

# Test Tema 2 de Percepción

ETSINF, Universitat Politècnica de València, Febrero de 2017

Apellidos:  Nombre:

Profesor: ☒ Jorge Civera ☐ Carlos Martínez

## Cuestiones (0.25 puntos, 15 minutos, con apuntes)

☐ A ¿Cuál de las siguientes representaciones requiere un mayor espacio en memoria?

- A) Representación directa global de una imagen RGB de  $48 \times 48$  píxeles a color real (24 bits por color básico)  $48 \cdot 48 \cdot 3 \cdot 3 = 20736$  bytes RGB es  $3 \cdot 3$
- B) Representación mediante histograma de una imagen de  $512 \times 512$  en escala de grises de 1024 niveles  $3 \text{ bytes por nivel} \cdot 1024 \text{ niveles} = 3072$  bytes
- C) Representación de una imagen de  $20 \times 20$  píxeles en escala de grises de 512 niveles mediante características locales de  $15 \times 15$  píxeles extraídas una cada dos píxeles  $(6 \cdot 6) / 2 \cdot 15 \cdot 15 \cdot 2 = 8100$  bytes res =  $n \cdot s$ ,  $n = (6 \cdot 6) / 2$ ,  $s = 15 \cdot 15 \cdot \log_2(\text{niveles}) / 8$
- D) Representación de una imagen de  $25 \times 25$  píxeles en escala de grises de 256 niveles mediante características locales de  $17 \times 17$  píxeles extraídas una cada tres píxeles y representada cada una de ellas mediante histograma  $(9 \cdot 9) / 3 \cdot 2 \cdot 256 = 13824$  bytes  $n \cdot \log_2(n+1) / 8 \cdot l$  (niveles)

☐ A La representación de voz presenta el problema de la variabilidad temporal no lineal (elasticidad de los sonidos). Una posible solución para reducir el impacto de este problema en la clasificación es:

- A) Emplear modelos de clasificación, como modelos ocultos de Markov, que den cuenta de esta variabilidad
- B) Usar técnicas de paso de dominio temporal a frecuencial, como la transformada de Fourier
- C) Realizar una cuantificación vectorial de la representación vectorial de la señal de voz
- D) Recurrir a la representación temporal de la señal de voz

☐ C Indica qué característica **no** presenta la representación *bag of words* de documentos de texto:

- A) Sus vectores de características son binarios o de números naturales
- B) La dimensión de sus vectores de características es igual al tamaño del vocabulario a representar
- C) Cada documento se representa por varios vectores de características
- D) Suelen usarse funciones de ponderación local y global en la representación final

# Test Tema 2 de Percepción

ETSINF, Universitat Politècnica de València, Febrero de 2017

Apellidos:

Nombre:

Profesor: ☐ Jorge Civera ☒ Carlos Martínez

Cuestiones (0.25 puntos, 15 minutos, con apuntes)

☐ D ¿Cuál de las siguientes representaciones requiere un mayor espacio en memoria?

- A) Representación directa global de una imagen RGB de  $32 \times 32$  píxeles a color real (24 bits por color básico)  
 $32 \cdot 32 \cdot 3 \cdot 3 = 9216$  bytes
- B) Representación mediante histograma de una imagen de  $8192 \times 8192$  en escala de grises de 2048 niveles 4 bytes por nivel  $\cdot 2048$  niveles = 8192 bytes  $\rightarrow 1 \cdot \log_2(n+1)/8 = 2048 \cdot \log_2(8192 \cdot 8192 + 1)/8 = 2048 \cdot 4$
- C) Representación de una imagen de  $24 \times 24$  píxeles en escala de grises de 256 niveles mediante características locales de  $15 \times 15$  píxeles extraídas una cada cuatro píxeles  $(10 \cdot 10)/4 \cdot 15 \cdot 15 \cdot 1 = 5625$  bytes  $\rightarrow$  repr. directa
- D) Representación de una imagen de  $20 \times 20$  píxeles en escala de grises de 256 niveles mediante características locales de  $13 \times 13$  píxeles extraídas una cada píxel y representada cada una de ellas mediante histograma  $8 \cdot 8 \cdot 256 = 16384$  bytes  
 $n = (20 - 13 + 1) \cdot (20 - 13 + 1) = 8 \cdot 8 = 64 \rightarrow n \cdot s = 64 \cdot 1 \cdot \log_2(n+1)/8 = 64 \cdot 256 \cdot 0.75$  (redon. 1) =  $64 \cdot 256$

☐ D La representación de voz presenta el problema de la variabilidad temporal no lineal (elasticidad de los sonidos). Una posible solución para reducir el impacto de este problema en la clasificación es:

- A) Aplicar el banco de filtros de la escala de Mel
- B) Recurrir a la representación frecuencial de la señal de voz
- C) Emplear técnicas para obtener los marcos (frames) de la señal, como la ventana de Hamming
- D) Usar técnicas de regularización del tamaño de la representación, como segmentación de traza

☐ A Indica que característica **no** presenta la representación *bag of words* de documentos de texto:

- A) Tienen en cuenta la información contextual por cada token del documento representado
- B) De cada documento se deriva un vector de características
- C) Se le pueden aplicar funciones de atenuación para disminuir el peso de los tokens más frecuentes
- D) Si la representación es por presencia o ausencia del token, es un vector binario