

Metodología para la implementación de un sistema de reconocimiento de objetos mediante técnicas de visión artificial aplicado a manipuladores robóticos

Sergio Andrés Zabala Vargas

MSc (c) en Ingeniería Área Electrónica,
Universidad Industrial de Santander
Docente Tiempo Completo,
Director Grupo GPS, Universitaria
de Investigación y Desarrollo UDI
Bucaramanga, Colombia
sergiozabala@udi.edu.co

Jhonny Marvin Maldonado

Ingeniero Electrónico,
Universitaria de Investigación y Desarrollo
Investigador Grupo GPS,
Universitaria de Investigación y Desarrollo UDI
Bucaramanga, Colombia
jmar433@hotmail.com

Gianni Yasser Rivera

Ingeniero Electrónico
Universitaria de Investigación y Desarrollo
Investigador Grupo GPS
Universitaria de Investigación y Desarrollo UDI
Bucaramanga, Colombia
giasser@hotmail.com

Resumen— En el presente artículo, se realiza una breve descripción de una metodología implementada para el funcionamiento de un sistema de visión artificial para la identificación de objetos, basado en criterios como la forma, el color y el tamaño. Se realiza la presentación de técnicas y resultados, se destacan las estrategias de visión artificial utilizadas, los principales descriptores para las imágenes y su aplicación a un manipulador robótico. En este orden de ideas, cabe resaltar que este documento ilustra un proyecto de investigación de la Universitaria de Investigación y Desarrollo (UDI), donde se realizaron y aplicaron técnicas de pre-procesamiento y adecuación de la imagen, entrenamiento y uso de redes neuronales, así como la gestión de las señales para el control del manipulador. Es pertinente en este resumen destacar el uso de las Redes Neuronales Artificiales, como eje central del desarrollo del prototipo; lo que permite a través de estas estrategias el reconocimiento de las figuras basado en diferentes criterios y descriptor. Finalmente, pero no menos importante, es necesario resaltar que la implementación del sistema de posicionador no es descrita en este artículo, enfocándose solamente al procesamiento y adecuación de las imágenes.

Palabras clave— Iluminación, manipulador robótico, reconocimiento de formas, redes neuronales artificiales, ubicación espacial, visión artificial.

Abstract— This article is a brief description of a Artificial Vision System for the objects identification based on criteria such as shape, color and size. We carried out the technical and presentation of results, highlighting the strategies used artificial vision, key descriptors for images and its application to a robotic manipulator. In this vein, it is

worth noting that this document illustrates a research project of the Universitaria de Investigación y Desarrollo (UDI), where were implemented pre-processing techniques and adequacy of the image, training and use of neural networks, as well as managing the signals to control the manipulator. It is relevant in this summary to emphasize the use of artificial neural networks, as the core of the development of the prototype, allowing these strategies through the recognition of the figures based on different criteria and descriptors. Last but not least, it is necessary to emphasize that the implementation of the positioner is

not described in this article, focusing only on processing and adequacy of the images.

Keywords— Lighting, robotic manipulator, shape recognition, artificial neural networks, spatial location, machine vision.

I. INTRODUCCIÓN

Cualquier sistema manipulador autónomo diseñado para implementarse en un entorno desconocido, necesita de múltiples sensores (mecánicos, ultrasónicos e infrarrojos) e implementar diferentes estrategias de control para evitar colisiones con obstáculos y detectar con certeza los objetivos de su entorno para desplazarse a un punto específico. Una alternativa utilizada actualmente, son los sistemas de procesamiento digital de imágenes como los presentados en los trabajos [1-3], para la detección y reconocimiento de objetos en el es-

pacio de trabajo. Estos modelos de percepción le permiten al sistema manipulador tener un grado de autonomía mayor y le ayudan en la toma de decisiones. Una de las consideraciones principales al momento de diseñar los sistemas de reconocimiento, es la sensibilidad a los cambios de iluminación, requiriendo que los algoritmos de filtrado, detección, y segmentación sean ajustados según las condiciones del entorno donde se quiera experimentar. Algunos de los sistemas más avanzados se basan en conceptos de inteligencia artificial e implementan técnicas basadas en lógica Fuzzy y redes neuronales [4-5]. En este orden de ideas, el presente artículo propone una metodología de localización de uno o varios objetos, con el fin de controlar la locomoción del manipulador para llegar hasta él. La investigación está basada principalmente en la visión artificial por computador, para manipular las imágenes de un entorno, digitalizadas por medio de una sencilla cámara Web.

En la sección II se realiza la descripción de la metodología utilizada para el reconocimiento de formas y colores de los objetos, la cual detalla los principales parámetros considerados. En dicha sección se presentan también los métodos de aprendizaje a partir de redes neuronales, a las cuales se les asigna como parámetros de entrada un conjunto de valores numéricos que son el resultado de diferentes propiedades de forma y color de las figuras analizadas (círculos, cuadrados, triángulos, rectángulos).

Por otra parte, la aplicación implementada, así como la descripción del software de reconocimiento se proporcionan en la Sección III. Finalmente, se presenta un conjunto de conclusiones y comentarios del trabajo realizado, se busca que el lector comprenda, de manera global, qué consideraciones se deben tener en cuenta para el desarrollo de este tipo de sistemas de reconocimiento.

II. RECONOCIMIENTO DE COLORES Y FORMAS

En esta sección, se realiza una breve descripción de la metodología utilizada para el reconocimiento e identificación de patrones implementado en el presente trabajo. Es importante recalcar que se han seleccionado criterios de color, forma y tamaño; los cuales permiten un reconocimiento básico de un objeto. Cabe destacar, que para efectos de la implementación del programa de reconocimiento, se ha recurrido a la plataforma MATLAB® 7.0,

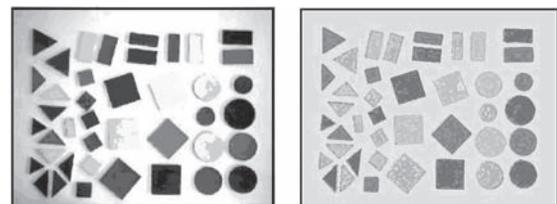
se resalta el uso de las herramientas suministradas para el tratamiento de imágenes y redes neuronales. Sin embargo, el objeto central de este artículo es presentar la metodología implementada, así como los resultados obtenidos en aplicaciones específicas, sin entrar en detalles de la programación del mismo.

A. Reconocimiento de Colores

Una de las mayores dificultades de muchas tareas de visión es adaptar el sistema automáticamente a los cambios de iluminación. En el caso del proyecto, la imagen capturada por la cámara Web presenta un alto contenido de ruidos y sombras, como se muestra en la Fig. 1(a), las cuales afectan los resultados de las etapas de segmentación, binarización y reconocimiento de forma. Para reducir este inconveniente, se realizó una etapa de pre-procesamiento, la cual se fundamenta en una conversión del espacio RGB al espacio HSV y en el posterior uso de umbrales para cada componente. En este proceso, la componente de intensidad (V) se homogeniza con un valor de 0.9 para eliminar el efecto de la sombra.

El ruido se minimiza modificando la componente de saturación, niveles inferiores a 0.2 se igualan a cero y valores mayores se igualan a 0.6. La componente de matiz (H) no se modifica porque contiene la información del color. Un ejemplo de dicho procedimiento se evidencia en la Fig. 1(b).

FIG. 1. PROCESAMIENTO INICIAL DE LA IMAGEN. (A) IMAGEN ORIGINAL (B) IMAGEN FILTRADA



(a) (b) Fuente: Autores del Proyecto

Para seleccionar las figuras presentes en el escenario de trabajo que compartan el mismo color, se entrenó una red neuronal probabilística PNN [4], con el comando *newpnn* de la *toolbox* de MATLAB. Para su entrenamiento se utilizó como dato de entrada un vector constituido por una paleta de 256 niveles de color y la salida es un vector de igual tamaño, con valores de cero y uno, donde un uno representa la posición correspondiente al nivel de color que se quiere reconocer y un cero

lo que se desea descartar (Fig. 2). Se debe ser cauteloso al momento de seleccionar los parámetros de la red neuronal, como es el caso del *spread*, el cual depende de las características del entorno en el que se trabaje como el tipo de iluminación e intensidad de los colores que serán segmentados.

FIG. 2. DATOS DE ENTRENAMIENTO PARA UNA RED PNN

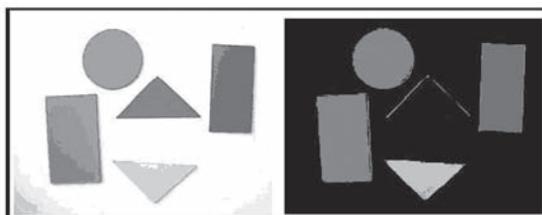


Fuente: Autores del Proyecto

Para observar los efectos causados por la iluminación sobre la segmentación de los colores se tomaron imágenes con luz natural y luz amarilla producida por bombillas de 100 Watts. En la Fig. 3(a)-(b) se observa que la técnica implementada bajo esta condición no responde de forma óptima, ya que no identifica el color verde.

El problema se centra en la etapa de filtrado, debido al alto contenido de luz blanca la figura de color verde toma una tonalidad cercana al nivel de gris, y al realizar las modificaciones en el espacio HSV se elimina, pero se resalta que la red responde de forma adecuada para los demás colores: rojo, azul y amarillo.

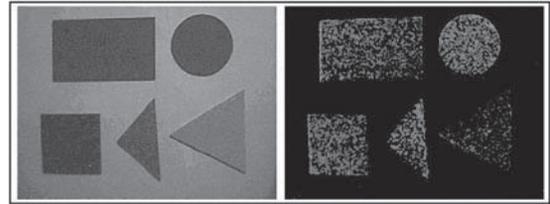
FIG. 3. ILUMINACIÓN CON LUZ NATURAL, A) IMAGEN ORIGINAL B) IMAGEN SEGMENTADA



(a) b) Fuente: Autores del Proyecto

En el segundo caso, al recurrir a la iluminación con luz amarilla, se percibe un cambio en el color de las figuras en la imagen captada por la cámara Web, el verde toma una tonalidad cercana al azul y el amarillo se acerca al color naranja; al pasarla por el filtro este comportamiento se resalta y la red neuronal no va responder de forma satisfactoria (Fig. 4).

FIG. 4. ILUMINACIÓN CON LUZ MARILLA, A) IMAGEN ORIGINAL, B) IMAGEN SEGMENTADA



a) b) Fuente: Autores del Proyecto

B. Reconocimiento de formas

La función de los algoritmos de reconocimiento es detectar e identificar cada objeto segmentado de una escena y asignarle una etiqueta. Una limitante en la mayoría de los sistemas de visión, es que los niveles de reconocimiento trabajan objetos que se supone han sido segmentados como unidades individuales. Otra limitación común es que las imágenes se deben tomar con una geometría de visión conocida, a menudo perpendicular al espacio. Esto hace que disminuya la diversidad de características de forma, lo que simplifica la segmentación y la descripción, y reduce a su vez las posibilidades de ocultación de detalles en las imágenes [6].

El problema de la descripción se centra en extraer características de un objeto para reconocerlo. Por lo general, los descriptores deben ser independientes del tamaño, la localización y orientación del objeto, además deben contener suficiente información de discriminación para distinguir un objeto de otro [7]. En esta etapa lo que se hace básicamente, es reducir la tarea de descripción de una imagen en dos dimensiones a la descripción de una función unidimensional por medio de la signatura de las figuras, que es una representación o modelo funcional unidimensional de un borde a la cual se le extraen propiedades características como son los coeficientes de Fourier aplicando el algoritmo de la Transformada rápida de Fourier (FFT), parámetros topológicos como la redondez, compacidad y excentricidad y otros datos estadísticos como el valor medio, la media aritmética, media geométrica, media armónica, desviación estándar y la mediana. En la Fig. 5 se muestra un ejemplo de la extracción de descriptores para un cuadrado.

FIG. 5. DESCRIPTORES DE FORMA PARA UN CUADRADO



Descriptor	Cuadrado
Coef. Fourier 1	0,012852609
Coef. Fourier 2	0,017550328
Coef. Fourier 3	0,021406995
Coef. Fourier 4	0,063029502
Coef. Fourier 5	0,014836957
Coef. Fourier 6	0,002034224
Coef. Fourier 7	0,00607452
Excentricidad	0,230126774
Redondez	0,73792832
Compacidad	0,859027543
Med. Aritmética	0,835982655
Med. Geométrica	0,832062168
Med. Armónica	0,828226026
Desv. Estándar	0,826130863
Mediana	0,082521374

Fuente: Autores del Proyecto

Los datos extraídos de la signatura son introducidos a una red back-propagation [6-7] previamente entrenada con un número de 41 figuras planas en las que se encuentran 12 cuadrados, 8 círculos, 12 triángulos y 9 rectángulos, en diferentes tamaños y orientación. El vector de salida se construye asignándole a cada figura un número que la identifique, donde un uno representa un cuadrado, un dos un rectángulo, el tres es un círculo y el cuatro un triángulo. Con el fin de validar la efectividad de la metodología utilizada para el reconocimiento de formas, se analizó el comportamiento de la red, con una imagen en la cual se encuentran figuras superpuestas de diferente color y forma (Fig. 6 y Fig. 7). Para las diferentes figuras se realizaron decenas de pruebas, donde se observa que el sistema responde de forma adecuada. Los resultados erróneos se presentan en figuras con dimensiones inferiores a 5cm en sus lados, debido a la resolución que tiene la cámara, que elimina el detalle de los bordes y esquinas en cada figura, aproximando, por ejemplo, un cuadrado a un círculo. En el caso que se unan dos figuras del mismo color, la red la aproxima a una de las cuatro figuras que se utilizaron para el entrenamiento, ya que las toma como un solo objeto.

FIG. 6. RECONOCIMIENTO DE FORMAS EN FIGURAS UNIDAS DE DIFERENTE COLOR.



Fuente: Autores del Proyecto

FIG. 7. RECONOCIMIENTO DE FORMAS EN FIGURAS UNIDAS



Fuente: Autores del Proyecto

III. DESCRIPCIÓN DEL PROTOTIPO FINAL

Para los propósitos de la aplicación de las pruebas y desarrollos descritos en un ambiente real, se ha planteado y trabajado sobre las características motrices de un manipulador tipo pórtico con cuatro grados de libertad, tres de desplazamiento y uno de orientación.

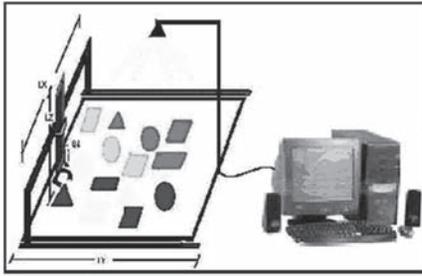
Al trabajar con este tipo de diseños fue necesario crear un algoritmo de control motriz flexible, el cual permitiera ajustar las características físicas (longitud y área) del manipulador, ya que la dimensión del área de trabajo puede variar según la ubicación de la cámara Web. La interfaz gráfica, como ya se comentó, fue desarrollada en el entorno de MATLAB 7.0, en ella el algoritmo reconoce y ubica automáticamente la posición y orientación y área de una figura que selecciona el usuario. En el aparte A. de la presente sección se describirá con mayor detalle este aplicativo.

En este orden de ideas, se plantea entonces un escenario iluminado con luz blanca fluorescente, uniformemente distribuida, de tal forma que se presente la menor cantidad de ruido y sombras posibles en la imagen.

El entorno, en el que se trabajó abarca un área de 1.10m de ancho por 0.80m de largo, pintada de color blanco, en el se encuentran distribuidas de forma aleatoria, figuras planas de diferentes tamaños con forma circular, cuadrada, triangular o rectangular, de color rojo, verde, azul o amarillo.

El escenario lo complementa una cámara WEB ubicada a una altura de 1.40m, que captura imágenes a una velocidad de 15 cuadros por segundo, con una resolución de 640x480 píxeles (Fig. 8).

FIG. 8. ESQUEMA DEL ESCENARIO DE TRABAJO.



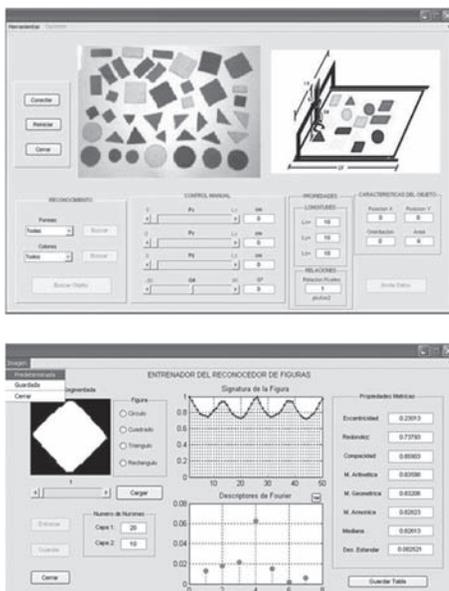
Fuente: Autores del Proyecto

C. Descripción aplicativo software de reconocimiento y control de posición

Luego de realizar los análisis y pruebas descritas sobre el espacio trabajado, se desarrolló el aplicativo software, el cual cuenta con una interfaz de usuario constituida por una pantalla principal, presentada en Fig. 9(a) y con el subprograma para el entrenamiento de la red neuronal con los objetos utilizados y registrados por la imagen. Fig. 9(b).

En síntesis, la interfaz principal permite visualizar el número de figuras, ya sea por forma, color, tamaño; a su vez, permite combinar criterios de elección.

FIG. 9. APLICATIVO SOFTWARE. (A) VENTANA PRINCIPAL PARA RECONOCIMIENTO Y CONTROL DEL MANIPULADOR. (B) VENTANA DEL ENTRENADOR DE LA RED NEURONAL



(a) (b) Fuente: Autores del Proyecto

También cuenta con un control manual del elemento posicionador, el cual envía las señales a un puerto serie convencional, encargado de la gestión del movimiento del manipulador descrito con anterioridad. Finalmente, se cuenta con una herramienta denominada Entrenador para el reconocimiento de figuras, la cual permite realizar la adecuación de la red neuronal, basado en el análisis de las figuras existentes fundamentándose en la forma, en las propiedades métricas descritas en la Fig. 5 y en el número de neuronas por capa.

IV. CONCLUSIONES

En el presente trabajo se ha socializado la metodología para la implementación de una herramienta práctica para el reconocimiento de figuras, basado en descriptores de forma, tamaño y color; aplicado básicamente a posicionadores y controladores robóticos. En general, se pueden destacar los siguientes elementos como conclusiones del presente trabajo:

- Es evidente que las diferentes clases de iluminación en el escenario de trabajo afectan directamente a los algoritmos de reconocimiento de color y forma, debido a que la etapa previa de acondicionamiento de la imagen no genera los resultados esperados. Estos problemas se muestran, en mayor medida, en las imágenes con demasiada iluminación, donde no es conveniente homogenizar la componente de intensidad, pues ocasiona saturación y eliminación de figuras.
- Uno de los mayores inconvenientes en las metodologías de reconocimiento de formas es su dependencia directa con el algoritmo de segmentación de colores, porque al presentarse problemas en las condiciones de iluminación no se identificarían de forma clara los bordes de las figuras, generándose datos erróneos en los descriptores de forma utilizados. Una estrategia interesante para este inconveniente es la implementación de algoritmos de segmentación y reconocimiento totalmente independientes.
- Se deben tener en cuenta todas las características mecánicas del manipulador a la hora de desarrollar los algoritmos de control y locomoción porque la imprecisión de

- su diseño afecta de forma considerable sus movimientos.
- Una de las principales dificultades de operar en escenarios que suponen un entorno con cámaras en posición fija, es que el espacio de trabajo se ve restringido al rango de visión de la cámara implementada, que reduce el área de exploración. Se plantea, como trabajo futuro ubicar una cámara móvil sobre el manipulador, para tener una visión periférica del ambiente, además se propone la realización de comunicación vía inalámbrica.
 - En el tratamiento de las imágenes, se convierte en una herramienta fundamental el uso de filtros adaptativos para el mejoramiento de la iluminación, se busca siempre mejorar y aumentar la capacidad de operación del manipulador en ambientes no controlados.

REFERENCIAS

- [1] Y. D. Amaya. Universidad Industrial de Santander, Bucaramanga. Facultad de Ingenierías Físico-Mecánicas. Escuela de ingenierías eléctrica, electrónica, y telecomunicaciones. Localización dinámica de móviles y obstáculos en una escena controlada para aplicaciones en robótica. Marzo 2005.
- [2] J.F.Flórez. Universidad del Cauca, Colombia. Herramienta para la experimentación con robots móviles aplicando control inteligente y visión por ordenador.
- [3] M.D. Ramos, Marcos Daniel. Escuela Politécnica Universidad de Alcalá. Micro-robot futbolista GOLIATTOR.
- [4] S. Haykin, Neural networks a comprehensive foundation. 2ed. Ed. Prentice-Hall. 842 p. 1999. ISBN 0139083855.
- [5] M.A. Zamora, L.M. Grupo de Visión, Robótica y Proy. Universidad de Murcia. H.Martínez, A.G.Skarmeta. Dpto. Inf., Inteligencia artificial y electrónica Facultad de informática Universidad de Murcia. Navegación planificada de un robot móvil en entornos interiores desconocidos.
- [6] G. Pajares. Redes neuronales artificiales: fundamentos modelos y aplicaciones. Ed. Addison-Wesley Iberoamericana 390p. 1995. ISBN 0-201-87895-X.
- [7] G. Pajares. Visión por computador: Imágenes digitales y aplicaciones. Ed. Alfaomega. 764p. 2002. ISBN 970-15-0804-1.