VISIÓN POR COMPUTADOR:

CUESTIONARIO 1

Jesús Moyano Doña - 76626194-S
jesusmoyano97@correo.ugr.es
Grupo: 4ºA

Ejercicio 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 principal de la visión por computador es extraer información de los pixeles de una imagen. De cara a los algoritmos es que estos mismos pueden reconocer objetos, personas, lugares o actividades mediante la interpretanción de los pixeles de una imagen. Por lo que se puede crear un algoritmo con percepción e interpretación.

Ejercicio 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 convolución es asociativa y conmutativa.
Tanto convolución como correlación, son operaciones lineales, invariantes a la traslación. Ambas se escriben como una multiplicación de matriz $\star$ vector, cumplen el principio de superposición:

$$h(f0 + f1) = h(f0) + h(f1)$$

y la asociatividad con multiplicación escalar:

$$a(f * g) = (af) * g = f * (ag)$$

La convolución se usa para generar filtrados sobre las imágenes. En cuanto a la correlación, es usada para el template matching (reconocimiento de modelos)

Ejercicio 3:

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

Si, porque la convolución cumple el principio de superposición.

$$h(f0 + f1) = h(f0) + h(f1)$$

Otra forma de explicar porque la convolución define funciones lineales seria mediante la definición de función linear y sus propiedades:

1. Las propiedades de una función lineal son:
   -Propiedad aditiva: $f(x + y) = f(x) + f(y)$
   -Propiedad homogénea: $f(ax) = af(x)$
2. Por definición una convolución es una media móvil, entonces si comprobamos que la media es una función lineal por implica que la convolución lo será.

3. Tenemos dos conjuntos $X = \{x_1, x_2, ..., x_i\}$ e $Y = \{y_1, y_2, ..., y_i\}$, cuyas medias son:

$$Media(X) = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad Media(Y) = \frac{1}{n}\sum_{i=1}^{n} y_i$$

4. Vamos a comprobar si se cumple la propiedad aditiva:

$$Media(X+Y) = Media(X) + Media(Y) = \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n} x_i + y_i = Media(X+Y)$$

5. Ahora comprobamos la propiedad homogénea:

$$Media(X) = \frac{1}{n}\sum_{i=1}^{n} x_i * \lambda = \frac{1}{n}\lambda * \sum_{i=1}^{n} x_i$$

6. Como la función media cumple las dos propiedades se demuestra que es una función linear lo que implica que la convolución también lo sea.

La mediana no es función linear porque no cumple la propiedad aditiva: Si tenemos dos conjuntos de datos $X = \{-4, 5, -2\}$ e $Y = \{8, -5, 9\}$:
$Mediana(X) = -2 \qquad Mediana(Y) = 8 \qquad Mediana(X+Y) = M(4, 0, 7) = 4$

$Mediana(X) + Mediana(Y) = Mediana(X+Y) \rightarrow -2 + 8 = 4 \rightarrow 6 \neq 4$

Ejercicio 4:

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

Utiliza información de tipo local, porque estas accediendo a la información local de una imagen. Coges un píxel y sus vecinos. No puede ser de tipo global ya que no se podrían calcular los bordes.

Ejercicio 5:

¿De qué depende que una máscara de convolución pueda ser implementada de forma separable por filas y columnas? Justificar la respuesta

Lo primero la mascara a debe ser separable en un vector fila y un vector columna. Para saber si una mascara es separable debemos expresarla, ya que es una matriz, en forma de descomposición de valores singulares, entonces si el primer valor de la DSV es distinto de 0 mascara es separable.

Ejercicio 6:

Para implementar una función que calcule la imagen gradiente de una imagen cabe 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 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.

La estrategia más adecuada sería la a) porque:
Desde el punto de vista teórico, el gradiente genera 2 imágenes (Gradiente en x, gradiente en Y), por lo tanto si suavizamos primero, haremos el gradiente sobre una imagen unicamente. En caso caso de calcular el gradiente de la imagen primero, tendremos que suavizar las 2 imágenes generadas por el gradiente. Tras esta comparación es más eficiente alisar primero y luego calcular las derivadas.

Desde el puto de vista practico: si primero alisamos eliminamos el ruido por lo tanto cuando apliques las derivadas la imagen saldrá sin ruido. Sin embargo, al derivar sin alisar, lo que hacemos es intensificar el ruido, por lo que el alisamiento de después no hará grandes cambios

Ejercicio 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.

Vamos a derivar la formula del filtro gaussiano para demostrarlo:

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

Calculamos la derivada:

Comenzamos por la derivada en x:

$$G'_x(x,y) = \frac{d}{dx} \cdot \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2} \frac{d}{dx} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} \frac{d}{dx}\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

$$= \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \frac{d}{dx}\left(\frac{-x^2}{2\sigma^2} + \underbrace{\cancel{\frac{-y^2}{2\sigma^2}}}_{0}\right) = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} \left(\frac{-x}{\sigma^2}\right)$$

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

Ahora hacemos la derivada en y:

$$G'_y(x,y) = \frac{d}{dy} \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2} \frac{d}{dy} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} \frac{d}{dy}\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

$$= \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \frac{d}{dy}\left(-\underbrace{\cancel{\frac{x^2}{2\sigma^2}}}_{0} + \frac{-y^2}{2\sigma^2}\right) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \left(\frac{-y}{\sigma^2}\right)$$

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

Tras obtener ambas derivadas las separamos en dos factores que uno depende de x y otro de y.

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

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

Cada una de las derivadas nos dará una componente del gradiente asociado a la derivada. Por lo tanto vemos que se puede separa en dos núcleos, uno para las filas y otro para las columnas. Si este filtro lo aplicamos a una imagen vemos como se suavizan los bordes horizontales o los verticales, dependiendo e la derivada aplicada.

Ejercicio 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 laplaciana de la gaussiana es la suma de las segundas derivadas, por lo que derivamos las primeras derivadas obtenidas en el ejercicio anterior

Derivada en X:

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

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

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

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

Separamos las derivadas en dos factores uno dependiente de x y dependiente de y

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

$$= \underbrace{\left[\frac{-1}{2\pi\sigma^4}\left(\frac{-x^2}{\sigma^2}+1\right)\left(e^{-x^2/2\sigma^2}\right)\right]}_{H_1(x)}\underbrace{\left[\left(e^{-y^2/2\sigma^2}\right)\right]}_{H_2(y)}$$

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

$$= \underbrace{\left[\frac{-1}{2\pi\sigma^4}\left(\frac{-y^2}{\sigma^2}+1\right)\left(e^{-y^2/2\sigma^2}\right)\right]}_{H_3(y)}\underbrace{\left[\left(e^{-x^2/2\sigma^2}\right)\right]}_{H_4(x)}$$

Como se observa la Laplaciana de Gaussiana es $H1(x) * H2(y) + H3(y) * H4(x)$

Ejercicio 9:

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

Para reducir el tamaño de una imagen, dos operaciones que tenemos que hacer son:

1. Suavizado: Aplicamos un filtro Gaussiano para suavizar la imagen.

2. Reducción: Suprimimos las ultimas filas y columnas.

Este método funciona de las siguiente forma: Suavizamos primero para que la información de la imagen y los pixeles vecinos se distribuya, para así cuando reduzcamos la imagen, la perdida de información dada por la supresión de filas y columnas no sea tan mayor ya que la información de esos pixeles ha sido distribuida a sus vecinos.

Ejercicio 10:

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

En el ejercicio anterior hemos explicado las operaciones básicas sobre la reducción de una imagen, que es lo mismo que subir niveles en una piramide gaussiana. Por lo tanto teniendo en cuenta lo explicado anteriormente, al ir realizando un filtro gaussiano a medida que reducimos la imagen lo que vamos obteniendo son las frecuencias bajas y a medida que vamos atenuando las altas.

Como ejemplo: en las imágenes híbridas, cuando mezclamos las frecuencias bajas y altas, a medida que se va reduciendo la imagen, las altas van desapareciendo y solo quedan las bajas.

Ejercicio 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.

Aclaración sobre como he entendido la pregunta: Explicar que información extraemos de la piramide Gaussiana por una parte y por otra parte la Laplaciana.

Para comenzar la piramide Gaussiana, ya explicada en el ejercicio 9 y 10, a medida que disminuimos la imagen se conservan las frecuencias bajas mientras que las altas de van perdiendo.
En el caso contrario, la Laplaciana, construida a partir de la piramide gaussiana y a continuación un ampliado progresivo de la imagen. Se va comparado cada imagen ampliada con la original, quedándonos con las diferencias entre ambas, el resultado de esas diferencias da la piramide Laplaciana, siendo esta una representación de las frecuencias altas.

Para resumir, la piramide Gaussiana es una representación de las frecuencias bajas y la Piramide Laplaciana representa las frecuencias altas.

Ejercicio 12:

¿Podemos garantizar una perfecta reconstrucción de una imagen a partir de su pirámide Laplaciana? Dar argumentos y discutir las opciones que considere necesario.

Si, sin perdidas. La manera de reconstruir una imagen a partir de la piramide laplaciana consiste en primero reducimos la imagen mediante la piramide Gaussiana, a continuación volvemos a aumentarla. Cuando tengamos la imagen en la resolución inicial, las restamos con su original, obteniendo así una imagen que será una aproximación de la Laplaciana, la cual si la sumamos con el nivel mas bajo de la piramide Gaussiana nos dará la imagen original sin perdidas.

Ejercicio 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

Ejercicio 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 Las dos propiedades más relevantes del algoritmod de Canny son:

1. La supresión falsos máximos: Esta técnica se usa para afinar los bordes encontrados en las fases anteriores del algoritmo. Consiste en eliminar los pixeles de la imagen que no formen parte de los bordes.

2. Umbral de histéresis: Tras hacer la supresión de máximos, puede ser que quedan algunos pixeles provenientes del ruido. Para llevar acabo este paso se fijan dos valores de umbral $minV$ y $maxV$. Los pixeles con un gradiente de intensidad mayor que el máximo son aceptados como pertenecientes a bordes, mientras los inferiores al mínimo los descartamos. Los pixeles que se encuentren el el medio del umbral son considerados como débiles y serán aceptados dependiendo de su colectividad.

En el paso en el que se busca el gradiente de intensidad de la imagen, una vez aplicado el filtro Gaussiano, calculamos el gradiente. Para eso se filtra la imagen de nuevo, pero usando un kernel Sobel tanto en la dirección vertical como horizontal.

Ejercicio 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.

Este ejercicio es parecido al 2A de la practica 1, el cual nos pedían implementar una función para aplicar la derivada sobre un filtro gaussiano.
Los pasos son los siguientes:

1. Obtenemos los Kernels de la mascara derivada, especificamos si queremos la 1ª o 2ª derivada.

2. Aplicamos el filtro Gaussiano a la imagen para alisarla y eliminar el ruido.

3. Aplicamos la mascara derivada sobre la imagen alisada.

Bibliografía

- Computer Vision: Algorithms and Applications. Richard Szeliski
- Wikipedia:
  - Principio de Superposición
  - Convolución
  - Función Lineal
- Alg Canny