

Reconocimiento de patrones numéricos para vuelo controlado de un AR drone utilizando redes neuronales artificiales

Juan Carlos Rodríguez-Sánchez, Víctor Manuel Landassuri-Moreno, José Martín Flores Albino

Universidad Autónoma del Estado de México, Centro Universitario UAEM Valle de México
México

{jcrsylatp, jose_martin_70}@yahoo.com, vmlandassurim@uaemex.mx,

Resumen. En el presente trabajo se utilizó la segmentación de letras de fotografías tomadas mediante la cámara del Parrot AR Drone, con el objetivo de establecer una interacción de estímulo-respuesta, donde la imagen original en formato de combinación de colores Rojo, Verde y Azul (RGB) se segmentó por color (escogiendo el canal rojo). Una vez que se reconoce el carácter, el Drone ejecuta la acción correspondiente. Se utilizaron inicialmente patrones numéricos libres de ruido y posteriormente se agregaron algunos pixeles a la imagen con el objetivo de hacer más robusto este conjunto de patrones, los cuales proporcionaron el conjunto de entrenamiento para la red neuronal y de esta forma se pudieron interpolar patrones nuevos. Para la segmentación de imágenes se utilizaron técnicas de detección de bordes que incluyen el filtro de Sobel así como filtros para eliminación de ruido basados en filtrado de la mediana que es un filtro pasa baja. Todo lo anterior se llevó a cabo en un entorno cerrado y se espera ampliar este trabajo para su aplicación en diferentes entornos.

Palabras clave: Segmentación de caracteres, vehículos aéreos no tripulados (UAVs), procesamiento de imágenes, filtro Sobel, detección de bordes, redes neuronales artificiales.

Recognition of Numeric Patterns for a Flight Controlled by an AR Drone Using Artificial Neural Networks

Abstract. In this paper the letter segmentation of photographs was used, taken from a Parrot AR Drone's camera with the aim of establishing a stimulus-response, where the original picture formed by Red, Green and Blue (RGB) colors was segmented by color (choosing the red channel). Once the character is recognized, the Drone executes the corresponding action. Noise-free number patterns were initially used and then some pixels were added in the image in order to make a set of patterns more robust, which provided the training set for neural network and thus are able to interpolate new patterns. Edge techniques detection were used for image segmentation including Sobel filter and filters for noise removal based on the median filtering, that is a low pass filter. All this took place in a closed environment, expecting to extend this to different environments.

Keywords: Character segmentation, unmanned aerial vehicles, image processing, Sobel filter, border detection, artificial neural networks.

1. Introducción

El ojo de los mamíferos es un órgano complejo capaz de procesar señales luminosas que el cerebro interpreta para llevar a cabo diversas tareas donde, en el caso particular del ojo humano, éste es capaz de detectar figuras a diferentes distancias, con diferentes rotaciones y niveles de ruido. De esta forma en un lugar donde se proyecta sombra, aún es posible reconocer formas e incluso colores. Por otro lado, hasta hace poco el reconocimiento e interpretación de caracteres escritos tales como letras y números resultaba ser una tarea de gran complejidad para una computadora, ya que ésta solo procesa señales digitales. Por otra parte los colores corresponden a diferentes frecuencias del espectro electromagnético, las cuales son señales continuas en el tiempo, i.e. son analógicas y por consiguiente requieren un procesamiento previo antes de ser manipuladas por medio de procesamiento digital. A pesar de ello, el desarrollo de diversas herramientas de software ha facilitado considerablemente las tareas de procesamiento de imagen, donde estos programas han visto mejorado su rendimiento mediante la incorporación de las Redes Neuronales Artificiales (RNAs).

En lo referente al control de los parámetros de vuelo de los UAVs explicados a detalle en [1, 2, 3], se requiere de un control adecuado y las RNAs en combinación con las técnicas de procesamiento digital de imagen proporcionan una herramienta eficiente para este propósito. Una amplia revisión de aplicaciones de RNAs para procesamiento digital de imágenes se menciona en [4] donde además se muestran las etapas de nivel de procesamiento que son: pre-procesamiento, Segmentación, Detección, Clasificación y Análisis. Otras técnicas involucran el uso de estadísticas para el análisis minimizando el riesgo de hacer una clasificación incorrecta de los patrones [5] o mediante técnicas matemáticas de rechazo de umbrales con aprendizaje no supervisado [6]. Lo anterior demuestra que las redes neuronales resultan ser muy útiles al momento de trabajar con imágenes.

Por otra parte, el uso cada vez más frecuente de robots de servicio aplicados en diferentes áreas ha motivado el interés por el mejoramiento de la interacción Hombre-Máquina y por ello se han llevado a cabo algunos experimentos de reconocimiento de patrones utilizando robots terrestres en entornos cerrados, utilizando odometría y codificadores ópticos como se señala en [7] y otros analizando la interacción con las máquinas en entornos cambiantes [8]. Así, la adquisición de imágenes mediante cámaras conectadas a robots, busca establecer una relación de estímulo-respuesta procesando comandos para ejecutar acciones concretas en entornos reales, apoyándose a su vez en simulaciones 3D virtuales, como lo señalado por [9]. En el caso particular de los vehículos aéreos no tripulados (UAVs), algunas técnicas involucran la elaboración de mapas vectorizados basados en imágenes tomadas vía satélite mediante la detección de contornos y objetos [10], mientras que otros analizan los resultados de la interacción del AR Drone mediante el diseño de una interfaz para el reconocimiento de órdenes gestuales [11]. Un trabajo similar se muestra en [12], donde ya se utiliza el AR Drone para aterrizaje y despegue autónomo, implementando

técnicas de cotejamiento de patrones buscando pequeñas secciones dentro de una imagen o mediante técnicas de detección de aspectos y cotejamiento que basa su funcionamiento en la descripción de puntos mediante vectores utilizando sensores de ultrasonido instalados en el Parrot AR Drone. Así, éste artículo presenta la “segmentación de caracteres de fotografías tomadas mediante la cámara del Parrot AR Drone”, los algoritmos programados se basaron en el uso de operadores morfológicos, todos éstos enfocados al pre-procesamiento de la imagen. Se generaron patrones a partir de los caracteres segmentados con distintos niveles de ruido, mediante el uso de RNAs con un algoritmo evolutivo para adaptar sus arquitecturas durante evolución.

Este trabajo está organizado en seis secciones. La primera sección corresponde a los antecedentes. La segunda sección corresponde a la descripción de los métodos propuestos, posteriormente en la cuarta sección se muestra la configuración experimental, los resultados obtenidos se muestran en la sección cinco. Finalmente en la sección seis se describen las conclusiones y trabajos futuros.

2. Antecedentes

En esta sección se describirán las técnicas utilizadas para aislar el carácter del resto de la imagen. En primer lugar se tienen los llamados métodos de filtrado que permiten resaltar o suprimir de forma selectiva con la finalidad de destacar algunos elementos que forman parte de una imagen. Algunos filtros de imagen son los llamados filtros pasa baja que incluyen el filtro de la media, el filtro de la media ponderada, el filtro de la mediana, filtros adaptativos y filtros Gaussianos. También se tienen los llamados filtros pasa alta que incluyen la sustracción de la media, los filtros basados en la derivada y los filtros Laplacianos. El método utilizado en este trabajo es el de filtrado de la mediana el cual aplica una matriz de 3x3 a cada uno de los pixeles en una imagen en escala de grises, la cual funciona a partir de dos máscaras, una horizontal en función de “x” y otra vertical en función de “y” (figura 1).

1	1	1
1	1	1
1	1	1

 $DIV = 9$

Fig. 1. Filtro de la media.

Existen otros métodos que consideran que los bordes de la imagen no podrán procesarse debido a que se excederían los bordes de la imagen perdiendo N-2 Filas X N-2 Columnas.

3. Métodos propuestos

El software National Instruments - LabVIEW proporciona un método adecuado de trabajo para el desarrollo de instrumentación virtual que posee las capacidades de manipular hardware desde una tableta, computadora o teléfono celular. Debido a su flexibilidad en lo que se refiere al procesamiento de imagen IMAQ, es una herramienta muy utilizada en el ámbito industrial, tal y como se muestra en [13]. En una primera etapa se diseñaron interfaces para control de vuelo utilizando el LVH AR Drone Toolkit, posteriormente se enlazaron las cámaras del AR Drone con la cámara de una tableta. Esto permitió tomar fotografías de los caracteres impresos con la cámara del Drone, para lo cual se utilizaron herramientas de LabVIEW. Este procesamiento incluye herramientas basadas en el toolbox IMAQ dentro de la opción Vision and Motion/Vision Utilities. Las herramientas de este toolbox incluyen al IMAQ Create para reservar un espacio en memoria para la imagen. Dentro de la opción Vision and Motion se selecciona NI-IMAQdx y una vez ahí las herramientas Open, Configure Grab y Close se encargan de abrir la imagen, mantenerla en memoria y cerrar la sesión. Además la herramienta “Vision Acquisition” permite registrar los bits de inicio en la entrada. Posteriormente se almacenaron las imágenes adquiridas en una carpeta convirtiendo los valores RGB de la imagen de entrada a valores numéricos enteros. La figura 2 muestra un código para capturar la imagen con el AR Drone.

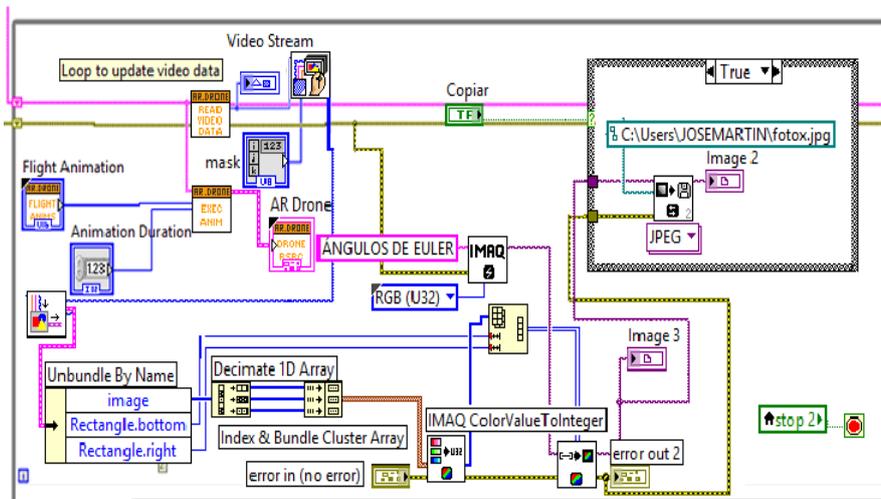


Fig. 2. Código para captura de Imagen.

En la siguiente etapa del desarrollo se utilizó Matlab para llevar a cabo la segmentación de los caracteres y su posterior reconocimiento y procesamiento. Con el propósito de eliminar la mayor cantidad de ruido posible de las fotografías tomadas con la cámara del AR Drone, los números se imprimieron en color Rojo por lo que en el algoritmo utilizado se separa la imagen RGB en sus correspondientes canales Rojo, Verde y Azul para eliminar el fondo y toda la información innecesaria. Esto se muestra en la figura 3.

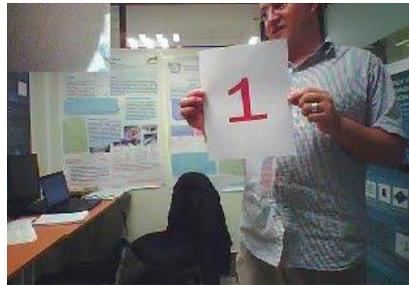


Fig. 3. Imagen Original con un comando numérico.

Aún es necesario quitar ciertas regiones que están formadas por diferentes grupos de píxeles por lo que fue necesario aplicar el filtro de la mediana y el filtro de imagen binaria. De esta forma se obtiene una imagen con una menor cantidad de píxeles, reduciendo considerablemente el ruido, tal y como se puede apreciar en la Figura 4.



Fig. 4. Imagen con Ruido (izquierda) e imagen con filtro de la mediana y filtro de imagen binaria (derecha).

Algunos grupos de píxeles se pueden apreciar en la imagen por lo que se utilizaron funciones que permiten la eliminación de grupos de píxeles menores a un umbral cuyo valor en este caso se fijó a 100 (Figura 5).

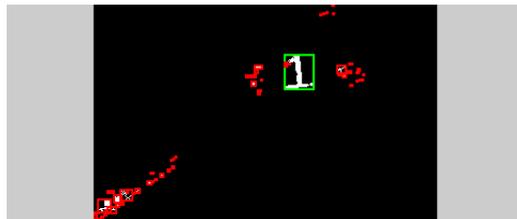


Fig. 5. Eliminación de los grupos de píxeles con un valor menor al umbral.

En la figura 6 se muestra el resultado final de la segmentación.



Fig. 6. Resultado final de la segmentación.

Para generar los patrones que servirán como datos de entrada para la red neuronal se rotaron las imágenes a 45° cada una con respecto de la anterior hasta completar los 360°. En la siguiente imagen se muestra la secuencia de rotación para el número 1.



Fig. 7. Secuencia de rotaciones.

Considerando todos los números del 1 al 8 se obtuvo este primer conjunto de patrones que consta de un total de 64 patrones. Una vez que se tienen estos datos se separa la imagen en sus columnas y se concatenan de forma horizontal con la etiqueta que servirá para tener un entrenamiento supervisado con las redes neuronales. Esta etiqueta es un vector de bits el cual sirve como identificador para cada patrón. En la figura 8 se puede ver la imagen antes y después de agregar el identificador, cabe destacar que la imagen final binarizada fue concatenada con todas sus filas, así como con la etiqueta, dando un tren de bits con el identificador concatenado al final, como se muestra en la figura 8.

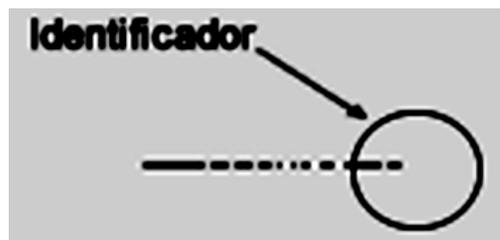


Fig. 8. Se añade el identificador a la secuencia de bits.

Un conjunto de patrones ya concatenados se muestra en la figura 9, donde cada línea representa un patrón de entrada para la red neuronal. Ahí es posible apreciar que las últimas columnas de la Figura 9 son iguales, ya que todas corresponden al dígito 1.



Fig. 9. Bits concatenados para el patrón de 1.

Con el objetivo de lograr un conjunto de entradas para la red neuronal más robusta, se generó ruido aleatorio en cada uno de estos patrones. En la figura 10 se muestra uno de los patrones girado y con ruido. La figura 11 muestra en algoritmo de flujo del programa implementado en este trabajo.

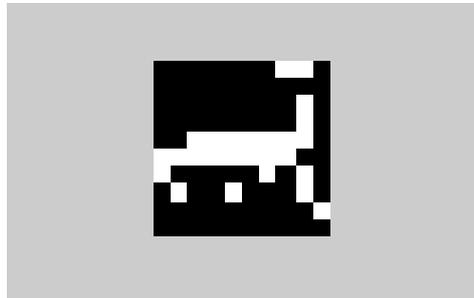


Fig. 10. Ejemplo de Imagen con ruido.

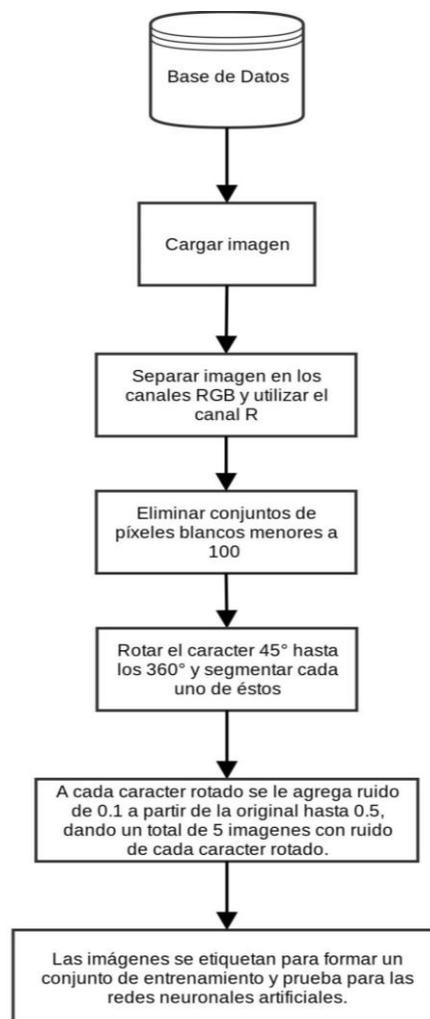


Fig. 11. Diagrama de Flujo del algoritmo.

4. Configuración experimental

Cada patrón (dígito) fue girado 8 veces en ángulos incrementales de 45°, dando 64 patrones de cada uno de los dígitos usados (1-8), posteriormente se insertó ruido con una probabilidad de 0.01 a 0.05, a bloques de imágenes del mismo tamaño, lo que resultó en un conjunto de 384 patrones. Fue utilizado un algoritmo evolutivo de redes neuronales artificiales para hacer evolucionar la arquitectura de ellas. Se tuvieron 10 individuos por población, evolucionados por 100 generaciones como máximo, donde la evolución puede ser terminada si se llega a un valor de clasificación de cero. Los individuos obtuvieron 144 neuronas de entradas y 8 de salida, dado que las imágenes fueron redimensionadas a matrices binarias de 12x12 bits, con un código binario de 8 bits para cada patrón. Los experimentos fueron repetidos 30 veces por cuestiones estadísticas. El conjunto de prueba de las RNAs consistió en el último bloque de 64 patrones, los cuales presentan la mayor cantidad de ruido (0.05), y el resto fueron usados para diseñar y entrenar las redes neuronales.

5. Resultados

Los resultados obtenidos del experimento realizado muestra que en promedio, las redes neuronales artificiales tardaron 6 generaciones en poder clasificar correctamente todos los patrones del conjunto de prueba. El método usado para determinar el nivel de clasificación fue “Winner takes all”. En las figuras 12–15 se muestra el error, conexiones, mutaciones y nodos ocultos en promedio de las redes neuronales evolucionadas durante 200 generaciones de evolución. Como se puede ver en la figura 12, el error fue disminuyendo conforme avanzaba la evolución. Donde cabe resaltar que el error usado en el NRMSE, Raíz cuadrada del error medio cuadrático normalizado, encargado de determinar el error de adaptabilidad de las redes y como se comentó anteriormente, el error de clasificación fue de cero a partir de la sexta generación.

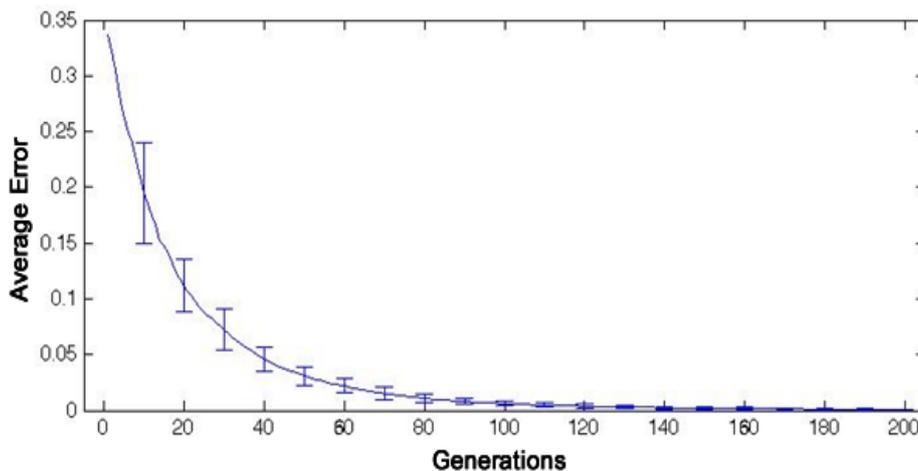


Fig. 12. Error NRMSE promedio durante evolución.

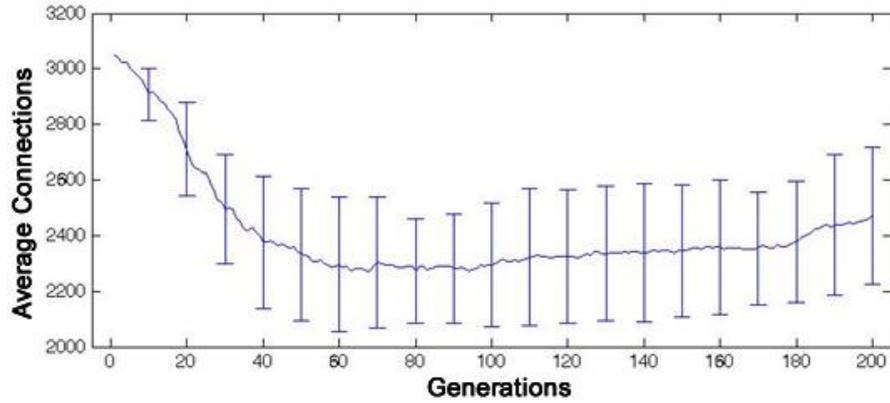


Fig. 13. Número de conexiones en promedio durante evolución.

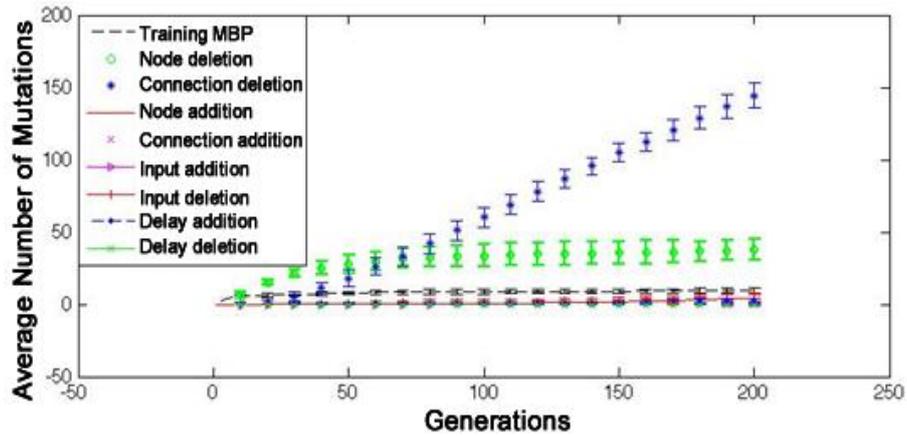


Fig. 14. Número de mutaciones aplicadas a la población en las 200 generaciones de evolución.

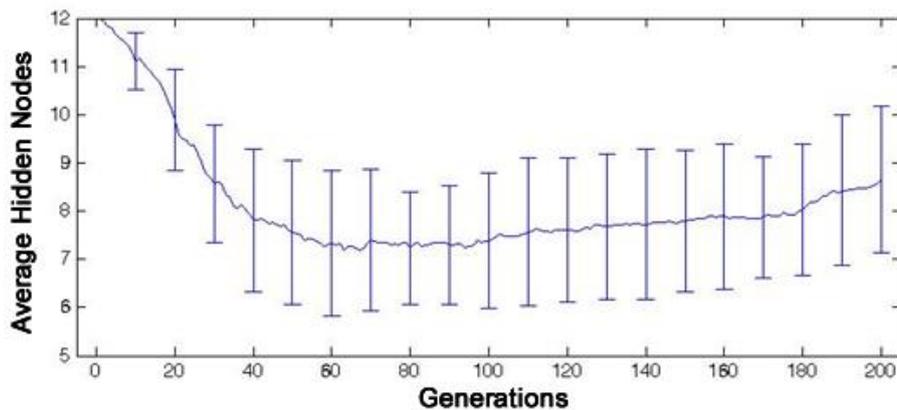


Fig. 15. Número de nodos ocultos en promedio por población por generación.

6. Conclusiones y trabajo a futuro

En este trabajo se presenta una configuración experimental, la cual con ayuda de redes neuronales artificiales fue posible reconocer 8 dígitos los cuales son usados como comandos gráficos para un Drone, donde este puede despegar y desplazarse. Las redes neuronales no mostraron ninguna dificultad para clasificar correctamente los datos, dada la simplicidad de ellos, en comparación con el reconocimiento de caracteres escritos, la cual es una tarea más complicada de resolver. Como trabajo a futuro, se tiene contemplada la aplicación de estos algoritmos para desarrollar el vuelo completamente autónomo del dispositivo y su posterior adaptación y aplicación en diversos entornos. Estos entornos pueden ser ambientes con mucho ruido ya sea en interiores o en exteriores. También se espera presentar más resultados en futuras publicaciones.

Agradecimientos. Se agradece al laboratorio de Cómputo de Alto rendimiento del Centro Universitario UAEM Valle de México por las facilidades prestadas para el uso del Clúster de computadoras para llevar a cabo estos experimentos.

Referencias

1. García Carrillo, L.R.: Modeling the Quad-Rotor Mini-Rotorcraft. *Advances in Industrial Control*, Springer-Verlag London, Vol. 46. pp. 23–35 (2013)
2. Raffo Guilherme, V.: Modelado y Control de un helicóptero Quadrotor. *Universidad de Sevilla Technology*, pp. 35–53 (2007)
3. Peña Giraldo, M.V.: Modelamiento Dinámico y Control LQR de un Quadrotor, *Revista Avances, Investigación en Ingeniería*. Universidad Nacional de Colombia, No. 13, pp. 71–86 (2010)
4. Ramírez Q., J.A., Chacón M., M.I.: Redes Neuronales para el procesamiento de imágenes, una revisión de la última década, *Revista de Ingeniería Eléctrica y Computación*, Vol. 9, No. 1, pp. 7–16 (2011)
5. Bishop, C.M.: *Neural Networks for Pattern Recognition*. Oxford University Press (2005)
6. Coates, A.: Text Detection and Character Recognition in Scene Images with Unsupervised Feature Language, *IEEE Computer Society Washington, DC, USA*, pp. 440–445 (2011)
7. Sánchez Bautista: Diseño y construcción de un Robot Móvil Autónomo contra Incendios. *Universidad Autónoma de Puebla* (2003)
8. Djkshoorn, N.: Simultaneous localization and Mapping with the AR. Drone. Master Thesis for the graduation in Artificial Intelligence, *Universiteit Van Amsterdam* (2012)
9. Castejón Nieto A.: Reconocimiento Visual de Instrucciones de un entorno 3D para un Robot Autónomo. *Universidad Complutense de Madrid Facultad de Informática* (2005)
10. Bielecki, A.: Recognition of two-dimensional representation of urban environment for autonomous flying agents. *Expert Systems with applications*, Elsevier, pp. 3623–3633 (2012)

11. Soto Guerrero, D.: Interacción Hombre-Robot con vehículos aéreos no tripulados basada en visión. Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Laboratorio de Tecnologías de la Información. Tesis de Maestría en Ciencias en Computación (2012)
12. Barták, R.: On Autonomous Landing of AR Drone: Hands-on Experience. Charles University in Prague, Faculty of Mathematics and Physics, No. 25 pp. 400–405 (2011)
13. Gordillo Erazo: Sistema de Reconocimiento de Patrones Visuales Basado en Técnicas de Procesamiento de Imágenes y Redes Neuronales, Telematique, Revista Electrónica de Estudios Telemáticos, Vol. 3 (2004)