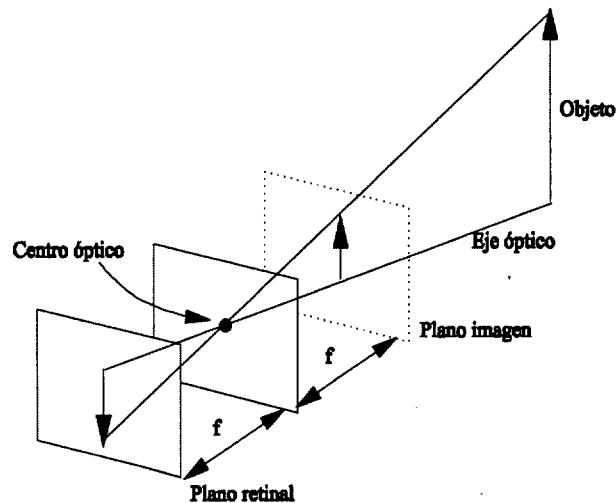


Figura A.1. Modelo geométrico de la cámara *pin-hole*.

Las ecuaciones (A.1) y (A.2) modelan el proceso de formación de imágenes o modelo geométrico de la cámara, que reflejan la conversión de un punto tridimensional en otro 2D. Las matrices  $\mathbf{R}$  y  $\mathbf{T}$  (parámetros extrínsecos) transforman las coordenadas de un objeto de la escena dadas en el sistema de referencia externo a coordenadas en el sistema de referencia de la cámara. Dicha matriz tiene seis parámetros independientes tres para la matriz de rotación y otros tres para el vector de traslación. Si se modifica la posición relativa entre un objeto de la escena y la cámara, las matrices  $\mathbf{R}$  y  $\mathbf{T}$  también se modifican. El resto de parámetros son inherentes a la óptica y al ajuste concreto en una cámara, por lo que se denominan parámetros intrínsecos. Los parámetros intrínsecos son constantes en tanto no varíen las características y posiciones relativas entre la óptica y el sensor imagen.

Si los componentes ópticos de la cámara fuesen perfectos, la transformación 3D-2D sería perfectamente lineal y fácil de resolver. Pero las lentes de la cámara producen aberraciones ópticas o distorsiones que no son lineales y afectan a la precisión de la transformación. Las ecuaciones que modelan estas distorsiones son:

$$\begin{aligned} x' &= x + x(k_1 r^2 + k_2 r^4) + 2p_1 xy + p_2(r^2 + 2x^2) \\ y' &= y + y(k_1 r^2 + k_2 r^4) + p_1(r^2 + 2y^2) + 2p_2 xy \end{aligned} \quad (\text{A.3})$$

los coeficientes  $k_1$ ,  $k_2$ ,  $p_1$  y  $p_2$  se consideran, también, parámetros intrínsecos.

## A.2. Calibración multi-imagen.

Los métodos clásicos de calibración de los parámetros de una cámara están basados en el procesamiento de una única imagen. La captura del patrón de calibración debe incluir numerosos puntos de los cuales se conoce su posición 3D. Hay que dejar patente que dichos métodos introducen errores en la medida de las coordenadas de los puntos posteriormente utilizados en el algoritmo de optimización.

Estos métodos presentan problemas cuando la distancia focal es muy pequeña (caso de objetivos de gran angular) no corrigiendo las distorsiones que presentan dicho tipo de objetivos. Además, se suponen conocidas las coordenadas reales de los elementos del patrón formando parte del proceso de optimización, por lo que los errores e incertidumbres de las medidas serán arrastrados en todos los cálculos posteriores. Así, los métodos tradicionales que hacían uso de una única imagen han dejado paso a métodos donde se considera una secuencia de imágenes del mismo patrón, permitiendo tener más ecuaciones, lo que se traduce en una mayor precisión en la calibración.

La calibración de cámara multi-imagen utiliza el modelo de proyección perspectiva pinhole. La matriz extrínseca  $\mathbf{RT}$ , matriz de rotación/traslación, permite un cambio entre los distintos sistemas de referencia involucrados: sistema de referencia del patrón/objeto  $\mathbf{R}_O$  y sistema de referencia de la cámara  $\mathbf{R}_C$ . Como existen múltiples capturas también existen diferentes  $\mathbf{RT}$ , para cada posición que ocupa el patrón en la escena.

El proceso de optimización debe obtener los siguientes valores: parámetros intrínsecos de la cámara, punto principal  $(u_0, v_0)$  y las distancias focales según el tamaño de píxel horizontal y vertical  $(f_x, f_y)$ , los parámetros de corrección de la distorsión ocasionada por el sistema de lentes, distorsión radial  $(k_1, k_2)$  y distorsión tangencial  $(p_1, p_2)$ , las coordenadas de los puntos del patrón con respecto al sistema de referencia del patrón  $\mathbf{R}_O$  y los parámetros de todas las matrices  $\mathbf{RT}$  ( $\alpha^i, \beta^i, \gamma^i, T_x^i, T_y^i, T_z^i$ ), existiendo una matriz  $\mathbf{RT}$  por cada imagen que toma el sistema visual. Todo lo comentado anteriormente se muestra en la figura A.2.



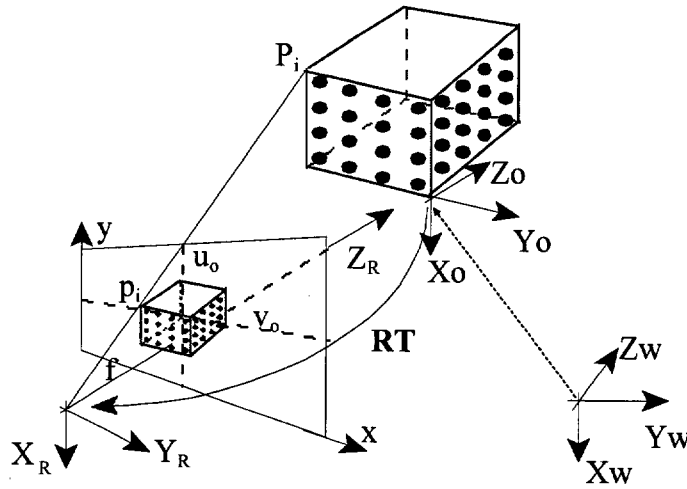


Figura A.2. Sistemas de referencia utilizados en la calibración de la cámara.

Un punto  $P$  de la escena y con coordenadas 3D  $[X, Y, Z]^T$  se proyecta en una imagen en el punto  $p$  de coordenadas píxelicas  $(u, v)$ , dadas, según (A.2), por:

$$\begin{aligned} u &= \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \\ v &= \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \end{aligned} \quad (\text{A.4})$$

Reestructurando (A.4) se obtiene:

$$\begin{aligned} m_{11}X + m_{12}Y + m_{13}Z + m_{14} - u \cdot m_{31}X - u \cdot m_{32}Y - u \cdot m_{33}Z - u \cdot m_{34} &= 0 \\ m_{21}X + m_{22}Y + m_{23}Z + m_{24} - v \cdot m_{31}X - v \cdot m_{32}Y - v \cdot m_{33}Z - v \cdot m_{34} &= 0 \end{aligned} \quad (\text{A.5})$$

Cada punto 3D proporciona dos ecuaciones como (A.5), de forma que para los  $n$  puntos del patrón de calibración tendremos  $2 \cdot n$  ecuaciones, las cuales se pueden expresar en forma matricial como:

$$\begin{bmatrix}
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_i & -u_i X_i & -u_i Y_i & -u_i Z_i & \cdot \\
 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_i & -v_i X_i & -v_i Y_i & -v_i Z_i & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
 \end{bmatrix}
 \begin{bmatrix}
 m_{11} \\
 m_{21} \\
 \cdot \\
 \cdot \\
 \cdot \\
 m_{32} \\
 m_{33} \\
 m_{34}
 \end{bmatrix}
 =
 \begin{bmatrix}
 0 \\
 0 \\
 \cdot \\
 \cdot \\
 \cdot \\
 0 \\
 0 \\
 0
 \end{bmatrix}
 \quad (A.6)$$

El sistema dado en (A.6) es de la forma  $QA=0$ , siendo necesario utilizar el algoritmo SVD (*Singular Value Decomposition*). Con los valores de los coeficientes  $m_{ij}$  se obtiene los parámetros intrínsecos y extrínsecos de la cámara. Éstos van a ser utilizados como valores iniciales del proceso recursivo propuesto en [Lavest 92], que permiten obtener de una forma más precisa de los parámetros intrínsecos y extrínsecos de la cámara, incluidos los de la distorsión. Este método se desarrolla a continuación.

De la ecuación (A.1) se pueden obtener las ecuaciones de colinealidad en fotogrametría:

$$\begin{aligned}
 x &= \frac{r_{11}X + r_{12}Y + r_{13}Z + T_X}{r_{31}X + r_{32}Y + r_{33}Z + T_Z} \\
 y &= \frac{r_{21}X + r_{22}Y + r_{23}Z + T_Y}{r_{31}X + r_{32}Y + r_{33}Z + T_Z}
 \end{aligned}
 \quad (A.7)$$

donde  $(x, y)$  corresponden a las coordenadas del punto  $p$  de la imagen dado en unidades métricas, siendo su relación con las coordenadas  $(u, v)$ :

$$x = \frac{u - u_0}{f_x} \quad y = \frac{v - v_0}{f_y}
 \quad (A.8)$$

Debido a la distorsión, las coordenadas  $(x, y)$  se convierten en  $(x', y')$  dadas por:

$$x' = \frac{u + e_u - u_0}{f_x} \quad y' = \frac{v + e_v - v_0}{f_y}
 \quad (A.9)$$

donde  $e_u$  y  $e_v$  reflejan los términos del error en la obtención de las coordenadas píxeles de la proyección del punto  $P$ . A su vez, la ecuación (A.3) se puede expresar como:

$$\begin{aligned}x' &= x + do_x \\y' &= y + do_y\end{aligned}\tag{A.10}$$

donde  $do_x$  y  $do_y$  son las componentes de distorsión en  $x$  e  $y$ , cuyos valores son:

$$\begin{aligned}do_x &= x(k_1 r^2 + k_2 r^4) + 2p_1 xy + p_2(r^2 + 2x^2) \\do_y &= y(k_1 r^2 + k_2 r^4) + p_1(r^2 + 2y^2) + 2p_2 xy\end{aligned}\tag{A.11}$$

donde los valores de  $x$  e  $y$  vienen dados por la ecuación (A.10).

De (A.7), (A.9) y (A.10) se obtiene:

$$\begin{aligned}u + e_u &= u_0 + f_x \frac{r_{11}X + r_{12}Y + r_{13}Z + T_x}{r_{31}X + r_{32}Y + r_{33}Z + T_z} + f_x \cdot do_x \\v + e_v &= v_0 + f_y \frac{r_{21}X + r_{22}Y + r_{23}Z + T_y}{r_{31}X + r_{32}Y + r_{33}Z + T_z} + f_y \cdot do_y\end{aligned}\tag{A.12}$$

Para cada uno de los  $n$  puntos de la plantilla de calibración se obtiene dos ecuaciones como (A.12). El proceso a seguir consistirá en la obtención de los parámetros (intrínsecos y extrínsecos) que minimicen  $e_x$  y  $e_y$ . Para ello se define un vector  $\Phi$  (vector de parámetros) que contiene todas las incógnitas del sistema:

$$\Phi = (u_0, v_0, f_x, f_y, k_1, k_2, p_1, p_2, \alpha, \beta, \gamma, T_x, T_y, T_z)\tag{A.13}$$

Utilizando (A.13), la ecuación (A.12) se puede escribir como:

$$\left. \begin{aligned}e_u &= F(\Phi) - u \\e_v &= G(\Phi) - v\end{aligned} \right\} Q(\Phi)\tag{A.14}$$

El problema consiste en estimar el vector  $\Phi$  que minimice:

$$\sum_{i=1}^n (e_u^2 + e_v^2) \quad (\text{A.15})$$

En (A.14)  $F(\Phi)$  y  $G(\Phi)$  son dos funciones no lineales de  $\Phi$ , para su linealización se utiliza el método de Newton-Raphson:

$$Q(\Phi) = v(\Phi_0) + \left. \frac{\partial Q(\Phi)}{\partial \Phi_i} \right|_{\Phi=\Phi_0} \cdot \Delta(\Phi) \quad (\text{A.16})$$

que, de forma simplificada, se puede expresar:

$$Q = Q_0 + J \cdot \Delta \Phi \quad (\text{A.17})$$

con:

$$Q_0 = Q(\Phi_0) \quad y \quad J = \left. \frac{\partial Q(\Phi)}{\partial \Phi_i} \right|_{\Phi=\Phi_0} \quad (\text{A.18})$$

Los incrementos del vector de parámetros ( $\Delta\Phi$ ) vienen dados por:

$$\Delta \Phi = (J^T \cdot W \cdot J)^{-1} (J^T \cdot W \cdot Q_0) \quad (\text{A.19})$$

Para mejorar el resultado de la calibración se pueden tomar varias imágenes del patrón de calibración situado en diferentes posiciones. En este caso, los parámetros intrínsecos no cambian, no sucede lo mismo con los extrínsecos. Así, para un conjunto de  $m$  imágenes el vector de parámetros vendrá dado por:

$$\Phi = (u_0, v_0, f_x, f_y, k_1, k_2, p_1, p_2, \alpha^1, \beta^1, \gamma^1, T_X^1, T_Y^1, T_Z^1, \dots, \alpha^m, \beta^m, \gamma^m, T_X^m, T_Y^m, T_Z^m) \quad (\text{A.20})$$

La calibración de cámara multi-imagen obtiene los valores de los parámetros del modelo de formación de las imágenes presentado más arriba. El número de incógnitas es el siguiente:

- 11 parámetros intrínsecos (teniendo en cuenta 5 parámetros de distorsión radial y dos de distorsión tangencial).

- 6 parámetros extrínsecos para cada imagen del patrón captada.

Así, el número de incógnitas a determinar aumenta a: “ $11+6m+3n$ ” donde  $m$  representa el número de imágenes que se han capturado y  $n$  el número de puntos de los que consta el patrón. Para poder resolver el sistema de ecuaciones sobredimensionado se ha de cumplir la relación A.21 entre número de ecuaciones e incógnitas.

$$2.n.m \geq 11 + 6.m \quad (\text{A.21})$$

Si se aplica la técnica de auto-calibración también los puntos del patrón se han de incluir como incógnitas en el vector  $\Phi$ , quedando la ecuación A.20 como:

$$\Phi = (u_0, v_0, f_x, f_y, k_1, k_2, p_1, p_2, \alpha^1, \beta^1, \gamma^1, T_X^1, T_Y^1, T_Z^1, \dots, \alpha^m, \beta^m, \gamma^m, T_X^m, T_Y^m, T_Z^m, X^p, Y^p, Z^p) \quad (\text{A.22})$$

Ahora la relación A.21 queda:

$$2.n.m \geq 11 + 6.m + 3.n \quad (\text{A.23})$$

### A.2.1. Detección sub-pixélica de los círculos del patrón.

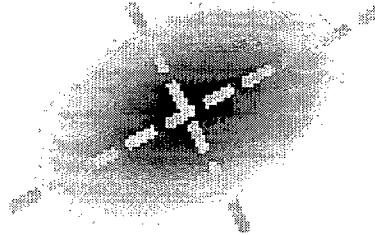
La extracción de los puntos del patrón involucra principalmente dos pasos. En primer lugar se ha de identificar cada punto o marca en cualquier imagen del patrón. Una vez identificado el punto se ha de refinar su posición de manera que se localice el centro del círculo con una precisión por debajo de una décima de píxel.

Hay que hacer notar que si cambia la orientación de captura del patrón de calibración respecto a la cámara la marca se deforma pero el algoritmo es capaz de recuperar correctamente la posición que ocupa el centro del círculo utilizando técnicas de ajuste en perspectiva para no introducir errores en el algoritmo posterior de calibración. La marca circular se deforma como consecuencia de la perspectiva en una forma elipsoidal. El algoritmo de detección obtiene los parámetros de la forma elipsoidal que más se aproxima al círculo deformado.

La determinación de la posición de los puntos del patrón de calibración, con el mínimo error posible, se realiza con una transformación fotométrica a partir de la luminancia de la zona teniendo en cuenta que la deformación de un círculo tiene una forma elipsoidal. En nuestro

caso, se han optimizado seis parámetros para asegurar que se tiene un modelo correcto y que se ajusta a la forma elipsoidal: niveles de luminancia mínimo y máximo, radio de curvatura de los arcos que conforman el pico y el valle de la luminancia, perfil de la línea que une dichos arcos y un parámetro adicional que permite trabajar con puntos de un patrón de calibración mayor o menor.

La transformación geométrica tiene en cuenta la posición del patrón de calibración con respecto a la cámara y así parametrizar la forma del punto circular del patrón captado conforme a la perspectiva que se tiene. En la figura A.3 se muestra el modelo de luminancia descrito.



**Figura A.3.** Modelo de luminancia de un punto circular.



---

## Bibliografía

---

- [Adán 02] Adán A. Feliu V. Ramos F. A New Approach For Mobil Robot Self-Location In Three Dimensional Environments. *IFAC. 15th Trienal Word Congres.* Barcelona. 2002.
- [Agrawal 03] Aggarwal M. Ahuja N. Camera Center Estimation, *Int. Conference on Pattern Recognition, ICPR 2000.* Vol. 1. pp. 876-880. September. Barcelona 2000.
- [Aitenbichler 03] Aitenbichler E. Mühlhäuser M. An IR Local Positioning System for Smart Items and Devices. *Proceedings of the 23rd IEEE International Conference on Distributed Computing Systems Workshops (IWSAWC03). 3rd International Workshop on Smart Appliances and Wearable Computing.* pp. 334-339. 2003.
- [Althaus 01] Althaus P. Christensen H. I. and Hoffmann F. Using the dynamical system approach to navigate in realistic real-world environments, *Proceedings of IEEE International Conference on Intelligent Robots and Systems IROS'2001.* pp. 1023-1029. (Maui, Hawaii). November. 2001.



- [Antonelli 03] Antonelli G. Chiaverini S. Fusco G. An odometry calibration method for mobile robots based on the least-squares technique. *American Control Conference, 2003. Proceedings of the 2003* . Vol. 4. pp.3429 - 3434. June 2003.
- [Arai 03] Arai Y. Sekiai M. Absolute position measurement system for mobile robot based on incident angle detection of infrared light. *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on* . Vol. 1. pp.986 - 991. October 2003.
- [Baker 03] Baker P. T. Aloimonos Y. Calibration of a Multicamera Network *2003 Conference on Computer Vision and Pattern Recognition Workshop*. Vol. 7. pp-728-234. Madison, Wisconsin. June 2003.
- [Barreto 04] Barreto J. P. and Daniilidis K. Wide Area Multiple Camera Calibration and Estimation of Radial Distortion. *OMNIVIS 2004 - Workshop on Omnidirectional Vision and Camera Networks*, Prague, Czech Republic. May 2004.
- [Bobick 97] A. Bobick. S. Intille. J. Davis. F. Baird. C. Pinhanez. L. Campbell. Y. Ivanov. A. Schütte. A. Wilson. The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment, *Technical report 398, MIT Media Laboratory, Perceptual Computing Section*. 1997.
- [Borenstein 96a] Borenstein J. Everet H.R. Feng L. Where an I?. Sensor and Methods for Mobile Robot Positioning. *The University of Michigan*. April 1996.
- [Borenstein96b] Borenstein J. Feng L. Gyrodometry. A New Method for Combining Data from Gyros and Odometry in Mobile Robots. *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*. Minneapolis. Vol. 1. pp. 423-428. April 1996.
- [Bostan 03] Bostan I. Lita I. Franti E. Dascalu M. Moldovan C. Goschin S. Systematic odometry errors compensation for mobile robot positioning. *CAD Systems in Microelectronics, 2003. CADSM 2003. Proceedings of the 7th International Conference. The Experience of Designing and Application of*. Vol.1. pp.574 - 576. 18-22 February. 2003.
- [Bouguet 99] Bouguet J.Y. Visual methods for three-dimensional modeling *Ph.D. Thesis. California Institute of Technology*. Pasadena California. 1999.

- [Boyd 98] Boyd J.E. Hunter E. Kelly P.H. Tai Li-Cheng. Clifton B. P. Ramesh C. J. MPI-Video infrastructure for dynamic environments. *In Proc. IEEE Int. Conf. Multimedia Computing and Systems*. Vol. 1. pp. 249–254. 1998.
- [Brown 71] Brown D.C. Close-Range Camera Calibration. *Photogrammetric Engineering*, XXXVII. Vol. 1. Pp.855 - 866. August 1971.
- [Burschka 97] Burschka D. Farber G. Active controlled exploration of 3D environmental models based on a binocular stereo system. *Advanced Robotics, 1997. ICAR '97. Proceedings., 8th International Conference on*. Vol. 1. pp:971 - 977.7-9 July 1997.
- [Caffery 98] Caffery J. and Stuber G. Subscriber Location in CDMA Cellular Networks. *IEEE Transactions on Vehicular Technology*, Vol. 47. N° 2. pp. 406-416. 1998.
- [Chen 00] Chen, X. Davis, J. Slusallek, P. Wide area camera calibration using virtual calibration objects. *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*. Vol. 2. pp.520 - 527. 13-15 June 2000.
- [Chen 01] Chen Y.S. Shih S.W. Hung Y.P. Fuh C.S. Simple and Efficient Method of Calibrating a Motorized Zoom Lens. *Journal of Image and Vision Computing*. Vol.19. N°14. pp. 1099-1110. 2001.
- [Cipolla 98] Cipolla R. Boyer E. 3D Model Acquisition from Uncalibrated Images. *Proceedings of IAPR Workshop on Machine Vision Applications*. Vol. 1. pp. 559-568. Chiba (Japan) November 1998.
- [Coen 98] Coen M.H. Design Principles for Intelligent Environments. In *Proceedings of the AAAI Spring Symposium on Intelligent Environments (AAAI98)*. Vol. 1. pp. 547-551. Standford, C.A. 1998..
- [Collins 01] Collins R.T. Lipton A.J. Fujiyoshi H. Kanade T. Algorithms for Cooperative Multisensor Surveillance. *Proceedings of the IEEE*. Vol. 89. N°. 10. pp.14156-1477. 2001.
- [DeMenthon 95] DeMenthon D. F. and Davis L. S. Model--based object pose in 25 lines of code. *International Journal of Computer Vision*. Vol.15. Pp.123-141. 1995.
- [Duckett98] Duckett T. Nehmzow U Mobile robot sel-loclisation and measurement of performance in middle-scale environments. *Robotics and Autonomous Systems*. Vol. 24. pp. 57-69. 1998.

- [Eren 97] Eren, H.; Fung, C.C.; Newcombe, D.; Goh, J. Operation of mobile robots in a structured infrared environment. *Instrumentation and Measurement Technology Conference, 1997. IMTC/97. Proceedings. 'Sensing, Processing, Networking'. IEEE*. Vol. 1. Pp.20 - 25. 19-21 May 1997.
- [Essa 98] I. Essa. G. Abowd. C. Atleson. Ubiquitous Smart Spaces, *A white paper submitted to DARPA*. February 1998.
- [Esteves 03] Esteves, J.S., Carvalho, A., Couto, C. Generalized geometric triangulation algorithm for mobile robot absolute self-localization. *Industrial Electronics, 2003. ISIE '03. 2003 IEEE International Symposium on*. Vol. 1. pp.346 - 351 9-11 June 2003.
- [Everet 94] Everett H.R. Gage D.W. G.A. Gilbreth. R.T. Laird. R.P. Smurlo. Real-World Issues in Warehouse Navigation. *Proceedings SPIE Mobile Robots IX*. Vol. 2352. pp 2-4. Boston. MA. November 1994.
- [Faugeras 93] Faugeras O. D. Three-Dimensional Computer Vision. *The MIT Press*. Cambridge.MA. 1993.
- [Faugeras 01] Faugeras O. Luong Q.T. The Geometry of Multiple Images, The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications. *MIT Press*. Cambridge. Massachusetts. 2001.
- [Fiala 02] Fiala M. and Basu A. Robot Navigation Using Panoramic Landmark Tracking. *Proceedings of Vision Interface*. Vol.1. pp.27-29. Calgary. Canada. May.2002.
- [Foux 93] Foux G. Heymann M. Bruckstein A. Two-Dimensional Robots Navigation Among Unknow Stationary Obstacles. *IEEE International Conference on Robotics and Automation*. Vol.1. pp. 96-101. February 1993.
- [Fraundorfer 02] Fraundorfer F. A map for mobile robots consisting of a 3D model with augmented salient image features. *26th Workshop of the Austrian Association for Patern Recognition(ÖOAGM/AAPR)*. Vol.2. pp. 249-256. 2002.
- [Fujiyoshi 03] Fujiyoshi H., Shimizu S., Nishi T., Nagasaka Y., Takahashi T. Fast 3D position measurement with two unsynchronized cameras. *Computational Intelligence in Robotics and Automation, 2003. Proceedings. 2003 IEEE International Symposium on*. Vol. 3. pp.1239 - 1244. 16-20 July 2003.

- [García 01] García J.C. Sistema de posicionamiento y autocalización para sillas de ruedas autónomas. *Tesis Doctoral. Escuela Politécnica. Universidad de Alcalá.* 2001.
- [García 93] García R. Contribución a un sistema integral de navegación, de robots móviles multisensoriales, mediante infrarrojos. *Tesis Doctoral. Facultad de Informática. Universidad Politécnica de Madrid.* 1993.
- [García 00] García J.C. Mazo M. Ureña J. Marrón M. Escudero M. Sistema de Posicionamiento Absoluto Mediante Marcas Artificiales. *Proceedings of TELECOM'2000 International Conference.* Universidad de Alcalá (España). Vol.1.pp. 1-10. July 2000.
- [Gardel 04] Gardel A. Calibración de una cámara con zoom y movimiento pan-tilt *Tesis Doctoral. Escuela Politécnica. Universidad de Alcalá.* 2004.
- [Gardel 03] Gardel A. Lázaro J. L. Lavest J. M. Camera Auto-Calibration with Virtual Patterns. *Proceedings of the ETFA 2003.* Vol.2. pp. 566-572. Lisboa. Septiembre 2003.
- [Gijeong 03] Gijeong J., Sungho K., Wangheon L., Inso K. Robust self-localization of mobile robots using artificial and natural landmarks. *Computational Intelligence in Robotics and Automation, 2003. Proceedings. 2003 IEEE International Symposium on.* Vol. 1. pp. 412 - 417. 16-20 July 2003.
- [Hartley93] Hartley R.I. Euclidean reconstruction from uncalibrated views. *Proceeding of the DARPA-ESPRIT workshop on Applications of Invariants in Computer Vision.* Vol.1. pp 187-202. Azores (Portugal) October. 1993.
- [Haya 04] P.A. Haya. G. Montoro. X. Alamán. A prototype of a context-based architecture for intelligent home environments. *International Conference on Cooperative Information Systems (CoopIS 2004).* Vol. 1. pp. 477-491. Larnaca, Cyprus. October 2004.
- [Hazas 03] Hazas M. Ward A. A high performance privacy-oriented location system. *Pervasive Computing and Communications, 2003. (PerCom 2003). Proceedings of the First IEEE International Conference on.* Vol.1. pp. 216 - 223. 23-26 March 2003.
- [Heikkila 97] Heikkila J. Silven O. A four-step camera calibration procedure with implicit image correction. *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on.* Vol.2. pp. 1106 - 1112. June 1997.

- [Heikkilä 00] Heikkilä J. Geometric Camera Calibration Using Circular Control Points. *IEEE Transactions Pattern Analysis and Machine Intelligence*. Vol. 1, pp 1066-1076. 2000.
- [Heren 97] Heren H. Fung C. C. Position Estimation of Mobile Robots Based on Coded Inferred Signal Transmission. *IEEE T Transactions on Instrumentation and Measurement*. Vol 46, N°6, December 1997.
- [Hernandez 03] Hernandez, S.; Morales, C.A.; Torres, J.M.; Acosta, L. A new localization system for autonomous robots. *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, Vol. 2. pp.1588 - 1593. 14-19 Sept. 2003.
- [Hoover 99] Hoover A. Olsen B. D. Path Planning for Mobile Robots Using a Video Camera Network, *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. pp. 890-895. Atlanta GA. 1999.
- [Hoover 00] Hoover A. Olsen B. D. Sensor network perception for mobile. *IEEE Conference on Robotics and Automation*. Vol. 2. pp 965-980. April 2000.
- [Hoover01] Hoover A. Olsen B. D. Calibrating a Camera Network Using a Domino Grid, *Pattern Recognition (journal)*, Vol. 34. N° 5. May 2001.
- [Heraud 94] Heraud R. Christy S. Dornaika F. Object pose: The link between weak perspective, para perspective, and full perspective. *Technical Report 2356, INRIA*, September 1994.
- [Ito 91] Ito M. Robot vision modelling- camera modelling and camera calibration. *Advanced Robotics*. Vol.5. pp 321-335. 1991.
- [Jae 03] Jae Seok P. Myung Jin C. Path planning with uncalibrated stereo rig for image-based visual servoing under large pose discrepancy. *Robotics and Automation, IEEE Transactions on*. Vol. 19. Issue. 2. pp.250 - 258. April 2003.
- [Kanayama 86] Kanayama Y. Miyake M. Trajectory Generation for Mobile Robots. *Robotic Research. MIT Press*. Vol. 3. pp. 330-340. 1986.
- [Kee 00] Kee C. Jun H. Yun D. Kim B. Kim Y. Parkinson B. Lenganstein T. Pullen S. and Lee J. Development of Indoor Navigation System using Asynchronous Pseudolites. *Proceedings of ION GPS-2000*. pp.121.125. Salt Lake Utah, (USA). September 2000.

- [Krumm 00] Krumm J. Harris S. Meyers B. Brumitt M. Hale M. Shafer S. Multi-Camera Multi-Person Tracking for EasyLiving. *Proceeding of the Third IEEE International Workshop on Visual Surveillance*. pp. 3-10. Dublin, Ireland July 2000.
- [Kruse 98] Kruse E. Wahl F.M. Camera-based monitoring system for mobile robot guidance. *Proceedings., 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Vol.2. pp. 1248 - 1253. 1998.
- [Kudoh 00] Kudoh E. Shibuya A. Ogawa T. Uchida D. Nakatsugawa M. Suda H. Kubota S. Picocell network for local positioning and information system. *Industrial Electronics Society, 2000. IECON 2000. 26th Annual Conference of the IEEE*. Vol. 2. pp.1165 - 1170. 22-28 Oct. 2000.
- [Ladd 02] Ladd A. M. Bekris K. E. Rudys A. Marceau G. Kavraki L. E. Wallach D. S. Robotics Based Location Sensing using Wireless Ethernet. in *Proc. 8th ACM Int. Conf. Mobile Computing and Netwrking*. pp. 227-238. Atlanta GA. September. 2002,
- [Latombe 91] Latombe J.C. Robot Motion Planning. *Kluwer Academic Publishers*, 1991.
- [Lavest, 99] Lavest J. M. Viala M. Dhome M. Quelle précision pour une mire d'étalonnage?. *Traitment du Signal*. Vol.7. N° 1. pp 241-254. 1999.
- [Lavest, 92] Lavest J. M. Stéréovision axiale par zoom pour la robotique, *These de Doctorat de l'Université Blaise-Pascal. Clermont-Femand*, 1992.
- [Lavest 98] Lavest J.M. Viala M. Dhome M. Do We Really Need an Accurate Calibration Pattern to Achieve a Reliable Camera Calibration?. *ECCV'1998*. Vol: 1. pp. 158-174 Freiburg, Germany. June 1998.
- [Lázaro 00a] Lázaro J.L. Mataix C. Gardel A. Martín E. Calibración de cámaras CCD con corrección bidimensional de ópticas de gran angular. *Información Tecnológica*. Vol: 11. N° 1. pp. 129-134. Chile. 2000.
- [Lázaro 00b] Lázaro J.L. Mataix C. Gardel A. Mazo M. Modelado de lentes radiales con distorsión y calibración de detectores CCD que las incluyen. *Información Tecnológica*. Vol. 11. N° 1. pp. 237-242. Chile. 2000.
- [Lee 03] Joo-Ho Lee. K. Morioka.. H. Hashimoto. Human-following mobile robot in a distributed intelligent sensor network. *Industrial Electronics, IEEE Transactions on*. Vol. 51. pp.229 - 237. February 2004.

- [Lee 01] Lee J. Ando N. Yakushi T. Nakajima K. Adaptive Guidance for Mobile Robots in Intelligent Infrastructure. *Proceedings 2001 IEEE/RSJ International Conference on Robots and Systems*, Vol.2. pp: 90–95. 2001.
- [Lee 99] Lee J. Ando N. Hashimoto H. Intelligent Space for Human and Mobile robot. *Proceedings of IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. pp: 784 . 1999.
- [Lee 04] Joo-Ho Lee. K. Morioka. N. Ando. H. Hashimoto. Cooperation of distributed intelligent sensors in intelligent environment. *Mechatronics, IEEE/ASME Transactions on*. Vol. 9. pp.535 - 543. September 2004 .
- [LeMaster 02] LeMaster E.A. Matsuoka M. Rock S.M. Field demonstration of a Mars navigation system utilizing GPS pseudolite transceivers *Position Location and Navigation Symposium, 2002 IEEE*. Vol.2. pp.150 - 155.15-18 April 2002.
- [Liao 99] Liao P.S. Chen T.S. Chung P.C. A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering*. Vol.2. pp 713-723. 1999.
- [Lilienthal 03] A. Lilienthal. T. Duckett. An Absolute Positioning System for 100 Euros. *Proceedings of the IEEE International Workshop on Robotic Sensing (ROSE 2003)*. June 2003.
- [Liu 01] Liu Y. Hoover A. I.D. Walker. Sensor Network Based Workcell for Industrial Robot. *Proceeding of the International Conference on Intelligent Robots and Systems*. Vol.2. pp. 1434-1439. Maui, Hawaii (USA). 2001.
- [Lozano 83] Lozano-Perez T. Wesley A. An algorithm for Planning Colision freePath Among Polyedral Obstacles. *Communications of de ACM*. pp.383-393. 1983.
- [Maaref 02] Maaref H. Barret C. Sensor-based navigation of a mobile robot in an indoor environment. *Robotics and Autonomous Systems*. Vol. 38. Issue 1. pp. 1-18 January 2002.
- [Marrón 02] Marrón M. Bueno E. J. Mazo M. Diseño de un algoritmo de visión para la navegación de un robot móvil en interiores estructurados. *SAAEI'02*. Vol. 1. pp. 343-350. Alcalá de Henares. 2002.

- [Martinez 03] Martinez E. Ohya A. Yuta S. Recognition of people's positioning by cooperative mobile robots for human groups steering. *Computational Intelligence in Robotics and Automation, 2003. Proceedings. 2003 IEEE International Symposium on*. Vol. 2. pp.758 - 763.16-20 July 2003.
- [Mata 01] Mata M. Armingol J.M. Escalera A. de la. Salichs M.A. A visual landmark recognition system for topological navigation of mobile robots Robotics and Automation. *Proceedings 2001 ICRA. IEEE International Conference on*. Vol. 2. pp. 1124 -1129. 2001.
- [McNamee 03] McNamee L. P. Photogrammetric Calibration of Mobile Robot Kinematics. *Doctoral dissertation, Ottawa-Carleton Institute for Electrical and Computer Engineering. School of Information Technology and Engineering (Electrical and Computer Engineering)*. June 2003.
- [Melen 94] Melen T. . Geometrical Modelling and Calibration of Video Cameras for Underwater navigation. *Doctoral dissertation, Norwegian Univ. of Science and Technology*. Trondheim, Norway. 1994.
- [Mizoguchi 96] H. Mizoguchi. T. Sato. T. Ishikawa. Robotic office room to support office work by human behavior understanding function with networked machines. *In Proceedings of Robotics and Automation, IEEE International Conference on*, Vol.4. pp: 2968 - 2975. 22-28 April 1996.
- [Morioka 02] Morioka K. Lee J. Hashimoto H. Human Centered Robotics in Intelligent Space. *Proceedings IEEE/ International Conference on Robotics and Automation*, Vol.3. pp: 2110 -2115. 2002.
- [Mouaddib 00] Mouaddib E. Marhic B. Geometrical Matching for Mobile Robot Localization. *IEEE Transactions on Robotics and Automation*. Vol. 16. Nº. 5. pp.542-552. October. 2000.
- [Murray 00] Murray D. and Little J. Using real-time stereo vision for mobile robot navigation. *Autonomous Robots*. Vol. 8. Nº 2. pp. 161.171. April 2000.
- [Nakju 03] Nakju D., Choset, H., Wan Kyun C. Accurate relative localization using odometry. *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*. Vol. 2. pp.1606 - 1612.14-19 September 2003.



- [Nguyen 03] Nguyen X. D., Bum-Jae Y., Sang-Rok O., Myung H. Visual self-localization for indoor mobile robots using natural lines. *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*. Vol. 2. pp.1252 - 1257. 27-31 October 2003.
- [Nilsson 69] Nilsson N.J. A Mobile Automaton: An Application of Artificial Intelligence Techniques. *Proc. of the 1st. International Joint Conference on Artificial Intelligence*. Vol.3. pp 509-520.1969.
- [Ocaña 04] Ocaña M. Bergasa L.M. Sotelo M.A. Robust Navigation Indoor Using Wifi Localization. *10th IEEE International Conference on Methods and Models in Automation and Robotics*30. Miedzyzdroje, Poland. August - September 2004.
- [Ohya 01] Ohya, A.; Miyazaki, Y.; Yuta, S. Autonomous navigation of mobile robot based on teaching and playback using trinocular vision. *Industrial Electronics Society, 2001. IECON '01. The 27th Annual Conference of the IEEE*. Vol. 1. pp.398 - 403. 29 November.-2 December 2001.
- [Okada 03] Okada K., Inaba M., Inoue H. Integration of real-time binocular stereo vision and whole body information for dynamic walking navigation of humanoid robot. *Multisensor Fusion and Integration for Intelligent Systems, MFI2003. Proceedings of IEEE International Conference on*. Vol.2. pp.131 - 136.30 July-1 August 2003.
- [Otsu 79] Otsu N. A threshold selection method from gray-level histograms. *IEEE Transactions On-Systems, Man, and Cybernetics*. Vol. 8. pp.66-69. 1979.
- [Paz 97] E. Paz. Sistema de posicionamiento basado en visión y planos CAD para navegación de robots móviles de interiores. *Tesis Doctoral. E.T.S. de Ingenieros Industriales y Minas. Universidad de Vigo*. Octubre 1997.
- [Peixoto 00] Peixoto P. Batista J. Araújo H. Integration of information from several vision systems for a common task of surveillance. *Robotics and Autonomous Systems*. Vol. 31. pp.99-108. 2000.
- [Pentland 97] A. Pentland. Smart rooms, desks and clothes. *Acoustics, Speech, and Signal Processing, ICASSP-97. IEEE International Conference on*. Vol. 1 pp.171 - 174. 21-24 April 1997.

- [Ponweiser 01] Ponweiser, W. Ayromlou, M. Vincze, M. Beltran, C. Madsen, O. Gasteratos, A. RobVision: vision based navigation for mobile robots *Multisensor Fusion and Integration for Intelligent Systems, 2001. MFI 2001. International Conference on*. Vol.2. pp:109 - 114.20-22 August 2001.
- [Prigge 02] Prigge, E.; How, J. Signal architecture for a Distributed Magnetic Local Positioning System Sensors, 2002. *Proceedings of IEEE*, Vol. 2. pp.1497 - 1504. 12-14 June 2002.
- [Rajagopalan 97] Rajagopalan R. A generic kinematic formulation for wheeled mobile robots. *Journal of Robotics System*. Vol.2. pp 77-91.1997.
- [Rander 98] Rander P. A Multi-Camera Method for 3D Digitalization the Dynamics, Real-World Events. *Doctoral dissertation, tech. report CMU-RI-TR-98-12, Robotics Institute, Carnegie Mellon University*, May 1998
- [Rao 88] Rao N.S.V. Iyengar S.S. deSaussure G. The Visit Problem: Visibility Graph-Based Solution. *Proc. of IEEE Conference on Robotics and Automation*. Vol.1. pp.1.650-1.655.1988.
- [Rombaut 91] Rombaut M. Segovia A. Meziel D. Preciado A. Displacements of a Mobile Robot in a Known Environment. *IMACS Symposium MCTS*. 1991.
- [Romero 00] Romero J. A. Contribución al control de un Robot Móvil Guiado Mediante Visión Artificial. *Tesis Doctoral. Facultad de Informática. Universidad Politécnica de Madrid*. 2000.
- [Rupp 00] Rupp T. Levi P. Optimized landmark arrangement for absolute localization-a practical approach. *Proceedings of Intelligent Robots and Systems. (IROS 2000). 2000 IEEE/RSJ International Conference on*. Vol. 1 pp.448 -453. 2000.
- [Saeedi 03] Saeedi P. Lowe D.G. Lawrence P.D. 3D localization and tracking in unknown environments. *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*. Vol. 1. pp.1297 - 1303.14-19 September 2003.
- [Saeedi 00] Saeedi P. Lawrence P. Lowe, D. 3D motion tracking of a mobile robot in a natural environment. *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on*. Vol. 2. pp.1682 - 1687.24-28 April 2000.

- [Saito 03] Saito H. Baba S. and Kanade T. Appearance-Based Virtual View Generation From Multicamera Videos Captured in the 3-D Room. *IEEE Transactions On Multimedia*. Vol. 5. Nº. 3. pp.298-303 September 2003.
- [Salvi 97] Salvi J. An approach to coded structured light to obtain three dimensional information. *Ph.D. Thesis Universitat de Girona. Departament d'Electrònica. Informàtica i Automàtica*. 1997.
- [Salvi 02] Salvi J Armangué X. Batle J. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*. Vol. 35. pp. 1617-1635. 2002.
- [Santiso 03] Santiso E. Posicionamiento Absoluto de un Robot Móvil a partir del Reconocimiento de Marcas. *Tesis Doctoral. Escuela Politécnica. Universidad de Alcalá*. 2003.
- [Santos 01] Santos-Victor J. Visual servoing of celular robots. *VisLab-TR. European Control Conference. ECC2001*. Porto (Portugal). September. 2001.
- [Shafer 98] Shafer S. Krumm J. Brumitt B. Meyers B. Czerwinski M. and Robbins D. The new easyliving project at microsoft research. *In Proceedings of Joint DARPA/NIST Smart Spaces Workshop*, Gaithersburgh, MD, July 1998.
- [Skrzypczynski98] Skrzypczynsk P. Localisation of a mobile robot based on natural landmarks. *Intelligent Autonomous Vehicles (IAV98). Proceedings from the 3rd IFAC Symposium*. Vol.1. pp 171-176. Madrid (Spain) 1998.
- [Sogo 99] Sogo T. Ishiguro H. Ishida T. Nakashima H. and Zhang C. Mobile robot navigation by distributed vision agents Approaches to Intelligent Agents, *Lecture Notes in Artificial Intelligence*, Vol. 1733. pp. 96-110. Springer-Verlag, Berlin 1999.
- [Steinhaus 04] Steinhaus P. Walther M. Giesler B. Dillmann R. 3D global and mobile sensor data fusion for mobile platform navigation. *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*. Vol. 4. pp.3325 - 3330. April, May 2004.
- [Stone 98] Stone J.M. Powell J.D. Precise positioning with GPS near obstructions by augmentation with pseudolites. *Position Location and Navigation Symposium, IEEE 1998*. Vol.2. pp:562 - 569. 20-23 April 1998.

- [Sukthankar 04] Sukthankar G. Mandel M. Sycara.K. Hodgins J. Modeling Physical Capabilities of Humanoid Agents Using Motion Capture Data. *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS)*. Vol. 1. N° 1. pp.344-351. July 2004.
- [Svoboda 03] Svoboda T. Quick guide to multi-camera self-calibration. *Technical Report 263, Computer Vision Lab, Swiss Federal Institute of Technology, Zurich*, July 2003.
- [Svoboda 02] Svoboda T. Hug H. and Van Gool L. ViRoom -- low cost synchronized multicamera system and its self-calibration. *In Pattern Recognition, 24th DAGM Symposium, number 2449 in LNCS*. Vol.1 pp.515-522. Springer, September 2002.
- [Taeseok 03] Taeseok J., Soomin P., JangMyung L. A study on position determination for mobile robot navigation in an indoor environment. *Computational Intelligence in Robotics and Automation, 2003. Proceedings. 2003 IEEE International Symposium on*. Vol. 2. pp.636 - 640. 16-20 July 2003.
- [Theobalt 02] Theobalt C. Magnor M. Schueler P. Seidel H.P. Multi-Layer Skeleton Fitting for Online Human Motion Capture . *Proceedings of 7th International Fall Workshop on Vision, Modeling and Visualization (VMV 2002)*. Vol.1. pp.471-478. Erlangen, Germany 2002.
- [Toscani 87] ToscaniG. Faugeras O. D. Camera Calibration for 3D computer vision. *Proc. International Workkshop on Industrial Applications of Machine Vision and Machine Intelligence.1987*. Vol.1. pp. 240-247. Silken, Japan 1987.
- [Toscani 86] Toscani G. Faugeras O. D. The calibration problem for stereo. *Proceedings of the IEEE Computer Vision and Patern Recognition*. Vol.1. pp15-20. 1986.
- [Tsai 87] Tsai T. A camera calibration technique for haigh-accuracy 3D machine vision metrology ussing off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation.RA-3*. Vol.1. pp323-344. 1987.
- [Ueshiba 03] Ueshiba T. Tomita F. Plane-based calibration algorithm for multi-camera systems via factorization of homography matrices *Computer Vision, 2003 Proceedings. Ninth IEEE International Conference on*. Vol.2. pp:966 - 973.13-16 October 2003.

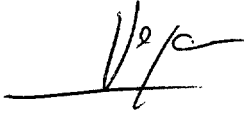
- [Valin 03] Valin J.M. Michaud F. Rouat J. Letourneau D. Robust sound source localization using a microphone array on a mobile robot. *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on* Vol. 2. pp.1228 - 1233. 27-31 October. 2003.
- [Villadangos 04] Villadangos J.M. Ureña J. Mazo M. Hernández A. Martín M. Marziani C.M. Global positioning of mobile robots by codified ultrasonic beacons. *Conferencia Internacional en Telecomunicación, Electrónica y Control. TELEC-04.* pp 47. Santiago de Cuba (Cuba). 2004.
- [Wang 01] Wang J. Integration of GNSS and Pseudo-Satellites: New Concepts for Precise Positioning, *IAG Scientific Assembly*, 2-8 September 2001, Budapest, Hungary.
- [Wang 88] Wang C.M. Location estimation and uncertainty analysis for mobile robots. *Proceedings of the IEEE International Conference On Robotics and Automation.* Vol.2. pp.1230-1235.1988
- [Want 88] R. Want. B. Schilit. N. Adams. R. Gold. K. Petersen. D. Goldberg. J. Ellis. M. Weiser. The ParcTab Ubiquitous Computing Experiment. *Xerox Parc technical report.* 1988
- [Want 92] R. Want. A. Hopper. V. Falcao. J. Gibbons The active badge location system. *ACM Transactions on Information Systems.* Vol:1. pp:91-102, January 1992.
- [Weckesser 95] Weckesser P. Dillmann R. Elbs M. Hampel S. Multiple sensor processing for high-precision navigation and environmental modeling with a mobile robot. *Intelligent Robots and Systems 95. 'Human Robot Interaction and Cooperative Robots', Proceedings. 1995 IEEE/RSJ International Conference on.* Vol.1. pp.453 - 458. 5-9 August 1995 .
- [Weiser 91] M. Weiser. The computer of the 21st century. *Scientific American*, Vol 3. pp.66-75. September 1991.
- [Welch 01] Welch G. Bishop G. An Introduction to the Kalman Filter. *SIGGRAPH 2001. Department of Computer Science, University of North Carolina at Chapel Hill.* 2001.
- [Weng 92] Weng J. Cohen P. Herniou M. Camera calibration with distortion models and accuracy evaluatio. *IEEE Transactions Pattern Analysis and Machine Intelligence.* Vol.14. pp 965-980. 1992.

- [Wiles 99] Wiles C. Davison A. Calibrating a multi-camera system for 3D modelling. *Multi-View Modeling and Analysis of Visual Scenes, 1999. (MVIEW '99) Proceedings. IEEE Workshop*. Vol.1.pp:29 - 36. 26 June 1999.
- [Xu 96] Xu G. and Zhang Z. Epipolar Geometry in Stereo, Motion and Object Recognition. A Unified Approach. *Kluwer Academic Publishers*, pp 313. ISBN 0-7923-4199-6.1996.
- [Youssef 02] Youssef M. Agrawala A. Shankar A. U and Noh S. H: A Probabilistic Clustering-Based Indoor Location Determination System. *Tech. Report, University of Maryland at College Park*, CS-TR 4350, March 2002
- [Yuen 02] David C.K. Yuen and Bruce A. MacDonald. Natural landmark based localisation system using panoramic images. *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'02)*. Vol. 1.pp.915-20. Washington DC May 2002.
- [Yun 04] Jae-mu Yun, Jin-woo Park, Ho-seek Choi, Jang-myung Lee. Absolute Positioning System for Mobile Robot Navigation in an Indoor Environment. *ICCAS2004*. August 25-27 2004.
- [Zhang 99] Zhang Z. Flexible Camera Calibration By Viewing a Plane From Unknown Orientations. *International Conference on Computer Vision (ICCV'99)*. pp. 666-673. Corfu, Greece. September 1999.
- [Zimmerman 97] Zimmerman. Experimental demonstration of an indoor GPS-based sensing system for robotic applications. *Navigation*. Vol. 43.Nº 4. pp 375-395. winter 1996-1997.

UNIVERSIDAD DE ALCALÁ  
SERVICIO DE POSTGRADO

DILIGENCIA PARA HACER CONSTAR QUE EL  
PRESENTE EJEMPLAR DE LA TESIS PRESENTADA  
POR D. Ignacio Fernández Lorenz  
CONSTA DE 274 PAGINAS Y HA SIDO ENTREGADA  
CON FECHA 27 de julio de 2005  
A EFECTOS DE DEPOSITO DE TESIS.

EL FUNCIONARIO.



Reunido el Tribunal que suscribe en  
el día de la fecha acordó otorgar  
a la presente Tesis Doctoral la  
calificación de  sobresaliente  
 cum laude (por unanimidad)

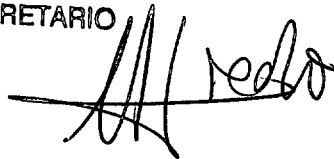
Alcalá de Henares, 03 de NOVIEMBRE de 2005

EL PRESIDENTE



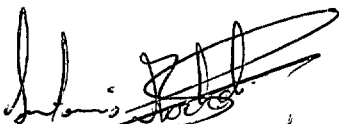
Fdo.: JESÚS UREÑA UREÑA

EL SECRETARIO



Fdo.: ALFREDO GARDEL VICENTE

EL VOCAL



Fdo.: ANTONIO NOCKLI SALCEDO

EL VOCAL



Fdo.: RAMÓN CENEA QUIZ

EL VOCAL



Fdo.: LEOPOLDO CALDERÓN ESTEVÉZ