

5



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

FACULTAD DE INGENIERÍA

**"MANIPULACIÓN ESTABLE DE OBJETOS RÍGIDOS  
UTILIZANDO DOS ROBOTS INDUSTRIALES  
Y UN SISTEMA DE VISIÓN"**

T E S I S

QUE PARA OBTENER EL TÍTULO DE  
INGENIERA ELÉCTRICA Y ELECTRÓNICA  
PRESENTA:

CARINA MONIQUE BENTRUP VALLADARES

DIRECTOR: DR. MARCO ANTONIO ARTEAGA PÉREZ



CIUDAD UNIVERSITARIA

SEPTIEMBRE 2002

Autorizo a la Dirección General de Bibliotecas de la  
UNAM a difundir en formato electrónico e impreso el  
contenido de mi trabajo recepcional.

NOMBRE: Bentrup Valladares

Carina Monique

FECHA: 28/8/2002

FIRMA: Bentrup

TESIS CON  
FALLA DE ORIGEN



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

PAGINACIÓN  
DISCONTINUA

**DEDICO MI TESIS Y MI CARRERA**

**A MIS QUERIDOS PADRES**

**a mi mamá**  
**POR SU CARÍÑO**  
**Y APOYO INCONDICIONAL**

**a mi papá**  
**POR ESTAR JUNTO A MÍ**  
**A CADA INSTANTE**

## Agradecimiento

ESTE TRABAJO ES UNO DE MIS TANTOS SUEÑOS HECHO REALIDAD Y QUIERO DIRIGIR UN ESPECIAL AGRADECIMIENTO AL Sr. Wolfgang Ehrensberger.

EXPRESO MI MAYOR GRATITUD AL Dr. Marco Antonio Arteaga Pérez DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO DE LA FACULTAD DE INGENIERÍA, POR HABER SIDO UN EXCELENTE PROFESOR DE CONTROL Y MI ASESOR EN ESTA TESIS. GRACIAS DR. ARTEAGA POR TODAS SUS ENSEÑANZAS Y GUÍA, POR SU INTERÉS EN ESTE TRABAJO Y POR HABERME AYUDADO SIEMPRE (EN LAS BUENAS Y EN LAS MALAS).

UN SINCERO RECONOCIMIENTO PARA EL Dr. Arturo Espinosa-Romero DEL INSTITUTO DE INVESTIGACIONES MATEMÁTICAS APLICADAS Y DE SISTEMAS, POR HABERME ASESORADO EN TODO LO QUE RESPECTA A LA VISIÓN COMPUTACIONAL.

QUIERO MENCIONAR DE FORMA MUY ESPECIAL AL Dr. Francisco García Ugalde, MI TUTOR Y PROFESOR; POR LA VALIOSA AYUDA Y LOS CONSEJOS QUE RECIBÍ DE ÉL A LO LARGO DE TODA MI CARRERA.

A TODOS LOS **catedráticos de la Facultad de Ingeniería** Y A MIS MAESTROS DEL Colegio Alemán Alexander von Humboldt LES AGRADEZCO SUS ENSEÑANZAS, YA QUE CON SUS ASIGNATURAS ME HAN DADO HERRAMIENTAS PARA ESTE TRABAJO Y PARA LA VIDA PROFESIONAL.

QUIERO NOMBRAR PARTICULARMENTE A LOS PROFESORES PROTAGONISTAS DE MOMENTOS IMPORTANTES DE MI CARRERA, QUIENES, ADEMÁS DE SER NOTABLES MAESTROS, SON UNA PARTE MUY VALIOSA DE MI FORMACIÓN COMO INGENIERA Y COMO PERSONA:

- Ing. Roberto Macías Pérez, POR SU DEDICACIÓN, CORAJE Y DISCIPLINA.  
-ELECTRÓNICA DE POTENCIA
- M. en I. Larry H. Escobar Salguero, POR SU FORMA DE APRECIAR LA CIENCIA.  
-PROCESAMIENTO DIGITAL DE SEÑALES Y TEORÍA ELECTROMAGNÉTICA
- Ing. Alfredo López Tagle, POR HABER SIDO LA PRINCIPAL FUENTE PARA ENTENDER Y

ADQUIRIR UN GUSTO ESPECIAL POR LA POTENCIA.

-TRANSFORMADORES Y MOTORES DE INDUCCIÓN Y MÁQUINAS SÍNCRONAS Y DE CD

- Ing. José de Jesús González Franco, POR SUS ENTRETENIDAS CLASES CON LAS QUE PUDE ENTENDER MÁS DE LA NATURALEZA HUMANA.  
- TEMAS SELECTOS DE ÉTICA APLICADA
- Ing. Rodolfo Peters Lammel, POR SU FORMA DE VER LA VIDA.  
- MEDICIÓN E INSTRUMENTACIÓN
- Ing. Gloria Mata Hernández, POR SER UN EXCELENTE EJEMPLO DE UNA VERDADERA INGENIERA.  
- ANÁLISIS DE CIRCUITOS ELÉCTRICOS
- Ing. Gabriel Jaramillo, POR CÓMO ME ENSEÑÓ A ESTUDIAR.  
- ELECTRICIDAD Y MAGNETISMO
- Quím. Félix Núñez, POR SU FORMA TAN PECULIAR DE DISCUTIR.  
- QUÍMICA Y TERMODINÁMICA

GRACIAS POR SER PARTE DE LO QUE HOY SOY.

ADEMÁS DE LOS ANTERIORES, APROVECHO ESTE ESPACIO PARA RECONOCER A LOS MAESTROS QUE ME IMPARTIERON EXCELENTE CÁTEDRAS:

- Dr. Pablo Mulás del Pozo  
- ENERGÍA E IMPACTO AMBIENTAL
- Ing. Octavio Estrada Castillo  
- CALIDAD
- Ing. Roberto Mandujano Wild  
- DISEÑO DIGITAL
- Ing. Juan Ocariz Castelazo  
- ESTÁTICA
- Lic. Ma. Fernanda Azuela  
- TEMAS SELECTOS DE FILOSOFÍA, CIENCIA Y TECNOLOGÍA
- Ing. David Vázquez Ortiz  
- SISTEMAS ELÉCTRICOS DE POTENCIA I
- Ing. Erick Castañeda de Isla Puga  
- ECUACIONES DIFERENCIALES Y MATEMÁTICAS AVANZADAS
- Lic. Ma. Teresa Yebra García  
- COSTOS Y EVALUACIÓN DE PROYECTOS

- Ing. Yukihiko Minami Koyama  
- GEOMETRÍA DESCRIPTIVA

A TODOS LOS QUE FUERON Y HAN SIDO MIS amigos DURANTE ESTOS CINCO AÑOS DE LA CARRERA (DENTRO Y FUERA DE LA UNIVERSIDAD): GRACIAS POR LOS MOMENTOS QUE DISFRUTAMOS JUNTOS. LOS VIAJES, LOS CONCURSOS, LAS DISCUSIONES, LAS REFLEXIONES, LOS E-MAILS Y LOS CAFÉS CON USTEDES HICIERON DE ESTE TIEMPO ALGO INOLVIDABLE.

FINALMENTE AGRADEZCO EL APOYO DE:

- Fundación TELMEX
- PROBETEL

RECONOZCO AL Programa de Alto Rendimiento Académico TODO LO QUE ME DIO CULTURAL Y ACADÉMICAMENTE. GRACIAS POR HABER SIDO UNO DE LOS MOTIVOS MÁS GRANDES PARA ESTUDIAR EN LA FACULTAD DE INGENIERÍA DE LA UNAM.

APROVECHO TAMBIÉN PARA AGRADECERLE AL Dr. Jesús Savage Carmona LAS TUTORÍAS Y SU INTERÉS EN EL TEMA QUE AQUÍ SE PRESENTA.

GRACIAS Universidad Nacional Autónoma de México POR TODO LO QUE HE APRENDIDO EN TUS AULAS; POR LOS CONCIERTOS, LAS EXPOSICIONES, LAS CONFERENCIAS; POR TU GENTE Y POR LO QUE SIGNIFICAS PARA MÍ, PARA EL PAÍS Y PARA AMÉRICA.

*Carina Monique Bentrup V.  
Septiembre 2002*

# Índice General

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introducción</b>   | <b>1</b>  |
| <b>2</b> | <b>Antecedentes</b>   | <b>5</b>  |
| 2.1      | Visión humana vs. visión computacional . . . . .                      | 5         |
| 2.2      | Estructura del ojo humano y formación de imágenes . . . . .           | 6         |
| 2.3      | Cámaras y sus características . . . . .                               | 8         |
| 2.4      | Percepción humana de la profundidad . . . . .                         | 10        |
| 2.5      | Procedimiento sencillo para fotografías en tres dimensiones . . . . . | 11        |
| 2.6      | Matrices de transformación . . . . .                                  | 13        |
| <b>3</b> | <b>Procesamiento y análisis de imágenes digitales</b>                 | <b>15</b> |
| 3.1      | Elementos del sistema . . . . .                                       | 15        |
| 3.2      | Digitalización . . . . .  | 16        |
| 3.3      | Filtrado . . . . .  | 17        |
| 3.4      | Segmentación y extracción de características . . . . .                | 19        |
| 3.4.1    | Detección de bordes, líneas y regiones . . . . .                      | 20        |
| 3.5      | Modelo geométrico de una cámara . . . . .                             | 23        |
| 3.5.1    | Trasformaciones Proyectivas . . . . .                                 | 25        |
| 3.6      | Visión Tridimensional . . . . .                                       | 27        |

|          |   |           |
|----------|---|-----------|
| 3.6.1    | Localización de un punto en el espacio . . . . .                        | 27        |
| 3.7      | Calibración de las cámaras . . . . .                                    | 29        |
| 3.7.1    | Eliminación de la distorsión . . . . .                                  | 30        |
| 3.7.2    | Estimación de la longitud focal . . . . .                               | 35        |
| 3.7.3    | Determinación de la posición y orientación . . . . .                    | 36        |
| <b>4</b> | <b>Agarre estable de un objeto . . . . .</b>                            | <b>39</b> |
| 4.1      | Cálculo del agarre estable . . . . .                                    | 39        |
| 4.2      | Cálculo de los centros de masa . . . . .                                | 43        |
| 4.3      | Sistema de dos robots industriales . . . . .                            | 44        |
| 4.4      | Cinemática de los robots . . . . .                                      | 46        |
| 4.4.1    | Cinemática Directa . . . . .  | 46        |
| 4.4.2    | Cinemática Inversa . . . . .  | 51        |
| 4.5      | Trabajo simultáneo . . . . .  | 60        |
| <b>5</b> | <b>Sujeción de un objeto con dos robots industriales . . . . .</b>      | <b>65</b> |
| 5.1      | Control PID de los robots industriales . . . . .                        | 65        |
| 5.2      | Diseño del sistema de control . . . . .                                 | 67        |
| 5.3      | Reconocimiento de un objeto por medio de un sistema de visión . . . . . | 69        |
| 5.3.1    | Coincidencia de patrones . . . . .                                      | 69        |
| 5.3.2    | Espacio de características . . . . .                                    | 70        |
| 5.4      | Calibración intrínseca . . . . .  | 71        |
| 5.5      | Programa Visual.prj . . . . .   | 72        |
| 5.5.1    | Diagrama del programa . . . . .   | 73        |
| 5.5.2    | Ventana inicial . . . . .   | 73        |
| 5.5.3    | Librerías del código fuente . . . . .                                   | 75        |

|          |   |            |
|----------|---|------------|
| 5.5.4    | Variables globales . . . . .  | 76         |
| 5.5.5    | Botones y funciones . . . . .   | 76         |
| 5.5.6    | Resultados experimentales . . . . .   | 85         |
| <b>6</b> | <b>Conclusiones y trabajo a futuro</b>  | <b>91</b>  |
| <b>A</b> | <b>Características principales de los manipuladores del Laboratorio de Robótica de la División de Posgrado de la Facultad de Ingeniería</b> | <b>95</b>  |
| <b>B</b> | <b>Características principales de las Cámaras CCD</b>   | <b>99</b>  |
| <b>C</b> | <b>Características principales de las tarjetas PCI-FlexMotion-6C</b>  | <b>103</b> |
| <b>D</b> | <b>Características principales de la tarjeta PCI-IMAQ-1408</b>  | <b>105</b> |
| <b>E</b> | <b>Librerías de programa Visual.prj</b>   | <b>107</b> |



# Índice de Figuras

|      |   |    |
|------|---|----|
| 1.1  | Telecirugía transatlántica . . . . .  | 2  |
| 2.1  | Partes del ojo humano . . . . .   | 6  |
| 2.2  | Comparación del ojo y una cámara . . . . .                                      | 8  |
| 2.3  | Sensibilidad luminosa ocular . . . . .  | 9  |
| 2.4  | Campo Visual del ojo humano . . . . .   | 10 |
| 2.5  | Convergencia . . . . .  | 11 |
| 2.6  | Ejemplo del procedimiento para una fotografía tridimensional sencilla . . . . . | 12 |
| 2.7  | Para una buena fotografía estereoscópica . . . . .                              | 12 |
| 3.1  | Imagen original FI a 256 colores . . . . .                                      | 17 |
| 3.2  | Arreglo de cuantización . . . . .   | 18 |
| 3.3  | Tipos de función de borde . . . . .   | 20 |
| 3.4  | Dirección del borde y del gradiente . . . . .                                   | 21 |
| 3.5  | Imagen original de prueba . . . . .   | 22 |
| 3.6  | Imagen después del detector de bordes de Canny . . . . .                        | 22 |
| 3.7  | Imagen comparativa . . . . .  | 23 |
| 3.8  | Modelo básico de la proyección . . . . .  | 24 |
| 3.9  | Relaciones de triángulos . . . . .  | 25 |
| 3.10 | Localización de las cámaras . . . . .   | 28 |

|   |    |
|---|----|
| 3.11 Punto en 3D . . . . .  | 28 |
| 3.12 Rejilla original para calibración . . . . .                            | 31 |
| 3.13 Imagen de la rejilla con ambas cámaras . . . . .                       | 31 |
| 3.14 Imagen de la rejilla después de pasar por el algoritmo Canny . . . . . | 32 |
| 3.15 Parámetros de la imagen ideal . . . . .                                | 33 |
| 3.16 Punto vecino más próximo . . . . .                                     | 34 |
| 3.17 Imágenes sin distorsión . . . . .                                      | 34 |
| 3.18 Imágenes del marco de calibración . . . . .                            | 35 |
| 3.19 Caja de calibración . . . . .  | 36 |
| 3.20 Calibración extrínseca . . . . .                                       | 37 |
| 4.1 Ejes coordenados de los contactos . . . . .                             | 40 |
| 4.2 Interpretación geométrica de los modelos de fricción . . . . .          | 41 |
| 4.3 Elemento diferencial $dm$ de un cuerpo . . . . .                        | 44 |
| 4.4 Robots A465 y A255 del Laboratorio . . . . .                            | 45 |
| 4.5 Parámetros de la asignación Denavit-Hartenberg . . . . .                | 47 |
| 4.6 Marcos de referencia del robot A465 . . . . .                           | 48 |
| 4.7 Marcos de referencia del robot A255 . . . . .                           | 50 |
| 4.8 Vector $p_c$ . . . . .  | 52 |
| 4.9 Proyección de $p_c$ . . . . .   | 53 |
| 4.10 Ángulo $\theta_3$ . . . . .  | 54 |
| 4.11 Ángulos reales del robot . . . . .                                     | 55 |
| 4.12 Proyección de $p_c$ . . . . .  | 58 |
| 4.13 Ángulos $\theta_2$ y $\theta_3$ . . . . .                              | 58 |
| 4.14 Ángulo $\theta_3$ y $z_m$ . . . . .                                    | 59 |

|  |     |
|--|-----|
| 5.21 Esquema Experimento 1 . . . . .     | 86  |
| 5.22 Calibración Experimento 1 . . . . . | 87  |
| 5.23 Caja 1 . . . . .                    | 88  |
| 5.24 Experimento 2 . . . . .             | 88  |
| 5.25 Experimento 3 . . . . .             | 89  |
| A.1 Robot A255 . . . . .                 | 95  |
| A.2 Robot A465 . . . . .                 | 97  |
| B.1 Cámara CCD . . . . .                 | 99  |
| B.2 Sistema sensorial CCD . . . . .      | 100 |

|      |  |    |
|------|--|----|
| 4.15 | Relación de marcos de referencia . . . . .                                 | 60 |
| 4.16 | Objeto del agarre desde dos perspectivas . . . . .                         | 61 |
| 4.17 | Objeto modelo . . . . .  | 63 |
| 4.18 | Círculo-Trayectoria de radio $r$ . . . . .                                 | 63 |
| 5.1  | Estructura básica de un sistema de control con retroalimentación . . . . . | 66 |
| 5.2  | Sistema de control del laboratorio . . . . .                               | 67 |
| 5.3  | Controlador digital PID . . . . .  | 68 |
| 5.4  | Base para las cámaras . . . . .  | 72 |
| 5.5  | Nivel de tipo gota . . . . .   | 72 |
| 5.6  | Ambiente de trabajo . . . . .  | 73 |
| 5.7  | Diagrama funcional del programa . . . . .                                  | 74 |
| 5.8  | Ventana de inicio . . . . .  | 75 |
| 5.9  | Puntos de calibración de la cámara 1 . . . . .                             | 77 |
| 5.10 | Puntos de calibración de la cámara 2 . . . . .                             | 78 |
| 5.11 | Imágenes del objeto en la pantalla principal . . . . .                     | 80 |
| 5.12 | Marcas (manchas) reconocidas . . . . .                                     | 81 |
| 5.13 | Probables esquinas . . . . .   | 82 |
| 5.14 | Enumeración de esquinas según la posición de la caja . . . . .             | 82 |
| 5.15 | Caso imposible . . . . .   | 82 |
| 5.16 | Coordenadas espaciales desde el robot A465 . . . . .                       | 83 |
| 5.17 | Pantalla del programa rutina.prj . . . . .                                 | 84 |
| 5.18 | Robots en posición origen . . . . .  | 84 |
| 5.19 | Posición deseada . . . . .   | 85 |
| 5.20 | Caja de calibración . . . . .  | 86 |

# Capítulo 1

## Introducción

No es bueno dejarse arrastrar por los sueños  
y olvidarse de vivir.  
– Harry Potter, J. K. Rowling

Es mejor vivir de tal forma  
que esos sueños se hagan realidad.

El término robot apareció en 1917 cuando el checo Karel Capek lo introdujo en una de sus obras teatrales, llamada *Robots Universales de Rossum*. *Robota* significa servidumbre, trabajador forzado o trabajo en checo (Groover *et al.* 1990: pp. 7). En esta obra Karel describe a los robots como máquinas que aligeran la vida de los hombres, pero que finalmente se rebelan y aniquilan a la raza humana. El término fue adoptado por muchos escritores y así nacieron los robots con forma humana, inteligencia y personalidad. Sin embargo, la tecnología no ha podido, hasta hoy, llegar a las grandes creaciones de la ciencia ficción (un robot con sentimientos, por ejemplo).

El RIA (Instituto Americano de Robots) define *robot industrial* como (Groover *et al.* 1990: pp. 6):

manipulador multifuncional reprogramable diseñado para mover material por medio de movimientos programados de gran variedad.

Aunque desde mediados del siglo XVIII se construyeron máquinas que realizaban algunas tareas de forma autónoma, el estudio formal de la robótica comenzó después de la Segunda Guerra Mundial. Los llamados teleoperadores realizaban manipulaciones remotas para manejar materiales radioactivos. Simultáneamente aparecieron las computadoras de control numérico. El primer robot instalado, en 1961, era ya reprogramable y se le podían cambiar las herramientas para que realizara funciones diversas dentro de una planta automotriz. Esta fue una de las primeras ramas de la industria que aceptó la utilidad de estas máquinas.

En 1963 apareció por primera vez un robot con visión binaria que le permitía responder a obstáculos en su entorno y en 1969 uno que podía caminar. No fue hasta 1973 que se diseñó un lenguaje de



programación de alto nivel (WAVE) ligado a estas máquinas (Groover *et al.* 1990: pp. 11). A finales de los 70's ya existían manipuladores famosos como el PUMA (Máquina Universal Programable para Ensamblado) y el SCARA (Robot Articulado para Ensamblado Selectivo, por sus siglas en inglés).

En los 80's se hicieron adelantos en el control de robots con controladores de lazo cerrado, además de cambios en la mecánica de las máquinas. En esta misma época se comenzó a estudiar la dinámica de robots que actúan conjuntamente sujetando algún objeto, lo que dio paso al estudio de manos mecánicas. Éstas tienen muchas ventajas sobre las pinzas convencionales: pueden llevar a cabo una manipulación diestra de objetos y, en ciertos casos, se puede llegar a pensar en utilizarlas como prótesis para personas que hayan perdido alguna de sus manos.

Durante los años 90's se intentó introducir robots a otras industrias y a otros ambientes. Aparecieron así manipuladores en la industria electrónica y en otros planetas, como Marte. Hoy en día se trabaja con teleoperadores en muchas áreas como la medicina, la aeronáutica y la astronomía, entre otras, pero se estima que éstos se convertirán en robots autónomos en el mañana.

El avance de la tecnología ha permitido también el uso de sistemas de visión integrados a los robots. Las cámaras que componen el sistema visual (que pueden ser una o más) se utilizan para detectar o evadir objetos. Un ejemplo notable del uso de robots trabajando coordinadamente con un sistema de visión son las cirugías trasatlánticas. En septiembre del año 2001, doctores en Estados Unidos hicieron una operación a un paciente en Francia, por medio de un robot y cámaras de video (IEEE 2001: pp. 9).

Para llevar a cabo una operación médica remota se necesita mucha precisión, tanto en los instrumentos que la efectúan como en las imágenes que el médico utiliza como referencia de la realidad (realidad virtual), del otro lado del océano (Figura 1.1) (<http://news.bbc.co.uk>).

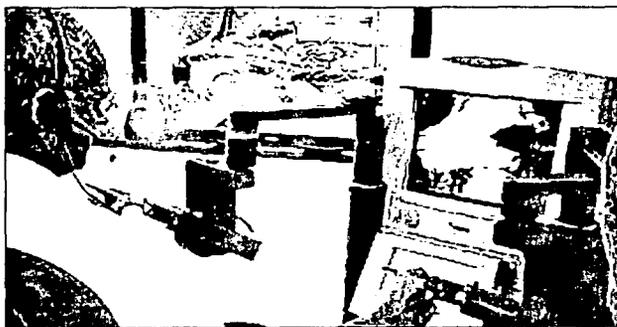


Figura 1.1: Telecirugía transatlántica

El término realidad virtual se refiere a la creación remota de un ambiente artificial que contiene imágenes, sonidos y fuerzas, para manipular objetos de forma remota (telepresencia). El propósito de simular un ambiente remoto es el de permitir a una persona interactuar con él. Una de las partes más importantes dentro de la realidad virtual es lo que percibe el sentido de la vista, siendo la visión computacional la que tiene como finalidad la reproducción artificial de este sentido, transformando la energía luminosa que reflejan o emiten los objetos y las escenas del mundo tridimensional en que vivimos a imágenes que

TESIS CON  
FALLA DE ORIGEN

los representen de la forma más real posible.

La visión computacional o entendimiento de imágenes es el estudio de cómo construir máquinas inteligentes que puedan reconocer su entorno por medio de la visión; su estudio empezó hace más de veinte años (Kanatani 1995: pp. 1). Es un área que ha evolucionado, pero las técnicas iniciales como las de transformación, segmentación y caracterización de imágenes siguen siendo válidas y usadas en estudios actuales. En un principio se trabajó sólo con imágenes bidimensionales (Bryant y Bryant 1989), pero el campo de estudios fue ampliándose tiempo después al espacio tridimensional (Xiao-Wei y Dubuisson 1990).

Uno de los objetivos más importantes es el de identificar y clasificar objetos, generalmente con procedimientos estadísticos. Desde la aparición del estudio de imágenes se pensaba que el reconocimiento podría conducir a sistemas de visión inteligentes. Actualmente ya existen algunos sistemas de este tipo. Con imágenes bidimensionales y transformaciones adecuadas es posible llegar a un reconocimiento tridimensional. El problema es que los objetos tridimensionales se ven de formas muy distintas según el punto desde donde se observan. Por ello, para conocer un objeto por medio de imágenes bidimensionales es necesario tener conocimiento *a priori* de la escena. De este modo se han empezado a utilizar como herramientas leyes ópticas y físicas, aunque en algunos casos siguen siendo necesarias ciertas características en iluminación, distorsión y reflejos. El esquema de trabajo general consiste en:

1. Extraer algunas características por medio de detección de bordes y segmentación
2. Aproximar las formas y la orientación
3. Empalmar el resultado con modelos tridimensionales

Aunque cada caso tiene sus variantes. El problema de este esquema general es que es muy sensible al ruido, el cual es muy difícil de aislar de forma totalmente eficaz cuando se implementan los algoritmos computacionales. Actualmente se lleva a cabo investigación con cámaras con movimiento controlado o con varias cámaras fijas (el caso que se estudia en el presente trabajo). La visión computacional tiene aplicaciones en áreas como compresión de imágenes, medicina y robótica, entre otras. En el área de robótica se utiliza para la inspección industrial, el mapeo de superficies y como guía para manipuladores y vehículos. La visión computacional se convierte así en los ojos de un robot.

El propósito de esta tesis es crear un sistema de visión confiable, que permita localizar objetos en el espacio y reconocerlos como tales, con el fin de dar los primeros pasos hacia un sistema visual inteligente. La información debe ser analizada y procesada para permitir el agarre estable del objeto por medio de dos robots. Una vez sujetado el objeto se debe realizar alguna manipulación diestra con él. El sistema de visión está formado por dos cámaras de video blanco y negro, mientras que el sistema de robots se compone de dos manipuladores industriales.

El presente trabajo está estructurado en dos secciones principales: la primera parte se refiere a los sistemas de visión y la segunda se centra en los robots industriales.

El Capítulo 2 trata de cómo se conforma un sistema visual e incluye la comparación entre el sistema de visión del ser humano y el que se utiliza para una máquina. Además, se exponen los motivos del uso de dos cámaras en vez de sólo una. El Capítulo 3 contiene los detalles para todo lo que tiene que ver con el procesamiento y análisis de las imágenes adquiridas. Se explican los elementos básicos del procesamiento digital de imágenes (adquisición, almacenamiento y filtrado, entre otros), los detalles de la segmentación y la localización de puntos específicos en ellas. Este capítulo contiene también los detalles de la calibración de las cámaras que constituyen el sistema experimental.

El Capítulo 4 explica la teoría y definición del agarre estable, la cinemática directa e inversa de cada uno de los manipuladores y cómo se trabaja con un sistema que incluye a dos robots industriales en una tarea conjunta. El Capítulo 5 engloba la sujeción del objeto con ambos robots; ahí se presenta un controlador PID, la forma en que se hace la detección y manipulación de los objetos y los resultados experimentales. Finalmente, en el Capítulo 6 se exponen las conclusiones de este trabajo, así como las mejoras y extensiones que se derivan del mismo.

# Capítulo 2

## Antecedentes

Una máquina es realmente útil  
hasta que es independiente del conocimiento  
que llevó a su creación.  
-Die Physiker, Dürrenmatt

### 2.1 Visión humana vs. visión computacional

El sentido de la vista permite al ser humano percibir la luz, formas y colores de objetos y escenas de su entorno. Se considera a este sentido como el más elevado y perfecto, ya que por medio de él nos damos cuenta de la existencia de objetos que no se hallan en inmediato contacto con nosotros. Se estima que aproximadamente tres cuartas partes de la información que maneja el ser humano es visual. (Aunque un ser humano desprovisto del sentido de la vista puede desarrollar totalmente su capacidad mental y llevar una vida plena.) Los sentidos del oído y del olfato son también de gran importancia, pero en los seres humanos no se han desarrollado de una manera tan sensible como los ojos.

La visión computacional, visión de máquina o visión artificial, tiene como finalidad la reproducción ficticia del fenómeno de visión. Podría hacerse una analogía entre un sistema artificial de procesamiento de imágenes y un sistema visual humano, si consideramos a la cámara como los ojos y a la computadora como el cerebro (Maravall 1993: pp. 219).

Para poder explicar en que consiste un sistema de visión artificial es necesario tener un conocimiento básico de la anatomía y del funcionamiento del ojo, así como del sistema visual humano en general.



## 2.2 Estructura del ojo humano y formación de imágenes

El sistema ocular es de alta complejidad, con millones de conexiones nerviosas y sanguíneas, y una extensa red celular. El ojo es casi una esfera con un radio promedio de 20 mm que puede rotar en su órbita bajo el control de seis músculos extrínsecos. Cada ojo ve una imagen distinta, y el cerebro procesa la información recibida combinando ambas imágenes para tener una visión completa (Maryon-Davis 1985: pp. 16).

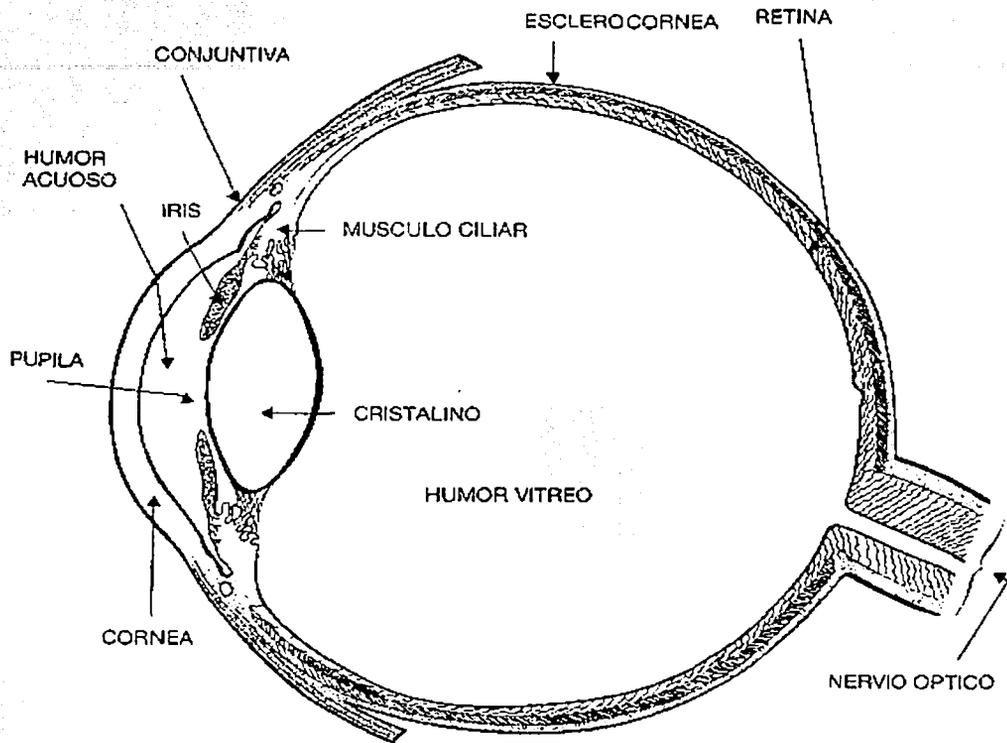


Figura 2.1: Partes del ojo humano

El ojo está compuesto de tres capas y tres cámaras. Las capas son la esclerocórnea, la úvea y la retina, mientras que las cámaras son la anterior, la posterior y la vítrea. La esclerocórnea es la capa formada por la esclerótica y la córnea. La esclerótica es la más exterior de las membranas; es la parte fibrosa y blanca del ojo. Tiene una función de protección; es resistente y está destinada a mantener la forma del globo, además de que soporta grandes presiones. En su zona exterior está recubierta por una mucosa transparente llamada conjuntiva.

La córnea es la parte transparente de la capa externa y es la *ventana óptica* del ojo. Ésta tiene una curvatura diferente del resto del globo del ojo y por ella atraviesan los rayos luminosos. La transparencia

TESIS CON  
FALLA DE ORIGEN

es una importante característica de la córnea, que no tiene vasos sanguíneos en ella (pero sí a su alrededor, que son los que la nutren). La córnea está llena de nervios, la mayoría de los cuales se dirigen hacia su superficie externa para hacerla sensible.

La función del iris es regular la cantidad de luz que entra en el interior del ojo, para lo cual varía su tamaño según la intensidad de luz. Controla la abertura de la pupila, dependiendo de la luminosidad y de la distancia de los objetos que se observan. La pupila conduce a la cámara oscura del ojo. Antes de que la luz llegue a la pupila, ésta debe atravesar un espacio lleno de humor acuoso entre la cara posterior de la córnea y el iris. Inmediatamente atrás del iris, unido a los músculos ciliares, está el cristalino: la lente del ojo. El cristalino es de un tejido transparente y elástico, con forma de lente biconvexa. Su curvatura aumenta o disminuye, y por tanto su potencia dióptrica, por medio de las contracciones de los músculos ciliares. Esto es lo que permite enfocar a distintas distancias. Atrás del cristalino está la mayor parte del globo ocular, llena de un líquido más espeso, llamado humor vítreo, que debe su nombre a que es como un cristal.

En la sección más interna del ojo está la retina, que tiene más de diez capas. La retina es la zona sensible a la luz. Es donde se deberían formar las imágenes para poder verlas con nitidez. Pero antes de llegar a la capa de las verdaderas células de la visión, la luz tiene que atravesar por lo menos ocho capas de diferente estructura, debido a que su parte anterior es ciega, y su sensibilidad va en aumento conforme se va alejando de la zona anterior. En ella se extienden las terminaciones del nervio óptico. El punto de máxima sensibilidad es una pequeña hendidura llamada fovea, que es donde se encuentra una mayor concentración de las células responsables de la sensibilidad de la retina: conos y bastones. Hay más bastones (aproximadamente unos 12 millones) que conos (alrededor de 7 millones). Éstos últimos son los que nos permiten distinguir los colores, mientras que los bastones sirven para ver con poca luz.

La percepción de los objetos en la retina se produce por la excitación de los receptores luminosos, que transforman la energía luminosa en impulsos eléctricos que se decodifican en el cerebro. Las vías ópticas constituyen la transmisión de los impulsos nerviosos desde la retina hasta la corteza cerebral a través del nervio óptico. Como las imágenes se ven como una serie de puntos discretos brillantes, el ojo debe poder discriminar las diferentes intensidades.

Los anexos del aparato visual son: el sistema óculo-motor, compuesto por seis músculos externos que provocan la movilidad del globo ocular; el sistema de protección, compuesto por párpados, vías y glándulas lagrimales, entre otros.

El proceso visual consta de cuatro etapas:

1. Formación de la imagen en la retina a través del sistema óptico (córnea, humor acuoso, cristalino y humor vítreo)
2. Transmisión del impulso nervioso a través del nervio óptico
3. Interpretación del impulso nervioso en el cerebro.

Es importante que se mencione lo relacionado con el enfoque al infinito (a partir de unos 5 metros), ya que las cámaras de este trabajo están enfocadas de esta forma. El ojo humano en el enfoque al infinito

está en reposo. Aparte de la posible contracción del iris para regular la cantidad de luz, el cristalino está en reposo. Es decir, que el ojo humano no necesita esfuerzo adicional para ver a lo lejos. Por supuesto, un ojo enfocado al infinito, si no varía algo de su sistema óptico, verá borroso a una distancia próxima. Lo que varía es el grosor del cristalino. Cuando se precisa enfocar a una distancia próxima, los músculos ciliares entran en acción y provocan un aumento de grosor del cristalino, aumentando en consecuencia su potencia (al fin y al cabo es una lente biconvexa) y consiguiendo el enfoque correcto.

Una de las diferencias más importantes entre las lentes del ojo y las lentes ópticas es que las primeras son flexibles. Su curvatura está controlada por la tensión de fibras a su alrededor, la distancia focal del centro de las lentes y la retina, varía desde 14 mm hasta 17 mm.

### 2.3 Cámaras y sus características

Las primeras cámaras aparecieron cuando se observó que, si existe un pequeño agujero en una pared de un cuarto oscuro, en el muro contrario se proyectan los objetos del mundo exterior, pero de forma invertida. Actualmente, todas las cámaras basan su funcionamiento en este fenómeno. Cuentan con un objetivo (ojo), que consta de una o varias lentes convexas que consiguen proyectar los rayos de luz que lo atraviesan en un punto llamado foco. Éste se puede alejar o acercar gracias, normalmente, a un mecanismo para obtener una imagen nítida. La distancia entre la lente y el foco es la distancia focal, mientras que la distancia entre el objetivo y la lente se le llama distancia objetivo. Si los rayos luminosos llegan desde el infinito, por ejemplo del sol, se concentran en el foco (Figura 2.2).

Se puede graduar la cantidad de luz que entra en una cámara con otro dispositivo al que se conoce como diafragma. Para cualquier toma de imágenes se conjugan estos tres elementos: el enfoque (distancia focal), el diafragma (cantidad de luz) y el obturador (velocidad de disparo). Si se compara el funcionamiento del ojo con el de una cámara, la pupila actúa de diafragma, la retina de película, la córnea de lente y el cristalino se equipara a acercar o alejar la cámara del objeto para conseguir un buen enfoque.

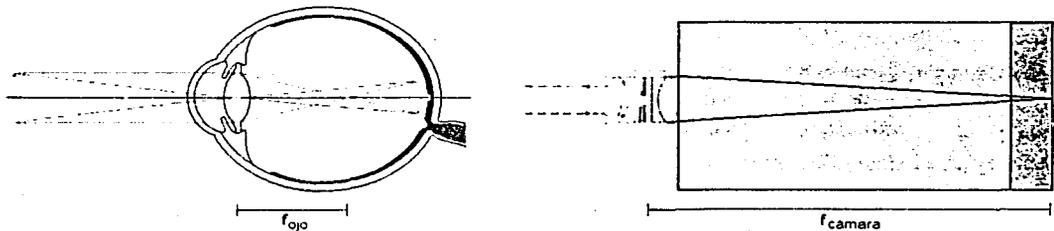


Figura 2.2: Comparación del ojo y una cámara

La cantidad de luz juega un papel importante en la visión. La unidad internacional para la intensidad luminosa, es decir la claridad de una fuente luminosa, es la candela (cd), que es proporcional a la intensidad luminosa de una vela con una flama de 4 cm. Actualmente la medición se hace con un tubo que contenga platino a 1769°C y un agujero de 1 cm<sup>2</sup>. La luz que se emite por el agujero equivale a

60 cd. La unidad de lux (lx) se utiliza en relación con la claridad de alguna superficie iluminada, y es equivalente a lo que puede iluminar una fuente de 1 candela a un área de  $1 \text{ m}^2$  que dista 1 metro de la fuente luminosa (Born *et al.* 1992: pp. 86).

En condiciones de buena iluminación (más de  $3 \text{ cd/m}^2$ ), como ocurre de día, la visión es nítida, detallada y se distinguen muy bien los colores; es la visión fotópica. Para niveles inferiores a  $0.25 \text{ cd/m}^2$  desaparece la sensación de color y la visión es más sensible a los tonos azules y a la intensidad de la luz (visión escotópica). En situaciones intermedias, la capacidad para distinguir los colores disminuye a medida que baja la cantidad de luz pasando de una gran sensibilidad hacia el amarillo a una hacia el azul (visión mesiópica).

En estas condiciones, se definen unas curvas de sensibilidad del ojo a la luz visible (Figura 2.3) para un determinado observador patrón que tiene un máximo de longitud de onda de 555 nm (amarillo verdoso) para la visión fotópica y otro de 480 nm (azul verdoso) para la visión escotópica. Toda fuente de luz que emita valores cercanos al máximo de la visión diurna (555 nm) tendrá un rendimiento energético óptimo porque producirá la máxima sensación luminosa en el ojo con el mínimo consumo de energía.

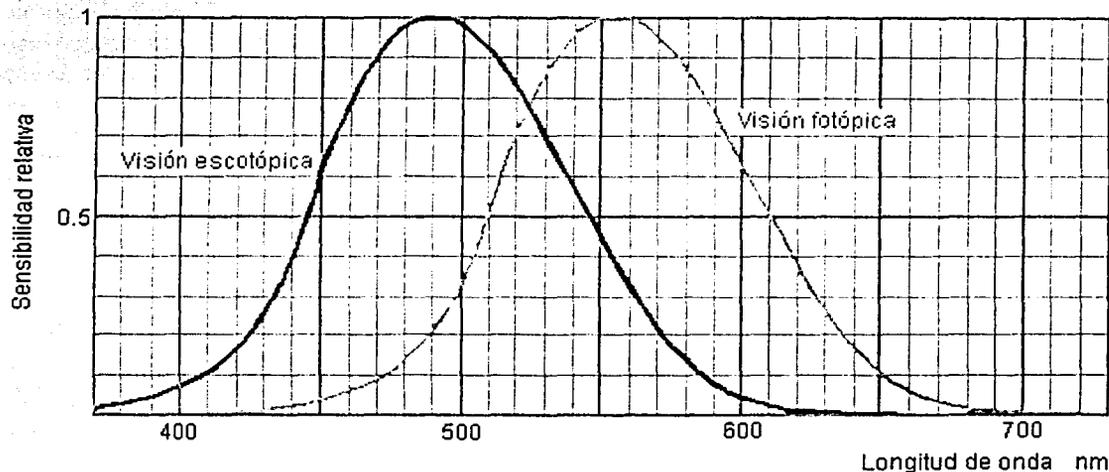


Figura 2.3: Sensibilidad luminosa ocular

La adaptación es la facultad del ojo para ajustarse automáticamente a cambios en los niveles de iluminación. Para pasar de ambientes oscuros a luminosos el proceso es muy rápido, pero el caso contrario es mucho más lento. La adaptación en una cámara varía según el mecanismo con el que cuente; en el caso de este trabajo la adaptación se hace de forma manual. La iluminación juega un papel muy importante en el momento de tomar las imágenes.

El ojo humano, al igual que la cámara, dispone de un campo visual. Cada ojo ve aproximadamente  $150^\circ$  sobre el plano horizontal, y con la superposición de ambos se abarcan los  $180^\circ$ . Sobre el plano vertical sólo son unos  $130^\circ$ ,  $60^\circ$  por encima de la horizontal y  $70^\circ$  por debajo. El campo visual de cada ojo es de tipo monocular (sin sensación de profundidad) siendo la visión en la zona de superposición de

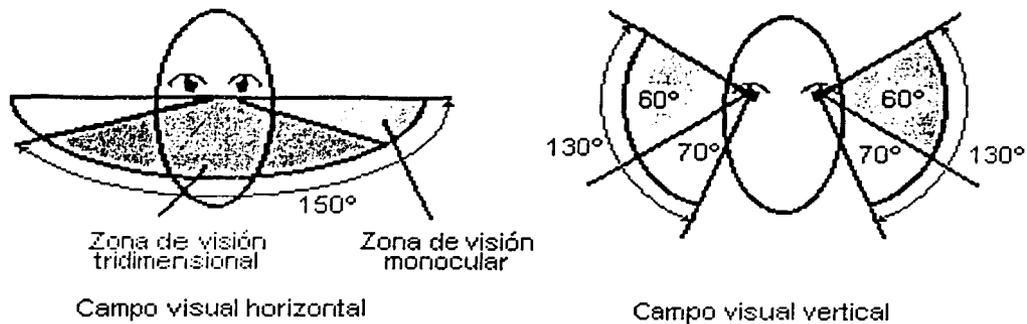


Figura 2.4: Campo Visual del ojo humano

ambos campos del tipo binocular (Figura 2.4). La percepción de la profundidad o visión tridimensional se produce en el cerebro cuando éste superpone e interpreta ambas imágenes. En una cámara el campo visual también está limitado y depende del modelo de cámara con que se cuente. El campo visual para las cámaras que se utilizan en este trabajo es de  $43^\circ$  sobre el plano horizontal y  $32^\circ$  sobre el plano vertical.

Teniendo claras las limitaciones de una cámara, se puede empezar a adquirir imágenes. Para ello se toma en cuenta que la señal de video debe ser muestreada y digitalizada para poder ser tratada en la computadora. El término imagen se refiere a una función bidimensional de intensidad luminosa, que se denota con  $f(x, y)$ , donde la amplitud de  $f$  en las coordenadas espaciales  $(x, y)$  nos da la intensidad de brillantez de la imagen en ese punto.

El nivel de resolución espacial de la imagen es el efecto combinado del número de líneas y el de la frecuencia de muestreo, lo que da como resultado una matriz de  $N \times M$  elementos que se conocen como píxeles (*picture elements* o elementos de la imagen). Cada pixel incorpora el nivel de luminosidad del punto correspondiente a la escena digital (niveles de gris). De este modo se puede enviar la imagen a la computadora para su procesamiento y decodificación.

## 2.4 Percepción humana de la profundidad

Gracias a que se poseen dos ojos es posible obtener una visión tridimensional o estéreo del mundo. La sensación de profundidades y distancias entre los objetos son producto del fenómeno estéreo, así como de la experiencia. Toda la información que recibe el hombre es por la vía de los sentidos, por ello el ser humano hace asociaciones basadas en la experiencia. Una de esas asociaciones se refiere al grado de convergencia (Figura 2.5), es decir, que la posición de los objetos se puede estimar con la inclinación de los ojos al observarlo (Bruce *et al.* 1997: pp. 138).

Los ojos se encuentran a una distancia de unos cuantos centímetros. Esta distancia se llama base

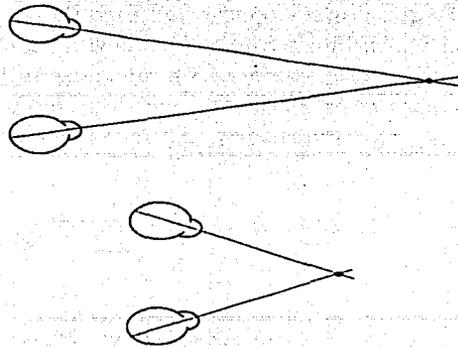


Figura 2.5: Convergencia

estéreo, la cual en promedio es de 6.3 cm a 7.0 cm en un adulto. Como ya se mencionó, cada uno de los ojos es capaz de obtener una imagen independiente del mundo exterior. El cerebro combina ambas en una (imagen estereoscópica) y, con cierta información adicional, reporta la profundidad, la que permite distinguir qué está cerca y qué está lejos, así como relacionar los tamaños de los objetos y la distancia a la cual se está de ellos.

Como el mecanismo de una cámara es muy parecido al del ojo humano, si se quiere ver una imagen en tres dimensiones, es necesario contar por lo menos con dos cámaras. Se requiere que tanto la cámara derecha como la izquierda, reciban una imagen del mismo objeto. De esta forma la computadora se encargará de procesarlas en una sola imagen en tres dimensiones. Esto significa que un punto en una sola imagen no determina la localización de éste en el espacio. Como ya se mencionó, para lograr un reconocimiento de objetos tridimensionales con dos imágenes bidimensionales es necesario tener una cámara con movimiento controlado, o varios sensores que ayuden a una *fusión* del ambiente. Normalmente, para obtener imágenes estereo simples se tienen que obtener dos imágenes desde diferentes puntos, como se describe a continuación.

## 2.5 Procedimiento sencillo para fotografías en tres dimensiones

Un procedimiento muy simple que permite obtener imágenes tridimensionales con una cámara es el siguiente (<http://3eradimension.com/iart/photography.htm>). Primero se enfoca el objetivo con ayuda de un ojo y se toma una foto, luego se debe mover la cámara horizontalmente hasta el otro ojo y tomar una segunda fotografía (Figura 2.6a). Debido a que se está realizando un movimiento horizontal al tomar las fotografías, existirán dos imágenes adicionales en los bordes del par estereoscópico (Figura 2.6b). Esto se conoce como *parallax* y es causado por el corrimiento de la cámara. Se debe reducir este



efecto al mínimo y para ello, hay que poner el centro del foco de la cámara apuntando a algún sitio que sea fácil de reconocer, y se toma la primera fotografía. Luego se mueve la cámara horizontalmente y hay que cuadrar nuevamente el foco de la cámara en el mismo punto de la foto anterior. Normalmente se tiene que inclinar la cámara para conseguirlo (aproximadamente hay que girar la cámara unos  $4^\circ$  o  $5^\circ$ ), este procedimiento es el análogo a la convergencia mencionada anteriormente (Figura 2.7).

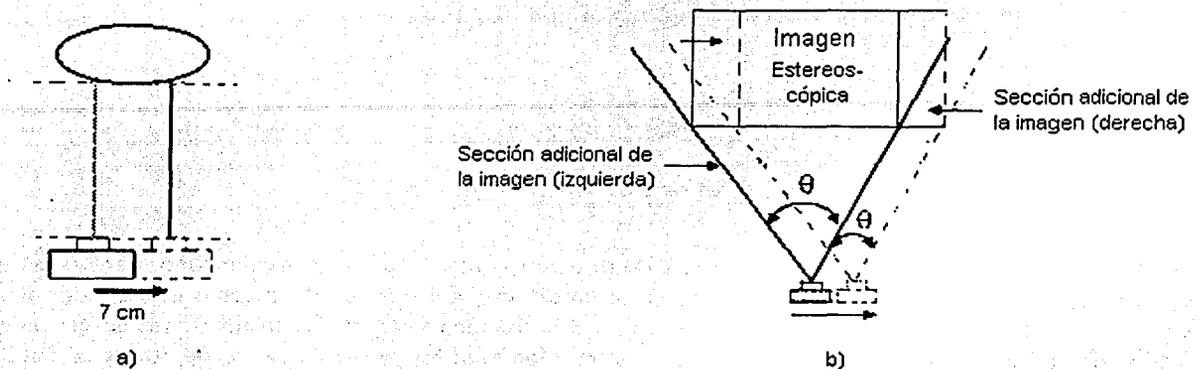


Figura 2.6: Ejemplo del procedimiento para una fotografía tridimensional sencilla

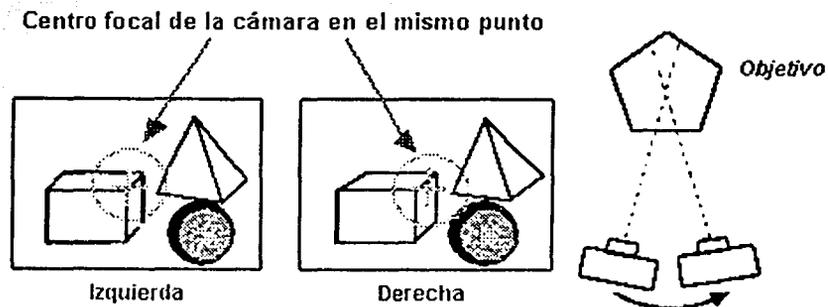


Figura 2.7: Para una buena fotografía estereoscópica

Existe un tipo de imágenes que no es posible tomar con los métodos indicados anteriormente: las imágenes en movimiento. En este caso no se tiene tiempo de pasar la cámara de un ojo a otro mientras el objeto que se desea tomar está moviéndose. Para que se produzca el efecto estéreo las imágenes deben ser las mismas. La solución consiste en utilizar dos cámaras, de tal forma que se puedan disparar ambas a la vez.

## 2.6 Matrices de transformación

Para poder interpretar una imagen, y en particular para poder determinar la posición de un objeto en el espacio por medio de imágenes, es necesario contar con un marco de referencia inercial. Sin embargo, representar cualquier posición y orientación en este marco inercial puede resultar bastante complejo. Una solución consiste en contar con múltiples sistemas coordenados, donde una posición y orientación sean más fáciles de representar, y luego convertir la información al marco de referencia inercial o a cualquier otro sistema de coordenadas. Esto se lleva a cabo utilizando matrices de transformación.

Todas estas transformaciones se expresan en sistemas coordenados tridimensionales, y son válidas tanto para puntos de imágenes como para cualquier punto tridimensional. En general, una transformación se compone de una traslación y de una rotación. Si se desea trasladar un punto con coordenadas  $(X, Y, Z)$  a una nueva posición con los desplazamientos  $(x_0, y_0, z_0)$ , se obtiene el nuevo punto  $(X^*, Y^*, Z^*)$  (Gonzalez y Woods 1993: pp. 52):

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.1)$$

La notación con matrices cuadradas simplifica los procesos de transformación, por tanto se reescribe la ecuación (2.1) de la siguiente forma:

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.2)$$

Equivalentemente se tiene:

$$v^* = Av, \quad (2.3)$$

donde  $A$  es una matriz de transformación, y  $v$  es el vector que contiene las coordenadas del punto original. Existen tres transformaciones básicas para desplazamientos puros a los largo de los ejes coordenados cartesianos.  $Trans_{x,a_i}$  denota la traslación de una distancia  $a_i$  en dirección del eje  $x$ :

$$Trans_{x,a_i} = \begin{bmatrix} 1 & 0 & 0 & a_i \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.4)$$

$Trans_{y,b_i}$  es la traslación de una distancia  $b_i$  en dirección del eje  $y$ :

$$Trans_{y,b_i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & b_i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.5)$$

Finalmente,  $Trans_{z,c_i}$  es la traslación de una distancia  $c_i$  en dirección del eje  $z$ :

$$Trans_{z,c_i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & c_i \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2.6)$$

El caso de la rotación es un poco más complicado y, al igual que en el caso de las traslaciones, existen rotaciones básicas.  $Rot_{z,\theta}$  es la rotación  $\theta$  alrededor del eje  $z$  (Gonzalez y Woods 1993: pp. 52):

$$Rot_{z,\theta} = \begin{bmatrix} c_\theta & -s_\theta & 0 & 0 \\ s_\theta & c_\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2.7)$$

El ángulo  $\theta$  se mide en sentido antihorario viendo hacia el origen desde un punto en la parte positiva del eje  $z$ . Esta transformación sólo afecta a los valores en los ejes  $x$  e  $y$ .  $Rot_{x,\alpha}$  es la rotación  $\alpha$  alrededor del eje  $x$ .

$$Rot_{x,\alpha} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_\alpha & -s_\alpha & 0 \\ 0 & s_\alpha & c_\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2.8)$$

Finalmente,  $Rot_{y,\phi}$  es la rotación  $\phi$  alrededor del eje  $y$ .

$$Rot_{y,\phi} = \begin{bmatrix} c_\phi & 0 & s_\phi & 0 \\ 0 & 1 & 0 & 0 \\ -s_\phi & 0 & c_\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2.9)$$

La matriz de transformación homogénea más general  $T$  toma la forma:

$$T = \begin{bmatrix} n & s & a & d \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (2.10)$$

donde  $n$  es un vector en dirección del eje  $x$ ,  $s$  del eje  $y$  y  $a$  del eje  $z$ ; siendo todos unitarios.

# Capítulo 3

## Procesamiento y análisis de imágenes digitales

¿Quién se atreverá a poner límites  
al ingenio de los hombres?  
– Galileo

### 3.1 Elementos del sistema

Las distintas etapas del procesamiento digital de imágenes son (Gonzalez y Woods 1993: pp. 10):

1. Adquisición
2. Almacenamiento
3. Procesamiento y
4. Despliegue

#### Adquisición de imágenes

Para adquirir imágenes es necesario un dispositivo físico que sea sensible a la luz y que produzca una señal eléctrica de salida proporcional al nivel de energía sensada. Además se necesita un dispositivo capaz de convertir la salida eléctrica en una versión digital. El dispositivo con el que se trabaja en el sistema de visión de esta tesis está compuesto por dos cámaras CCD. En el Apéndice B se hace una descripción más detallada de las cámaras.

La escena que se proyecta en los sensores se convierte en una estructura discreta de valores enteros. Esta estructura es una retícula rectangular o cuadrada, normalmente expresada en forma de matriz con

$M$  renglones y  $N$  columnas. Los valores enteros que aparecen en ella son producto de una cuantización que está definida en  $K$  intervalos discretos; estos están normalmente definidos por el número de bits  $BN$  que se desean asignar a cada pixel ( $BN = 2^K$ ).

### Almacenamiento de imágenes

Una sola imagen de  $N \times M$  pixeles de  $BN$  bits requiere de cierto espacio para su almacenamiento, ya sea que se desee conservar la imagen por largo tiempo o sólo para poder hacer manipulaciones con la información.

En el almacenamiento de las imágenes digitales que se hace en esta tesis se utilizan dos categorías principales, según los procesos a los que se tengan que someter las imágenes. El almacenamiento a corto plazo se utiliza cuando las cámaras están en la función de grabado y cuando se hace la manipulación de imágenes dentro de funciones, como la corrección de la distorsión. En estas tareas las matrices de valores enteros se encuentran en memoria de corto plazo. El almacenamiento a largo plazo se utiliza para guardar ciertas imágenes resultantes de algunas partes del proceso.

### Procesamiento

El procesamiento que se les hace a las imágenes digitales está generalmente estipulado por algoritmos. Con excepción de la adquisición y del despliegue, la mayor parte del procesamiento de la imagen se puede implementar en software, es decir que es programable. Las implementaciones que se hacen en hardware son sólo para procesos especializados, que generalmente no pueden ser elaborados en computadoras personales por los recursos que se requieren. Es importante resaltar que muchas veces en el procesamiento de imágenes se tiene que recurrir a procesadores digitales de señales, porque la cantidad de cálculo que se maneja es realmente exclusiva para un procesador convencional.

### Despliegue

La información de las cámaras llega hasta la PC por medio de una tarjeta de adquisición de imágenes de National Instruments (ver Apéndice D) y el despliegue se hace por medio de programas que permitan la manipulación de la información visual: LabView o LabWindows.

El despliegue se refiere a la forma en que el usuario puede ver las imágenes (procesadas o no). El dispositivo que las muestra puede ser un monitor o una impresión. En este caso el despliegue de las imágenes se hace en pantalla.

## 3.2 Digitalización

En el diseño y análisis de los sistemas de procesamiento digital es necesario hacer una caracterización matemática de la imagen a procesar. Cuando se adquiere una imagen se digitaliza para tener la in-

formación luminosa de la escena a interpretar, es decir, se hace una cuantización para obtener una matriz de valores enteros. La cuantización es la operación de conversión de una señal analógica en un número digital con resolución finita de representación. La cuantización depende de la cantidad de bytes disponibles para cada pixel. Normalmente se hace una cuantización de un byte por pixel, por tanto existen 8 bits disponibles para  $K$  intervalos, donde  $K = 2^8 = 256$  (Pratt 1991: pp. 3).

La cuantización se hace en niveles uniformes. Se selecciona el número de bytes que se utilizarán para determinar los niveles de cuantización y luego se hace la conversión a los valores correspondientes. En este caso los valores más pequeños son colores oscuros y los más grandes claros, siendo el negro 0 y el blanco 255. Como los valores resultantes deben ser enteros, éstos representan sólo una aproximación del valor real, por lo que existe un error de cuantización, que es la diferencia entre la señal luminosa analógica de la escena y la salida del cuantizador. Este error no es de gran importancia en este trabajo, porque los objetos que se utilizan poseen suficiente contraste como para considerar que el error de cuantización es despreciable (Couch II 1993: pp. 137).

El número de líneas y columnas que hay en la imagen, junto con la frecuencia de muestreo que tiene la cámara de video, especifica el nivel de resolución espacial de la imagen. Cada pixel incorpora el nivel de luminosidad del punto correspondiente en la escena digital.

En la Figura 3.1 se muestra una imagen de  $13 \times 13$  pixeles (sin considerar el borde) a 256 colores. La Figura 3.2 expone la cuantización de esta imagen (sin borde). Se ve claramente que cada pixel se representa por medio de un número entero de cuantización, donde el 255 es el valor correspondiente al blanco. Mientras más oscuros sean los tonos, el valor va disminuyendo hasta 0 (negro). Las zonas de las letras *F* e *I* se pueden distinguir por los números menores (104 y 76).

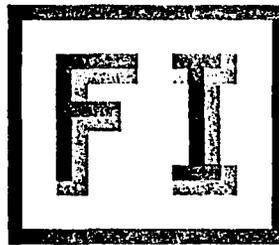


Figura 3.1: Imagen original FI a 256 colores

### 3.3 Filtrado

El proceso de cambiar las amplitudes relativas de las componentes de frecuencia en una señal, o el de eliminar por completo algunas componentes de frecuencia, se conoce como filtrado (Oppenheim y Willsky 1998: pp. 230). El filtrado sirve esencialmente para separar señales (eliminación de ruido), para modelarlas o para estimar parámetros de éstas. Existen varios tipos de filtros según (Escobar 2002):

|     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 |
| 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 |
| 255 | 76  | 76  | 76  | 76  | 76  | 255 | 255 | 76  | 76  | 76  | 104 | 255 |
| 255 | 76  | 104 | 104 | 104 | 104 | 255 | 255 | 104 | 76  | 104 | 104 | 255 |
| 255 | 76  | 104 | 255 | 255 | 255 | 255 | 255 | 255 | 76  | 104 | 255 | 255 |
| 255 | 76  | 76  | 76  | 76  | 255 | 255 | 255 | 255 | 76  | 104 | 255 | 255 |
| 255 | 76  | 104 | 104 | 104 | 255 | 255 | 255 | 255 | 76  | 104 | 255 | 255 |
| 255 | 76  | 104 | 255 | 255 | 255 | 255 | 255 | 255 | 76  | 104 | 255 | 255 |
| 255 | 76  | 104 | 255 | 255 | 255 | 255 | 255 | 255 | 76  | 104 | 255 | 255 |
| 255 | 76  | 104 | 255 | 255 | 255 | 255 | 255 | 76  | 76  | 76  | 104 | 255 |
| 255 | 104 | 104 | 255 | 255 | 255 | 255 | 255 | 104 | 104 | 104 | 104 | 255 |
| 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 |
| 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 |

Figura 3.2: Arreglo de cuantización

- La continuidad
- La banda de paso
- La variancia en el tiempo
- La linealidad

El filtrado espacial es un método utilizado para conocer la distribución de la intensidad luminosa de los píxeles sobre una imagen y para detectar y acentuar discontinuidades en las fronteras. La frecuencia espacial se define por cambios de los valores de intensidad a lo largo de una dirección determinada en la imagen; estas diferencias se expresan como un número de ciclos por unidad de distancia (<http://www.euromet.met.ed.ac.uk>). Una imagen con una única frecuencia espacial está compuesta por rayas igualmente espaciadas. En general, las imágenes tienen varias frecuencias espaciales dominantes.

Técnicas matemáticas utilizadas para separar una imagen en sus varias frecuencias espaciales son el análisis de Fourier y el filtrado de convolución. Un ejemplo de éste último es el reemplazo del valor de cada píxel por un promedio sobre un área cuadrada, centrada en dicho píxel. En este caso es como si se aplicara un filtro pasobajas; las desviaciones sobre las medias locales tienden a reducirse y por tanto la imagen se suaviza. Es importante hacer notar que el filtrado es esencial en el proceso de detección de bordes que se expone a continuación (Gonzalez y Woods 1993), y que el método de filtrado que se utiliza es el de convolución.

### 3.4 Segmentación y extracción de características

El objetivo principal dentro del sistema de visión es el reconocimiento de objetos dentro de un espacio para poderlos manipular por medio del sistema de robots industriales. Para llegar a esta meta es necesario, en el caso de este trabajo, conocer las características de cada uno de los objetos en las imágenes para poder clasificarlos y localizarlos. Como características de los objetos pueden usarse:

- Formas o contornos
- Colores o texturas
- Señalizaciones de ayuda

Estas características se pueden extraer si se conoce la forma exacta del objeto, tratando de que una plantilla encuentre coincidencia dentro de la imagen. Generalmente esto es muy difícil ya que los objetos se ven distorsionados o leves rotaciones hacen que se vean distintos.

La segmentación es un método útil para extraer características específicas de las imágenes. Por medio de técnicas de segmentación se pueden hallar niveles de grises o bordes que describan algún segmento específico. La segmentación es la parte más importante y delicada de cualquier sistema de visión artificial, porque de ella depende la correcta interpretación de la imagen. Segmentar significa dividir una imagen digital en zonas específicas y es el primer paso del reconocimiento. Una región, en una imagen, es un grupo de píxeles conectados que tienen propiedades similares. Una imagen compleja puede ser sustituida por otra en donde cada sección está perfectamente descrita. Una vez segmentada la imagen se puede formar una lista de objetos agrupando sólo algunos píxeles.

El proceso de segmentación es en sí bastante complicado, por lo que en este trabajo se utilizarán sólo objetos con niveles de intensidad uniformes o casi uniformes, es decir, sin texturas, así como objetos no traslapados; también se intentará que exista un alto contraste en las zonas que se desean analizar. Debido a los errores que se generan de la segmentación, la correspondencia entre regiones y objetos no es perfecta; por ello se deben utilizar generalmente ciertos conocimientos específicos en etapas posteriores para la correcta interpretación de las imágenes. Lo anterior significa que, para una autonomía más grande hace falta una mejor inteligencia artificial.

El primer paso en el análisis para entender una imagen es partirla de tal forma que las regiones que representan diferentes objetos sean marcadas explícitamente. Estas particiones pueden ser obtenidas a partir de las características de los niveles de gris en la imagen. El nivel adecuado lo determina la aplicación (cuando se logra aislar el objeto deseado). Los algoritmos de segmentación para imágenes monocromáticas están basados en dos propiedades de los niveles de gris: la discontinuidad y la similitud. La primera hace la partición en cambios abruptos; sirve para detectar puntos aislados, líneas y bordes en la imagen. La segunda es para regiones por umbrales: dos píxeles pueden ser asignados a la misma región si tienen características similares de intensidad y si están próximos. Al parecer los principios de similitud y proximidad hacen que los puntos de un mismo objeto produzcan en la imagen píxeles con niveles de gris similares y próximos espacialmente; pero esto no es cierto en muchas situaciones (Gonzalez y Woods 1993: pp. 414).

### 3.4.1 Detección de bordes, líneas y regiones

Las fronteras son bordes unidos que caracterizan la forma de un objeto y son útiles para calcular rasgos geométricos, como tamaño u orientación. El objetivo final es alcanzar al menos una segmentación parcial, es decir, agrupar bordes locales en una imagen donde estén presentes sólo cadenas de bordes con correspondencia con los objetos en la imagen. Los algoritmos detectores de bordes están basados en los cambios de intensidad luminosa, ya que los bordes están precisamente donde la función de intensidad cambia drásticamente. Los bordes se distinguen fácilmente según su función de gradiente, pues generalmente muestra un perfil de ascenso precipitado, como se puede ver en la Figura 3.3.

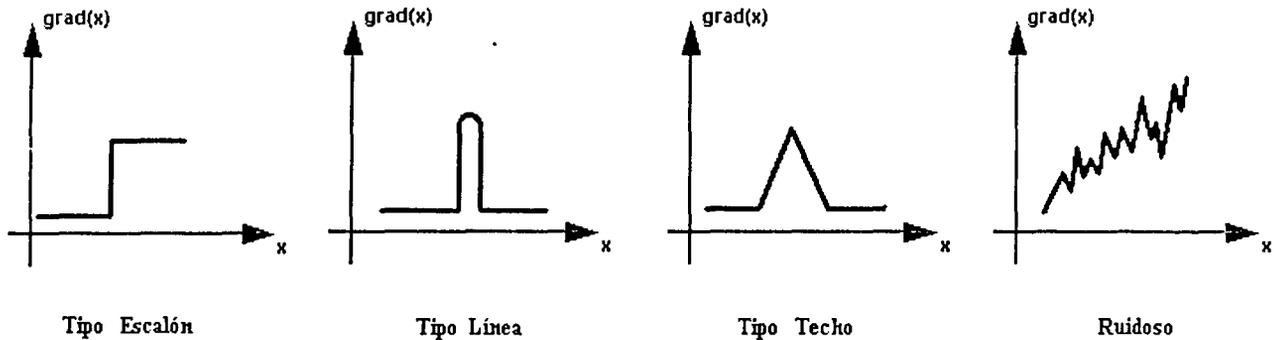


Figura 3.3: Tipos de función de borde

El cálculo diferencial que se hace en los detectores de bordes describe a los cambios de funciones continuas usando las derivadas. Como la función de intensidad de la imagen depende de dos variables, se deben usar derivadas parciales. Un borde es un vector variable con dos componentes: magnitud y dirección (Sonka *et al.* 1993: pp. 76). La magnitud se refiere a la magnitud del gradiente, y la dirección a la dirección  $\phi$  del borde, que está rotada con respecto a la dirección del gradiente  $\psi$ ,  $-90^\circ$ . El gradiente da la dirección hacia el máximo crecimiento de la función, como se muestra en la Figura 3.4.

Dentro de los detectores de bordes existen varios operadores: Operador de Robert, de Laplace, de Prewitt, de Robinson, de Kirsch, etc. (Sonka *et al.* 1993: Capítulo 4). El detector de bordes de Canny que se utilizará en la sección de calibración tiene como objetivo principal localizar los bordes en una imagen contaminada con ruido blanco. Su ventaja radica en la utilización de tres criterios principales:

1. Detección
2. Localización y
3. Respuesta única

El criterio de detección expresa que los bordes importantes no se deben perder, y que no deben existir respuestas dudosas. La localización es el criterio que dice que la distancia entre la posición actual y

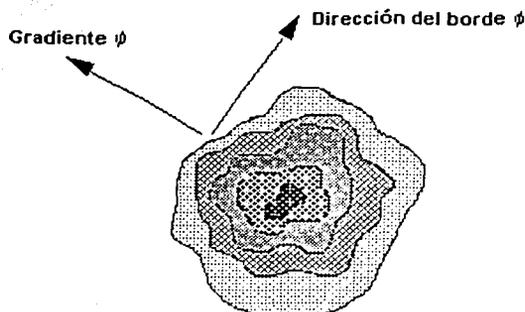


Figura 3.4: Dirección del borde y del gradiente

la posición localizada debe ser mínima. Finalmente, el criterio de respuesta única minimiza múltiples respuestas a una sola. Cuando hay dos respuestas una se debe tomar como falsa y con ello se resuelve el problema de bordes ruidosos.

El detector de bordes de Canny es una herramienta bastante robusta y, dentro de las técnicas de detección de bordes, ha sido de gran trascendencia. Se basa en distintas ideas que han ido evolucionando:

- Detección en señales univariadas con cálculos de variaciones
- Implementación del criterio de respuesta única con una efectividad mayor al 80%
- Umbralización con histéresis, entre otras.

Canny propuso una aproximación de todas las herramientas sintetizadas. Los bordes más significativos se marcan primero y se va haciendo una selección según la desviación estándar de las gaussianas. El algoritmo consiste a grandes rasgos en convolucionar la imagen con una gaussiana de distintas escalas. Así se puede estimar la dirección del borde e iniciar la búsqueda de los bordes. Con ello se calcula su magnitud y luego se les pasa una máscara de umbralización con histéresis para eliminar las impurezas. Finalmente se agrega la información faltante, es decir, se completan las líneas que sean discontinuas.

Hay otros algoritmos que sirven para detectar figuras geométricas predeterminadas; el método de la transformada de Hough es uno de ellos. Este método es robusto y poco sensible a datos imperfectos. Inicialmente fue diseñado para hallar líneas rectas y curvas, pero se puede utilizar para cualquier figura si sus bordes son susceptibles de ser representados por medio de ecuaciones analíticas (Hsu y Huang 1990, McKenzie y Protheroe 1990, Mitiche y Habelrih 1989).

La idea básica del método se puede explicar con la detección de una línea recta (Dean *et al.* 1995: Sección 9.5). El método de la transformada de Hough consiste en aplicar un detector de bordes. Con éste se remarcan los límites de las figuras y así se puede proceder a la detección de líneas. La detección se

hace por medio de un barrido de la imagen en función de un espacio de parámetros elegido (según el tipo de coordenadas que se utilicen). Cuando se encuentra un pixel que cumple con lo asignado en este espacio de parámetros se le agrega un valor al espacio mencionado. De esta forma, las líneas que existen en la imagen irán acentuándose dentro del espacio de parámetros. Las líneas de la imagen se pueden localizar como los picos del espacio de parámetros. Con esto se entiende la poca sensibilidad del método al ruido, pues líneas incompletas sólo resultan en picos de menor valor (Sonka *et al.* 1993: pp. 149). Es importante hacer notar que el tamaño del mapa de transformación tiene mucho peso en los resultados que se deseen obtener. Por ejemplo, un mapa de  $80 \times 80$  pixeles para  $\theta$  proporciona una precisión de  $2.25^\circ$ . En cambio, si el mapa es de  $800 \times 800$  pixeles la precisión será de  $0.225^\circ$ . La Figura 3.5 es una imagen que contiene líneas que se desean localizar.



Figura 3.5: Imagen original de prueba

Se aplica el detector de bordes de Canny como se describió anteriormente, y se obtiene la Figura 3.6 del mismo tamaño que la anterior. Después de la asignación al espacio de parámetros y la localización de



Figura 3.6: Imagen después del detector de bordes de Canny

los picos más intensos, se detectan ciertas líneas. La superposición de la imagen original con la imagen

de detección de líneas (Figura 3.7) muestra que las líneas discontinuas siguen siendo detectadas, pero que las líneas más pequeñas son despreciadas. En esta tesis se trabajó inicialmente con el método de



Figura 3.7: Imagen comparativa

la transformada de Hough para la localización de los bordes de los objetos. Debido a que el mapa de parámetros necesita un nivel de umbralización bastante preciso y a que los cruces entre las líneas que se detectaban (con las herramientas con que se contaban) no eran lo suficientemente confiables, se decidió cambiar de método de segmentación.

El método que resulta más fácil, en ciertos casos como el que se maneja en este trabajo, es el basado en umbralización de un histograma. El histograma de una imagen permite conocer la frecuencia relativa de aparición de cada uno de los posibles niveles de gris. Una forma de extraer los objetos del fondo es seleccionar un umbral que los separe. Se puede segmentar la imagen en regiones de píxeles con nivel de gris en un conjunto, y el resto como fondo. También se pueden utilizar umbrales múltiples, lo que produce imágenes no binarias con un conjunto muy limitado de niveles de gris. Sin embargo, este tipo de umbralización multinivel es generalmente menos fiable que la umbralización en dos regiones. La razón es que resulta difícil establecer múltiples umbrales que aislen regiones correctamente, especialmente cuando hay muchos máximos locales. Para hacer la segmentación más robusta, el umbral debería ser seleccionado automáticamente por el sistema. Tanto el conocimiento de la escena como el del problema a resolver y cualquier otra información deberían ser usados para fijar el umbral. La información de la que se dispone puede venir proporcionada por las características de los objetos: su tamaño, la fracción de la imagen que ocupan y los diferentes tipos de objetos que pueden aparecer en la imagen.

### 3.5 Modelo geométrico de una cámara

Para poder manejar la información que se obtiene de las imágenes de una cámara es necesario definir un modelo geométrico y matemático que facilite la manipulación de los datos. Los parámetros que definen las características geométricas de una cámara son:

1. Longitud focal  $f$

2. Coordenadas del centro óptico  $(X_0, Y_0)$  y
3. Razón entre las escalas (tanto verticales como horizontales) de la imagen  $(k_x, k_y)$ .

El término de longitud o distancia focal, en óptica, se relaciona con el índice de refracción de las lentes y los radios de curvatura de sus superficies (Halliday *et al.* 1998: pp. 377). Como el modelo de cámara que se maneja en este trabajo es de tipo *pinhole*, esta distancia se refiere sólo a la separación entre el centro de la lente y el plano donde se forma la imagen.

Las coordenadas del centro óptico se refieren a las coordenadas donde se encuentra el centro de la imagen. Se encuentra a través del foco de expansión de un conjunto de puntos espaciales de referencia, mientras la cámara se traslada a lo largo de su eje óptico. Finalmente, la razón entre las escalas incluye una combinación de los factores de escala y longitud focal.

La ecuación que se maneja en este trabajo para describir el modelo de la cámara es:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (3.1)$$

donde  $(x, y, z)$  es la proyección del punto  $P$  en el plano de la imagen (Figura 3.8).

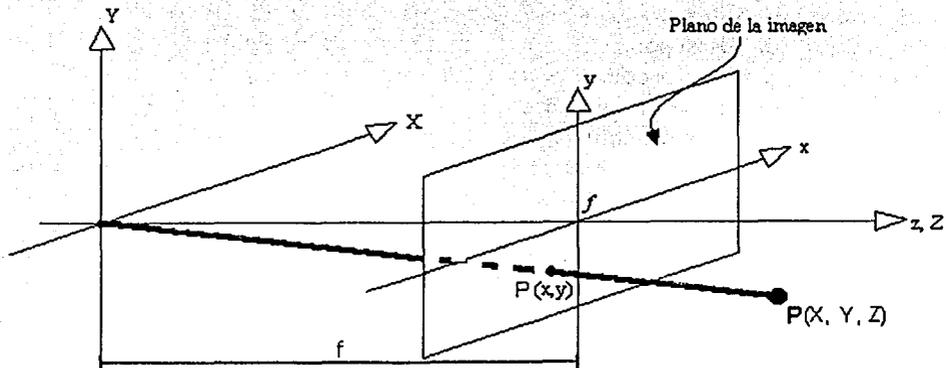


Figura 3.8: Modelo básico de la proyección

Este modelo tiene como premisa que la cámara está en coincidencia con el sistema coordenado espacial. Se asume que la lente de la cámara está enfocada al infinito, de tal forma que la distancia focal es la distancia entre el plano de la imagen y el origen. Por facilidad se utiliza la distancia focal en unidades de píxeles, ya que la información de la densidad de píxeles en el sensor no está siempre disponible para hacer la conversión a unidades de longitud (Kanatani 1995: pp. 20).

### 3.5.1 Transformaciones Proyectivas

Las transformaciones de perspectiva o transformaciones proyectivas son las que proyectan puntos tridimensionales en planos. Estas transformaciones juegan un papel muy importante en el procesamiento de las imágenes, porque dan la forma aproximada de una visión en tercera dimensión. Las transformaciones son esencialmente distintas a las anteriores porque son no lineales e involucran divisiones entre valores de las coordenadas (Gonzalez y Woods 1993: pp. 56).

El modelo de la cámara que se presentó en la sección anterior (Figura 3.8) es de gran utilidad para la creación de las matrices de transformación que se verán a continuación. En la Figura 3.8 se muestra el modelo de la formación de la imagen, donde  $f$  representa la distancia focal. Se asume que la cámara y el sistema coordenado están alineados con el sistema coordenado del espacio. El centro del plano de la imagen es  $(0, 0, f)$ , mientras que el centro del sistema de coordenadas es  $(0, 0, 0)$ .

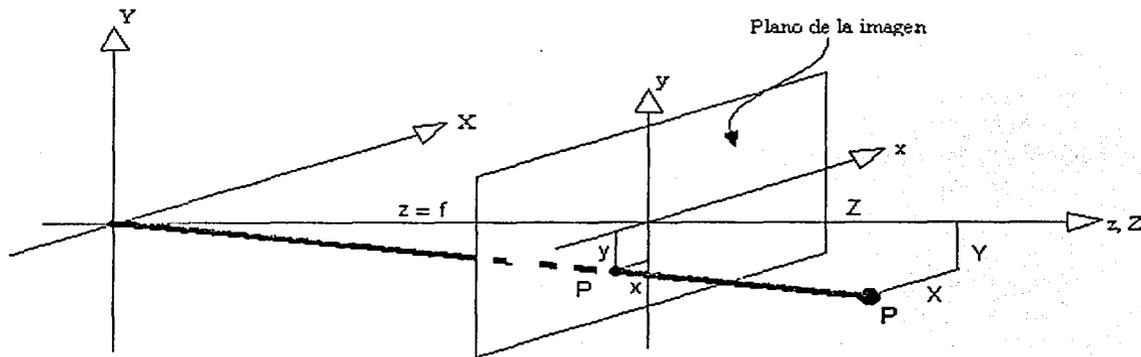


Figura 3.9: Relaciones de triángulos

Si  $(X, Y, Z)$  son las coordenadas de cualquier punto en el espacio (localizado al frente de la lente como se muestra en la figura), es necesario conocer las coordenadas  $(x_i, y_i)$  de las proyecciones de este punto en el plano de la imagen ( $z_i$  siempre es cero). Se obtienen entonces, por relaciones de triángulos semejantes, las siguientes igualdades:

$$x = f \frac{X}{Z} \quad (3.2)$$

y

$$y = f \frac{Y}{Z} \quad (3.3)$$

Estas igualdades se refieren precisamente al modelo de la cámara de la sección anterior. Se observa que estas ecuaciones son no lineales, porque están divididas entre la variable  $Z$ . Aunque se pueden

utilizar directamente, es conveniente expresarlas en forma matricial con coordenadas homogéneas. Las coordenadas homogéneas para  $(X, Y, Z)$  se definen como  $(kX, kY, kZ)$ , donde  $k$  es una constante arbitraria, diferente de cero (usualmente 1). De ahí que un vector  $w$  dado por:

$$w = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (3.4)$$

tenga la contraparte homogénea:

$$w_h = \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix}. \quad (3.5)$$

Nótese que el vector  $w$  se obtiene al dividir los tres primeros elementos de  $w_h$  entre el cuarto. La matriz de transformación de perspectiva se define como:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{Z}{f} \end{bmatrix} \quad (3.6)$$

Si se hace el producto  $Pw_h$  se obtiene  $c_h$  que son las coordenadas de la cámara de forma homogénea:

$$c_h = \begin{bmatrix} kX \\ kY \\ kZ \\ \frac{kZ}{f} \end{bmatrix} \quad (3.7)$$

Estas coordenadas se pueden convertir a cartesianas con una simple división, obteniéndose:

$$c = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{fX}{Z} \\ \frac{fY}{Z} \\ \frac{fZ}{Z} \end{bmatrix} \quad (3.8)$$

Es fácil ver que los valores de  $x$  e  $y$  coinciden con los de la ecuación (3.1) del modelo de la cámara.

La transformación inversa de un punto del plano de la imagen a un punto de la imagen se puede obtener de la siguiente forma:

$$w_h = P^{-1}c_h. \quad (3.9)$$

Si se tiene un punto  $(x_0, y_0, f)$  en el plano de la imagen, expresado en el sistema coordenado general, se obtiene que:

$$c_h = \begin{bmatrix} kx_0 \\ ky_0 \\ kf \\ k \end{bmatrix} \quad (3.10)$$

como  $P^{-1}$  es:

$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{f}{Z} \end{bmatrix} \quad (3.11)$$

entonces se tiene

$$w_h = \begin{bmatrix} kx_0 \\ ky_0 \\ kf \\ \frac{kf}{Z} \end{bmatrix}, \quad (3.12)$$

lo que en coordenadas cartesianas es:

$$w = \begin{bmatrix} \frac{Zx_0}{f} \\ \frac{Zy_0}{f} \\ Z \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (3.13)$$

Este resultado es bastante conveniente, porque para cualquier par de coordenadas se tendrá algún valor en  $Z$ . Es importante hacer notar aquí que, aunque aparecen las tres coordenadas espaciales ( $X, Y, Z$ ), no es posible determinar la localización del punto en el espacio si se conoce sólo  $x_0$  e  $y_0$ . Para recuperar un punto en el espacio, con coordenadas del plano de la imagen, es necesario conocer al menos una de las coordenadas espaciales del punto original. Este problema se resuelve más adelante utilizando de ambas cámaras de forma simultánea.

## 3.6 Visión Tridimensional

Un sistema estéreo es el que sirve para medir la profundidad de las superficies de los objetos utilizando dos cámaras. Como la traslación de la escena es equivalente a la traslación de una cámara en la escena, un sistema estéreo se puede tratar como un problema de este tipo. Cuando se identifican los dos planos de la cámaras, las líneas imaginarias que conectan los puntos correspondientes en cada imagen se llaman *epipolares*.

### 3.6.1 Localización de un punto en el espacio

Para lo que corresponde a la visión estereoscópica se necesitan dos imágenes. En el laboratorio se cuenta con dos cámaras de video fijas que toman las imágenes simultáneamente para obtener el efecto estéreo. Las cámaras están en posición de ejes paralelos y separadas horizontalmente una distancia



Figura 3.10: Localización de las cámaras

definida (Figura 3.10). El objetivo es encontrar las coordenadas  $(X, Y, Z)$  de un punto  $P$  en el espacio (Figura 3.11). Se asume que los sistemas coordenados de ambas cámaras son iguales y que difieren sólo en la localización del origen del sistema.

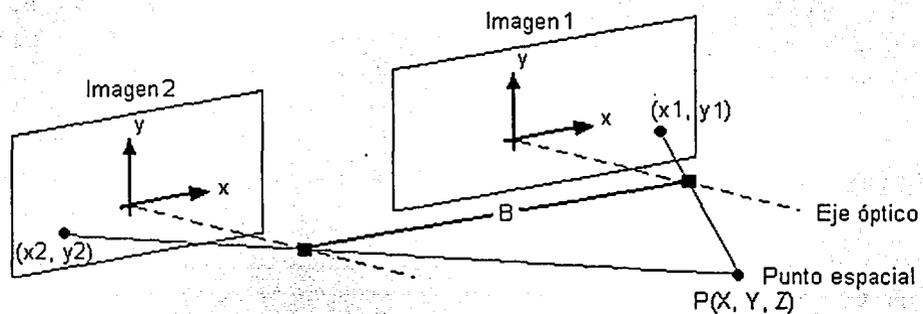


Figura 3.11: Punto en 3D

El punto  $P$  tiene las coordenadas  $(x_1, y_1)$  en la imagen 1 y  $(x_2, y_2)$  en la imagen 2. Como las dos cámaras son paralelas se puede asumir que la coordenada  $Z$  es igual para las dos cámaras. De este modo se deduce fácilmente lo siguiente:

$$X_2 = X_1 + B \quad (3.14)$$

Si se toma en cuenta la distancia focal  $f$ , se puede obtener, como en las secciones anteriores, por

relaciones en triángulos semejantes, la siguiente ecuación para la coordenada  $X_1$ :

$$X_1 = \frac{Z_1 x_1}{f_1} \quad (3.15)$$

Esto es válido si se hace al marco de referencia de la cámara 1 el sistema coordenado del espacio en general. Como

$$Z_1 = Z_2 = Z \quad (3.16)$$

Se obtiene:

$$X_1 = \frac{Z x_1}{f_1} \quad (3.17)$$

y

$$X_1 + B = \frac{Z x_2}{f_2} \quad (3.18)$$

Finalmente se obtiene:

$$Z = \frac{B}{\frac{x_2}{f_2} - \frac{x_1}{f_1}} \quad (3.19)$$

La ecuación (3.19) indica que, si se conocen las coordenadas  $x_1$  y  $x_2$ , la línea base  $B$  y las distancias focales  $f_1$  y  $f_2$  de cada cámara, entonces es trivial conocer la coordenada  $Z$ . Las coordenadas  $X$  y  $Y$  se pueden calcular de las ecuaciones (3.2 y 3.3) como:

$$X_i = \frac{Z x_i}{f_i} \quad (3.20)$$

y

$$Y_i = \frac{Z y_i}{f_i} \quad (3.21)$$

utilizando cualquier par de puntos  $(x_1, y_1)$  de la imagen 1 o  $(x_2, y_2)$  de la imagen 2. El problema más grande es el de encontrar los dos puntos correspondientes en las distintas imágenes. Para esto se utilizan técnicas de correlación. El procedimiento es más sencillo si las imágenes tienen puntos distintivos, como esquinas. La calibración hace que las cámaras se puedan tratar de forma independiente y, con ello, que las imágenes estéreo sean más fácilmente manejables, como se explica en la siguiente sección.

### 3.7 Calibración de las cámaras

Las cámaras de video tienen varios parámetros internos y externos, independientes entre sí, que se deben conocer para manipular los datos que se obtienen de ellas. El proceso mediante el cual se adquieren estos parámetros se llama *calibración*. Ésta sirve para conocer la correspondencia entre el mundo exterior de

los objetos y las imágenes digitales que se obtienen. La calibración representa una parte fundamental de todo el sistema de visión, pues es la responsable de la correcta interpretación de los datos. Errores mínimos en su interpretación pueden convertirse en grandes fallas del sistema.

La primera parte del trabajo experimental consiste en la calibración de las dos cámaras del laboratorio que forman parte del sistema de visión. Las cámaras de fotografía y video caseras no necesitan ningún tipo de calibración, ya que no es necesario hacer mediciones exactas con estos instrumentos. En general, los sistemas que trabajan con elementos bidimensionales tampoco necesitan ser calibrados; pero cuando se trabaja con interpretaciones tridimensionales es importante conocer todos los parámetros de los equipos a fin de poder interpretar las imágenes de la mejor manera. El procedimiento de calibración utilizado en este trabajo está basado en el método de calibración que se encuentra en Kanatani (1995) con algunas modificaciones para hacer más simple su uso.

Los parámetros internos y externos determinan las características geométricas y ópticas de la cámara, así como su orientación y posición respecto a un marco de referencia. Los parámetros internos o intrínsecos son las características geométricas que se presentan en la Sección 3.5, y los parámetros externos o extrínsecos se refieren a la posición y orientación de la cámara con respecto a cierto sistema coordinado. Se eligió un método con el cual se pueden calibrar las cámaras en poco tiempo (después de que se tienen las herramientas informáticas necesarias) y que es además lo suficientemente confiable para el sistema de visión, porque toma en cuenta la incertidumbre inherente del mismo proceso de captura. Esto es algo muy importante pues para el movimiento preciso del robot se deben tener intervalos de confianza muy grandes.

### 3.7.1 Eliminación de la distorsión

Lo primero al realizar la calibración de una cámara es la eliminación de la distorsión de la imagen producida por la lente de la cámara, es decir, una corrección de la imagen. La corrección de la distorsión se hace para que la información que se procese posteriormente sea lo más semejante a la fuente. En particular, en el presente trabajo es muy importante eliminar la distorsión, porque se tienen dos cámaras con sensores diferentes e independientes (aunque sean del mismo modelo y marca).

Para compensar los cambios geométricos se utiliza un modelo polinomial. La distorsión se modela con funciones que mapean cada pixel  $(x, y)$  en el plano de la imagen capturada, la cual se asume distorsionada, a un plano de una imagen ideal  $(x', y')$ . La calibración se lleva a cabo utilizando ecuaciones polinomiales de segundo y cuarto orden (Schalkoff 1990: pp. 52):

$$x' = \sum_{i=0}^m \sum_{j=0}^m a_{ij} x^i y^j \quad y' = \sum_{i=0}^m \sum_{j=0}^m b_{ij} x^i y^j, \quad (3.22)$$

es decir con  $m = 2$  y  $m = 4$ .

Para encontrar los coeficientes de los polinomios es necesario tomar una imagen de figuras regulares *perfectas*. Se utiliza en este caso la imagen de una rejilla regular de cuadros negros sobre fondo blanco (Figura 3.12) tomada con cada cámara perpendicular a ella. Las imágenes que resultan de este

procedimiento se muestran en la Figura 3.13. Las coordenadas de los puntos de la plantilla se conocen

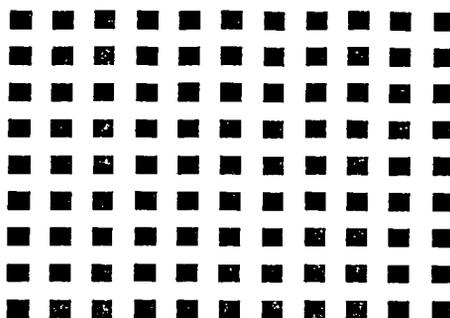


Figura 3.12: Rejilla original para calibración

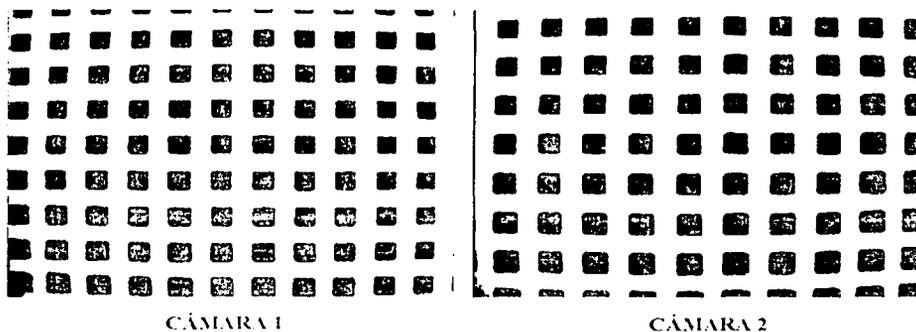


Figura 3.13: Imagen de la rejilla con ambas cámaras

debido a su regularidad. Estas son las esquinas de los cuadros que se usan como puntos de referencia. Con algún algoritmo detector de bordes (como los que se mencionaron en la Sección 3.4.1), se extraen las esquinas de cada uno de los cuadros. Las coordenadas de éstas servirán para formar un sistema de ecuaciones que servirá para la estimación de los coeficientes. El algoritmo que se utiliza en este trabajo es el detector de bordes de Canny. Las imágenes resultantes se observan en la Figura 3.14.

Ahora bien, conocer la distancia focal requiere de un método iterativo. Por tanto, inicialmente se necesita una distancia focal aproximada para la detección de bordes. Para hacer la estimación de la distancia focal se utiliza la siguiente relación, que tiene relación directa con el modelo de la cámara:

$$f = x \frac{Z}{X}, \quad (3.23)$$

donde  $f$  es el foco aproximado,  $x$  es la distancia en pixeles de un borde de algún cuadro en la imagen obtenida,  $Z$  es la distancia de la cámara hasta la imagen real y  $X$  es la distancia real del mismo borde en la imagen real (ver Figura 3.9).

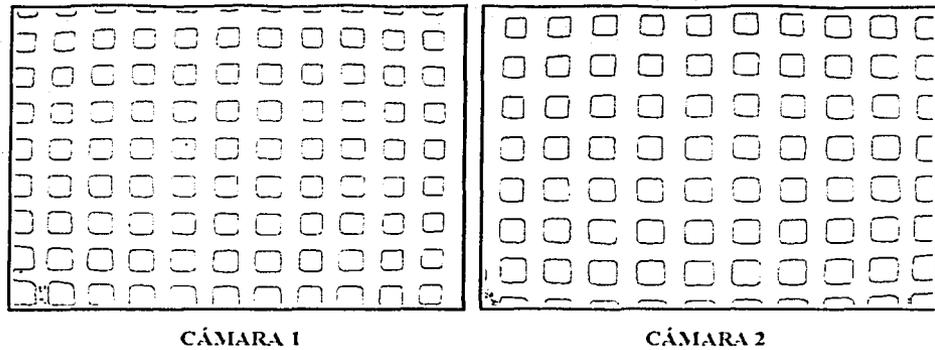


Figura 3.14: Imagen de la rejilla después de pasar por el algoritmo Canny

Las distancias focales aproximadas que se proponen para iniciar la iteración son:

| Cámara   | Distancia focal |
|----------|-----------------|
| Cámara 1 | 583 pixeles     |
| Cámara 2 | 634 pixeles     |

Para que el proceso de calibración sea más fácil y exacto se tienen que emparejar los puntos de la imagen ideal con la imagen de referencia, obtenida de la cámara. Esto es con el objeto de evitar cualquier desplazamiento. Una opción simple sería estimarlo o no tomarlo en cuenta en este momento, pero no es conveniente, porque se complica la corrección de la distorsión después. Por otro lado, lo único que se sabe realmente de la imagen original es que la rejilla debe ser regular, por lo que la imagen ideal se crea basándose en un punto central del cual se hacen cuadros regulares. Utilizando una malla de puntos de referencia se puede correlacionar cada malla de puntos resultantes de las imágenes de las cámaras. Para la malla de las imágenes de las cámaras se toma cada imagen de la Figura 3.14 y se van eligiendo punto a punto cada una de las esquinas de los cuadros que constituyen la rejilla.

Para encontrar el punto central de la imagen ideal se utiliza la distancia Euclídeana  $d$  entre los puntos de la imagen ideal  $(x', y')$  y puntos de la imagen adquirida  $(x, y)$ :

$$d^2 = (x - x')^2 + (y - y')^2 \quad (3.24)$$

Se supone una imagen ideal que tiene como centro a  $X_o$  y  $Y_o$  y a  $I$  como la distancia ideal entre los cuadros (Figura 3.15). De este modo se obtiene:

$$x' = X_o + nI \quad y \quad y' = Y_o + kI \quad (3.25)$$

donde  $n$  y  $k$  son elementos enteros.

Cualquier punto de un cuadro de la imagen estará a  $nI$  o  $kI$  intervalos del centro ideal, sabiendo que  $I$  es la distancia que hay entre cuadros y también la longitud propia de los lados de los cuadros. Se utiliza el método de mínimos cuadrados para relacionar ambas imágenes:

$$f(x, y, I, X_o, Y_o) = d^2 = (x - (X_o + nI))^2 + (y - (Y_o + kI))^2 \quad (3.26)$$

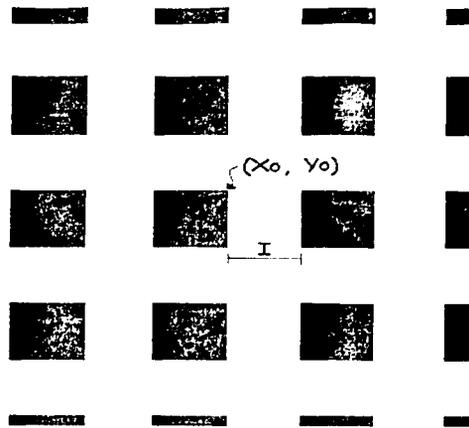


Figura 3.15: Parámetros de la imagen ideal

Para minimizar se deriva parcialmente con respecto a cada una de las variables, obteniéndose así las ecuaciones siguientes:

$$\frac{\partial f(x, y, I, X_o, Y_o)}{\partial X_o} = -2x + 2X_o + 2nI \quad (3.27)$$

$$\frac{\partial f(x, y, I, X_o, Y_o)}{\partial Y_o} = -2y + 2Y_o + 2kI \quad (3.28)$$

$$\frac{\partial f(x, y, I, X_o, Y_o)}{\partial I} = -2xn + 2X_on + 2n^2I - 2yk + 2Y_ok + 2k^2I. \quad (3.29)$$

Con ellas se puede calcular el centro de la imagen ideal y el intervalo  $I$ . Éstos son ahora los parámetros para crear la malla perfecta y, con base en ella, una malla que se acople a la imagen que se tiene. Es como una máscara que permitirá trabajar con las imágenes de la cámara, ya sin desplazamientos.

Después de esto se puede hacer el ajuste para encontrar los coeficientes del polinomio. Para ambas imágenes se obtienen tantas ecuaciones como puntos se seleccionen y se sustituyen en las ecuaciones propuestas de polinomios de segundo o cuarto orden. Resolviendo estas ecuaciones por mínimos cuadrados se encuentran las funciones que describen el mapeo. Con las cuatro ecuaciones resultantes se construyen mapas de transformación que registran la correspondencia entre cada pixel desde la imagen distorsionada hasta la nueva imagen corregida. Los mapas de transformación deben hacerse con algún tipo de interpolación (puntos de vecindad mínima, interpolación bilineal, aproximaciones, etc.) para mejorar la precisión. Los *puntos de la vecindad mínima* se obtienen por medio de la distancia mínima dentro del plano de la imagen, semejante a la distancia Euclídeana (Figura 3.16). Es decir, que a un punto dado se le asigna la intensidad del punto discreto que se encuentre más cercano.

Para el proceso de calibración, en este caso, se utiliza la de interpolación bilineal, que consiste en que dados los cuatro *puntos de vecindad mínima* de un punto  $(x', y')$  se les asigna un peso por medio de

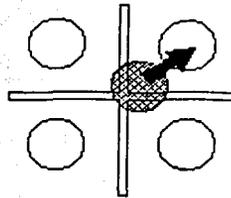


Figura 3.16: Punto vecino más próximo

la siguiente relación (Schalkoff 1990: pp. 56):

$$f(x', y') = (1-a)(1-b)f(x, y) + a(1-b)f(x+1, y) + (1-a)b f(x, y+1) + abf(x+1, y+1), \quad (3.30)$$

donde  $a$  y  $b$  son las distancias respectivas del punto  $(x', y')$  hasta el punto ideal de la rejilla  $(x, y)$ . El procedimiento anterior se puede resumir como una función tal que:

$$f(x', y') = g_s(\text{round}(x), \text{round}(y)). \quad (3.31)$$

Por tanto, el máximo error podrá ser sólo de medio pixel. La desventaja de esta interpolación es que es tajante y causa un efecto de ruptura. Con la interpolación bilineal se consigue desaparecer las rupturas creando promedios de intensidad y, aunque la imagen se vuelve un poco borrosa, la aproximación es más precisa. En la Figura 3.17 se muestran las imágenes corregidas con el polinomio de segundo grado.

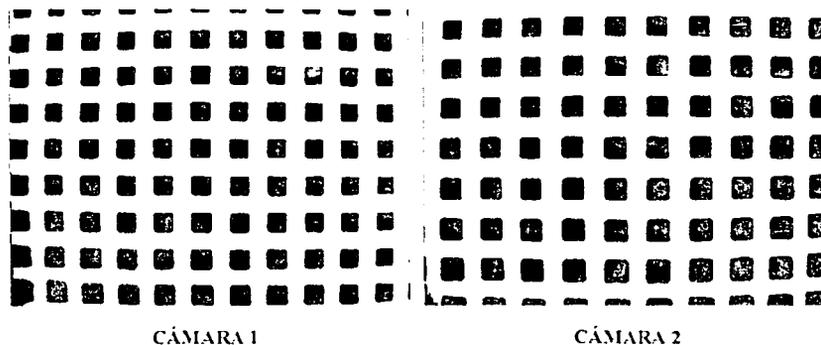


Figura 3.17: Imágenes sin distorsión

Como se puede observar, los resultados del polinomio de segundo orden son bastante buenos. Aunque la corrección en las imágenes de los polinomios de cuarto orden es mejor, se elige el polinomio de segundo orden porque representa menor tiempo de procesamiento y la utilización de menos memoria, tomando en cuenta que la corrección es de un nivel aceptable para los fines que se persiguen.

### 3.7.2 Estimación de la longitud focal

El procedimiento para estimar la longitud focal consiste en tomar imágenes de un marco ideal de calibración (Figure 3.18) en diferentes posiciones. Después de eliminar la distorsión producida por la lente se obtienen los bordes del marco con el detector de bordes Canny. Con estas nuevas imágenes se ajustan los bordes del marco de calibración y se determinan sus matrices de covarianza. Las matrices de covarianza se obtienen de la teoría de colinearidad de puntos o concurrencia de líneas (Kanatani 1995: pp. 20).

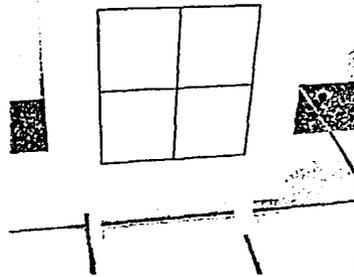


Figura 3.18: Imágenes del marco de calibración

Con las matrices de covarianza se calculan los puntos en el infinito (o *vanishing points*) con coordenadas homogéneas  $m$  y  $m'$ . Como estos puntos no caen necesariamente en la imagen capturada por el sensor de la cámara, la mejor manera de representarlos es con vectores normalizados. Como las coordenadas homogéneas normalizadas no proveen una expresión explícita para su equivalente en el plano de la imagen (porque todavía no se conoce el valor de la longitud focal), no se puede utilizar directamente una ecuación que relacione los puntos conjugados y la longitud focal. Como alternativa se expresan los puntos en términos de la razón de cambio del vector normalizado  $m$  al variar la longitud focal del sistema. Se obtiene así para  $m'$ , vector que representa un punto con respecto a  $f$  (mientras que  $m$  es el vector que representa al mismo punto con respecto a la longitud focal  $\hat{f}$ ):

$$m = \pm N \begin{bmatrix} m_1 \\ m_2 \\ \frac{f}{\hat{f}} m_3 \end{bmatrix} \quad (3.32)$$

Estas representaciones dan como resultado la ecuación siguiente (Kanatani 1995: pp. 59):

$$m_1 m'_1 + m_2 m'_2 + \left(\frac{f}{\hat{f}}\right)^2 m_3 m'_3 = 0, \quad (3.33)$$

donde  $m$  y  $m'$  son vectores de los puntos en el infinito de dos líneas ortogonales en el espacio, definidas con respecto a la  $\hat{f}$  estimada, mientras que  $f$  es la verdadera longitud focal. Así se obtiene:

$$f = \hat{f} \left( -\frac{m_1 m'_1 + m_2 m'_2}{m_3 m'_3} \right)^{\frac{1}{2}} \quad (3.34)$$

Para las cámaras del laboratorio se obtienen las siguientes distancias focales:

| Cámara   | Distancia focal |
|----------|-----------------|
| Cámara 1 | 814 pixeles     |
| Cámara 2 | 896 pixeles     |

Nótese que la precisión en el cálculo de la distancia focal también depende de la configuración espacial de las líneas usadas para determinar los puntos en el infinito. Mientras más lejos se encuentren del centro de la imagen, más impreciso será el resultado.

### 3.7.3 Determinación de la posición y orientación

La calibración concluye con la determinación de la situación espacial de las cámaras respecto a un sistema general fijo en la escena, con origen  $\hat{O}$ . Lo que realmente interesa en este caso, es que dado un punto  $P$  representado por  $(x_1, y_1)$  en la imagen de la cámara 1 y por  $(x_2, y_2)$  en la imagen de la cámara 2, se desea saber las coordenadas espaciales  $(\hat{X}, \hat{Y}, \hat{Z})$  con respecto al origen  $\hat{O}$ . Para ello es necesario conocer una matriz homogénea que transforme las coordenadas espaciales desde el marco de referencia de una cámara al sistema coordenado general. Las coordenadas espaciales  $(X_C, Y_C, Z_C)$  del punto de las imágenes se obtienen con las ecuaciones de la Sección 3.5.1. La representación matemática es entonces:

$$\begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \\ 1 \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{21} & \hat{A}_{31} & \hat{A}_{41} \\ \hat{A}_{12} & \hat{A}_{22} & \hat{A}_{32} & \hat{A}_{42} \\ \hat{A}_{13} & \hat{A}_{23} & \hat{A}_{33} & \hat{A}_{43} \\ \hat{A}_{14} & \hat{A}_{24} & \hat{A}_{34} & \hat{A}_{44} \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} \quad (3.35)$$

Para conocer los elementos de  $\hat{A}$  se sitúa a un objeto de calibración (Figura 3.19) en cierta posición conocida dentro del espacio (Figura 3.20).

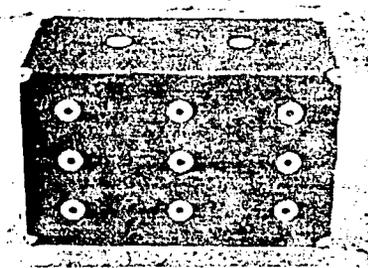


Figura 3.19: Caja de calibración

El objeto, al ser de calibración, tiene dimensiones conocidas y, por medio de dos imágenes (una de cada cámara) se hace la relación de puntos. La cantidad mínima de puntos que se utiliza es 9, debido a que la

matriz de transformación sólo requiere la solución para 9 incógnitas. En el caso de este trabajo se hace la calibración externa con 13 puntos, ya que 9 definen el plano del objeto más cercano a las cámaras y los 4 restantes son elementos que ayudan a calibrar la profundidad.

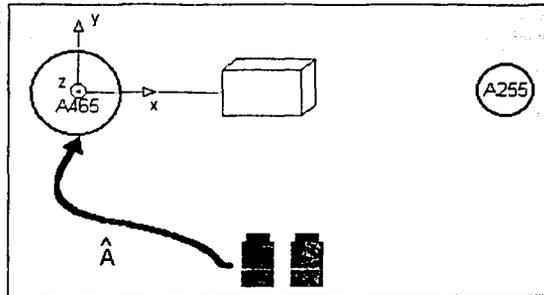


Figura 3.20: Calibración extrínseca

Con la posición conocida del objeto y sus dimensiones se obtiene una matriz de datos para el objeto de calibración. Una segunda matriz se obtiene con los datos de las imágenes. Es decir, que de la siguiente ecuación:

$$\hat{V} = \hat{A} V_C, \quad (3.36)$$

se conocen las matrices  $\hat{V}$  y  $V_C$ . Siendo  $\hat{V}$  de  $4 \times N$  y  $V_C$  de  $4 \times N$ , donde  $N$  es la cantidad de puntos que se desean utilizar. Conociendo éstas matrices se resuelve para  $\hat{A}$  y con ello se adquieren los 16 elementos de la matriz homogénea. Ésta debe cumplir con las propiedades de las matrices homogéneas que se vieron en la Sección 2.6.

Esta ecuación toma la forma (Kanatani 1995: pp. 58):

$$x_1 x_2 + y_1 y_2 + f^2 = 0. \quad (3.37)$$



## Capítulo 4

# Agarre estable de un objeto

Siempre recuerda que eres único.  
Como todos los demás.  
– Eddington

### 4.1 Cálculo del agarre estable

Alcanzar y/o agarrar un objeto representan diferentes metas para un manipulador, y con ello también distintas aplicaciones. Si se necesita sólo el alcance, lo que se pretende es un contacto momentáneo; pero una tarea de larga duración generalmente tendrá necesidad de un agarre o sujeción del objeto. Las metas del manipulador pueden ser (Skowronski 1986: pp. 129):

- Alcance por instantes
- Alcance en intervalos
- Captura o sujeción
- Seguimiento y
- Combinaciones de los anteriores.

En este trabajo se hará la manipulación de los objetos por medio de la captura o sujeción de los mismos. Sujetar un objeto rígido con dos robots manipuladores (dedos) no es una labor sencilla. Una vez que se tiene la descripción física del objeto y los puntos de agarre, se tienen que obtener las velocidades y fuerzas necesarias para agarrar realmente el objeto, asegurar que no se resbale, caiga ni realice giros indeseables. El problema del agarre estable consiste en determinar puntos de contacto que permitan la resistencia a fuerzas externas y la manipulación diestra del objeto. Para el análisis se asumirán dos cosas (Murray *et al.* 1994: pp. 214):

1. El objeto es un cuerpo rígido en contacto con un eslabón rígido del manipulador.
2. Se conoce el modelo del objeto y el de los dedos.

El contacto entre un dedo y el objeto se puede describir como un mapeo de la fuerza ejercida por el dedo en el punto de contacto y los pares resultantes en algún punto de referencia en el objeto. Normalmente se elige como punto de referencia el centro de gravedad del objeto. El contacto se describe con su posición y orientación relativas con respecto a un marco de referencia en el objeto mismo y este marco de referencia se puede conocer fácilmente con respecto a otro por medio de una transformación (Figura 4.1).

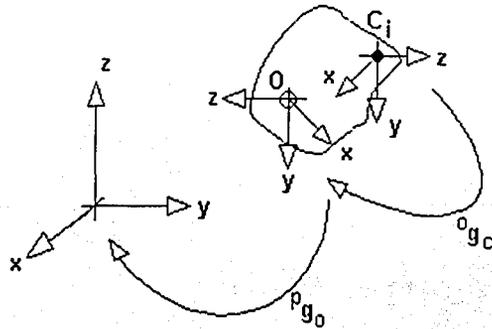


Figura 4.1: Ejes coordenados de los contactos

Según se puede observar en la Figura 4.1, el sistema coordenado  $C_i$  se conoce con respecto al marco de referencia  $O$ , gracias a la transformación  ${}^0g_{C_i}$ . La fuerza que aplica un contacto se modela como una fuerza-par<sup>1</sup>  $F_{C_i}$  aplicada en el origen del marco de contacto,  $C_i$ .

Primero se pensará en que el contacto de los dedos y el objeto es puntual, es decir, se ignorarán las posibilidades de que los dedos resbalen o de que el objeto gire. Si se supone que no existe fricción entre el dedo y el objeto, y que sólo se pueden aplicar fuerzas  $F_{C_i}$  en dirección normal a la superficie del objeto. Estas fuerzas se representan como (Murray *et al.* 1994: pp. 214):

$$F_{C_i} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} f_{C_i} \quad (4.1)$$

donde  $f_{C_i}$  es la magnitud de la fuerza aplicada, que es positiva porque empuja al objeto y no lo jala.

<sup>1</sup>Un par de fuerzas es un conjunto formado por dos fuerzas no colineales, tales que sean paralelas, de igual magnitud y de sentidos contrarios (Solar 1993: pp. 46).

Aunque tomar en cuenta sólo la fuerza normal a la superficie permite una representación sencilla del contacto del dedo con ella, conviene considerar la fricción entre los puntos de contacto. Para ello se utiliza un modelo que permite describir la cantidad de fuerza que se debe aplicar de forma tangente a una superficie, como función de la fuerza normal. La teoría de la fricción seca o fricción de Coulomb es empírica y predice las fuerzas de fricción máximas que pueden ejercerse entre superficies secas en contacto, ya sea en reposo o movimiento relativo (Bedford y Fowler 1996: pp. 504). Esta teoría establece que el deslizamiento comienza cuando (Murray *et al.* 1994: pp. 216):

$$|f^t| > \mu f^n, \quad (4.2)$$

donde  $f^t$  es la fuerza tangencial,  $f^n$  es la fuerza normal y  $\mu$  es el coeficiente estático de fricción. Por tanto, el intervalo de fuerzas tangenciales, de interés en nuestro caso es el que incluye todas las que sean menores o iguales a:  $\mu f^n$ . El conjunto de fuerzas que se pueden aplicar al contacto deben estar en un cono alrededor de la fuerza normal, llamado *cono de fricción*. El ángulo  $\alpha$  del cono con respecto a la normal está dado por:

$$\alpha = \arctan \mu, \quad (4.3)$$

como se muestra en la Figura 4.2. Puesto que los valores de  $\mu$  son generalmente menores que 1,  $\alpha$  es casi siempre menor que  $45^\circ$ .

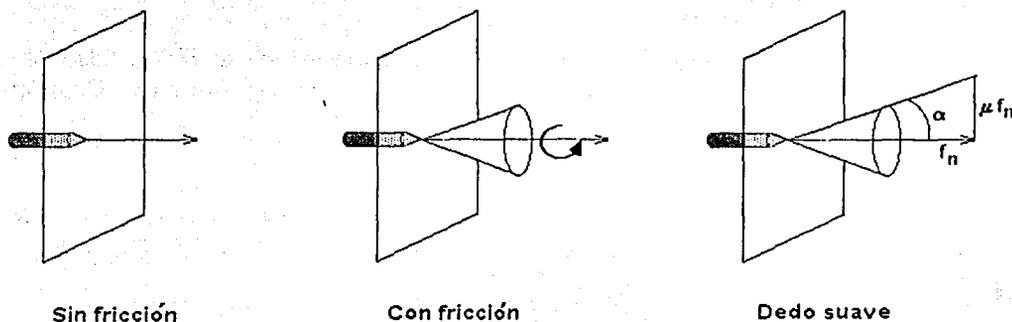


Figura 4.2: Interpretación geométrica de los modelos de fricción

Cuando se usa el modelo con fricción, se tiene que representar  $F_{C_i}$  con respecto a la base de las direcciones de la fricción, es decir:

$$F_{C_i} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_{C1} \\ f_{C2} \\ f_{C3} \end{bmatrix}. \quad (4.4)$$

$f_{C1}$ ,  $f_{C2}$  y  $f_{C3}$  deben ser todas positivas. Un modelo más realista es el de dedo suave (*soft finger*), donde se toman en cuenta también posibles giros alrededor de la normal y se modelan los pares como

limitados por un coeficiente de fricción torsional  $\gamma$ .

$$F_{C_i} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} f_{C_i} = B_{C_i} f_{C_i}, \quad (4.5)$$

donde  $f_{C_i} \in \mathbb{R}^4$  es elemento del cono de fricción  $FC_{C_i}$ :

$$FC_{C_i} = [f_c : (f_1^2 + f_2^2)^{\frac{1}{2}} \leq \mu f_3, |f_4| \leq \gamma f_3] \quad (4.6)$$

En general, se modela el contacto usando una base de fuerzas  $B_{C_i}$ , junto con el cono de fricción. La dimensión de la base de momentos  $m_i$  indica el número de fuerzas independientes que se pueden aplicar en el contacto. De ahí que se necesite que el cono de fricción satisfaga las siguientes propiedades:

1. El cono de fricción debe ser un subconjunto no vacío de  $\mathbb{R}^{m_i}$ .
2.  $f_1$  y  $f_2$ , elementos del cono, deben cumplir con la propiedad de cerradura para  $af_1 + bf_2$ , si  $a$  y  $b$  son mayores que cero.

Como se había mencionado, una propiedad importante del agarre estable es la habilidad de resistir giros externos del objeto aplicando los torques adecuados en los puntos de contacto. Una definición más formal es:

**Definición 4.1.1** Agarre estable (Murray et al. 1994: pp. 223): *Un agarre es estable si al aplicar cualquier fuerza-par externa  $F_e \in \mathbb{R}^p$  al objeto, existen fuerzas de contacto  $f_c \in FC$  que se encuentren dentro del cono de fricción.*

Las fuerzas internas se utilizan para asegurar que las fuerzas de contacto anulan a las fuerzas de fricción del cono, y no causan giros del objeto. Una fuerza interna es el conjunto de fuerzas de contacto, representada con respecto a la base de torque en cada contacto. La existencia de una fuerza interna es una condición necesaria para el agarre estable.

**Teorema 4.1.1** Agarre plano antipodal: *Un agarre plano con dos puntos de contacto con fricción es un agarre estable si y sólo si la línea que conecta ambos puntos de contacto está dentro de ambos conos de fricción.*

Para el caso que se estudia en esta tesis, se deben establecer las condiciones de un agarre estable en tres dimensiones. El siguiente teorema proporciona estas condiciones.

**Teorema 4.1.2** Agarre estable espacial: *Un agarre espacial es estable si y sólo si la línea que conecta a los puntos de contacto de los dos dedos suaves está dentro de ambos conos de fricción.*

## 4.2 Cálculo de los centros de masa

El peso de un objeto no actúa en un solo punto sino que está distribuido sobre su volumen total. Sin embargo, este peso se puede representar con una sola fuerza equivalente actuando en un punto llamado centro de masa (Bedford y Fowler 1996: pp. 333). Para poder hallar este punto se debe encontrar el centroide del volumen total, en el caso de que el objeto sea de densidad uniforme. Por ello, primero se debe establecer un sistema coordenado, donde la coordenada del centroide ( $\bar{x}$ ) de volumen de  $x$  estará dada por:

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}, \quad (4.7)$$

donde  $N$  es el número total de puntos que se utiliza para el cálculo. Las coordenadas medias  $\bar{y}$  y  $\bar{z}$  pueden calcularse de la misma forma. Cuando se tienen pesos asociados a ciertos puntos se utiliza:

$$\bar{x} = \frac{\sum_{i=1}^N c_i x_i}{\sum_{i=1}^N c_i}. \quad (4.8)$$

Esta ecuación sirve para determinar la posición media de cualquier conjunto de cantidades a las que se les pueda asociar una posición. La posición media de esta ecuación es el centroide, es decir el peso asociado con las posiciones  $x_i$ ,  $y_i$ ,  $z_i$  y  $c_i$ .

Cuando se consideran volúmenes se obtienen las siguientes ecuaciones para la posición del centroide (Bedford y Fowler 1996: pp. 339):

$$\bar{x} = \frac{\int_V x dV}{\int_V dV} \quad (4.9)$$

$$\bar{y} = \frac{\int_V y dV}{\int_V dV} \quad (4.10)$$

$$\bar{z} = \frac{\int_V z dV}{\int_V dV} \quad (4.11)$$

El centro de masa de un cuerpo es el centroide de su masa:

$$\bar{x} = \frac{\int_m x dm}{\int_m dm} \quad (4.12)$$

$$\bar{y} = \frac{\int_m y dm}{\int_m dm} \quad (4.13)$$

$$\bar{z} = \frac{\int_m z dm}{\int_m dm} \quad (4.14)$$

donde  $x$ ,  $y$  y  $z$  son las coordenadas del elemento diferencial de masa  $dm$  (Figura 4.3). La integración se hace sobre toda la masa del cuerpo.

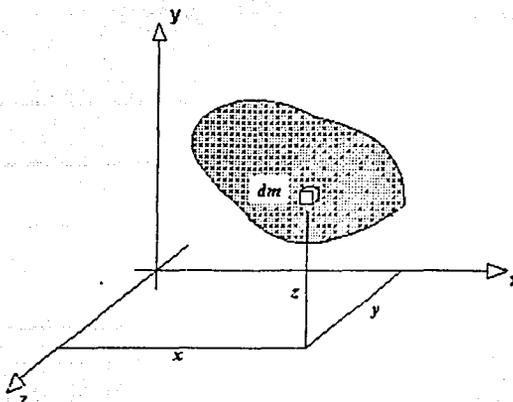


Figura 4.3: Elemento diferencial  $dm$  de un cuerpo

Como se mencionó al inicio de esta sección: el peso de un cuerpo se puede representar con una sola fuerza equivalente que actúa en su centro de masa. Por ello es de interés sólo la fuerza y el momento totales ejercidos por el peso del cuerpo.

Un punto interesante es cuando se considera la densidad  $\rho$ .

$$m = \int_m dm = \int_V \rho dV. \quad (4.15)$$

Para un cuerpo homogéneo se tiene:

$$m = \rho \int_V dV = \rho V. \quad (4.16)$$

En este trabajo se manipulan sólo cuerpos homogéneos y simples, por lo que el centro de masa de ellos siempre coincide con el centroide de su volumen. La teoría para cuerpos no homogéneos y compuestos es similar, pero necesita de un desarrollo un poco más extenso.

### 4.3 Sistema de dos robots industriales

Los robots industriales pueden realizar labores verdaderamente complejas; se componen de tres partes fundamentales:

1. El manipulador
2. El controlador
3. Los sistemas sensitivos o de medición

El sistema sensitivo que se maneja en el presente trabajo es de visión, aunque hay también sistemas de sonido, fuerza, encoders, etc. Este apartado se refiere a los manipuladores, mientras que el Capítulo 5 trata de la parte del control.

Un manipulador en general consta de las siguientes partes:

1. Elementos rígidos o eslabones que están conectados mediante uniones a elementos de movimiento, de forma translacional o rotacional.
2. Elementos de movimiento o articulaciones, que son sistemas motores con dispositivos neumáticos, hidráulicos o eléctricos, que proporcionan la energía mecánica.
3. Dispositivos de agarre que poseen la capacidad de orientar, operar y sujetar las piezas manipuladas.

Los robots A255 y A465 del laboratorio se componen de cinco y seis elementos de movimiento, lo que los hace ser de cinco y seis grados de libertad, respectivamente (Figura 4.4). Esto significa que son cinco o seis las variables de articulaciones que es necesario conocer para determinar la posición del robot.

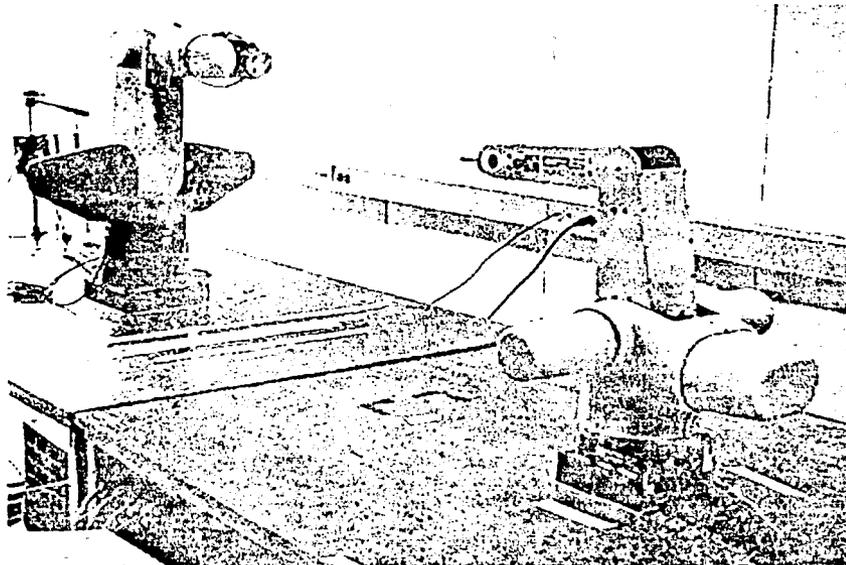


Figura 4.4: Robots A465 y A255 del Laboratorio

Cuando se operan los dos robots conjuntamente<sup>2</sup> no es posible aún controlar todas las articulaciones de ambos debido a la falta de líneas de comunicación (Castillo 2002). Este problema se trata con más detalle en el capítulo siguiente. Debido a esta limitante se decide tener control sobre tres articulaciones de cada uno, para un total de seis. La elección de las articulaciones a controlar es libre, pero por conveniencia se controlan las primeras tres.

<sup>2</sup>Los robots están conectados a una computadora personal por medio de dos tarjetas de adquisición de datos Flex Motion de National Instruments (Apéndice C)

## 4.4 Cinemática de los robots

### 4.4.1 Cinemática Directa

La manipulación se reduce siempre a giros y traslaciones realizados en un espacio tridimensional. Por ello, para poder trabajar con los manipuladores del laboratorio es necesario conocer su cinemática. Se puede hacer un análisis con un marco de referencia arbitrario en cada eslabón, pero un trabajo sistemático simplifica en mucho los cálculos. En lo que se refiere a la robótica, comúnmente se utiliza la convención Denavit - Hartenberg (Spong y Vidyasagar 1989: pp. 65). Ellos propusieron en 1955 la utilización del álgebra matricial para describir la geometría espacial de cualquier elemento con respecto a un sistema de referencia fijo. La idea general es asignar los sistemas coordenados de cada articulación siguiendo algunas reglas preestablecidas; una vez fijos los sistemas de coordenadas, cada transformación homogénea entre los eslabones  $i - 1$  e  $i$  se representa como el producto de cuatro transformaciones básicas:

$${}^{i-1}A_i = Rot_{z,\theta} Trans_{z,d_i} Trans_{x,a_i} Rot_{x,\alpha_i} \quad (4.17)$$

donde  ${}^{i-1}A_i$  es la transformación homogénea desde el marco de referencia  $i - 1$  al  $i$ ,  $Rot_{z,\theta}$  es la rotación  $\theta$  alrededor del eje  $z$ ,  $Trans_{z,d_i}$  es la traslación de una distancia  $d_i$  en dirección del eje  $z$ ,  $Trans_{x,a_i}$  es la traslación de una distancia  $a_i$  en dirección del eje  $x$  y  $Rot_{x,\alpha_i}$  es la rotación  $\alpha_i$  alrededor del eje  $x$  (ver Figura 4.5). Llevando a cabo el producto se obtiene

$${}^iA = \begin{bmatrix} c\theta_i & -s\theta_i c\alpha_i & s\theta_i s\alpha_i & a_i c\theta_i \\ s\theta_i & c\theta_i c\alpha_i & -c\theta_i s\alpha_i & a_i s\theta_i \\ 0 & s\alpha_i & c\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.18)$$

Los parámetros  $\alpha_i$ ,  $\theta_i$ ,  $d_i$  y  $a_i$  reciben generalmente nombres característicos:

- $\alpha_i$ : ángulo de torsión, que es el ángulo entre  $z_{i-1}$  y  $z_i$  medido desde  $x_i$ ;
- $\theta_i$ : ángulo entre eslabones, que es el ángulo entre  $x_{i-1}$  y  $x_i$  medido desde  $z_{i-1}$  (este ángulo es variable si la articulación es rotacional);
- $d_i$ : desplazamiento o distancia entre eslabones, que es la distancia medida a lo largo del eje  $z_{i-1}$  desde el origen  $O_{i-1}$  a la intersección de los ejes  $x_i$  y  $z_{i-1}$  (es variable si la articulación  $i$  es prismática) y
- $a_i$ : longitud, que es la distancia medida a lo largo del eje  $x_i$  desde  $O_i$  hasta la intersección de los ejes  $x_i$  y  $z_{i-1}$ .

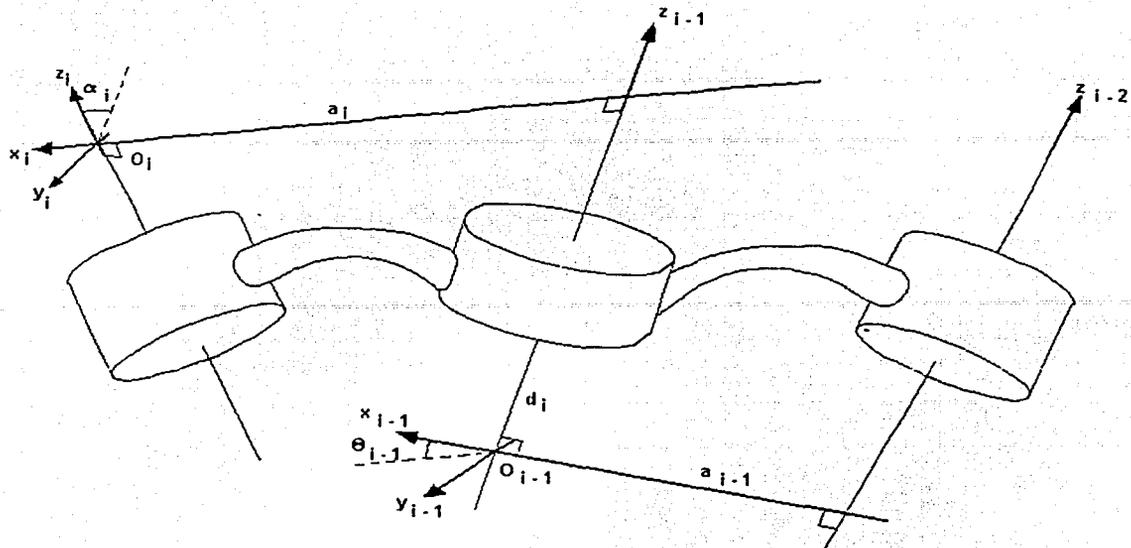


Figura 4.5: Parámetros de la asignación Denavit-Hartenberg

*Robot de 6 grados de libertad:*

Cuando se aplica el algoritmo D-H para describir al manipulador industrial de seis grados de libertad del laboratorio (el robot de la izquierda en la Figura 4.4), los ejes se colocan como se muestra en la Figura 4.6. Se crea a continuación la tabla de parámetros de eslabonamiento:

| Articulación | Tipo       | $a_i$ [cm] | $\alpha_i$  | $d_i$ [cm] | $\theta_i$ |
|--------------|------------|------------|-------------|------------|------------|
| 1            | Rotacional | 0          | $90^\circ$  | 33         | $\theta_1$ |
| 2            | Rotacional | -30.5      | $0^\circ$   | 0          | $\theta_2$ |
| 3            | Rotacional | 0          | $90^\circ$  | 0          | $\theta_3$ |
| 4            | Rotacional | 0          | $-90^\circ$ | 33         | $\theta_4$ |
| 5            | Rotacional | 0          | $90^\circ$  | 0          | $\theta_5$ |
| 6            | Rotacional | 0          | $0^\circ$   | $d_6$      | $\theta_6$ |

La magnitud de  $d_6$  depende de la longitud del efector final con el que se esté trabajando. Con esta tabla se calculan las matrices de transformación  ${}^{i-1}A_i$  (en metros) con lo que se obtiene:

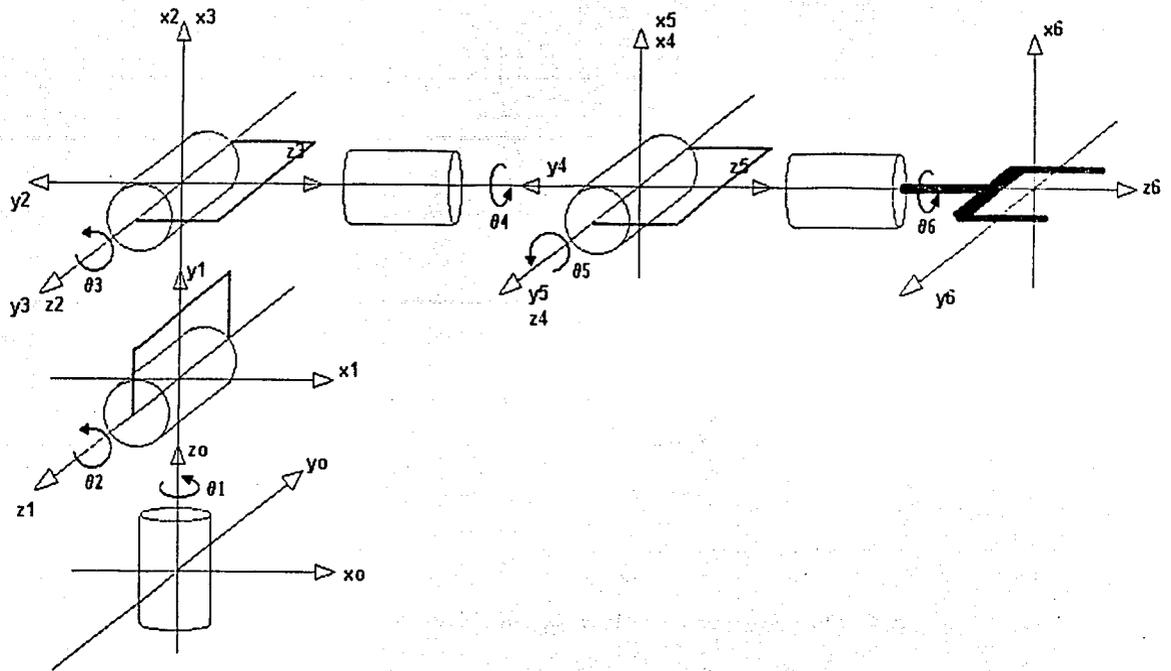


Figura 4.6: Marcos de referencia del robot A465

$${}^0A_1 = Rot_{z,\theta_1} = \begin{bmatrix} c_{\theta_1} & 0 & s_{\theta_1} & 0 \\ s_{\theta_1} & 0 & -c_{\theta_1} & 0 \\ 0 & 1 & 0 & 0.33 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.19)$$

$${}^1A_2 = \begin{bmatrix} c_{\theta_2} & -s_{\theta_2} & 0 & -0.305c_{\theta_2} \\ s_{\theta_2} & c_{\theta_2} & 0 & -0.305s_{\theta_2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.20)$$

$${}^2A_3 = \begin{bmatrix} c_{\theta_3} & 0 & s_{\theta_3} & 0 \\ s_{\theta_3} & 0 & -c_{\theta_3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.21)$$

$${}^3A_4 = \begin{bmatrix} c\theta_4 & 0 & -s\theta_4 & 0 \\ s\theta_4 & 0 & c\theta_4 & 0 \\ 0 & -1 & 0 & 0.33 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.22)$$

$${}^4A_5 = \begin{bmatrix} c\theta_5 & 0 & s\theta_5 & 0 \\ s\theta_5 & 0 & -c\theta_5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} y \quad (4.23)$$

$${}^5A_6 = \begin{bmatrix} c\theta_6 & -s\theta_6 & 0 & 0 \\ s\theta_6 & c\theta_6 & 0 & 0 \\ 0 & 0 & 1 & d_6 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.24)$$

Finalmente se forma la matriz de transformación del manipulador de 6 grados de libertad:

$${}^0T_6 = {}^0A_1 {}^1A_2 {}^2A_3 {}^3A_4 {}^4A_5 {}^5A_6. \quad (4.25)$$

*Robot de 5 grados de libertad:*

El robot A255 es el que se muestra del lado derecho en la Figura 4.4. Para calcular la cinemática directa de éste se tiene que elegir una posición inicial o de reposo (por ejemplo, se escoge la misma que en el manipulador anterior), tomando en cuenta que ambos robots se encuentran cara a cara. El procedimiento para conocer las matrices de transformación no es el mismo que en el caso anterior, porque con el algoritmo D - H toma como primicia que cada variable de articulación se mide con respecto al eslabón inmediato anterior. Éste no es el caso para el robot A255, dado que el ángulo de las articulaciones 3 y 4 se mide con respecto al plano de la horizontal. Por ello, las matrices de transformación homogénea se tienen que calcular de forma convencional, por inspección. La elección de los sistemas de coordenadas se muestra en la Figura 4.7.

Los sistemas coordenados son arbitrarios. Por facilidad en el caso de las articulaciones 1 y 2 se mantiene la convención del manipulador A465, pero a partir de la articulación 3 ya no se puede proceder de la misma forma. Se obtienen las siguientes matrices de transformación:

$${}^0A_1 = Rot_{z,\theta_1} = \begin{bmatrix} c\theta_1 & 0 & s\theta_1 & 0 \\ s\theta_1 & 0 & -c\theta_1 & 0 \\ 0 & 1 & 0 & 0.254 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.26)$$

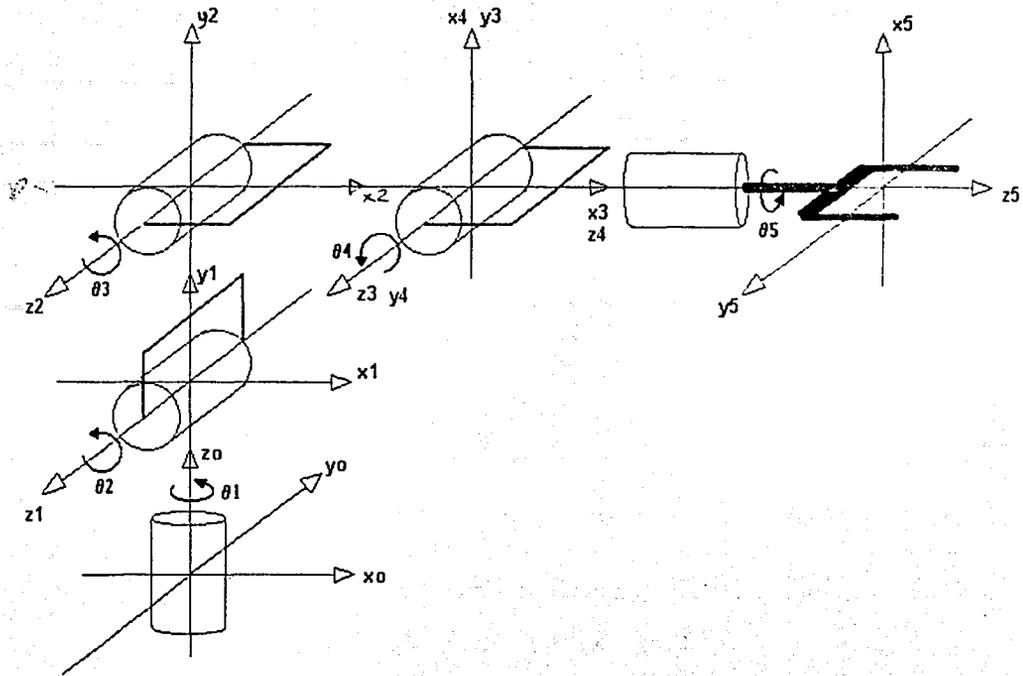


Figura 4.7: Marcos de referencia del robot A255

${}^1A_2$  representa una rotación alrededor del eje  $z_1$ , donde se tiene que tomar en cuenta la distancia desde el origen, igual que para el manipulador de seis grados de libertad. Sin embargo, a diferencia del caso anterior, el sistema de coordenadas 2 no cambia su orientación con respecto al sistema de coordenadas 1, por lo que la matriz de rotación es la identidad.

$${}^1A_2 = \begin{bmatrix} 1 & 0 & 0 & 0.254c_{\theta_2} \\ 0 & 1 & 0 & 0.254s_{\theta_2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.27)$$

Igualmente,

$${}^2A_3 = \begin{bmatrix} 1 & 0 & 0 & 0.254s_{\theta_3} \\ 0 & 1 & 0 & 0.254c_{\theta_3} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.28)$$

Finalmente,

$${}^3A_4 = \begin{bmatrix} -s_{\theta_4} & 0 & c_{\theta_4} & 0 \\ c_{\theta_4} & 0 & s_{\theta_4} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.29)$$

y

$${}^4A_5 = \begin{bmatrix} c_{\theta_5} & -s_{\theta_5} & 0 & 0 \\ s_{\theta_5} & c_{\theta_5} & 0 & 0 \\ 0 & 0 & 1 & d_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.30)$$

Se puede calcular la matriz de transformación desde el marco de referencia base hasta el efector final:

$${}^0T_5 = {}^0A_1 {}^1A_2 {}^2A_3 {}^3A_4 {}^4A_5. \quad (4.31)$$

#### 4.4.2 Cinemática Inversa

La cinemática directa de un manipulador permite conocer la posición y orientación del efector final de un robot con respecto a la base y como función de los ángulos de las articulaciones. Sin embargo, cuando se desea llevar al manipulador a una posición y orientación deseada, surge el problema de la cinemática inversa. En este caso se desea conocer los ángulos que corresponden a una posición y orientación deseadas (Spong y Vidyasagar 1989: pp. 65). Se puede plantear el problema general de la cinemática inversa como sigue: dada una transformación homogénea

$$H = \begin{bmatrix} R & -d \\ 0 & 1 \end{bmatrix}, \quad (4.32)$$

se tienen que encontrar una o todas las soluciones de la ecuación:

$${}^0T_n(\theta_1, \theta_2, \dots, \theta_n) = H \quad (4.33)$$

donde

$${}^0T_n(\theta_1, \theta_2, \dots, \theta_n) = {}^0A_1 {}^1A_2 \dots {}^{n-1}A_n. \quad (4.34)$$

Esta ecuación se puede escribir como:

$$({}^0T_n)_{ij}(\theta_1, \theta_2, \dots, \theta_n) = h_{ij} \text{ con } i = 1, 2, 3 \quad j = 1, 2, 3, 4$$

donde  ${}^0T_n$  y  $h_{ij}$  se refieren a las 12 soluciones no triviales de  ${}^0T_n$  y  $H$ , respectivamente. Las cuatro soluciones restantes se refieren al renglón de  $(0,0,0,1)$ .

Como se trabaja con dos manipuladores se necesita conocer la cinemática inversa de ambos.

## Análisis del manipulador con seis grados de libertad

En este caso se tiene la matriz de transformación dada por (4.25).

$${}^0T_6 = \begin{bmatrix} r_{11} & r_{12} & r_{13} & d_x \\ r_{21} & r_{22} & r_{23} & d_y \\ r_{31} & r_{32} & r_{33} & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (4.35)$$

que se obtuvo con la cinemática directa. Como ya se explicó, para la cinemática inversa se tienen que encontrar las variables  $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$  y  $\theta_6$  correspondientes a una posición y orientación arbitraria. Para ello se tendrían que resolver 12 ecuaciones trigonométricas simultáneas, lo que es muy difícil, sobre todo si se intenta obtener la solución de forma cerrada. Sin embargo, puesto que el robot A465 es un manipulador de seis grados de libertad con una muñeca esférica, se puede separar el problema en dos más sencillos (desacoplamiento cinemático). El efector final se puede manejar como una muñeca esférica, donde los ejes  $z_4, z_5$  y  $z_6$  se intersecan en  $O$ . Se denota a  $p_c$  como el vector desde el marco base de referencia hasta el centro de la muñeca (Figura 4.8). Para tener al efector final del manipulador en la posición final  $d$  con la orientación dada  $R = (r_{ij})$  ( $i, j = 1, 2, 3$ ), es necesario y suficiente colocar el centro de la muñeca  $O$  en el punto  $p_c$ :

$$p_c = d - d_6 R k. \quad (4.36)$$

Esto se obtiene de  $O_6 - O = d_6 R k$ , donde el origen  $O_6$  es una translación  $d_6$  sobre el eje  $z_5$  desde  $O$ .

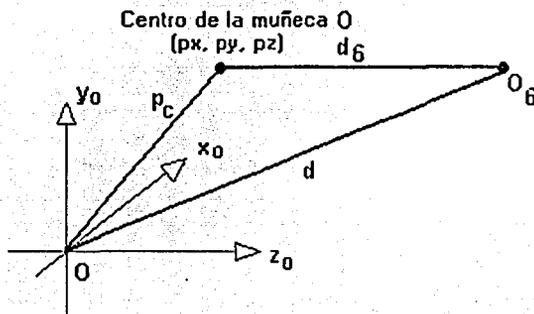


Figura 4.8: Vector  $p_c$

Por tanto, si a las componentes de  $d$ , se les denota como  $d_x, d_y$  y  $d_z$ , y a las de  $p_c$ , como  $p_x, p_y$  y  $p_z$ , se obtiene la siguiente relación:

$$p_c = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} = \begin{bmatrix} d_x - d_6 r_{13} \\ d_y - d_6 r_{23} \\ d_z - d_6 r_{33} \end{bmatrix}. \quad (4.37)$$

La relación anterior significa que la posición del centro de la muñeca no variará si cambian las variables  $\theta_1$ ,  $\theta_2$  y  $\theta_3$ . Es decir, estos tres valores dependen exclusivamente de  $p_c$ , que es dato del problema. Por ello se pueden calcular independientemente de  $\theta_4$ ,  $\theta_5$  y  $\theta_6$ . Una vez calculados estos valores, se puede determinar la transformación  ${}^0R_3$ . Como la orientación deseada debe satisfacer:

$$R = {}^0R_3 {}^3R_6, \quad (4.38)$$

se obtiene

$${}^3R_6 = ({}^0R_3)^{-1} R = ({}^0R_3)^T R. \quad (4.39)$$

Los ángulos finales se encontrarán con el conjunto de ángulos de Euler correspondientes a  ${}^3R_6$  ( $\theta_4$ ,  $\theta_5$  y  $\theta_6$ ).

Como los 3 primeros ángulos dependen exclusivamente de  $p_c$ , se puede determinar su valor por medio de relaciones trigonométricas. Para ello se proyecta  $p_c$  en el plano  $x_0 - y_0$ , como se muestra en la Figura 4.9. Se concluye que

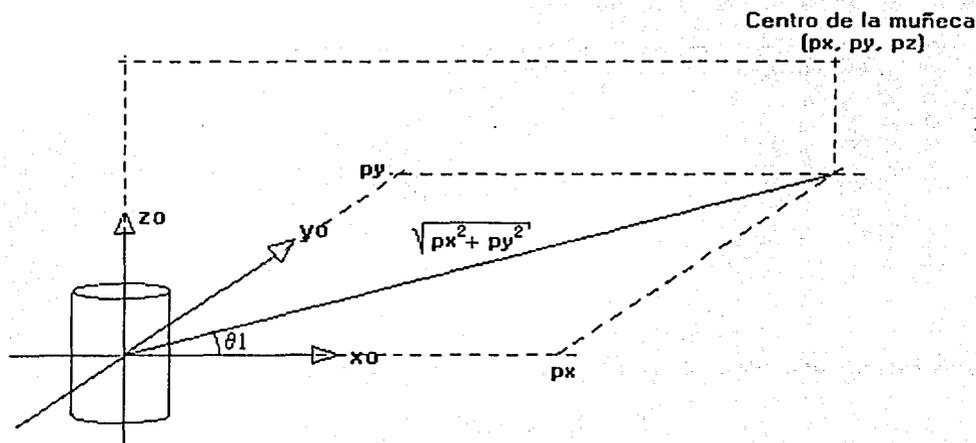


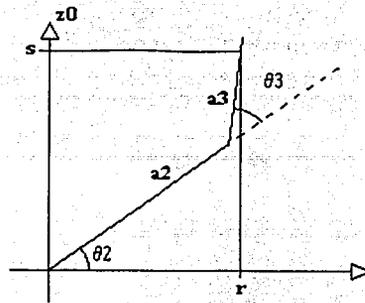
Figura 4.9: Proyección de  $p_c$

$$\theta_1 = \arctan(p_y, p_x) \quad (4.40)$$

donde  $\arctan(y, x)$  está definida para todos los pares  $(y, x)$  diferentes de cero e iguales a un ángulo único  $\theta$ , tal que:

$$\cos \theta = \frac{x}{(x^2 + y^2)^{\frac{1}{2}}} \quad y \quad \sin \theta = \frac{y}{(x^2 + y^2)^{\frac{1}{2}}}. \quad (4.41)$$

Una segunda solución válida puede ser:  $\theta_1 = \pi + \arctan(p_y, p_x)$ , tomando en cuenta que la solución  $\theta_2$  correspondiente debe ser  $\pi + \theta_2$ . Estas soluciones para  $\theta_1$  son válidas excepto si  $p_x = p_y = 0$ , ya que

Figura 4.10: Ángulo  $\theta_3$ 

en este caso hay un número infinito de soluciones. Para encontrar los ángulos  $\theta_2$ ,  $\theta_3$  dado que se tiene  $\theta_1$  se considera el plano formado por el segundo y tercer eslabón:

Se sabe que:

$$D = \cos \theta'_3 = \frac{r^2 + s^2 - a_2^2 - a_3^2}{2a_2a_3} = \frac{p_x^2 + p_y^2 + (p_z - d_1)^2 - a_2^2 - d_4^2}{2a_2d_4} \quad (4.42)$$

donde

$$r^2 = p_x^2 + p_y^2 \quad (4.43)$$

y

$$s = p_z - d_1 \quad (4.44)$$

Entonces  $\theta'_3$  estará dada por:

$$\theta'_3 = \arctan(\pm(1 - D^2)^{\frac{1}{2}}, D). \quad (4.45)$$

Las dos soluciones corresponden a posiciones del codo hacia arriba o hacia abajo, respectivamente. De forma similar se tiene  $\theta_2$ :

$$\theta_2 = \arctan(p_z - d_1, (p_x^2 + p_y^2)^{\frac{1}{2}}) - \arctan(d_4 \sin \theta'_3, a_2 + d_4 \cos \theta'_3). \quad (4.46)$$

Finalmente se puede decir que:

$$\theta_3 = \theta'_3 + \frac{\pi}{2}. \quad (4.47)$$

Vale la pena resaltar que  $\theta'_3$  no se calcula directamente con la ecuación (4.45) porque ésta asume que para  $\theta_2 = 0$  y  $\theta'_3 = 0$  se tiene alineación de los dos eslabones. Sin embargo, como se muestra en la Figura 4.11, suponer  $\theta_2 = 0$  y  $\theta_3 = 0$  implica que los eslabones asociados forman un ángulo de  $90^\circ$  (ver ecuación (4.47)).

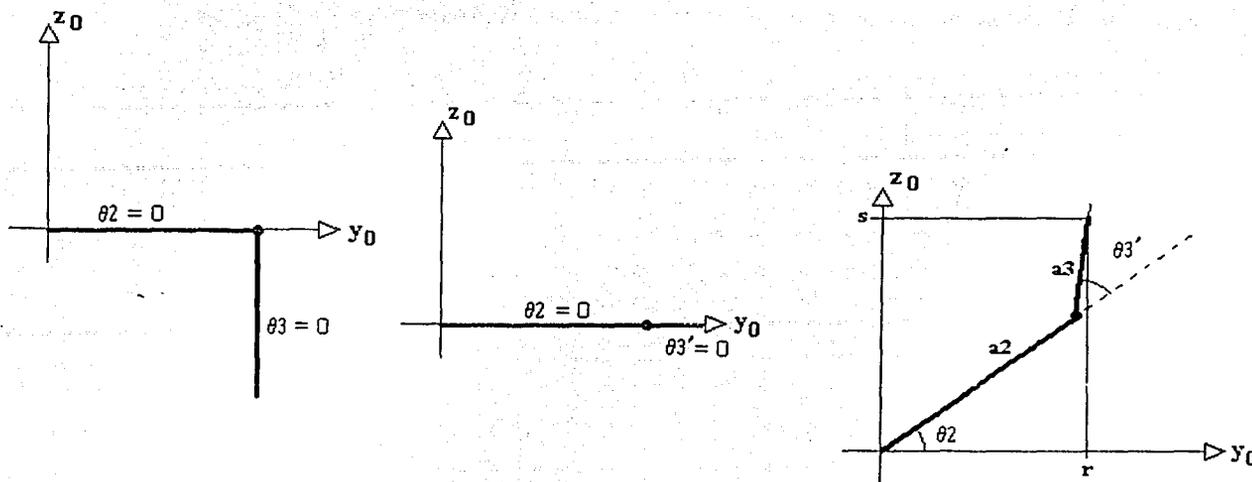


Figura 4.11: Ángulos reales del robot

Una vez que se conocen los valores de  $\theta_1$ ,  $\theta_2$  y  $\theta_3$ , se procede a calcular  $\theta_4$ ,  $\theta_5$  y  $\theta_6$  como los ángulos de Euler correspondientes a las matriz de rotación  $({}^0R_3)^T R_6$ , dada por la ecuación (4.39). El método de los ángulos de Euler consiste en hallar una matriz de rotación en términos de cantidades independientes. Los ángulos de Euler permiten especificar la orientación de un marco de referencia 1 relativo a otro marco de referencia 0 por medio de los ángulos  $(\theta, \phi, \psi)$ , que se obtiene de tres rotaciones sucesivas:

1. Alrededor del eje  $z$  por un ángulo  $\phi$
2. Alrededor del eje  $y$  actual por un ángulo  $\theta$
3. Alrededor del eje  $z$  actual por un ángulo  $\psi$

Con rotaciones básicas se obtiene:

$$\begin{aligned}
 {}^0R_1 &= R_{z,\phi} R_{y,\theta} R_{z,\psi} & (4.48) \\
 &= \begin{bmatrix} c_\phi & -s_\phi & 0 \\ s_\phi & c_\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_\theta & 0 & s_\theta \\ 0 & 1 & 0 \\ -s_\theta & 0 & c_\theta \end{bmatrix} \begin{bmatrix} c_\psi & -s_\psi & 0 \\ s_\psi & c_\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} c_\phi c_\theta c_\psi - s_\phi s_\psi & -c_\phi c_\theta s_\psi - s_\phi c_\psi & c_\phi s_\theta \\ s_\phi c_\theta c_\psi + c_\phi s_\psi & -s_\phi c_\theta s_\psi + c_\phi c_\psi & s_\phi s_\theta \\ -s_\theta c_\psi & -s_\theta s_\psi & c_\theta \end{bmatrix}
 \end{aligned}$$

Se puede tomar esta matriz para el análisis, ya que lo que se desea es  ${}^3R_6$ . Si se compara esta matriz

con la que se obtuvo en la sección de cinemática directa, es decir con

$$U = {}^3R_6 = \begin{bmatrix} c_{\theta_4}c_{\theta_5}c_{\theta_6} - s_{\theta_4}s_{\theta_6} & -c_{\theta_4}c_{\theta_5}s_{\theta_6} - s_{\theta_4}c_{\theta_6} & c_{\theta_4}s_{\theta_5} \\ s_{\theta_4}c_{\theta_5}c_{\theta_6} + c_{\theta_4}s_{\theta_6} & -s_{\theta_4}c_{\theta_5}s_{\theta_6} + c_{\theta_4}c_{\theta_6} & s_{\theta_4}s_{\theta_5} \\ -s_{\theta_5}c_{\theta_6} & -s_{\theta_5}s_{\theta_6} & c_{\theta_5} \end{bmatrix} \quad (4.49)$$

se observa que son idénticas, por lo que se concluye que los ángulos de Euler corresponden precisamente al problema que se analiza.

Los elementos de  $U$  se pueden definir como  $u_{ij}$ , donde  $i, j = 1, 2, 3$ . Se puede suponer que los elementos  $u_{13}$  y/o  $u_{23}$  son diferentes de cero, de ahí que  $s_{\theta}$  es diferente de cero y que  $u_{33}$  es diferente de  $\pm 1$ . De lo anterior se concluye que:

$$c_{\theta} = u_{33} \quad (4.50)$$

y que:

$$s_{\theta} = \pm(1 - u_{33}^2)^{\frac{1}{2}}. \quad (4.51)$$

Por tanto se obtiene

$$\theta = \arctan(\pm(1 - u_{33}^2)^{\frac{1}{2}}, u_{33}). \quad (4.52)$$

Si se elige el valor positivo, se obtiene:

$$s_{\theta} > 0$$

$$\phi = \arctan(u_{23}, u_{13}) \quad (4.53)$$

$$\psi = \arctan(u_{32}, -u_{31}).$$

Si se elige el valor negativo, se obtiene:

$$s_{\theta} < 0$$

$$\phi = \arctan(-u_{23}, -u_{13}) \quad (4.54)$$

$$\psi = \arctan(-u_{32}, u_{31}).$$

Con la comparación de matrices anterior se tiene la siguiente correspondencia:  $\theta = \theta_5$ ,  $\phi = \theta_6$  y  $\psi = \theta_4$ . En este trabajo se eligió  $s_{\theta} > 0$ , debido a la geometría del robot, por lo que se tiene:

$$\theta_4 = \arctan(u_{23}, u_{13}) \quad (4.55)$$

$$\theta_5 = \arctan((1 - u_{33}^2)^{\frac{1}{2}}, u_{33}) \quad (4.56)$$

$$\theta_6 = \arctan(u_{32}, -u_{31}). \quad (4.57)$$

Por otro lado, si  $u_{13} = u_{23} = 0$ , se tendrá que  $u_{31} = u_{32} = 0$ , porque es ortogonal. Asimismo, se tendrá  $u_{33} = \pm 1$ . De esta forma se tiene:

$$U = {}^3R_6 = \begin{bmatrix} c_{\theta_4}c_{\theta_5}c_{\theta_6} - s_{\theta_4}s_{\theta_6} & -c_{\theta_4}c_{\theta_5}s_{\theta_6} - s_{\theta_4}c_{\theta_6} & 0 \\ s_{\theta_4}c_{\theta_5}c_{\theta_6} + c_{\theta_4}s_{\theta_6} & -s_{\theta_4}c_{\theta_5}s_{\theta_6} + c_{\theta_4}c_{\theta_6} & 0 \\ 0 & 0 & \pm 1 \end{bmatrix} \quad (4.58)$$

Dado que  $s_{\theta_5} = 0$ . Si  $u_{33} = 1$ , entonces  $c_{\theta_5} = 1$  y  $\theta_5 = 0$ . Además es fácil concluir que  $\theta_4 = 0$  y  $\theta_6 = \arctan(-u_{12}, u_{11})$ . Si  $u_{33} = -1$  y  $c_{\theta_5} = -1$ , entonces  $\theta_5 = \pi$ . De la misma forma se obtiene  $\theta_4 = 0$  y  $\theta_6 = -\arctan(-u_{12}, -u_{11})$ .

En resumen, las variables de las articulaciones para el robot A465 son:

| Angulo      |   |
|-------------|---|
| $\theta_1$  | $\arctan(p_y, p_x)$   |
| $\theta_2$  | $\arctan(p_z - d_1, (p_x^2 + p_y^2)^{\frac{1}{2}}) - \arctan(d_4 \sin \theta'_3, a_2 + d_4 \cos \theta'_3)$ |
| $\theta'_3$ | $\arctan(\pm(1 - D^2)^{\frac{1}{2}}, D)$  |
| $\theta_3$  | $\theta'_3 + \frac{\pi}{2}$   |
| $\theta_4$  | $\arctan(u_{23}, u_{13})$   |
| $\theta_5$  | $\arctan((1 - u_{33}^2)^{\frac{1}{2}}, u_{33})$   |
| $\theta_6$  | $\arctan(u_{32}, -u_{31})$  |

#### Análisis del manipulador de cinco grados de libertad

El análisis de este manipulador lleva al mismo resultado para el ángulo  $\theta_1$  como se desprende de la Figura 4.12. Es decir, se obtiene:

$$\theta_1 = \arctan(p_y, p_x). \quad (4.59)$$

Sin embargo, para calcular los ángulos  $\theta_2$  y  $\theta_3$  se debe utilizar otro procedimiento, ya que el ángulo  $\theta_3$  no se mide con respecto al eslabón 2, sino con respecto al plano horizontal  $x_0 - y_0$ . Esto se muestra en la Figura 4.13. Claramente se tiene

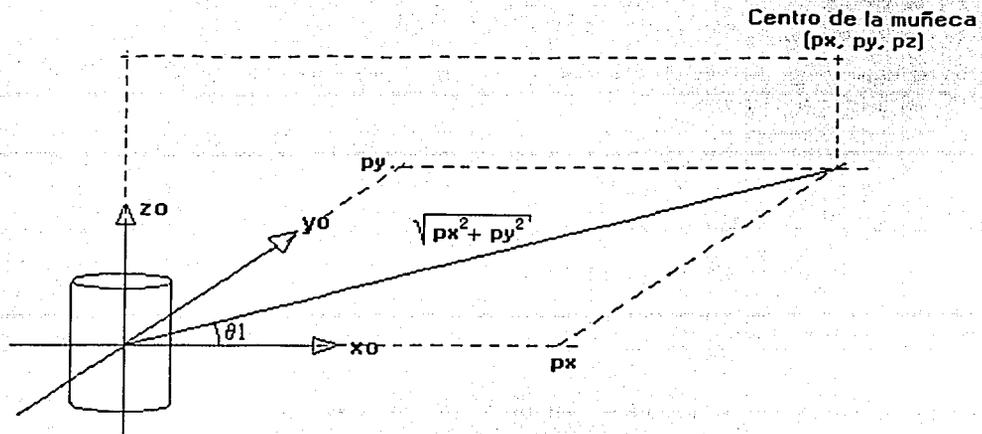
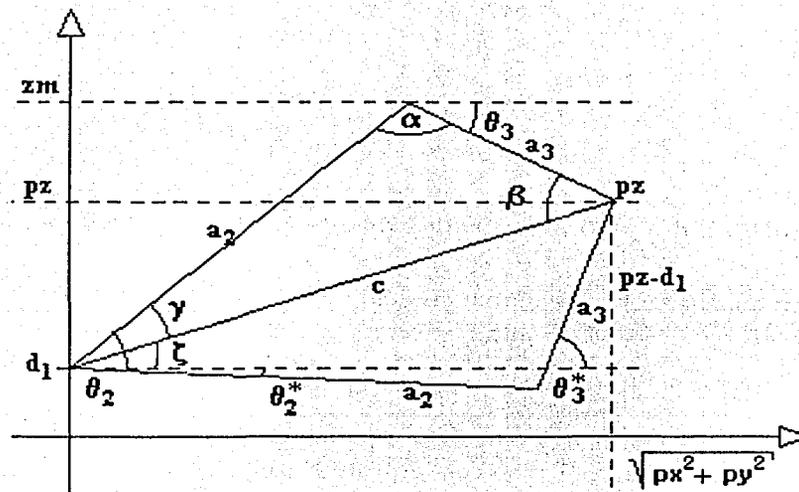
$$\theta_2 = \gamma + \xi \quad (4.60)$$

donde

$$a_3^2 = a_2^2 + c^2 - 2a_2c \cos \gamma \quad (4.61)$$

$$\cos \gamma = \frac{a_2^2 + c^2 - a_3^2}{2a_2c} \quad (4.62)$$

$$\xi = \arctan \frac{p_z - d_1}{(p_x^2 + p_y^2)^{\frac{1}{2}}}. \quad (4.63)$$

Figura 4.12: Proyección de  $p_c$ Figura 4.13: Ángulos  $\theta_2$  y  $\theta_3$ 

Entonces

$$\theta_2 = \arccos \frac{a_2^2 + c^2 - a_3^2}{2a_2c} + \arctan \frac{p_z - d_1}{p_x^2 + p_y^2} \quad (4.64)$$

Además se tiene que:

$$z_m = a_2 \sin \theta_2 + d_1. \quad (4.65)$$

Entonces se puede formar la Figura 4.14, de donde se puede apreciar que:

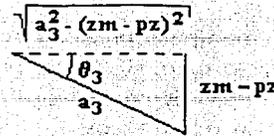


Figura 4.14: Ángulo  $\theta_3$  y  $z_m$

$$\theta_3 = \arctan(p_z - z_m, (a_3^2 - (p_z - z_m)^2)^{\frac{1}{2}}). \quad (4.66)$$

La orientación del efector final relativo al marco de referencia está dada por:

$$R = {}^0R_3 {}^3R_5, \quad (4.67)$$

o bien,

$${}^3R_5 = {}^0R_3^{-1} R = {}^0R_3^T R. \quad (4.68)$$

En este caso  ${}^3R_5$  depende de  $\theta_4$  y  $\theta_5$ , por lo que no es posible tener una orientación arbitraria. Ahora bien, desarrollando  ${}^3R_5$  se tiene

$${}^3R_5 = \begin{bmatrix} -s_{\theta_4} c_{\theta_5} & s_{\theta_4} s_{\theta_5} & c_{\theta_4} \\ c_{\theta_4} c_{\theta_5} & -c_{\theta_4} s_{\theta_5} & s_{\theta_4} \\ s_{\theta_5} & c_{\theta_5} & 0 \end{bmatrix}. \quad (4.69)$$

Es necesario encontrar los valores de los ángulos  $\theta_4$  y  $\theta_5$ . Claramente se observa que el problema tendrá solución si y sólo si el elemento  $u_{33}$  de  $U = {}^0R_3^T R$  es cero. Si este es el caso, se puede seguir exactamente el mismo procedimiento que el utilizado para la muñeca del robot de seis grados de libertad, con lo que se obtiene:

$$\theta_5 = \arctan(u_{31}, u_{32}) \quad \text{o} \quad \theta_5 = \arctan(-u_{31}, -u_{32}), \quad (4.70)$$

$$\theta_4 = \arctan(u_{23}, u_{13}) \quad \text{o} \quad \theta_4 = \arctan(-u_{23}, -u_{13}). \quad (4.71)$$

En resumen, para el manipulador de cinco grados de libertad se tienen los ángulos que siguen:

| Ángulo     |   |
|------------|---|
| $\theta_1$ | $\arctan(p_y, p_x)$   |
| $\theta_2$ | $\arccos \frac{a_2^2 + c^2 - a_1^2}{2a_2c} + \arctan(p_z - d_1, (p_x^2 + p_y^2)^{\frac{1}{2}})$ |
| $\theta_3$ | $\arctan(p_z - z_m, (a_3^2 - (p_z - z_m)^2)^{\frac{1}{2}})$                                     |
| $\theta_4$ | $\arctan(u_{23}, u_{13})$   |
| $\theta_5$ | $\arctan(u_{31}, u_{32})$   |

## 4.5 Trabajo simultáneo

Hecho el análisis cinemático de ambos manipuladores por separado, el paso siguiente consiste en que ambos hagan referencia a un mismo marco inercial para el trabajo simultáneo. En principio este marco o sistema de coordenadas es arbitrario, pero es conveniente la elección de alguno de los dos sistemas base de los robots. La relación entre estos dos marcos de referencia se muestra en la Figura 4.15.

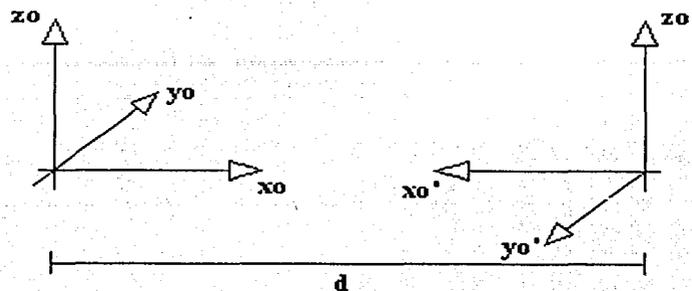


Figura 4.15: Relación de marcos de referencia

Cualquier punto, con coordenadas  $(x'_0, y'_0, z'_0)$  en el sistema del robot A255 se puede expresar en el sistema de la base del robot A465 por medio de la siguiente transformación homogénea:

$${}^0R_{0'} = \begin{bmatrix} -1 & 0 & 0 & d \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.72)$$

La transformación inversa está dada por:

$${}^{0'}R_0 = \begin{bmatrix} -1 & 0 & 0 & d \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.73)$$

Es decir, ambas transformaciones son iguales. Para poder utilizar un espacio de trabajo común, se elige el sistema a la base del robot A465. En adelante, cualquier trayectoria será referenciada a este marco de referencia. Por tanto, para trabajar con el robot A255 se utilizará la transformación  ${}^{0'}R_0$ .

Para poder calcular las trayectorias que deben seguir los efectores finales de los robots para sujetar un objeto, considérese la Figura 4.16.

Las coordenadas  $(x_{01}, y_{01}, z_{01})$  representan la posición del efector final del manipulador 1 en el tiempo  $t = 0$ , lo mismo que las coordenadas  $(x_{02}, y_{02}, z_{02})$  para el manipulador 2. Los puntos  $p_c$  son los puntos

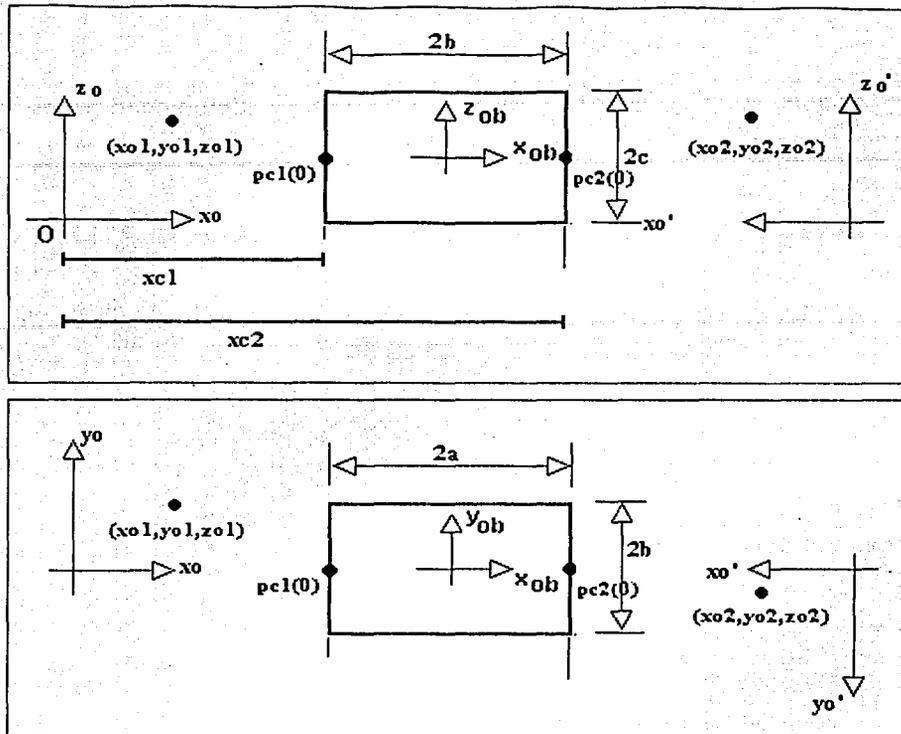


Figura 4.16: Objeto del agarre desde dos perspectivas

de contacto para el agarre. Para sujetar el objeto hay que mover  $p_{0i}$ , que es el punto espacial donde está el efector final en  $t=0$ , hasta  $p_{ci}(0)$ . Se debe generar de una trayectoria suave, es decir, una trayectoria que comience y termine con velocidad y aceleración angular nulas. El movimiento tendrá que realizarse en  $t_c$  segundos. De forma conveniente se puede elegir un polinomio de quinto orden cuyos coeficientes se calculen para satisfacer las posiciones inicial y final deseadas (para  $t = 0$  y  $t = t_c$ , respectivamente), además de las condiciones de velocidad y aceleración. El análisis es similar para  $x_{di}$ ,  $y_{di}$  y  $z_{di}$  con  $i = 1, 2$ . Por ejemplo, para  $x_{di}$  se tiene:

$$x_{di} = ax_{0i} + ax_{1i}t + ax_{2i}t^2 + ax_{3i}t^3 + ax_{4i}t^4 + ax_{5i}t^5. \quad (4.74)$$

La velocidad y la aceleración son:

$$v_{di} = \dot{x}_{di} = ax_{1i} + 2ax_{2i}t + 3ax_{3i}t^2 + 4ax_{4i}t^3 + 5ax_{5i}t^4 \quad (4.75)$$

$$a_{di} = \dot{v}_{di} = \ddot{x}_{di} = 2ax_{2i} + 6ax_{3i}t + 12ax_{4i}t^2 + 20ax_{5i}t^3, \quad (4.76)$$

respectivamente. Las condiciones iniciales son:

$$x_{di}(0) = x_{0i} \quad (4.77)$$

$$\dot{x}_{di}(0) = 0 \quad (4.78)$$

$$\ddot{x}_{di}(0) = 0 \quad (4.79)$$

lo que significa que el manipulador está en su posición inicial  $x_{0i}$ . Como condiciones finales se eligen:

$$x_{di}(t_c) = x_{ci} \quad (4.80)$$

$$\dot{x}_{di}(t_c) = 0 \quad (4.81)$$

$$\ddot{x}_{di}(t_c) = 0. \quad (4.82)$$

Es decir, que cuando el manipulador llegue a su posición final  $x_{ci}$  al tiempo  $t_c$ , no deberá tener ya velocidad ni aceleración alguna. De las ecuaciones anteriores se puede calcular:

$$a_{x_{0i}} = x_{0i} \quad (4.83)$$

$$a_{x_{1i}} = 0 \quad (4.84)$$

$$a_{x_{2i}} = 0 \quad (4.85)$$

Entonces se tiene, hasta el momento:

$$x_{di} = x_{0i} + ax_{3i}t^3 + ax_{4i}t^4 + ax_{5i}t^5 \quad (4.86)$$

$$v_{di} = \dot{x}_{di} = 3ax_{3i}t^2 + 4ax_{4i}t^3 + 5ax_{5i}t^4 \quad (4.87)$$

$$a_{di} = \dot{v}_{di} = \ddot{x}_{di} = 6ax_{3i}t + 12ax_{4i}t^2 + 20ax_{5i}t^3. \quad (4.88)$$

Para calcular los coeficientes restantes se forma el siguiente sistema matricial:

$$B a = x$$

donde

$$B = \begin{bmatrix} t_c^3 & t_c^4 & t_c^5 \\ 3t_c^2 & 4t_c^3 & 5t_c^4 \\ 6t_c & 12t_c^2 & 20t_c^3 \end{bmatrix}, \quad (4.89)$$

$$a = \begin{bmatrix} a_{x_{3i}} \\ a_{x_{4i}} \\ a_{x_{5i}} \end{bmatrix} \quad (4.90)$$

y

$$x = \begin{bmatrix} x_{ci} - x_{0i} \\ 0 \\ 0 \end{bmatrix}. \quad (4.91)$$

Con este sistema es posible calcular los tres coeficientes faltantes:

$$a = B^{-1}x$$

Como se mencionó anteriormente,  $y_{di}$  y  $z_{di}$  se pueden calcular de manera similar. Cabe hacer notar que, como  $t_c$  es igual en todos los casos,  $B$  y  $B^{-1}$  sólo se tienen que calcular una vez.

Después de que ya se ha sujetado el objeto, es necesario mantenerlo detenido para que no caiga. La forma más fácil de lograr esto es que los efectores finales de los robots se mantengan paralelos todo el tiempo. Si, por ejemplo, se quiere hacer un movimiento circular con el objeto, se procede como sigue. Para hacer un círculo en el plano  $y_o - z_o$  se necesita seguir la trayectoria siguiente (Figura 4.18):

$$(y - y_r)^2 + (z - z_r)^2 = r^2, \quad (4.92)$$

donde  $r$  es el radio del círculo y  $(y_r, z_r)$  su centro. Resulta claro que:

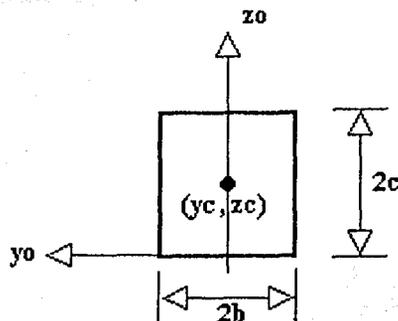


Figura 4.17: Objeto modelo

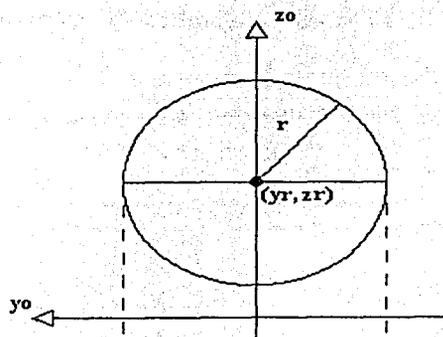


Figura 4.18: Círculo-Trayectoria de radio  $r$

$$y_r - r \leq y \leq y_r + r \quad (4.93)$$

Cualquier valor de  $y$  que satisfaga (4.93) se puede elegir como trayectoria deseada. Por ejemplo, se puede hacer:

$$y_d = y_r + r \sin f(t), \quad (4.94)$$

donde  $f(t)$  puede ser cualquier función del tiempo.  $z_d$  debe calcularse para satisfacer la ecuación (4.92), es decir,

$$z_d = z_r \pm (r^2 - (y_d - y_r)^2)^{\frac{1}{2}}, \quad (4.95)$$

o bien

$$z_r = z_d + r \cos f(t). \quad (4.96)$$

Lo mismo se puede hacer si se quisiera variar en el eje  $x$ :

$$x_d = x_r + r \cos f(t). \quad (4.97)$$

Es posible moverse desde  $p_c$ , el punto inicial de contacto con el objeto, hasta  $p_r$ , el centro del círculo, con la misma trayectoria polinomial utilizada anteriormente (haciendo los cambios correspondientes en las condiciones iniciales). Las trayectorias para el robot A255 son básicamente las mismas, sólo la coordenada  $x$  aumenta en un factor  $2a$  como se muestra en la Figura 4.16. De esta forma, el punto donde se iniciará el movimiento circular estará dado por:

$$x_{r2} = x_{r1} + 2a \quad (4.98)$$

$$y_{r2} = y_{r1} \quad (4.99)$$

$$z_{r2} = z_{r1} \quad (4.100)$$

Para evitar que el objeto choque con el suelo se necesita que:

$$z_r \geq r + (b^2 + c^2)^{\frac{1}{2}} \quad (4.101)$$

De la misma forma, las trayectorias deseadas de ambos efectores finales se relacionan por:

$$x_{d2} = x_{d1} + 2a \quad (4.102)$$

$$y_{d2} = y_{d1} \quad (4.103)$$

$$z_{d2} = z_{d1} \quad (4.104)$$

## Capítulo 5

# Sujeción de un objeto con dos robots industriales

Todo buen científico no repite un mismo experimento en condiciones diferentes esperando un mismo resultado.  
– Wilde, O.

### 5.1 Control PID de los robots industriales

El problema de control de manipuladores trabajando aisladamente es el de determinar las entradas requeridas para cada articulación a fin de que el efector final llegue a la posición deseada, siguiendo una trayectoria predeterminada. Estas entradas pueden ser voltaje, corriente, fuerza, torque, flujo neumático, flujo hidráulico, etc. Normalmente el movimiento del efector final es una trayectoria determinada por una secuencia de posiciones y orientaciones.

En el caso de robots sujetando un objeto, las trayectorias deben ser elegidas de tal forma que éste no caiga. Esto implica a su vez que no sólo se deben determinar las trayectorias deseadas para los efectores finales de los robots, sino que también debe determinarse la fuerza necesaria para sostener el objeto.

El diseño de un controlador que contemple la dinámica de ambos robots sujetando a un objeto y las fuerzas internas que se derivan del agarre del mismo está fuera del alcance de este trabajo. En lugar de ello, se programará un controlador para cada manipulador como si ambos trabajaran aisladamente y se elegirán trayectorias para los efectores finales que permitan sujetar un objeto cuando los dos robots sean puestos a trabajar simultáneamente.

La dinámica de un robot rígido de  $n$  grados de libertad es no lineal (Arteaga Pérez 1998). El acoplamiento entre la dinámica de los  $n$  actuadores del robot se da por medio de fuerzas gravitacionales, centrífugas, de Coriolis, etc., así como por medio de una matriz de inercia. Cuando el robot se mueve a altas velocidades, el efecto de la dinámica no lineal y del acoplamiento de la dinámica de las articulaciones es muy grande y no puede despreciarse, por lo que debe de ser tomado en cuenta en el diseño del controlador. Sin embargo, cuando los actuadores son motores de un radio de engranaje grande y el movimiento que debe realizar el robot es lento, es posible diseñar un controlador para cada actuador sin tomar en cuenta efectos de acoplamiento con el resto del sistema. Debido a que este es el caso particular del presente trabajo, se hará la implementación de controladores para cada articulación que sólo tomen en cuenta la variable de la misma articulación. Como se mencionó, la trayectoria deseada para cada articulación se elegirá de tal forma que los dos robots simultáneamente manipulen un objeto.

Se pueden aplicar muchas técnicas y metodologías para el control de los motores. Tanto el método elegido como factores mecánicos (fricción, deslizamientos, etc.) tienen un impacto importante en la eficiencia del motor, y con ello en las diferentes aplicaciones que se le puede dar al manipulador. Lo mejor será siempre elegir una estrategia de control robusta con respecto a los errores en las condiciones iniciales, ruido de los sensores y errores del modelo. Para llevar a cabo el control de articulaciones independientes, cada eje del manipulador se controla por un sistema de entrada-salida, y cualquier efecto debido al movimiento de otros ejes se ignora o se maneja como una perturbación. La estructura básica de un sistema de control con retroalimentación de una entrada - una salida (SISO) es:

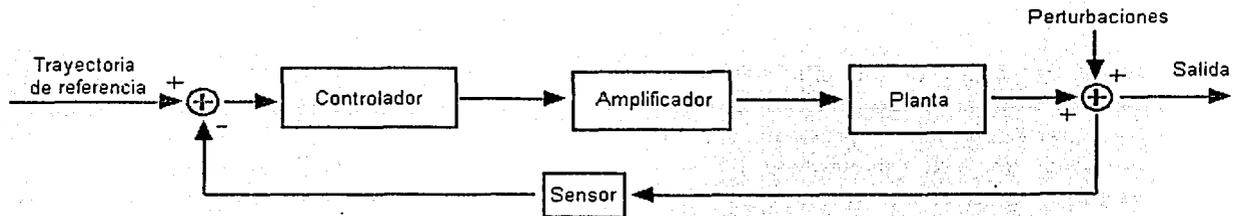


Figura 5.1: Estructura básica de un sistema de control con retroalimentación

El controlador debe elegirse de tal forma que la salida de la planta siga una referencia deseada. La señal de control no es la única señal que actúa en el sistema, ya que las perturbaciones también tienen influencia en la salida. Por ello es necesario que en el diseño se tomen en cuenta los efectos de las perturbaciones en la planta.

## 5.2 Diseño del sistema de control

El enfoque tradicional de diseño es el de emplear la retroalimentación de las señales de salida hacia las señales de referencia para construir señales de error y tratar de minimizarlas. En los sistemas de control de datos discretos existen muchas formas de implementar controladores: en el dominio de la frecuencia  $z$  o de variables de estado. El diseño tradicional de sistemas de control se caracteriza por el concepto de configuración fija, en la que se selecciona primero la composición de todo el sistema, incluyendo el proceso controlado y controlador. En la práctica existen muchas configuraciones para realizar el control de los procesos:

- Sistema de control digital con controlador de datos continuos en cascada
- Sistema de control digital con controlador digital en cascada
- Sistema de control digital con controlador de datos continuos en una trayectoria de retroalimentación
- Sistema de control digital con controlador digital en una trayectoria de retroalimentación

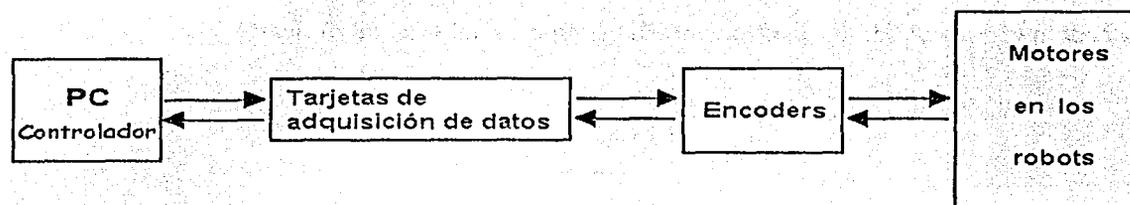


Figura 5.2: Sistema de control del laboratorio

La selección de la configuración se hace según la facilidad de implementación de la misma. En este trabajo se maneja el sistema de control digitalizado con el controlador en cascada (Figura 5.5). Se decidió utilizar este tipo de sistema porque encontrar una función de transferencia para la planta es realmente difícil. Como se ha mencionado, la planta o proceso que se quiere controlar es cada uno de los motores de las diferentes articulaciones del robot (3 para cada robot).

Los controladores digitales representan la manera más versátil de compensar un sistema de control. Se pueden implementar con redes digitales, computadoras, PDS o microcontroladores. El controlador digital efectúa operaciones lineales sobre una secuencia de valores y genera una segunda secuencia. La función de transferencia de un controlador digital se describe mediante:

$$D(z) = \frac{b_m z^m + b_{m-1} z^{m-1} + \dots + b_0}{a_n z^n + a_{n-1} z^{n-1} + \dots + a_0}, \quad (5.1)$$

donde  $m$  y  $n$  son enteros positivos.

La realizabilidad física de cualquier controlador digital implica que no aparezca alguna señal de salida antes de recibir una entrada, es decir, que la función de transferencia del controlador satisfaga la condición de que  $m$  sea menor o igual a  $n$ . Otra forma de escribir la ecuación (5.1) cuando el controlador tiene el mismo número de polos y ceros es (Kuo 1997: pp. 475):

$$D(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_m z^{-m}}{a_0 + a_1 z^{-1} + \dots + a_n z^{-n}}. \quad (5.2)$$

En el análisis se utilizan retenedores de orden cero (ROC) porque no existen muestreadores ideales físicamente. El ROC tiene como función de transferencia:

$$\mathcal{Z}\{G_{ROC}\} = \mathcal{Z}\left\{\frac{1 - e^{-sT}}{s}\right\} = (1 - z^{-1})\mathcal{Z}\left\{\frac{1}{s}\right\} = (1 - z^{-1})\frac{z}{z - 1} = 1. \quad (5.3)$$

En este caso particular se recibe la señal de los encoders de cada uno de los motores (en forma de trenes de pulsos). Esta señal entra a una tarjeta de adquisición de datos (Flex-Motion) y se transmite al controlador programado en LabWindows. El controlador programado maneja los errores y define el tiempo de muestreo (que en general es precisamente la duración del mismo programa controlador). Para obtener los errores se hace una comparación con una señal de referencia.

Para cada motor se utiliza un controlador PID (proporcional-integral-derivativo) como el que sigue:

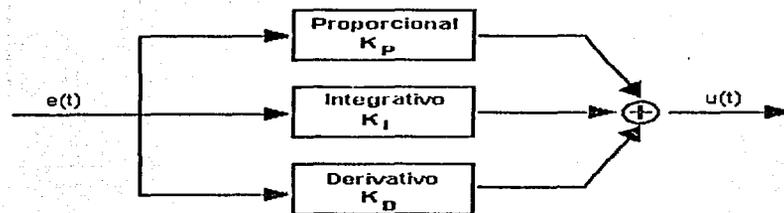


Figura 5.3: Controlador digital PID

Este controlador actúa sobre las señales de error  $e(t)$ : Para la acción proporcional, la señal de error  $e(k)$  sólo se multiplica por su respectiva ganancia  $K_P$ , que es constante:

$$R_P = K_P e(k). \quad (5.4)$$

En el caso de la acción derivativa se toma en cuenta una acción anticipativa (que reduce las oscilaciones y el sobreimpulso). Para su función de transferencia se toma el estado de la señal de error antes y después de cierto tiempo, obteniéndose así:

$$R_D = K_D \frac{e(k) - e(k - 1)}{h}, \quad (5.5)$$

donde  $h$  es el periodo de muestreo. Finalmente, la función integrativa es proporcionar una acción que disminuya el error bajo  $e(k)$ . La función de transferencia que corresponde al esquema de integración con base en la transformación trapezoidal es (Kuo 1997: pp. 482):

$$R_I = K_I R_I(k-1) + K_I \frac{h(e(k) + e(k-1))}{2}, \quad (5.6)$$

siendo  $e(k)$  la señal de error en un tiempo  $t = kh$  y  $e(k-1)$  la señal de error en un tiempo anterior  $t = (k-1)h$ .

Las constantes del controlador PID se pueden obtener utilizando un análisis matemático sólo en el caso de que la planta del sistema sea bien conocida. Como este no es el caso, se obtuvieron experimentalmente utilizando el método de Ziegler-Nichols (Åström y Wittenmark 1990).

El programa de LabWindows genera las salidas codificadas que por medio de las tarjetas de adquisición de datos se convierten en voltajes. Éstos atraviesan una etapa de amplificación y llegan a los motores de las articulaciones correspondientes.

Un problema fundamental en todo el proceso de control digital es el de mantener bajo el tiempo de muestreo, que idealmente debe tender a cero. Debido al tipo de tarjetas de adquisición disponibles en el laboratorio, el tiempo de muestreo se aproxima a 6 ms. Este tiempo debe ser suficiente para controlar seis motores de forma simultánea. Los 6 ms pueden parecer demasiado grandes para un sistema no lineal como un manipulador mecánico, pero no es en realidad una desventaja en este trabajo, porque los movimientos de ambos robots serán siempre relativamente lentos.

## 5.3 Reconocimiento de un objeto por medio de un sistema de visión

Un sistema de reconocimiento automático de formas opera siempre con un universo de trabajo previamente definido. De hecho, la visión computacional se apoya mucho en el trabajo de inteligencia artificial, pues es esta área la que se ocupa de obtener la definición de los universos, de ampliarla y de automatizarla. Los universos se pueden formar por redes neuronales, lenguajes y gramáticas formales, reglas de producción, etc. El universo está formado por todos los objetos que se reconocerán y estos objetos son a los que se les denomina patrones. Los objetos aparecen normalmente mezclados entre sí, y la operación de extracción de los objetos individuales es la segmentación, como se describió en la Sección 3.4.

### 5.3.1 Coincidencia de patrones

Si el objeto que se quiere localizar es de forma y tamaño totalmente conocidos, *aparentemente* se puede hallar de manera fácil con una máscara, pero por desgracia en la práctica no sucede así. El objeto puede

tener distinta orientación con respecto a la máscara o, por distorsiones leves de la imagen, puede diferir de ella. Para solucionar el problema existen métodos de localización de formas, como la transformada de Hough expuesta anteriormente o herramientas de la inteligencia artificial.

### 5.3.2 Espacio de características

Para iniciar el reconocimiento de los objetos se utiliza una transformación que crea un vector  $x$  cuyas componentes se denominan *características*. Cada vector de características se compara con un conjunto de vectores preestablecidos, compuesto por los vectores de propiedades de los objetos del universo de trabajo. Los objetos individuales se convierten en vectores numéricos  $x$  antes de ser reconocidos (*diferenciación*).

Una vez calculado el vector  $x$  asociado a un objeto individual, su reconocimiento automático se basa en determinar su grado de semejanza con los vectores prototipos. La primera etapa del proceso consiste en el establecimiento de los patrones del sistema. La siguiente es la elección del vector de características. Esta etapa es crítica, porque todo depende de los rasgos elegidos, además de que éste es el elemento clave en el sistema de reconocimiento (Maravall 1993: pp. 14). En general, las características seleccionadas tienen que ser:

1. Discriminantes
2. Fiables
3. No correlacionadas
4. Calculables en tiempos "aceptables" y
5. Observables con sensores económicos

Una vez seleccionado el vector de características del sistema, la etapa final se centra en el cálculo de las correspondientes funciones discriminantes. Las funciones discriminantes sirven para separar las diferentes secciones que corresponden a cada grupo. Son regionalizaciones del espacio abstracto de los patrones, basadas en el uso de funciones de distancia y consisten en combinaciones lineales o no lineales de las características. Para el caso de  $n$  rasgos discriminantes, una función lineal es:

$$f(x) = W_1X_1 + W_2X_2 + \dots + W_nX_n + W_{n+1} = \mathbf{W}^T \mathbf{X} \quad (5.7)$$

donde  $w$  es el vector de coeficientes de discriminación y  $x$  es el vector de características ampliado.

$$\mathbf{w} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \\ W_{n+1} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \\ 1 \end{bmatrix} \quad (5.8)$$

La función discriminante se aplica sobre  $x$  y así se puede constituir una regionalización lineal o no del espacio de características. Las regionalizaciones ayudan a determinar de qué tipo de objeto se trata. Mientras menos características se tomen en cuenta la función discriminante será más sencilla.

## 5.4 Calibración intrínseca

En esta sección se exponen los detalles y resultados de la calibración interna de las cámaras. Como se ha mencionado a lo largo de todo el trabajo, este procedimiento es en gran parte responsable de la precisión en todos los resultados basados en imágenes. Las cámaras CCD se utilizan en el formato NTSC<sup>1</sup> y las imágenes que se capturan son de 640x480 píxeles, con una escala de intensidad en grises de 256 niveles.

El proceso para obtener los parámetros internos de la cámara es el que se describe en la Sección 3.7. y se hace por medio de algunos programas en MATLAB. Se utiliza MATLAB debido a que los programas son puramente matriciales y a que ésta sólo se realiza una sola vez (mientras no se muevan las lentes de las cámaras). La programación en otro lenguaje de alto nivel, como C, no es tan conveniente debido a la extensión tan grande de código y a que todas las funciones matriciales deben de ser programadas.

| Parámetro                     | Cámara 1 | Cámara 2 | Unidades |
|-------------------------------|----------|----------|----------|
| Número de puntos              | 14 x 20  | 14 x 18  | [ ]      |
| Distancia focal estimada      | 582.86   | 634.28   | píxeles  |
| Distancia focal calculada $f$ | 814      | 896      | píxeles  |
| Variación de $f$              | 1.4796   | 0.396    | [ ]      |
| Desviación Estándar de $f$    | 1.2164   | 0.6293   | [ ]      |

Tabla 5.1: Parámetros internos de las cámaras

En la Tabla 5.1 se muestran los resultados de la calibración interna de las cámaras; estos resultados se utilizan en todo el resto de este trabajo y seguirán vigentes mientras no se muevan las lentes de las cámaras, o mientras no se decida volver a realizar la calibración interna con el fin de mejorar los resultados antes presentados.

El número de puntos se refiere a la cantidad de puntos que se utiliza para hacer la corrección de la distorsión. Como se mencionó anteriormente, para la calibración se hace una estimación inicial de la distancia focal que en este caso resulta ser mucho menor a la que se calcula; esta primera estimación sirve como punto de partida para el cálculo de la nueva distancia focal. Como el proceso es iterativo, se puede volver a calcular la distancia focal tomando como dato ésta última.

Es importante resaltar que, aunque las cámaras son aparentemente iguales, el resultado de la calibración es distinto para las distancias focales. Esto puede deberse a la primera estimación de la distancia focal,

<sup>1</sup>NTSC es la abreviación del Comité Americano de Sistemas Televisivos (por sus siglas en inglés) que denota también al video estándar de 60 Hz, conocido como RS-170. EL video NTSC tiene una frecuencia de 30 marcos por segundo con 480 líneas activas, una frecuencia de 59.94 Hz en la vertical y de 15.734 kHz en el barrido horizontal. Tiene un ancho de banda de 4.2 MHz.

pero también a que las rejillas de sensores CCD no sean realmente iguales. Además, las cámaras utilizan lentes que probablemente tampoco sean iguales en su geometría. El resultado en las distancias focales nos muestra que la imagen de la cámara 2 será ligeramente más grande.

La variancia que se muestra en la Tabla 5.1 se refiere a una medida de dispersión estadística que determina la probabilidad de error en el valor propuesto. Se toma como función de distribución de probabilidad a una distribución normal y se obtienen errores de 1.5 y 2.8 píxeles para un intervalo de confianza del 99%. Estas variaciones son despreciables cuando se está trabajando con imágenes grandes, ya que un pixel representa un error en la vertical menor al 0.5% y menor al 0.37% en la horizontal.

## 5.5 Programa Visual.prj

El programa *Visual.prj* es el programa de calibración externa, reconocimiento del objeto, y rutina con los manipuladores. Está escrito en lenguaje de programación *C* en *LabWindows Versión 5.0.1* y puede correr en *Windows 95, 98* o *NT*, aunque para compilarlo y ensamblarlo es necesario contar con el software *IMAQ* que viene con la tarjeta de National Instruments PCI 1408. Es importante resaltar que también se debe contar con el software de MATLAB, cualquier versión.

Para utilizar este programa es necesario tener el objeto de calibración y el objeto que se desee manipular. Ambas cámaras deben estar puestas en una base que asegure el paralelismo entre los ejes ópticos (Figura 5.4). La base está soportada por dos tripies y se verifica que estén perfectamente horizontales con un nivel de tipo gota (Figura 5.5).

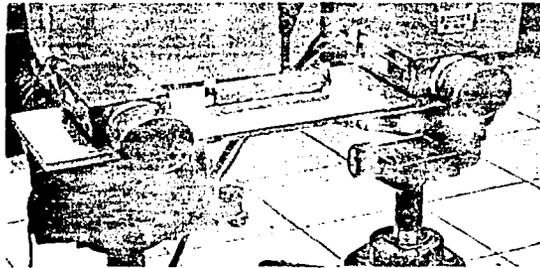


Figura 5.4: Base para las cámaras

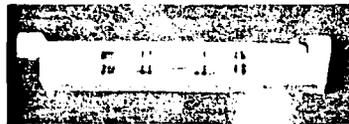


Figura 5.5: Nivel de tipo gota

Los robots están a una distancia variable de las cámaras, pero lo suficientemente alejados de ellas como para que su área máxima de trabajo no llegue a los aparatos de visión; esto con el fin de evitar daños a

cualquiera de los dos sistemas. Los manipuladores se encuentran sobre una mesa a 75 cm del piso y los objetos que localiza el sistema de visión se ponen sobre ella.

El ambiente de trabajo se modifica para obtener la iluminación más adecuada con el fin de que los algoritmos de extracción de características funcionen de la mejor forma. Se cubre el fondo de la escena con una tela de color uniforme y se deja que en la imagen aparezcan sólo los manipuladores y el objeto en cuestión (Figura 5.6).

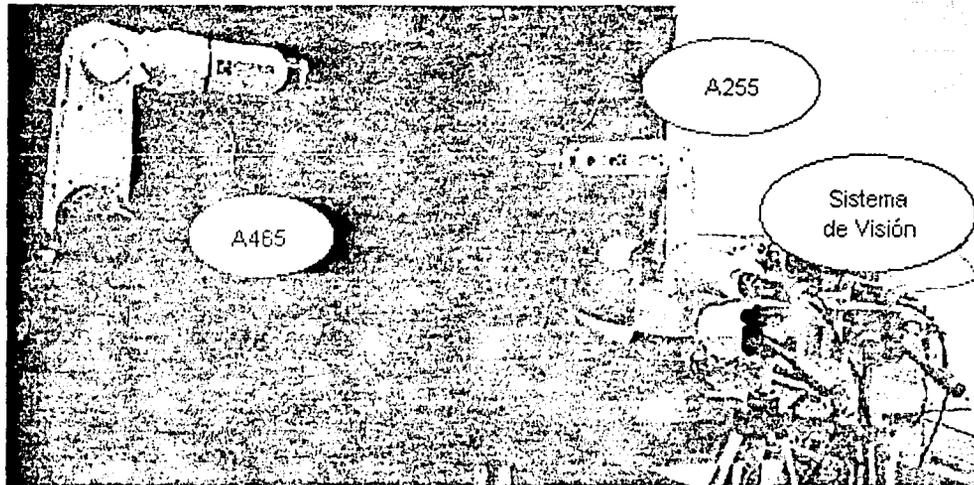


Figura 5.6: Ambiente de trabajo

### 5.5.1 Diagrama del programa

El diagrama funcional general del programa se muestra en la Figura 5.7. Es importante hacer notar que todos los botones, al oprimirse, tienen banderas que activan las funciones que van ligadas a ellos. Por falta de espacio no se consideran estas banderas en el diagrama de la figura, pero se describen detalladamente en la Sección 5.4.4.

### 5.5.2 Ventana inicial

Cuando se hace la corrida del programa lo primero que aparece es una ventana que le permite al usuario interactuar con el sistema. La pantalla inicial (Figura 5.8) es la que guía al usuario a lo largo de todo el programa por medio de cuadros de diálogo y una ventana de texto. La ventana de inicio está compuesta por:

1. El panel

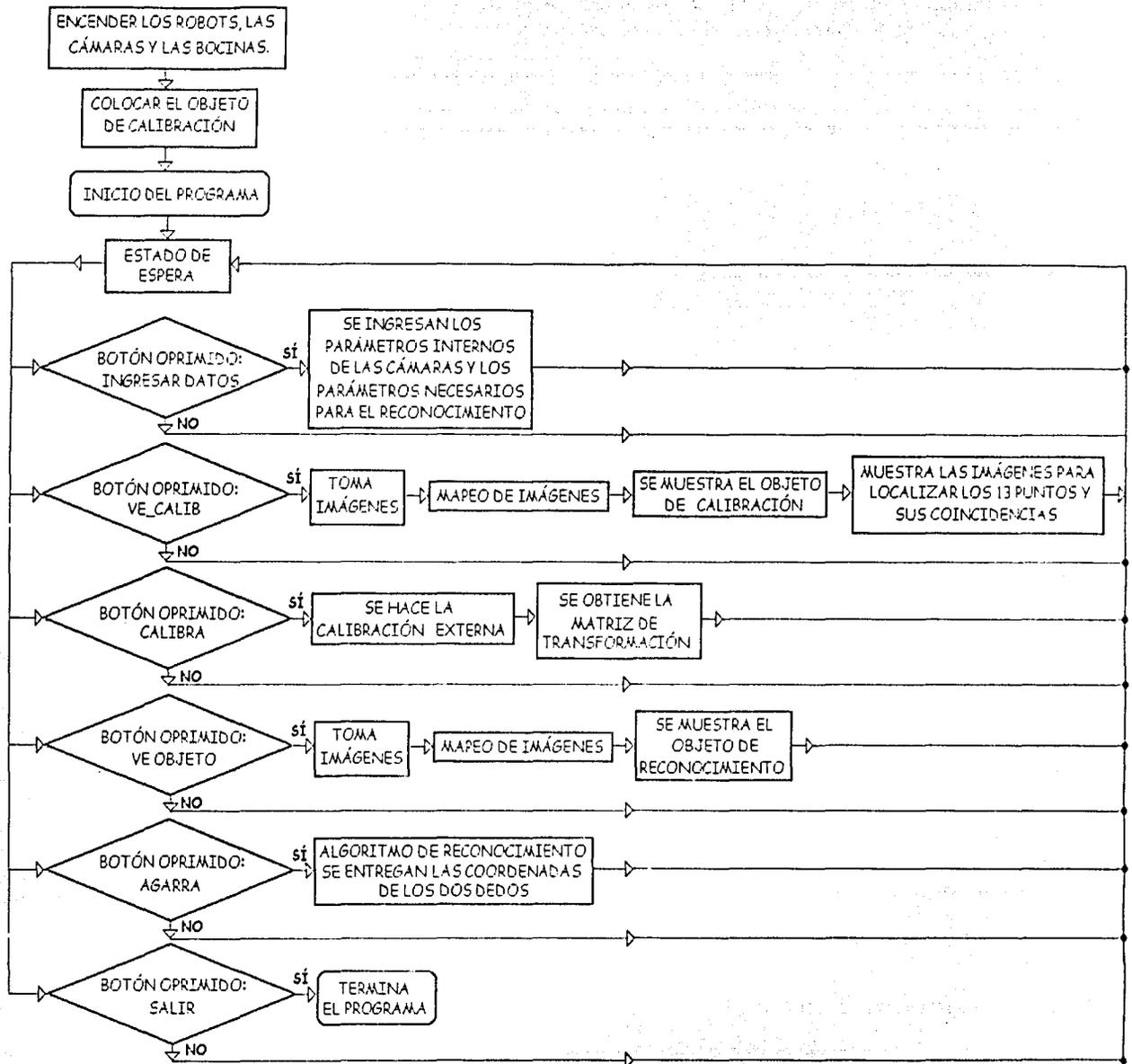


Figura 5.7: Diagrama funcional del programa

2. El lienzo para la cámara de la izquierda y su canal correspondiente.
3. El lienzo para la cámara de la derecha y el canal que le corresponde.

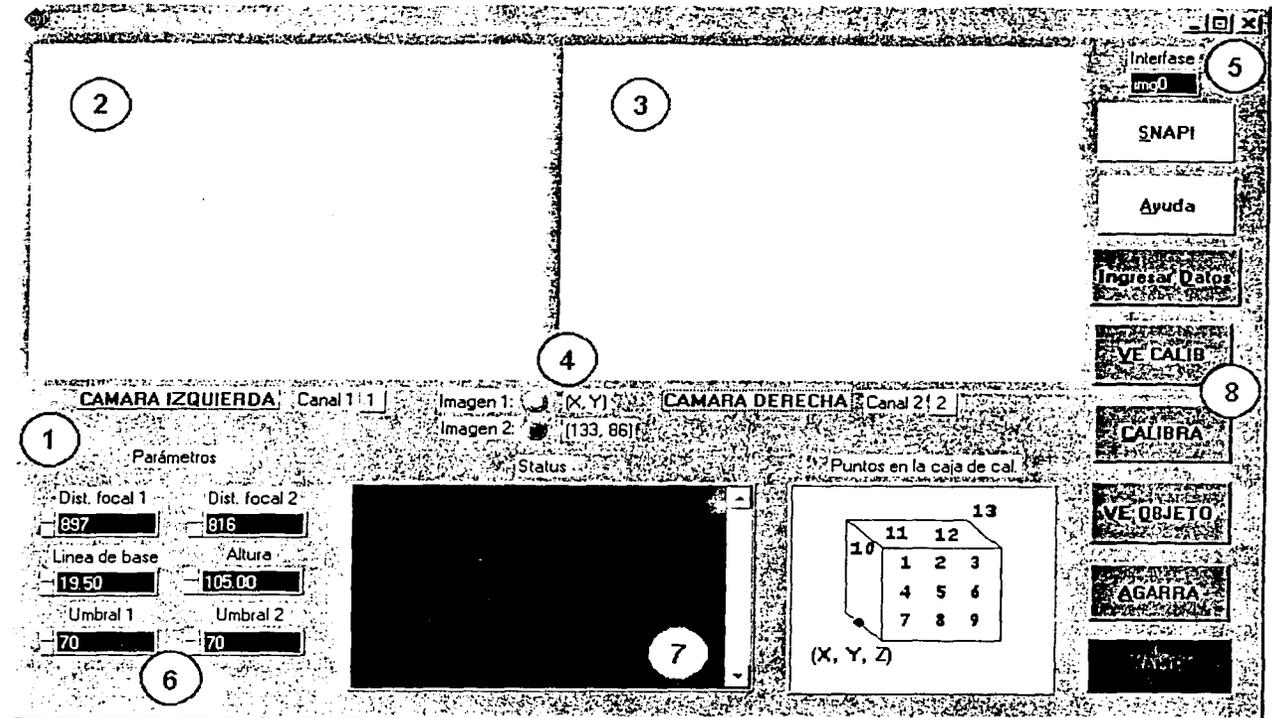


Figura 5.8: Ventana de inicio

4. Dos leds que indican si el cursor está sobre alguna de las imágenes, así como dos indicadores de coordenadas.
5. Nombre de la interfaz para las cámaras.
6. Datos necesarios para poder realizar la detección, el reconocimiento del objeto y la rutina con los manipuladores.
7. Ventana de texto en la que aparecen mensajes para el usuario.
8. Botones de comandos para realizar diferentes secciones del programa que se muestran en el diagrama funcional (Figura 5:7) y se describen en la Sección 5.4.5.

### 5.5.3 Librerías del código fuente

El código del programa inicia con la inclusión de las librerías que contienen todas las funciones que necesita el programa. Existen dos tipos de librerías: las que se incluyen con el programa de LabWindows (la ANSI estándar, las del manejo de los puertos y las cámaras, etc.) y las que se escribieron especialmente

para el programa `Visual.prj`. Una explicación más detallada de lo que contienen estas librerías se encuentra en el Apéndice E.

#### 5.5.4 Variables globales

En el programa existen tres variables bandera, de tipo entero, que indican el estado del proceso:

1. **DATA**: indica el ingreso de datos de la calibración interna (distancias focales), así como distancia entre las cámaras, umbrales, etc. Esta bandera es necesaria para poder realizar cualquier función, ya que los datos mencionados se ocupan en todas ellas.
2. **LIST\_C**: es la bandera que lleva la cuenta de los puntos de calibración que se han ingresado, hasta que se tienen los 13 puntos (con su correspondencia) se puede hacer la calibración.
3. **CALIB**: esta bandera muestra si la calibración externa de las cámaras se ha hecho o no. Cuando se calibran las cámaras queda la matriz de transformación correspondiente a disposición del programa. Mientras el programa no se reinicie, se podrán hacer todos los experimentos que se deseen con la misma calibración externa.

También se tienen seis variables a nivel global, que se utilizan en muchas secciones del programa.

- **B**: la distancia línea base, que es la separación entre las cámaras.
- **f1** y **f2**: son las distancias focales.
- **umbral1** y **umbral2**: son las variables que determinan el nivel de umbralización para la detección del objeto. Estas variables son muy importantes, porque el ambiente en que se realizan experimentos no tiene iluminación controlada. El correcto manejo de estas variables y de los diafragmas en las cámaras determinan en gran medida la precisión de los resultados.
- **altura**: determina la altura a la que se encuentran las cámaras desde el piso.

Existen también cuatro estructuras tipo `Matrix_f`, que son matrices de flotantes de tamaño variable que contienen la información de la calibración externa (los 13 puntos de la imagen 1 (**P1**), de la imagen 2 (**P2**), los 13 puntos con sus coordenadas espaciales (**PC1**) y la matriz de transformación (**A**)).

#### 5.5.5 Botones y funciones

Los botones que aparecen en la ventana inicial (Figura 5.7, indicador 8) sirven para que el usuario vaya accediendo a los diferentes niveles del programa en un orden lógico. Existe un botón **Ayuda** que muestra las instrucciones y funciones principales del programa, y un botón **Snap** que permite tomar

fotos con las cámaras (casi simultáneamente), las cuales son reducidas y desplegadas en las pantallas correspondientes. El botón Snap sirve para hacer ajuste de iluminación, diafragmas, y posición de cámaras y objetos. En la parte inferior derecha está el botón Salir que permite abandonar el programa en cualquier momento. Cuando se presiona se libera la memoria de las variables globales y de las estructuras.

A continuación se describen detalladamente las funciones de los botones restantes:

### Botón VE CALIB

Cuando se presiona este botón, el objeto de calibración debe estar en la posición correcta para poder iniciar el proceso de calibración externa. Al momento de presionar este botón se toman dos fotos del objeto (una cada cámara). Estas fotos se guardan como archivos PGM. Las imágenes de estos archivos son mapeadas con los polinomios que se obtuvieron en la corrección de la distorsión en la Sección 3.7.1. Las imágenes nuevas se vuelven a almacenar en archivos de tipo PGM.

A continuación aparece una nueva pantalla (Figura 5.9) que permite ver la imagen ya mapeada de la cámara de la izquierda, o cámara 1, con dimensiones de  $640 \times 480$  píxeles. La imagen que aparece es la que se guardó después del mapeo. Se decidió tomar las fotos, mapearlas y finalmente usar el resultado para ser mostrado porque el tiempo de mapeo es bastante grande y repetir la operación 13 veces es inútil. En esta pantalla se seleccionan los 13 puntos de calibración por medio del *mouse*. Después de cada punto seleccionado aparece una tercera pantalla (Figura 5.10) que muestra la imagen mapeada de la cámara 2 o de la derecha, para encontrar el punto de coincidencia. El proceso se repite 13 veces, para los 13 puntos de la imagen 1 que se guardan en *Matrix.f P1* y 13 puntos de la imagen 2 que se almacenan en *Matrix.f P2*. Al finalizar la selección se regresa a la pantalla inicial.

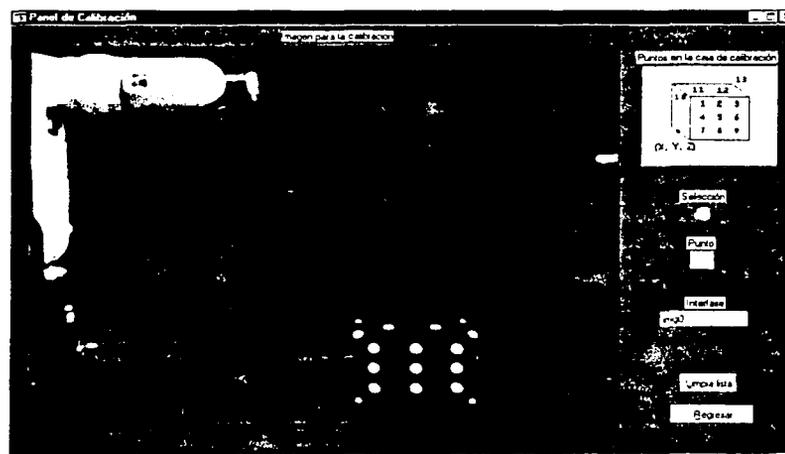


Figura 5.9: Puntos de calibración de la cámara 1

Cabe resaltar que existen botones en las pantallas nuevas que permiten borrar las listas de puntos

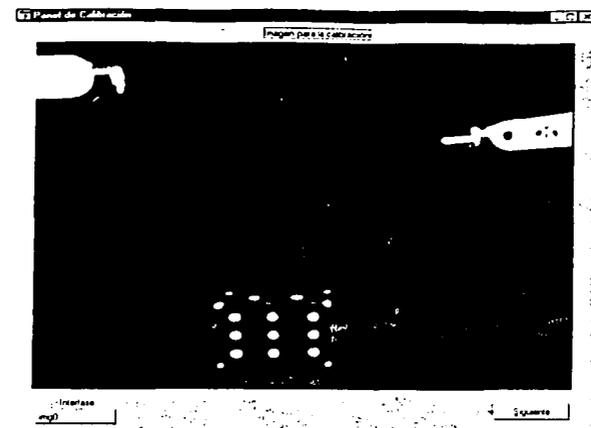


Figura 5.10: Puntos de calibración de la cámara 2

seleccionados y regresar a la pantalla principal; así como cuadros de diálogo que permiten aceptar o rechazar alguna selección.

### Botón INGRESAR DATOS

Llama a la siguiente función: `int CVICALLBACK Insertdata (int panel, int control, int event, void *callbackData, int eventData1, int eventData2)`

```
{
switch (event)
{
case EVENT_COMMIT:
GetCtrlVal (panel, PANEL_BASE_LINE, &B);
GetCtrlVal (panel, PANEL_FOCO, &f1);
GetCtrlVal (panel, PANEL_FOCO_2, &f2);
GetCtrlVal (panel, PANEL_ALT, &altura);
GetCtrlVal (panel, PANEL_UMBRAL, &umbral);
GetCtrlVal (panel, PANEL_UMBRAL_2, &umbral2);
DATA=1;
break;
}
return 0;
}
```

Esta función sirve para asignar los datos que el usuario escribe a las variables que se utilizan dentro del programa. Los datos pueden ser modificados en cualquier momento, aunque se tiene que tomar en cuenta que, después de la calibración externa, las distancias focales tendrán que permanecer constantes.

### Botón CALIBRA

Este botón es el que inicia el algoritmo para calcular las matrices de transformación necesarias para conocer la situación espacial de un punto con respecto al marco de referencia general, situado en el centro de la base del robot A465. Cuando se presiona el botón (si las variables bandera cumplen con  $DATA=1$  y  $LIST\_C=13$ ) se inicia la función *calibra*. Esta función convierte los datos de los puntos en las imágenes P1 y P2 en puntos con coordenadas espaciales desde el marco de referencia de la cámara 1 (los datos se guardan en PC1, que es de tipo Matrix.f). La forma en que se convierte la información de dos a tres dimensiones es la que se expone en la Sección 3.5.1.

Después de la conversión a coordenadas espaciales es necesario solamente encontrar los coeficientes de la matriz de transformación A. Como se menciona en la Sección 3.5.1, el objeto de calibración es totalmente conocido y por tanto se cuenta a priori con los datos de la matriz que lo describe *Matrix.f* PCR. Para hallar los 12 coeficientes desconocidos de la matriz de transformación es necesario utilizar MATLAB. La función *calibra* llama abrir entonces al programa de MATLAB. Se exportan las dos matrices de datos: PCR y PC1 y se obtienen los coeficientes de A. Finalmente se almacena la matriz de transformación A para su uso posterior.

Nota: La matriz A (calibración externa) deberá ser calculada cada vez que se reinicie el programa y cada vez que se muevan las cámaras de posición u orientación.

### Botón VE OBJETO

Antes de presionar este botón, el objeto a manipular debe estar ya en la posición deseada. El botón VE OBJETO funciona de una forma muy similar al botón VE CALIB. Al momento de presionarlo se toman las dos fotos del objeto (una desde cada cámara) y se muestran en los lienzos respectivos de la pantalla principal (Figura 5.11), después de pasar por un algoritmo de reducción de tamaño. Las imágenes, que originalmente son de  $640 \times 480$  pixeles, se muestran como imágenes de  $320 \times 240$ . Las fotos originales, igual que en el caso anterior, se guardan como archivos PGM y se mapean con los polinomios que se obtuvieron en la corrección de la distorsión (Sección 3.7.1). Las imágenes nuevas se vuelven a almacenar en archivos de tipo PGM y son las que se utilizan para el proceso de reconocimiento.

### Botón AGARRA

Este es el último botón de funciones dentro del programa y hace un llamado a la función:

```
object_recog(B, f1, f2, umbral1, umbral2, &Xd1, &Yd1, &Zd1, &Xd2, &Yd2, &Zd2, panel).
```

B, f1 y f2 son las variables globales de parámetros de las cámara. *umbral1* y *umbral2* son los límites para hacer la umbralización en cada imagen. *Xdi*, *Ydi* y *Zdi* son las coordenadas del punto donde se

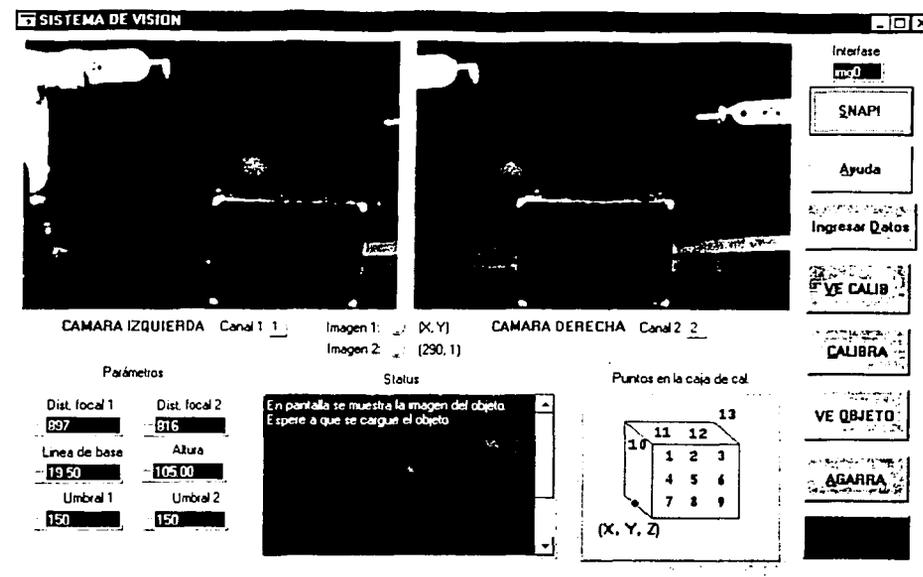


Figura 5.11: Imágenes del objeto en la pantalla principal

va a agarrar el objeto en el sistema coordenado de la imagen y *panel* es el identificador de la ventana que está en uso.

Esta función contiene el algoritmo para reconocer el objeto. Cabe resaltar que el algoritmo que se programó se hizo de forma heurística y que tiene como premisas varios aspectos:

- El objeto que se trata de identificar es una caja.
- La caja tendrá que estar pintada de negro con las esquinas en blanco.
- El objeto debe estar colocado sobre la mesa de forma natural, es decir, con una de sus caras completamente sobre la mesa.
- Si la caja fuera más alta que el robot A255 (60 cm) no será reconocida como tal.
- Si la caja es muy pequeña (menor a 30 cm), muy grande (mayor a 45 cm) o está en una posición donde los manipuladores no la pueden agarrar será reconocida como caja, pero no se realizará la rutina.
- Es importante tomar en cuenta que el algoritmo de control no trabaja con retroalimentación de fuerza y que el agarre que se hace está limitado a ciertos pesos del objeto.
- El ambiente tiene que ser modificado; el fondo de la escena debe ser cubierto con una tela de color uniforme (Figura 5.6).

- Las cámaras deben estar perfectamente horizontales y con sus ejes ópticos paralelos (Figura 5.4).
- Los ejes ópticos deben estar en posición perpendicular a la línea que une las bases de los robots.

El algoritmo de reconocimiento consiste primero en hacer una umbralización de las imágenes (Sección 3.4.1). El valor de los umbrales cambia de experimento a experimento y, en este caso, sólo el usuario podrá determinar su valor. El umbral sirve para encontrar las marcas blancas (esquinas del objeto) dentro de la imagen. Como se menciona en la Sección 3.2, a cada nivel de gris se le asigna un valor discreto dentro de la imagen, siendo 255 el más brillante y 0 el más oscuro. Las variables *umbral* marcan el límite de la intensidad para hacer distinciones de lo que se desea encontrar.

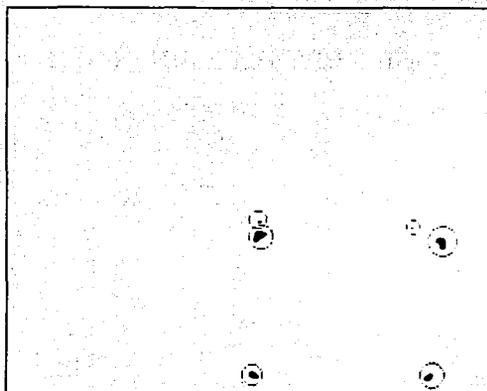


Figura 5.12: Marcas (manchas) reconocidas

Las marcas son en realidad manchas (Figura 5.12) que deben ser reducidas a un punto que signifique, por ejemplo, una esquina o cualquier otro punto conocido del objeto. La reducción de las manchas a un solo punto se hace por medio de un algoritmo que lleva a cabo un análisis de zonas con un cierto radio y luego hace un promedio. La media proporciona las coordenadas del punto que se toma como esquina o punto en cuestión (Figura 5.13).

Para este trabajo se toman sólo 2 casos: cuando se pueden ver 6 ó 7 esquinas, es decir 2 ó 3 caras de la caja. Con base en este dato se decide que si existen más de 7 esquinas en la imagen, el objeto "no es una caja". En caso de presentarse esto, el resultado es anunciado al usuario. Se puede volver a realizar el experimento cambiando de objeto o modificando la posición del mismo. En caso de que la imagen contara con 6 ó 7 puntos que se pudieran asumir como esquinas, comienza un algoritmo de detección de posición del objeto. Este algoritmo calcula el tamaño de las aristas y con ellas, los puntos donde los dedos de los manipuladores tienen que tomar al objeto. Después se ordenan los puntos de la imagen asumiendo la configuración de la Figura 5.14, según sea el caso. El ordenamiento se hace primero considerando las coordenadas en dirección vertical y luego en dirección horizontal. El caso de la Figura 5.15 no se puede presentar debido a que una de de las premisas del programa dice que "El objeto debe estar colocado con una de sus caras completamente sobre la mesa".

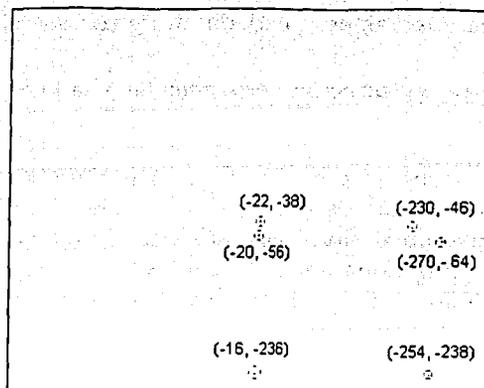


Figura 5.13: Probables esquinas

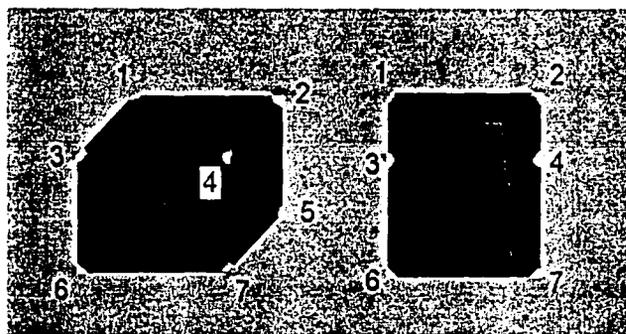


Figura 5.14: Enumeración de esquinas según la posición de la caja

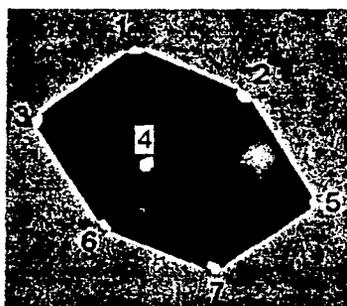


Figura 5.15: Caso imposible

Tanto la umbralización como el ordenamiento se hace en las imágenes obtenidas con ambas cámaras. Con ello se puede saber la correspondencia de las esquinas en las distintas imágenes. Es importante resaltar que la esquina 5 de la configuración de 7 esquinas juega un papel muy importante. Cuando

alguna de las imágenes muestra sólo 6 esquinas la número 5 de la imagen con 7 no se considera al hacer las coincidencias.

Conociendo las esquinas correspondientes en cada imagen y su posición, se puede proceder al cálculo de los valores en el espacio. Por medio de las transformaciones de la proyección (Sección 3.5.1) se encuentran las coordenadas espaciales de cada una de las esquinas con respecto al sistema coordenado de la cámara 1. Cuando se tienen estas coordenadas espaciales se calcula el ángulo de rotación de la caja con respecto a los manipuladores por medio de los puntos 1, 3, 4 y algo de geometría. Como se sabe que los ejes ópticos de las cámaras están en una posición perpendicular a la línea que une las dos bases de los robots (premisa), un triángulo rectángulo en alguna de las imágenes asegurará que la caja tenga sus caras perpendiculares al eje  $y$  del sistema coordenado general. El ángulo en cuestión es el de la esquina 3 y sirve para determinar si los manipuladores son capaces de agarrar y manipular el objeto. El intervalo permisible para este ángulo es realmente muy pequeño, ya que de por sí los datos de las imágenes no son precisos y un ángulo grande evita totalmente la idea del agarre estable (Sección 4.1).

Si se decide que los robots pueden agarrar y manipular el objeto se procede al cálculo de las aristas. Por medio de las coordenadas espaciales se calculan las longitudes de las aristas visibles de la caja. Las aristas horizontales (en la imagen) deben ser lo más semejantes posible; igualmente las aristas verticales. En el caso de las aristas "diagonales", es decir, las que denotan la profundidad, también debe existir una similitud. Los valores de aristas que se toman son los promedios. Finalmente, con los datos de las aristas se calculan los puntos a donde los robots deben llegar para iniciar la rutina. Los puntos que se obtienen en este momento son con respecto al sistema de la cámara 1, pero con la matriz de transformación  $A$ , almacenada previamente, se obtienen, por medio de una multiplicación, las coordenadas con respecto al robot (sistema general) (véase la Figura 5.16).

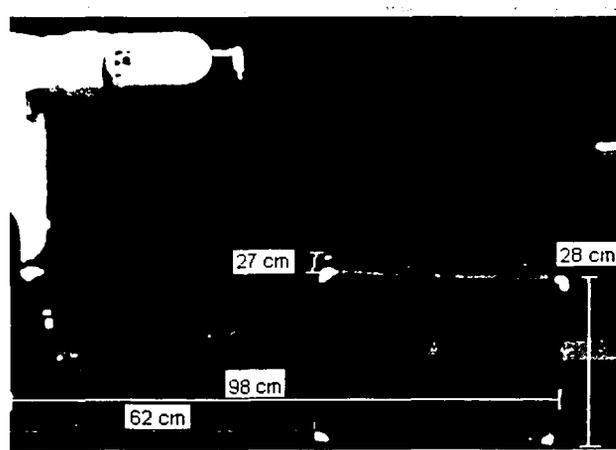


Figura 5.16: Coordenadas espaciales desde el robot A465

Las coordenadas obtenidas se pasan al programa rutina.prj (Figura 5.17) y se inicia la rutina de sujeción del objeto. Los robots se encuentran al principio en la posición *origen* (Figura 5.18), y están alimentados con un voltaje tal que no se muevan de esta posición mientras no se les indique lo contrario.

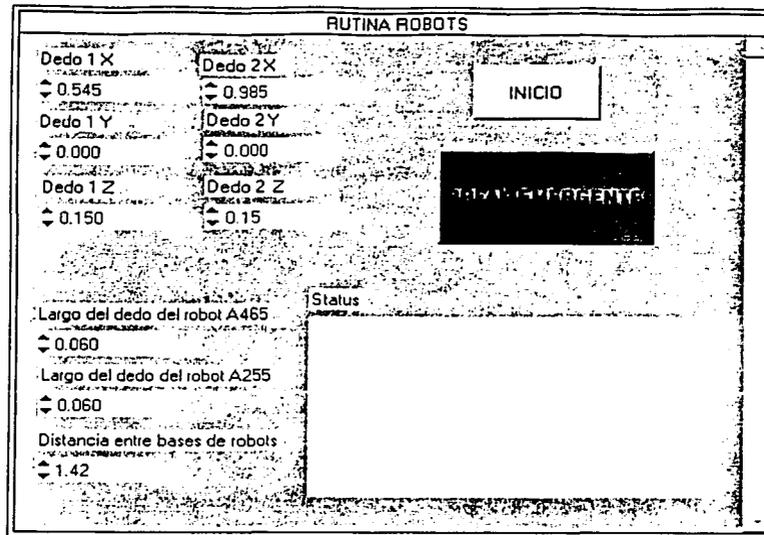


Figura 5.17: Pantalla del programa rutina.prj



Figura 5.18: Robots en posición origen

Al oprimir el botón Inicio, se leen las variables de posición de la pantalla, así como algunos parámetros de los robots (largo de los dedos, separación entre ellos). La función que realiza esto es muy similar a la que llama el botón Insertar Datos del programa visual.prj. Por medio del controlador PID programado (Sección 5.1) llegan ambos robots a la posición deseada (Figura 5.19).

Después de ello, hacen la rutina que consiste en un círculo con radio de 50 cm en un plano, trabajando de forma simultánea (Sección 4.5). Finalmente regresan a los puntos marcados en el programa (para devolver el objeto a su lugar inicial, si éste fuera el caso) y como última etapa vuelven a su posición de origen.

El programa rutina.prj cuenta con un botón de Salida Emergente en caso de que hubiera problemas con la rutina y como seguridad para el usuario. Al presionar este botón, inmediatamente pierden cualquier

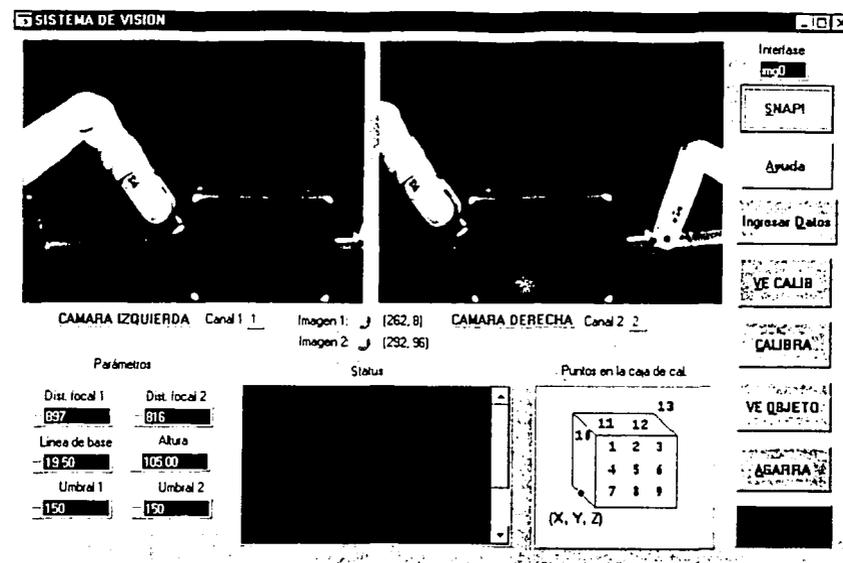


Figura 5.19: Posición deseada

voltaje los manipuladores.

### 5.5.6 Resultados experimentales

Los experimentos consisten en utilizar el programa `vision.prj` para hacer la calibración externa y así conocer las matrices de transformación de posición y orientación de la cámara con respecto al robot. Después, con los parámetros internos de las cámaras, el programa calcula los puntos a los que el robot tendrá que llegar por medio de las imágenes de las dos cámaras. Finalmente, recibe el robot estos datos (programa `rutina.prj`), llega a los puntos calculados y hace una rutina, que en este caso es un círculo en el aire con el objeto (Sección 5.5).

#### Experimento 1

El primer experimento consiste sólo en verificar la precisión de los parámetros calculados de las cámaras, por medio del sistema de visión. Se utiliza la caja de calibración (Figura 5.20) y se le coloca a 142 cm de las cámaras, con una de sus caras de forma perpendicular a los ejes ópticos (Figura 5.21).

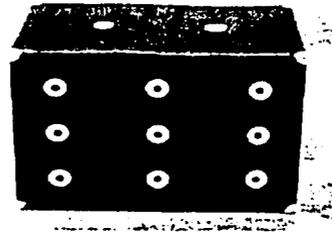


Figura 5.20: Caja de calibración

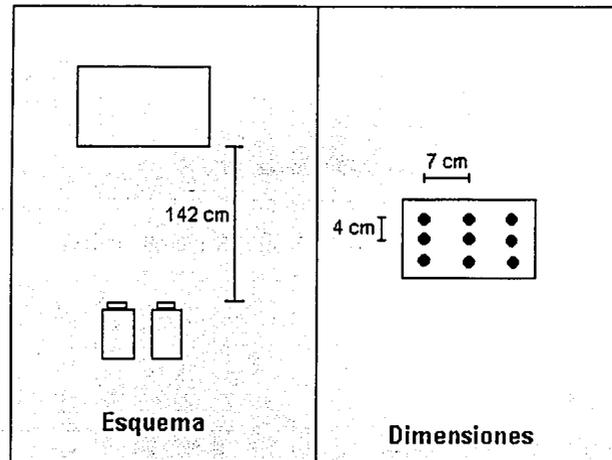


Figura 5.21: Esquema Experimento 1

Se obtienen los resultados de la Tabla 5.2.

| Distancia                      | Real [cm] | Calculada [cm] | Error porcentual |
|--------------------------------|-----------|----------------|------------------|
| Cámara-Caja                    | 142       | 141.99         | 0.007            |
| Segmento vertical en la caja   | 4         | 3.7305         | 6.74             |
| Segmento horizontal en la caja | 7         | 7.2308         | 3.30             |

Tabla 5.2: Relación de distancias para la calibración

Como se aprecia en este primer experimento, los parámetros de la calibración interna son bastante precisos. Los errores que se obtienen son muy pequeños y las relaciones de dimensiones son muy buenas, lo que significa que el procedimiento de calibración general es aceptable.

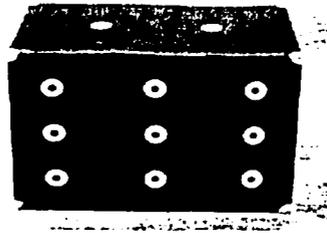


Figura 5.20: Caja de calibración

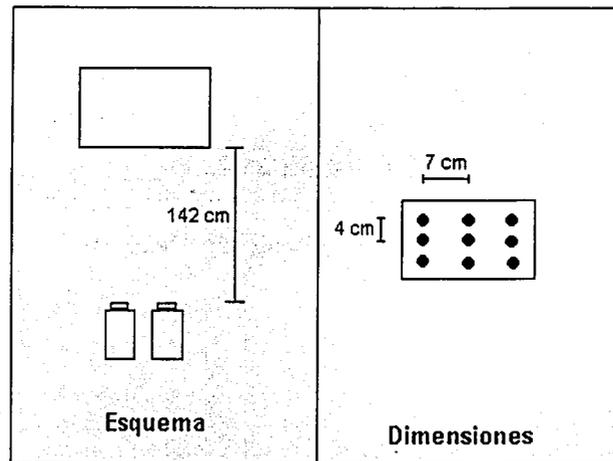


Figura 5.21: Esquema Experimento 1

Se obtienen los resultados de la Tabla 5.2.

| Distancia                      | Real [cm] | Calculada [cm] | Error porcentual |
|--------------------------------|-----------|----------------|------------------|
| Cámara-Caja                    | 142       | 141.99         | 0.007            |
| Segmento vertical en la caja   | 4         | 3.7305         | 6.74             |
| Segmento horizontal en la caja | 7         | 7.2308         | 3.30             |

Tabla 5.2: Relación de distancias para la calibración

Como se aprecia en este primer experimento, los parámetros de la calibración interna son bastante precisos. Los errores que se obtienen son muy pequeños y las relaciones de dimensiones son muy buenas, lo que significa que el procedimiento de calibración general es aceptable.

## Experimento 2

Este experimento se realiza con las cámaras en una posición donde se pueda detectar objetos para ser manipulados por los robots. Es decir, que las cámaras tienen que subirse, para que los objetos más grandes que la caja de calibración puedan ser detectados. Lo que ven las cámaras al hacer la calibración externa se aprecia en la Figura 5.22. Se desea manipular una caja (a la que se le conocerá como caja 1

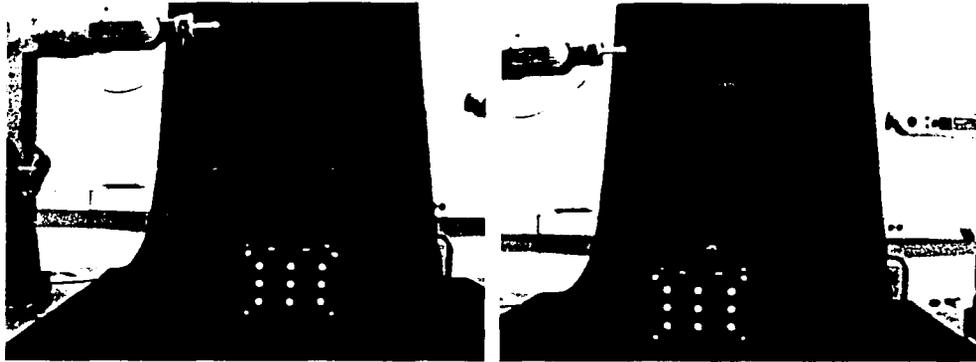


Figura 5.22: Calibración Experimento 1

(Figura 5.23)) con las dimensiones de la Tabla 5.3.

|       |    |
|-------|----|
| Largo | 35 |
| Ancho | 35 |
| Alto  | 29 |

Tabla 5.3: Dimensiones caja 1

Las cámaras se encuentran a 110 cm del piso y a 96 cm de la mesa que sostiene a los robots y a las cajas. Las imágenes que ven las cámaras del objeto a manipular se pueden apreciar en la Figura 5.24.

Por medio del programa *visual.prj* se detectan sólo 4 esquinas (esto se puede apreciar fácilmente en la Figura 5.24); por lo tanto no se puede conocer la profundidad del objeto. El programa marca que no se trata de una caja y no se realiza la rutina con los manipuladores.

La posición de las cámaras es entonces inadecuada. A pesar de que el objeto es de tamaño aceptable para los manipuladores no se puede realizar la rutina. Algunas posibles soluciones para este experimento son:

- Cambiar la posición y/u orientación de las cámaras; por ejemplo subirlas con el fin de que puedan ver las esquinas posteriores del objeto, separarlas más o modificar su ángulo con respecto a la horizontal.
- Agregar una cámara más al sistema de visión para mejorar la perspectiva

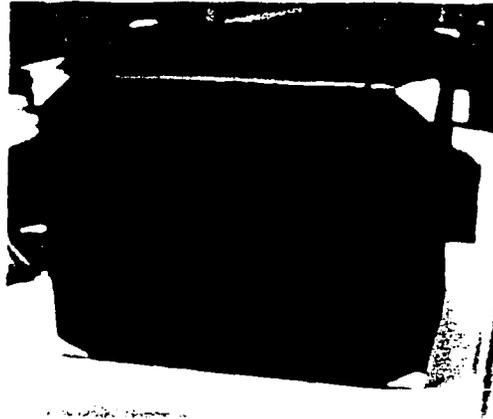


Figura 5.23: Caja 1



Figura 5.24: Experimento 2

### Experimento 3

Para el experimento 3 se trata de corregir la posición de las cámaras. Éstas se colocan a 125 cm del piso y a 99 cm de la mesa de los robots. Se vuelve a hacer la calibración extrínseca y se coloca la misma caja 1 para la manipulación. El objeto es detectado como caja y se realiza la rutina con los manipuladores (Figura 5.25).

| Distancia                      | Real [cm] | Calculada [cm] | Error porcentual |
|--------------------------------|-----------|----------------|------------------|
| Cámara-Caja                    | 145       | 151            | 4.13             |
| Segmento vertical en la caja   | 4         | 5.43           | 35               |
| Segmento horizontal en la caja | 7         | 8.98           | 28.29            |

Tabla 5.4: Relación de distancias para la calibración

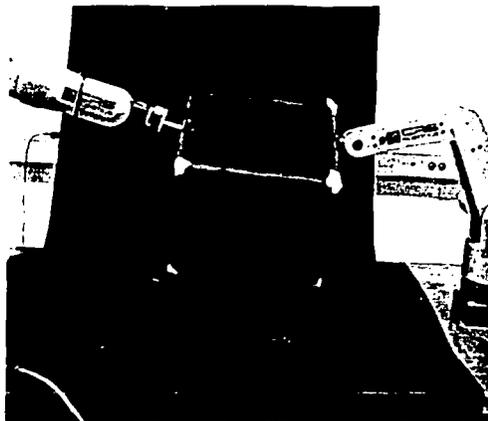


Figura 5.25: Experimento 3

Como se puede observar de la Tabla 5.4, la distancia de las cámaras juega un papel importante. Mientras más se alejen, el sistema se vuelve más impreciso, pero a pesar de ello se obtienen resultados más o menos aceptables. La caja es detectada y se realiza la rutina aunque el objeto rota durante la ejecución de la misma. Para asegurar que el objeto se agarre realmente se hace un ajuste de los puntos de agarre, es decir, se eligen puntos de agarre dentro del objeto. La caja puede ser deformada ligeramente pero el agarre es más seguro.

Realizar un experimento con este sistema de visión que logre un agarre estrictamente estable de forma automática es muy difícil, debido a que para que la calibración externa sea buena el objeto de calibración debe estar en el centro de la imagen y para que los robots puedan manipular un objeto éste debe ser considerablemente más grande que la caja de calibración, lo que implica que las cámaras deben tener otra posición para ver todas las esquinas del objeto.

El problema que se presenta en este sistema de visión es en el campo de visión de las cámaras y en su posición y/u orientación. Ampliar el campo de visión de los aparatos implica un cambio en las lentes. La variación de la posición y/u orientación de los aparatos lleva consigo un análisis matemático más profundo. Ambas soluciones se pueden realizar; para la primera se tienen que volver a hallar los parámetros internos de las cámaras y para la segunda se tiene que reprogramar la función de calibración dentro del programa *visual.prj*.



# Capítulo 6

## Conclusiones y trabajo a futuro

El contenido de la física le incumbe a los científicos,  
sus consecuencias a toda la humanidad.  
—*Die Physiker*, Dürrenmatt

El objetivo de esta tesis es el de implementar un sistema robusto de visión que permita distinguir y reconocer un objeto en el espacio. El objeto debe ser sujetado por dos manipuladores y realizar una rutina definida. En este sentido el propósito de la tesis ha sido cumplido y los resultados son satisfactorios.

Para tener un sistema de visión robusto se necesita que los parámetros tanto intrínsecos como extrínsecos de las cámaras sean lo más exactos posibles; por ello, la calibración es un punto fundamental en la precisión de todo el sistema. La primera parte de un sistema de visión consiste en contar con un método de calibración confiable, para lo que existen muchas técnicas.

La parte más difícil del proceso de calibración se encuentra en la programación de los algoritmos para hallar los parámetros. Una vez programados estos algoritmos la calibración es realmente sencilla y segura (tomando en cuenta las limitaciones generales del sistema: campo de visión de las cámaras, nitidez de las imágenes, iluminación, etc.). Esta tesis se centró en gran parte en la creación de estos códigos con el objetivo de que fueran precisos y portables. Los resultados son satisfactorios con errores en la distancia focal de 2 píxeles aproximadamente. Como los métodos que se manejan son iterativos se asegura que, si se necesitan mejores resultados, los que se obtuvieron aquí son una muy buena base para un recálculo.

El tiempo que dura el proceso de calibración no es importante, ya que se lleva a cabo previamente al inicio de cualquier experimento. Además, la calibración interna se hace sólo una vez para cada cámara mientras no se muevan las lentes; puesto que las lentes están enfocadas al infinito no es necesario moverlas continuamente.

La calibración externa se hace también sólo una vez al inicio de cualquier experimento y, si las cámaras se dejan fijas, no se necesitará volver a hacerla. El programa *visual.prj* que se presenta da la facilidad de mover las cámaras a voluntad y realiza la calibración externa en un par de minutos.

El análisis que se les hace a las imágenes en esta tesis para lograr el reconocimiento de objetos es bastante sencillo dado que parte de la heurística. En este caso se trata sólo de realizar un ejemplo práctico y simple. Para experimentos más bastos existen muchas técnicas con las que se puede trabajar; pero elementos como iluminación, formas y bases de datos rápidas siguen siendo motivos de estudio para su perfección.

El tiempo de procesamiento de imágenes es importante cuando se intenta trabajar en *tiempo real*. Las herramientas que se presentan aquí distan mucho de lo que se espera para un manejo en este campo, debido a que tan solo la corrección de la distorsión incluye muchos cálculos que hacen lento el proceso.

La determinación de los niveles de umbralización, que aquí se simplifica por medio de la modificación del ambiente, es bastante complicada con ambientes diversos; alternativas para su obtención óptima se hacen por ejemplo, por medio de redes neuronales o de lógica difusa. Para contar con una base de objetos más amplia, en vez de trabajar sólo con una forma en el presente trabajo, es necesario introducirse en técnicas de clasificación: grafos, árboles, etc.

Esta tesis es un paso en dirección de la visión artificial; hay todavía muchos campos para trabajar. La gran aportación que se hace aquí son las herramientas base: los programas para la calibración, detección de bordes, umbralizaciones, y hasta la transformada directa e inversa de Hough. Éstos son instrumentos fundamentales para continuar con la perfección del sistema.

Los experimentos muestran que el procedimiento de calibración programado es bastante confiable cuando el objeto de referencia se sitúa en el centro de la imagen y a una distancia en la que se puedan distinguir los puntos de calibración. Esto demuestra también que los parámetros internos de las cámaras (que permanecen constantes, si no se mueven las lentes) tienen un alto grado de precisión. Aunque una segunda calibración de éstos parámetros daría resultados mucho mejores.

La detección del objeto a manipular es también muy buena, si se tiene una buena iluminación y el ambiente se encuentra libre de manchas que puedan confundir al sistema. Un simple punto en el ambiente puede hacer que el sistema falle, es decir, que este programa es muy sensible al ruido. Es importante también recalcar que el tamaño de las marcas en las esquinas del objeto puede ser un factor determinante en el buen funcionamiento del algoritmo de detección y reconocimiento.

En general, los resultados de los experimentos son buenos si se toma en cuenta que las cajas que se utilizaron son de materiales rígidos, lo que hace que el coeficiente de fricción con los dedos sea pequeño y más fácil de presentar giros.

Puesto que el estudio de los sistemas de visión integrados a sistemas de robótica es relativamente nuevo, hay muchas áreas para realizar trabajos innovadores. Es excitante pensar en la tecnología robótica y su futuro, debido a los desarrollos que se hacen día a día en técnicas de control, electrónica, procesamiento de señales, realidad virtual, teleoperación, etc.

Un punto a considerar es que el empleo de robots en la industria ha alcanzado un nivel de introducción exitoso en muchas áreas; sin embargo, existen todavía muchas otras en las que los robots están en un estado inmaduro.

Finalmente, aparece una pregunta de carácter ético:

*¿Que tan responsables seremos los seres humanos al dotar de inteligencia a una máquina sin perder de vista lo que esto significa?*

Capek podría convertirse en profeta.



## Apéndice A

# Características principales de los manipuladores del Laboratorio de Robótica de la División de Posgrado de la Facultad de Ingeniería

Robot A255 de cinco grados de libertad

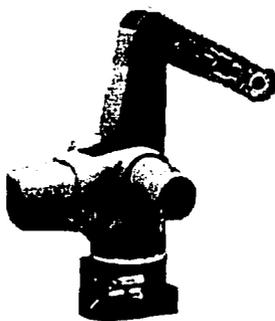


Figura A.1: Robot A255

El A255 es un robot industrial accionado por servomotores. Los eslabones están enlazados por cadenas, y tiene un sistema que compensa el movimiento del antebrazo con la posición del efector final, manteniendo su relación con el plano horizontal. Esencialmente está formado por el brazo, los cables que lo proveen de energía eléctrica para los motores y la comunicación con el controlador. Consiste mecánicamente en un cuerpo que gira sobre un plano horizontal, gracias a la articulación 1; un brazo posterior acoplado al cuerpo (o base) mediante la articulación 2; un antebrazo acoplado al brazo posterior por la articulación 3 y una muñeca. La muñeca se dobla (articulación 4) y tuerce (articulación 5)

con respecto al antebrazo. A la muñeca se le pueden montar una gran variedad de efectores finales como puntas, pinzas, etc. Cada articulación cuenta con encifradores incrementales que proveen información continua de la posición del motor.

El A255 se controla por medio de una interfaz directa con la PC. Es posible controlar el movimiento de cada una de las articulaciones mediante la PC, a cualquier velocidad (dentro de los límites mecánicos y del tiempo de muestreo). Aparte existe también un panel de control donde se encuentran los interruptores para el control y los motores. Cuando la base o cualquiera de las dos secciones del brazo sobrepasan los límites especificados del movimiento, actúan los interruptores límites y los motores pierden la alimentación a la vez que actúan los frenos de emergencia y se emite una señal auditiva. Esta protección es necesaria para evitar la destrucción de los componentes mecánicos del sistema; además cuenta con un botón que es el freno de emergencia para el usuario y la protección del personal que esté laborando cerca del robot.

### Especificaciones técnicas

|                         |                                 |
|-------------------------|---------------------------------|
| Nombre                  | CRS Robotics A255 Arm           |
| Fabricante              | CRS Robotics Corporation        |
| País de origen          | Canadá                          |
| Número de ejes          | 5                               |
| Peso                    | 17 kg                           |
| Capacidad de carga      | 1 kg                            |
| Alcance                 | 0.559 m                         |
| Frenos                  | En las articulaciones 2,3,4 y 5 |
| Voltaje de alimentación | 100/115/230 VAC $\pm 10\%$      |
| Frecuencia              | 50-60 Hz                        |
| Potencia máxima         | 1000 W                          |
| Temperatura             | 10° a 40°C                      |
| Humedad                 | 80%                             |

### Articulaciones del robot A255

| Art. | Rotación de...  | Rangos de movimiento | Velocidad máxima | Aceleración normal     | Relación de transmisión | Torque  |
|------|-----------------|----------------------|------------------|------------------------|-------------------------|---------|
| 1    | Cuerpo          | $\pm 175^\circ$      | 210° / s         | 498° / s <sup>2</sup>  | 72 : 1                  | 6.4 Nm  |
| 2    | Brazo posterior | 0° a 110°            | 210° / s         | 498° / s <sup>2</sup>  | -72 : 1                 | 6.4 Nm  |
| 3    | Brazo anterior  | -125° a 0°           | 210° / s         | 498° / s <sup>2</sup>  | 72 : 1                  | 6.4 Nm  |
| 4    | Muñeca (doblar) | $\pm 180^\circ$      | 675° / s         | 2040° / s <sup>2</sup> | -16 : 1                 | 1.4 Nm  |
| 5    | Muñeca (torcer) | $\pm 180^\circ$      | 1350° / s        | 4490° / s <sup>2</sup> | 8 : 1                   | 0.71 Nm |

## Robot A465 de seis grados de libertad



Figura A.2: Robot A465

El A465 es un robot industrial también accionado por servomotores, los eslabones están enlazados por cadenas. Consiste en un cuerpo que gira sobre un plano horizontal (articulación 1); un brazo posterior acoplado al cuerpo o base mediante la articulación 2; un antebrazo acoplado al brazo posterior por la articulación 3; la muñeca que está formada por la articulación giratoria 4 en el antebrazo, la articulación que se dobla, la 5 y otra que se tuerce (articulación 6) con respecto al antebrazo, como sucede en el manipulador de cinco grados de libertad.

El A465 se controla exactamente de la misma forma que el manipulador de cinco grados de libertad, si se opera individualmente. Asimismo cuenta con frenos de emergencia para la seguridad del usuario y de los componentes mecánicos. También tiene encifradores incrementales y sensores que informan de la posición de los motores.

## Especificaciones técnicas

|                         |                             |
|-------------------------|-----------------------------|
| Nombre                  | CRS Robotics A465 Arm       |
| Fabricante              | CRS Robotics Corporation    |
| País de origen          | Canadá                      |
| Número de ejes          | 6                           |
| Peso                    | 31 kg                       |
| Capacidad de carga      | 2 kg                        |
| Alcance                 | 0.710 m                     |
| Frenos                  | En las articulaciones 2 y 3 |
| Voltaje de alimentación | 100/115/230 VAC $\pm 10\%$  |
| Frecuencia              | 50-60 Hz                    |
| Potencia máxima         | 1000 W                      |
| Temperatura             | 10° a 40°C                  |
| Humedad                 | 80%                         |

## Articulaciones del robot A465

| Art. | Rotación de...              | Rangos de movimiento | Velocidad máxima | Aceleración normal | Relación de transmisión | Torque  |
|------|-----------------------------|----------------------|------------------|--------------------|-------------------------|---------|
| 1    | Cuerpo                      | $\pm 175^\circ$      | $180^\circ / s$  | $720^\circ / s^2$  | 100 : 1                 | 39.5 Nm |
| 2    | Brazo posterior             | $\pm 90^\circ$       | $180^\circ / s$  | $720^\circ / s^2$  | 100 : 1                 | 39.5 Nm |
| 3    | Brazo anterior              | $\pm 110^\circ$      | $180^\circ / s$  | $720^\circ / s^2$  | -100 : 1                | 39.5 Nm |
| 4    | Rotación del brazo anterior | $\pm 180^\circ$      | $171^\circ / s$  | $1430^\circ / s^2$ | -101 : 1                | 6.9 Nm  |
| 5    | Muñeca (doblar)             | $\pm 105^\circ$      | $173^\circ / s$  | $1430^\circ / s^2$ | -100 : 1                | 6.9 Nm  |
| 6    | Muñeca (torcer)             | $\pm 180^\circ$      | $171^\circ / s$  | $1430^\circ / s^2$ | -101 : 1                | 2.5 Nm  |

## Apéndice B

# Características principales de las Cámaras CCD

Las cámaras del Laboratorio de Robótica de la División de Posgrado de la Facultad de Ingeniería son cámaras de video CCD en blanco y negro (Marca: SONY, Modelo: XC-75CE). Una de estas cámaras se muestra en la Figura B.1. Los dispositivos CCD son los más utilizados cuando se trata de sensores de estado sólido. CCD por sus siglas en inglés quiere decir que se trata de un Dispositivo de Cargas Acopladas y generalmente requiere de un sistema óptico para enfocar la luz entrante al chip.



Figura B.1: Cámara CCD

Un dispositivo de cargas acopladas consta de tres funciones básicas: colección de carga, transferencia de carga y conversión de carga en un voltaje que se pueda medir. La señal de control se refiere a la energía que llega al chip. El sensor está compuesto de elementos semiconductores que convierten la luz en información binaria por medio de un proceso de cuantización. Tiene dos compuertas de transferencia que sirven para controlar los tiempos de los elementos de la imagen a los registros de transporte. También consta de una compuerta de salida para controlar el tiempo de transporte desde los registros hasta el amplificador. Las salidas del amplificador muestran voltajes proporcionales a las señales energéticas recibidas (Figura B.2). Las características principales del modelo XC-75CE se resumen en las tablas siguientes.

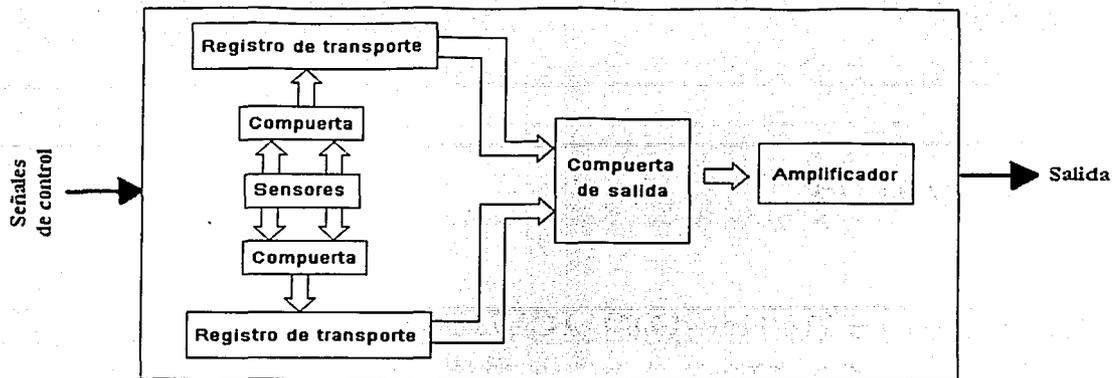


Figura B.2: Sistema sensorial CCD

## Especificaciones técnicas

|        |         |
|--------|---------|
| Marca  | SONY    |
| Modelo | XC-75CE |

## Sistema de imagen

|                                  |                         |
|----------------------------------|-------------------------|
| Frecuencia de barrido vertical   | 15.7 kHz $\pm 1\%$      |
| Frecuencia de barrido horizontal | 14.3 kHz $\pm 1\%$      |
| Tamaño de la célula              | 8.6 x 8.3 $\mu\text{m}$ |
| Tamaño del chip                  | 7.95 x 6.45 mm          |

## Sistema óptico

|                             |  |
|-----------------------------|--|
| Longitud focal              | 17.526 mm  |
| Sincronización              | Interna/Externa  |
| Sistema de adquisición      | 625 líneas, 2:1 entrelazadas/no entrelazadas (según la señal de entrada) |
| Resolución horizontal       | 560 líneas-TV  |
| Líneas verticales efectivas | 575 líneas (con 2:1 entrelazado)   |
| Sensibilidad                | 400 lux = 400 $\frac{\text{cd}}{\text{m}^2}$                             |
| Iluminación mínima          | 3 lux = 3 $\frac{\text{cd}}{\text{m}^2}$                                 |
| Relación señal a ruido      | 54 dB  |
| Disparo                     | de $\frac{1}{1000}$ a $\frac{1}{100}$ s                                  |

## Especificaciones eléctricas

|                         |                             |
|-------------------------|-----------------------------|
| Voltaje de alimentación | 12 VDC (Rango: 10.5 a 15 V) |
| Potencia                | 1.6 W                       |

## Ambiente de operación

|             |            |
|-------------|------------|
| Temperatura | -5° a 45°C |
| Humedad     | 20% a 80%  |

## Especificaciones adicionales

|           |             |
|-----------|-------------|
| Vibración | 11 a 200 Hz |
| Peso      | 140 g       |

Algunas desventajas de este tipo de cámaras son:

**Manchas.** Ocurren cuando se observan objetos muy luminosos. Se produce porque se induce una carga eléctrica infraroja al sensor.

**Aliasing vertical.** Cuando se observan líneas verticales, pueden aparecer melladas.

**Defectos.** El sensor consiste de un arreglo de elementos sensoriales individuales; si uno de ellos no funciona habrá defectos en algún pixel de la imagen.

**Manchas blancas.** Si se observa un objeto oscuro a altas temperaturas pueden aparecer manchas blancas en la imagen.



## Apéndice C

# Características principales de las tarjetas PCI-FlexMotion-6C

Las tarjetas PCI-FlexMotion-6C son la interfaz entre los encoders del robot y la PC. Son de tipo PCI y son de utilización sencilla pues con cinco bloques de programación pueden controlar una articulación en: posición, velocidad, aceleración, inicio y reporte de error, ya que originalmente contienen un controlador PID. Este controlador, en hardware, ha sido desconectado para poder hacer pruebas por medio de controladores programados. Los encoders trabajan con diferentes frecuencias de cuentas:

- Encoders 1 y 2                    16 MHz
- Encoders 3 y 4                    1 MHz
- Encoders 5 y 6                    2 MHz

Tienen 8 entradas analógicas multiplexadas y 6 salidas de tipo analógico también. Los voltajes en estas entradas/salidas pueden variar de  $\pm 10$  V, y la corriente máxima de salida es de  $\pm 10$  mA.

### Especificaciones técnicas

|                                  |                           |
|----------------------------------|---------------------------|
| Nombre                           | Tarjeta PCI-FlexMotion-6C |
| Fabricante                       | National Instruments      |
| País de origen                   | E.U.A.                    |
| Voltaje de alimentación          | 5V                        |
| Corriente máxima de alimentación | 2A                        |
| Potencia máxima                  | 14.2 W                    |
| Temperatura de operación         | 0° a 55°C                 |
| Humedad                          | 90%                       |



## Apéndice D

# Características principales de la tarjeta PCI-IMAQ-1408

La tarjeta PCI-1408 es la que se utiliza para tener acceso a los datos de las cámaras de video CCD. Esta tarjeta es para señales monocromáticas y puede manejar videos de tipo RS-170, CCIR, NTSC, y PAL. La tarjeta 1408 adquiere los marcos en tiempo real y los transfiere a la memoria del sistema.

### Especificaciones técnicas

|                                |                                   |
|--------------------------------|-----------------------------------|
| Marca                          | National Instruments              |
| Modelo                         | PCI-IMAQ-1408                     |
| Número de entradas de video    | Cuatro monocromáticas             |
| Frecuencia de respuesta        | 20 MHz                            |
| Ganancia                       | Programable a 1, 1.33 o 2         |
| Referencia de blancos y negros | Programable de 0 a 1.26 V         |
| Niveles de gris                | 256 (8 bits)                      |
| Relación SNR                   | 48 dB typ                         |
| Frecuencia de muestreo         | 5 a 20 MHz (manipulación externa) |
| Ancho de pulso                 | 50 ns                             |
| Voltaje para la tarjeta        | 5 V                               |
| Corriente máxima               | 1.34 A                            |
| ROM de expansión               | 4 KB                              |
| Temperatura                    | 0-55°C                            |
| Humedad relativa               | 5-90%                             |
| Emisiones EM                   | EN 55011:1991 Grupo 1 Clase A     |



# Apéndice E

## Librerías de programa Visual.prj

### Librerías de LabWindows

- *< ansi\_c.h >*: contiene todas las librerías estándar de C, como son: *stdio.h*, *conio.h*, *string.h*, *stdlib.h*, etc.
- *"niimaq.h"*: es la librería que contiene todas las funciones para acceder a la información de la tarjeta de adquisición de imágenes.
- *< userint.h >*: es necesaria si se está utilizando una interfase amigable con ventanas y controles de botones. Contiene funciones para utilizar el mouse, manejo de cuadros de texto y de diálogo, así como funciones para crear nuevas ventanas.
- *< utility.h >*: sirve para trabajar con archivos, directorios, fechas, horas, puertos, interrupciones y el teclado.
- *"matlabsvr.h"*: esta librería sirve para abrir el programa MATLAB y poder hacer operaciones en él. Se pueden exportar e importar matrices, cadenas, enteros, etc.

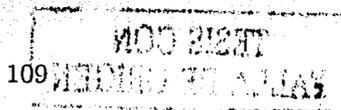
### Librerías creadas para este trabajo

- *"box.h"*: es la librería que contiene el algoritmo de reconocimiento de una caja. Calcula el tamaño de las aristas y calcula el centro de masa del objeto.
- *"calib\_panel.h"*: es el archivo que contiene todos los identificadores del panel de calibración.
- *"canny.h"*: contiene el algoritmo detector de bordes de Canny. Este algoritmo fue escrito en el *Robot Vision Group* de la Universidad de Edinburgo. Sus autores son: Bob Fisher, Dave Croft y A. Fitzgibbon en septiembre de 1986. Para este programa se le hicieron algunas modificaciones.

- “*listas.h*”: contiene las funciones necesarias para la creación y manejo de listas, así como su transformación a arreglo o estructuras.
- “*mapeo.h*”: es la librería que calcula los coeficientes de los polinomios de corrección de la distorsión y que hace precisamente la corrección de las imágenes que llegan de las cámaras.
- “*matlab.h*”: contiene funciones que fueron programadas y probadas inicialmente en MATLAB, debido a que utilizan muchos cálculos con matrices.
- “*matrix.h*”: esta librería encierra muchas funciones para hacer cálculos con matrices de diferentes tamaños. Funciones básicas de sumas o multiplicaciones, productos internos, determinantes, normas, etc.
- “*memory.h*”: contiene las funciones que reservan y liberan espacio para los diferentes tipos de estructuras que se manejan en el programa (Matrix\_f, Matrix\_i, CVMat, Puntos, etc.).
- “*misc.h*”: es una librería que contiene funciones de todo tipo: para hacer ordenamientos, reunir zonas, encontrar máximos y mínimos, hacer histogramas, hallar medias, calcular distancias Euclidianas, etc.
- “*panel.h*”: es el archivo que contiene todos los identificadores del panel inicial.

# Bibliografía

- Arteaga Pérez, M. A. 1998. On the properties of a dynamic model of flexible robot manipulators. *ASME Journal of Dynamic Systems, Measurement, and Control* 120. 8-14.
- Åström, K. J. y B. Wittenmark. 1990. *Computer-Controlled Systems Theory and Design, 2nd edition*. New Jersey: Prentice-Hall.
- Bedford, A. y W. Fowler. 1996. *Mecánica para Ingeniería: Estática*. México: Addison-Wesley Iberoamericana.
- Born, G., H. Hübscher y H. Lochhaas. 1992. *Querschnitt Physik und Technik*. Braunschweig: Westermann Schulbuchverlag.
- Bruce, V., P. Green y M. A. Georgeson. 1997. *Visual Perception: Physiology, Psychology and Ecology*. United Kingdom: Psychology Press.
- Bryant, A. y J. Bryant. 1989. Recognizing shapes in planar binary images. *Pattern Recognition* 22(2). 155-164.
- Castillo, A. 2002. *Adaptación de dos robots industriales para su utilización en el desarrollo de nuevas técnicas y algoritmos de control*. México: Tesis de licenciatura: ENEP Aragón, U.N.A.M.
- Couch II, L. 1993. *Digital and Analog Communication Systems, 5th edition*. New Jersey: Prentice-Hall.
- Dean, T., J. Allen y Y. Aloimonos. 1995. *Artificial Intelligence: Theory and Practice*. U.S.A.: The Benjamin/Cummings Publishing Company.
- Escobar, L. 2002. *Apuntes de clase: Procesamiento Digital de Señales*. México: Semestre 02-1, Facultad de Ingeniería, U. N. A. M.
- Gonzalez, R. y R. Woods. 1993. *Digital Image Processing*. Perceptics Corporation: Addison-Wesley.
- Groover, M., M. Weiss, N. Odrey y R. Nagel. 1990. *Robótica Industrial: Tecnología, programación y aplicaciones*. México: McGraw-Hill.
- Halliday, D., R. Resnick y K. Krane. 1998. *Física, 4a edición*. México: CECSA: Compañía Editorial Continental, S. A. de C. V.



- Hsu, C. y J. Huang. 1990. Partitioned hough transform for ellipsoid detection. *Pattern Recognition* 23. 275-282.
- IEEE 2001. 2001: Bad news, and good. *IEEE-Spectrum* December 2001. 9.
- Kanatani, K. 1995. *Geometric Computation for Machine Vision*. New York: Oxford University Press, Inc.
- Kuo, B. C. 1997. *Sistemas de Control Digital*. México: CECSA: Compañía Editorial Continental, S. A. de C. V.
- Maravall, D. 1993. *Reconocimiento de Formas y Visión Artificial*. España: RA-MA Editorial.
- Maryon-Davis, A. 1985. *Nuestro Asombroso Cuerpo*. México: Fernández Editores.
- McKenzie, D. y S. Protheroe. 1990. Curve description using the inverse hough transform. *Pattern Recognition* 23. 283-290.
- Mitiche, A. y G. Habelrih. 1989. Interpretation of straight line correspondences using angular relations. *Pattern Recognition* 22. 299-308.
- Murray, R. M., Z. Li y S. S. Sastry. 1994. *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, Florida, U. S. A.: CRC Press.
- Oppenheim, A. y A. Willsky. 1998. *Señales y Sistemas, 2a edición*. México: Prentice-Hall.
- Pratt, W. 1991. *Digital Image Processing, 2 edición*. New York: John Wiley & Sons.
- Schalkoff, R. 1990. *Digital Image Processing and Computer Vision*. New York: John Wiley & Sons.
- Skowronski, J. M. 1986. *Control Dynamics of Robotic Manipulators*. U. S. A.: Academic Press, Inc.
- Solar, J. 1998. *Cinemática y Dinámica Básicas para Ingenieros*. México: Trillas.
- Sonka, M., V. Hlavac y R. Boyle. 1993. *Image Processing, Analysis and Machine Vision*. United Kingdom: Chapman & Hall.
- Spong, M. W. y M. Vidyasagar. 1989. *Robot Dynamics and Control*. New York: John Wiley & Sons.
- Xiao-Wei, T. y B. Dubuisson. 1990. 3-D information derivation from a pair of binocular images. *Pattern Recognition* 23. 223-235.

TESIS CON  
FALLA DE ORIGEN