Cuestionario T1

1.- ¿Cuáles son los objetivos principales de las técnicas de visión por computador? Poner algún ejemplo si lo necesita.
Los objetivos de las técnicas de visión por computador son la interpretación de las imágenes para extraer de manera automática información de éstas mediante el tipo de dato que por defecto son píxeles. Esta información puede ser de carácter semántico o geométrico, lo que implica reconocimiento de patrones, reconstrucción de escenarios, etc.

2.- ¿Una máscara de convolución para imágenes debe ser siempre una matriz 2D? ¿Tiene sentido considerar máscaras definidas a partir de matrices de varios canales como p.e. el tipo de OpenCV CV_8UC3? Discutir y justificar la respuesta.
Una máscara de convolución no necesariamente ha de ser una matriz 2D. La información de una máscara del filtro gaussiano, por ejemplo, se puede almacenar en un vector o matriz 1D. De esta forma podríamos aplicar este filtro por filas o columnas según el objetivo que se desee; en el caso de simular una matriz 2D con un vector, aplicaríamos primero la convolución por filas y después por columnas o al revés.

$I \otimes M = I \otimes (M1 \otimes M2)$ Donde I es la imagen y M la máscara cuadrada.

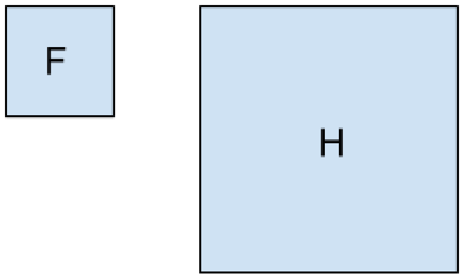
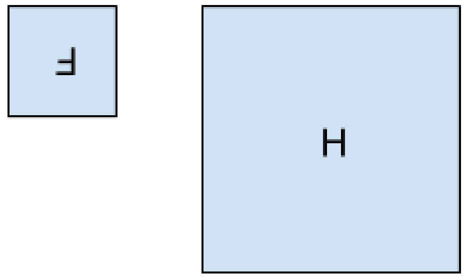
No obstante el proceso con un vector es más óptimo que con una matriz 2D pues la eficiencia del vector es de $O(MN^2)$ mientras que el de la matriz 2D es de $O(M^2N^2)$.

Normalmente los filtros que se suelen usar realizan una transformación continua de los píxeles por lo que no tendría sentido una máscara de tres canales. Sin embargo si se quisieran realizar operaciones concretas sobre los colores de los píxeles si podría tener sentido.

3.- Expresar y justificar las diferencias y semejanzas entre correlación y convolución. Justificar la respuesta.
Correlación y convolución son técnicas para el procesamientos de imágenes que tienen una misión diferente. A pesar de ello, cuando nos encontramos con un filtro simétrico, tanto correlación como convolución tendrán el mismo efecto.

Las diferencias entre ambos radica en que la convolución describe transformaciones de sistemas lineales o lo que sería la aplicación de un filtro de suavizado a una imagen (con la eliminación del ruido existente), mientra que con la correlación se intenta medir la similitud de patrones y detección de objetos.

Y tal y como antes se ha mencionado, si el núcleo no es isotrópico entonces correlación y convolución no es lo mismo:

Correlación:	Convolución:
	

4.- ¿Los filtros de convolución definen funciones lineales sobre las imágenes? ¿y los de mediana? Justificar la respuesta.
La función de convolución alisa la imagen comportándose de la misma forma para cada valor o pixel de la imagen, realizando un promedio de los valores que rodean el píxel principal y dependiendo de los pesos de la máscara por lo que sí sería una función lineal.

El filtro de media no es lineal ya que éste se encarga de eliminar ciertos tipos de ruido como el de sal y pimienta, modificando el valor de los píxeles por la mediana de los valores de la máscara y no una operación promedio como el caso de la convolución.

5.- ¿La aplicación de un filtro de alisamiento debe ser una operación local o global sobre la imagen? Justificar la respuesta.
La aplicación de un filtro de alisamiento se realiza mediante la convolución con una máscara que realiza el promedio de los píxeles que rodean al principal. Como hemos dicho anteriormente es una función lineal y por tanto también será una operación local.

La aplicación de un filtro de alisamiento trabaja sobre los extremos de la derivada de la función intensidad, es decir, donde hay un cambio rápido en la intensidad. La derivada como es una operación local, también lo serán los filtros de alisamiento al usarlas.

6.- Para implementar una función que calcule la imagen gradiente de una imagen dada puede plantearse dos alternativas:

a) **Primero alisar la imagen y después calcular las derivadas sobre la imagen alisada.**
b) **Primero calcular las imágenes derivadas y después alisar dichas imágenes.**

Discutir y decir qué estrategia es la más adecuada, si alguna lo es, tanto en el plano teórico como de implementación. Justificar la decisión.
En el plano teórico tendríamos que ambas cosas equivaldrían a lo siguiente:

a) $\frac{\delta}{\delta x}(h \star f), \frac{\delta}{\delta y}(h \star f)$

b) $\frac{\delta h}{\delta x} \star f,\ \frac{\delta h}{\delta y} \star f$

Si h es Gaussiana, por el derivative theorem of convolution sabemos que entonces la siguiente igualdad es cierta:

$$\frac{\delta}{\delta x}(h \star f) = \frac{\delta h}{\delta x} \star f$$

Por tanto teóricamente ambas soluciones son válidas al poder tener el mismo resultado.

En el plano de la implementación tenemos que:

a) Se realiza una sola vez la convolución y posteriormente dos derivadas.
b) Se realizan dos convoluciones y posteriormente dos derivadas.

Esto nos hace ver que la primera va a ser más eficiente realizando menos tarea de cómputo. Otro dato importante es que la opción b) al calcular las derivadas primero, está acentuando el ruido que pueda existir en dicha imagen y que no se mejorará al alisar posteriormente. Por tanto nos quedaríamos que la opción a) es la más adecuada.

7.- Verificar matemáticamente que las primeras derivadas (respecto de x e y) de la Gaussiana 2D se puede expresar como núcleos de convolución separables por filas y columnas. Interpretar el papel de dichos núcleos en el proceso de convolución.

Tenemos que la Gaussiana es la siguiente:

$$G\sigma(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Donde las derivadas parciales de x e y son:

$$\frac{\delta}{\delta x} G\sigma(x,y) = \left(- \frac{1}{2\pi\sigma^4} x e^{-\frac{x^2+y^2}{2\sigma^2}} \right)$$

$$\frac{\delta}{\delta y} G\sigma(x,y) = \left(- \frac{1}{2\pi\sigma^4} y e^{-\frac{x^2+y^2}{2\sigma^2}} \right)$$

Esto puede expresarse como el producto de una función x y una función y (la misma forma tanto para la derivada en y):

$$\frac{\delta}{\delta x} G\sigma(x,y) = \left(- \frac{1}{2\pi\sigma^4} x e^{-\frac{x^2}{2\sigma^2}} \right) \cdot \left(\frac{1}{2\pi\sigma^4} e^{-\frac{y^2}{2\sigma^2}} \right)$$

$$\frac{\delta}{\delta y} G\sigma(x,y) = \left(- \frac{1}{2\pi\sigma^4} e^{-\frac{x^2}{2\sigma^2}} \right) \cdot \left(\frac{1}{2\pi\sigma^4} y e^{-\frac{y^2}{2\sigma^2}} \right)$$

Para el proceso de convolución ambos núcleos son simétricos y se resaltarán los bordes en dirección vertical para la derivada del eje X mientras que para el eje Y se resaltarán los bordes horizontales.

8.- Verificar matemáticamente que la Laplaciana de la Gaussiana se puede implementar a partir de núcleos de convolución separables por filas y columnas. Interpretar el papel de dichos núcleos en el proceso de convolución.

Tenemos que la Laplaciana de la Gaussiana es la siguiente:

$$\nabla^2 G\sigma(x,y) = \frac{\delta^2 G\sigma(x,y)}{\delta x^2} + \frac{\delta^2 G\sigma(x,y)}{\delta y^2}$$

Esto es lo mismo que decir:

$$\frac{\delta^2}{\delta x^2}\ G\sigma(x,y) = \frac{\delta}{\delta x}\left(-\frac{1}{2\pi\sigma^4} x e^{-\frac{x^2+y^2}{2\sigma^2}}\right) = \left(-\frac{1}{2\pi\sigma^6} x e^{-\frac{x^2+y^2}{2\sigma^2}}\right)\cdot\left(\sigma^2 - x^2\right)$$

$$\frac{\delta^2}{\delta x^2}\ G\sigma(x,y) = \frac{\delta}{\delta y}\left(-\frac{1}{2\pi\sigma^4} y e^{-\frac{x^2+y^2}{2\sigma^2}}\right) = \left(-\frac{1}{2\pi\sigma^6} y e^{-\frac{x^2+y^2}{2\sigma^2}}\right)\cdot\left(\sigma^2 - y^2\right)$$

Una vez que ya tenemos las segundas derivadas hay que aplicar la primera fórmula(∇^2) en la que definimos la Laplaciana de la Gaussiana como suma de productos de funciones en x por funciones en y:

$$\nabla^2 G\sigma(x,y) = \frac{\delta^2 G\sigma(x,y)}{\delta x^2} + \frac{\delta^2 G\sigma(x,y)}{\delta y^2} =$$

$$= \left[-\left(\frac{1}{2\pi\sigma^6} x e^{-\frac{x^2+y^2}{2\sigma^2}}\right)\cdot(\sigma^2 - x^2)\right] + \left[-\left(\frac{1}{2\pi\sigma^6} y e^{-\frac{x^2+y^2}{2\sigma^2}}\right)\cdot(\sigma^2 - y^2)\right] =$$

$$= -\left(\frac{1}{2\pi\sigma^4} x\left(1 - \frac{x^2}{\sigma^2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}\right) - \left(\frac{1}{2\pi\sigma^4} y\left(1 - \frac{y^2}{\sigma^2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}\right)$$

9.- ¿Cuáles son las operaciones básicas en la reducción del tamaño de una imagen? Justificar el papel de cada una de ellas.

Las dos operaciones básicas son el alisamiento y el submuestreo. Cualquier persona sin conocimientos técnicos podría afirmar que para reducir una imagen simplemente se eliminan filas y columnas de la imagen original y se soluciona el problema pero no es así. Haciendo esto podríamos encontrarnos problemas de ruido en la imagen reducida generados por la pérdida de información al eliminar filas y/o columnas. Este problema se soluciona aplicando un filtro de alisamiento a la imagen antes de realizar el submuestreo ya que de esos píxeles que perdemos vamos a guardar su información en los que no se eliminan (al realizar el promedio).

10.- ¿Qué información de la imagen original se conserva cuando vamos subiendo niveles en una pirámide Gausssiana? Justificar la respuesta.

En la primera imagen de una pirámide Gaussiana vamos a observar las frecuencias altas y a medida que aumentemos los niveles y la imagen reduzca su tamaño iremos perdiendo estas frecuencias y nos iremos quedando con las bajas. Un ejemplo de esto podemos verlo en la siguiente imagen:

11.- ¿Cuál es la diferencia entre una pirámide Gaussiana y una Pirámide Laplaciana? ¿Qué nos aporta cada una de ellas? Justificar la respuesta. (Mirar en el artículo de Burt-Adelson)
La pirámide Gaussiana consiste en aplicar el alisamiento gaussiano y posterior submuestreo de la imagen resultante del alisamiento (todo esto se realiza tantas veces como niveles tenga la pirámide) con el objetivo de conservar las frecuencias bajas.

La pirámide Laplaciana a diferencia de la pirámide Gaussiana se basan en conservar las frecuencias medias y altas de la imagen ya que lo que se realiza es ir sustrayendo las frecuencias bajas.

Para la codificación eficiente de las imágenes la eficiencia de compresión es mayor en imágenes con frecuencias más altas, que se debe a que las imágenes de la pirámide Laplaciana tienen entropía y varianza pequeña, es por ello que es eficiente.

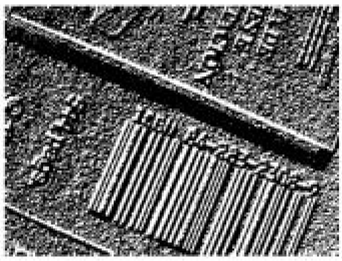
12.- Cuál es la aportación del filtro de Canny al cálculo de fronteras frente a filtros como Sobel o Robert. Justificar detalladamente la respuesta.
El filtro de Canny utiliza la derivada de la Gaussiana para calcular el gradiente de la imagen. Aplica la supresión de los no máximos en el que por cada píxel de la imagen gradiente se establece una región para comparar la intensidad de los píxeles vecinos que tienen la misma dirección gradiente y si es mayor, se conserva el píxel del borde y en caso contrario lo elimina. A continuación la histéresis y enlazado donde se establecen el umbral alto y bajo. Aquellos píxeles con una intensidad menor al umbral bajo se eliminan mientras que los píxeles con intensidad mayor al umbral alto serán un borde.

En cambio el filtro de Sobel y Roberts sólo aplican dos máscaras de convolución a la imagen (eje x e y) para obtener el gradiente donde uno detecta los bordes horizontales y verticales y el otro obtiene con eficacia los bordes diagonales.

13.- Buscar e identificar una aplicación real en la que el filtro de Canny garantice unas fronteras que sean interpretables y por tanto sirvan para solucionar un problema de visión por computador. Justificar con todo detalle la bondad de la elección.

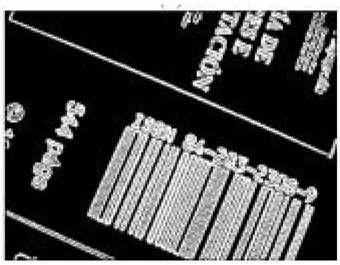
Una aplicación para el filtro de Canny es la detección de los bordes de los códigos de barras:
En las siguientes imágenes se podrá ver los pasos que se llevan a cabo que son la obtención del gradiente (2), supresión no máxima (3) y histéresis del umbral (4).

(1) Imagen Original



(2) Obtención del gradiente

(3) Supresión de no máximos



(4) Histéresis del umbral

De esta forma resalta y son más detallados los bordes y la distancia entre ellos pudiendo diferenciarlos para posteriormente analizarlos y obtener el código.

Fuente que explica con detalle la aplicación: [UJI Vicente Castellon]

14.- Suponga que necesita implementar un detector de distintas características de la imagen, p.e. bordes y cruces por cero, usando filtros gaussianos. Explique cómo abordaría la implementación de los filtros para garantizar su perfecto funcionamiento en todos los casos.

Si el filtro que se quiere implementar es el de la primera derivada Gaussiana, hay que aplicarle al filtro Gaussiano (g) [-1 1] h es decir $g \otimes h$. Aquí lo que se quiere obtener son los máximos locales, es decir, serán bordes los picos de la primera derivada.

En caso de la segunda derivada de la Gaussiana, o lo que sería la Laplaciana de la Gaussiana habría que pasar primero un filtro de suavizado a la imagen ya que es muy sensible al ruido. Se consideran bordes a los que producen un máximo local al detectar el cruce por cero en la segunda derivada. Hay que aplicarle al filtro Gaussiano (g) [1, -2, 1] h es decir $g \otimes h$