

## Procesamiento Audiovisual 3º II/ITIS

# Enunciado de la práctica 2008

---

### DESCRIPCIÓN GENERAL

- **Desarrollar y documentar** una aplicación que permita la entrada/salida de imágenes y vídeo, pudiendo aplicar sobre las mismas algunos de los principales tipos de operaciones, filtros y procesamientos estudiados en la asignatura.
- La aplicación deberá desarrollarse usando Borland C++ Builder® 6 y las librerías de procesamiento IPL y OpenCV.
- **Requisitos mínimos** de la aplicación:
  - La aplicación deberá poder abrir varias imágenes simultáneamente, pudiéndose aplicar las operaciones sobre cualquiera de ellas.
  - Se deberá trabajar con imágenes en color, admitiendo los mismos formatos que los reconocidos por OpenCV.
  - El menú principal deberá incluir, como mínimo, las opciones: **Archivo**, **Edición**, **Herramientas**, **Transformar**, **Efectos** y **Vídeo**. Cada una de ellas tendrá, como mínimo, los comandos que se detallan a continuación.
- Todo el trabajo desarrollado por los alumnos deberá documentarse. La **documentación** de la práctica deberá incluir: a) el diseño de la aplicación, b) una descripción de cómo se han resuelto los problemas más relevantes, c) el listado del código (sólo de los módulos de procesamiento, no de la parte de interface), y d) algunas pruebas y ejemplos de uso.

### MENÚ ARCHIVO

- El menú **Archivo** contiene las opciones típicas de entrada/salida de imágenes. En concreto, debe tener como mínimo los siguientes comandos:
  - **Nueva**. Crear una nueva imagen, especificando el tamaño y el color de fondo.
  - **Abrir**. Leer una imagen de un archivo existente, en cualquiera de los formatos admitidos por OpenCV.
  - **Guardar**. Guardar una imagen, sin necesidad de indicar el nombre.
  - **Guardar como...** Guardar una imagen, especificando el nuevo nombre.
  - **Cerrar**. Cerrar una imagen abierta. En caso de haber sido modificada, se dará la opción de guardarla.
  - **Capturar de cámara**. Capturar la imagen de una cámara de vídeo conectada al ordenador. Se podrá previsualizar la entrada de la cámara, pulsando una tecla o un botón cuando se quiera capturar una imagen.

- **Capturar de AVI.** Abrir una imagen que corresponda a un frame concreto, especificado por el usuario, dentro de un archivo de vídeo AVI. Se podrá previsualizar el frame, y colocarse en cualquier posición dentro del vídeo.
- **Nueva desde el portapapeles.** Opcional. Si el portapapeles de Windows contiene una imagen, la abre para editarla (ver TClipboard en la ayuda de Builder).
- **Salir.** Terminar el programa, liberando toda la memoria y dando opción a guardar las imágenes que hayan sido modificadas.

## MENÚ EDICIÓN

- De todas las ventanas de imágenes abiertas, estas operaciones del menú **Edición** (y también las de los restantes menús) se aplicarán sobre la ventana activa actualmente.
  - **Deshacer.** Opcional. Deshacer la última, o las **n** últimas, operaciones aplicadas sobre la imagen.
  - **Seleccionar todo.** Seleccionar toda la imagen, de cara a las operaciones en las que se use una región de interés (ROI).
  - **Copiar a nueva.** Crear una nueva imagen, con el ROI de la imagen actual.
  - **Copiar al portapapeles.** Opcional. Copiar el ROI de la imagen actual al portapapeles de Windows.
  - **Ver histograma.** Mostrar el histograma de una imagen, ya sea del valor de gris o de los canales por separado (R, G o B). Utilizar un formato parecido a de los histogramas del tema 2 (es decir, mostrando la escala horizontal y la vertical). Los histogramas deben estar normalizados de forma adecuada.
  - **Ver histograma 2D.** Opcional. Mostrar histogramas 2D agrupando canales de 2 en 2 (R-G, R-B o G-B).
  - **Ver información.** Mostrar la información relevante de la estructura `IpImage`: tamaño, profundidad, número de canales, memoria ocupada, etc.
  - **Opciones.** Configurar las opciones del programa, como el color actual, y otras que puedan surgir.

## MENÚ HERRAMIENTAS

- Las **Herramientas** son operaciones que se aplican al pinchar con el ratón sobre las imágenes en un punto concreto. El programa tendrá definido en todo momento un **radio** (tamaño de la transformación), un **difuminado** (suavizado en la transformación) y un **color actual**, que se aplicarán con estas operaciones. Siempre habrá una herramienta seleccionada en cada momento, que será la que se use. Como mínimo deben haber las siguientes herramientas:
  - **Puntos.** Pintar puntos en una imagen, con el tamaño, color y nivel de difuminación dados.
  - **Continuo.** Igual que la anterior, pero dibujando líneas entre cada par de posiciones donde se pinta con el ratón (en lugar de puntos sueltos).
  - **Líneas.** Pintar líneas en una imagen, dando un punto de inicio y otro de fin.
  - **Elipses.** Pintar elipses en una imagen, señalando el centro y el tamaño. La elipse puede ser rellena o sólo el borde (será rellena si el radio actual es 0).

- **Arco iris.** Igual que la herramienta “puntos”, dibujará puntos con distinto radio y difuminado. Pero el color del pincel irá cambiando de manera automática y rotativa según se pinta, siguiendo una paleta como en la página 78 del tema 2.
- **Rellenar.** Opcional. Rellenar una región de tono uniforme con el color actual (ver cvFloodFill). En este caso, el radio indica el rango de tolerancia del relleno. El difuminado indicará si el color relleno es opaco o más o menos transparente.
- **Suavizado.** Opcional. Suavizar un trozo circular de la imagen (usando media o gaussiana), según el sitio donde se pulse, con el radio y el valor de difuminado actual.
- **Copiar.** Opcional. Esta herramienta permite copiar un trozo de una imagen en un trozo de otra. Funciona de la siguiente forma. Primero pinchamos sobre una imagen, para seleccionar la imagen desde la que se copia y el punto de origen de la copia. En los siguientes sitios donde se pinche (en la misma imagen o en otra) se copiará un trozo circular de la imagen de origen, según el tamaño y el difuminado seleccionados.
- **Seleccionar.** Seleccionar una región de interés (ROI) dada, de forma rectangular. La ROI se usará en las operaciones que se indique.

## MENÚ TRANSFORMAR

- Las **transformaciones** son operaciones que se aplican sobre toda la imagen. Normalmente el resultado será una nueva imagen, de forma que no se modifica la imagen de entrada, sino que se produce una nueva.
- Opcionalmente, se puede **previsualizar** el resultado que tendrá la operación antes de realizarla. En las operaciones donde sea adecuado, se puede mostrar cómo será el histograma de la imagen resultante.
- Se deben incluir las siguientes operaciones en el menú **Transformar**:
  - **Ajustar brillo/contraste/gama.** Modificar una imagen, indicando el brillo (valor sumado) y el contraste (valor multiplicado) y la gama.
  - **Ajustar matiz/saturación/luminosidad.** Realizar un ajuste de color, usando internamente el espacio HLS, consistente en modificar el matiz, la saturación y la luminosidad de la imagen.
  - **Ecualización del histograma.** Opcional. Ecualizar el histograma de una imagen, permitiendo que sea de forma independiente por canales o conjunta.
  - **Ajuste lineal del histograma.** Operación de ajuste o estiramiento lineal del histograma. Se debe poder indicar el porcentaje de píxeles que se dejan a la derecha y a la izquierda en el histograma.
  - **Rotar imagen y reescalar.** Opcional. Rotar una imagen en un ángulo cualquiera y al mismo tiempo reescalarla al tamaño indicado por el usuario. La imagen de salida deberá estar centrada y tener un tamaño suficiente para caber entera.
  - **Perspectiva y bilineal.** Crea una nueva imagen, con el tamaño indicado por el usuario. En esta imagen se permite que el usuario pinche en 4 puntos, y hace una transformación perspectiva o bilineal de la imagen actual en esos 4 puntos.
  - **Transformación de malla.** Opcional. Realizar una transformación de malla, pudiendo especificar el tamaño de la malla usada. Se deben poder modificar de forma gráfica los puntos de la malla, previsualizando el resultado después de cada modificación. Se deben controlar los posibles problemas debidos a la concavidad de los rectángulos de la malla.

- **Convertir a escala de color.** Convertir una imagen a escala de grises o escala de color, según el color seleccionado actualmente.
- **Convertir a color falso.** Opcional. Conversión de escala de grises a una paleta de color falso (estilo “espectro”).
- **Media ponderada de dos imágenes.** Calcular una nueva imagen con la media de dos imágenes, pudiendo especificar la ponderación. Si ambas imágenes no tienen el mismo tamaño, se deberá escalar una al tamaño de la otra.
- **Curva tonal.** Opcional. Realizar una transformación de curva tonal genérica, introducida de forma gráfica por el usuario.
- **Modelos de color.** Opcional. Aplicar transformaciones del modelo de color de la imagen, entre RGB, HLS, HSV, XYZ, YUV, etc.
- **Espectro.** Opcional. Dada una imagen cualquiera, obtiene su espectro de intensidad (magnitud de la transformada de Fourier centrada).

## MENÚ EFECTOS

- Los **Efectos** son operaciones que se puede realizar sobre las regiones de interés seleccionadas o bien sobre toda la imagen. Además, se opera sobre las propias imágenes sin necesidad de crear imágenes nuevas.
- Opcionalmente, se puede **previsualizar** el resultado que tendrá la operación antes de realizarla. También se puede incluir un cuadro de chequeo para que la imagen se cree en una ventana nueva o no.
- El programa deberá incluir por lo menos los siguientes **Efectos**:
  - **Suavizado: gaussiano, media o mediana.** Aplicar un suavizado, según el tipo y el tamaño indicados.
  - **Perfilado.** Opcional. Perfilar la imagen, especificando el radio y el porcentaje de perfilado aplicado. Se debe poder previsualizar el resultado.
  - **Texto.** Añadir un texto a la imagen, indicando el contenido del texto, tamaño, color y el sitio. El texto deberá tener un efecto de sombra difusa y semitransparente.
  - **Morfología matemática.** Opcional. Aplicar las diferentes operaciones de morfología matemática (dilatación, erosión, cerrar y abrir), indicando el número de iteraciones.
  - **Buscar patrón.** Opcional. Dada una imagen grande y otra imagen pequeña (o un ROI de la misma), busca las apariciones más probables del patrón, señalándolas con un recuadro de color rojo.
  - **Pinchar/estirar.** Opcional. Transformaciones geométricas de pinchar/estirar una imagen. Se debe poder indicar el punto central, el radio del efecto y la cantidad a aplicar. En este caso no se usará la ROI.
  - **Balance de blancos.** Ajustar automáticamente el balance de blancos de una imagen.
  - **Geométricos.** Aplica sobre una imagen una transformación geométrica. El usuario debe poder decidir entre: cilíndrica, aleatoria, acristalado, y superficie deformante.
  - **Bajorrelieve.** Aplica un efecto de bajorrelieve a la imagen, pudiendo seleccionar la dirección y cantidad del efecto.

## MENÚ VÍDEO

- Este menú contiene algunas operaciones adicionales usando imágenes y vídeo, desde cámara o desde AVI. Las imágenes serán algunas de las abiertas actualmente y los archivos de vídeo los seleccionará el usuario al abrir la opción.
- Las operaciones “Vídeo a...” se pueden implementar usando entrada de cámara o de AVI. Se pide que haya por lo menos una operación usando entrada de AVI y otra usando cámara (además de las que ya están hechas en el programa base). En caso de usar AVI, se deben poder especificar los frames a utilizar (inicio y fin).
- Se deben incluir por lo menos las siguientes operaciones:
  - **Imagen a Vídeo.** Operaciones que dada una o varias imágenes de entrada, producen un archivo AVI. En todas ellas se debe poder indicar el número de frames a generar.
    - **Rotar imagen.** Dada una imagen, produce una secuencia de vídeo con una rotación completa de la imagen.
    - **Transición entre imágenes.** Dadas dos imágenes, realizar una transición entre las mismas. Se deben escalar si no tienen el mismo tamaño. Deben ofrecerse por lo menos dos métodos distintos, aplicando diferentes técnicas. Por ejemplo, aplicar un zoom de aumento sobre la primera imagen, y luego un zoom de decremento de la segunda hasta llegar al tamaño original.
    - **Guerra de las Galaxias.** Opcional. Dada una imagen y un texto introducido por el usuario, genera un vídeo donde se va mostrando el texto de abajo hacia arriba, con una transformación perspectiva (similar a los títulos iniciales de las películas de la Guerra de las Galaxias).
  - **Vídeo a Imagen.** Operaciones que dada una secuencia de vídeo (de archivo o de cámara) producen una imagen.
    - **Media** Obtener la media en cada píxel de los frames considerados.
    - **Mínima, máxima.** Opcional. Obtener dos imágenes, con el mínimo y el máximo en cada píxel de los frames considerados.
    - **Movimiento.** Produce una imagen en la que señalan las zonas donde hay más movimiento en el vídeo. Para cada par de frames consecutivos, se calculará la diferencia y se irá acumulando en la imagen resultante.
  - **Vídeo a Vídeo.** Operaciones que dada una secuencia de vídeo (de archivo o de cámara) producen un archivo AVI.
    - **Copiar ajustado.** Copiar una secuencia a un archivo, añadiendo en una esquina de las imágenes un logotipo pequeño (fijado por los alumnos). Antes de añadir el logotipo, deberá aplicarse la operación de ajuste lineal del histograma a todos los frames escritos en la salida.
    - **Suavizado temporal.** Dada una secuencia de vídeo, produce un vídeo de salida donde cada frame es una media de los  $n$  frames anteriores en la entrada. El usuario debe poder seleccionar el valor de  $n$ , entre 1 y 30.
    - **Detección de fondo.** Opcional. Dado un modelo de fondo (que se creará con los dos primeros frames de la secuencia de entrada), producir otro vídeo donde el fondo se sustituya por la imagen activa actualmente.

## MINIPROYECTOS

- Las técnicas fundamentales para resolver estos miniproyectos son las que se estudian en la asignatura. No obstante, se considera que estos miniproyectos implican por parte de los alumnos cierta **labor de investigación**: saber combinar y adaptar las técnicas estudiadas en clase, hacer experimentación y pruebas con distintas posibilidades, ajustar los parámetros para conseguir los mejores resultados, e incluso buscar bibliografía adicional. Los alumnos deberán ir informando de sus progresos al profesor.
- Antes de empezar a trabajar, los alumnos deben hablar con el profesor, que buscará un reparto uniforme de los temas. Se proponen los siguientes.
  - **Mosaicos de imágenes.**
    - **Descripción.** Un mosaico es una imagen que se forma juntando trozos pequeños de otras imágenes. Tenemos una imagen objetivo (la que queremos formar) y una base de datos de imágenes. La base de datos de imágenes se puede obtener de un directorio (todos los archivos gráficos contenidos) y/o de uno o varios vídeos. Se debe poder especificar el tamaño del resultado y el número de trozos que lo componen. En cada trozo se debe poner la imagen de la BD que más se asemeje a ese fragmento del objetivo. Se permitirán varios modos: sin variar las imágenes de la BD, pudiendo variar su intensidad, o variando su color; y usando las imágenes de la BD completas o trozos de ellas. Se deben manejar tamaños grandes de las imágenes resultantes. Se pondrá un énfasis especial en la creación de la base de datos de imágenes.
    - **Técnicas a usar.** Transf. geométricas, comparación de imágenes, matching de patrones, transf. de color.
  - **Protección de testigos y falsa identidad.**
    - **Descripción.** Dada una secuencia de vídeo en la que pueden aparecer varias caras humanas, crear otra secuencia de vídeo en la que se oculten las caras, mediante suavizado, pixelado, banda negra, etc. Esta es la protección de testigos. La falsa identidad consistirá en sustituir la zona de la cara en el vídeo por otra cara que estará en una imagen estática (escalada y ajustada de manera adecuada). Se deberá buscar estabilidad (evitar oscilaciones), un buen ajuste a la cara, y un resultado de la mayor calidad posible (realismo, en el caso de la falsa identidad, es decir, que la nueva cara esté en la misma posición y orientación de la antigua).
    - **Técnicas a usar.** Detección de caras humanas (ver cvHaarDetect-Objects), seguimiento de objetos, transf. geométricas, transf. de color, suavizado temporal.
  - **Croma-key y detección de fondo.**
    - **Descripción.** Dado un vídeo donde aparecen cosas delante de un fondo, segmentar el fondo de la escena y sustituirlo por otro fondo dado. En el sistema croma-key, el fondo es una pared de cierto color uniforme y destacado (aunque realmente no es tan uniforme como se esperaría, pudiendo variar la intensidad). En detección de fondo puede haber

- cualquier fondo, lo único que se sabe es que no cambia. El nuevo fondo puede ser una imagen fija o un vídeo (que se repite indefinidamente). El sistema croma-key también puede ocupar sólo una parte de la imagen (una región rectangular) a la cual se debe mapear el nuevo fondo, a modo de un proyector de vídeo. Se intentará conseguir la mayor calidad y estabilidad posible (evitar falsos positivos y falsos negativos).
- **Técnicas a usar.** Espacios de color, segmentación por color, modelos de fondo, transf. geométricas, detección de bordes, morfología matemática.
- **Aplicaciones de interacción visual.**
    - **Descripción.** El mecanismo y las técnicas subyacentes a este miniproyecto son las mismas que las del sistema de croma-key. En especial, se trabajará con técnicas de modelado y detección de fondo, o bien con las técnicas de detección de color. Sin embargo, ahora el resultado no se usará para sustituir por un fondo distinto, sino para permitir cierta interacción con el usuario; por ejemplo, en un juego de tipo Arkanoid. Por lo tanto, será importante conseguir una ejecución rápida de los algoritmos implementados.
    - **Técnicas a usar.** Espacios de color, segmentación por color, modelos de fondo, transf. geométricas, detección de bordes, morfología matemática.
  - **Composición de vídeo por barrido.**
    - **Descripción.** Tenemos una escena estática y una cámara móvil. Si la cámara gira sobre sí misma o se mueve hacia un lado, vemos un nuevo trozo de la escena. Juntando las dos imágenes (con los desplazamientos adecuados) obtenemos una visión más amplia de la escena. Juntando todos los frames de un vídeo, con la cámara haciendo un giro completo de 360°, podemos tener una panorámica completa de toda la escena. Si la velocidad de giro fuera constante, un sistema de fotofinish podría valer para obtener la composición. Pero como esto es difícil de conseguir, habrá que calcular el desplazamiento entre cada par de imágenes. Se deben poder crear composiciones de 360° u otras de menor tamaño (moviendo la cámara en cualquier dirección). Se buscará una buena calidad en la composición, haciendo ajustes de intensidad en caso necesario.
    - **Técnicas a usar.** Combinación de imágenes, cálculo del flujo óptico (Optical Flow), transf. geométricas, transf. de color.
  - **Construcción de panoramas.**
    - **Descripción.** El objetivo es parecido a la composición de vídeo, obtener una imagen juntando trozos de otras imágenes. Pero en este caso partimos de un conjunto más reducido de imágenes (entre 2 y 6) que están más espaciadas entre sí. El problema clave es seleccionar los puntos equivalentes en las imágenes. Se debe permitir una selección manual o automática. La composición debe intentar que las líneas de corte entre las imágenes sean lo más suaves posible, haciendo una media ponderada progresiva, o cortando por el sitio donde más se parecen las imágenes.

- **Técnicas a usar.** Búsqueda de patrones, comparación de imágenes, transf. geométricas, filtros, combinación de imágenes, transf. de color.
- **Metamorfosis o morphing de imágenes.**
  - **Descripción.** Dado un par de imágenes y un conjunto de puntos equivalentes en ambas, crear un vídeo en el que haga una transición suave de una a otra. Los puntos serán seleccionados por el usuario de manera manual, pinchando en las imágenes. El efecto implica una transformación geométrica y una mezcla de imágenes. Probar inicialmente con una malla rectangular (de tamaño variable) y transf. bilineales, y después con otras técnicas que faciliten la selección de correspondencias y mejoren la calidad del resultado.
  - **Técnicas a usar.** Transf. geométricas, combinación de imágenes, interpolación.
- **Herramientas en el dominio frecuencial.**
  - **Descripción.** La manipulación de imágenes en el dominio frecuencial es una técnica muy interesante y ampliamente tratada en libros de procesamiento de imágenes. Sin embargo, las aplicaciones de retoque fotográfico no suelen incluirlas (al menos de forma explícita). Se trata de añadir operaciones en el dominio frecuencial: convertir una imagen espacial a frecuencial (en modo parte real/parte imaginaria, o en modo magnitud/gradiente), hacer la conversión inversa (de frecuencial a espacial), editar la imagen en el dominio frecuencial (con las herramientas de pintar puntos o líneas), filtros en el dominio frecuencial y estudiar las técnicas de deconvolución y filtros de Wiener.
  - **Técnicas a usar.** Transf. lineales (especialmente DFT, DCT y transf. inversas).
- **Marcos virtuales.**
  - **Descripción.** Un rectángulo detectado en una imagen puede ser sustituido (mediante una transformación perspectiva) por otra imagen cualquiera. El rectángulo podría ser, por ejemplo, un recuadro blanco hueco. Repetido el proceso sobre una secuencia de vídeo, se obtiene el efecto de un marco “porta fotos” virtual. En este miniproyecto se trata de diseñar un método fiable para detectar rectángulos blancos en una imagen, y aplicarlo a una secuencia de vídeo obteniendo el efecto de marco virtual. Se puede trabajar con varios marcos simultáneamente. Por ello, es importante el problema de seguimiento: saber qué recuadros del frame anterior corresponden a qué recuadros del siguiente.
  - **Técnicas a usar.** Detección de bordes, detección de segmentos (transformada de Hough), detección de rectángulos (ver `OpenCV\c\samples\squares.c`), seguimiento de objetos, transf. geométricas.
- **Indexación de galerías de imágenes.**
  - **Descripción.** La búsqueda de imágenes por contenido es una herramienta muy interesante, que permitiría dar una enorme potencia a los grandes almacenes de contenidos multimedia. El usuario realiza una consulta que puede ser por una imagen, un trozo de una imagen, o bien alguna categoría simbólica. El sistema debe buscar las imágenes más

parecidas de su base de imágenes, las que contengan el trozo dado, o las que se encuentren en la misma categoría. Se pueden usar distintas técnicas para la comparación: diferencia entre imágenes, análisis de histogramas de color, matching de patrones, etc.

- **Técnicas a usar.** Histogramas, matching de patrones, transf. geométricas, transf. globales, etc.

## EVALUACIÓN

- La práctica contará un 80% en la nota final de la asignatura. El examen contará un 20%. La asistencia a clase podrá sumar hasta 2 puntos en la nota de examen. Para ello, los alumnos deben tener un mínimo de 6 asistencias a clase. Sea  $na$  el número de asistencias del alumno, y  $nc$  el número de clases impartidas, la nota será:

$$\text{Nota de clase} = 2 \cdot (na - 5) / (nc - 5)$$

- Para los alumnos que no hagan el miniproyecto, el 60% de la nota de prácticas corresponde a las operaciones obligatorias y el 40% a las opcionales. Las partes opcionales se valorarán según su dificultad. De forma orientativa, para alcanzar la máxima puntuación en los opcionales basta con hacer la mitad de los mismos. Se entregará un programa base a los alumnos, que pueden usar en la resolución de la práctica.
- Para los alumnos que hagan un miniproyecto, su evaluación será el 100% de la nota de prácticas. Además, los alumnos podrán obtener hasta un +1 (a sumar en la nota total) exponiendo en clase el miniproyecto realizado.
- Para aprobar la práctica es imprescindible que todas las operaciones implementadas funcionen correctamente. Además, la evaluación también tendrá en cuenta la documentación y la defensa de la práctica realizada en la entrevista (que será obligatoria).
- La práctica se podrá realizar en grupos de dos alumnos.
- Los alumnos deberán entregar la documentación **en papel**, incluyendo el código de los módulos de proceso (no de los de interface). También deberán entregar el código (todo lo necesario para compilar el programa) y la documentación en **formato electrónico** (se admite el uso de email a ginesgm@um.es).
- La fecha tope de entrega de la práctica es el martes 1 de julio de 2008.
- Sobre mediados de mayo se convocarán entrevistas de seguimiento con los alumnos, para comprobar el estado de resolución de la práctica.
- **Nota:** todas las prácticas serán sometidas a un **sistema automático de detección de copias**, que contiene también las prácticas de años anteriores. La copia de la práctica supondrá el suspenso **fulminante** en la asignatura.