

PERCEPCIÓN ACTIVA PARA LA IDENTIFICACIÓN DE PANELES DE SALIDA DE EMERGENCIA

A. Astigarraga, E. Lazkano, B. Sierra, J. M. Martínez-Otzeta
Departamento de Ciencias de la Computación e Inteligencia Artificial
Facultad de Informática
Euskal Herriko Unibertsitatea
Paseo Manuel Lardizábal,1
C.P. 20018 Donostia-San Sebastián
E-mail: {ccbaspaa,ccplaore,ccpsiarb,ccbmaotj}@si.ehu.es

Resumen

La navegación basada en identificación de marcas visuales puede ser mejorada de manera sustancial mediante un ajuste dinámico de los parámetros de la cámara, a fin de seleccionar las regiones más interesantes de la escena. En este artículo se presenta un método para seleccionar de manera dinámica, basada en las lecturas de los sónares, los valores del zoom y la inclinación de la cámara. Los datos medidos desde diferentes distancias y con diferentes valores de los parámetros, así como la salida del Perceptrón Multicapa usado como reconocedor de marcas, son la base para la selección de las funciones de control de la cámara. Los resultados experimentales muestran que la metodología adoptada es adecuada para la tarea en cuestión, así como capaz de generalizar en nuevos entornos.

Palabras clave: Visión con propósito, Robótica basada en el Comportamiento, Marcas Visuales.

1. INTRODUCCIÓN

A pesar de los grandes progresos que se han realizado, la navegación de robots móviles es todavía un área

de investigación con muchos retos abiertos. Las capacidades navegacionales de muchos agentes biológicos (hormigas, abejas y pájaros, por ejemplo), han sido fuente de inspiración para los investigadores en robótica en su búsqueda de mecanismos susceptibles de ser implementados en robots autónomos móviles. Uno de estos enfoques biológicos para la navegación es la respuesta desencadenada por el reconocimiento, la cuál proporciona un medio para conectar diferentes localizaciones por medio de algoritmos de navegación local. Mediante la apropiada secuenciación de diferentes pares reconocimiento-respuesta, el robot adquiere la capacidad de seguir una ruta predefinida [13]. Se ha observado que las abejas utilizan este método de navegación para aprender sus rutas, es decir, asocian decisiones de movimiento con marcas visuales. Por lo tanto, las marcas en el entorno pueden ser usadas por el robot como pistas para identificar diferentes lugares, bien para situarse en el entorno, bien para desencadenar diferentes comportamientos. Aunque las marcas pueden ser de diferente naturaleza, los sistemas de visión proporcionan una gran fuente de información para la identificación de marcas menos sensibles al *aliasing*. En el caso de que el robot esté dotado con un sistema de visión, el problema consiste en determinar qué información

de la escena debería ser usada para optimizar la implementación de la tarea que el robot debe realizar. Aunque se trata de marcas diseñadas por los humanos, los paneles de salida de emergencia pueden ser considerados como marcas naturales internacionales obligatorias en todo edificio público, que deben cumplir ciertos estándares en cuanto a forma, color y situación (European “Council Directive 92/58/EEC of 24 June 1992 on the minimum requirements for the provision of safety and/or health signs at work”. Official Journal L 245 , 26/08/1992 P. 0023 - 0042). Por tanto, se pueden usar para implementar estrategias de navegación como la respuesta desencadenada por el reconocimiento. Nuestro enfoque consiste en usar los paneles de salida de emergencia para la navegación basada en marcas. En [12] usábamos los paneles junto con los pasillos como marcas para la navegación de un robot móvil B21, mediante una arquitectura de control basada en el comportamiento [4], desarrollada usando SORGIN, un marco software [2] especialmente diseñado para la construcción de módulos de comportamiento. El comportamiento emergente surgía de la cooperación entre los módulos que se orientaban según los pasillos y la brújula, y los subsistemas de identificación de marcas actuaban cambiando la orientación de brújula que el robot debía seguir dependiendo de las marcas identificadas en cada momento. La interacción entre el robot y su entorno conduce al robot a través de caminos nominales, pero según van apareciendo obstáculos dinámicos, los cambios en la trayectoria del robot distorsionan la visión del panel, complicando su identificación y haciendo que el robot pierda la señal de emergencia y falle en su tarea de navegación. Un comportamiento similar fue observado en abejas por Collet et al. [5]. Una vez éstas habían aprendido una secuencia de marcas, mostraban un comportamiento inconsistente al cambiarse el orden de las mismas. Nuestra investigación actual está motivada por la búsqueda de una mejora en el comportamiento de reconocimiento de los paneles de salida de emergencia, permitiendo la adaptación dinámica de los grados de libertad de la cámara.

Los procesos visuales están directamente relacionados con la tarea a realizar por el observador. La visión

animada o con propósito implica que la percepción está ligada a la acción y que la visión debe estar basada en expectativas [3, 1]. El conocimiento sobre el entorno puede, de alguna manera, condicionar la interpretación de lo que se está viendo. Ambas, la visión activa y, más recientemente, las capacidades de zooming, son utilizadas en el campo de la robótica móvil principalmente para fijar la atención del sistema visual en alguna característica del entorno [10, 18, 14]. Se ha comprobado que los sistemas biológicos usan visión activa en bucles percepción-acción, dando lugar en cierto sentido a un conjunto de trucos visuales.

El sistema de visión Cannon VCC4 montado en la parte superior de nuestro robot Pioneer3 está dotado con características de zoom, inclinación de cámara y ángulo panorámico. El robot es una máquina holónoma de direccionamiento dual, que ejecuta el paquete software Player [8] (desarrollado conjuntamente en USC Robotics Research Lab y HRL Labs, y disponible libremente de acuerdo con la licencia GNU en la dirección <http://playerstage.sourceforge.net>) para acceder a los sensores del robot (en realidad sólo un anillo frontal de sónares) y los actuadores. Por tanto, el sistema permite adaptar los parámetros de la cámara para ajustar el punto de vista según la tarea a llevar a cabo.

En este artículo presentamos el diseño e implementación de un comportamiento de posicionamiento activo de cámara, que explota las capacidades de los sensores del robot para cambiar sus parámetros de zoom e inclinación, a fin de identificar de manera fiable las señales de salida de emergencia que haya en la escena durante su tarea enfocada a la navegación.

En la sección 2 se describe el proceso de identificación de paneles de emergencia y los resultados obtenidos durante la trayectoria nominal a distintas distancias manteniendo fija la elevación de la cámara y también el zoom. La sección 3 describe los procesos seguidos para definir las funciones de ajuste de los parámetros. Los resultados experimentales que demuestran la mejora sustancial en el rendimiento del proceso de identificación se describen en la sección cuarta. Para terminar, la última sección incluye algunas conclusio-

nes junto con el trabajo futuro.

2. RECONOCIMIENTO DE PANELES DE SALIDA DE EMERGENCIA

A fin de distinguir el panel del fondo de la imagen, primero aplicamos un algoritmo de realce de color (proporcionado por la herramienta GNU de procesamiento de imágenes The Gimp [11]) a la imagen para obtener una mejor segmentación. Este algoritmo ejecuta una fuerza de saturación automática sobre los tres canales de la imagen dentro del espacio de color HSV [16], preservando el *hue*. Después de que la imagen ha sido realzada un simple filtrado por umbral es suficiente para segmentar de manera apropiada las áreas verdes de la imagen (ver figura 1).



(a)



(b)

Figura 1: Imagen original y segmentada

Una vez que la imagen ha sido segmentada necesitamos un método para distinguir imágenes que contienen un panel de las que no; en vez de usar funciones

de correlación de imágenes elegimos utilizar un perceptrón multicapa (MLP) [15] entrenado usando como entrada la suma cuadrática ponderada de bloques de tamaño 20×20 , dando lugar a 28 neuronas de entrada [12]. El conjunto de imágenes de entrenamiento contenía 320 imágenes, tomadas en la trayectoria nominal, y con valores constantes de zoom, inclinación y panorámica. La red tiene una única capa oculta de tres neuronas y la salida es un vector binario que denota la clase de la imagen como conteniendo o no una señal de emergencia. Por lo tanto, se aborda la identificación de paneles como un problema de clasificación supervisada. Aplicando la técnica de validación leaving-one-out (LOO), y después de un periodo de entrenamiento de 1000 ciclos por cada conjunto, obtuvimos una tasa de aciertos del 96,25%. Hemos mejorado el MLP buscando pesos apropiados para las entradas usando un *Algoritmo Genético* (GA) [9] y asociando a los individuos, como función de adaptación, la tasa de aciertos LOO del MLP, consiguiendo elevar la precisión hasta un 97,5%.

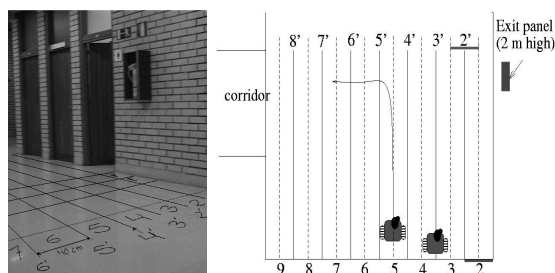
Para añadir robustez al proceso de reconocimiento de marcas, usamos un nivel de confianza (cl) proporcional al número de imágenes consecutivas identificadas positivamente. Esta cantidad afecta a la velocidad de traslación de acuerdo a la siguiente expresión: $v' = v(1 - cl)$. La motivación tras esta reducción de velocidad consiste en ralentizar al robot cuando un panel de emergencia está siendo reconocido para, de esa manera, no perderlo de vista.

El propósito principal de este proceso de identificación de marcas consiste en cambiar la orientación según la brújula que el robot debe seguir para llegar a cierto lugar. El robot Pioneer usado para estos experimentos no está dotado con un sensor de brújula, por lo que para simular el comportamiento de seguir la brújula utilizamos la odometría local. Después de inicializar los ejes de coordenadas a la posición inicial proporcionamos al robot, en vez de un ángulo relativo de la brújula, un desplazamiento angular deseado con respecto al eje X . De esta manera, el robot debe inicialmente ir hacia adelante y, cuando reconozca el panel, debe girar para seguir la orientación de brújula simulada.

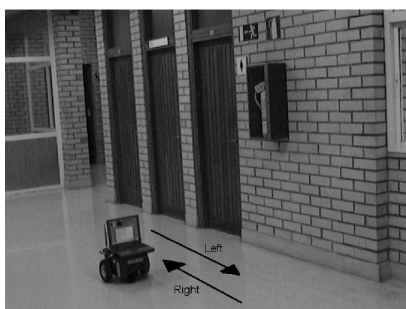
Para testear el comportamiento del MLP en el robot, pusimos al robot a diferentes distancias con respecto a la señal de emergencia y consideramos que el MLP fallaba cuando el robot perdía el panel y no giraba.

Cuadro 1: Rendimiento medio tras cinco experimentos desde cada posición

| | 4 | 4' | 5 | 5' | 6 | 6' | 7 | 7' |
|---------|---|-----|------------|------------|------------|----|----|----|
| Right % | 0 | 100 | 100 | 100 | 80 | 80 | 0 | 0 |
| Left % | 0 | 0 | 80 | 100 | 100 | 60 | 60 | 0 |



(a)



(b)

Figura 2: Entorno experimental

Se programó el robot para que siguiera el eje X cuando identificara la señal. El experimento se llevó a cabo en ambas direcciones, dejando la pared a la derecha o a la izquierda, como se muestra en la figura 2. Cada ejecución se repitió cinco veces. La tabla 1 muestra los resultados obtenidos usando valores fijos de zoom e inclinación, los mismos usados para capturar imágenes para el proceso de entrenamiento del MLP. Cada columna representa media baldosa del suelo (aproximadamente 20 cm), y las columnas en negrita representan el área de camino nominal.

3. AJUSTES DEL ZOOM Y LA INCLINACIÓN

La sección anterior mostró cómo el robot no ve el panel cuando éste cae fuera de su trayectoria nominal. Si el robot está muy cerca o muy lejos del panel, éste queda fuera del campo de visión o es demasiado pequeño para que el MLP lo reconozca de manera fiable. Obviamente, la adaptación dinámica de los valores de zoom e inclinación de la cámara debería llevar a una mejora del proceso de reconocimiento. Estudiando el comportamiento interno de los sistemas de visión activa, encontramos que muchos de ellos buscan relaciones geométricas entre la posición de las marcas y el robot [6, 14]. Los valores de zoom e inclinación pueden ser ajustados definiendo los valores adecuados para cada variable según la distancia a la pared, utilizando el anillo de sónicas que posee el robot. Pero las limitaciones de los sensores ultrasónicos son bien conocidas. Los valores dependen mucho de las propiedades de reflexión de las superficies. Además, la distancia según los sónicas tiende a sobreestimarse debido a reflexiones especulares y ecos perdidos durante la navegación [7]. En vez de imponer restricciones geométricas al sistema de percepción del robot, estimamos que es mejor adaptar esas relaciones geométricas según la percepción que tiene el robot del entorno. Debido a que el MLP es quien debe identificar los paneles de emergencia, confiamos en su proceso de identificación y obtenemos del mismo los valores apropiados para cada parámetro. En otras palabras, acomodamos los valores de zoom e inclinación de acuerdo con las necesidades del MLP. Las siguientes subsecciones describen estos dos procesos.

3.1. UNA APROXIMACIÓN LINEAL PARA EL AJUSTE DE LA INCLINACIÓN

Para establecer la relación entre el ángulo de inclinación de la cámara y la distancia a la pared, medida mediante los sónares del robot, simplemente pusimos el robot en diferentes posiciones y consideramos un ángulo de inclinación válido si existía un rango de zoom para el que el MLP era capaz de identificar la señal de una manera robusta. Aunque los datos recogidos no mostraron una relación lineal exacta entre los parámetros, aproximamos la curva usando una función lineal con valor mínimo acotado, según la ecuación 1. La figura 3 muestra el resultado de la recogida de datos, junto con la relación geométrica teórica y la aproximación lineal elegida (ecuación 1). Obsérvese la diferencia entre el modelo teórico y el modelo adquirido. Esta diferencia confirma cierta tendencia a que el panel de emergencia aparezca en la parte superior de la imagen, extraída por red neuronal de las imágenes de entrenamiento.

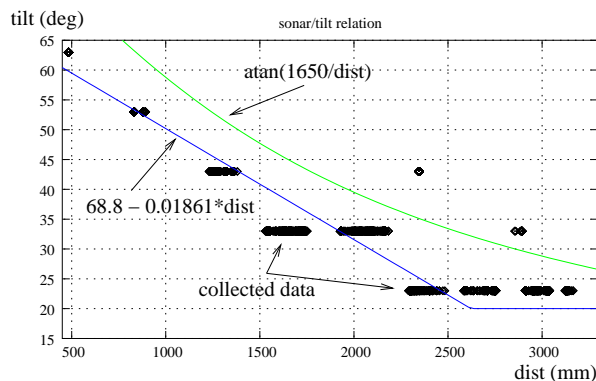


Figura 3: Datos recogidos, relación teórica y aproximación lineal obtenida a partir de los datos

$$tilt = \max(68,8 - 0,01861 * dist), 20) \quad (1)$$

3.2. AJUSTE DEL ZOOM MEDIANTE INTERPOLACIÓN UTILIZANDO UNA B-SPLINE UNIFORME

Los paneles de salida de emergencia son pequeños y el robot puede pasarlos de largo incluso aunque la cámara esté adecuadamente enfocada al área donde se supone que está el panel. Esto se ve reflejado en la figura 4, en la cuál podemos observar al robot enfocando el panel de salida de emergencia mediante el uso de un comportamiento de ajuste dinámico de la inclinación. Para identificar de manera fiable las señales de emergencia del resto de la escena, el sistema MLP necesita paneles de un cierto tamaño (véase la figura punteada en el lado derecho de la figura 4). Cuando el robot está demasiado lejos de la pared, la visión que proporciona la cámara de la señal de salida es demasiado pequeña para poder ser correctamente identificada por el MLP. Por otra parte, cuando el robot está muy cerca de la pared la imagen de emergencia ocupa toda la pantalla y el MLP también falla. El tamaño apropiado del panel puede ser obtenido dinámicamente adaptando el parámetro de zoom de la cámara según las necesidades del MLP.

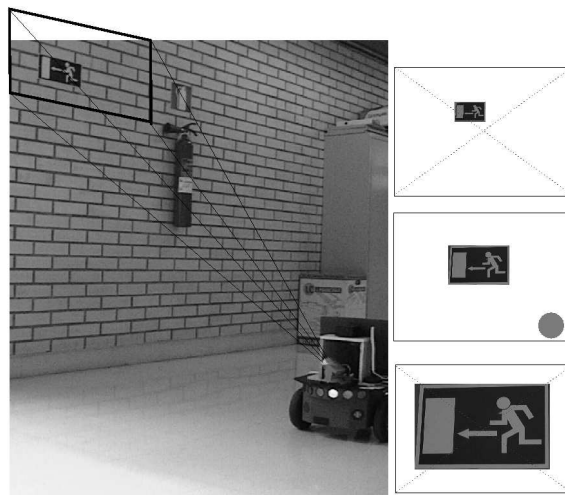


Figura 4: Influencia del valor de zoom en la percepción del panel

Para poder ajustar el valor del zoom, procedimos de la siguiente manera: mientras se ejecutaba el comportamiento de ajuste dinámico de la inclinación en el robot, recogimos datos para los valores del zoom de manera similar. Permitiendo al robot seleccionar el ángulo de inclinación y cambiando el valor del zoom usando intervalos uniformes (desde 0 hasta 1300, con un salto de 20), recogimos los resultados del MLP para cada valor del zoom y la distancia a la pared según el s3nar. Intuitivamente, debera haber una clara relaci3n entre la distancia de la c3mara a la se3al y el valor de zoom necesario para poder reconocerlo correctamente. Mayor distancia implica mayor zoom; y viceversa, menor zoom debera ser suficiente cuando el robot se acerca al panel. El an3lisis de los datos recogidos para el rango v3lido de zoom se muestra en la figura 5, y confirma que el rango v3lido de zoom disminuye con la distancia.

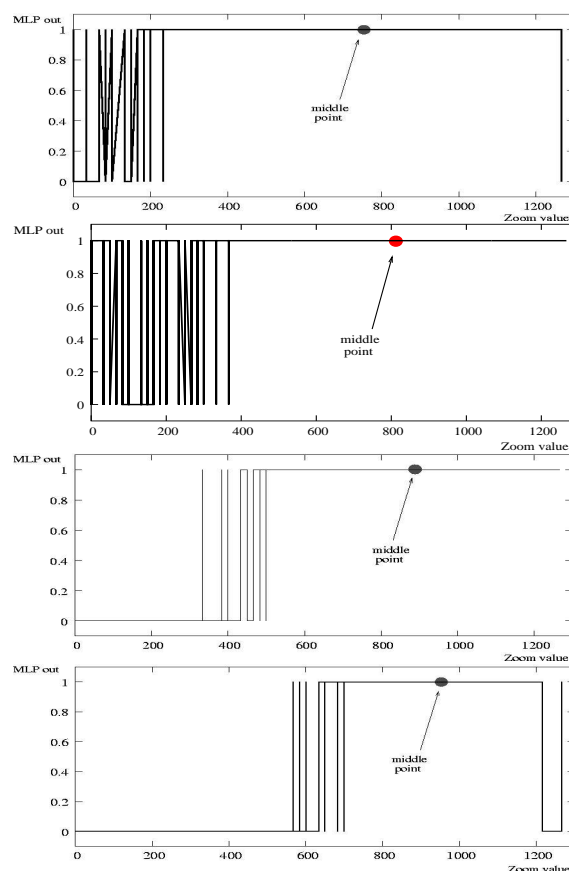


Figura 5: Respuesta del MLP ante diferentes valores de zoom a diferentes distancias: 1.5m, 2.5m y 3m

Usando como centroide la media del rango aceptable de zoom para cada distancia medida, interpolamos la curva usando una B-spline, una definici3n matem3tica param3trica de una serie de segmentos de curva conectados. Estas curvas pertenecen a la familia de las spline, usadas en la industria del dise3o asistido por ordenador y m3s ampliamente a3n en computaci3n de geometr3a 3D y modelado [17].

La selecci3n de los puntos medios del rango v3lido del zoom hace el comportamiento final m3s robusto porque permite un intervalo de confianza para lecturas con ruido. La figura 6 muestra la curva que se obtiene. El panel se sit3a en la pared a dos metros de altura. Debido al peque3o tama3o del robot Pioneer, seg3n se acerca a la pared la perspectiva que obtiene se distorsiona. Esto explica por qu3 el valor del zoom se incrementa cuando las distancias son muy peque3as. Obs3rvese tambi3n la llanura entre 1.5m y 2m. Este efecto es debido a que la red fue entrenada con im3genes capturadas a esa distancia de la pared, aproximadamente desde el centro del pasillo.

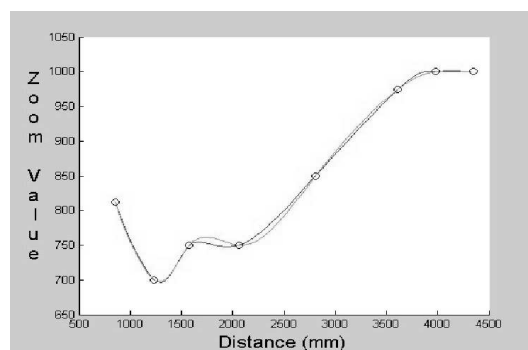


Figura 6: Curva de relaci3n entre la distancia medida por los ultrasonidos y el valor de zoom

4. RESULTADOS EXPERIMENTALES

Se realizaron dos conjuntos de experimentos a fin de testear el comportamiento del enfoque presentado en lo que respecta al reposicionamiento dinámico de los parámetros de zoom e inclinación de la cámara. En primer lugar pusimos el robot en el entorno en el que recogimos los datos para el proceso de ajuste del zoom y la inclinación, como en la sección 2, haciéndole seguir el eje X cuando la señal es identificada en ambas direcciones. El segundo conjunto de experimentos consistió en situar al robot en un entorno completamente nuevo, y de nuevo distinguir el lado en el que estaba situado el panel. Las tablas 2 y 3 muestran los resultados de los experimentos, obtenidos en el primer y segundo entorno, respectivamente.

Cuadro 2: Resultados experimentales en el primer entorno

| | | | | | | | | |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|
| | 2 | 2' | 3 | 3' | 4 | 4' | 5 | 5' |
| Right % | 60 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Left % | 0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | 6 | 6' | 7 | 7' | 8 | 8' | 9 | |
| Right % | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |
| Left % | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |

Cuadro 3: Capacidad de generalización del sistema

| | | | | | |
|---------|-----|-----|-----|-----|-----|
| | 2 | 3 | 4 | 5 | 6 |
| Right % | 100 | 100 | 100 | 100 | 100 |
| Left % | 100 | 100 | 100 | 100 | 100 |
| | 7 | 8 | 9 | 10 | 11 |
| Right % | 100 | 100 | 100 | 100 | 100 |
| Left % | 100 | 100 | 100 | 100 | 100 |

Se debe resaltar el hecho de que la fase de experimentación consistió en 250 ejecuciones y sólo 7 fallaron, dando una tasa de aciertos del 97,8%. También queremos hacer notar que el sistema de reconocimiento de marcas sólo falló en la baldosa número dos, debido a las dificultades inherentes al robot para obtener una buena perspectiva del panel cuando se halla cerca de la pared, dada su morfología.

5. CONCLUSIONES Y TRABAJO FUTURO

El proyecto principal en el que estamos inmersos es el diseño e implementación de una arquitectura robusta basada en el comportamiento, para ser utilizada en la navegación de robots móviles. Ya mostramos en [12] la viabilidad de la navegación basada en marcas usando señales de salida de emergencia y pasillos como marcas de un sistema de navegación biomimética más complejo. Este artículo viene a mejorar el sistema de reconocimiento de paneles de salida, mostrando cómo la percepción del entorno por el robot mediante un sistema MLP puede ser usado para implementar un control activo de cámara. Con ese proceso activo es posible detectar paneles de salida de emergencia incluso aunque el robot no pase a través de su trayectoria nominal. Esto enfatiza el principio de que la visión no debe funcionar de manera aislada, sino como parte de un sistema que interactúa con su entorno para llevar a cabo su tarea.

Actualmente, este módulo de comportamiento está ya integrado en la arquitectura de control global, y permite al robot identificar los paneles aún cuando se desvía de su trayectoria nominal. Como trabajo futuro, se propone adecuar el ángulo horizontal de la cámara de forma similar, para que la orientación del robot con respecto de la pared no perturbe el proceso de identificación.

Agradecimientos

Queremos dar las gracias al Catedrático Man Graña por el desinteresado préstamo de su robot Pioneer3. Este trabajo ha sido subvencionado por los proyectos OF761/2003 de la Gipuzkoako Foru Aldundia y 1/UPV00140.226-E-15412/2003 de la Euskal Herriko Unibertsitatea.

Referencias

- [1] J. Aloimonos. *Active Perception*. Lawrence Erlbaum Assoc Inc, 1993.

- [2] A. Astigarraga, E. Lazkano, I. Rañó, B. Sierra, and I. Zarautz. SORGIN: a software framework for behavior control implementation. In *CSCS14*, volume 1, pages 243–248, 2003.
- [3] D. H. Ballard. Animate vision. *Artificial Intelligence*, 48:57–86, 1991.
- [4] R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of robotics and automation*, RA-26:14–23, 1986.
- [5] T. S. Collett, S.Ñ. Fry, R. Wehner. Sequence learning by honeybees. *Journal of computational Physiology*, 172:693–706, 1993.
- [6] A. J. Davison. *Mobile Robot Navigation Using Active Vision*. PhD thesis, University of Oxford, 1999.
- [7] H. R. Everett. *Sensors for Mobile Robots. Theory and Applications*. A. K. Peters, Ltd., 1995.
- [8] B. P. Gerkey, R. T. Vaughan, and A. Howard. The Player/Stage project: tools for multi-robot and distributed sensor systems. In *Proc. of the International Conference on Advanced Robotics (ICAR)*, pages 317–323, 2003.
- [9] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.
- [10] E. Hayman, T. Thórhallsson, and D. Murray. Tracking while zooming using affine transfer and multifocal tensors. *International Journal of Computer Vision*, 51, 2002.
- [11] S. Kimball and P. Mattis. The GIMP: GNU image manipulation program, 1997. <http://www.xcf.berkeley.edu/~gimp/gimp.html>.
- [12] E. Lazkano, A. Astigarraga, B. Sierra, and I. Rañó. On the adequateness of emergency exit panel and corridor identification as pilot scheme for a mobile robot. In *Intelligent Autonomous Systems 8*, volume 1, pages 915–924, 2004.
- [13] H. A. Mallot and M. A. Franz. Biomimetic robot navigation. *Robotics and Autonomous System*, 30:133–153, 2000.
- [14] F. Michaud. Teaching a robot how to read symbols. In J. P. Müller, E. Andre, S. Sen, and C. Frasson, editors, *Proceedings of the Fifth International Conference on Autonomous Agents*, pages 184–185, Montreal, Canada, 2001. ACM Press.
- [15] T. Mitchell. *Machine Learning*. McGraw-Hill, 1997.
- [16] J. C. Russ. *The image processing handbook (2nd edition)*. IEEE Press, 1995.
- [17] W. Tiller L. A. Piegl. *The NURBS Book*. Springer Verlag, 1997.
- [18] B. Tordoff. *Active Control of Zoom for Computer Vision*. PhD thesis, University of Oxford, 2002.