

Detección de puertas mediante información multisensorial y su aplicación a la navegación de robots móviles *

Eugenio Aguirre, Miguel García-Silvente, Antonio González, Rafael Muñoz,
Departamento de Ciencias de la Computación e Inteligencia Artificial
E.T.S. de Ingeniería Informática. Universidad de Granada
18071-Granada tlfno.: +34.958.240807 fax: +34.958.243317
e-mail: {eaguirre, mgs, gonzalez, salinas}@decsai.ugr.es

Resumen

En este trabajo se presenta un nuevo método para llevar a cabo la detección visual de las puertas en un entorno de interiores, teniendo en cuenta distintas condiciones ambientales de iluminación y puntos de vista con diferentes perspectivas. Para ello se utiliza la transformada de Hough y se estudian las relaciones espaciales entre las líneas detectadas. El método es robusto frente a la deformación producida por la perspectiva, a cambios de escala y a oclusión parcial. Una vez detectada visualmente la posible puerta el robot es capaz de dirigirse hacia ella y utilizar un modelo perceptual basado en ultrasonidos para confirmar si el objeto es una puerta abierta o no. La implementación se ha realizado para un robot móvil Nomad 200 mediante un sistema multi-agente capaz de llevar a cabo una navegación segura en entorno de interiores utilizando una serie de comportamientos difusos y un mapa topológico del entorno.

Palabras clave: Robots móviles, Visión artificial, Integración multisensorial

1. Introducción

La percepción del entorno constituye uno de los principales problemas a los que hay que hacer frente cuando se plantea el diseño e implementación de agentes físicos típicos como los robots móviles. Mediante un sistema de sensores, que pueden ser de diferente naturaleza, los robots obtienen información del entorno. Esta información además suele estar afectada por ruido debido a la propia imperfección de los sensores y está sometida a los cambios de las condiciones ambientales propias de los entornos del mundo real. La información obtenida se utiliza para controlar el robot, estableciendo una relación percepción-acción, según los objetivos a conseguir, y para crear modelos del mundo a partir de los cuales se puede razonar sobre la relación del robot con el mundo. En este sentido, en trabajos anteriores hemos desarrollado un modelo de percepción para sensores de ultrasonidos [3] que permite tanto la construcción de mapas topológicos del entorno como la navegación basada en comportamientos difusos usando la información representada en dichos mapas. Por otra parte en [1] se desarrolló un sistema multi-agente que combinaba agentes de navegación que utilizaban información de sensores de ultrasonidos con otros que utilizaban solamente información visual. A partir de la experiencia obtenida en estos trabajos previos, resulta de sumo interés, el poder combinar la información procedente de varios sistemas sensoriales, tanto para controlar el movimiento del robot como para construir posibles modelos de representación del entorno, ya que unos sistemas sensoriales pueden suplir la carencia de otros. Es más, también interesa introducir un componente ajeno a los sistemas de sensores como es la propia capacidad de movimiento del robot. Es decir, el agente debe ejercer un papel activo en el proceso de percepción, lo que nos permite verificar hipótesis obtenidas por el sistema perceptual y facilitar también el tratamiento de la información percibida.

Este trabajo es un paso más en el proceso de la integración de la información visual con el sistema de navegación basado en comportamientos difusos ya desarrollado [2]. En concreto, nos centraremos en la detección de las

*Este trabajo ha sido parcialmente financiado por el proyecto TIC2003-04900 del MCYT

puertas de un entorno de interiores sin la inclusión de ninguna marca artificial en el entorno. La ventaja que ofrece la integración es la confrontación de hipótesis. Por sí solo, el sonar es capaz de detectar puertas como huecos en paredes de habitaciones o pasillos. Estos huecos en determinadas condiciones pueden ser clasificados erróneamente como puertas. Un caso típico, es cuando el robot detecta mediante los ultrasonidos una zona de paso estrecha entre el mobiliario de una habitación. En este caso la situación perceptual detectada por los ultrasonidos es muy parecida a la producida por la presencia de una puerta abierta. Una forma de evitar este problema sería contar con un mapa del entorno donde se tenga una idea aproximada de la posición de la puerta en la habitación, y tener un conocimiento aproximado de la posición del robot. En este trabajo, la navegación que realiza el robot se fundamenta en mapas de tipo cualitativo [10, 11] y comportamientos con capacidades reactivas por lo que no va a ser preciso conocer con exactitud la posición del robot ni utilizar un sistema métrico de coordenadas en el entorno. A partir de información visual el robot podrá extraer unas hipótesis que serán confrontadas con las hipótesis del modelo de percepción basado en ultrasonidos. En cualquier caso, el robot utilizará su propio movimiento para dirigirse a la puerta y poder incrementar su creencia en la posible existencia de la puerta, o bien descartarla, recogiendo nuevos datos desde su nueva posición.

Centrándonos en el problema concreto de la detección visual de las puertas, en la literatura se pueden encontrar algunos enfoques para la resolución de este problema. El trabajo desarrollado por [6] presenta una técnica basada en redes neuronales para la detección de puertas a partir de sus componentes. Su sistema se compone de dos redes neuronales, una para la detección de las esquinas de la puerta y otra para la detección de los marcos lateral y superior. Cada red analiza para cada pixel subventanas de tamaño 18x18 y decide si en ella hay presente una esquina de una posible puerta, un marco o ninguno de estos elementos. La entrada de la red son las componentes de tono y saturación de las subventanas. Cada red tiene un total de 648 neuronas de entrada, una de salida y una capa intermedia. Tras clasificar cada pixel en la imagen, se realiza un proceso de análisis de componentes. Consideran que en una región hay una esquina si el número total de pixels clasificados como tal supera un umbral. Para detectar los marcos laterales se realiza un proceso similar y tras esto se analiza si los elementos detectados se combinan de forma adecuada para formar una puerta. El sistema es capaz de detectar la puertas en condiciones de oclusión parcial y desde distintos ángulos, pero presenta el problema de requerir un alto coste computacional.

En [9] se presenta un método de localización de objetos genéricos para la navegación de robots. Utilizan un sistema trinocular de visión con el que detectan segmentos en imágenes. Los segmentos son emparejados usando la información proveniente de las tres cámaras. La detección de la puerta se realiza buscando un conjunto de segmentos con forma rectangular. Sobre el posible conjunto de candidatos, se analiza si cumplen unas restricciones de altura y anchura estimada de las puertas de su entorno. Gracias al sistema trinocular son capaces de conocer la distancia entre los segmentos para determinar si realmente cumplen las restricciones impuestas. Este sistema cuenta con el inconveniente del coste del equipo perceptual. Además el sistema no garantiza la efectividad en el reconocimiento puesto que otros objetos (por ejemplo armarios) pueden ser confundidos con puertas.

En [13] se emplea un método para la detección de puertas en pasillos. En este trabajo se limita la detección de puertas a pasillos en entornos de interiores y se realiza teniendo en cuenta las dimensiones del mismo. En primer lugar se captura una imagen del pasillo y se resaltan los bordes, tras una dilatación seguida de una erosión se detectan bandas verticales. Teniendo en cuenta la distancia y dirección de las paredes del pasillo respecto del robot, se detectan las posibles puertas basándose en las dimensiones esperadas de las mismas. Sin embargo no queda muy claro cómo son clasificadas como puertas las bandas verticales detectadas en la imagen. Además la técnica limita la detección de las puertas a pasillos y no tiene en cuenta las deformaciones debidas a los cambios de perspectiva.

En la aproximación llevada a cabo por [12] se hace uso de la detección de puertas para ayudar a un robot móvil a atravesarla. La detección de la puerta se hace basándose en los marcos de la misma. Tras capturar una imagen del entorno se resaltan bordes y tras esto se realiza una dilatación. Sobre esta imagen se aplica un algoritmo para realzar las columnas y se asume que hay una puerta si existe una columna más ancha de 35 pixels. El principal problema que presenta este método es que es fácil que confunda una puerta con otro objeto del entorno.

Respecto a la integración multisensorial para la detección de puertas también encontramos trabajos anteriores en la bibliografía como el trabajo de Cañas y otros [4]. En este caso se plantea un sistema de percepción activa que integra visión con sonars para detectar la existencia y posición de puertas en el entorno del agente. Estos autores utilizan un enfoque probabilístico para llevar a cabo la fusión de información visual y de sensores de rango. En su caso, la detección visual de las puertas se realiza a partir de la detección de las líneas verticales del marco mediante

un tratamiento visual sencillo mientras que nosotros proponemos un método que tiene en cuenta el marco completo y realiza un proceso más elaborado de la información visual.

En este trabajo presentamos una nueva aproximación para localizar puertas en entornos de interior. Las puertas son detectadas en una imagen estática captada por la cámara del robot, agrupando las líneas que forman los marcos laterales y superior de las mismas. Para detectar las líneas se hace uso de la transformada de Hough [7] sobre los bordes extraídos de la imagen. Estas líneas son analizadas para determinar si se cortan en la imagen formando una posible puerta. El método es capaz de detectar la existencia de puertas en entornos de iluminación variable usando una sola cámara. Es robusto frente a la deformación debida a la perspectiva, a cambios de escala y a oclusión parcial. Una vez detectada visualmente una puerta, el robot se dirigirá hacia ella para confirmar su existencia. En este proceso se plantea como un objetivo que el robot no pierda la localización de la posible puerta por los movimientos al dirigirse hacia ella. Para ello la visión adoptará un papel activo, y se pretende que el sistema oriente la cámara de forma que la puerta sea siempre visible. Una vez que el robot esté cerca de la posible puerta se utilizará un modelo perceptual basado en ultrasonidos para confirmar la presencia del hueco en la puerta y poder concluir así que existe una puerta abierta en el entorno o en caso contrario, descartar esa posibilidad. Es importante remarcar que el trabajo que aquí presentamos está aún en desarrollo. Nos centraremos en explicar cómo se realiza el reconocimiento visual de las puertas analizando tanto información proveniente de la transformada de Hough como de un detector de fronteras. Comentaremos brevemente las ideas que seguiremos para llevar a cabo la integración con el sistema de ultrasonidos que nos permitirá confirmar si efectivamente el objeto visualmente detectado es una puerta abierta.

Nuestro robot es un robot móvil Nomad 200 al que se le ha añadido un sistema de procesamiento de información visual. Así, cuenta con una cabeza robótica (PTU) sobre la que está montada la cámara y que tiene movimientos independientes de los del robot.

A continuación se comenta brevemente la arquitectura del sistema multi-agente en el que se encuentra incluido el agente de detección de puertas. En la Sección 3 se describe con detalle el proceso de detección visual de las puertas y se comenta brevemente al final cómo se incorpora la información del sistema de ultrasonidos. La Sección 4 muestra algunos resultados concretos y finalmente se presentan algunas conclusiones y aspectos a desarrollar en un futuro.

2. Arquitectura del Sistema Multi-Agente

La arquitectura de navegación utilizada en este trabajo sigue la filosofía del sistema multi-agente presentado en [1] y que se resume brevemente a continuación. El sistema permite la navegación basada en comportamientos difusos utilizando un mapa topológico. En este mapa están representados como nodos los pasillos y habitaciones y como arcos las puertas que los unen. En el sistema multi-agente existen agentes tanto de tipo reactivo como deliberativo y se ha desarrollado de la forma más modular posible para facilitar su posterior expansión ya sea mediante la adición de nuevas competencias a un agente concreto o bien mediante la adición de nuevos agentes. El conjunto de agentes diseñado es el siguiente:

- **Planner:** Utilizando el mapa topológico calcula el camino óptimo que une el origen y el destino. Realiza también la ejecución y monitorización del plan controlando al resto de agentes.
- **Navigation:** Utiliza diversos comportamientos difusos cuya activación le permite llevar a cabo una navegación segura.
- **Nomad200:** Su función es enviar las órdenes al robot y recibir los datos de los sensores. Hace de interfaz entre el robot y el resto de agentes.
- **PTU:** Este agente se utiliza para controlar tanto el giro horizontal como vertical de la PTU. Ello hace posible que el resto de los agentes puedan controlar la dirección hacia la que apunta la cámara.
- **Door Finder:** Realiza el procesamiento sensorial necesario para detectar la puerta y mantener la cámara enfocada hacia ella siempre que las condiciones del entorno lo permitan.

En la Figura 1, se muestra el conjunto de agentes empleado. Las líneas que unen los agentes implican que hay algún tipo de comunicación entre ellos o prestación de servicios de unos a otros. A continuación, nos centraremos en describir solamente en profundidad el agente Door Finder para la detección de puertas.

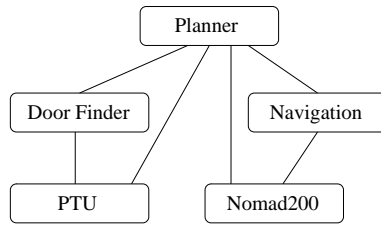


Figura 1: Arquitectura del sistema multi-agente.

3. Diseño del agente Door Finder para la detección de puertas

En esta sección nos centramos en el problema de la creación de un agente capaz de detectar puertas sin incluir ninguna marca artificial en el entorno. El objetivo de este agente puede enunciarse como: *Detectar una puerta en el entorno del robot y mantener la cámara apuntando hacia ella pese a los movimientos del mismo*. Con este objetivo podemos descomponer las subtareas del agente como: a) detectar una posible puerta b) mantenerla siempre localizada en la imagen pese a los movimientos del robot.

En la primera tarea, que será la que expliquemos, se usarán algoritmos para la detección que se mostrarán más adelante. Una vez localizada la posible puerta en la imagen, podrán aplicarse otros algoritmos de menor costo computacional para realizar el fijado, que seguirán la misma filosofía de los empleados en [1]. Esto nos permitirá obtener un comportamiento que reaccione con mayor velocidad.

El principal problema que presenta la detección de objetos en entornos no controlados proviene de la iluminación variable, esto es, en una misma imagen la iluminación no será constante en todas las zonas debido a las fuentes de luz. Además hemos de tener en cuenta que la cámara que estamos utilizando realiza un proceso de balance de blancos, lo que provoca que pequeños cambios en la posición de la cámara puedan provocar cambios bruscos en la iluminación de la escena. Bajo estas condiciones el uso de un detector de fronteras, ha presentado en nuestras pruebas mayor robustez que el uso de algoritmos de segmentación adaptativa. Nuestro objetivo es detectar los marcos de la puerta y para ello lo que haremos es detectar las líneas de los marcos en la imágenes tomadas con la cámara.

Para detectar las líneas se utilizará la transformada de Hough [7] sobre la imagen resultado de un extractor de fronteras. La transformada de Hough nos indica en el espacio de parámetros los votos recibidos por cada línea en la imagen. Una línea estará representada por el par $(\rho, \phi) \in [-\sqrt{2}N, \sqrt{2}N] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \subset \mathcal{R}$. Dónde N es el tamaño de la imagen analizada.

Definiremos el objeto puerta como un par de rectas verticales con una recta horizontal en la parte superior. Hemos de tener en cuenta que, por efectos de la proyección, las líneas verticales no lo serán desde todos los puntos de vista. Tampoco las horizontales aparecerán como tales en su proyección en la imagen.

Para poder determinar si los puntos de una línea horizontal están entre dos líneas verticales, será necesario analizar sus puntos en la imagen. Por tanto, el análisis implica usar tanto información del espacio de parámetros como del resultado de la extracción de bordes. El algoritmo que se explica a continuación permite la detección de puertas en imágenes salvando estos problemas.

El proceso se divide en fases. En primer lugar se estiman los bordes de la imagen. Después se detectan aquellas líneas verticales que ocupen un alto porcentaje de la imagen. Tras esto se detectan aquellas líneas horizontales que pueden ser parte de un marco superior. Después, se analizará cada pareja de líneas verticales con cada línea horizontal para determinar: a) si los puntos de la horizontal están en la imagen entre las verticales y b) si los límites de las líneas están unidos formando esquinas. A continuación se explica el algoritmo para la detección de las puertas en detalle.

1. Estimar los bordes de la imagen tomada del entorno, obteniendo un conjunto de puntos frontera $I(p)$. Denotaremos $p = (p_x, p_y)$ a cada uno de estos puntos. Dicha imagen se supone cuadrada de tamaño $N \times N$ para

facilitar la notación. Para realizar la extracción de bordes hemos utilizado el detector de fronteras de Canny [5].

2. Aplicar la transformada de Hough obteniéndose la matriz HT donde se almacenan los votos de los puntos de $I(p)$ para cada línea. El elemento $HT(\rho^i, \phi^i)$ representa los votos recibidos por la línea L^i de coordenadas (ρ^i, ϕ^i) en el espacio de parámetros. Debido a la discretización del espacio de parámetros, $HT(\rho^i, \phi^i)$ no representa solamente la línea L^i sino a todo un conjunto de líneas próximas. Por esta razón podrían votar por ella puntos que no corresponden a la línea exacta L^i . Para evitar esto, en el proceso de votación no se agregarán votos de líneas cercanas. Todos los puntos que votaron por la recta L^i en la imagen $I(p)$ se almacenan y se denotan como $PL^i = \{p_1^i, p_2^i \dots p_n^i\}$ siendo $p_j^i = (p_{j,x}^i, p_{j,y}^i)$ el punto en $I(p)$. Definimos $|PL^i|$ como el cardinal del conjunto PL^i .
3. Seleccionar en el espacio de parámetros las n líneas verticales más interesantes. Queremos detectar las posibles líneas pertenecientes a los marcos laterales de las puertas. Es decir, aquellas que por sus propiedades son más propensas a pertenecer a los marcos de una puerta. Una línea se considera interesante o no teniendo en cuenta varios criterios. Si una línea cumple los criterios que se explicarán a continuación se dice que cumple la propiedad *LineaVertical*.

a) En primer lugar el ángulo ϕ en el espacio de parámetros debe representar una línea de dirección vertical en la imagen. La línea completamente vertical en el espacio de parámetros es aquella con $\phi = 0$. Pero para dotar de un mayor grado de flexibilidad se permitirá seleccionar líneas con cierta inclinación.

b) En segundo lugar, dado que las puertas son objetos grandes en el entorno de oficina, deben aparecer como líneas largas en las imágenes. Por lo tanto se exigirá un tamaño mínimo que se analiza con el número de puntos que votó por ella.

c) Por otro lado, las líneas de los marcos son líneas continuas, es decir se mostrarán como grandes segmentos en la imagen. Por ello definiremos una medida de continuidad de una línea en función del número de huecos que tiene en la imagen.

$$Cont(L^i) = \frac{NH^i}{|PL^i|} \quad (1)$$

donde NH^i es el número de huecos de la línea L^i . Esta función toma valor 0 cuando no hay ningún hueco y tiende a 0'5 conforme aumenta el número de huecos. El valor máximo es 0'5 y se alcanza cuando hay dos puntos y un hueco.

Una vez definidos todos estos conceptos se puede definir la propiedad *LineaVertical*. Diremos que una línea L^i es *LineaVertical* si y solo si cumple:

$$\phi_i \in [-\phi_1, \phi_1] \wedge \frac{|PL^i|}{N} > \sigma_1 \wedge Cont(L_i) < \sigma_2 \quad (2)$$

donde $\phi_1 \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ y $\sigma_1 \in [0, 1]$ $\sigma_2 \in [0, 0'5]$. El primer elemento de la Ecuación 2 acota el ángulo de las líneas verticales. No sólo se tratan aquellas líneas que son completamente verticales, sino también aquellas con cierto grado de inclinación. El segundo parámetro mide el tamaño de la línea en la imagen. Y el último acota la continuidad de la línea. Es posible que al analizar las líneas en el intervalo $[-\phi_1, \phi_1]$ haya más de n líneas interesantes. Entonces deberán elegirse aquellas que tenga un mayor número de votos.

4. Seleccionar en el espacio de parámetros aquellas m líneas horizontales más interesantes. Este concepto es más difícil de aproximar ya que no podremos acotar tanto el espacio de búsqueda como en el caso anterior. Estamos buscando aquellas líneas horizontales que pertenezcan a la zona superior del marco de una puerta. El principal problema radica en que si no tenemos información sobre la posición relativa del observador a la puerta, no es posible conocer a priori la dirección que tiene esta línea. Si la vemos desde una posición completamente perpendicular esta línea será horizontal respecto de la imagen. Pero conforme nos alejamos, en una dirección no perpendicular, las direcciones que toma esta línea varían en gran medida. Por ello el margen angular que definiremos para buscar posibles líneas horizontales tiene que ser mayor. Igual ocurre con el tamaño, no es posible definir a priori cuál será debido a la perspectiva. Pero sí sabemos que debe ser menor que el de las líneas verticales de los marcos laterales. Una característica que sí es discriminante es que estas líneas están sobre las verticales, es decir, deberán estar en las zonas superiores de la imagen. Por tanto, solo serán válidas aquellas líneas que cumplan que la media en el eje y de los puntos que votaron por ella esté en la zona superior de la imagen. Finalmente, para acotar aún más el espacio de búsqueda, se

impondrá que las líneas cumplan un valor de continuidad mínimo que evite seleccionar líneas formadas por puntos dispersos en la imagen.

Definiremos pues la propiedad *LineaHorizontal*. Diremos que una línea L^i es *LineaHorizontal* si y sólo si cumple:

$$\phi_i \in \left[-\frac{\pi}{2}, -\frac{\pi}{2} + \phi_2\right] \cup \left[\frac{\pi}{2} - \phi_2, \frac{\pi}{2}\right] \wedge Cont(L_i) < \sigma_3 \wedge \frac{1}{|PL^i|} \sum_{j=0}^{|PL^i|} p_{j,y}^i < \sigma_4 N \quad (3)$$

donde $\phi_2 \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ y $\sigma_3 \in [0, 0.5]$ y $\sigma_4 \in [0, 1]$. El primer término de la ecuación acota el conjunto de líneas a las horizontales con cierto grado de flexibilidad. El segundo impone de nuevo el criterio de continuidad y el último especifica que las líneas horizontales deben estar localizadas en la zona superior de la imagen, por encima de $\sigma_4 N$.

Como en el caso de las líneas verticales, habrá más de m líneas que cumplan esta propiedad, nos quedaremos con aquellas que tengan un mayor número de votos.

5. Determinar los marcos. El objetivo ahora es determinar cuáles de las líneas anteriormente seleccionadas forman marcos. Para ello analizaremos el trío formado por dos líneas verticales y una horizontal. Debemos analizar todas las posibles combinaciones entre las líneas seleccionadas en los puntos 3 y 4.

Para determinar si este trío forma un marco impondremos una serie de condiciones. En primer lugar será necesario que la línea horizontal tenga un alto número de puntos entre las verticales. También exigiremos que las líneas verticales tengan un gran porcentaje de sus puntos bajo la intersección con la horizontal. Con esto evitaremos tomar líneas horizontales que no corten por la zona superior a las verticales. Y por último, exigiremos que la distancia entre los puntos de cada línea más cercanos a las intersecciones cumplan un criterio de distancia entre sí. Con ello se pretende seleccionar aquellas líneas que estén unidas formando esquinas. Si un trío cumple todas las condiciones impuestas se dice que tiene la propiedad *marco*.

A continuación explicaremos como se define la propiedad *marco* a partir de los puntos de las líneas y de sus intersecciones en la imagen. En la Figura 2(a) se muestra un esquema con dos líneas verticales (L^{left} y L^{right}) y una horizontal (L^{sup}). Las intersecciones entre ambas se muestran en la figura y las llamaremos IL e IR . En la figura están indicados los puntos de L^{left} y L^{right} más cercanos a la intersección con L^{sup} (L y R respectivamente). De la misma forma $S1$, $S2$ son los puntos de L^{sup} más cercanos a las intersecciones.

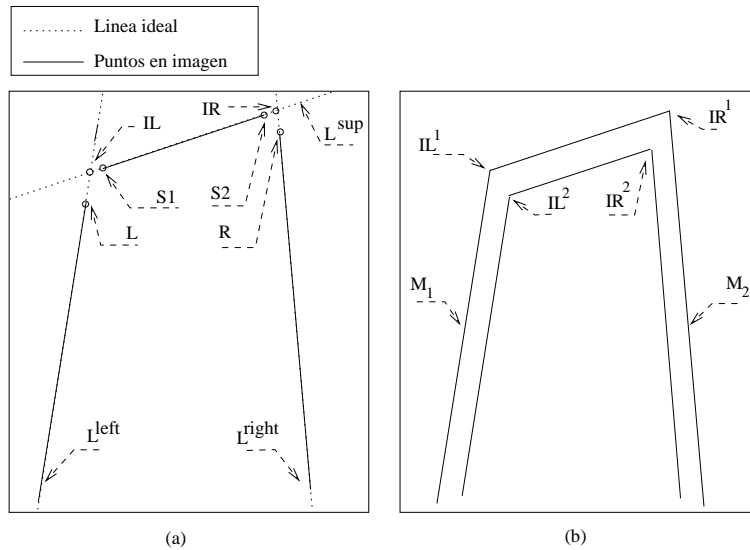


Figura 2: Trío de líneas en la imagen. (a) Puntos necesarios para determinar la propiedad marco (b) Esquema de marco contenido dentro de otro marco

Una vez mostrados los puntos que se requieren para definir la propiedad *marco*, explicaremos como la hemos definido. Sean las líneas seleccionadas L^{left} , L^{sup} , L^{right} . Siendo L^{sup} la línea horizontal y L^{left} , L^{right} las verticales. Comenzamos determinando la intersección entre las líneas $Int(L^{left}, L^{sup}) = IL$ siendo

$IL = (IL_x, IL_y)$. Igualmente $Int(L^{right}, L^{sup}) = IR$ donde $IR = (IR_x, IR_y)$. Asumiremos que las líneas han sido nombradas de forma que $IL_x < IR_x/IL_x \in L^{left} \wedge IR_x \in L^{right}$.

Calculamos el conjunto de puntos de L^{sup} que está entre las intersecciones y el conjunto de puntos de L^{left} y L^{right} que está bajo las intersecciones.

$$\begin{aligned} C^{sup} &= \{p_j^{sup} \in PL^{sup}/p_{j,x}^{sup} \in (IL_x, IR_x)\} \\ C^{left} &= \{p_j^{left} \in PL^{left}/p_{j,x}^{left} > IL_y\} \\ C^{right} &= \{p_j^{right} \in PL^{right}/p_{j,x}^{right} > IR_y\} \end{aligned}$$

Obteniendo los conjuntos $C^{sup} = \{c_0^{sup}, \dots, c_n^{sup}\}$, $C^{left} = \{c_0^{left}, \dots, c_n^{left}\}$ y $C^{right} = \{c_0^{right}, \dots, c_n^{right}\}$. Ordenamos los puntos de C^{left} y C^{right} de forma ascendente por la componente y . De forma análoga se ordena C^{sup} por la componente x . También determinamos el punto de cada línea vertical que está más cerca del punto de intersección con la horizontal. Esta información es fácilmente determinable con los conjuntos ordenados como:

$$L = c_0^{left}; \quad R = c_0^{right}$$

y de forma análoga los puntos de la línea horizontal más cercanos a las intersecciones.

$$S1 = c_0^{sup}; \quad S2 = c_n^{sup}$$

La utilidad que tienen estos puntos es poder medir la distancia que hay entre el punto más superior de una línea vertical con el punto más cercano de la línea horizontal. Si esta distancia es muy grande es por que ambas líneas están muy separadas en la imagen.

Podemos pues definir la propiedad *marco*. Se dice que el trío $L_{left}, L_{sup}, L_{right}$ tienen la propiedad marco si y sólo si cumple:

$$\frac{|C^{left}|}{|L^{left}|} > \sigma_5 \wedge \frac{|C^{right}|}{|L^{right}|} > \sigma_5 \wedge \frac{|C^{sup}|}{d(IL, IR)} > \sigma_6 \wedge d(L, S1) < \sigma_7 N \wedge d(R, S2) < \sigma_7 N \quad (4)$$

donde:

$$d(a, b) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}$$

En los términos primero y segundo de la ecuación, $\sigma_5 \in [0, 1]$ indica el porcentaje de puntos de las líneas verticales que debe haber bajo la intersección con la horizontal. El tercer término, $\sigma_6 \in [0, 1]$ especifica el grado de ocupación que debe haber entre los puntos (IL, IR) por puntos de la línea horizontal. En los dos últimos términos, $\sigma_7 \in [0, 1]$, se especifica la separación que se acepta entre los puntos límite de la línea horizontal con los más cercanos a las intersecciones de las verticales. Con esto se evitan que líneas horizontales separadas de la puerta (que estén por encima o debajo del marco superior real) se tomen como posibles marcos.

6. Agrupar los marcos seleccionados para formar objetos puerta. Tras el proceso anterior pueden quedar unos marcos contenidos en otros (Figura 2(b)). Este proceso realiza la agrupación de los marcos en objetos de tipo puerta. Esto se realiza de la siguiente manera: sean dos marcos $M_1 = (L_1^{left}, L_1^{sup}, L_1^{right})$, $M_2 = (L_2^{left}, L_2^{sup}, L_2^{right})$. Y sus correspondientes puntos de intersección calculados en el paso anterior, IL^1, IR^1 los de M_1 y IL^2, IR^2 los de M_2 . Se dice que el marco M_2 está dentro del marco M_1 en la imagen, si cumple la siguiente condición:

$$M_2 \subset M_1 \Leftrightarrow IL_x^2 \in [IL_x^1, IR_x^1] \wedge IL_y^2 \in [IL_y^1, IR_y^1].$$

Obsérvese que con esta definición un marco siempre está contenido en sí mismo. De esta manera definimos el objeto puerta como el conjunto de marcos que están contenidos dentro de un marco envolvente como $P(M_i) = \{M_j / M_j \subset M_i \forall j\}$. Donde M_i es el marco envolvente y el resto son el conjunto de marcos que están dentro de él.

3.1. Integración multisensorial

El proceso explicado hasta ahora nos permite detectar en una imagen una posible puerta, pero es necesario incorporar información de otros sistemas sensoriales para confirmar la presencia de la puerta y comprobar realmente si está abierta. El robot debe buscar puertas abiertas pues son las que le permiten alcanzar otras habitaciones. Para confirmar la existencia de la puerta abierta el robot se acerca, mediante el agente Navigation y siguiendo la dirección marcada por la cámara, a la zona en donde se detectó visualmente la posible puerta. Una vez alcanzada cierta distancia se utilizan los sensores de ultrasonidos para detectar el hueco que supone la presencia de la puerta entre las paredes. A continuación se comentan brevemente las ideas subyacentes de la detección de las puertas mediante ultrasonidos. Este proceso utiliza el modelo perceptual explicado en [3] para establecer la presencia o ausencia de las paredes del entorno. La filosofía de este proceso es definir una serie de variables difusas para determinar la variación que se produce entre las medidas de los sensores de ultrasonidos que rebotan en las paredes y de aquellos que atraviesan el hueco de la puerta. Debido a que se producen situaciones de ruido sensorial en este contexto, se utiliza también un grado de estabilidad y certeza de las medidas sensoriales. Además se exige cierta repetición continuada de los valores apropiados en estas variables para llegar a establecer un grado de certeza en la existencia del hueco creado por la puerta abierta.

Como resultado de este proceso, se puede llegar a concluir que efectivamente la información suministrada por el sistema de visión corresponde a una puerta abierta. En caso de que no se llegue a detectar el hueco de la puerta, el objeto puerta visualmente detectado puede ser una puerta cerrada o un objeto diferente a una puerta pero parecido a ella. En la Sección 4 se presenta un caso de este tipo que corresponde realmente a una estantería de dimensiones parecidas a las puertas. Por otra parte, también podría tratarse de una puerta cerrada, pero para diferenciar las puertas cerradas de otros objetos parecidos, tendríamos que incorporar más información al sistema.



Figura 3: Detección visual de posibles marcos

4. Experimentación

En esta sección queremos por un lado mostrar los resultados de la aplicación del método de detección de puertas en las imágenes de nuestro entorno y por otro realizar un análisis del rendimiento del sistema de visión.

En la Figura 3 se muestra, en la parte superior, las imágenes capturadas por nuestra cámara a las que se ha incorporado el resultado del extractor de fronteras (en negro). En las imágenes que aparecen debajo (Figura 3), respectivamente, se muestran los marcos detectados con el método explicado en este trabajo. En las tres imágenes se muestra como el método detecta adecuadamente el marco de la puerta ante condiciones de iluminación variable.

Sin embargo, como vemos en la Figura 4, se puede observar como éste método puede llegar a interpretar como una posible puerta otros objetos del entorno que no lo son en realidad. En el caso mostrado por la Figura 4 el método asigna la propiedad *marco* a una estantería de parecidas dimensiones a las puertas de nuestro entorno. Este problema de falsos positivos puede ser superado mediante la integración de información procedente de los sensores de ultrasonidos como se comentó anteriormente. El robot intentará detectar la presencia del hueco de la puerta para confirmar la existencia de la misma. Así, en este caso, el robot será capaz de detectar que no existe hueco para poder pasar y no clasificará el objeto como una puerta abierta.



Figura 4: Falso positivo en la detección visual de marcos

Realizaremos a continuación un estudio del rendimiento del sistema. La transformada de Hough es un algoritmo de complejidad cuadrática. Depende por un lado del número de puntos que analicemos y por otro del número de elementos que se toman para discretizar el ángulo ϕ . En la Figura 5 se muestran las curvas de tiempo de nuestra implementación del algoritmo. Cada línea representa el tiempo empleado por el mismo considerando una discretización concreta, haciendo variar el número de puntos. Se observa como el tiempo de cómputo del algoritmo aumenta al incrementar el número de puntos a analizar y el número de elementos en la discretización de ϕ , pero observamos que para un tamaño de la discretización concreto, el crecimiento es lineal respecto del número de puntos.

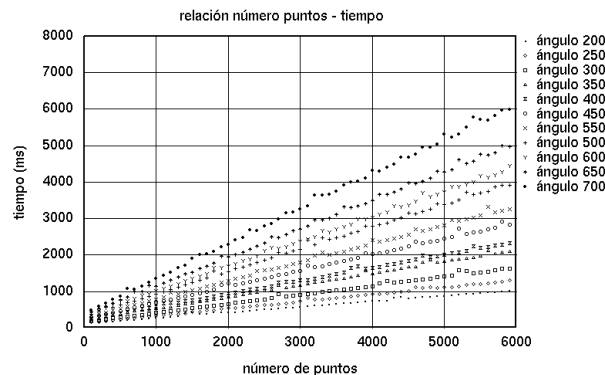


Figura 5: Tiempos de la transformada de Hough en función del número de puntos y el número de discretizaciones en ϕ

En la gráfica 6(a), se muestra el tiempo de cómputo del proceso completo de detección, variando el número de elementos en la discretización. Este ejemplo se ha realizado con una imagen de 250x250 pixels que contenía una puerta. Tras aplicar el detector de Canny quedaron 1541 puntos frontera. La gráfica 6(b), muestra el reparto de tiempo entre las principales operaciones del proceso, discretizando el ángulo en 350 intervalos. Como se observa, aproximadamente un 16 % se emplea para realizar la extracción de bordes con Canny (punto 1 explicado en la Sección 3). El 50 % del tiempo de procesamiento se emplea en realizar la transformada de Hough. Otro 16 % se emplea en llevar a cabo el procesamiento correspondiente a los puntos del 3 al 6 del método. Y el resto del tiempo se emplea en otras operaciones como capturar la imagen con la cámara y operaciones intermedias. Teniendo en cuenta esta distribución de tiempos, los elementos que inciden sobre el tiempo de cómputo del algoritmo pueden ser ajustados para reducirlo sin perder demasiada fiabilidad, especialmente la transformada de Hough. En las pruebas realizadas utilizando tamaños de discretización de 350 hemos obtenido tiempos de detección alrededor de 1,5

segundos. Es por tanto un tiempo adecuado para utilizar la técnica en tiempo real en nuestro sistema.

Podríamos mejorar el rendimiento de la transformada de Hough seleccionando aleatoriamente una fracción del total de puntos [8]. Al emplear esta técnica, los picos que representan líneas en el espacio de parámetros adquieren un carácter más pronunciado, lo que facilita la detección de los mismos. Otra ventaja es que al tomar una fracción $\frac{1}{f}$ de los puntos de análisis, conseguiríamos un incremento de velocidad de f .

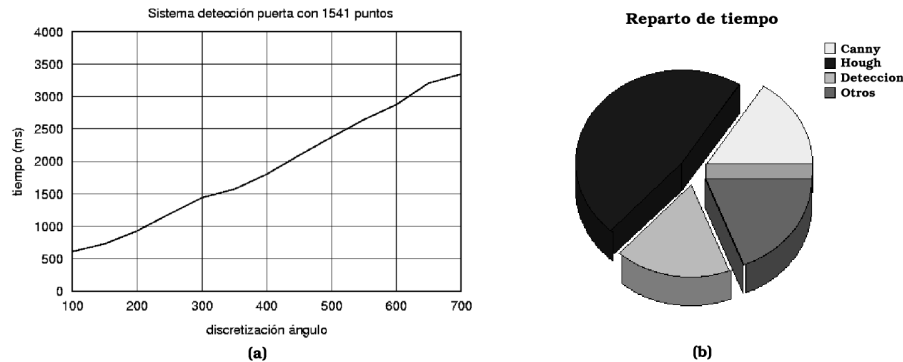


Figura 6: (a) Tiempos del sistema de detección para 1541 puntos con distintas discretizaciones de ϕ (b) Reparto del tiempo entre las distintas operaciones. Canny) Extractor de bordes de Canny. Hough) transformada de Hough. Detección) Análisis de las líneas en la transformada de Hough. Otros) Otras operaciones

Todas las pruebas han sido realizadas sobre imágenes de 250x250 pixels en el equipo de procesamiento del robot. Éste es un ordenador portátil Pentium III a 1100 MHz con 512 Mb de RAM que usa WindowsXP.

5. Conclusiones y Trabajo Futuro

En este trabajo se plantea el interés de integrar diferentes fuentes de información sensorial para llevar a cabo la detección de puertas abiertas en un entorno de interiores. El trabajo se centra en el reconocimiento visual de las puertas, sin añadir ninguna marca artificial en el entorno y teniendo en cuenta diferentes condiciones ambientales de iluminación y perspectivas. Para llevar a cabo la detección de los marcos de la puerta en una imagen se utiliza la transformada de Hough aplicada a los bordes detectados en dicha imagen. Estas líneas son analizadas teniendo en cuenta tanto la información de la transformada de Hough como la relación espacial existente entre los puntos que las forman. El método es robusto frente a la deformación producida por la perspectiva, a cambios de escala y a oclusión parcial. Una vez detectada visualmente la posible puerta el robot es capaz de dirigirse hacia ella y utilizar un modelo perceptual basado en ultrasonidos para confirmar si el objeto es una puerta abierta o no. Esta integración multi-sensorial nos permite descartar falsos marcos que pueden llegar a detectarse en las imágenes. La implementación se ha realizado para un robot móvil Nomad 200 mediante un sistema multi-agente capaz de llevar a cabo una navegación segura en entorno de interiores utilizando una serie de comportamientos difusos y un mapa topológico del entorno.

Tal y como se dijo anteriormente este es un trabajo en desarrollo. Por tanto aún quedan varios aspectos sobre los que avanzar. Respecto a la integración multisensorial, aunque el modelo perceptual basado en ultrasonidos ya está implementado, aun falta realizar su integración completa con el sistema de visión mediante un modelo formal adecuado. Respecto al tratamiento de la información visual, nuestro agente será capaz de mantener la puerta en el campo de visión de la cámara pese a los movimientos del robot mientras se dirige hacia ella. También se pretende flexibilizar el concepto de marco y puerta introduciendo técnicas basadas en lógica difusa. Ello nos permitiría contemplar situaciones en las que sería posible la detección aunque la puerta no esté completamente abierta. Finalmente también se tendrán en cuenta otras características visuales como el estudio de la entropía de la imagen o detección de esquinas para determinar si la puerta está abierta o cerrada.

Referencias

- [1] E. Aguirre, M. Gómez, R. Muñoz, y C. Ruiz. Un sistema multi-agente que emplea visión activa y ultrasonidos aplicado a navegación con comportamientos difusos. *Actas del IV Workshop en Agentes Físicos (WAF'2003)*, páginas 63–74, Alicante, 2003.
- [2] E. Aguirre y A. González. Fuzzy behaviors for mobile robot navigation: Design, coordination and fusion. *International Journal of Approximate Reasoning*, 25:255–289, 2000.
- [3] E. Aguirre y A. González. A fuzzy perceptual model for ultrasound sensors applied to intelligent navigation of mobile robots. *Applied Intelligence*, 19(3):171–187, 2003.
- [4] J.M. Cañas, R. Simmons, y M. C. García-Alegre. Detección probabilística de puertas con visión monocular activa. *Actas del II WorkShop Hispano-Luso en Agentes Físicos (Waf'2001)*, páginas 113–128, 2001.
- [5] J. Canny A computational approach to edge-detection *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 8, No. 6, Nov 1986.
- [6] G. Cicirelli, T. D'orazio, y A. Distanto. Target recognition by component for mobile robot navigation. *Journal of Experimental and Theoretical Artificial Intelligence*, 15(3):281–297, 2003.
- [7] P.V.C. Hough. Method and means for recognizing complex patterns. *U.S. Patent 3069654*, 1962.
- [8] Nahum Kiryati, Heikki Kälviäinen and Satu Alaoutinen. Randomized or probabilistic Hough transform: unified performance evaluation. *Pattern Recognition Letters*, Volume 21, Issues 13-14, December 2000, Pages 1157-1164.
- [9] D. Kim y R. Nevatia. A method for recognition and localization of generic objects for indoor navigation. *Image and Vision Computing*, 16(11):729–743, 1998.
- [10] D. Kortenkamp y T. Weymouth. Topological mapping for mobile robots using a combination of sonar and vision sensing. *Proc. of the Twelfth National Conf. on AI (AAAI-94)*, páginas 979–984, Menlo Park, Calif., 1994.
- [11] M. Mataric. Integration of representation into goal-driven behavior-based robots. *IEEE Transactions on Robotics and Automation*, 8(3):304–312, 1992.
- [12] I. Monasterio, E. Lazkano, I. Rañó, y B. Sierra. Learning to traverse door using visual information. *Mathematics and Computer in Simulation*, 60:347–356, 2002.
- [13] S. A. Stoeter, F. L. Mauff, y N. P. Papanikolopoulos. Real-time door detection in cluttered environments. *Proceeding of the 15th IEEE International Symposium on Intelligent Control (ISIC 200)*, páginas 187–191, 2000.