

Cuestionario 1

David Gil Bautista

DNI: 45925424M

Email: davidbautista@correo.ugr.es

Universidad de Granada — October 26, 2018

- 1 Diga en una sola frase cuál cree que es el objetivo principal de la Visión por Computador. Diga también cuál es la principal propiedad de cara a los algoritmos que está presente en todas las imágenes.**

El objetivo de la Visión por Computador es que una máquina pueda llegar a analizar e interpretar una imagen.

Todos los algoritmos intentan mejorar (o cambiar) la imagen para que así sea más fácil trabajar con ella y extraer algún tipo de información.

- 2 Expresar las diferencias y semejanzas entre las operaciones de correlación y convolución. Dar una interpretación de cada una de ellas que en el contexto de uso en visión por computador.**

La principal diferencia entre correlación y convolución es la propiedad asociativa, cosa que se soluciona en el caso de que la máscara de correlación sea simétrica. En este caso las operaciones presentan las siguientes propiedades:

Superposition: El resultado de la suma de los filtros es igual a la suma de los filtros individuales.

$$R(x + y) = R(x) + R(y) \quad (1)$$

Scaling: El resultado de aplicar un filtro escalado es lo mismo que escalar el resultado

$$kR(x) = R(kx) \quad (2)$$

Shift invariance: El resultado de aplicar un filtro cambiado equivale a cambiar el resultado de aplicar ese filtro.

$$R(f(x)) = f(R(x)) \quad (3)$$

Ambas operaciones se definen como la aplicación de un filtro, en el caso de la correlación buscamos reconocer formas o patrones y en el de la convolución la posibilidad de combinar y operar con filtros.

3 ¿Los filtros de convolución definen funciones lineales sobre las imágenes? ¿y los de mediana? Justificar la respuesta.

Una función lineal es toda aplicación que cumple las siguientes propiedades: Sean V y W espacios vectoriales sobre el mismo cuerpo K . Una aplicación T de V en W es una transformación lineal si para todo par de vectores $u, v \in V$ y para todo escalar $k \in K$, se satisface que:

- $T(u + v) = T(u) + T(v)$
- $T(ku) = kT(u)$

Para comprobar si definen o no filtros lineales comprobaremos mediante un ejemplo si cumplen las propiedades.

Para demostrar si convolución presenta una aplicación lineal probaré con un filtro de alisamiento gaussiano. Siendo nuestra imagen de muestra:

$$I = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad (4)$$

Y u y v nuestros filtros gaussianos:

$$T(u) = [0.15 \quad 0.7 \quad 0.15] \quad (5)$$

$$T(v) = [0.1 \quad 0.8 \quad 0.1] \quad (6)$$

$$T(u + v) = [0.25 \quad 1.5 \quad 0.25] \quad (7)$$

Siendo $I2 = T(u + v)$ y $I3 = T(u) + T(v)$ comprobamos que sean iguales:

$$I2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * [0.25 \quad 1.5 \quad 0.25] = 3.5 \quad (8)$$

$$I3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * \begin{bmatrix} 0.15 & 0.7 & 0.15 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * \begin{bmatrix} 0.1 & 0.8 & 0.1 \end{bmatrix} = 1.7 + 1.8 = 3.5 \quad (9)$$

Comprobando $T(kv) = kT(v)$:

$$T(2v) = \begin{bmatrix} 0.2 & 1.6 & 0.2 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 3.6 = 2 * T(v) \quad (10)$$

Para el caso del filtro mediano probaremos con una matriz más grande para que sea más fácil de comprobar. Usando las siguientes matrices:

$$I = \begin{bmatrix} 1 & 1 & 2 & 15 & 34 \end{bmatrix} \quad (11)$$

$$J = \begin{bmatrix} 6 & 7 & 7 & 8 & 9 \end{bmatrix} \quad (12)$$

$$(I + J) = \begin{bmatrix} 7 & 8 & 9 & 23 & 43 \end{bmatrix} \quad (13)$$

Siendo $m(X)$ la mediana de una matriz X tenemos que:

$$m(I) = 2, \quad m(J) = 9 \quad (14)$$

$$m(I) + m(J) = 11 \quad (15)$$

$$m(I + J) = 9 \quad (16)$$

Sabiendo que un filtro de mediana no cumple la primera propiedad ya podemos determinar que no define una función lineal sobre una imagen.

4 Una operación de máscara que tipo de información usa, ¿local o global? Justificar la respuesta

Local. Cada vez que aplicamos la máscara estamos trabajando sobre una zona concreta de la imagen con unos píxeles determinados, tan sólo operamos con la información esos píxeles y al cambiar de zona trabajamos con otros distintos cuya información puede ser distinta.

- 5 ¿De qué depende que una máscara de convolución pueda ser implementada de forma separable por filas y columnas? Justificar la respuesta
- 6 Para implementar una función que calcule la imagen gradiente de una imagen cabe plantearse dos alternativas:
 - 6.1 Primero alisar la imagen y después calcular las derivadas sobre la imagen alisada
 - 6.2 Primero calcular las imágenes derivadas y después alisar dichas imágenes.

Discutir y decir cuál de las estrategias es la más adecuada, si alguna lo es, tanto en el plano teórico como en el de la implementación. Justificar la decisión.

Una imagen gradiente nos permite encontrar los saltos bruscos que hay en la imagen y de esta forma encontrar los bordes. En caso de tener una imagen con ruido, al hacer el gradiente vamos a resaltar el ruido y al alisar tan solo vamos a difuminar las manchas. En el caso de alisar primero quitamos todo el ruido haciendo una media de los valores que debería haber en esos puntos y al hacer el gradiente a esta imagen los errores que puedan llegar a percibirse van a ser menos intensos.

Para probar si esto es cierto o no lo comprobaremos con el siguiente ejemplo:



Figure 1: Original image

Para un filtro de desenfoque gaussiano con tamaño de kernel de 5×5 y un filtro de detección de bordes laplaciano con un tamaño de kernel de 5×5 hemos obtenido los resultados anteriores.

Aplicando la teoría, al alisar primero estamos quitando las imperfecciones, por lo que al detectar los bordes no deberíamos encontrar ruido. La realidad es que si la imagen es pequeña, al alisar estamos haciendo que el ruido ocupe más espacio, por lo que será más reconocible al detectar los bordes. En caso de hacerlo a la inversa, el ruido estará en píxeles determinados que estarán rodeados de un color plano el cual hará que al hacer la media ese píxel tome el valor de los contiguos y por tanto se eliminará casi al completo el ruido.



Figure 2: First smooth



Figure 3: then get sobels

- 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.**

Para demostrar esto tomaremos la expresión de una operación gaussiana y la descompondremos para probar que es separable.

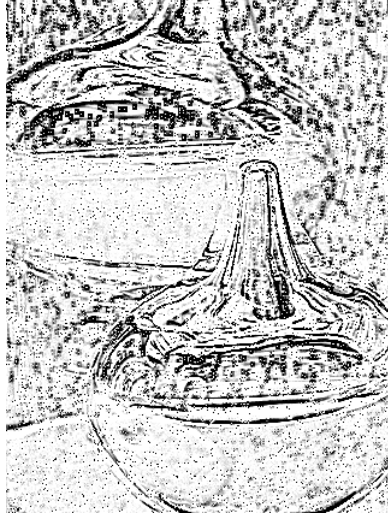


Figure 4: First get sobels

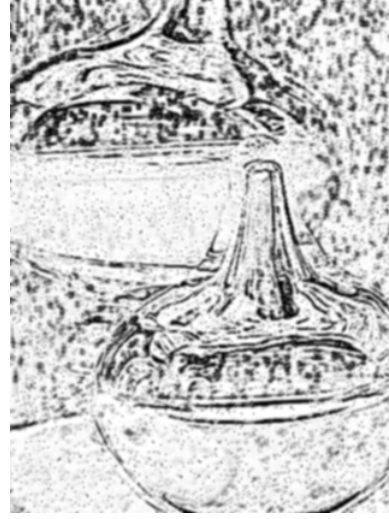


Figure 5: then smooth

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

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

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

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

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

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

$$G_{\sigma}(x, y) = \left(\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x^2}{2\sigma^2}} \right) \left(\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-y^2}{2\sigma^2}} \right)$$

$$G_{\sigma}(x, y) = G_{\sigma}(x)G_{\sigma}(y)$$

De esta forma podemos ver que podemos tratar la Gaussiana 2D como producto de la convolución de filas y columnas.

Haciendo las derivadas parciales de $G_\sigma(x, y)$ tenemos que:

$$\frac{\partial}{\partial x} G_\sigma(x, y) = -\frac{e^{-x^2/(2\sigma^2)-y^2/(2\sigma^2)} x}{2\pi\sigma^4} \quad (17)$$

$$\frac{\partial}{\partial y} G_\sigma(x, y) = -\frac{e^{-x^2/(2\sigma^2)-y^2/(2\sigma^2)} y}{2\pi\sigma^4} \quad (18)$$

Haciendo la primera derivada de $G_\sigma(x)$ y de $G_\sigma(y)$ obtenemos lo siguiente:

$$\frac{\partial}{\partial x} G_\sigma(x) = -\frac{e^{-\frac{x^2}{(2\sigma^2)}} x}{\sqrt{2\pi}\sigma^3} \quad (19)$$

$$\frac{\partial}{\partial y} G_\sigma(y) = -\frac{e^{-\frac{y^2}{(2\sigma^2)}} y}{\sqrt{2\pi}\sigma^3} \quad (20)$$

Al calcular las derivadas parciales podemos ver que la función que obtenemos es la misma para la derivada de x e y , pero cambia el término que multiplica a la exponencial en el numerador. Del mismo modo obtenemos lo mismo al derivar el filtro Gaussiano en función de x e y , en este caso cambiamos la x o la y en toda la función. Al representar las derivadas obtenemos la misma función con un eje cambiado.

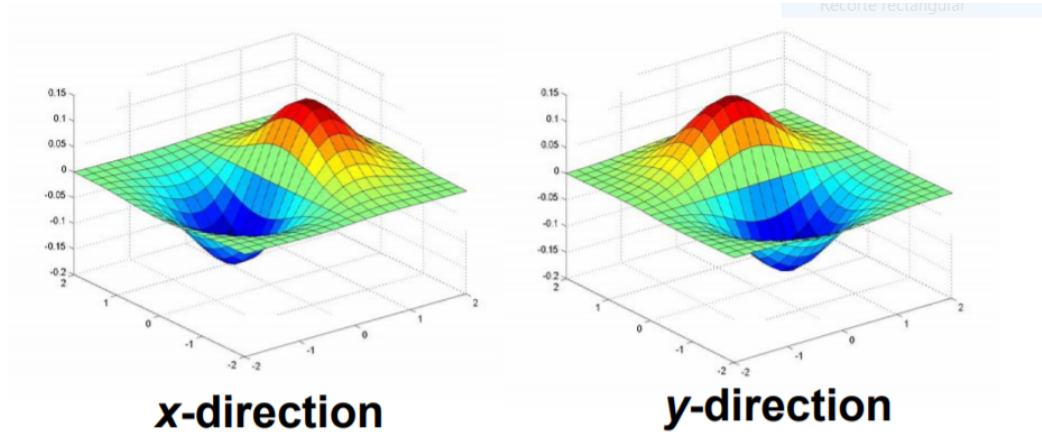


Figure 6: Gaussian derivatives

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.

La función Laplaciana de la Gaussiana se define como la suma de las segundas derivadas de la función Gaussiana. En el apartado anterior calculamos las derivadas parciales en x e y , por lo que volveremos a derivar estas y comprobarlas con la función Laplaciana para comprobar si son equivalentes. Siendo la función Laplaciana:

$$\nabla^2 G_\sigma(x, y) = \frac{\partial^2 G_\sigma(x, y)}{\partial x^2} + \frac{\partial^2 G_\sigma(x, y)}{\partial y^2} \quad (21)$$

Y las segundas derivadas parciales de la función Gaussiana:

$$\frac{\partial^2}{\partial x^2} G_\sigma(x, y) = \frac{(x^2 - \sigma^2)e^{-\frac{x^2+y^2}{2\sigma^2}}}{2\pi\sigma^6} \quad (22)$$

$$\frac{\partial^2}{\partial y^2} G_\sigma(x, y) = \frac{(y^2 - \sigma^2)e^{-\frac{x^2+y^2}{2\sigma^2}}}{2\pi\sigma^6} \quad (23)$$

Desarrollando podemos obtener lo siguiente:

$$\nabla^2 G_\sigma(x, y) = \frac{(x^2 - \sigma^2)e^{-\frac{x^2+y^2}{2\sigma^2}}}{2\pi\sigma^6} + \frac{(y^2 - \sigma^2)e^{-\frac{x^2+y^2}{2\sigma^2}}}{2\pi\sigma^6}$$

$$\nabla^2 G_\sigma(x, y) =$$

- 9 ¿Cuáles son las operaciones básicas en la reducción del tamaño de una imagen? Justificar el papel de cada una de ellas.**
- 10 - ¿Qué información de la imagen original se conserva cuando vamos subiendo niveles en una pirámide Gaussiana? Justificar la respuesta.**

Las bajas frecuencias, ya que al ir subiendo niveles en la pirámide gaussiana reescalamos la imagen de forma que en cada nivel tenemos menos información que en el anterior, y esto es lo mismo que eliminar las frecuencias altas.

- 11 ¿Qué información podemos extraer de la pirámide Gaussiana y la pirámide Laplaciana de una imagen? ¿Qué nos aporta cada una de ellas? Justificar la respuesta.**

La ambas pirámides nos aportan una imagen más pequeña que la original, la diferencia es que en el caso de la Laplaciana al reescalarla al tamaño original podemos recuperar las altas frecuencias que se pierden al reescalar la imagen. En el caso de la Gaussiana obtenemos tan solo las bajas frecuencias ya que no guardamos ninguna información y al reescalar al tamaño original tan sólo interpolamos los píxeles de la imagen pequeña.

- 12 ¿Podemos garantizar una perfecta reconstrucción de una imagen a partir de su pirámide Laplaciana? Dar argumentos y discutir las opciones que considere necesario.
- 13 En OpenCV solo se pueden calcular máscaras Sobel de hasta dimensión 7x7 ¿Por qué? De una explicación razonable a este hecho y diga cómo influye en el cálculo con máscaras de mayor tamaño. Justificar la respuesta
- 14 Cuales son las contribuciones más relevantes del algoritmo de Canny al cálculo de los contornos sobre una imagen?. ¿Existe alguna conexión entre las máscaras de Sobel y el algoritmo de Canny? Justificar la respuesta
- 15 Suponga que le piden implementar un algoritmo para el cálculo de la derivada de primer y segundo orden sobre una imagen usando un filtro gaussiano cualesquiera. Enumere y explique los pasos necesarios para llevarlo a cabo.
- 16 Algunas referencias
- Computer vision a modern approach
 - Separable convolutions
 - Herramientas usadas para ejercicio 6:
Edge detection
Blur images