

## Planificación del tema:

Fundamentos teóricos mapas cognitivos. planteamiento de la práctica. Entornos experimentales: Escuela [ B2-A3202/ B4-A3202] o domicilio particular.

Especificaciones del trabajo práctico de este tema:

-1.-creación del mapa topológico del entorno escogido [ grafo con nodos/landmarks + arcos con orientaciones y distancias ( $\theta, \delta$ ) entre los landmarks del entorno escogido.

2.- Desarrollo y testado del reconocedor de imágenes de los landmarks del entorno escogido: se generará un data set de al menos 60 imágenes per landmark para entrenar y testar el reconocedor del navegador. La validación del reconocedor se basará en las siguientes pruebas estática y dinámica:

**Prueba “ estática”** → diseño y test “estático” ( % de errores mediante “leave-one-out” contra el data set de entrenamiento(60-100 imágenes per landmark) del reconocedor de imágenes de los landmarks del entorno de navegación escogido ( clasificador 1-NN- $\delta$  con el histograma de grises normalizado).

**-Prueba dinámica** → test del reconocedor y del navegador contra vídeos de recorridos ( es suficiente pasar un par de vídeos para el test dinámico del reconocedor ) captados por los estudiantes en el entorno escogido.

La práctica se puede realizar individualmente o en equipo ( máximo 6 personas). Entregas finales de trabajos:

1.- Memoria en formato pdf describiendo el mapa topológico del entorno y el diseño y testado estático del reconocedor de imágenes.

2.- vídeo de un recorrido (en visión subjetiva o fpv) del entorno experimental en el que se visualice el reconocimiento de los landmarks del recorrido. Enviar ambas entregas directamente por correo electrónico antes de las 23:55 horas del viernes 27 de noviembre de 2020 al profesor del tema: [dario.maravall@upm.es](mailto:dario.maravall@upm.es)

# Animals' Spatial Cognitive Maps:

- Edward C. Tolman in his experiments with the spatial orientation and navigation of rodents in labyrinths introduced the concept of spatial cognitive map “as formed of two related pieces of information. First, locations and their connecting paths, and second and simultaneously, objects or events associated with the corresponding location (*their semantic meaning*)”. Cognitive map= a set of perceptual landmarks and their physical links.

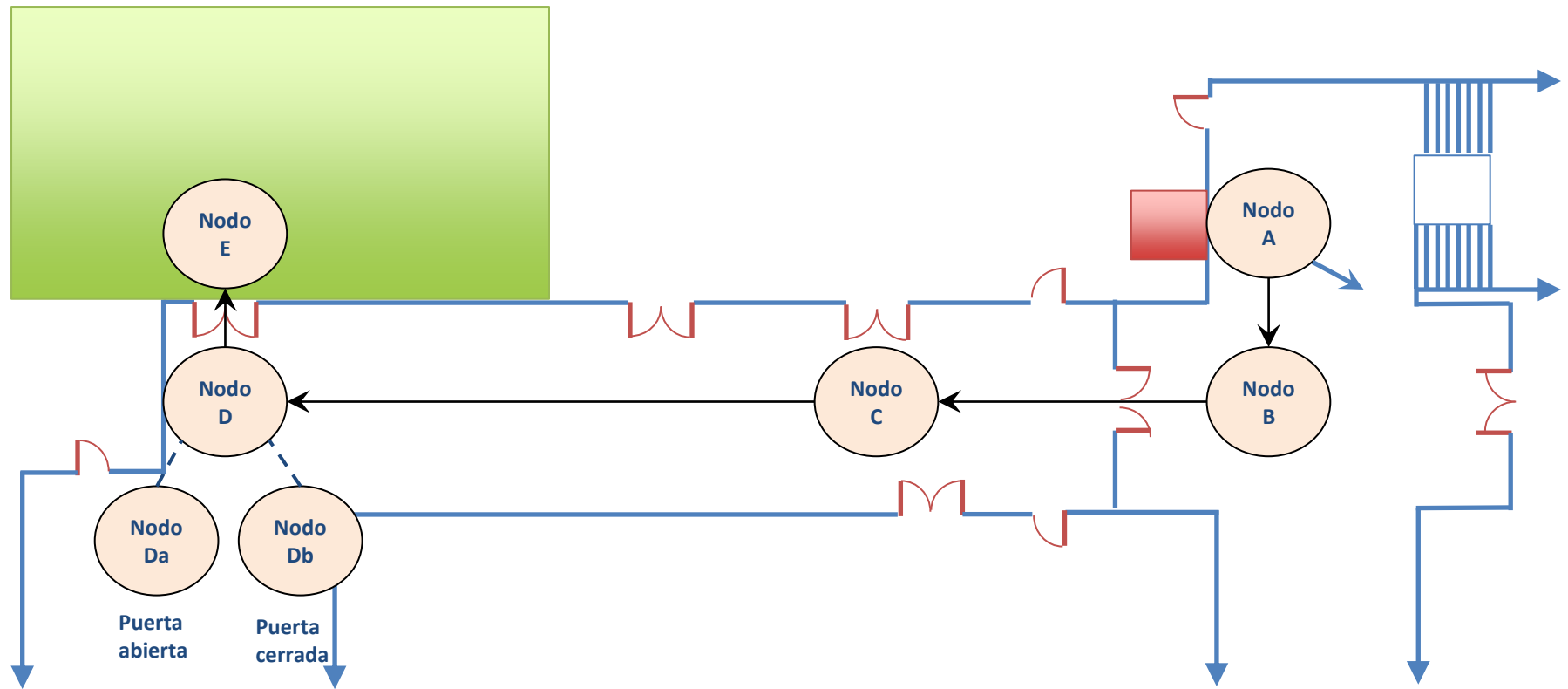
# Robot Sensorial Topological Maps:

- Robot sensorial topological maps are perception-based topological maps implemented as graphs with sensorial information associated to each node/landmark such as landmark's images plus information concerning the navigation maneuvers associated to the physical navigation between the connected nodes/ landmarks.

## Mapas Topológicos visuales:

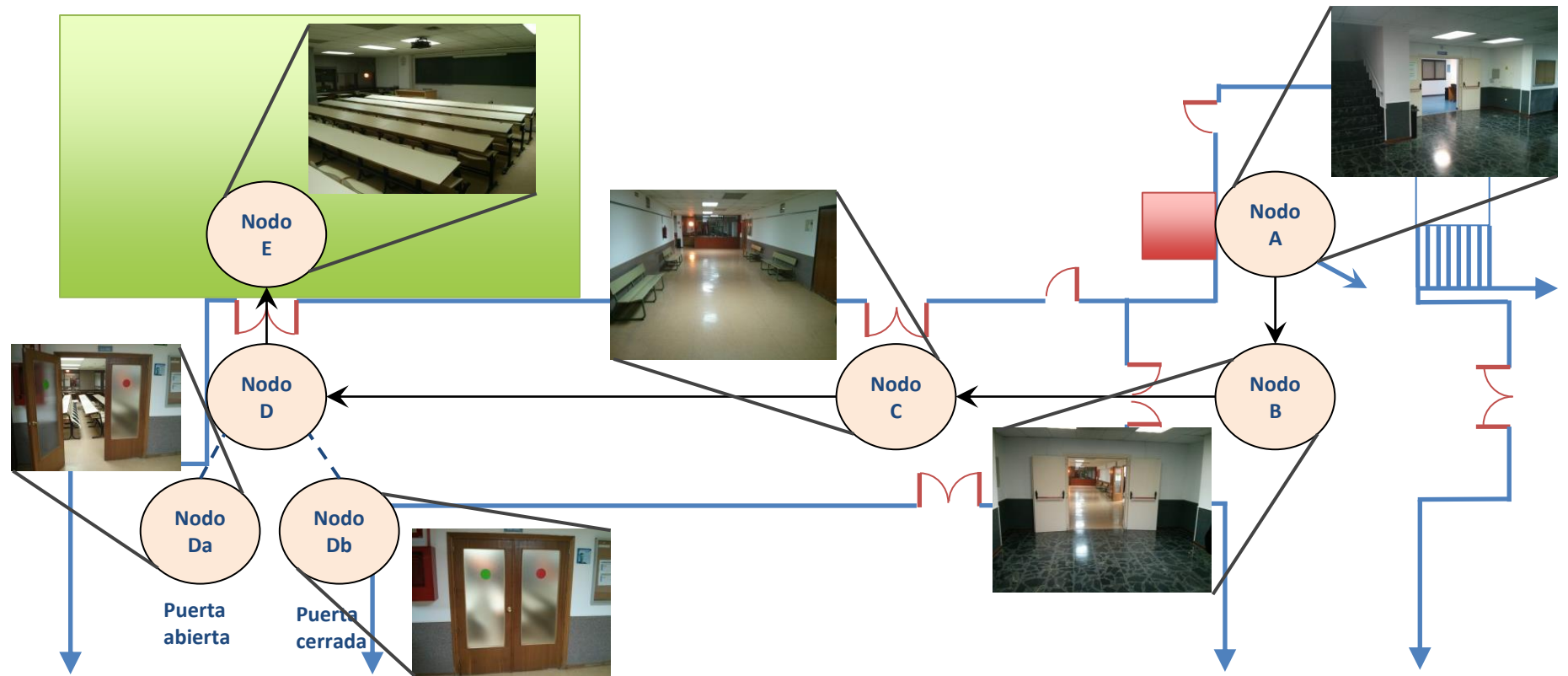
- El entorno se modela con un mapa topológico visual que utilizará el robot para navegar por el entorno
- -Un mapa topológico visual es un grafo en el que los nodos almacenan las imágenes representativas de los lugares relevantes o landmarks del entorno:
- Mapa topológico  $\rightarrow$  Grafo = { Nodos, Arcos }
- \_nodo  $\rightarrow$  se refiere a un lugar relevante o landmark ( del que almacena una imagen representativa)
- - Arco  $\rightarrow$  información para desplazarse entre los nodos conectados ( orientaciones y distancias vs procedimientos de navegación reactiva)

## Mapa Topológico Bloque 2- Aula 3202



# Navegación y reconocimiento de landmarks

## Bloque 2-Aula 3202



## Entorno mixto interior/ exterior:



## Modus Operandi del trabajo práctico de mapas topológicos visuales:

Una vez definido el mapa del entorno de navegación ( grafo de los landmarks del entorno ) , el primer paso consistirá en la creación de una base de datos con las imágenes individuales de todos los landmarks del mapa-grafo a reconocer : se tomarán al menos 60 imágenes por cada landmark para trabajar con data sets etiquetados de suficiente tamaño para el entrenamiento y testeo del reconocedor básico [1-NN con umbral ajustable+ histograma de grises normalizado ] :

$DS = \{ ( \mathbf{x}_1, L_1 ) , \dots , ( \mathbf{x}_p, L_p ) \}$  ;  $\mathbf{x}$  son los vectores de las variables discriminantes y  $L$  el correspondiente landmark ;  $P = (n^\circ \text{ landmarks}) \times 60$

### Funciones de opencv para la creación del data set :

Con la función `videocapture` `cv2videocapture` captamos señal de video y con la función `cv2imread_grayscale` podemos captar los frames directamente en niveles de grises de los landmarks del mapa,  $I_m$

- **Funciones de opencv para el cálculo de las variables discriminantes:**
- Con la función `calcHist` obtenemos el histograma de la imagen de grises de cada landmark y a continuación normalizamos el histograma dividiendo sus valores por el número de píxeles de la imagen:  $n \times m$  (  $n = n^\circ$  de filas;  $m = n^\circ$  de columnas):
- `CalcHist(  $I_m$ , hist )` ; deja en `hist` el histograma de la imagen( 256 valores) y después normalizamos estos valores:
- $hist / n \times m \rightarrow \mathbf{x} = (x_1, x_2, \dots, x_{256})$  es el vector de las variables discriminantes del data set de entrenamiento DS del reconocedor.

Finalmente, testeamos la eficiencia del reconocedor básico de la práctica [ 1-NN con umbral ajustable + histograma de grises normalizado] mediante validación cruzada tipo “leaving-one-out” ( % error): ver las siguientes transparencias.



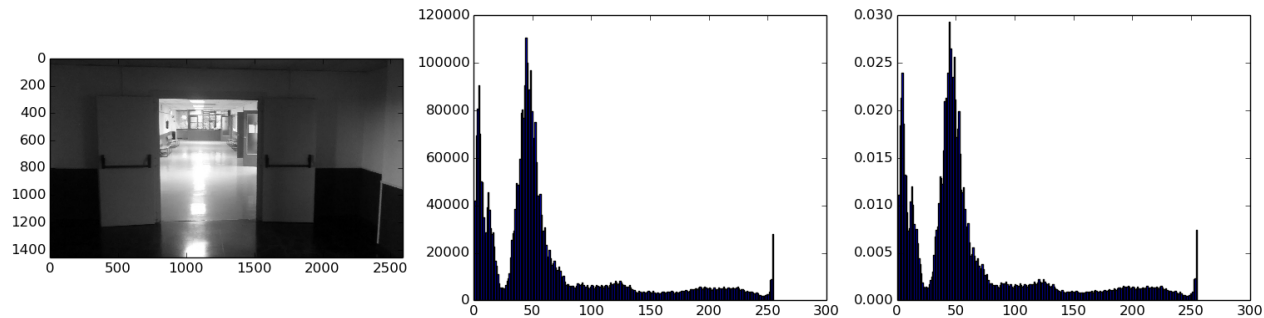
# Pruebas y Testeo del reconocedor de landmarks:

- **1.- Creación del Conjunto de Muestras Etiquetadas CME o data sets de entrenamiento:** se deben captar al menos 60 tomas para cada uno de los landmarks del mapa-grafo a reconocer, por lo que el data set de entrenamiento y test deberá contar con  $P$  datos, con  $P = (n^{\circ} \text{ landmarks del mapa}) \times 60$  :
- $DS = \{ ( \mathbf{x}_1, \alpha_1 ) \dots\dots\dots ( \mathbf{x}_p, \alpha_p ) \}$  ; siendo  $\mathbf{x}$  el vector de variables discriminantes ( histograma de grises normalizado) y  $\alpha$  es la etiqueta( landmark) del correspondiente dato
- **2.-Testeo estático del reconocedor básico:** se estimarán las tasas de errores por el método de validación cruzada “leave-one-out” del reconocedor básico [ 1-NN con umbral ajustable + histograma normalizado de grises ].
- **3.- Testeo dinámico del navegador:** una vez testeado en estática el reconocedor del navegador, pasamos al testeo dinámico contra vídeos de recorridos del entorno de navegación para estimar la eficiencia del navegador en el reconocimiento de los landmarks del entorno.

# Resumen reconocimiento de landmarks

## Bloque 2- Aula3202

- 1-NN
- Validación Leave one out
- Histograma de grises normalizado
- Dataset equilibrado, sin condensar



## Algoritmo 1-NN- $\delta$ con umbral ajustable + histograma normalizado:

### □ Procedure 1-NN with adjustable threshold

**Design parameter:** an adjustable distance threshold  $\delta \approx 0$  to be optimized

### □ Given a memory of P landmarks' images and their corresponding normalized histograms:

[ (Landmark  $l_1$  :  $l_1 \rightarrow h_1$ ) ..... ( $l_p \rightarrow l_p \rightarrow h_p$ ) ]

remind:  $h \rightarrow (x_1, x_2, \dots, x_{256})$  : the 256 normalized histogram values

### □ For a new image [ $l_{new} \rightarrow h_{new}$ ] to be recognized as a possible landmark **do**

#### □ 1<sup>st</sup> step :

### □ **Compute** the Euclidean distances of the new image's normalized histogram $h_{new}$ to the P normalized histograms stored in the recognizer's memory and let $d_{min}$ be the minimum of the P Euclidean distances

#### □ 2<sup>nd</sup> step :

### □ **If** [ $d_{min} \leq \delta$ ] **Then** [classify the new image as the landmark corresponding to the minimum **distance** $d_{min}$ ] **Else** [classify the new image as "unknown"]

Note: the adjustable distance threshold  $\delta$  is meant to avoid false alarms and simultaneously to keep a good efficiency in the classification of the real landmarks. For its experimental tuning it is advisable to implement a search starting with a small value near zero (  $\delta_{min} = 0,1$  ) up to a maximum value (  $\delta_{max} = 1,0$  ) by means of small scanning steps (  $\approx +0,1$  )

Note: use the landmarks training dataset DS as the memory of the recognizer.

# The Leave-one-out cross validation method

For a Dataset with P labeled exemplars perform P **steps**:

- For each **step** use P-1 labeled exemplars for training and the remaining one for testing;
- Compute the true error rate as the average error of the tested exemplars:
- $\text{'true error rate' } \epsilon\% = (\text{\#errors in the tested exemplars} / P) \times 100$
- Note : it is a particular case of the k-fold cross validation method in which the training data set is randomly divided into k equally sized groups such that you leave out one of these k groups and you train the classifiers with the remaining k-1 groups and test them with the group you left out. You have to repeat this operation k times and finally you must estimate the average accuracy of the classifiers.
- For huge training data sets is advisable to use the k-fold cross validation [ e.g. k= 5, 10]; but for small training data sets like in this *practical assignment* [ with just  $P \approx 500$  labeled instances] is compulsory to use the leave-one- out validation method.

# Pruebas estática y dinámica:

## □ **Prueba estática:**

Eficiencia ( % error) del reconocedor ( 1-NN-umbral, histograma de grises normalizado) con el data set de imágenes de los landmarks del mapa del entorno (al menos 60 imágenes por landmark)

## **Prueba dinámica:**

Reconocimiento online de los landmarks en un vídeo de recorrido del entorno.