

Estrategias de Exploración para la Construcción Autónoma de Modelos Topológicos en Robots Móviles

J. de Lope¹, D. Maravall², J.G. Zato¹
Dpto de Sistemas Inteligentes Aplicados¹ Dpto. de Inteligencia Artificial²
Universidad Politécnica de Madrid, España
e-mail: {jdllope,jzato}@eui.upm.es, dmaravall@fi.upm.es

Palabras clave: robots móviles autónomos, modelos topológicos, estrategias de exploración.
Keywords: autonomous mobile robots, topological modeling, exploration strategies.

RESUMEN. *La exploración puede definirse como el proceso mediante el cual el robot recorre el entorno detectando lugares de referencia y transiciones entre ellos sin otro objetivo que construir un modelo del entorno lo más completo posible. En este artículo, se presentan procedimientos o pautas a seguir para la construcción de modelos mediante exploración y se definen parámetros que permiten medir la calidad de los mismos.*

ABSTRACT. *Exploration can be defined as the process in which the robot crosses the environment and detects landmarks and transitions between them for constructing an environment model. The model must be as complete as possible for using it in a planning phase. In this paper, we present procedures and guidelines for constructing environment models using the exploration concept. Parameters that describe the quality of the guidelines are also defined.*

1.- INTRODUCCIÓN.

La exploración puede definirse como el proceso mediante el cual un robot móvil recorre el entorno detectando lugares de referencia y transiciones entre ellos sin otro objetivo que construir de forma autónoma un modelo del entorno lo más completo posible.

En ocasiones, el proceso de exploración no tiene lugar de forma independiente, construyéndose el modelo del entorno a medida de que el robot va completando misiones. De esta forma, cada misión lleva consigo la búsqueda de un camino que una la salida con la meta y una exploración de la zona del entorno en la que tiene lugar. El modelo del entorno así construido será incompleto ya que incluirá solamente las zonas del entorno por las que se ha pasado durante la búsqueda del objetivo, sin tener en cuenta todas las alternativas desde cada lugar de referencia. Además, como el camino puede que no sea reversible, el modelo construido puede no ser útil cuando se invierten los papeles entre los lugares de salida y de llegada. Por ello, este tipo de exploración precisa de una fase de refinamiento del modelo, que permita ampliarlo con zonas que todavía no se han visitado y completar las ya conocidas.

Otra desventaja de esta aproximación está relacionada con la forma en la que se define el objetivo. En ausencia de un modelo del entorno, el mecanismo para la definición del objetivo se reduce prácticamente al uso de las coordenadas de la posición que se quiere alcanzar. De no usarse este método, es necesario disponer de un sistema sensorial y algoritmos lo suficientemente complejos para que se admita una descripción más o menos detallada de la zona del entorno en la que se encuentra la meta.

En este artículo, se presentan procedimientos o pautas a seguir para la construcción de modelos mediante exploración y se definen parámetros que permiten medir la calidad de los mismos.

2.- ESTRATEGIAS DE EXPLORACIÓN.

La primera referencia al término *estrategia de exploración* como secuencia ordenada de acciones para la definición de un modelo del entorno se debe a Shannon (Shannon 52). Su aportación en este campo fue la de una máquina capaz de resolver un laberinto de 5×5 casillas por un método de ensayo y error y de recordar la solución desde todas las casillas por las que había pasado, incluida la inicial, hasta la casilla de destino.

La máquina tenía dos modos de funcionamiento a los que denominó estrategia de exploración y estrategia de objetivo. Durante la *estrategia de exploración* empleaba dos relés asociados con cada casilla capaces de recordar una de las cuatro posibles direcciones de salida, en concreto, la dirección de salida de la última vez que se había abandonado la casilla. Cuando se llegaba a una casilla, el sistema obtenía la dirección de salida de la última visita e incrementaba su valor con un giro de 90° a la izquierda antes de abandonar la casilla en esa dirección. Una mejora que permitía reducir el tiempo de exploración a la tercera parte era la de impedir que la dirección de salida fuese la de entrada en la visita actual. De esta forma se prevenía que los ensayos de ese camino se repitiesen demasiado.

Después de unas cuantas vueltas por el laberinto, se llegaba a la casilla destino. A partir de ese momento, la máquina operaba según la *estrategia de objetivo*. En ella se utilizaban las direcciones memorizadas en los relés de cada casilla para decidir la dirección de salida que condujese al objetivo.

Otra característica de la máquina era que podía detectar si la configuración del laberinto había cambiado, se había caído en un bucle sin fin o se había entrado en cualquier otra situación de *neurosis*. Esto se efectuaba mediante un circuito que contaba el número de casillas por las que se había pasado, cuando este número era mayor que veinticuatro y no se había alcanzado la casilla destino, se había presentado una situación de neurosis. El remedio era olvidar los valores memorizados en cada casilla y volver a la estrategia de exploración que garantizaba encontrar la solución.

Este ejemplo de la literatura clásica recoge conceptos precursores de los que se emplean actualmente. En primer lugar, tenemos que una estrategia de exploración ordenada con cualquier criterio puede resolver el problema de la construcción de modelos de entornos. En segundo lugar, pequeñas variaciones en los criterios pueden conducir a mejorar de forma sustancial las soluciones obtenidas o incluso a desistir ante situaciones comprometidas. Por último, la utilización del modelo para llegar a la casilla destino establece el precedente de los planes como comunicación (Agre & Chapman 91), de los campos de gradiente (Payton 91) y, por extensión, de nuestro enfoque basado en sugerencias (de Lope 98).

3.- CALIDAD DE LA EXPLORACIÓN.

De las consideraciones anteriores sobre la posibilidad de utilizar varios criterios durante el proceso de exploración se desprende inmediatamente que puede existir más de una estrategia y que cada una puede definirse como mejor o peor en cuanto a la calidad del producto obtenido con relación al esfuerzo necesitado para ello.

Para determinar la calidad de la exploración pueden seguirse dos procedimientos alternativos: comparar el nivel alcanzado en los modelos para un mismo coste o comparar el coste de la construcción completa. El más habitual es este último pero implica tener que construir todo el modelo. Tanto en uno como en otro caso, debemos definir el coste de exploración.

En ocasiones (Lumelsky *et al.* 90) se utiliza como coste de exploración la distancia que recorre el robot durante la exploración del entorno comparándola con la distancia total que tendría que recorrerse para detectar todos los obstáculos del entorno. En otras (Lee 96) se introduce una nueva variable para reflejar también el coste introducido por los sensores durante las acciones de captación de las características del entorno y define su coste como el tiempo total utilizado tanto

en los movimientos del robot como en las acciones de percepción. Otra alternativa, ésta desde el ámbito de estudio del comportamiento animal, es la asociar el coste de exploración con la energía consumida por el robot durante el proceso.

En nuestro caso, definimos el coste como el tiempo total utilizado por la exploración, pero a diferencia de los métodos tradicionales, determinamos la calidad de la exploración comparando el nivel de realización conseguido para un mismo coste.

Sea p_j el número de lugares de referencia detectados como nuevos durante el intervalo j de duración Δt . El número total de lugares creados en el modelo, P_i , hasta el intervalo i es:

$$P_i = \sum_{j=0}^i p_j \quad (1)$$

El número acumulado de lugares de referencia P_T durante los N primeros intervalos de la exploración viene determinado por la expresión:

$$P_T = \sum_{i=0}^N P_i \quad (2)$$

Si P_M es el número máximo de lugares de referencia que pueden detectarse en el entorno, δ_p se define de manera que indique la diferencia entre el número máximo de lugares y el número acumulado de los que se han ido detectando durante la exploración en los primeros N intervalos, es decir

$$\delta_p = N \cdot P_M - P_T \quad (3)$$

Para un mismo entorno, cuánto menor sea esta diferencia, más rápidamente se habrán detectado todos los nodos del entorno.

La calidad de la exploración en función del número de lugares de referencia puede describirse, de forma normalizada, mediante la expresión:

$$Q_p = 1 - \frac{\delta_p}{N \cdot P_M} \quad (4)$$

y operando

$$Q_p = \frac{1}{N \cdot P_M} \sum_{i=0}^N P_i \quad (5)$$

El valor de Q_p de una estrategia de exploración será tal que cuánto más se acerque a 1 mejor será dicha estrategia en la detección de los lugares. Si fuese $Q_p = 1$ indicaría que la estrategia ha sido capaz de encontrar todos los lugares del entorno en el primer intervalo considerado. Al estar normalizados sus valores en el intervalo $[0,1]$ podremos comparar los resultados de varias estrategias para elegir la que mejor se adapte a su criterio.

De forma análoga, podemos construir otros índices de la calidad de la exploración: el que establece la calidad en función del número de transiciones existentes entre los lugares de referencia del entorno

$$Q_T = \frac{1}{N \cdot T_M} \sum_{i=0}^N T_i \quad (6)$$

y el que determina la calidad conjunta, considerando lugares de referencia y transiciones

$$Q = \frac{1}{N(P_M + T_M)} \sum_{i=0}^N (P_i + T_i) \quad (7)$$

donde T_i se refiere al número total de transiciones creadas en el modelo hasta el intervalo i y T_M al número máximo de transiciones que pueden detectarse en el entorno.

El posible inconveniente en la utilización de estos índices de calidad se presenta al tener que conocer los valores de P_M y T_M antes de que se encuentre una estrategia de exploración que garantice la construcción completa del modelo del entorno. En nuestro caso, es sencillo obtener

estos valores ya que pueden precisarse de antemano considerando las características del entorno. Para el caso general, podrían determinarse los valores dejando concluir completamente una exploración.

Los lugares de referencia contemplados son en los que el robot debe cambiar su trayectoria para evitar una colisión y quedan reflejados en la Figura 2. Los tipos (a) y (b) se asocian a rincones del entorno, a la izquierda y a la derecha, respectivamente. En el tipo (c) existe un obstáculo frontal pero no obstáculos laterales. Las estrategias de navegación pueden ser de navegación libre de colisiones, de seguimiento del perímetro de obstáculos, de desvío (para reflejar órdenes del estilo *toma la primera a la derecha*) y de giro (de Lope *et al.* 98).

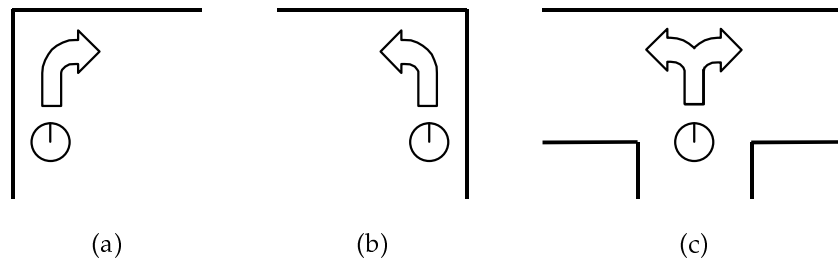


Figura 1: Tipos de lugares de referencia detectados en los entornos

4.- RESULTADOS EXPERIMENTALES.

Vamos a definir un conjunto de estrategias de exploración de las que se compararán sus índices de calidad. Cada estrategia de exploración está basada en diferentes patrones de decisión sobre el orden y número de veces que se aplican tanto los sentidos de giro en la salida de los lugares como las estrategias de navegación a activar. También se contemplan otros criterios referidos con la determinación de lugares sobre los que ensayar nuevas transiciones.

Las estrategias de exploración pueden ser dependientes del entorno que está siendo explorado o incluso de la posición desde la que se inicia la exploración. Por ello, cada estrategia se utilizará sobre 3 entornos diferentes desde 5 lugares de salida, repitiendo cada misión de exploración en 3 ocasiones. Se considerarán $N = 30$ intervalos de $\Delta t = 60$ segundos de duración.

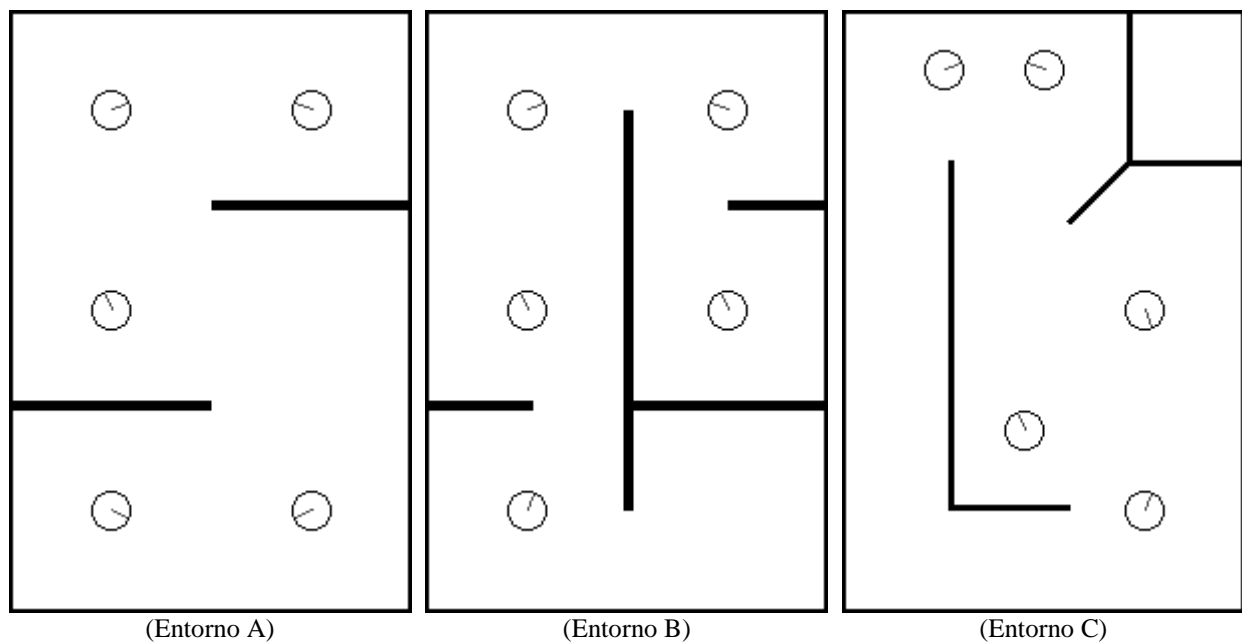


Figura 2: Entornos de los experimentos con las posiciones de salida del robot

En la Figura 2 se muestran los 3 entornos usados en los experimentos en los que se incluyen los lugares de salida y en la Tabla 1 los valores de P_M , número máximo de lugares, y T_M , número máximo de transiciones, de los entornos. Las figuras se corresponden con ventanas capturadas en el simulador de robots móviles Nomad-200 construido para realizar las pruebas de las estrategias de navegación y de exploración antes de introducirlas en el robot real. En este caso, el uso del simulador asegura cierto determinismo durante los experimentos ya que se descartan eventos externos no controlados.

Entorno	P_M	T_M
A	21	45
B	33	79
C	15	32

Tabla 1: Valores P_M y T_M de los entornos del experimento

4.1. Aleatoria.

La estrategia aleatoria selecciona al azar el tipo de estrategia de navegación a ensayar desde cada lugar de referencia dentro de los que pueden aplicarse en el tipo de lugar concreto. Para los lugares de tipos (a) y (b), que se corresponden con rincones en el entorno, las estrategias de navegación posibles son: navegación libre, seguimiento del perímetro, desvío y giro, con el sentido de giro que aleje al robot del obstáculo. Por último, para los de tipo (c) pueden aplicarse las estrategias de navegación libre y de desvío con dos sentidos de giro en la salida del lugar de referencia y que denominaremos en lo sucesivo el lado *corto*, lado con más obstáculos o más cercanos, y el lado *largo*, el contrario.

En la Figura 3 se muestra la curva que forma los valores de la calidad conjunta para cada entorno: en línea continua el A, en línea de puntos el B y en línea discontinua el C. Como puede verse, con esta estrategia no es posible completar la exploración de ninguno de los tres entornos en el periodo estudiado (por ejemplo, para el entorno A, el más sencillo de todos, son necesarios algo más de 40 minutos de exploración, momento en el que la calidad conjunta vale 1).

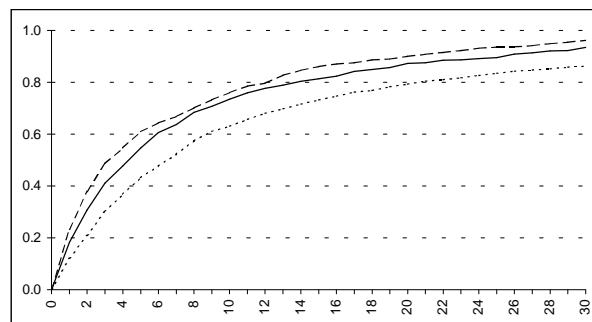


Figura 3: Gráficas de exploración de la estrategia aleatoria

La Tabla 2 muestra los índices de calidad y un porcentaje que indica el nivel alcanzado en el modelo del entorno. En general, los valores obtenidos son algo bajos, sobre todo en el entorno B en el que el número de lugares y transiciones a detectar es relativamente alto. En promedio, se completa aproximadamente un 92% del total del modelo.

Entorno	Q_P	Q_T	Q	% Total	Minutos
A	0,8691	0,6955	0,7507	93,43	>30
B	0,8360	0,5930	0,6646	86,25	>30
C	0,8893	0,7400	0,7887	96,23	>30
Promedio	0,8648	0,6762	0,7347	91,97	>30

Tabla 2: Parámetros de la estrategia aleatoria

4.2. Secuencial.

En este caso se establece un orden en el ensayo de las estrategias de navegación desde cada lugar de referencia con un criterio fijo relativo a probar primero estrategias asociadas a comportamientos menos costosos. Para los lugares de tipos (a) y (b) el orden de las estrategias es: navegación libre, seguimiento del perímetro, desvío y giro. Para los lugares de tipo (c) el orden es: navegación libre con sentido de giro al lado corto, navegación libre con sentido de giro al lado largo, desvío con sentido de giro al lado corto y desvío con sentido de giro al lado largo. Se toma en primer lugar el sentido de giro al lado corto ya que favorece que se continúe la exploración en la misma zona en la que se encuentra el robot aunque, obviamente, pueden presentarse excepciones que hagan que, al tomar el lado corto, el robot entre en grandes zonas aún sin visitar, dejando por un tiempo la exploración de la zona en la que se tomó la decisión.

Ensayar indefinidamente la misma secuencia de estrategias desde cada lugar de referencia puede llevar a la entrada en bucles en la exploración con el consiguiente estancamiento en la obtención del modelo —en cierta forma, equivalente al estado de neurosis que mencionaba Shannon—. Por ello, cada vez que se completa una secuencia de cuatro estrategias se selecciona aleatoriamente una de las disponibles para ser ensayada antes de comenzar de nuevo con la secuencia.

La Figura 4 muestra las curvas de exploración de esta estrategia. Obsérvese que solamente en el entorno A se completa el modelo para todos los lugares y transiciones. En el entorno B, con 79 transiciones, esto no es posible al tratarse de un número relativamente elevado para el tiempo que se dispone. En el entorno C tampoco se detectan todas las transiciones, ya que algunas de ellas parten de los lugares que se visitan en muy pocas ocasiones con este tipo de exploración.

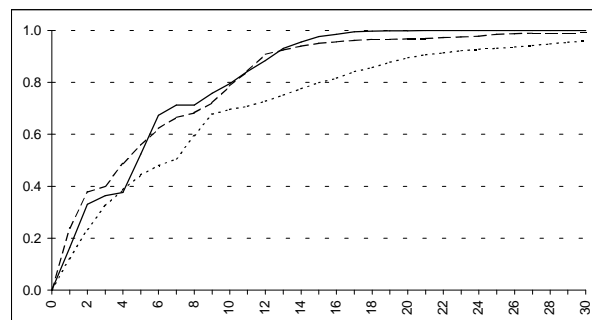


Figura 4: Gráficas de exploración de la estrategia secuencial

La Tabla 3 recoge los parámetros de la estrategia secuencial. Los índices de calidad son mejores que los de la estrategia aleatoria tanto en promedio como para cada entorno por separado. Sólo en el entorno A se consigue el modelo completo tras al menos 21 minutos de exploración.

Entorno	Q_P	Q_T	Q	% Total	Minutos
A	0,9181	0,7925	0,8325	100	21
B	0,8918	0,6604	0,7286	95,95	>30
C	0,9132	0,7826	0,8252	99,13	>30
Promedio	0,9077	0,7452	0,7954	98,38	>30

Tabla 3: Parámetros de la estrategia secuencial

Como comentario a esta estrategia de exploración puede decirse que introducir un orden de ensayo en las estrategias de navegación desde los lugares de referencia permite que a la larga se complete el modelo del entorno aunque en ocasiones se requiera demasiado tiempo. Para mejorar el tiempo de exploración se probaron también otras secuencias de ensayo que aumentaban el uso de la navegación libre frente a las de seguimiento. En general, los resultados eran bastante similares a los de la estrategia que se presenta ya que las mejoras son muy dependientes del entorno concreto o de zonas del entorno sobre el que se realiza la prueba.

Otra posible mejora en el tiempo de exploración podría consistir en *dar por ensayadas* ciertas estrategias que se sabe que no van a introducir nueva información en el modelo. Por supuesto, esto no puede realizarse con tan solo información local de cada lugar de referencia. Reglas sencillas del estilo «*si desde un lugar de tipo (b) se detecta otro lugar de tipo (b) con una estrategia de navegación libre, no probar las de seguimiento y las de desvío*» no son válidas aunque funcionen en la mayor parte de los casos. Podrían encontrarse entornos que, a modo de contraejemplo, inutilizaran todas las reglas de este tipo que pudieran proponerse.

4.3. Combinada.

El exceso en el tiempo de exploración está vinculado con el gran número de visitas que se hacen a lugares que ya se han visitado con anterioridad y en los que se han completado todos los ensayos. La estrategia de exploración combinada pretende eliminar este problema. Para ello, se intercalan fases de planificación que conducen al robot al lugar de referencia más cercano de los que no se dispone aún de toda su información. La activación de las fases de planificación se realiza en lugares en los que por el número de visitas recibidas se puede inferir que todas las estrategias posibles han sido probadas, es decir, en lugares desde los que el ensayo de cualquier transición no aporta nada al modelo.

Una vez alcanzado el lugar de referencia cuya información no ha sido completada, se pasa nuevamente a una fase de exploración que continúa desde ese lugar y que se mantendrá mientras que se estén añadiendo transiciones al modelo.

La Figura 5 muestra las gráficas de exploración de esta estrategia. Como puede verse el modelo se completa en los tres entornos. Con esta estrategia la primera vez que se entra en una fase de planificación en el entorno C tiene lugar a los 8 minutos que se refleja en el significativo cambio de pendiente en la curva de líneas discontinuas.

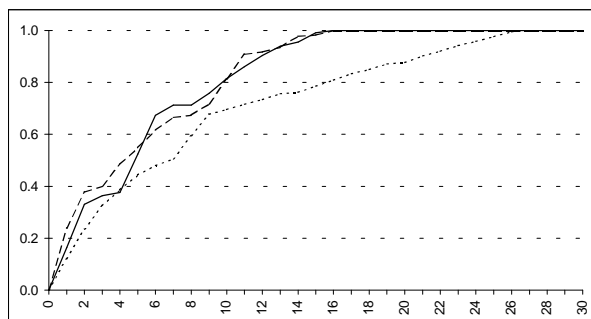


Figura 5: Gráficas de exploración de la estrategia de los no completados

En la Tabla 4 se recogen los parámetros de la estrategia. La mejora en la calidad conjunta Q en su mayor parte es debida a la mejora en la calidad asociada con las transiciones Q_T sobre todo en el entorno B en el que por primera vez se ha podido completar el modelo.

Entorno	Q_P	Q_T	Q	% Total	Minutos
A	0,9181	0,7975	0,8359	100	16
B	0,8919	0,6744	0,7385	100	27
C	0,9149	0,8061	0,8419	100	16
Promedio	0,9086	0,7594	0,8054	100	20

Tabla 4: Parámetros de la estrategia de los no completados

5.- CONCLUSIONES.

Hemos planteado la construcción autónoma de modelos de entornos a partir de estrategias de exploración. Cada estrategia de exploración, mediante un patrón o esquema, ensaya todas las posibles estrategias de navegación que pueden aplicarse en los lugares de referencia desde los

que se parte. Las estrategias de navegación que dan como resultado la detección de un nuevo lugar de referencia o una nueva transición con algún lugar ya existente contribuyen a la actualización del modelo del entorno.

Desde cierto punto de vista, este mecanismo puede ser comparable a los tradicionales métodos de aprendizaje por repetición en los robots manipuladores, e incluso robots móviles (Martin y Nehmzow 95), son guiados en la trayectoria que tienen que seguir para alcanzar el objetivo. En ambos casos, para la ejecución de la trayectoria durante la fase de explotación se repite una serie de movimientos que han sido comprobados previamente. La diferencia con nuestro enfoque radica en la manera en la que el robot aprende los movimientos. En los métodos tradicionales de aprendizaje por repetición, los movimientos a efectuar son mostrados por un agente externo al sistema. En el método de exploración utilizado es el propio robot quien prueba de forma autónoma qué movimientos, o más ampliamente sugerencias de movimientos, son los que le pueden ser de utilidad registrándolos en el modelo para su uso posterior.

Para determinar la calidad de las diferentes estrategias de exploración se han definido índices que permiten evaluar el progreso tanto en la detección de lugares de referencia como en la verificación de estrategias de navegación asociadas con las transiciones entre los lugares. Los índices junto con el nivel al que se ha completado el modelo demuestran las ventajas de la exploración combinada frente a las otras alternativas y cómo puede influir el orden en los ensayos de las estrategias de navegación desde los lugares de referencia.

Las exploraciones para construir modelos topológicos con un robot Nomad-200 en el Laboratorio de Visión Artificial y pasillos cercanos de la Facultad de Informática de la Universidad Politécnica de Madrid se utilizó la estrategia combinada. De esta forma, se pudo reducir el tiempo destinado a la exploración, disminuyendo también el consumo de baterías.

AGRADECIMIENTOS.

El trabajo experimental presentado en este artículo ha sido parcialmente financiado por la Comisión Interministerial de Ciencia y Tecnología (CICYT) bajo contrato TER96-1957-CO3.

REFERENCIAS.

- Agre, P.E. y Chapman, D. (1991): What are plans for? En P. Maes, editor, *Designing Autonomous Robots*, páginas 17-34. MIT Press, Cambridge, Massachusetts.
- De Lope, J. (1998): *Modelado de entornos con técnicas basadas en Redes de Petri Borrosas para la exploración y planificación de robots autónomos*. Tesis Doctoral. Facultad de Informática, Universidad Politécnica de Madrid.
- De Lope, J., Maravall, D. y Zato, J.G. (1998): Topological Modeling with Fuzzy Petri Nets for Autonomous Mobile Robots. En A.P. del Pobil, J. Mira and M. Ali, editores, *Proc. of the 11th Int'l. Conf. on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA-98-AIE*, Lecture Notes in Artificial Intelligence 1416, Springer-Verlag, Berlin, páginas 290-299.
- Lee, D. (1996): *The Map-Building and Exploration Strategies of a Simple Sonar-Equipped Mobile Robot*. Cambridge University Press, Cambridge.
- Lumelsky, V.J., Mukhopadhyay, S. y Sun, K. (1990): Dynamic path planning in sensor-based terrain acquisition. *IEEE Trans. on Robotics and Automation*, 6(4):462-472.
- Martin, P. y Nehmzow, U. (1995): "Programming" by teaching: Neural network control in the Manchester mobile robot. En *Proc. of IFAC Conf. on Intelligent Autonomous Vehicles (IAV'95)*, páginas 277-282.
- Payton, D.W. (1991): Internalized Plans: A representation for action resources. En P. Maes, editor, *Designing Autonomous Robots*, páginas 89-103. MIT Press, Cambridge, Massachusetts.
- Shanon, C.E. (1952): Presentation of a maze-solving machine. *Trans. Eighth Cybernetics Conference*. (Disponible en C.E. Shannon, N.J.A. Sloane y A.D. Wyner, *Claude Elwood Shannon: Collected papers*, páginas 681-687. IEEE Computer Society Press, New-York, 1993.)