

Ejercicios Tema 2

Percepción

Curso 2019/2020

1. Se desea representar dígitos manuscritos mediante representación geométrica. Los dígitos se almacenan en imágenes de 32×32 píxeles. Las imágenes son en escala de gris de 256 niveles. Calcula para los siguientes supuestos el número de bytes necesarios para almacenar la representación de un dígito:
 - a) Si se usa representación directa global
 - b) Si la representación es la cuenta de cuántas veces aparece cada uno de los niveles de gris en la imagen (histograma)
 - c) Si la representación es con características locales (CL) de 15×15 píxeles en el caso extremo que se extraigan todas las CL posibles (rejilla de tamaño 1)
 - d) Si la representación es la misma que la anterior salvo que de cada una de las características locales se extrae un histograma de los niveles de gris presentes en la misma
2. Para un problema de clasificación de imágenes, se tienen imágenes de 20×20 píxeles en 256 niveles de grises. Se pide el tamaño en bytes de una imagen para cada una de las siguientes representaciones:
 - a) Representación global
 - b) Representación por histograma
 - c) Representación local con ventanas de 5×5 , en donde se muestrea uno de cada dos píxeles en dirección horizontal
3. Calcula espacio requerido en memoria de las siguientes representaciones de imágenes:
 - a) Representación directa global de una imagen de 1024×1080 píxeles en escala de grises a 65536 niveles
 - b) Representación mediante histograma de una imagen RGB de 512×512 píxeles en escala de 256 niveles por color
 - c) Representación de una imagen de 100×100 píxeles en escala de grises de 256 niveles mediante características locales de 25×25 píxeles extraídas una cada 5 píxeles
 - d) Representación de una imagen de 80×80 píxeles en escala de grises de 512 niveles mediante características locales de 21×21 píxeles extraídas una cada 10 píxeles y representada cada una de ellas mediante histograma
4. Calcula el espacio requerido en memoria para las siguientes representaciones de imágenes y audio:
 - a) Imagen de 1024×768 píxeles en color RGB, 256 niveles por color, por representación directa global
 - b) Imagen de 1024×768 píxeles en escala de grises, 256 niveles, por representación por histograma
 - c) Imagen de 1024×768 píxeles en escala de grises, 256 niveles, por representación local, con ventanas de 33×33 píxeles extraídas cada 16 píxeles tanto en horizontal como en vertical
 - d) Señal de audio mono de 10 segundos de duración, con frecuencia máxima a reproducir de 8 KHz, 2 bytes por muestra
 - e) Señal de audio 5.1 de 1 minuto de duración, con frecuencia de muestreo de 44 KHz, 32 bits por muestra
5. Calcula espacio requerido en memoria de las siguientes representaciones:
 - a) Representación de una imagen de 27×50 píxeles en escala de grises de 256 niveles mediante características locales de 7×7 píxeles extraídas una cada 3 píxeles en horizontal y una cada 4 píxeles en vertical
 - b) Representación de una imagen de 60×60 píxeles en escala de grises de 256 niveles mediante características locales de 11×11 píxeles extraídas una cada 5 píxeles en ambas dimensiones y representada cada una de ellas mediante histograma

- c) Representación de la señal de audio en estéreo de un vídeo de Polimedia grabado a 16KHz que tiene un duración de 10 minutos donde cada muestra se representa mediante 16 bits
- d) Representación de la señal de audio del apartado anterior habiendo sido procesada para mono (1 canal) y almacenada mediante sus frames con un tamaño $W = 25$ ms y desplazamiento $S = 10$ ms. Considera que cada componente del frame se almacena mediante un *float* de 4 bytes
6. Calcular el tamaño en bytes que ocuparía cada una de estas señales acústicas, adquiridas en las condiciones indicadas:
- a) 3 minutos de una señal telefónica, adquirida a 8 KHz con 8 bits por muestra
- b) 10 segundos de señal vocal captada por micrófono, adquirida a 16 KHz con 16 bits por muestra
- c) 1 minuto de una señal de alta fidelidad para un sistema de audio 5.1, adquirida a 44 KHz con 16 bits por muestra
7. Dado el siguiente *codebook* de dos dimensiones, obtener la cadena asociada a la secuencia de puntos dada empleando distancia euclídea
- Codebook*: $a \rightarrow (0,0)$ $b \rightarrow (-1,1)$ $c \rightarrow (0,2)$ $d \rightarrow (2,-1)$
- Secuencia: (0,0.9), (0,1.3), (-1,1.3), (-1.5,0.5), (0,1.5), (1,0), (2,-0.5)
8. Se tiene un conjunto de muestras de audio procesadas para obtener una representación bidimensional de cada parte de la secuencia. Tras aplicar a esas muestras el algoritmo *c-medias* para 3 particiones, se han obtenido los siguientes agrupamientos:

| <i>a</i> | <i>b</i> | <i>c</i> |
|---------------|---------------|---------------|
| (0.00, 0.50) | (-1.00, 3.00) | (4.75, -1.25) |
| (1.00, 0.00) | (-2.00, 2.50) | (2.50, -0.75) |
| (-0.50,-0.50) | (-1.00, 3.50) | (3.25, -1.00) |
| (0.00, 0.00) | (-2.00, 3.00) | (5.00, -1.50) |
| (-1.00, 0.25) | | (4.50, -0.50) |
| (0.50,-0.25) | | |

Se pide:

- a) Obtener el *codebook* con las etiquetas *a*, *b* y *c*
- b) Calcular la codificación en cadena de dichas etiquetas para la secuencia: (0,00, 0,15), (-0,50, -0,20), (3,30, -0,50), (4,15, -0,65), (-0,75, 1,50), (-1,25, 1,65), (2,75, 0,25), (0,10, -0,70)
9. Sea una colección de documentos:
- comp.graphics:** i have a trident card and fullview works real gif jpg image
- comp.graphics:** hello , subject says it all . i need a gif to targa converter
- comp.graphics:** my scanner can do the job for you . gif , tiff , bmp . interested ?
- rec.autos:** i guess that makes altima the most generic car in the us .
- rec.autos:** the car might also need a front end alignment
- rec.autos:** is it ok to take the car out of gear without using the clutch
- sci.space:** your description sounds like one of the gravity probe spacecraft ideas .
- sci.space:** did the russian spacecraft send back any image of the martian moon ?
- sci.space:** where i might find a stereo image of planetary and satellite surfaces ?
- politics.guns** though i agree this is not the place to discuss guns
- politics.guns** i just heard on cnn that the texas rangers found m60 machine guns
- politics.guns** sniper injures 9 outside mca buildling in l.a. man arrested-suspect

Calcula *term frequency* por documento y las funciones globales *Normal*, *GfIdf* y *Idf* para los tokens *gif*, *the*, *of* y *image*.

10. Dada la siguiente matriz de *term frequency* que representa la frecuencia de aparición de cada token en cada documento

| x_{dt} | token 1 | token 2 | token 3 | token 4 | token 5 | token 6 |
|----------|---------|---------|---------|---------|---------|---------|
| doc 1 | 10 | 1 | 10 | 1 | 0 | 0 |
| doc 2 | 10 | 1 | 0 | 0 | 0 | 0 |
| doc 3 | 10 | 1 | 10 | 1 | 10 | 1 |
| doc 4 | 10 | 1 | 0 | 0 | 0 | 0 |
| doc 5 | 10 | 1 | 10 | 1 | 0 | 0 |

se pide:

- Calcular la función global *Normal* para todos los tokens
- Calcular la función global *GfIdf* para todos los tokens
- Calcular la función global *Idf* para todos los tokens utilizando logaritmo neperiano
- En este caso, ¿cuál es la principal diferencia que observas entre la función global *Normal* y *GfIdf*?, ¿y entre *GfIdf* e *Idf*?

11. Tenemos los siguientes textos para 3 clases de documentos:

| Número | Clase | Texto |
|--------|---------------|--|
| 1 | Internacional | May gana sin mayoría absoluta y Corbyn pide su dimisión |
| 2 | Economía | El FROB vigiló al Popular durante cuatro días antes de la intervención |
| 3 | Política | El Constitucional falla que la amnistía de Montoro avaló el fraude fiscal |
| 4 | Internacional | Theresa May fracasa en su apuesta y pierde la mayoría absoluta |
| 5 | Economía | Los ganadores y perdedores en la quiebra de Popular |
| 6 | Política | El varapalo del Constitucional al Gobierno de Rajoy cuestiona futuras amnistías fiscales |
| 7 | Internacional | May pierde la mayoría absoluta y Corbyn pide su dimisión |
| 8 | Economía | El BCE rebajó las exigencias al Banco Popular hace sólo seis meses |
| 9 | Política | ¿Qué consecuencias tiene el fallo que declara nula la amnistía de Montoro? |

Se pide (en todos los casos, considera indistinto mayúsculas y minúsculas):

- Realizar la representación *bag-of-words* con las siguientes palabras: May, mayoría, absoluta, dimisión, Popular, intervención, quiebra, Constitucional, amnistía, Rajoy, Montoro, fraude
 - Realizar la representación *bag-of-bigrams* con las siguientes secuencias de palabras: “Theresa May”, “mayoría absoluta”, “su dimisión”, “Banco Popular”, “la quiebra”, “la intervención”, “fraude fiscal”, “amnistías fiscales”, “el fallo”
 - Calcular el valor de las funciones globales *GfIdf* y *Idf* en el conjunto de documentos para los términos: y, mayoría, dimisión, de, la, intervención, el, fraude, en, quiebra
12. Asume que se dispone de una colección de $D > 0$ documentos con un token t_1 que ocurre con una frecuencia constante $k > 0$ en todos los documentos, y un token t_2 que ocurre únicamente en uno de los documentos con frecuencia $k > 0$. Se pide:
- Calcular las funciones globales *Normal*, *GfIdf* y *Idf* para el token t_1 y discutir qué función global le asignaría un menor valor.
 - Calcular las funciones globales *Normal*, *GfIdf* y *Idf* para el token t_2 y discutir qué función global le asignaría un menor valor.

Soluciones

- 1024, 512, 72900, 82944
- (Examen Marzo 2013) 400, 512, 3200
- (Examen Abril 2017) 2211840, 2304, 721875, 368640
- (Examen Recuperación Junio 2017) 2359296, 768, 3105828, 320000, 63360000
- (Examen Marzo 2018) 3773 bytes, 25600 bytes, 36.62 Mbytes, 91.55 Mbytes
- (Examen Recuperación Junio 2013) 1440000, 320000, 31680000
- (Examen Marzo 2013) acbbcad
- a (0,0), b (-1.5, 3), c (4, -1); cadena: $aacc(a|b)bca$
- Normal: $\text{gif} = 3^{-\frac{1}{2}}$, $\text{the} = 17^{-\frac{1}{2}}$, $\text{of} = 4^{-\frac{1}{2}}$, $\text{image} = 3^{-\frac{1}{2}}$; GfIdf: $\text{gif} = 1$, $\text{the} = \frac{11}{8}$, $\text{of} = 1$, $\text{image} = 1$; Idf: $\text{gif} = \log 4$, $\text{the} = \log \frac{3}{2}$, $\text{of} = \log 3$, $\text{image} = \log 4$
- (Examen Abril 2017)

| G (t) | token 1 | token 2 | token 3 | token 4 | token 5 | token 6 |
|----------|---------|---------|---------|---------|---------|---------|
| Normal | 0.04 | 0.4 | 0.06 | 0.6 | 0.1 | 1 |
| GfIdf | 10 | 1 | 10 | 1 | 10 | 1 |
| Idf (ln) | 0 | 0 | 0.51 | 0.51 | 1.63 | 1.63 |

La principal diferencia entre la función global Normal y GfIdf es que los tokens con frecuencias altas tienen un menor peso en la función Normal que en la GfIdf. La principal diferencia entre GfIdf e Idf es que la función Idf asigna el mismo peso a tokens con diferentes frecuencias (ya que solo considera la ocurrencia o no del token en un documento), mientras que GfIdf tiene en cuenta la frecuencia del token y le asigna un mayor peso.

- (Examen Recuperación Junio 2017)

a)

| t/d | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------------|---|---|---|---|---|---|---|---|---|
| May | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| mayoría | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| absoluta | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| dimisión | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Popular | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| intervención | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| quiebra | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Constitucional | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| amnistía | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| Rajoy | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| Montoro | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| fraude | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

b)

| t/d | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------------------|---|---|---|---|---|---|---|---|---|
| "Theresa May" | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| "mayoría absoluta" | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| "su dimisión" | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| "Banco Popular" | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| "la quiebra" | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| "la intervención" | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| "fraude fiscal" | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| "amnistías fiscales" | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| "el fallo" | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

| | | | |
|----|--------------|---------------|--------------------|
| | Término | GfIdf | Idf |
| | y | 1 | $\log \frac{9}{4}$ |
| | mayoría | 1 | $\log 3$ |
| | dimisión | 1 | $\log \frac{9}{2}$ |
| | de | 1 | $\log \frac{9}{5}$ |
| c) | la | 1 | $\log \frac{3}{2}$ |
| | intervención | 1 | $\log 9$ |
| | el | $\frac{6}{5}$ | $\log \frac{9}{5}$ |
| | fraude | 1 | $\log 9$ |
| | en | 1 | $\log \frac{9}{2}$ |
| | quiebra | 1 | $\log 9$ |

12. (Examen Marzo 2018)

a) Normal: $G(t_1) = (D \cdot k^2)^{-\frac{1}{2}}$, GfIdf: $G(t_1) = k$, Idf: $G(t_1) = 0$.
Idf es la de menor valor.

b) Normal: $G(t_2) = \frac{1}{k}$, GfIdf: $G(t_2) = k$, Idf: $G(t_2) = \log D$

En este caso, la función global de menor peso dependerá del valor de k y D . Asumiendo una colección de cientos de documentos, la función global *Normal* será menor o igual ($k = 1$) que las otras dos funciones.