

# Entorno para Visión Artificial Educativo

**Coral Calero Muñoz, Vicente Feliu Batlle**

**E.T.S. Ingenieros Industriales  
Universidad de Castilla-La Mancha  
Campus Universitario s/n  
13071 Ciudad Real  
E-mail: vfeliu@ind-cr.uclm.es**

*Se estima que aproximadamente las tres cuartas partes de la información que maneja un ser humano es visual; así, parece natural pensar que el objetivo de dotar a las máquinas del sentido de la vista supondrá un salto cualitativo en sus capacidades de actuación. La visión artificial puede ser definida como los procesos de obtención, caracterización e interpretación de información de imágenes tomadas de un mundo tridimensional. Esta comunicación presenta un proyecto cuyo objetivo es conseguir un entorno de trabajo para el desarrollo de unas prácticas de Visión Artificial para una asignatura de segundo ciclo de Ingeniería Superior Industrial. Por su carácter didáctico, se pretende conseguir que el usuario de este entorno realice unos tratamientos, siguiendo unas pautas que le conduzcan a unos resultados que pueda analizar. Es por esto que se han tenido muy en cuenta varios aspectos, fundamentalmente la planificación de estas prácticas que abarcan todos los elementos que forman un sistema de visión artificial, incluida la captura de imágenes desde una cámara de vídeo. Además, estos tratamientos deben tener una secuencialidad, de forma que haya un adentramiento progresivo en el mundo de la visión artificial, realizando tratamientos cada vez más complicados hasta llegar a una última práctica donde se realice el tratamiento objetivo de todo sistema de visión artificial que es el reconocimiento y clasificación de objetos de una escena.*

## 1. Planteamiento general

Cuando llegó el momento de planificar un programa de prácticas fue necesario tener en cuenta varios aspectos. De un lado había que conseguir que el alumno fuese progresando de forma paulatina en el mundo del procesamiento de imágenes de forma clara y sencilla. De otro, había que conseguir cubrir muchos

campos del tratamiento de imágenes lo cual podía llevar a un proceso demasiado largo y tedioso. Así pues se nos planteaba el problema de conseguir aglutinar todas las posibilidades en un número tratable sin que por ello las prácticas quedaran o demasiado complicadas o demasiado resumidas. Se optó entonces por seguir un proceso lógico respetando las fases en que se divide un sistema de visión artificial (fig.1) y que se muestran a continuación:

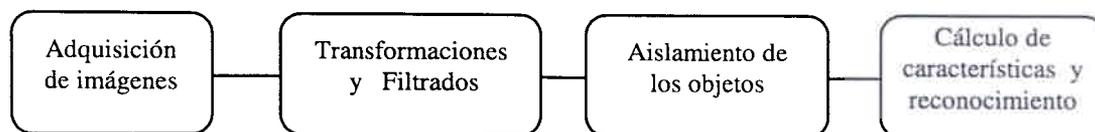


Figura 1. Etapas de un sistema de visión artificial

Por tanto parecía que una buena opción podría ser la siguiente:

- Dar posibilidad al alumno para capturar imágenes.
- Hacer una primera parte que abarcara todo el tema del preprocesado en el que el discente se fuera familiarizando con las características principales y definitorias de una escena digital.

- Realizar una segunda parte que tratara el procesamiento avanzado de imágenes en la que el alumno tratara con los distintos elementos de una imagen y trabajara con ellos para conseguir el objetivo de todo sistema de visión artificial, que no es otro que el reconocimiento de objetos dentro de una escena.

Llegado a este punto ya se tenía el esquema general que debía tener el programa. Ahora se trataba de desglosar cada una de las partes de forma que el proceso fuera desarrollándose de una manera paulatina. Para el desglose de la primera parte se quiso distinguir entre características de la imagen y tratamiento básico de las mismas; es decir, por un lado y como primer paso, el alumno trabajaría con las características de una imagen como resolución espacial, niveles de gris, componentes RGB o planos de bits, de forma que llegara a comprender qué es una imagen digital. A continuación, como segundo paso, se le haría trabajar con tratamientos básicos como la transformada de Fourier y algunos tipos de filtros, lo cual le ayudaría a entender cómo pueden manejarse las imágenes y qué resultado se obtiene ante los distintos tratamientos. Para el desglose de la segunda parte, la separación era aún más intuitiva puesto que para el reconocimiento de formas hay que hacer un paso previo que es el de extracción de características y el de detección del número de piezas, para lo cual es necesario tratar la imagen con el fin de obtener los bordes de los elementos de esta. Así pues, se decidió dividir esta parte en otras dos: la primera trabajaría con convolución, con realce, con los histogramas y con el contraste, técnicas orientadas todas ellas a la detección de objetos.

La segunda parte se ocuparía ya de la determinación del número de piezas, de la obtención de las características discriminantes y, por fin, de la clasificación.

Por tanto, el programa de prácticas que se propuso como solución fue el siguiente:

### **Parte 1ª.**

#### **Práctica 1:**

- Resolución espacial.
- Niveles de gris.
- Planos de bits.
- Componentes RGB.

#### **Práctica 2:**

- Transformada de Fourier.
- Filtrado en frecuencia.
- Filtrado lineal.
- Filtro de la mediana.

### **Parte 2ª.**

#### **Práctica 3:**

- Equivalencia entre convolución y Transformada de Fourier.
- Realce de Bordes.
- Modificación del Histograma.
- Modificación del contraste.
- Detección de bordes.

#### **Práctica 4:**

- Detección del número de piezas.
- Extracción de características.
- Clasificación.

Documentación necesaria: Resultaría conveniente poner a disposición de los equipos de prácticas un dossier resumen con los principales comandos MATLAB para procesamiento de imágenes de forma que pudieran consultarlo para la realización de las prácticas.

## **2. Elección del entorno de trabajo.**

Para poder cubrir todas las partes de un sistema de visión artificial había que conseguir capturar imágenes para posteriormente tratarlas. Así pues se debía tener un entorno de captura y otro de procesamiento. Para la captura se escogió el entorno MATROX, en concreto una tarjeta IP-8, por su facilidad de uso ya que dispone de librerías para C, lo cual resulta muy conveniente ya que es un lenguaje de alto nivel bastante implantado en el ámbito educativo. Para la fase de procesamiento se eligió el entorno MATLAB. Este entorno es indicado para este fin por varios motivos. Uno de ellos es el tipo de datos que maneja, las matrices, las cuales son también la forma clásica de almacenar imágenes. Por lo tanto, todas las operaciones en MATLAB se realizan con matrices por lo que todos los tratamientos podían realizarse sin tener que modificar para nada la estructura de representación de las imágenes. Pero además, existe otro aspecto fundamental: MATLAB provee de una herramienta ex-

clusiva de procesamiento de imágenes que contiene gran cantidad de operaciones para este fin. El uso del entorno con esta herramienta de procesamiento de imágenes y con la herramienta de procesamiento de la señal pone a disposición del usuario todas las condiciones necesarias para conseguir el objetivo deseado sin grandes complicaciones, una vez que se conoce el entorno. Para conseguir utilizar ambos entornos había que superar una pequeña dificultad. Las imágenes capturadas por Matrox tienen distinto formato que el que tienen las imágenes MATLAB. Así pues era necesario realizar un cambio de formato.

Una vez elegido el entorno de trabajo había que pasar a diseñar el entorno de usuario, tema este bastante crucial ya que de él depende, en gran manera, conseguir un resultado adecuado.

### 3. Diseño general

Cuando se planteó el problema de dar forma al programa de prácticas se quiso conseguir homogeneidad, sencillez y facilidad de uso. Así pues, había que conseguir un diseño compacto pero a la vez completo. Además se hacía necesario, en cualquier punto, la presencia de ayuda al usuario, de forma que éste pudiera seguir de manera correcta el proceso, sabiendo en todo momento a qué podía llevar la elección de una opción determinada.

Otro de los aspectos a tener en cuenta era que había que diseñar las prácticas para la versión del profesor y del alumno. La principal diferencia de ambas versiones debía ser el contenido y no el continente por lo que se optó por hacer diseños paralelos, es decir, ambas versiones tendrían el mismo reparto de menús y de botones y controles. La diferencia vendría en los resultados. Mientras que la versión del profesor tendría todas las opciones realizadas y funcionando, la versión del alumno se limitaría a tener el aspecto exterior que luego el propio alumno debía encargarse de rellenar. Por otro lado se creyó necesario dotar a la versión del alumno de una ayuda adicional orientada, en vez de al uso del entorno, a la realización de los procedimientos, o sea, de las prácticas. Se llegó, por tanto, al siguiente resultado:

- *Pantalla principal.* Constaria únicamente de un menú superior con cuatro opciones: Introducción, Procesamiento, Ayuda y Fin.

- *Pantallas de prácticas.* Compuestas por un menú superior constituido por: Imagen, Tratamiento, Ayuda y Volver a Principal.

- *Pantallas de tratamientos.* Compuestas por un cuadro de ejercicio, los botones necesarios para el tratamiento y un botón de fin.

A continuación se mostrará en detalle el diseño pasando a explicar más en profundidad cada uno de los componentes:

- *Pantalla principal.* Esta pantalla se diseñó de igual manera para ambas versiones (profesor y alumno). Se optó por tener una pantalla principal de comienzo (fig.2) desde la que se pudieran arrancar las dos partes principales en las que se dividió el programa de prácticas, es decir, una parte de Introducción y otra de Procesamiento. Además, habría una opción de Ayuda con la explicación de cada una de estas opciones. Por último, se añadiría un menú de Fin para terminar la sesión. Cada una de las dos partes de Introducción y Tratamiento tendrían a su vez otros submenús que serían los correspondientes a las distintas prácticas. En el menú Introducción se daría acceso a las dos primeras (“Conceptos básicos” y “Fourier y Filtrado”) y en el menú Procesamiento a la tercera y la cuarta práctica (“Procesamiento básico” y “Procesamiento avanzado”). Una vez que el usuario eligiera una de estas opciones pasaría directamente a la pantalla de la práctica.

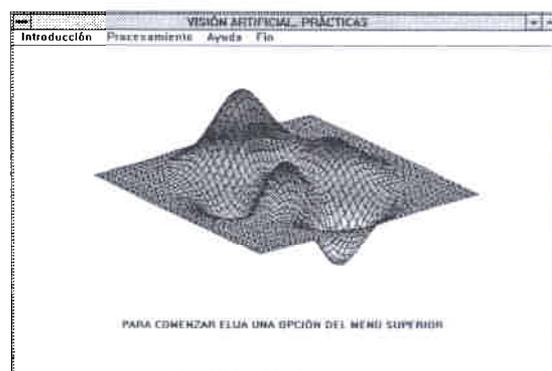


Figura 2. Pantalla Principal.

- *Pantallas de las prácticas.* También estas pantallas (fig.3) son iguales en ambas versiones. La explicación de esta decisión es sencilla: en estas pantallas aún no se hace ningún tratamiento sino que se muestran las partes en las que se divide la

práctica seleccionada. Para conseguir homogeneidad, se hizo un diseño común para todas las prácticas con cuatro menús: Imagen, Tratamiento, Ayuda (explicando los contenidos de los menús) y Volver a Principal (para volver a la pantalla principal y elegir una nueva práctica o finalizar la sesión). Los menús Imagen y Tratamiento se desplegarían a su vez en varios submenús. El menú Imagen, igual para todas las prácticas, se divide en: Capturar Imagen (captura una imagen con la cámara), Recuperar Imagen (recupera una imagen de memoria), Ayuda (sobre el submenú) y Salir (del submenú). El menú Tratamiento, desde el que se podrá elegir entre las diferentes opciones de que conste la práctica y que variarán para cada una. Además tendrá la opción de Ayuda sobre los tratamientos y la opción de Salir para la salida del submenú. Como complemento, y sólo cuando sea la primera vez (en la sesión actual) que se accede a una práctica, se visualizará un cuadro con el objetivo perseguido en cada caso y que indicará al usuario qué es lo que se pretende con la realización de los diversos tratamientos contenidos en la práctica elegida. Una vez que el usuario elige una de las opciones de tratamiento, pasa al diseño específico de la elección, el cual está integrado dentro de la misma pantalla de la práctica. Para evitar conflictos, cuando se entra en un tratamiento, todos los menús superiores quedan deshabilitados hasta que el usuario termine. Entonces vuelven a habilitarse automáticamente.

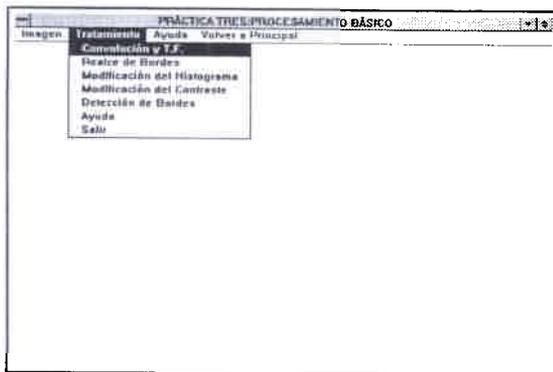


Figura 5. Pantalla de las prácticas

- **Pantallas de los tratamientos.** Aunque estas pantallas (fig.4) son muy similares en ambas versiones, la versión del alumno incorpora un botón de ayuda sobre la programación del tratamiento (fig.5). En estas pantallas no hay menús superiores ya que todos los componentes se encuentran dentro de la pantalla. Los componentes principa-

les y presentes en todas ellas, independientemente del tratamiento y de la práctica, son: Cuadro del ejercicio (ocupando toda la parte izquierda de la pantalla y donde se explica a grandes rasgos en qué consiste el tratamiento que se va a realizar y se indica el ejercicio que se debe llevar a cabo) y Botón de Fin (situado en la esquina inferior derecha, termina el tratamiento, limpia la pantalla y habilita el menú superior de las opciones de la práctica en la que se está). Además de éstas, cada opción incorporará otras componentes según el tratamiento. Estas serán botones que conllevarán unas acciones determinadas, menús de opciones que conllevarán distintos resultados... En todos los casos se situarán en la parte derecha de la pantalla.

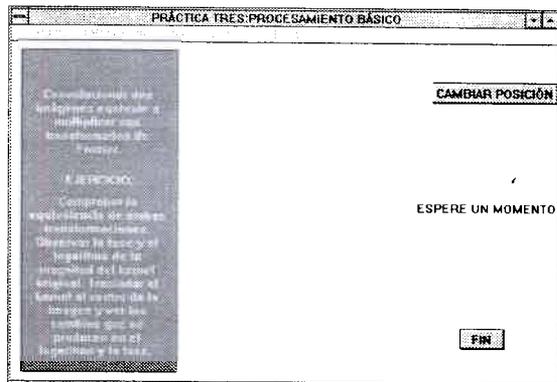


Figura 4. Pantalla tratamiento versión profesor

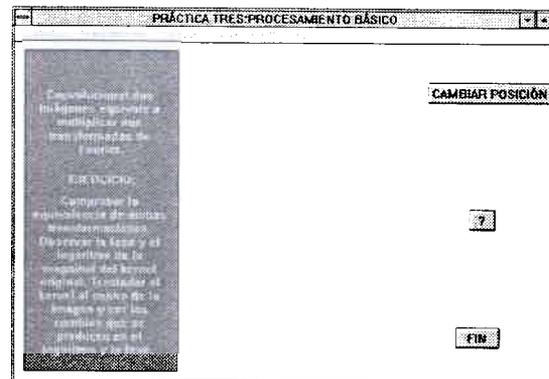


Figura 5. Pantalla tratamiento versión alumno

Para finalizar, queda por explicar la parte central de la pantalla, que es la que variará dependiendo de la versión. En la versión del profesor (fig.6) la parte central se utilizará para mostrar los resultados obtenidos en el tratamiento. Dependiendo de los casos se mostrará la imagen origen y la resultado, o se mostrarán dos tratamientos distintos. En la versión del alumno (fig.7) esta parte central estará ocupada por el cuadro de ayuda sobre la programación que se pone a su disposición para ayudarle en su tarea.

Para habilitar este cuadro de ayuda, la versión del alumno dispone de un botón en la parte derecha de la pantalla y que está etiquetado con el símbolo de interrogación (?).

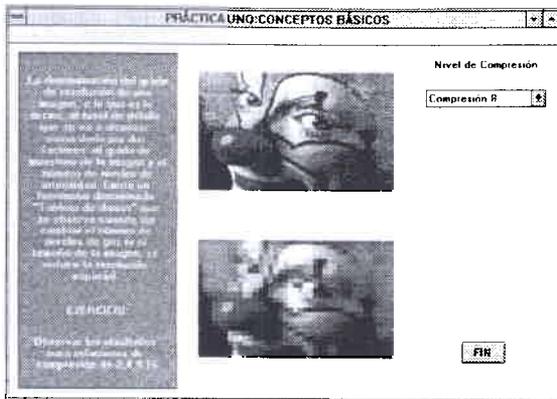


Figura 6. Parte central en la versión del profesor

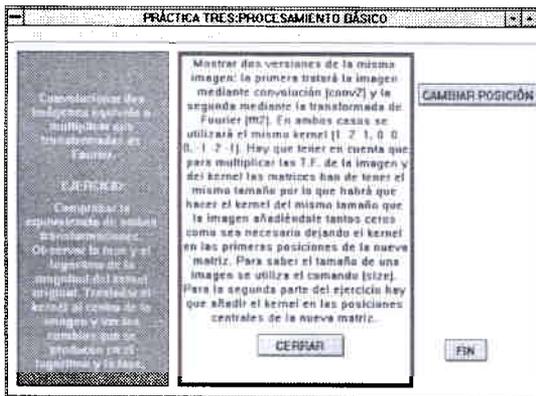


Figura 7. Parte central en la versión del alumno

## 4. Aportaciones

Con respecto a los cambios en relación a otros sistemas consultados, este trabajo aporta la captura de imágenes (de la cual carecen casi todos los examinados), lo que proporciona al sistema una ventaja importante en el hecho de dar opción al usuario a capturar una imagen en cualquier momento de la sesión desde el mismo entorno de trabajo, lo cual posibilita la experimentación de los mismos tratamientos en distintas imágenes, pudiéndose comparar los resultados obtenidos. La segunda aportación del sistema aquí presentado es la clasificación, que tampoco suele estar presente, ya que se sale del mero tratamiento de imágenes, a pesar de ser el objetivo de cualquier sistema de visión y que se ha realizado mediante el algoritmo de las *k*-medias, algoritmo que, a pesar de necesitar saber a priori el número de clases, es muy potente y bastante sencillo de programar.

## 5. Conclusiones

El intento de cubrir de forma estricta todas las etapas que definen un sistema de visión artificial ha dado lugar a un trabajo, cuando menos, preciso. Pero a la vez, creemos que también se ha conseguido la meta didáctica ya que se logró crear un interfaz cómodo y claro.

Se destaca, por otra parte, que dar la posibilidad de capturar imágenes contribuye en gran medida a estos objetivos ya que de esta forma el usuario siempre puede experimentar con distintas imágenes, que él mismo puede seleccionar, consiguiendo así una mayor comprensión en los tratamientos. Otro de los aspectos importantes, también ligado al interfaz gráfico con el que se trabaja, es el sistema de ayudas que proporciona tanto para la comprensión del entorno como para la programación de las prácticas. Esta posibilidad de ayudas garantiza una seguridad de movimiento por el entorno y una asistencia en la programación, integrada en el mismo, de forma que el usuario siempre puede utilizar la ayuda on-line del MATLAB basándose en las sugerencias dadas. Además, de esta manera, se da al alumno desde el mismo entorno de trabajo la primera idea "guía" con la que comenzar su tarea.

A pesar de ello hay que distinguir entre las características de una imagen y la detección y clasificación de objetos. Mientras en el primer campo se puede trabajar con cierta soltura, en el segundo las limitaciones son grandes y los resultados pocos. Este problema se agrava con referencia a la velocidad de procesamiento: El ordenador no es capaz de interpretar lo que "ve" en un tiempo razonable para que sea práctico.

Se ha observado que esta es una de las grandes trabas en este campo, por lo menos en las condiciones en las que se ha desarrollado el estudio. Lo que sí parece lógico es pensar que, debido a toda la carga computacional que las operaciones conllevan, resulta conveniente disponer de un equipo de trabajo bien provisto, tanto en memoria como en velocidad de procesador.

Una buena continuidad de esta actividad sería incluir la misma en una red, de forma que no hubiera necesidad de disponer de una cámara para cada usuario. La idea sería el realizar las capturas

desde una sola cámara a la que se pudiera acceder desde cualquier puesto. De esta forma se conseguiría un enfoque más de acuerdo con las dotaciones que normalmente disponen los laboratorios de prácticas. Otra posible mejora, que debería basarse en los propios usuarios del entorno, sería la ampliación o modificación del sistema de ayudas ya que, como se planteó previamente, se considera que es uno de

los puntos fuertes de este sistema de prácticas. El hecho de que se estime la posibilidad de reforma o mejora viene dado por nuestra propia condición de usuarios de programas informáticos. Generalmente los autores suelen pecar de parcos en las ayudas debido básicamente a su desenvoltura en el medio, resultando esta asistencia por lo general incompleta o escasa.

### **REFERENCIAS**

- Image Processing Toolbox de MATLAB.
- Fu, K.S.; González, R.C. y Lee, C.S.G. (1994). Robótica. Control, detección, visión e inteligencia. Ed. McGraw-Hill. México.
- Maravall Gómez-Allende, D. (1995). Reconocimiento de formas y visión artificial. Ed. RA-MA. España.
- Rodríguez Lozano, D. (1996). Proceso y análisis de imágenes digitales. Proyecto de fin de carrera. Facultad de Informática. Universidad de Sevilla.
- Manuales de MATLAB.
- Signal Processing Toolbox de MATLAB.
- Prácticas de visión artificial. Escuela Técnica Superior de Ingeniería Industrial. Universidad Carlos III. Madrid.

## **IV CONGRESO IBEROAMERICANO DE INFORMÁTICA EN LA EDUCACIÓN**

### **RIBIE'98**

**Hacia una educación libremente accesible, equitativa y de mejor calidad con  
las nuevas tecnologías de información y comunicación**



*Brasilia,  
20-23 de Octubre, 1998*

**SEDE DEL EVENTO:  
CENTRO DE CONVENÇÕES ULYSSES GUIMARÃES  
DE BRASILIA**

URL en el Internet: <http://penta.ufrgs.br/ribie98>

URL de la Ribie: <http://www-sce.fct.unl.pt/ribie>

E-Mail del Congreso: [ribie98@cesup.ufrgs.br](mailto:ribie98@cesup.ufrgs.br)