

# La Visión Artificial en la Robótica

Ing. Rogelio Rojas Hernández  
UPIITA

Dr. Ramón Silva Ortigoza ♣

M. en C. María Aurora Molina Vilchis. ♦  
Profesores del CIDETEC

**E**l vertiginoso avance de las computadoras ha permitido que la visión artificial se incorpore a una gran cantidad de disciplinas, apoyando e incluso sustituyendo a los sistemas que antaño se utilizaban. Este trabajo tiene como finalidad presentar un panorama general de las bases, trabajos, proyectos y avances sobre la visión artificial en aplicaciones de robótica.

---

## INTRODUCCIÓN

---

Con la cantidad de información disponible actualmente, no es de extrañar el enorme número de referencias sobre robots a la que una persona puede acceder (basta teclear *robot* en el popular buscador Google y se obtendrán más de 78 millones de enlaces). Sin embargo, pocas veces alguien puede explicar lo que verdaderamente son estos artefactos en el mundo real.

La palabra *robot* viene del vocablo checo *robota*, que significa “servidumbre”, “trabajo forzado” o “esclavitud”; el término fue utilizado por primera vez por Karel Čapek en su novela *R.U.R. (Rossum’s Universal Robots)*. Fue un científico

y divulgador de la ciencia, Isaac Asimov, quien introdujo por primera vez el término robótica con el sentido de disciplina científica encargada de construir y programar robots [1]; pero en definitiva no fue la comunidad científica y académica la que difundió de forma masiva este concepto.

Hollywood, como industria cinematográfica, ha hecho una gran labor en cuanto a la divulgación de la idea de un robot. La amplia filmografía realizada recurriendo a este tema proporciona una imagen mental de lo que son estos artefactos. Así pues, cualquiera sabe quienes son *C3PO* o *R2D2* (curiosamente mexicanizado como Arturito) e identifica claramente que son robots (ver la **Figura 1**).



Figura 1: C3PO y R2D2, robots emblemáticos de nuestra cultura.

Estos robots imaginarios están aún lejanos de lo que realmente existe, pero nos dan la idea de lo que es un robot; aunque ésta no es estrictamente correcta, es una idea al fin, que es lo importante. Afinar esta imagen mental y convertirla en un concepto general y universal es bastante complicado; lo cierto es que aún no nos hemos puesto de acuerdo del todo.

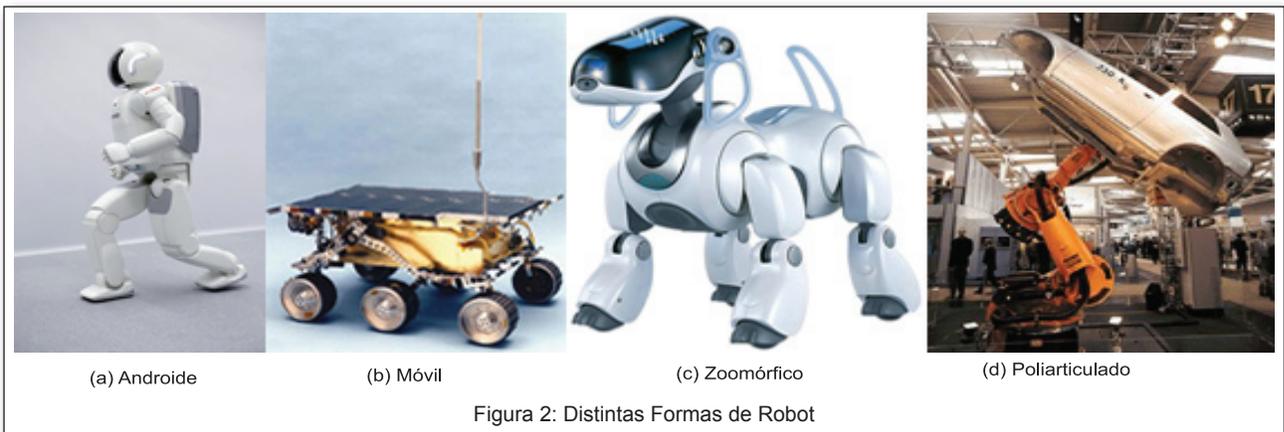
En este artículo se presenta un panorama de los elementos, desarrollo y aplicaciones de la visión artificial, específicamente en el campo de la robótica. El documento consta de seis secciones; en la Sección 2 se define lo que se puede calificar como un robot. En la Sección 3 se introduce al lector a las posibilidades de interacción con las que puede contar un robot. En la Sección 4 se detallan los conceptos fundamentales de la visión artificial. En la Sección 5 se tratan los trabajos más relevantes sobre la aplicación de la visión artificial en robótica y, finalmente, en la Sección 6 se presentan las conclusiones.

---

## DEFINIENDO UN ROBOT

---

Existen suficientes definiciones como para llenar una biblioteca entera, pero tomando una de las más aceptadas se considera que los robots son dispositivos mecánicos capaces de realizar tareas



que podrían desempeñar los seres humanos. Bajo esta definición, las máquinas que intervienen en una cadena de ensamblaje de automóviles son robots, aun cuando no son muy inteligentes en apariencia; hoy en día a estas máquinas las llamamos robots industriales.

Desde inicios de la década de los 90 y dependiendo del país, han aparecido con rapidez nuevos robots más inteligentes. A estos se les ha dotado de capacidades perceptivas (vista, oído y tacto principalmente), aunque reducidas en comparación con las capacidades de los seres vivos. Estas nuevas máquinas, que en apariencia son más inteligentes y que se suelen adaptar a entornos cotidianos, se denominan por lo general robots de servicio o robots inteligentes. Pero al analizarlos cuidadosamente, muchos robots de servicio no son tan inteligentes como parecen, e inclusive existen robots industriales desarrollados años atrás, que realizan tareas mucho más complicadas.

### COMO SOCIEDAD, ¿QUE ES UN ROBOT?

Habitualmente se debate sobre que máquina es un robot y cual no lo es, porque realmente es complicado llegar a un acuerdo cuando el tema es tratado por el público

general. La batidora, un electrodoméstico común, realiza una función que puede llevarse a cabo por un ser humano, y siguiendo estrictamente la definición, la batidora es un robot aún cuando no sea más que una herramienta eléctrica. De hecho, se pueden encontrar a la venta robots de cocina cuya única capacidad es triturar alimento.

También existen ejemplos de máquinas que se transforman en robots; un automóvil es una máquina, así está definido hoy en día [2]. Sin embargo, al combinarlo con un sistema informático, electrónico y mecánico que lo controle y le permita llegar a un destino sin la intervención de un ser humano, se convierte en un robot. Evidentemente, en las sociedades más desarrolladas tecnológicamente, una batidora no se considera un robot, si se le compara con los robots de última generación que se conocen. En realidad, se trata de un problema de contraste social; en el año 1985 nadie ponía en duda que una batidora fuera un robot de cocina y ahora resulta increíble decir que tanto una batidora como un humanoide son robots. Quizás en el año 2030 el conductor robotizado de un auto sea tan poco llamativo para el público como una batidora en el presente.

Lo más curioso de esta discusión es que no podemos definir a conciencia lo que es un robot, pero si los hemos clasificado en cuatro formas, básicamente diferenciados por sus capacidades [3].

1. **Androides:** son robots que imitan la forma y comportamiento humano; suelen ser solo de experimentación, pero existen a la venta modelos cuya función es ser juguete o una mascota. La principal limitante de este modelo es la implementación del equilibrio a la hora del desplazamiento, pues es bípedo (ver la **Figura 2(a)**).

2. **Móviles:** se desplazan mediante una plataforma rodante (ruedas u orugas); estos robots aseguran el transporte de piezas de un punto a otro (ver la **Figura 2(b)**).

3. **Zoomórficos:** logran la locomoción imitando a los animales. Estos robots son utilizados en ambientes específicos, tales como el estudio de volcanes o la exploración espacial (ver la **Figura 2(c)**).

4. **Poliarticulados:** suelen imitar un brazo humano, con algunas variantes y/o limitantes. Su utilidad es principalmente industrial, para desplazar elementos que requieren fuerza y cuidado en su manejo (ver la **Figura 2(d)**).

### LOS SENTIDOS DE UN ROBOT

En un ser humano, un sentido es una fuente de información sobre el mundo que lo rodea, para que en base a dicha información el cerebro tome una decisión. Los sentidos proveen datos sobre temperatura, luz, sonido, fuerza, olor, sabor, etc.; vista, tacto, oído, gusto y olfato son los sentidos que tradicionalmente se asocian a una persona, aunque popularmente se acepta la existencia

lados, que solo requieren de algún sentido en especial, incluso más desarrollado que su contraparte humana.

### LA VISIÓN ARTIFICIAL

Si solo se va a incorporar un sentido al robot, la selección se inclinará naturalmente hacia el que más información proporcione sobre el ambiente; en general, la información que suele requerirse

y su propósito es “entender” las características de una imagen y en consecuencia generar una acción. No se debe confundir la visión artificial con el procesamiento digital de imágenes (PDI); el PDI es un conjunto de técnicas que se aplican con el objetivo de mejorar la calidad ó facilitar la búsqueda de información en una imagen digital. El PDI es lo que suelen aplicar los laboratorios fotográficos al corregir “ojos rojos”, montar escenarios, restaurar, etc. La visión artificial está profundamente apoyada en el

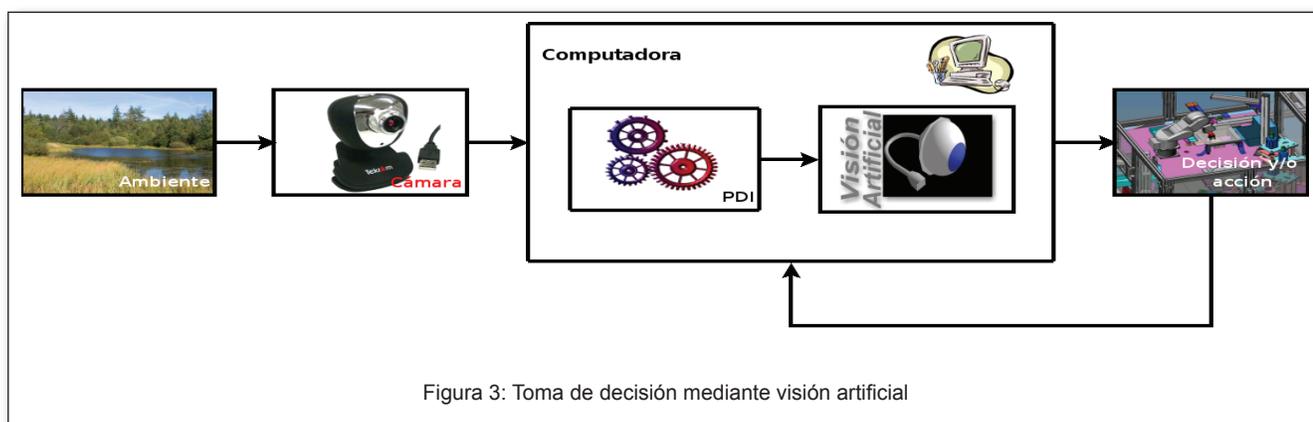


Figura 3: Toma de decisión mediante visión artificial

de otros sentidos como la empatía, la telepatía y otras clases de percepción extrasensorial.

Extrapolando la misma idea a los robots, ya se dispone de la tecnología para dotarlos de tales sentidos, mediante el uso de computadoras que sustituyen al cerebro, y micrófonos, cámaras, sensores, etc. que sustituyen a los ojos, la piel, los oídos, la lengua y la nariz. En general, resulta demasiado costoso e impráctico para los proyectos que se requieren y/o desarrollan dotar a un robot con los mismos sentidos con los que cuenta un ser humano; estas condiciones se presentan debido a lo corto que suelen ser los tiempos y presupuestos, y a que los robots normalmente se desempeñan en ambientes específicos y/o contro-

son dimensiones físicas (tamaño, volumen, posición, velocidad, aceleración, dirección, etc.), por lo que el sentido seleccionado suele ser la vista.

Una imagen puede suministrar mayor cantidad de información sobre las dimensiones de un lugar, un objeto o una persona, que la que podría obtenerse a través de un olor, un sonido, un sabor o una textura. En adelante se utilizarán los términos imagen e imagen digital de forma indistinta, puesto que las tecnologías utilizadas para dotar de vista a los robots solo pueden utilizar imágenes digitales.

La tecnología que se encarga de dotar del sentido de la vista a un robot se denomina *visión artificial* o visión por computadora,

PDI, pero de ninguna forma son lo mismo. El proceso de obtener información a través de la visión artificial para la toma de una decisión puede esquematizarse como muestra la **Figura 3**.

#### AMBIENTE

El ambiente es simplemente el lugar donde el robot desarrollará sus actividades. Puede ser una nave industrial, un laboratorio, una mesa de pruebas, etc.

#### CÁMARA

En visión artificial la cámara es el elemento que reemplaza al ojo humano. La cámara tiene como propósito convertir la luz presente en el lugar de trabajo, en una señal eléctrica que una computadora

pueda utilizar. Las cámaras ya no resultan extrañas, debido a la gran cantidad de modelos disponibles en el mercado, pero cabe mencionar que en el ambiente científico e industrial no cualquier tipo de cámara resulta útil para los fines que se persiguen. Los dispositivos que se han diseñado específicamente para las tareas en los ambientes antes mencionados son denominados cámaras de visión artificial. Las cámaras de visión artificial son más sofisticadas que las convencionales, ofreciendo un completo control de los tiempos y señales de la cámara [4], [5]. Existen infinidad de marcas y modelos de este tipo de dispositivos, cada uno especializado en alguna necesidad, pero todas basadas en dos tecnologías para transformar la luz en una señal eléctrica: CCD y CMOS.

### TECNOLOGÍA CCD

Un CCD (*Charge-Coupled Device*, Dispositivo de Cargas Eléctricas Interconectadas) es un circuito integrado que contiene un número determinado de condensadores enlazados o acoplados. Bajo el control de un circuito interno, cada condensador puede transferir su carga eléctrica a uno o a varios de los condensadores que estén a su lado en el circuito impreso. El CCD es uno de los elementos principales de las cámaras fotográficas y de video digitales; en éstas, el CCD trabaja conjuntamente con células fotoeléctricas diminutas que registran la imagen, formando así el sensor. Desde allí, la imagen es procesada por la cámara y registrada en la tarjeta de memoria. La capacidad de resolución o detalle de la imagen depende del número de células fotoeléctricas, dicho número se expresa en píxeles; dichos píxeles registran tres colores diferentes: rojo, verde y azul (*RGB*, del inglés Red, Green, Blue), por lo cual tres

píxeles, uno para cada color, forman un conjunto de células fotoeléctricas capaces de captar cualquier color en la imagen. Para conseguir esta separación de colores la mayoría de las cámaras del tipo CCD utilizan un filtro de Bayer [6], que proporciona una trama para cada conjunto de cuatro píxeles, de forma que un píxel registra luz roja, otro luz azul y dos más se reservan para la luz verde (por imitación al ojo humano, que es más sensible a la luz verde).

### TECNOLOGÍA CMOS

Un *Active Pixel Sensor (APS)* es un sensor que detecta la luz basado en tecnología CMOS y por ello es más conocido como sensor CMOS. Gracias a esta tecnología de construcción es posible integrar más funciones en un chip sensor, tales como control de luminosidad, corrección de contraste, conversión analógico-digital, etc. El APS, al igual que el sensor CCD, se basa en el efecto fotoeléctrico, ya que está formado por numerosos fotodiodos, uno para cada píxel, que producen una corriente eléctrica que varía en función de la intensidad de luz recibida. En el CMOS, a diferencia del CCD, se incluye un convertidor digital en el propio chip y se incorpora un amplificador de la señal eléctrica en cada fotodiodo mientras que en un CCD se tiene que enviar la señal eléctrica producida al exterior y amplificarla. Al igual que ocurre con el CCD, los fotodiodos captan únicamente intensidad lumínica, para lo que se suele emplear nuevamente el filtro de Bayer [6] para la distinción de los colores. Mediante este filtro, unos fotodiodos recogen solo la luz roja, otros la luz verde y unos más la luz azul.

### COMPUTADORA

En un sistema de visión artificial la computadora es la encargada de recibir y procesar la información que proviene de la cámara. El término computadora aplica para el típico CPU con su monitor, teclado, ratón, etc., y para computadoras pequeñas especializadas basadas en dispositivos FPGA, DSP y microcontroladores. En la computadora se llevará a cabo la etapa de PDI y de visión artificial.

### PDI

El análisis humano de imágenes u objetos presenta dificultades, pues el proceso de inspección visual consume una gran cantidad de tiempo y está sujeto a interpretaciones incoherentes e incompletas. El PDI, que comprende la manipulación por computadora de imágenes reales, resuelve estos problemas al poder automatizar el proceso de extracción de información útil de los datos contenidos en la imagen, y también permite perfeccionar la imagen, corregir distorsiones, etc. El PDI se efectúa dividiendo la imagen en un arreglo rectangular de elementos; cada elemento de la imagen así dividida se conoce como píxel.

El siguiente paso consiste en asignar un valor a la luminosidad promedio de cada píxel; así, los valores de cada píxel, junto con las coordenadas que indican su posición, definen completamente a la imagen. Todos estos números se almacenan en la memoria de una computadora. A continuación se alteran los valores de luminosidad mediante operaciones o transformaciones matemáticas, para resaltar los detalles de la imagen que sean convenientes. Al final se obtiene la representación de estos píxeles en un monitor, una fotografía impresa, etc.

La utilidad del procesamiento de imágenes es muy amplia y abarca muchos campos. Como aplicaciones típicas tenemos: detección de presencia de objetos, inspección visual automática, obtención de características geométricas y color de objetos, restauración de imágenes y mejoramiento en la calidad de las mismas. También se aplica en imágenes para diagnóstico médico, o en tomas aéreas empleadas para realizar exámenes del terreno, con el fin de analizar los recursos naturales, las fallas geológicas, etc. [7].

### VISIÓN ARTIFICIAL

Cuando el PDI termina de optimizar la información de la imagen, la visión artificial entra en acción. Los objetivos típicos de la visión artificial incluyen:

- Reconocimiento de ciertos objetos en imágenes, por ejemplo, rostros humanos.

- La evaluación de los resultados; por ejemplo, obtener las dimensiones de un objeto proveniente de una máquina CNC y comprobar que cumpla con las especificaciones de diseño.

- Registro de diferentes imágenes de una misma escena u objeto; por ejemplo, hacer concordar un mismo objeto en diversas tomas.

- Seguimiento de un objeto en una secuencia de imágenes. Se suele utilizar para medir posición, velocidad y aceleración.

- Mapeo de una escena para generar un modelo tridimensional; tal modelo puede ser usado por un robot para navegar por la escena.

- Estimación de las posturas tridimensionales de humanos.

Estos objetivos se consiguen por medio de reconocimiento de pa-

trones [8], aprendizaje estadístico [9], geometría de proyección [9], procesado de imágenes [10], teoría de gráficos [11] y otras técnicas.

---

### LOGROS Y AVANCES DE LA VISIÓN ARTIFICIAL EN EL CAMPO DE LA ROBÓTICA

---

Los logros alcanzados por la visión artificial son notables; entre todos ellos destacan por su complejidad los del robot ASIMO, pero existen también desarrollos de autos robot y sistemas de control de calidad, entre muchos otros.

### ASIMO

ASIMO (*Advanced Step in Innovative Mobility*), es un robot humanoide creado por la empresa Honda, capaz de moverse, interactuar con seres humanos y ayudarles; es una de las mayores

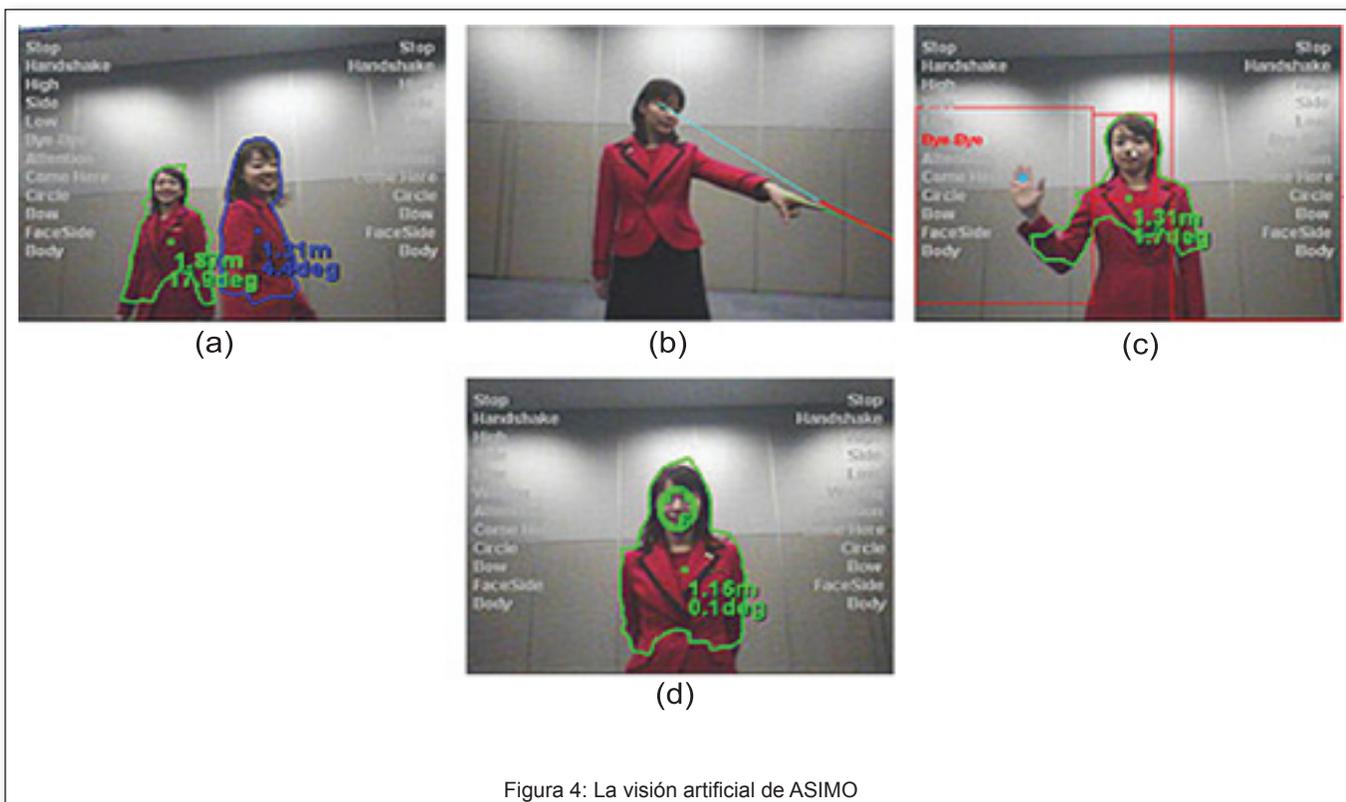


Figura 4: La visión artificial de ASIMO

proezas tecnológicas hasta ahora vistas. Aunque todavía se está lejos de poder atribuir roles concretos a los robots humanoides, estos podrían utilizarse, por ejemplo, para incrementar la autonomía de las personas con capacidades diferentes y personas mayores. Evidentemente, todavía tienen que transcurrir muchos años hasta que se pueda cumplir este objetivo, pero algunas empresas en Japón ya utilizan los servicios de ASIMO para funciones promocionales, como la recepción de visitantes.

La visión artificial que ASIMO incorpora puede determinar la distancia y dirección de movimiento de múltiples objetos o personas (ver la **Figura 4(a)**), es capaz de reconocer gestos como señalar objetos (ver la **Figura 4(b)**) y saludar (ver la **Figura 4(c)**), y además puede reconocer a una persona por sus rasgos faciales (ver la **Figura 4(d)**) [12].

### STANLEY

Stanley es un auto robot capaz de conducirse solo. Se trata de un VW Passat cuya dirección, acelerador y frenos están controlados completamente por una computadora (ver la **Figura 5**). Creado por un equipo de la Universidad

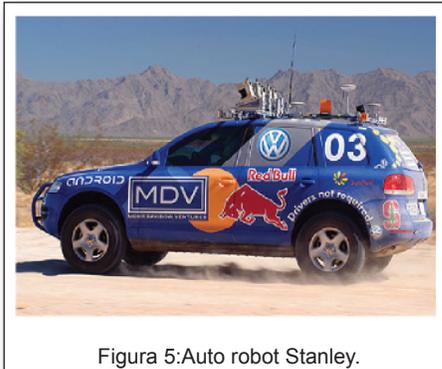


Figura 5: Auto robot Stanley.

sierto de Mojave, EU, guiado sólo por radares, dispositivos láser, y cámaras montadas en la carrocería. La visión artificial de Stanley, le permite evadir los objetos que se encuentran en su camino y estimar la calidad del terreno para determinar la velocidad de avance adecuada [13].

### SISTEMA DE CONTROL DE CALIDAD PARA EL LABORATORIO CIM DE LA UPIITA-IPN

Un laboratorio de manufactura integrada por computadora (CIM), consiste de una celda flexible de manufactura<sup>1</sup>, y el software de integración de administración para simular un ambiente real. El propósito de este laboratorio es complementar el proceso de enseñanza en las aulas a los estu-

diantes de ingeniería en sistemas modernos de manufactura e información gerencial.

Todo proceso de manufactura cuenta con un control de calidad, y en el laboratorio CIM de la UPIITA-IPN, dicho control se lleva a cabo mediante visión artificial. El proceso es bastante sencillo, ya que solo se fabrican dos piezas simples: un cilindro de aluminio y un cubo de acrílico (ver las **Figuras 6(a)** y **6(b)**); dichas piezas forman un ensamble en forma de T invertida. La visión artificial debe determinar si ambas piezas cumplen con las dimensiones que el diseño especifica; si es así, se ensamblan mediante un robot poliarticular, de lo contrario son desechadas (ver **Figura 6(c)**).

Todo el software que lleva a cabo el control de calidad fue desarrollado por alumnos de ingeniería mecatrónica de la UPIITA-IPN, al realizarse un proceso de rediseño, actualización y mantenimiento mayor en el laboratorio CIM [14].

### CONCLUSIONES

La visión artificial es un campo sumamente fértil para el desarrollo de nuevos proyectos y



Figura 6: Visión artificial en el laboratorio CIM de la UPIITA

de Stanford, ganó un premio de dos millones de dólares al lograr atravesar 212 kilómetros del de-

<sup>1</sup> Se considera flexible debido a que puede fabricar prácticamente cualquier elemento mecánico, a través de su maquinaria CNC.

aplicaciones, tanto en la industria, la academia y la vida diaria. Adicionando la robótica, es posible

producir resultados extraordinarios. Prácticamente no existen límites para lo que pueden lograr estas disciplinas unidas, desde artefactos que poseen tecnología de punta, con intenciones de incorporar la robótica a la vida diaria de la humanidad, hasta pequeñas aplicaciones escolares con fines puramente didácticos. Para finalizar, es conveniente mencionar que existe una condición muy alentadora en el desarrollo de estas disciplinas: lo único que puede limitarlas es la falta de ingenio, algo que afortunadamente a los mexicanos nos sobra.

---

### AGRADECIMIENTOS

---

♣ RSO agradece el soporte económico recibido por la Secretaría de Investigación y Posgrado del IPN (SIP-IPN), a través del proyecto 20071024 y del programa EDI, así como del Sistema Nacional de Investigadores (SNI-México).

♦ MAMV agradece el apoyo recibido del programa EDD del IPN.

---

### REFERENCIAS

---

- [1] <http://es.wikipedia.org/wiki/IsaacAsimov>
- [2] Real Academia Española, Diccionario de la Lengua Española, Espasa Calpe, 22va ed, 2002.
- [3] J. Craig, Robótica, Prentice Hall, 3ra ed, 2005.
- [4] [www.infaimon.com/catalogo/industria/index.html](http://www.infaimon.com/catalogo/industria/index.html)
- [5] [www.ni.com](http://www.ni.com)
- [6] <http://es.wikipedia.org/wiki/MosaicodeBayer>
- [7] J. González, *Visión por computador*, Paraninfo, Madrid, España, 2000.
- [8] J. L. Díaz de León Santiago, *Reconocimiento de patrones: avances y perspectivas*, IPN-CIC, México, 2002.
- [9] D. M. Gómez Allende, *Reconocimiento de formas y visión artificial*, RA-MA, 1994.
- [10] G. J. Awcock, *Applied image processing*, McGraw Hill, New York, EU, 1996.
- [11] O. Ore, *Teoría y aplicaciones de los gráficos*, Norma, Cali, Colombia, 1963.
- [12] <http://world.honda.com/ASIMO/technology/intelligence.html>
- [13] <http://cs.stanford.edu/group/roadrunner//old/index.html>
- [14] A. Vázquez, J. Cruz, R. Rojas, *Reacondicionamiento e implementación de un sistema de control de calidad para el módulo de visión artificial del CIM*, IPN-UPIITA, México, 2006.