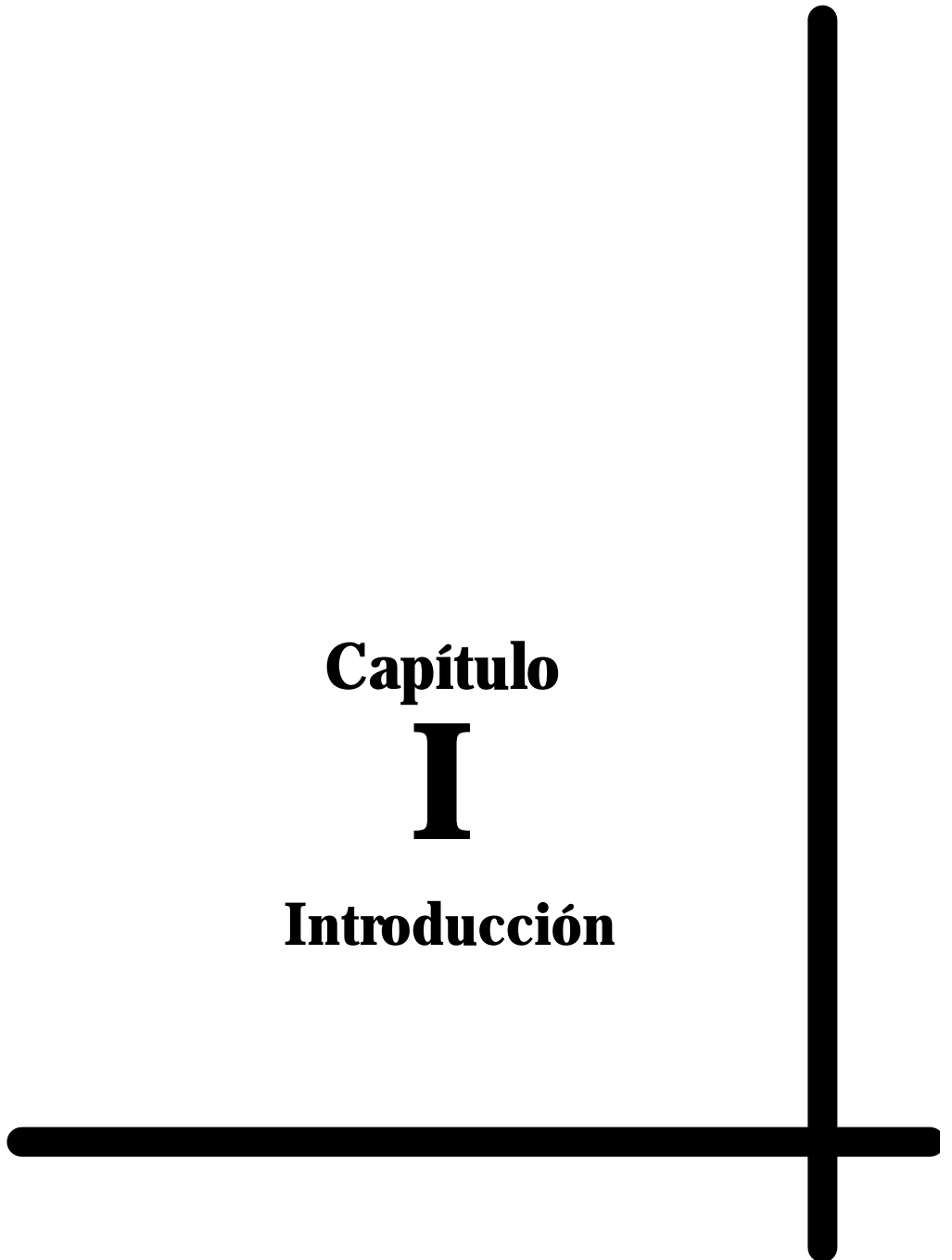


**Capítulo**

**I**

**Introducción**



## I

### 1.1 Introducción

**E**s hasta últimas fechas cuando el decreciente costo del poder de computación acoplado a la capacidad de una mayor producción han impulsado la introducción de mecanismos automáticos en las fábricas; pero además, han sido introducidos en otros campos como son: la medicina, la exploración y la biomedicina. Las clases de tareas a las que un autómata pueden ser orientado son aquellas que pueden ser peligrosas para los humanos, o aquellas en ambientes hostiles como el fondo del océano y el espacio exterior, y también las que resultan aburridas y repetitivas y que los humanos encuentran poco estimulantes y de igual manera aquellas que requieren gran precisión y exactitud.

Sin embargo tareas que son simples para el ser humano pueden llegar a ser extremadamente complejas y difíciles para un autómata, por ejemplo: un humano puede encontrar visualmente un objeto arbitrario, tomarlo y moverlo a otra posición evitando posibles obstáculos durante el trayecto y sin dañarlo si éste es frágil, esta tarea tan sencilla está mas allá de las capacidades de los robots en uso en nuestros días.

La mayoría de las tareas de los robots corresponden a secuencias de recoger y poner en un mundo no cambiante y totalmente conocido, estos robots son previamente enseñados para realizar una serie de movimientos que son equivalentes al movimiento humano. Los movimientos enseñados asumen un mundo que no cambia y en caso de presentarse, los robots no saben que hacer.

[Allen '87]

Idealmente un autómatas debe contar con características de inteligencia; esto quiere decir que debe ser capaz de adaptarse y desenvolverse en ambientes cambiantes que lo rodean, para que esto suceda debe tener receptores que monitoreen las variables del ambiente en el que se encuentre. Para poder “sentir” el ambiente, los robots se valen de los transductores también llamados sensores, éstos son los receptores capaces de convertir una forma de energía en otra, los transductores más comunes son los que convierten cualquier tipo de energía (por ejemplo: calorífica, mecánica, etc.) a energía eléctrica. Existe un sinnúmero de sensores, pero los más comúnmente usados en el campo de la robótica, son del tipo ultrasónico, de luz infrarroja, magnéticos y de tacto; pero sin lugar a duda una de las herramientas más poderosas con las cuales es posible obtener una mayor cantidad de información del medio ambiente que rodea es la visión; dicho en otras palabras, la adquisición de una imagen por medio de una cámara y su correspondiente procesamiento nos puede dar una cantidad de información sorprendente al aplicarle algoritmos que más se adecuen.

El campo de procesamiento digital de imágenes en estéreo es un área de Ciencias de la Computación relativamente nueva e inexplorada; de reciente estudio y su implantación ha sido muy escasa; la visión por computadora puede llegar a ser tan poderosa como para extenderse más allá del simple reconocimiento, por medio de ésta, se puede construir descripciones explícitas y con significado de los objetos físicos de una imagen, El rango de aplicaciones en las cuales se ha utilizado la visión por computadora, ha sufrido un incremento considerable en los últimos años y las áreas que más se han visto impactadas son las de medicina, por ejemplo: en la ayuda de los diagnósticos cuando se usa equipo como el ultrasonido y el tomógrafo para la predicción de anomalías; en la fotografía aérea y satelital para detectar zonas de interés; en el tráfico vehicular al detectar zonas de congestiónamiento y para la identificación de vehículos; en la seguridad en el correo al analizar y procesar documentos; aunque cabe señalar que estas tres últimas áreas todavía se encuentran en etapas iniciales de desarrollo y aún hay más disciplinas que pueden beneficiarse al utilizar la visión por computadora. [Boyle '88].

El desafío en la construcción de los sistemas robóticos es incorporar diferentes sensores en un sistema que haga uso de los datos provenientes de ellos.

Mucho del trabajo en robótica relacionado con el sensado del entorno, ha sido tratado usando un solo sensor en el sistema para determinar las propiedades del ambiente esto puede llegar a ser muy difícil,

debido a que no todos los sensores son capaces de determinar las propiedades del ambiente que son importantes. Por ejemplo: un sistema de visión utilizando proyecciones 2D tiene demasiadas dificultades para determinar las formas 3D. Muchos operadores de características han sido desarrollados para tratar de “aislar” la información del sistema que produzca información del fondo y superficie, ejemplo de esta separación son forma de contorno, forma de sombra, etc. Una idea prometedora es usar todos estos operadores de forma, en un sistema que integre los resultados.

[Allen '87]

Mucho han sorprendido los robots con lo que realizan en la actualidad; sin embargo, aún siguen siendo “tontos” y están muy lejos de la forma en que se desea que se desempeñen, este trabajo pretende contribuir al desarrollo de las áreas de percepción de ambiente por un autómata y modelado de mundos; para lograr lo anterior será necesario investigar, construir e implementar un prototipo que mediante la integración de diferentes sensores en un sistema, monitoree el ambiente que rodea a un autómata mediante el análisis y procesado de estos datos y con ayuda de técnicas de conversión 2D a 3D haciendo uso de rutinas de visualización 3D finalmente se obtendrá una representación en tercera dimensión del ambiente en una computadora.

De esta forma las investigaciones e implantaciones a realizar en este trabajo quedarán integradas por cuatro módulos o partes fundamentales que serán como a continuación se indica:

- a) Se investigará sobre métodos existentes utilizados en robótica que permitan percibir el entorno y se implantarán según sea necesario.
- b) Se investigarán, analizarán e implantarán los algoritmos que analicen la información recibida, especialmente algoritmos de procesamiento digital de imágenes que conviertan la secuencia de tomas provenientes de cámara(s) a datos tridimensionales de los objetos.
- c) Con los datos obtenidos mediante el procesamiento digital de imágenes y otros métodos auxiliares, analizando los obstáculos con sus coordenadas absolutas y relativas, se precisarán la posición y forma de los obstáculos.
- d) Se Implantarán rutinas de visualización 3D con las cuales se construya el mundo virtual en tercera dimensión del ambiente real en base a la posición y dimensión de los obstáculos previamente identificados.

## 1.2 Antecedentes

En los últimos años se ha concentrado un gran esfuerzo hacia las investigaciones en predicción de trayectorias y reconocimiento del ambiente en robótica y ya son muchas las universidades que manejan áreas de investigación orientadas en esta dirección por ejemplo:

En la Universidad de Purdue dentro del laboratorio de visión robótica “Robot Visión Laboratory” se trabaja en un robot conocido como Peter, en su primera versión (el ahora conocido old-peter) en el control fue usado un Motorola 68010-12MHz, con VxWorks y como elementos de adquisición de datos fueron usadas dos cámaras blanco y negro (fijas) y un sensor de ultrasonido.

Dentro de otros trabajos realizados en la universidad de Purdue en esta misma área, se puede citar el llamado “FINALE”; requiere de modelados geométricos simples del espacio interior. Estos modelos pueden ser creados simplemente con conocer la altura del techo y la localización de las principales esquinas de la habitación. El sistema puede llegar a funcionar aun cuando algunas de las características falten en el mapa de modelo.

Otro trabajo desarrollado, es uno basado en redes neuronales que permite mantener al robot centrado en la mitad de su trayectoria y genera secuencias de comandos cuando el robot necesita una corrección en su curso para evadir obstáculos. Únicamente es necesario proporcionarle al sistema modelos topológicos del espacio interior. Los modelos de esta forma son fácilmente creados. De hecho el sistema se ayuda de lógica difusa para un mejor control del mismo. En el procesamiento de la imagen es usada la detección de bordes, el filtrado y la segmentación basada en la diferencia de grises.

Una última mención acerca de esta universidad, la debemos situar en 1989 cuando realizaron el proyecto PSEIKI, que realmente fue ejercicio de la operación de máquinas usando interpretación de imágenes, lo importante de este sistema, es que no es necesario diseñar un sistema de navegación basado en la visión sino hacer una interpretación completa de la imagen entera que es necesaria para posicionar al robot. [Purdue ‘02].

Un robot con un notable desarrollo tecnológico y considerado hasta la fecha por muchos autores como el más avanzado es el que ha desarrollado la compañía Honda; este robot llamado Asimo, tiene una altura de 1.6 metros, y pesa aproximadamente 130 Kg. Una especie de mochila colocada en la “espalda” del mismo contiene la computadora de control y la batería. El robot ha demostrado ser capaz

de caminar y subir escaleras, y es llamado técnicamente P3, está equipado con girómetros, sensores de fuerza gravitacional, sensores de fuerza en sus rodillas y cámaras de visión. El robot es autónomo y es maniobrable tanto como lo permita la comunicación inalámbrica ethernet.

Poco se conoce respecto al sistema de control, pero se sabe que es controlado por una computadora, la cual contiene los algoritmos de procesamiento digital para analizar las imágenes provenientes de las cámaras. [Honda '02].

El Instituto Tecnológico de Masachussets es una de las universidades con mayor camino y experiencia en cuanto a autómatas se refiere, ellos han construido una gran variedad de robots, algunos de los cuales se mencionan a continuación:

Frankie, el propósito de este era la búsqueda de construcción de tecnología de punta a bajo costo, implementa evasión de colisiones para recorrido en el laboratorio, usa radio modems para comunicación bidireccional con una máquina con S.O. Unix.

Toto explora ambientes de oficina desconocidos, construye mapas y navega.

TJ es un robot móvil de navegación de oficinas y el objetivo es tratar de entender el lenguaje natural escrito con el que la gente nombra lugares y rutas, de esta forma se intenta comandar el destino del robot mediante lenguaje natural, es decir: se intenta hacer un puente entre las instrucciones dadas por una persona, el control del robot y la predicción de sus trayectorias, ya que esta forma de emitir instrucciones es muy distinta a las formas tradicionales en las cuales entienden el mundo mediante sus sensores. (Este proyecto fue parte de trabajo de Tesis de Maestría) [MIT '98]

Otro trabajo desarrollado en el MIT es el Sistema Robótico de Silla de ruedas, donde como su nombre lo indica, una silla de ruedas eléctrica es semi-autónoma en un ambiente de interior, su control es proporcionado mediante una Powerbook Macintosh. El usuario comanda una dirección general y la silla se encarga de rutinas como la evasión de obstáculos y seguimiento de paredes. La interfaz con el usuario permite operar en tres modos: manual, joystick e interfaz de usuario. En modo manual, la silla de ruedas opera como una silla eléctrica normal; en el modo joystick el usuario comanda las direcciones a través de un joystick mientras que el robot evita los objetos en respuesta a una trayectoria; en el de interfaz, el usuario interactúa con el robot únicamente a través de una interfaz adaptable.

Trabajos con esta orientación incluyen: navegación exterior usando sistema de visión, control compartido entre persona y robot, mediante el potencial electro-oculográfico.

El MIT trabaja en lo que son los algoritmos de visión con módulos de procesadores de información visual que es un hardware especializado para realizar operaciones de convolución de manera muy rápida.[MIT '97]

En la Universidad de Utah se ha desarrollado el robot EGOR; este robot compañero de laboratorio cuenta con un subsistema de proximidad, está muy bien equipado con 24 sonares, 8 sensores infrarrojos cámara de video Panasonic en blanco y negro conectada a través de una tarjeta capturadora de video Win/TV y como plataforma de procesamiento se encuentra una PC, ésta última es una Notebook y cuenta con "docking station".

En esta universidad se ha desarrollado XSim, éste es un software de simulación para el robot EGOR. La tarea que realiza dicho software es la de simular las acciones del robot en la pantalla, permitiendo pruebas extensivas de programas de control del robot.

En julio de 1994, EGOR obtuvo un séptimo lugar en navegación y sexta posición en visión durante la "AAAI Mobile Robotics Competition" en Seattle, Washington.[Utah '94]

Además de las anteriores, muchas otras universidades importantes en el extranjero han desarrollado y continúan con trabajos de investigación en el área de visión y robótica, entre estas universidades se encuentra: la Universidad de Rochester y la Universidad de Western Australia.

En la Universidad de Rochester se desarrollaron tres robots, los robots Cyclops y Map-maker, la investigación en el área de visión está orientada a la obtención de características en tiempo real mediante sistemas de adquisición de imágenes con 2 cámaras de video a color.[Rochester]

En cuanto a métodos alternativos de percepción del entorno y robótica, podemos mencionar la implantación de varios robots de la Universidad Iberoamericana como el Acamapixtli en donde el desplazamiento del robot es guiado mediante un haz infrarrojo. El robot cuenta con una corona de sensores infrarrojos para determinar la dirección de la cual proviene el haz. Otro robot, el Huitzilihuitl se vale de sensores colocados en la parte inferior del mismo para realizar tareas de seguimiento de trayectorias dibujadas en el piso y utiliza otra serie de pares emisor - receptor infrarrojos alrededor del mismo para realizar la evasión de obstáculos.

En el campo del procesamiento digital de imágenes en dos dimensiones, se puede citar el trabajo realizado en la Universidad Iberoamericana en donde se estudian varios conceptos sobre el análisis digital de imágenes aplicado al desplazamiento de objetos. Específicamente se trabaja con el reconocimiento de contornos de objetos y en base a la posición del objeto detectado, se comanda para realizar una rotación del mismo. Para el análisis de la imagen, primero es usada la segmentación para realzar la región de interés, una vez hecho esto se obtiene el contorno mediante el operador Sobel del método de Gradiente, un proceso intermedio implementado fue el adelgazamiento de los bordes mediante el algoritmo MAT (Medial Axis Transform) y el análisis de forma fue implementado mediante la transformada Hough.[Castellanos '99]

Ahora bien, en esta Universidad en lo que es el campo del procesamiento digital de imágenes orientado al análisis del medio ambiente que rodea, se cita el trabajo realizado por J. A. Chavez, el cual estudia técnicas para la exploración de ambientes de un robot móvil valiéndose del procesamiento de imágenes y de una manera eficiente y sencilla se lleva a cabo una implementación en donde se realiza el control de un robot Khepera para evadir obstáculos y además, se dibuja en forma de plano 2D el mapa del mundo explorado haciendo uso de datos provenientes de decoders del robot. La información es obtenida por medio de una cámara Blanco y negro y esta información es procesada; para este procesamiento digital de imágenes se hace un análisis del histograma y un conteo sobre el eje x de la imagen del número de píxeles de esa columna que corresponden a un objeto, así si un objeto es detectado a la izquierda de la imagen se enviará una señal para mover el robot hacia la derecha y viceversa.[Chavez '01]

### **1.3 Estado Actual**

En la Universidad de Purdue se encuentran trabajando en la segunda edición de Peter, mejor conocido como "New Peter", el robot es controlado por un sistema basado en PC, doble Pentium Pro a 200 Mhz con un Sistema operativo RT-Linux. Como hardware de monitoreo del ambiente que se explora fueron usadas dos cámaras de color montadas de tal forma que ayuden a percibir la volumetría de los objetos, a este último sistema se le llama "stereo head", un sensor de rango láser y un par de sensores de ultrasonido ayudan a determinar distancias y rangos. La principal razón de emigrar a Linux fue el costo, y la razón del uso de un sistema basado en una computadora es la facilidad de escalamientos futuros.[Purdue '02]



Referente a la parte de visión que es el tema que más nos interesa, cabe mencionar que un grupo de seis científicos de la universidad de Carnegie Mellon en los Estados Unidos están desarrollando el sistema de visión y navegación para el robot Honda. Principalmente están haciendo uso de determinadores de rango láser y visión estéreo para construir modelos 3D de los objetos y del ambiente. [Honda '02]

Proyectos para autómatas en visión se están desarrollando exitosamente; se han hecho sistemas para el reconocimiento de objetos curvos 3D y han reconocido color, texturas, movimiento y expresiones faciales. Estos sistemas están siendo usados en robots móviles autónomos y helicópteros. De igual manera se encuentran en desarrollo sensores computacionales VLSI, los cuales sensan y procesan información en un solo chip.

Otra universidad que se encuentra realizando trabajos en el área de procesamiento de imágenes y reconocimiento de objetos, incluyendo la reconstrucción 3D a partir de imágenes 2D es la Universidad de Cambridge, se encuentran realizando investigaciones en tópicos como son el reconocimiento de rostros y segmentación de imágenes.

En el campo de la robótica cubren aspectos como inspección de calidad usando la visión por computadora e interfaces hombre-máquina mediante el uso de gestos visuales, En el aspecto de la imagen médica, se incluye la adquisición, registro y segmentación de imágenes 3D de ultrasonido para el diagnóstico médico. [Cambridge '01]

Trabajos como el proyecto M4, tratan de simular animales y para el reconocimiento se basan en un sistema de visión dado por cuatro cámaras montadas 2 en cada ojo. Además hay proyectos robóticos interesantes en donde la construcción de robots tratan de asemejar el movimiento y comportamiento de los insectos que por supuesto son criaturas menos inteligentes y complejas que los humanos, pero no por esto se quiere decir que su construcción sea fácil, es relativamente mas sencilla que un sistema muy complejo que trate de asemejar al humano, esta nueva forma de evolución de la robótica nos podría ayudar a la construcción de máquinas más complejas y sofisticadas al estudiar y entender el comportamiento y movimientos de seres mas sencillos.

Interesantes proyectos se encuentran en desarrollo en la Universidad de Rochester, por ejemplo Randal Nelson y Raj Rao se encuentran trabajando en diferentes versiones de reconocimiento basado

en apariencia, estos modelos no requieren conocimiento anterior de modelos de los objetos. También Olac Fuentes en esta misma universidad junto con la universidad de Utah y el MIT han implementado una forma de manipulación a una amplia variedad de objetos sin el uso de modelos de los mismos.

Otro trabajo relacionado es el de Chris Brown, quien junto con un grupo de estudiantes están trabajando en visión para un robot móvil. Una aportación en esta área son las técnicas de manipulación sin calibración que permite mejoras a las interfaces hombre-máquina, por ejemplo: donde el robot es instruido por medio del señalamiento de objetos mediante el uso de guantes cibernéticos, el mouse sobre una fotografía o por el dibujo en la pantalla en un software para la manipulación.

Este paradigma de los métodos sin calibración es contrastante respecto a los sistemas actuales de uso, en donde un ambiente es desarrollado cuidadosamente y un modelo exacto a priori es requerido y la manipulación de la tarea es descrita en cuántos milímetros el robot necesita moverse en un mundo fijo en un marco de coordenadas.

Otro foco principal en el trabajo en esta misma Universidad de Rochester es el uso de visión avanzada o técnicas de realidad virtual en experimentos de psicofísica para aprender acerca del cerebro humano y el sistema visual humano. En uno de los proyectos se estudia la atención visual humana y la memoria a corto plazo en un mundo virtual en donde los objetos pueden hacerse aparecer y desaparecer, o cambiar de color en tiempo real controlado por la mirada fija del usuario. Usando estos métodos se ha encontrado que el humano hace uso de una cantidad mínima sorprendente de memoria interna durante tareas típicas de manipulación. Y en su lugar accesa información en el mundo cuando lo requiera; de esta forma se están aplicando estos conocimientos y requiriendo únicamente una mínima calibración y usando técnicas de visión para argumentar secuencias de video en tiempo real con gráficas simuladas de objetos reales.[Rochester]

A nivel nacional también existe investigación en el área, aunque pocos trabajos hay que hagan estudio de técnicas para el procesamiento estéreo; un trabajo notable que de hecho se encuentra en desarrollo y cuya mención es muy importante, se está desarrollando en la Universidad Nacional Autónoma de México donde se está desarrollando una alternativa de visualización del entorno explorado mediante la toma de fotogramas cada determinado intervalo y mediante algoritmos especiales de ensanchamiento o encogimiento de la imagen según sea necesario que se apliquen a las fotos dan un efecto de

tridimensionalidad al ojo humano del ambiente que se explora. Además de la UNAM, en este proyecto se encuentran trabajando otras instituciones como el ITAM y otras universidades en el extranjero.