Examen de Teoría de Percepción - Primer Parcial

ETSINF, Universitat Politècnica de València, Marzo de 2013

Apellidos: ______________________ **Nombre:** ____________

Profesor: ☐ **Carlos Martínez - Jorge Civera** ☐ **Roberto Paredes**

Cuestiones (3 puntos, 40 minutos, sin apuntes)

[B] La frontera de decisión entre dos clases se define como:

A) Aquellos puntos que dan la máxima diferencia entre las funciones discriminantes de ambas clases
B) Aquellos puntos que dan el mismo valor para las funciones discriminantes de ambas clases
C) Aquellos puntos en que la función discriminante de una clase da mayor valor que la de la otra clase
D) Aquellos puntos en que las funciones discriminantes de ambas clases dan un valor nulo

[C] El objetivo de la fase de preproceso en problemas de OCR es:

A) Encontrar la caja de mínima inclusión
B) Normalizar el aspect ratio
C) Conseguir invarianza a la posición y escala del carácter
D) Conseguir invarianza a la posición del carácter

[B] Dada una representación global de una imagen en 256 niveles de gris de 32x32 píxeles y una representación local de la misma imagen que emplea 10 ventanas de 9x9 píxeles, ¿qué afirmación es correcta?

A) La representación global ocupa más del doble de bytes que la local
B) La representación global ocupa más bytes que la local, pero menos del doble
C) La representación global ocupa los mismos bytes que la local
D) La representación global ocupa menos bytes que la local

[A] Cuál de las siguientes técnicas de detección de puntos de interés preserva mejor la propiedad de estar basada en escoger puntos con cierta información e invarianza:

A) Detección de contornos
B) Extracción por rejilla
C) Extracción aleatoria
D) Extracción aleatoria y por rejilla

[D] Se quiere adquirir una señal del canto de una persona que se sabe que obtiene armónicos de hasta 7000 Hz. ¿Qué frecuencia mínima de muestreo hay que emplear para obtener una señal que reproduzca fielmente ese canto?

A) 8000 Hz
B) 7000 Hz
C) 16000 Hz
D) 14000 Hz

[C] Uno de los problemas clásicos en el reconocimiento del habla consiste en el modelado de la variabilidad temporal no lineal; ¿en qué consiste ese fenómeno?

A) En la baja capacidad discriminativa de la señal temporal
B) En la dificultad de obtener las frecuencias representativas de una señal vocal
C) En la distinta duración de los sonidos en una señal vocal
D) En la imposibilidad de aplicar la transformada de Fourier sobre señales de una duración indeterminada

$\boxed{\text{A}}$ Dada una señal de audio de 10 segundos sobre la que se aplica una ventana de Hamming con tamaño W = 50 ms y solapamiento S = 20 ms, ¿cuántos marcos (frames) se obtendrán de la misma?

A) menos de 400 marcos
B) entre 400 y 499 marcos
C) entre 500 y 600 marcos
D) más de 600 marcos

$\boxed{\text{D}}$ El proceso de tokenización tiene, entre otras características

A) Ser independiente de la lengua tratada
B) Presentar una definición formal que permite automatizar fácilmente el proceso
C) Utilizar una serie de tokens especiales llamados "stop words"
D) Ser generalmente un proceso heurístico

$\boxed{\text{C}}$ Dado un espacio de representación de 100 dimensiones se desea reducir a 20 dimensiones mediante PCA, para ello:

A) Escogeremos 20 eigenvectores de la matriz de covarianza de los datos
B) Escogeremos 20 eigenvectores de la matriz intra-clase de los datos
C) Escogeremos 20 eigenvectores de mayor eigenvalor asociado de la matriz de covarianza de los datos
D) Escogeremos 80 eigenvectores de mayor eigenvalor asociado de la matriz de covarianza de los datos

$\boxed{\text{D}}$ Dado un problema de clasificiación en 15 clases donde los objetos se representan en un espacio de representación de 100 dimensiones. Se desea obtener una representación en un espacio reducido de 5 dimensiones. ¿ Cuál de las siguientes reducciones es la menos aconsejable?

A) Proyectar primero con PCA a 100 dimensiones y luego con LDA a 5
B) Proyectar primero con PCA a 20 dimensiones y luego con LDA a 5
C) Proyectar con LDA a 5 dimensiones
D) Proyectar con PCA a 5 dimensiones

Examen de Teoría de Percepción - Primer Parcial

ETSINF, Universitat Politècnica de València, Marzo de 2013

Apellidos: ____________________ **Nombre:** __________

Profesor: ☐ **Carlos Martínez - Jorge Civera** ☐ **Roberto Paredes**

Problemas (4 puntos, 75 minutos, con apuntes)

1. **(1 punto)** Se tiene un clasificador en dos clases basado en distribuciones gaussianas unidimensionales, de forma que para la clase 1 se tiene $\mu_1 = 0$, $\sigma_1 = 2$, y para la clase 2 se tiene $\mu_2 = 1$, $\sigma_2 = 1$. Se pide clasificar el punto $y = 0.5$ empleando $\arg\max P(c|y)$ teniendo probabilidades a priori idénticas para ambas clases.

   Solución:

   Para clase 1: $p(0.5|1) = \frac{1}{2\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{0.5-0}{2}\right)^2} = \frac{1}{2}\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{32}}$

   Para clase 2: $p(0.5|2) = \frac{1}{1\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{0.5-1}{1}\right)^2} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{8}}$

   Tomando log en clase 1: $\log p(0.5|1) = \log \frac{1}{2} + \log \frac{1}{\sqrt{2\pi}} - \frac{1}{32}$

   Tomando log en clase 2: $\log p(0.5|2) = \log \frac{1}{\sqrt{2\pi}} - \frac{1}{8}$

   Llamando $K = \log \frac{1}{\sqrt{2\pi}}$, tenemos para clase 1: $\log p(0.5|1) = -\log 2 - \frac{1}{32} + K = -0.69 - 0.03 + K = -0.72 + K$

   Y para clase 2: $\log p(0.5|2) = -\frac{1}{8} + K = -0.13 + K = -0.13 + K$

   Por tanto, al tener idénticas probabilidades a priori, se clasifica en clase 2

2. **(1 punto)** Para un problema de clasificación de imágenes, se tienen imágenes de 20x20 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 5x5, en donde se muestrea uno de cada dos píxeles en dirección horizontal

   Solución:

   a) $20\times20 = 400$ píxeles, 1 byte/píxel $\rightarrow$ 400 bytes

   b) $20\times20 = 400$ píxeles, dando 2 bytes por nivel (1 byte 0-255, 2 bytes 0-65535), por 256 niveles $\rightarrow$ 512 bytes

   c) Ventanas de 5×5 en 20×20 son 16×16=256 ventanas; al muestrearse cada dos píxeles, son 128 ventanas de 25 píxeles $= 3200$ píxeles, 1 byte/píxel $\rightarrow$ 3200 bytes

3. **(1 punto)** 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)

   Solución: acbbcad

4. **(1 punto)** Sean las siguientes 12 muestras en un espacio $\mathcal{R}^4$:

$$\mathbf{x}_0 = \begin{bmatrix}0\\0\\0\\0\end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \quad \mathbf{x}_4 = \begin{bmatrix}0\\0.5\\0\\0\end{bmatrix} \quad \mathbf{x}_6 = \begin{bmatrix}0\\0.5\\1\\0\end{bmatrix} \quad \mathbf{x}_8 = \begin{bmatrix}0\\0\\0\\2\end{bmatrix} \quad \mathbf{x}_{10} = \begin{bmatrix}0\\0.5\\1\\2\end{bmatrix}$$

$$\mathbf{x}_1 = \begin{bmatrix}0\\0\\0\\1\end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} \quad \mathbf{x}_5 = \begin{bmatrix}0\\0.5\\0\\1\end{bmatrix} \quad \mathbf{x}_7 = \begin{bmatrix}0\\0.5\\1\\1\end{bmatrix} \quad \mathbf{x}_9 = \begin{bmatrix}0\\0\\1\\2\end{bmatrix} \quad \mathbf{x}_{11} = \begin{bmatrix}0\\0.5\\0\\2\end{bmatrix}$$

siendo los valores y vectores propios de la matrix de covarianza de los datos:

$$\Lambda = \begin{bmatrix}0\\0.0625\\0.25\\0.667\end{bmatrix} \qquad W = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

a) Calcula la proyección PCA para todas las muestras de $\mathcal{R}^4$ a $\mathcal{R}^2$.

b) Dado que conocemos las etiquetas de clase de las muestras ($\{\mathbf{x}_0, \mathbf{x}_4, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8, \mathbf{x}_{10}, \mathbf{x}_{11}\} \in A$ y $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_5, \mathbf{x}_9\} \in B$), calcula el error de clasificación del siguiente clasificador lineal:

$$g_A(\mathbf{x}) = x - y + 1$$
$$g_B(\mathbf{x}) = x + y$$

Solución:

a) En la solución original de este examen no se hacía la resta de la media sobre los vectores a proyectar Proyectamos las muestras utilizando los vectores propios asociados a los valores propios de mayor magnitud

$$W' = \begin{bmatrix}0&0&0&1\\0&0&1&0\end{bmatrix}$$

Esto se reduce a quedarnos con la tercera y cuarta componente de cada una de las muestras restándoles previamente la media $\bar{\mathbf{x}}^t = [0, 0.25, 0.5, 1]$:

$$\mathbf{x}'_0 = \begin{bmatrix}-1\\-0.5\end{bmatrix} \quad \mathbf{x}'_2 = \begin{bmatrix}-1\\0.5\end{bmatrix} \quad \mathbf{x}'_4 = \begin{bmatrix}-1\\-0.5\end{bmatrix} \quad \mathbf{x}'_6 = \begin{bmatrix}-1\\0.5\end{bmatrix} \quad \mathbf{x}'_8 = \begin{bmatrix}1\\-0.5\end{bmatrix} \quad \mathbf{x}'_{10} = \begin{bmatrix}1\\0.5\end{bmatrix}$$

$$\mathbf{x}'_1 = \begin{bmatrix}0\\-0.5\end{bmatrix} \quad \mathbf{x}'_3 = \begin{bmatrix}0\\0.5\end{bmatrix} \quad \mathbf{x}'_5 = \begin{bmatrix}0\\-0.5\end{bmatrix} \quad \mathbf{x}'_7 = \begin{bmatrix}0\\0.5\end{bmatrix} \quad \mathbf{x}'_9 = \begin{bmatrix}1\\0.5\end{bmatrix} \quad \mathbf{x}'_{11} = \begin{bmatrix}1\\-0.5\end{bmatrix}$$

b) Calculamos para cada muestra la función discriminante de cada clase, etiqueta de clase estimada y comparamos con la etiqueta real para ver si se ha producido un error de clasificación:

Muestra	$g_A(\mathbf{x})$	$g_B(\mathbf{x})$	Estimada	Real	Error
$\mathbf{x}'_0$	0.5	-1.5	A	A	No
$\mathbf{x}'_1$	1.5	-0.5	A	B	Sí
$\mathbf{x}'_2$	-0.5	-0.5	?	B	?
$\mathbf{x}'_3$	0.5	0.5	?	B	?
$\mathbf{x}'_4$	0.5	-1.5	A	A	No
$\mathbf{x}'_5$	1.5	-0.5	A	B	Sí
$\mathbf{x}'_6$	-0.5	-0.5	?	A	?
$\mathbf{x}'_7$	0.5	0.5	?	A	?
$\mathbf{x}'_8$	2.5	0.5	A	A	No
$\mathbf{x}'_9$	1.5	1.5	?	B	?
$\mathbf{x}'_{10}$	1.5	1.5	?	A	?
$\mathbf{x}'_{11}$	2.5	0.5	A	A	No

En total se producen 2 errores de clasificación y 6 empates.