

Visión por computador (2016-2017)
GRADO EN INGENIERÍA INFORMÁTICA
UNIVERSIDAD DE GRANADA

Cuestionario 1

Ignacio Martín Requena

28 de octubre de 2016

Índice

1	¿Cuáles son los objetivos principales de las técnicas de visión por computador? Poner algún ejemplo si lo necesita	4
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.	4
3	Expresar y justificar las diferencias y semejanzas entre correlación y convolución. Justificar la respuesta.	5
4	¿Los filtros de convolución definen funciones lineales sobre las imágenes? ¿y los de mediana? Justificar la respuesta.	5
5	¿La aplicación de un filtro de alisamiento debe ser una operación local o global sobre la imagen? Justificar la respuesta	6
6	Para implementar una función que calcule la imagen gradiente de una imagen dada pueden 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 que estrategia es la más adecuada, si alguna lo es, tanto en el plano teórico como de implementación. Justificar la decisión.	6
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.	7
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.	7
9	¿Cuáles son las operaciones básicas en la reducción del tamaño de una imagen? Justificar el papel de cada una de ellas.	7
10	¿Qué información de la imagen original se conserva cuando vamos subiendo niveles en una pirámide Gaussiana? Justificar la respuesta.	7
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)	8

- 12 Cual es la aportación del filtro de Canny al cálculo de fronteras frente a filtros como Sobel o Robert. Justificar detalladamente la respuesta. 8
- 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. 8
- 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. 9

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 los de la resolución de problemas que a partir de los píxeles de una imagen sacan información de esta para poder percibir e interpretar objetos, personas. . . Además, es posible obtener formas geométricas mediante el cálculo de propiedades del mundo.

Por tanto, podríamos saber si una foto es de un paisaje mediante el reconocimiento de árboles, montañas. . . o incluso ver estudiando la geometría de la imagen si se ha tomado con alguna perspectiva.

Es decir, el objetivo de la visión por computador es el de interpretar imágenes.

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 tiene por qué ser siempre una matriz 2D, puede ser por ejemplo un vector 1D deducido a partir de una matriz 2D que es separable en dos vectores de 1D (como ocurre con el caso de una máscara gaussiana o si queremos pasar un filtro por filas o por columnas solamente). De hecho, si la matriz tiene la propiedad de ser separable, computacionalmente es más eficiente hacer la convolución de dos vectores (uno por filas y otro por columnas) que hacer la de la máscara 2D.

Tiene sentido definir máscaras a partir de matrices de varios canales siempre y cuando la correlación o convolución estén bien definidas, esto es, que cada canal de la máscara se utilice para hacer la correlación o convolución de cada canal de la imagen original.

3. Expresar y justificar las diferencias y semejanzas entre correlación y convolución. Justificar la respuesta.

Las **semejanzas** entre la correlación y la convolución es que ambas son operaciones lineales que cumplen el principio de superposición. Además ambas se comportan igual en toda la imagen, es decir, que el valor de salida depende del valor de los pixels de alrededor, no de dónde estén posicionados dichos pixels. Otra similitud es que ambos operadores pasan una máscara a través de una imagen, es decir, el valor devuelto en ambas es una combinación lineal de los valores vecinos de cada pixel obtenidos tras aplicar una máscara.

Por otra parte, las **diferencias** entre ambas técnicas radican en cómo se trata la máscara para pasarla por la imagen. En el caso de la correlación la máscara se pasa tal y como está por la imagen, mientras que en la convolución a la máscara se le da la vuelta de derecha a izquierda y de arriba a abajo antes de pasarla por la imagen. Por esto, en realidad la convolución es darle la vuelta a la máscara para después aplicar una correlación. Debido a este volteo, la correlación y la convolución devuelven los mismos valores si la máscara es gaussiana (ya que es simétrica con respecto a las dos diagonales principales). Otra de las diferencias entre ambas técnicas es a la hora de aplicarlas, mientras que la convolución se suele usar para oscurecer imágenes o elimina ruido, la correlación se suele usar para medir la similitud de patrones o detectar objetos en una imagen.

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

Si, ya que correlación y convolución son operaciones de desplazamiento lineal. La función de convolución se comporta de igual manera para cada pixel de la imagen, transformando la imagen ligeramente, como por ejemplo en un alisado. Dicho de otra forma, la convolución es lineal ya que se basa en productos escalares y el producto escalar es una operación lineal

Los filtros de mediana no definen funciones lineales, ya que la mediana de la suma de dos conjuntos no es la misma que la suma de la mediana de los dos mismos conjuntos.

5. ¿La aplicación de un filtro de alisamiento debe ser una operación local o global sobre la imagen? Justificar la respuesta

Local, los filtros de alisamiento se aplican a un conjunto de vecinos de un pixel. Ya que lo que se pretende es alisar la imagen, las operaciones hay que realizarlas de forma local porque dichos píxeles tendrán valores aproximados entre ellos, y se calcula una media.

6. Para implementar una función que calcule la imagen gradiente de una imagen dada pueden 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 que estrategia es la más adecuada, si alguna lo es, tanto en el plano teórico como de implementación. Justificar la decisión.

La opción de b) deberíamos descartarla por los efectos del ruido que hay en cualquier imagen, al calcular las derivadas acentuamos las altas frecuencias de la imagen y por tanto amplificamos el ruido, así que al derivar no obtendremos una señal limpia, lo que obtenemos será una señal con muchos picos que no estaban en la imagen y encima perdemos los que sí habían al mezclarlos con los que hemos añadido, así que aunque alisemos las derivadas después no vamos a obtener una salida correcta. Por tanto, con la opción b) obtendremos una imagen con mucho más ruido que la inicial, pero no la imagen gradiente.

La forma de solucionar esto es realizando la opción a), es decir, primero alisar la imagen original y eliminar el ruido impidiendo que se propague y después calcular las derivadas sobre la imagen alisada.

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

Las operaciones básicas en la reducción son el alisado y el submuestreo. El submuestreo consiste en seleccionar un cierto número de píxeles para reducir el tamaño de la imagen. Esto tiene como inconveniente que los píxeles seleccionados pueden ser representativos, perdiendo más información de la que deberíamos. Por tanto, alisando la imagen hacemos que cada píxel contenga también información de sus vecinos, con lo que mitigamos la información perdida en el submuestreo.

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

Lo que conservamos conforme vamos subiendo niveles en una pirámide Gaussiana serán las frecuencias bajas, ya que las altas se han ido perdiendo a cada nivel que ha pasado con la convolución. Esto pasa debido a que si vamos viendo algo cada vez más pequeño, lo que veremos tendrá que ser las formas más generales (bajas frecuencias) y no los pequeños detalles (altas frecuencias).

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 una imagen construida a partir de una serie de niveles en la que cada nivel es el resultado de aplicar un filtro de alisamiento gaussiano y un submuestreo de la imagen del nivel anterior, obteniendo en cada iteración una imagen de menor tamaño con una mínima pérdida de información. Por el contrario, cada elemento de la **pirámide laplaciana** almacena la diferencia entre cada dos niveles de la pirámide gaussiana, es decir, el proceso de construcción es el de aumentar la que tenemos en la gaussiana, hacer la diferencia y ese error es lo que almacenamos en la pirámide laplaciana. En decir, la pirámide laplaciana se compone del error que se comete en cada nivel de la pirámide gaussiana.

12. Cual es la aportación del filtro de Canny al cálculo de fronteras frente a filtros como Sobel o Robert. Justificar detalladamente la respuesta.

La principal aportación del filtro de Canny frente a los otros filtros es la implementación de dos umbrales distintos, uno para obtener los bordes más acentuados y otro para unir estos bordes. Además, se utiliza primero un filtro para difuminar la imagen y así reducir el ruido presente en la misma. Después utiliza el gradiente para encontrar las regiones con mayor derivada y suprimir cualquier pixel que no esté en cada máximo local, haciendo más finos los bordes encontrados. Por último, busca más bordes con los dos umbrales antes mencionados, utilizando el umbral alto para empezar los bordes y el bajo para unirlos.

Esto es distinto a los filtros como los de Sobel o Robert, que sólo utilizan el grandiente sobre la imagen original y se quedan con los pixeles donde el cambio de la luminosidad es brusco.

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 real del filtro de Canny podría ser la de a partir de una imagen de un paisaje de una montaña o un terreno con cierto relieve extraer el perfil de la forma de

la montaña y compararlo en tiempo real con los mismos perfiles tomados anteriormente para, por ejemplo, la detección de los cambios en el relieve debido a la erosión o el deslizamiento de tierra.¹

- 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.**

¹http://www.cirgeo.unipd.it/fp/francescopirotti_files/papers/pirotti_canny_grass2003_eng.pdf