

Silla de Ruedas Inteligente Controlada por Voz

*J. Minguez¹, J.M. Alcubierre¹, L. Montesano¹, L. Montano¹,
O. Saz² y E. Lleida²*

Instituto de Investigación en Ingeniería de Aragón

¹ **Departamento de Informática e Ingeniería de Sistemas**

² **Departamento de Ingeniería Electrónica y Comunicaciones**

Universidad de Zaragoza, España

Resumen

Esta comunicación muestra una silla de ruedas con capacidad de guiado autónomo controlada por voz. La silla robotizada está equipada con equipos informáticos y sensores para medir y procesar la información del entorno. Además, dispone de un sistema de movimiento autónomo y un sistema de reconocimiento de voz. El usuario puede introducir la localización hacia la que desea dirigir la silla o el movimiento que desea realizar por medio de órdenes verbales. La silla ejecuta las órdenes moviéndose autónomamente hacia las posiciones introducidas por el usuario. La ventaja de estas sillas de ruedas es que aumenta las posibilidades de los individuos con severas discapacidades motrices, especialmente en entornos donde resulta difícil maniobrar el vehículo.

1. Introducción

Una gran sección de la robótica e inteligencia artificial aplicada a la ayuda a los discapacitados está orientada al desarrollo de las sillas de ruedas inteligentes [1]. Estos equipos están orientados a mejorar la calidad de vida ofreciendo un amplio rango de aplicaciones como la ayuda a la manipulación o a la movilidad para personas con problemas motrices por ejemplo. Esta comunicación se centra en la ayuda a movilidad por medio de sillas robóticas inteligentes, las cuales dan un valor añadido a las sillas comerciales motorizadas convencionales. Hay al menos tres factores claves del desarrollo de estos equipos: (i) la silla robótica, (ii) el interfaz hombre - máquina y (iii) la inteligencia empotrada de guiado autónomo.



Organiza:



Patrocina:



Colaboran:



Las sillas de ruedas robotizadas son sillas convencionales que se han equipado con sistemas informáticos y sensores. Los sensores captan el entorno y el estado de la silla. Esta información es transmitida a los computadores de a bordo que la procesan y toman decisiones transmitiendo a los actuadores las órdenes para que sean ejecutadas. Usualmente, estas órdenes están relacionadas con el movimiento.

Los interfaces hombre-máquina son las entidades que permiten la comunicación del hombre con el dispositivo y viceversa. En el caso de personas discapacitadas esta comunicación suele ser bidireccional. En este proyecto, en una dirección el hombre da órdenes a la máquina por medio de un interfaz de voz. Muchos de los potenciales usuarios tienen problemas de dicción, lo que hace que el sistema de voz deba de ser robusto, flexible y adaptable a cada usuario. Dado que los sistemas comerciales no cumplen estas características, en este proyecto se ha desarrollado un sistema de reconocimiento de voz propio. Por otro lado, la máquina ofrece a la persona información por medio de un sistema visual en una pantalla.

El aspecto restante es la inteligencia de guiado empotrada. Los sistemas de movimiento autónomo son tecnologías que generan el movimiento de un vehículo entre localizaciones mientras se evitan las posibles colisiones con los obstáculos que puedan aparecer por el camino. La robustez de estos sistemas es de vital importancia dado que son los encargados de generar el movimiento final (sistemas de seguridad crítica). En este proyecto se ha desarrollado un sistema de movimiento con capacidad de planificar caminos entre posiciones y generar movimiento con evitación de obstáculos detectados por los sensores en ambientes estáticos y dinámicos [2, 3].

La aplicación silla de ruedas inteligente funciona de la siguiente forma. El usuario introduce las posiciones o movimientos deseados a donde desea mover la silla por medio del interfaz de voz. El reconocedor de voz interpreta y traduce los comandos verbales en órdenes para la máquina. El sistema de movimiento recibe las órdenes y conduce de forma autónoma el vehículo hasta la posición indicada.



Silla de ruedas inteligente controlada por voz

Este tipo de dispositivos mejora gratamente la calidad de vida de las personas discapacitadas y en muchos casos puede permitir manejar una silla de ruedas a personas que de otra manera serían incapaces de realizarlo. Por otro lado, el sistema de movimiento es capaz de conducir el sistema en situaciones de difícil maniobrabilidad, las cuales son complicadas para personas con patologías relacionadas con el parkinson o movimientos espasmódicos por ejemplo.

2. La silla de ruedas

La silla de ruedas robotizada es una silla de ruedas comercial motorizada que se ha equipado con computadores y sensores (Figura 1). La silla es rectangular con tracción diferencial (dos ruedas controlables y dos ruedas libres). Es decir, la silla dispone de dos motores eléctricos de corriente continua que impulsan las dos ruedas traseras (ruedas tractoras). Las dos ruedas delanteras de movimiento libre permiten el giro de la silla.



Figura 1. Silla de ruedas robotizada

El vehículo se ha equipado con dos ordenadores empotrados. Un ordenador, Pentium MMX 266 MHz con 64 MB de memoria RAM, encargado del control de la silla a bajo nivel. El sistema operativo instalado en este ordenador es VxWorks (dado que es el encargado del tiempo real). El otro es un Pentium III 850 MHz con 256 MB de memoria RAM, encargado de la navegación alto nivel de la silla. El sistema operativo instalado en este ordenador es Windows 2000. Este es el ordenador en el que se ejecutan todos los programas mencionados a continuación, como la interfaz de voz y gráfica y el sistema de movimiento.



Organiza:



Patrocinan:



Colaboran:



Ambos ordenadores están comunicados mediante puertos serie RS-232 y mediante una red ethernet.

El sensor principal es un láser de la marca SICK, cuyo cometido es el de medir la distancia de los objetos situados en torno a la silla. Para ello realiza un barrido en un ángulo máximo de 180° y a una distancia máxima de 8 metros. Además en el vehículo se dispone de una tarjeta ethernet que permite comunicación con ordenadores conectados a una red informática, pudiendo programar o teledirigir la silla sin estar conectado físicamente a ella. Finalmente la silla dispone de una pantalla conectada a la salida VGA del computador para tareas de visualización, y de un micrófono para introducir las órdenes.

3. El interfaz hombre-máquina

Esta sección describe brevemente el sistema de interfaz hombre - máquina de la silla de ruedas. Por un lado, el usuario introduce los sistemas pronunciando unas órdenes verbales. El sistema de reconocimiento de voz traduce e interpreta estas señales para convertirlas en movimiento (Subsección 3.1). El sistema informático proporciona información al usuario por medio de un interfaz visual (Subsección 3.2).

3.1 El reconocedor de voz

El objetivo del reconocedor de voz es transformar las órdenes de movimiento pronunciados por el usuario en posiciones del espacio. Es decir, hace de interfaz entre la orden verbal del usuario y la orden a ejecutar por la máquina. Este módulo tiene dos partes: el sistema de reconocimiento de voz [4], y la traducción de las órdenes en movimiento. En primer lugar se describe el sistema de reconocimiento automático del habla, que tiene los siguientes bloques funcionales básicos (Figura 2):

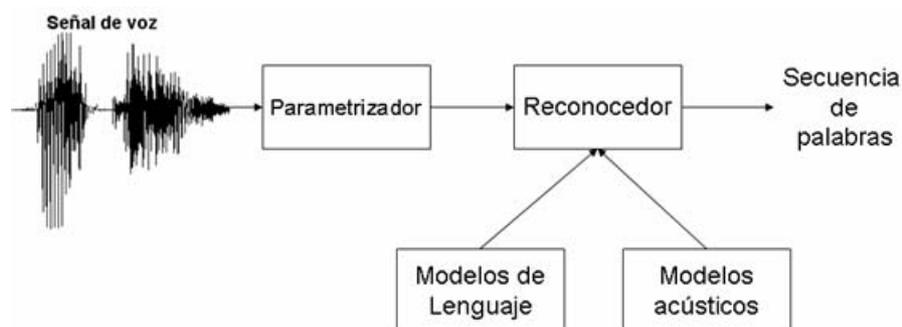


Figura 2. Esta figura muestra el diagrama de bloques funcional del reconocedor de voz.

Modelos de Lenguaje

El modelo de lenguaje representa el conjunto de palabras que se pueden decir (vocabulario) y cómo se pueden decir (gramática). Para el caso de la silla, las palabras del vocabulario son: **Dusila**¹, **Adelante**, **Detrás**, **Izquierda**, **Derecha**, **Lejos**, **Medio**, **Cerca**, **Anda**, **Para**, **Inicia** y **Termina**. La gramática de comandos se muestra en la Figura 3.

Modelos acústicos

Estos modelos se obtienen en una fase de entrenamiento. Para ello se crea una base de datos (corpora) con repeticiones de las diferentes unidades existentes en la lengua castellana. hasta obtener una representación estadística de la señal de voz para cada unidad. Se han usado modelos de subfonemas, unidad lingüística más pequeña que el fonema y que contiene información contextual. En este proyecto se utiliza un reconocedor adaptado al locutor, es decir, que utiliza la representación más próxima a la persona que lo va a usar. La estrategia utilizada es la Maximum A Posteriori (MAP) que parte de unos modelos generales con unas pocas frases de entrenamiento dichas por la persona a las que se quiere adaptar.

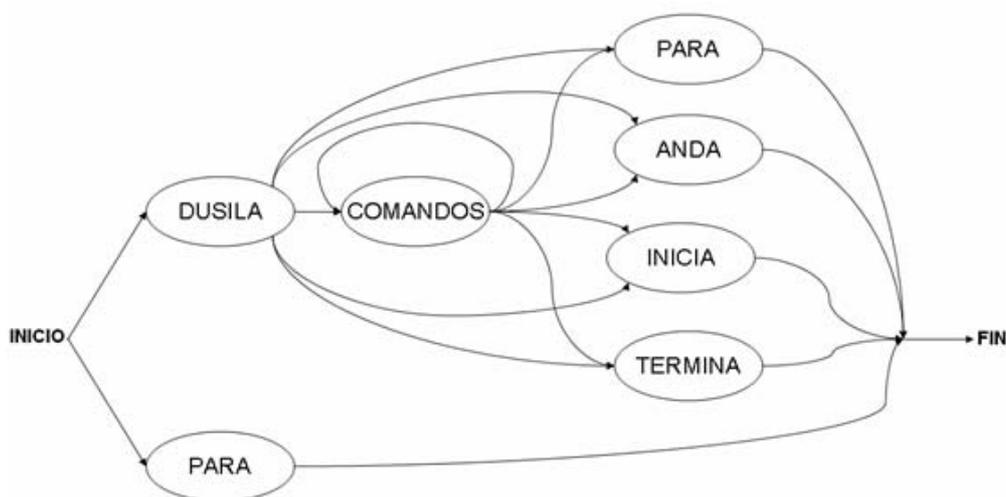


Figura 3. La palabra de activación es **Dusila**. Posteriormente se puede decir cualquiera de los comandos: **Adelante**, **Detrás**, **Izquierda**, **Derecha**, **Lejos**, **Medio** o **Cerca**. Después de estas órdenes, la orden **Anda** pone en marcha la silla. La orden **Para** detiene la silla en cualquier momento sin necesidad de orden de activación. Finalmente, las ordenes **Inicia** y **Termina** inician y finalizan respectivamente todo el sistema.

¹ Dusila en honor a Livia Drusila, 58 a.c.- 29 d.c. esposa del cesar Augusto fundador de la ciudad de Zaragoza, y madre de Tiberio precursor del teatro romano de dicha ciudad.

Parametrizador

Este bloque transforma la señal de voz obtenida con el micrófono en un conjunto de parámetros que representan dicha señal y sus variaciones. Para ello se han usado los parámetros MFCC (Mel-Frequency Cepstrum Coefficients) [5], por su capacidad de reducir al menor número de parámetros la información más relevante de la señal de voz.

Reconocedor

El reconocedor realiza un reconocimiento de patrones de la señal parametrizada de entrada con respecto a una serie de modelos que representan los diferentes fonemas. Se utilizan Modelos Ocultos de Markov [6] por su capacidad para modelar procesos aleatorios como la señal de voz que varían con el tiempo. La estrategia de búsqueda a través de estos modelos es el algoritmo de Viterbi. Para robustecer el reconocedor se implementa una medida de confianza de la palabra reconocida [7], que evitará que, cuando se pronuncien palabras fuera del vocabulario, se reconozcan como palabras pertenecientes al vocabulario, a costa de que algunas veces, cuando sí se pronuncie una palabra del vocabulario el reconocedor la rechace.

Una vez reconocidas las órdenes, la siguiente etapa es su interpretación para convertirlos en una posición en el espacio. La estrategia utilizada es similar a la de manejar el volante de un coche. Alrededor de la silla se establece un mallado polar (Figura 4). Las órdenes **Lejos**, **Medio**, **Cerca** colocan la posición final en uno de los tres círculos concéntricos. Las órdenes **Adelante** y **Detrás** colocan la localización en la parte delantera o trasera de la silla. Las órdenes **Izquierda**, **Derecha** mueven la localización dentro del círculo un arco hacia la izquierda o derecha respectivamente. Finalmente **Anda** y **Para** ejecutan o detienen el movimiento. Las órdenes pueden ser reintroducidas en todo momento en el sistema.

3.2 Interfaz visual

La interfaz visual permite al usuario recibir una realimentación del estado de la silla (Figura 4). Se distinguen tres partes:



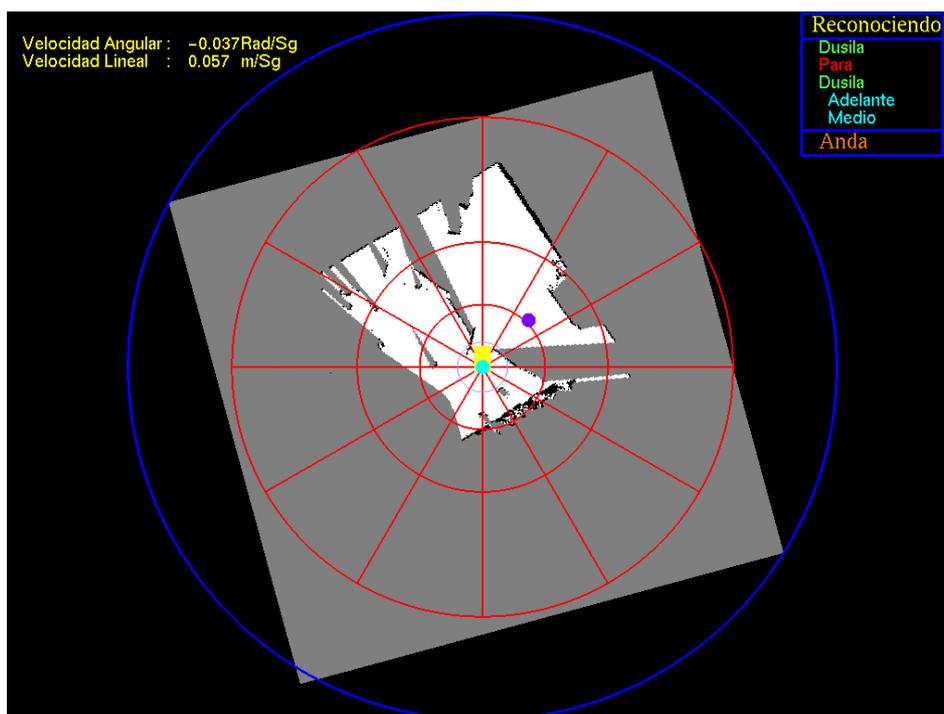


Figura 4. Esta Figura la interfaz visual del sistema. En ella se aprecian los comandos de voz reconocidos hasta el momento (parte superior derecha) el mapa del entorno con los obstáculos detectados hasta el momento (parte central), el mallado polar donde se coloca la posición de destino (parte central).

- El visualizador de órdenes reconocidas hasta el momento permite al usuario confirmar las órdenes dadas a la silla.
- El mallado polar de localización permite al usuario ver donde ha colocado la posición final que debe de alcanzar el vehículo.
- El mapa del entorno permite al usuario ver el mapa generado por la silla hasta el momento, viendo así los obstáculos detectados y la posición final ordenada en el mallado dentro del mapa.

Con este interfaz, el usuario tiene información de las órdenes interpretadas (lista de órdenes), hacia donde va la silla (posición final) y del entorno que le rodea (mapa generado). A su vez, el interfaz se actualiza en tiempo real durante el movimiento proporcionando información al usuario constantemente.

4. El Sistema de Movimiento Autónomo

El objetivo del sistema de movimiento es conducir el vehículo hasta una posición destino mientras evita las colisiones con los obstáculos detectados por los sensores. El sistema desarrollado incorpora las funcionalidades de modelado del entorno, planificación y evitación de obstáculos (Figura 5):



Organiza:



Patrocinan:



Colaboran:



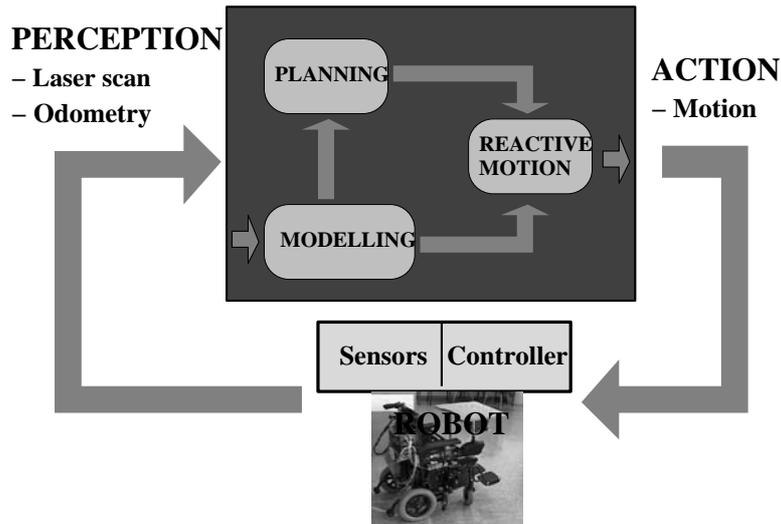


Figura 5. Esta figura muestra los diferentes módulos del sistema de movimiento autónomo: modelado, planificación y evitación reactiva de obstáculos.

Modelado del entorno

Construye un modelo del entorno para incrementar el dominio espacial de la planificación y usado como memoria local para la evitación de obstáculos. Se utiliza un mapa basado en una discretización del espacio de trabajo en celdas binarias (ocupado/libre), dado que es una estructura muy eficiente de representación en tiempo real de mapas en entornos densos y no estructurados. Para corregir la posición local del vehículo en el mapa se utiliza una técnica de alineamiento de barridos del sensor láser.

Planificador táctico

Calcula la dirección táctica de movimiento para evitar los movimientos cíclicos y las situaciones de atrapamiento. La idea detrás del planificador es computar un camino hasta la posición de destino utilizando sólo las zonas del espacio que han cambiado desde a última medida sensorial. Esta estrategia es muy eficiente para implementaciones tiempo real.

Evitación de obstáculos

Genera y ejecuta el movimiento hasta la posición de destino evitando los obstáculos detectados por los sensores. Se ha desarrollado una técnica caracterizada por ser robusta en entornos confinados y de difícil maniobrabilidad.

En general el sistema funciona como sigue. Dada una medida del sensor láser y la odometría del vehículo, el módulo de modelado actualiza el modelo del entorno. A continuación el planificador utiliza los cambios en el entorno para calcular la dirección táctica de movimiento. Finalmente, usando el modelo del entorno, el método de evitación de obstáculos computa la orden de movimiento para evitar obstáculos mientras alinea el vehículo con la dirección táctica de movimiento. La orden es ejecutada por el vehículo y el proceso vuelve a empezar. En este esquema la posición de destino puede ser fija o modificada en cualquier momento, lo cual está perfectamente adaptado a los requerimientos del sistema de voz y de las necesidades de conducción de la silla de ruedas.

5. Demostración del sistema

Esta sección muestra un ejemplo de utilización de la silla de ruedas. El entorno de aplicación es una oficina. Notar que no se conoce ninguna información a priori del entorno como su planta o estructura, además éste podría tener cualquier naturaleza como estructurado/desestructurado, dinámico/estático o incluso si se conociese el mapa del mismo éste podría ser utilizado.

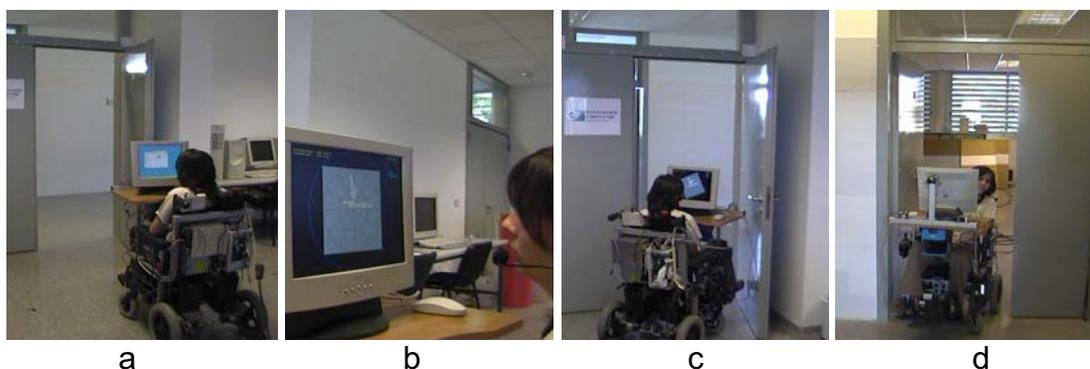


Figura 6. Esta figura muestra cuatro instantáneas de una prueba del sistema. El objetivo es cruzar la puerta. El usuario introduce la posición final por medio de la voz. La silla ejecuta las órdenes maniobrando para cruzar la puerta hasta el pasillo cumpliendo su misión.

La intención del usuario es abandonar la estancia en la que se encuentra. Para ello debe de conseguir cruzar una puerta. Éste es un caso difícil a nivel de navegación dado que la puerta es estrecha y el sitio para maniobrar es reducido sobre todo si no se dispone de un control fino del movimiento (Figura 6a). El usuario verbalmente coloca la posición final al otro lado de la puerta. En este proceso utiliza la interfaz gráfica para ver el mapa actual del entorno construido por el sistema de movimiento y saber que efectivamente la posición

final está colocada en el sitio adecuado (Figura 6b). Automáticamente el sistema de movimiento recibe la posición final hacia la que debe de dirigir el vehículo y procede a realizarlo.

La situación más complicada a nivel de movimiento aparece cuando la silla está cruzando la puerta (Figura 6c). En esta situación es importante que el sistema de movimiento tenga en cuenta la forma del vehículo para evitar colisiones en sitios estrechos. La cinemática es importante dado que es importante tener en cuenta cómo se mueve el vehículo para evitar colisiones en estas zonas. Finalmente, es importante tener en cuenta la dinámica, dado que se debe de tener en cuenta lo que el vehículo puede hacer para generar movimiento. El resultado de tener en cuenta todos estos factores en la capa de movimiento resulta en una aplicación robusta capaz de mover el vehículo en entornos de difícil maniobrabilidad, además que resulta en moviendo confortable y suave para el usuario. La demostración termina cuando la silla abandona autónomamente la estancia y alcanza el pasillo exterior, cumpliendo el objetivo marcado.

6. Conclusiones y futuras vías de trabajo

En esta comunicación se han presentado una silla de ruedas con capacidad de guiado autónomo controlada por voz. El trabajo es el resultado de varios años de investigación en los dominios de la navegación en robótica y en el de reconocimiento de la voz. Se enmarca en proyectos de investigación que tienen entre otros objetivos la aplicación de las técnicas mencionadas al diseño de sistemas y dispositivos a la mejora de la calidad de vida de las personas. Se sigue trabajando en la ampliación de las funcionalidades descritas, incorporando otras tales como la capacidad de reconocer el entorno para localizarse, la de planificar automáticamente el camino a seguir para ir a un destino determinado, el control por medio de sistemas remotos como PDAs y más a largo plazo la incorporación de un sistema de visión para reconocer órdenes como gestos o movimientos.

Desde el punto de vista del uso de la voz, para la plena funcionalidad de la silla será necesaria la implementación de sistemas de verificación del locutor. De esta forma, la silla sólo atenderá a los comandos pronunciados por el



Organiza:



Patrocinan:



Colaboran:



propietario de la silla, desechando las palabras pronunciadas por cualquier persona ajena a la silla.

En la actualidad se colabora con el colegio de educación especial Alborada de Zaragoza para niños discapacitados. La colaboración está siguiendo dos direcciones dentro del proyecto. En primer lugar se está utilizando y adaptando el interfaz de voz para reconocer órdenes verbales de personas con severas deficiencias orales. En segundo lugar se está estudiando como afecta esto a los sistemas de movimiento. En su conjunto, todo este trabajo está permitiendo la adaptación de las interfaces y del sistema a las discapacidades específicas, tanto motoras como cognitivas, a la vez que sirve para la puesta a punto final en un ambiente real y complejo.

Bibliografía

- [1] Research on Autonomous Robotic Wheelchairs in Europe. IEEE Robotics and Automation Magazine. March. 2001.
- [2] J. Minguez. Integration of Planning and Reactive Obstacle Avoidance for Sensor-Based Navigation. IEEE International Conference on Intelligent Robot and Systems. 2005. Edmonton, Canada.
- [3] J. Minguez L. Montano. Autonomous Sensor-Based Motion Control in Unknown, Dynamic and Troublesome Scenarios. Journal of Robotics and Autonomous Systems. Volume 52, Issue 4 , 30 September 2005, Pages 290-311. 2005.
- [4] Xuedong Huang, Alex Acero, Hsiao-Wuen Hon: "Spoken Language Processing: A Guide to Theory, Algorithm and System Development", Prentice Hall, 2001.
- [5] Furui, S.: "Speaker-independent isolated word recognition using dynamic features of speech spectrum", IEEE Transactions on Speech, Audio and Signal Processing, vol. 34 (1), pp. 52--59, 1986.
- [6] Jelinek, F.: "Continuous speech recognition using statistical methods", Proceedings of the IEEE, vol. 64(4), pp. 532--556, 1976



Organiza:



Patrocinan:



Colaboran:



- [7] Lleida, E., Rose, R.C.: "Utterance verification in continuous speech recognition: decoding and training procedures" IEEE Transactions on Speech and Audio Processing, vol. 8(2), pp. 126--139, 2000.



Organiza:



Patrocinan:



Colaboran:

