



DESARROLLO DEL SISTEMA DE PERCEPCIÓN DE UNA PLATAFORMA MÓVIL PARA ENTORNOS EXTERIORES

CRISTINA CASTEJÓN¹, DOLORES BLANCO²,
BEATRIZ LÓPEZ BOADA¹, LUIS E. MORENO²

Universidad Carlos III de Madrid

¹Departamento de Mecánica

²Departamento de Ingeniería de Sistemas y Automática

Avda. de la Universidad 30, 28911 Leganés, Madrid, España

(Recibido 24 de noviembre de 2003, para publicación 10 de febrero de 2004)

Resumen – *Un objetivo tecnológico que cada día cobra un mayor auge es dotar de autonomía a los sistemas móviles, desde robots a vehículos de uso particular o industrial. La autonomía conlleva tener capacidad de acción y reacción en un entorno que puede ser desconocido a priori. Por ello, es necesario modelar el entorno a partir de la información obtenida por los sensores. Esta etapa es fundamental, ya que será la base de las decisiones futuras que deberán realizarse para cumplir una tarea específica. Para alcanzar esta funcionalidad, debe obtenerse información referente al mundo exterior a través del **sistema de percepción**. En este artículo se presenta el sistema sensorial incorporado a un robot móvil para entornos exteriores que permitirá obtener información tridimensional útil para posteriores etapas de modelado.*

1. INTRODUCCIÓN

La obtención de modelos del entorno a partir de la información proporcionada por los sensores es una funcionalidad básica para cualquier sistema autónomo. En robótica, esta tarea es previa e indispensable para la navegación autónoma en un entorno que no siempre es conocido o que a lo largo del desarrollo de la tarea puede modificarse. El tipo de modelo elegido para representar el espacio de trabajo varía en función tanto de las características de la plataforma móvil, como de las características propias del entorno. En este artículo presentamos la implementación del sistema de percepción desarrollado para una plataforma robótica móvil diseñada para trabajar en entornos exteriores. El modelo de este tipo de entornos debe recoger aquellas características necesarias para la posterior toma de decisiones en lo relativo a la navegación segura del vehículo, evitando los obstáculos que pueda encontrarse en su camino.

El tipo de entorno condiciona la elección e implementación del sistema de percepción. En la primera parte del trabajo que aquí presentamos se ha llevado a cabo un análisis de las condiciones del área de trabajo y de las limitaciones que impone sobre el sistema sensorial. Es necesario identificar el tipo de información que debe extraerse del área de trabajo y qué sensor es el más adecuado en las diferentes condiciones que puede encontrarse el robot. Al tratarse de un entorno de exteriores, el abanico de condiciones en las que debe ser capaz de actuar nuestro sistema sensorial es mucho más amplio que en el caso de robots que actúan en interiores.

Como se explicará posteriormente, se hace necesario recoger información tridimensional que permita discernir los obstáculos, que no siempre se encuentran en el plano del suelo (por ejemplo, obstáculos aéreos), y las irregularidades del terreno que pueden dificultar el movimiento de la plataforma. Además el modelo tendrá en cuenta las características propias del robot móvil, que determinan su capacidad de movimiento en diferentes entornos.

Existe una gran variedad de técnicas y tecnologías de sensores (sistemas infrarrojos, sensores láser, sonda, cámaras CCD o giróscopos). En aplicaciones tales como el modelado del entorno, los sistemas de percepción más convenientes son aquellos que proporcionan imágenes de profundidad, ya sea por medio de sensores de profundidad activos (telémetro láser, sónares) o pasivos (visión, estereovisión) [1]. En

particular, cuando la necesidad es modelar entornos exteriores, será conveniente utilizar información tridimensional para detectar y evitar obstáculos, en lugar de la información bidimensional y unidimensional utilizada generalmente en entornos interiores. La información tridimensional ayudará a resolver el problema de desconocimiento de la irregularidad del terreno y de la geometría de los obstáculos, que presentan la mayoría de los entornos exteriores. En interiores pueden asumirse ciertas consideraciones sobre la estructura del entorno, generalmente compuesto por planos verticales perpendiculares al plano de movimiento, que no son aplicables en entornos exteriores.

Existe un gran desarrollo tecnológico en el campo de los sensores de profundidad. Sin embargo, la mayoría de las soluciones han sido realizadas para aplicaciones industriales tales como metrología, topografía o inspección. Por tanto, se hace necesario identificar los requerimientos que son específicos en aplicaciones robóticas, para conocer el verdadero estado de desarrollo en este campo. Estos requerimientos son los siguientes:

- Nos enfrentamos a un problema **dinámico**, ya sea porque el robot está en movimiento, o porque está observando un entorno que cambia dinámicamente. En cualquiera de los dos casos, la velocidad de adquisición de datos será crítica, y muchos de los conceptos desarrollados por las áreas de metrología y topología no pueden ser aplicados a la robótica móvil de forma directa.
- El **conjunto físico** del sistema sensorial. Los sensores van integrados en el robot, por lo que sus características físicas influirán considerablemente en el diseño del sistema sensorial. Teniendo en cuenta que la tendencia actual es la de construir robots cada vez más ligeros, con una capacidad de almacenamiento muy limitada, los problemas de peso y combustible son críticos. Como resultado, la mayoría de los sistemas sensoriales existentes en el mercado no pueden utilizarse, a menos que se hagan serios avances en la integración sensorial.
- Finalmente, la capacidad para la **operación en entornos difíciles**, como condiciones de iluminación malas, contaminación, niebla, etc., se hace crítica en muchas aplicaciones de campo. Serán necesarios nuevos desarrollos orientados a la robustez de los sensores, antes de que muchas de las tecnologías propuestas puedan ser utilizadas en los robots.

En los trabajos encontrados en la literatura, se observa un especial interés en el desarrollo de sistemas sensoriales para una navegación específica, esto es, dependientes de las características del entorno de trabajo. En el ámbito que nos ocupa, se ha demostrado que el telémetro láser [2,3,4] y la visión estéreo [5,1,6], son los sistemas más eficientes para la navegación en entornos exteriores cuando se trata con grandes distancias [7].

Las secciones 2, 3 y 4 describen los sistemas y tecnologías empleados comúnmente para el modelado de entornos exteriores encontrados en la literatura, comentando las ventajas e inconvenientes de cada dispositivo. Posteriormente se aborda la presentación de la plataforma experimental, estableciendo los criterios de selección del sistema de percepción. Por último, se expone la técnica de obtención de información 3D propuesta.

2. LA TELEMETRÍA LÁSER COMO SISTEMA DE PERCEPCIÓN

Los sistemas láser, también conocidos como *láser radar* o *lidar*, son sensores activos. Aparecieron por primera vez en el Jet Propulsory Laboratory, Pasadena, en los años 70 [8]. Por lo general, los sensores láser radar o telémetro láser (*laser range-finders LRF*) miden la distancia d entre el sensor y la superficie de tres modos diferentes: Si la radiación es pulsada, se utiliza el tiempo de vuelo (Time of Flight TOF) y, si la radiación es una onda continua (CW Continuous Wave) y modulada, se utiliza una medida de la modulación. La modulación de la radiación láser CW puede ser en amplitud (AMCW) o en frecuencia (FMCW) [9]. Las técnicas más comunes para aplicaciones en robótica son TOF y AMCW.

2.1. Utilización del sensor

Con independencia de la tecnología láser empleada, para su utilización como sensor de medida de profundidad, es necesario definir los sistemas de referencia del espacio cartesiano y del espacio de la imagen.

Además, presenta cierta sensibilidad a las condiciones del entorno como niebla o polvo. En los últimos años se han introducido técnicas de "medida del último pulso" [11] para resolver este problema de sensibilidad. Cuando operamos en entornos hostiles tales como niebla, polvo o humo, esta técnica garantiza que la distancia medida de retorno es la de la superficie en lugar de la producida por la dispersión del medio.

2.3. Aplicabilidad en entornos exteriores

En Europa, las primeras aplicaciones de telemetría laser tridimensional en entornos exteriores se deben al grupo RIA de Francia que participó en 1987 en el proyecto Eureka AMR (Advanced Mobile Robot), con la construcción del sistema de percepción del robot ADAM, y posteriormente en el proyecto EDEN en 1996, para modelizar entornos exteriores [12]. El sistema de percepción estaba formado por un telémetro láser de dos ejes, que proporcionaba una imagen de puntos tridimensionales. Actualmente se está utilizando el laser para obtener información 3D en entornos exteriores. Por ejemplo, en [13] se consiguen construir modelos tridimensionales utilizando técnicas de realidad virtual gracias a la disposición de dos telémetros bidimensionales, uno colocado en posición horizontal y el segundo apuntando hacia arriba.

3. TRIANGULACIÓN ACTIVA / LUZ ESTRUCTURADA

La triangulación activa se basa en la utilización de un sensor activo, llamado emisor, que produce una iluminación de forma estructurada sobre una superficie o escena, y un sensor (normalmente una cámara CCD) que adquiere la imagen del patrón distorsionado, reflejado por la superficie o escena. Existen tecnologías muy diferentes para la generación de patrones de luz: emisores láser, proyectores de luz blanca y proyectores de video. La posibilidad de poder elegir entre una gran variedad de patrones diferentes (puntos de luz, líneas o mallas regulares), permiten realizar modelos geométricos muy precisos, siendo sus aplicaciones más comunes en medicina, y en general, para modelado y diseño de objetos. Por ejemplo, Rocchini et al. en [14] construyen un scanner 3D, basado en luz estructurada, con un proyector de video y una cámara digital. El proyector de video se utiliza para generar los patrones de luz, y la cámara adquiere las imágenes. Este sistema global permite obtener un mapa de profundidades denso para modelar obras de arte con precisión geométrica.

3.1. Principio básico

El principio básico es el siguiente: un patrón de luz (que puede ser un único punto) es proyectado sobre la escena, y un sensor, normalmente una cámara CCD, ve dicha escena [11]. Las coordenadas (X, Y, Z) de un punto de la escena respecto a un sistema de coordenadas situado en el centro óptico de la cámara, pueden obtenerse por triangulación con la ecuación 4 obtenida a partir de la figura 2:

$$(X, Y, Z) = \frac{B}{f \cos \theta - u} (u, v, f) \quad (4)$$

donde B es la línea base o distancia de separación entre los centros ópticos del emisor y la cámara, f es la distancia focal de la lente de la cámara, u y v corresponden a la posición del punto en el plano de la imagen, y θ es el ángulo de proyección del haz respecto al eje x .

3.2. Ventajas e inconvenientes

Un aspecto atractivo de esta técnica es la simplicidad de su implementación. Es muy útil para el modelado y reconocimiento de objetos y, la posibilidad de elegir diferentes patrones de luz, permite adaptarse a cualquier superficie.

Sin embargo, en el campo de la robótica de entornos exteriores presenta varios inconvenientes. Como puede observarse en la ecuación 4, la resolución del sistema de triangulación es directamente proporcional a la línea base. Esto implica que la obtención de medidas de profundidad para grandes distancias sólo

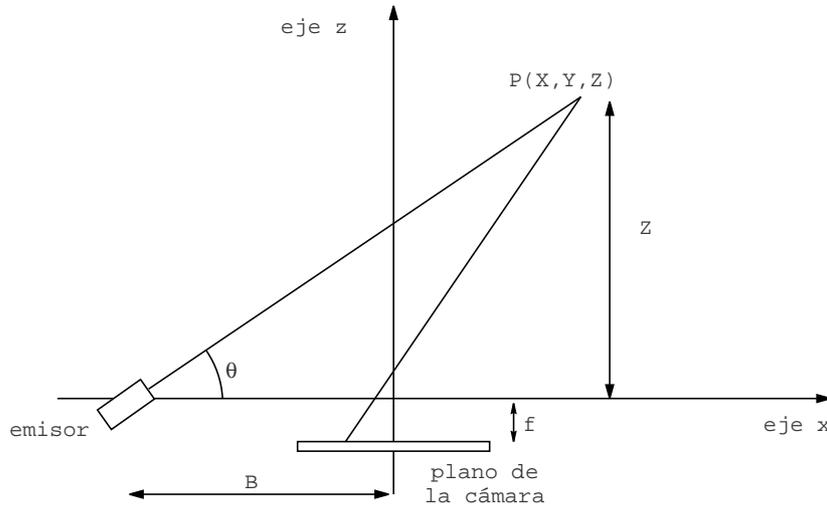


Fig. 2. Obtención de profundidad mediante luz estructurada.

puede ser alcanzada incrementando la línea base. Con ello también se incrementa la aparición de oclusiones y la pérdida de datos, y hace que el tamaño del sistema sensorial sea impracticable para sistemas que se encuentran a bordo de un robot. Otro aspecto a destacar es el incremento de ruido que introduce el emisor en entornos exteriores [15].

Además, para facilitar la detección del patrón en la imagen, las condiciones de iluminación deben ajustarse para que el patrón proyectado genere la característica más brillante de la escena [16]. Por ello esta técnica se aplica fundamentalmente en entornos interiores, donde la iluminación puede ser controlada.

4. LA VISIÓN ESTÉREO COMO SISTEMA DE PERCEPCIÓN

El término *estéreo* o *visión estéreo*, se utiliza para definir al proceso de combinar múltiples imágenes de una escena, para obtener información geométrica tridimensional. Es decir, la visión estéreo permite extraer información de profundidad a partir de dos o más imágenes.

Existen aplicaciones de visión estéreo para robots móviles autónomos, que navegan en entornos desconocidos estructurados [17], parcialmente estructurados [18] o naturales [19]. También tienen aplicación en la detección de obstáculos a partir de mapas generados con datos estéreo [20,18].

4.1. Fundamentos de la visión estéreo

Con la visión estéreo se recupera la información de la posición tridimensional de los objetos (información no contenida en una simple imagen) a partir de dos o más imágenes diferentes, tomadas de una misma escena. Las diferencias en las imágenes se originan: por un movimiento relativo del sensor, por desplazamiento relativo en dos o más sensores, o por movimiento de los objetos de la escena. Estas diferencias contienen la suficiente información para reconstruir la estructura de la escena tridimensional a partir de imágenes bidimensionales.

El principio general de la visión estéreo, es el de establecer un conjunto de emparejamientos entre las características adecuadas extraídas de las imágenes (primitivas), con el fin de llegar a construir una representación tridimensional del espacio.

El algoritmo de visión estéreo sigue las etapas que se detallan a continuación [21,22]:

a) Detección de características

Los algoritmos de visión estéreo se clasifican por el tipo de primitiva que se busca para el emparejamiento. La elección de las primitivas depende del entorno considerado y de la potencia de cálculo dispo-

nible. La elección de los elementos puestos en correspondencia es una de las fases más importantes: la primitiva ideal será aquella cuyas propiedades intrínsecas le permita tener una medida de similitud fuertemente discriminativa, y permita determinar el valor tridimensional de una posición a partir de la geometría del sistema. Al Haddad en [20] clasifica las posibles primitivas de emparejamiento en tres grupos:

- Puntos característicos como son: puntos de interés, elementos de contorno, uniones, etc.
- Segmentos como los segmentos rectos, arcos de círculo, etc..
- Regiones, es decir, zonas homogéneas que siguen un cierto criterio.

Salvo excepciones, las regiones se adaptan mal en entornos exteriores. Por el contrario, los segmentos rectos tienen mayor rango de aplicación, pero siempre en entornos estructurados. En un medio natural, que es normalmente no estructurado, los puntos (o pixels en el caso de imágenes) resultan ser las primitivas mejor adaptadas para la visión estéreo. El único problema que impone la utilización de puntos, como primitivas de base para el emparejamiento, es la elección de una medida de similitud. Las diferentes medidas para conseguir la correspondencia se desarrollan en esta misma sección, en el siguiente apartado.

Información detallada sobre los tipos de primitivas con las que diferentes autores han trabajado, y los algoritmos propuestos para su detección y emparejamiento, puede encontrarse desarrollada con claridad por Hemayed y Farag en [22].

b) Fase de emparejamiento

Una vez que las características han sido detectadas en las imágenes, es necesario establecer la correspondencia. Lo más usual es utilizar en el emparejamiento primitivas de bajo nivel, como el pixel de imagen. El método de correspondencia buscará una medida de similitud, basado en el valor de su intensidad o nivel de gris. A continuación se explican los principios básicos de este método de correspondencia, que aparece con frecuencia en la literatura de visión estéreo [18,20,23,24].

Métodos de correspondencia estéreo basados en la intensidad

La única información utilizada para efectuar la correspondencia es la intensidad de los pixels de las imágenes [25]. Esta estrategia supone que todos los puntos de la escena deben tener la misma intensidad en cada imagen. Como es evidente, en la mayoría de los casos no ocurre así, ya que las imágenes se toman desde distintos puntos de vista, y no captan la escena con la misma iluminación.

Por otra parte, en una misma imagen, existe un gran número de puntos con intensidades similares. Para poder minimizar la búsqueda y evitar falsos emparejamientos, es necesario elegir una medida de similitud entre patrones de intensidad. Estos patrones, generalmente, son regiones de la imagen, definidas como pequeñas ventanas alrededor del pixel (ver figura 3).

Considerando que el sistema estéreo capta dos imágenes representadas por $I_i(i, j)$ (imagen izquierda) e $I_d(i, j)$ (imagen derecha), el principio general de la búsqueda de emparejamientos es el de comparar una

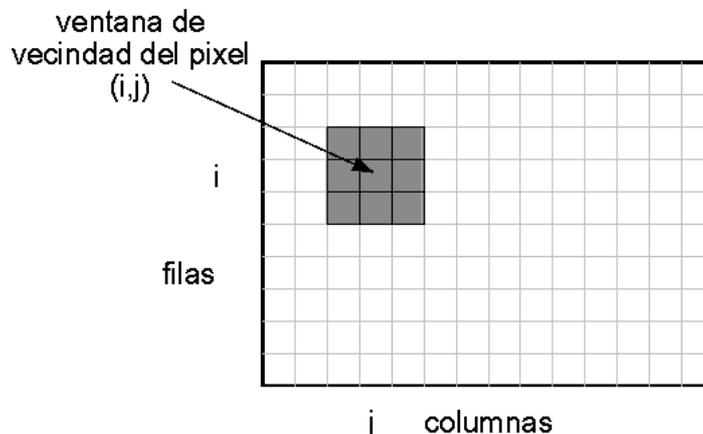


Fig. 3. Patrón de emparejamiento, ventana de vecindad de un pixel.

cierta vecindad de un pixel de la imagen $I_i(i, j)$, con una vecindad del mismo tamaño alrededor de cada uno de los pixels de la correspondiente imagen $I_d(i, j)$. La comparación se realiza, en el caso más general, con los siguientes **criterios de similitud**:

- la **suma de los cuadrados de las diferencias (SDC)**, que se define mediante la ecuación 5:

$$SDC(\Delta m, \Delta n) = \sum_{i,j} [I_i(i, j) - I_r(i - \Delta m, j - \Delta n)]^2 \quad (5)$$

- y la **correlación cruzada (CC)**, que se define con la ecuación 6:

$$CC(\Delta m, \Delta n) = \sum_{i,j} I_i(i, j) I_r(i - \Delta m, j - \Delta n) \quad (6)$$

donde $(\Delta m, \Delta n)$ es la disparidad entre las posiciones de las imágenes de dos regiones emparejadas, en las direcciones i y j del plano de la imagen respectivamente. Se entiende por **disparidad** la diferencia en posición (número de filas y columnas) existente entre un punto de la imagen $I_i(i, j)$ y el punto homólogo en la imagen $I_d(i, j)$.

Se obtiene, para cada pixel, una curva de resultados en función de las disparidades; la disparidad que se almacena será aquella que corresponda al valor máximo en el caso de CC y al valor mínimo en el de



Imagen izquierda



Imagen derecha



Imagen de profundidades

Fig. 4. Imagen de disparidad, resultado del emparejamiento.

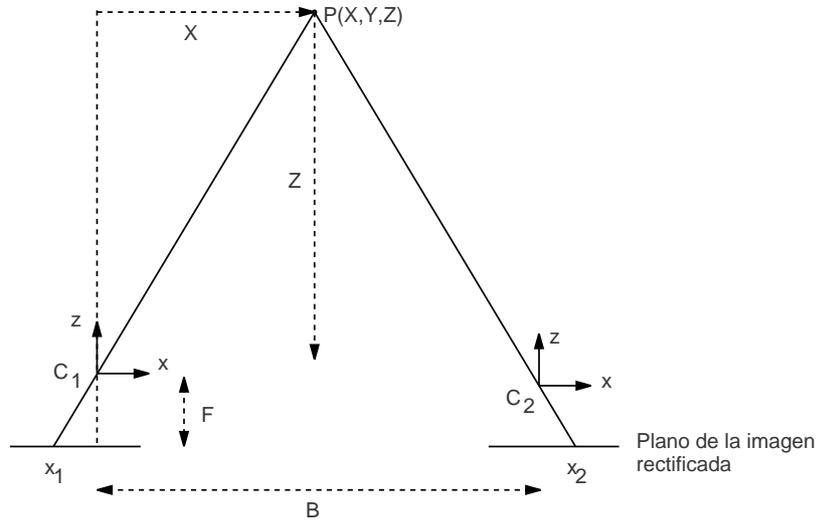


Fig. 5. Triangulación.

SDC. El resultado de esta operación es una imagen de niveles de gris, donde se representa la disparidad. En la figura 4 se observa el resultado de aplicar la correlación cruzada a dos imágenes. Se comprueba en la imagen de profundidades (o de disparidad) que los niveles de gris más claros corresponden a los objetos más cercanos.

Puede encontrarse información más detallada sobre algoritmos de correlación en [26].

c) Obtención de los datos tridimensionales

Tras completar la fase de emparejamiento, la última fase es el proceso de interpolación de la superficie. Se efectúa a partir de la imagen de disparidad utilizando una relación de triangulación. Con la triangulación pasiva, las imágenes tomadas desde distintos puntos de vista son analizadas para extraer el mapa de profundidades de la escena.

Las relaciones de la triangulación de base para dos cámaras perfectamente alineadas son las siguientes (notaciones definidas a partir de la figura 5):

$$\frac{X}{Z} = \frac{x_1}{F} \quad (7)$$

$$\frac{Z}{B} = \frac{F + Z}{B + (x_1 - x_2)} \quad (8)$$

La profundidad Z se hace por relación de referencia (C_1, x, y, z) para la ecuación 9:

$$Z = \frac{BF}{d(x_1, y_1)} \quad (9)$$

donde

- B es la distancia entre los dos centros C_1 y C_2 , definida como línea base.
- F es la distancia focal.
- $d(x_1, y_1)$ es la disparidad asociada al punto (x_1, y_1) .
- (x_1, y_1) es la proyección del punto P en el plano de la imagen rectificada número 1.

Las coordenadas X e Y se deducen a partir del modelo *pin-hole* de la cámara por las siguientes ecuaciones:

$$X = x_1 \frac{Z}{F}; Y = y_1 \frac{Z}{F} \quad (10)$$

4.2. Ventajas e inconvenientes de la visión estéreo

Los robots autónomos actuales que navegan en entornos no estructurados, utilizan sónares y telémetros láser como sensores espaciales primarios. Como ventajas de los sistemas estéreo frente a estos sistemas tradicionales podemos citar los siguientes:

- Permiten producir mapas de profundidad densos, y la imagen de barrido se representa con una velocidad interesante para las aplicaciones de la robótica móvil.
- Permite detectar mejor los objetos pequeños y estrechos.
- La visión estéreo es un sensor pasivo, por ello no es necesario utilizar una iluminación especial o fuentes de energía. Adquieren los datos de forma no invasiva, no alterando el entorno.
- Por último, puede usarse en aplicaciones más específicas en las que la información visual juega un papel básico, sin requerir ninguna modificación.

A pesar de estas ventajas, en la actualidad muy pocos robots móviles utilizan la visión estéreo para adquirir datos en tiempo real. Fundamentalmente se debe a lo siguiente:

- Son menos robustos a condiciones de niebla o iluminación inadecuada que los telémetros.
- Son sistemas difíciles de calibrar.
- Requieren un elevado coste computacional para obtener datos de profundidad densos y exactos en tiempo real [17]. Un camino obvio para mejorar los tiempos de ejecución, es desarrollar un hardware específico o realizar un procesamiento paralelo como el propuesto por Majumdar en [23].
- No son capaces de detectar profundidades en superficies sin textura. Por ejemplo, el robot Nomad de la CMU utilizó inicialmente un equipo estéreo para la detección de obstáculos, que fue posteriormente sustituido por un telémetro láser debido a la falta de textura del hielo [27].
- Trabajar con imágenes en niveles de gris, limita a 256 los niveles disponibles para representar las profundidades de un entorno. El reparto de estos niveles no es lineal. En la ecuación 9 se observa que la disparidad (profundidad en niveles de gris) es inversamente proporcional a la profundidad en metros.

Con el progreso de las técnicas informáticas y electrónicas se espera un aumento significativo de la explotación de este tipo de captadores, sobre todo para entornos naturales, ya que se resuelve una de las mayores limitaciones, que es su alto coste computacional.

4.3. Aplicabilidad en entornos exteriores

Para comprobar la viabilidad del sensor, trabajando en entornos exteriores, se realizaron una serie de pruebas con un equipo de visión estéreo TRICLOPS montado sobre una torreta pan/tilt que permite realizar un barrido del entorno, en nuestro caso horizontal, y obtener así una información similar a la del láser tridimensional. Los resultados se muestran a continuación:

En la imagen 7 se muestran los puntos tridimensionales obtenidos de la imagen real 6 con el equipo estéreo. Puede observarse claramente como a grandes distancias existe muy poca discriminación en las profundidades. Por ejemplo, la pared del fondo del laboratorio, aparece en el mapa de profundidades como circular, a unos 13 m. del sensor. También se observa que para objetos continuos con gran variación de profundidades, como es el caso de la pared lateral, se produce una rotura (en la figura 7 rodeado con un círculo) al no ser capaz de realizar una discretización más homogénea.

Interpolación sub-pixel

Para conseguir mejoras significativas en la precisión de un mapa de profundidades estéreo y resolver el problema de discretización planteado, puede utilizarse en el cálculo de las disparidades una precisión sub-



Fig. 6. Imagen real. Laboratorio de Robótica de la Universidad Carlos III de Madrid.

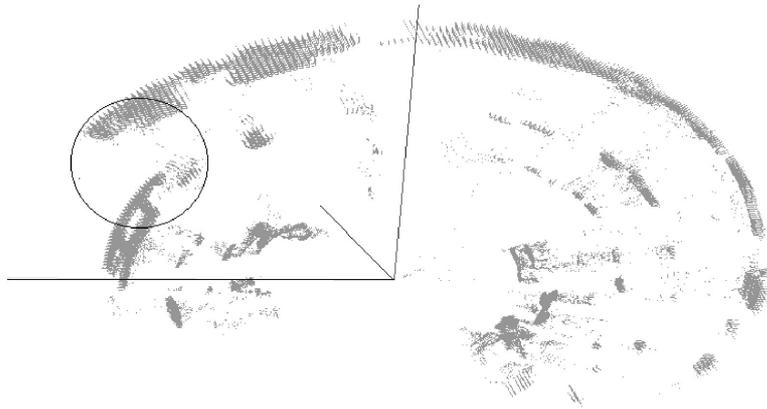


Fig. 7. Información tridimensional obtenida con el equipo estéreo TRICLOPS.

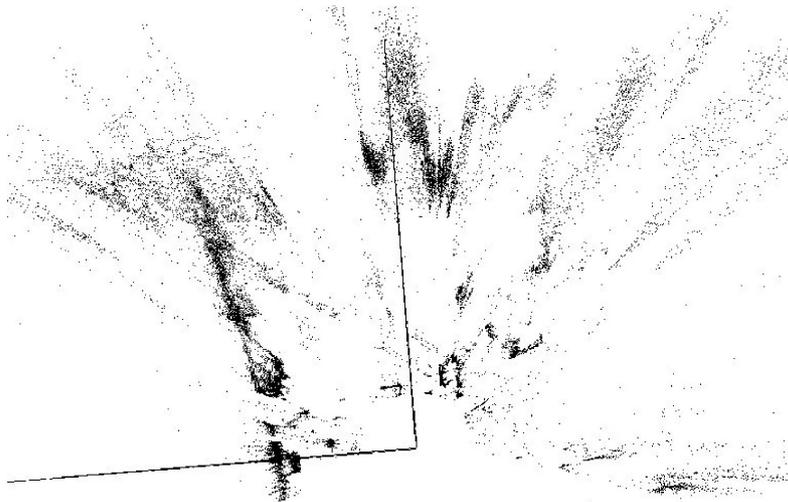


Fig. 8. Información tridimensional obtenida con el equipo estéreo TRICLOPS, utilizando interpolación sub-pixel.

pixel. La precisión sub-pixel se basa en la estimación, a partir de la interpolación de una función, de los valores de emparejamiento [26].

En la figura 8 se muestra el cálculo de los puntos tridimensionales de la figura 6 utilizando la interpolación sub-pixel. Se resuelve así el problema anterior; las paredes del laboratorio ya no adoptan apariencia circular y los objetos no presentan discontinuidades en función de la distancia.

5. DESARROLLO DEL SISTEMA SENSORIAL

Las tendencias y necesidades de la robótica móvil actual a la hora de diseñar una plataforma se dirigen a conseguir mayor autonomía a la vez que robustez. La forma en que se definan ambos factores depende de las tareas que vaya a realizar y del entorno en el que ejecuten esas tareas.

Existen una serie de criterios a tener en cuenta a la hora de diseñar o elegir una plataforma de experimentación. Estos criterios son:

- Operación flexible y robusta. Un robot debe adaptar su comportamiento al entorno de ejecución.
- Utilización de recursos. Recursos tales como la potencia, el almacenamiento de datos, y el ancho de banda de comunicación son restricciones con las que el robot se va a encontrar. Para representar sus objetivos científicos y completar las misiones con éxito, el robot necesita hacer uso completo de los recursos disponibles, incluso cuando cambian los valores esperados.
- Los sistemas que se instalan en la plataforma deberán ser robustos para que puedan adaptarse a las diferentes condiciones y cambios del entorno. Además de ser resistentes a tensiones mecánicas y térmicas.
- También es necesario una interfaz hombre-máquina amigable.

5.1. Plataforma experimental

La plataforma experimental presentada en este artículo es un robot de exteriores llamado *GOLIAT* construido e integrado gracias al proyecto denominado *Sistema de percepción y modelado del entorno para vehículos autónomos de entornos exteriores*, de referencia TAP1997-0296. Este robot ha sido desarrollado en la Universidad Carlos III de Madrid en el Departamento de Ingeniería de Sistemas y Automática [28].

GOLIAT es un robot autónomo capaz de trabajar en un terreno natural no estructurado, puede desplazarse con un alto grado de autonomía y un conocimiento mínimo previo del entorno de trabajo (ver figura 9). La plataforma del robot es un vehículo con tracción por orugas de 1.5 Ton MD-15 de la empresa Modulaire. Este vehículo, con motor diesel, se mueve gracias a los motores hidráulicos que hacen rotar las



Fig. 9. Robot de exteriores *GOLIAT*.

cadena de las orugas. Sobre la plataforma se han colocado diferentes dispositivos sensoriales que forman su sistema de percepción. El sistema de percepción está formado por tres sub-sistemas: un telémetro láser tridimensional (que detallamos en secciones posteriores), y dos sistemas de visión estéreo; un sistema tricámara TRICLOPS de Pointgrey Ltd. y una cabeza binocular con dos cámaras en blanco y negro. Además de estos sistemas, capaces de tomar medidas tridimensionales, el robot dispone de otros sistemas sensoriales necesarios para el modelado del entorno, como son una brújula, un sistema de medida inercial (que no se utiliza en este trabajo) y un sistema GPS para la obtención de medidas de posicionamiento global.

Para el control de todos los dispositivos que se encuentran a bordo del robot, se le ha dotado de una arquitectura hardware y software, capaz de procesar la información que recibe del sistema sensorial.

5.2. Arquitectura hardware

La arquitectura hardware sobre la que se ha desarrollado el robot *GOLIAT* se compone de tres ordenadores embarcados en el robot, y un cuarto ordenador que realizará las funciones de estación base. La comunicación entre computadores se realiza gracias a un conmutador (para la comunicación a bordo del robot) y un enlace de radio para comunicar al robot con el exterior, donde se encontrará la estación base. De los computadores embarcados, dos están dedicados a tareas relacionadas con el control del sistema de percepción y procesamiento de los datos obtenidos. El tercero se dedica al control del sistema. Los equipos que se disponen, así como sus características son los siguientes [29]:

- Conmutador. Empleado para la interconexión de los equipos dentro del robot. Es un conmutador FastEthernet dual, de la casa Compaq, con un ancho de banda en el crossbar interno de 1.6 Gbps, lo que permite un ancho de banda por puerto de 200 Mbps.
- Enlace de radio. Empleado para la comunicación entre el robot y la estación base. Es un punto de acceso Ethernet de la marca Aironet con una tasa de transferencia bruta de 2 Mbps. La tasa de transferencia sostenida (empleando Ethernet y TCP/IP) conseguida es de 850 Kbps.
- PC1. Equipo dedicado al procesamiento de imágenes. En él se encuentran conectados una tarjeta de procesamiento de imágenes Matrox Meteor RGB y una cámara de visión estereoscópica Triclops. El equipo se compone de un Pentium MMX a 200 Mhz con 32 MB, con sistema operativo Windows NT.
- PC2. Equipo encargado de la parte de control del sistema. Es también un Pentium MMX a 200 Mhz con 32 MB. El sistema operativo que posee es Linux. Para las tareas de control dispone de una tarjeta PC-LAB PCL-812, para la que se ha realizado un driver.
- PC3. Encargado principalmente de las tareas de procesamiento tales como las de modelado. Tiene asociado una torreta PAN/TILT para el control en posición del sistema estéreo y un telémetro láser SICK, además de una tarjeta controladora de dos ejes DMC-1000 de la casa GALIL. Es un ordenador biprocesador Pentium MMX II a 266 Mhz con 128 MB de memoria RAM. Este equipo funciona bajo sistema operativo Linux con soporte *SMP* (*Symmetric Multi-Processors*) para el multiprocesamiento simétrico.
- Estación Base. Se encarga principalmente de las tareas de representación de los datos adquiridos y procesados por el robot. Se trata de una estación Silicon Graphics modelo O2, 256 MB de memoria RAM. Todos los equipos poseen tarjetas de comunicaciones FastEthernet a 100 Mbps Duales de la empresa 3Com que permiten tasas de transferencia de 100 Mbps simultáneas tanto en transmisión como en recepción.

En la figura 10 se muestra el esquema de la arquitectura hardware del sistema.

5.3. Arquitectura software

La arquitectura software del robot tiene una estructura modular. Para cada elemento del robot (sensores, tareas) se ha preparado un módulo y un conjunto de librerías comunes entre los módulos. Cada uno de estos módulos ha sido realizado por diferentes integrantes del equipo de investigación dentro del grupo de

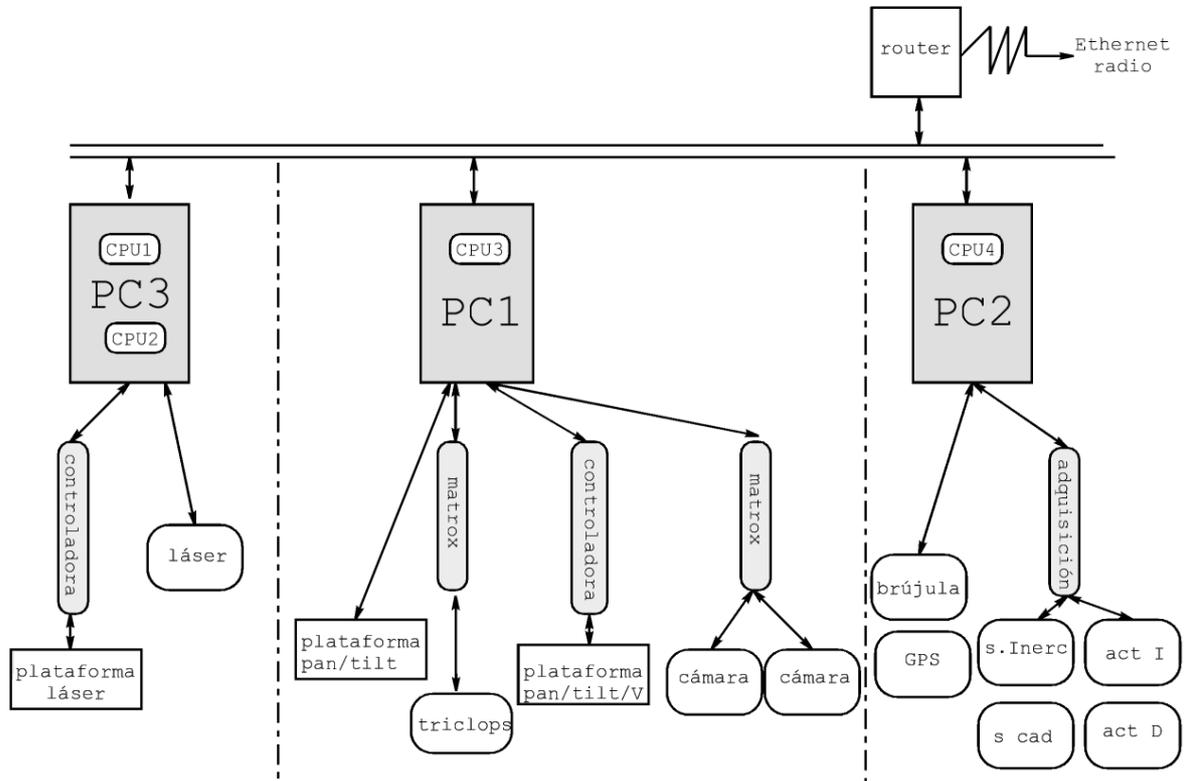


Fig. 10. Arquitectura Hardware.

investigación. Se ha diseñado como un sistema distribuido con computadores heterogéneos. Se utiliza PVM como software de comunicaciones, y el compilador de interfases GlueSoft para los módulos. La finalidad de esta herramienta es la de enlazar, pegar y unir distintos módulos software.

5.4. Sistema sensorial elegido

Teniendo en cuenta las pruebas de aplicabilidad que se han mostrado en las secciones 2.3 y 4.3, se ha optado por la utilización de la telemetría láser como fuente de información para el modelado del entorno. Se ha elegido la tecnología láser frente a otras metodologías, tales como la visión estéreo, por su independencia a las condiciones de iluminación y porque proporciona información tridimensional del entorno para distancias de hasta 25 m con gran fiabilidad, siendo estas características muy adecuadas para su aplicación en entornos exteriores. Además, la telemetría láser reduce la necesidad de procesamiento de los datos procedentes del sensor para la obtención de un mapa de profundidades.

Dado el escaso desarrollo tecnológico en los sensores tridimensionales para el modelado en entornos exteriores sobre robots móviles autónomos, ha sido necesario diseñar una plataforma sensorial propia. Se ha construido el telémetro láser tridimensional presentado en la figura 11.

La plataforma sensorial está formada por un scanner láser bidimensional PLS de Sick preparado para capturar 361 medidas en un barrido horizontal de 180 grados (se tomarán medidas cada $0,5^\circ$) y un motor de corriente continua más una reductora que se encarga de realizar el barrido vertical.

El PLS garantiza unos errores en la medida de profundidad que no superan los 20 milímetros. Este error se va a ver influenciado por dos parámetros, la distancia a medir y el ángulo de disparo del haz láser (de 0° a 180°).

En cuanto al conjunto motor-reductora, el motor es de corriente continua (Motor DC 3557K048CR de MINIMOTOR) con una masa de 290 gramos, potencia de salida de 72.7 Watios y velocidad de giro máxima de 5300 rpm. A este motor se le añade un reductor planetario 38/1S(43:1) con coeficiente de

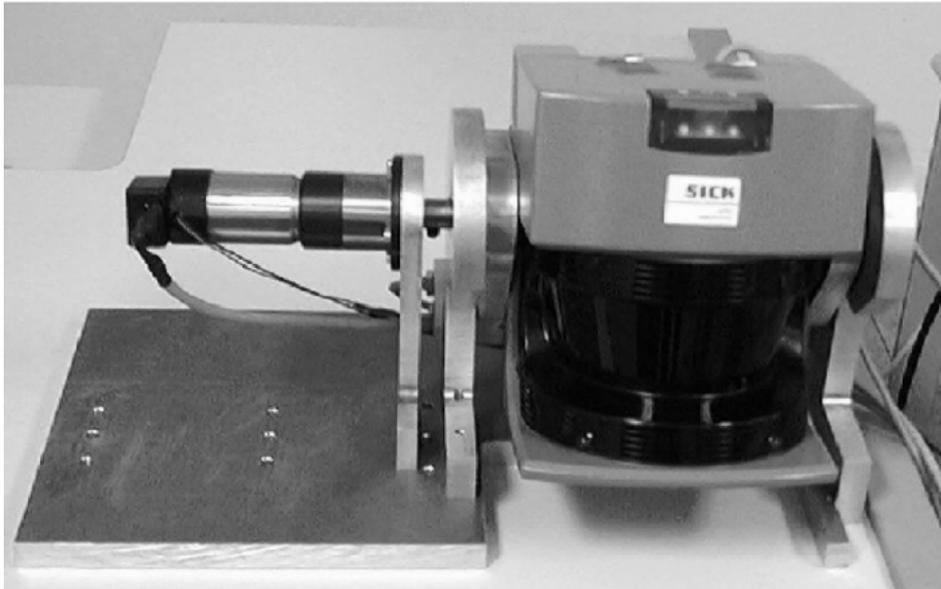


Fig. 11. Sistema sensorial utilizado.

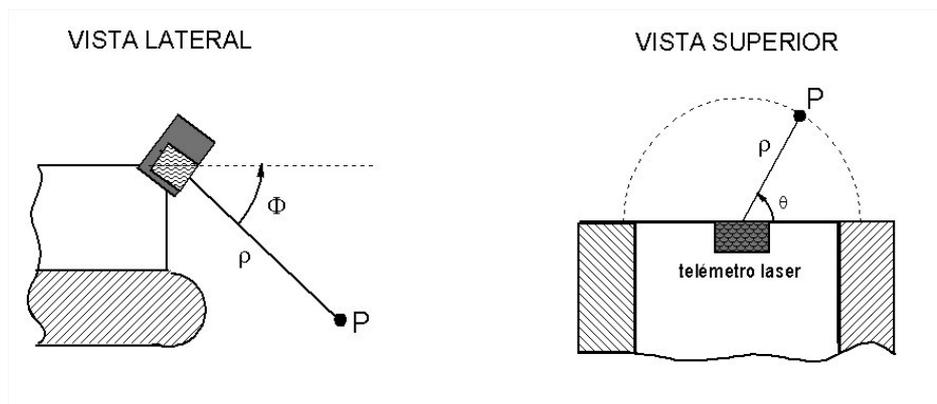


Fig. 12. Información obtenida por el láser.

reducción 1/43 y un encoder incremental HEDS 5540A14 de MINIMOTOR que mide 2000 pulsos por vuelta.

Todo el sistema motor es controlado desde el PC3 en Linux con una tarjeta controladora de 2 ejes DMC-1000 de GALIL. Para encontrar la posición inicial del sistema (HOME) se utiliza un sensor de proximidad CONTRINEX DW-AD-613-M8 unido a la plataforma y conectado a la tarjeta.

6. OBTENCIÓN DE LA INFORMACIÓN TRIDIMENSIONAL

Para la obtención de la información del entorno, se dispone del sensor láser tridimensional montado sobre la plataforma del robot, que proporciona una imagen de profundidad. Para cada ángulo vertical ϕ (proporcionado por el motor acoplado al telémetro) y cada ángulo horizontal θ (de barrido del láser) se obtiene una distancia a los objetos ρ utilizando la medida del tiempo de vuelo (TOF) del haz láser. Se consigue así una imagen tridimensional del entorno, formada por puntos $P(\rho, \theta, \phi)$ como el de la figura 12.

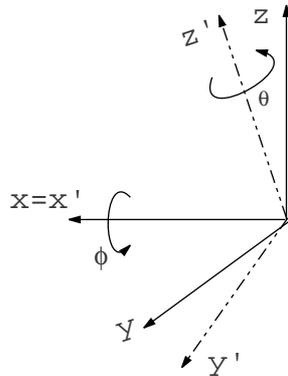


Fig. 13. Rotación de los ejes del sistema sensor.

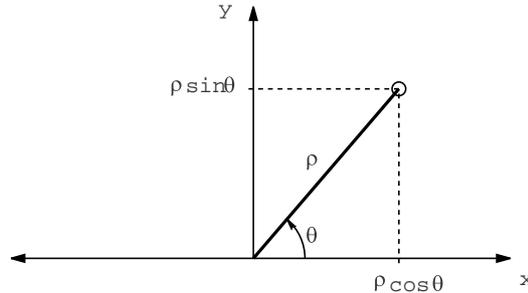


Fig. 14. Transformación de polares a cartesianas en 2D.

La información en coordenadas polares se transfiere al ordenador de control mediante comunicación puerto serie RS-232 a una velocidad de 58800 baudios.

En las siguientes secciones se obtendrá una representación del entorno percibido por el robot, en coordenadas cartesianas, referidas a un sistema de referencia (SDR) local situado en la base del robot. Se realizará también la transformación de la información a un sistema de referencia global, para su integración con la información del entorno, almacenada por el robot en percepciones sucesivas.

7. CAMBIO EN EL SISTEMA DE REFERENCIA

El sistema telémetro láser tridimensional produce una imagen esférica del entorno de la forma $f(i, j) = (\rho, \theta, \phi)$. Para poder obtener una representación del entorno, en el sistema de referencia de la base del robot, es necesario realizar las transformaciones que se muestran a continuación.

7.1. Transformación de coordenadas polares a cartesianas en el sistema de referencia del láser

El sensor mide distancias en barridos verticales, giros en torno a los ejes z , y horizontales, giro en torno al eje x , proporcionado por el motor de corriente continua. El funcionamiento de los dos ejes es independiente (el barrido horizontal se realiza para un ángulo vertical fijo) por lo que se considera que el sistema trabaja en modo correlado, como se representa en la figura 13. Es decir, el barrido horizontal se realiza alrededor del eje z' , y no alrededor del eje z .

La transformación a coordenadas cartesianas en el SDR del sensor se realiza de la forma siguiente:

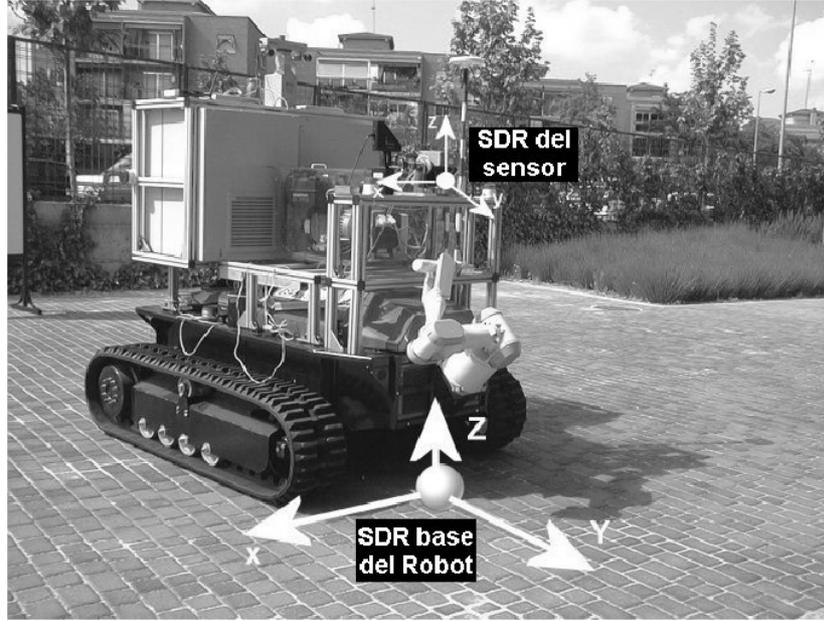


Fig. 15. Sistemas de referencia del sensor y del robot.

1. En primer lugar, se considera el caso plano de láser bidimensional, y se realiza la transformación a coordenadas cartesianas. Se considera la transformación para el plano $z = 0$. Las ecuaciones deducidas de la figura 14 son las siguientes:

$$\begin{aligned} x &= \rho \cos \theta \\ y &= \rho \sin \theta \\ z &= 0 \end{aligned} \quad (11)$$

2. Esto es cierto para cada barrido horizontal, considerando que no existe el barrido vertical ($\phi = 0$). En el caso de láser tridimensional con barrido vertical ϕ , para conseguir la tercera dimensión, es necesario realizar una rotación en el eje x un ángulo de barrido vertical ϕ (figura 13). Con ello se obtienen las coordenadas cartesianas en el sistema de referencia del sensor que se representa en la figura 15.

La transformación se realiza con la ecuación 12:

$$v' = R_x v = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi & 0 \\ 0 & \sin \phi & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (12)$$

siendo (R_x) la matriz de rotación a utilizar en este caso.

7.2. Transformación de coordenadas cartesianas al sistema de la base del robot

Para poder obtener la información sensorial en el sistema de referencia del robot, situado en la base (figura 15) necesitamos realizar una traslación en el eje z una distancia H , igual a la existente entre el sistema de referencia del láser y el de la base del robot. Las traslaciones en el eje x e y no se consideran por encontrarse ambos sistemas en la misma vertical. También es necesario incorporar los ángulos de cabeceo (*pitch*, β) y alabeo (*roll*, γ) del robot, es decir, sus inclinaciones laterales y frontales (ver figura 16). Esta información es proporcionada por una brújula magnética instalada en el robot.

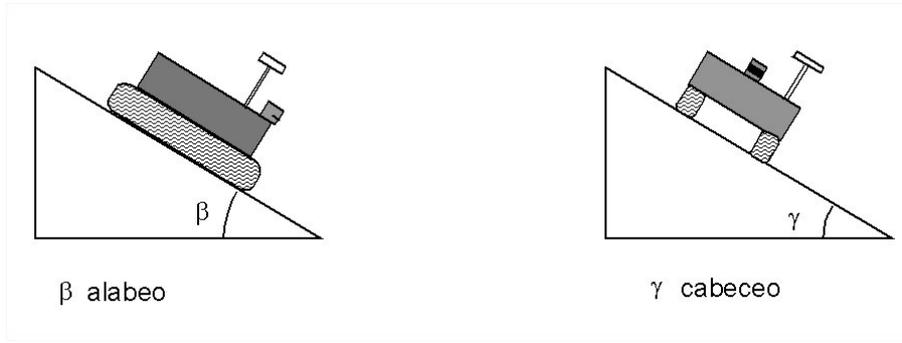


Fig. 16. Ángulos de cabeceo y alabeo del robot medidos con la brújula.

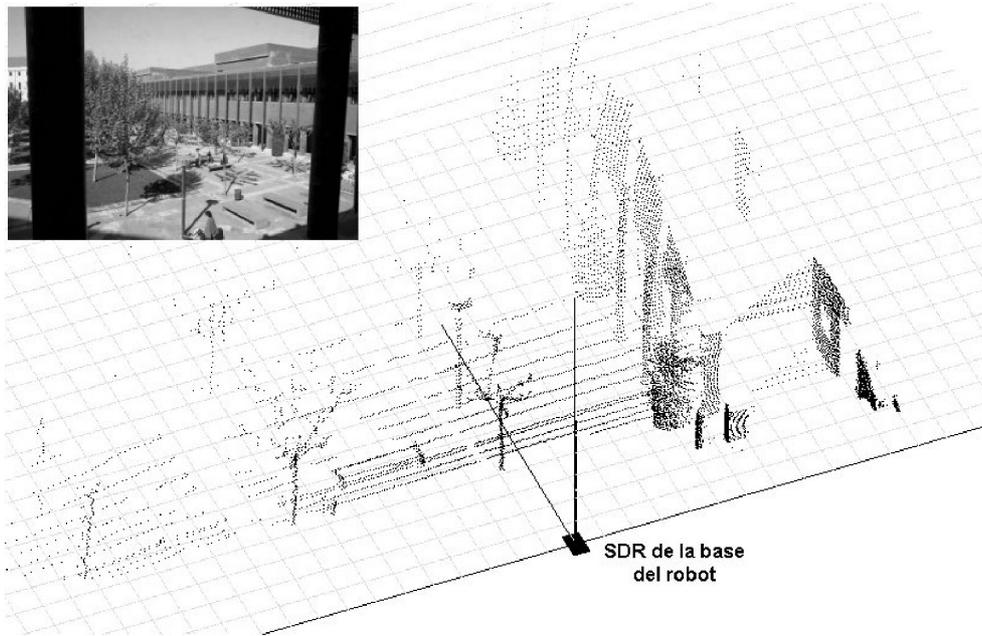


Fig. 17. Coordenadas cartesianas tridimensionales en el SDR de la base del Robot.

Para obtener los datos en función del nuevo sistema de referencia, es necesario realizar una rotación un ángulo γ igual al valor del alabeo, en el eje y y una rotación un ángulo β igual al valor del cabeceo en el eje x , además de la traslación H en el eje z . Las matrices de transformación utilizadas para este caso son:

$$T_z(H) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & H \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad R_y(\gamma) = \begin{pmatrix} \cos \gamma & 0 & -\sin \gamma & 0 \\ 0 & 1 & 0 & 0 \\ \sin \gamma & 0 & \cos \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad R_x(\beta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \beta & -\sin \beta & 0 \\ 0 & \sin \beta & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (13)$$

y la ecuación de transformación:

$$v'' = R_y(\gamma) \cdot R_x(\beta) \cdot T_z(H) \cdot v' \quad (14)$$

En la figura 15 se representan los sistemas de referencia locales utilizados en este trabajo, y en la figura 17 se muestran los resultados obtenidos al realizar la transformación de los datos del láser al sistema de referencia de la base del robot.

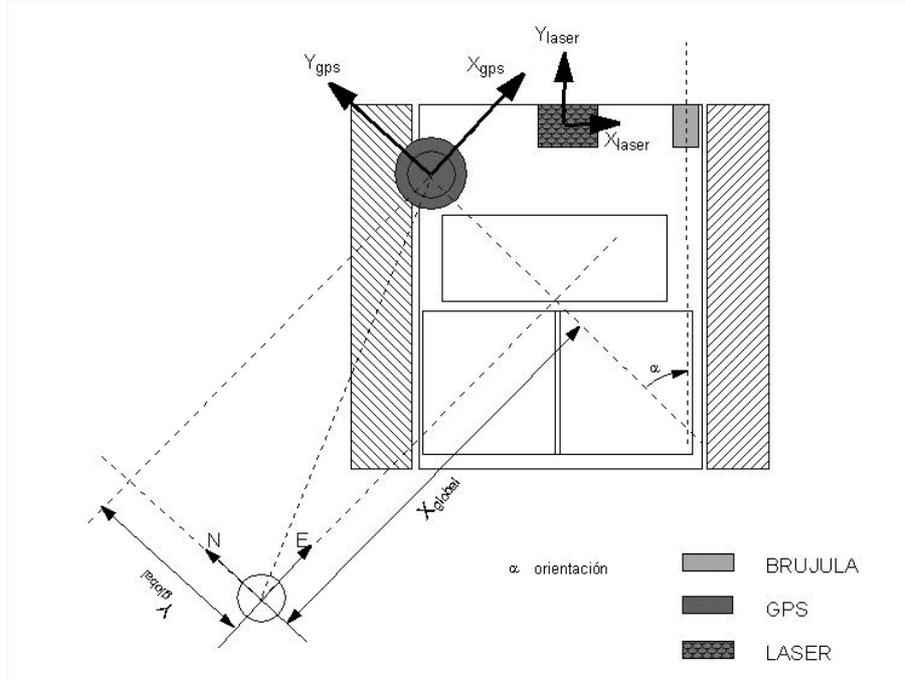


Fig. 18. Sistemas de referencia y localización de los sensores a bordo del robot.

8. TRATAMIENTO DE LA INFORMACIÓN PARA UN MODELO GLOBAL

El modelo local del entorno permite al robot evolucionar e interactuar con el mundo real más próximo a él. Pero para que un robot sea capaz de realizar una tarea completa, en un entorno de grandes dimensiones y desconocido, éste necesita construir de manera incremental, y a partir de percepciones sucesivas, un modelo global del entorno que le permita tener una *memoria* para construir su mundo. Este paso se consigue trasladando los diferentes mapas locales a una base de datos global. Para ello es necesario la información de dos sensores: la brújula y el GPS. Además de la brújula, utilizada en el apartado anterior para la transformación de coordenadas a un sistema de referencia local, la plataforma experimental presentada en este artículo dispone de un sistema de posicionamiento global por satélite (GPS), que proporciona el sistema de referencia Global. El sistema de referencia global está formado por tres ejes ortogonales (XYZ), con el eje X positivo en la dirección ESTE y el eje Y positivo en la dirección NORTE. En la figura 18 pueden observarse los diferentes sensores utilizados, su ubicación en el robot y las posiciones de los ejes en cada uno de ellos.

Las matrices de transformación son las siguientes:

$$T_{xyz}(x_g, y_g, z_g) = \begin{pmatrix} 1 & 0 & 0 & x_g \\ 0 & 1 & 0 & y_g \\ 0 & 0 & 1 & z_g \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad R_z(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 & 0 \\ \sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (15)$$

y la transformación:

$$v''' = T_{xyz}(x_g, y_g, z_g) \cdot R_z(\alpha) \cdot v'' \quad (16)$$

Donde (x_g, y_g, z_g) son las distancias en los ejes del GPS (X, Y y Z , respectivamente) entre el origen del sistema de referencia de la base del robot y el origen del sistema global, y α es el ángulo formado por el robot y la dirección Norte.



Fig. 19. Imagen real.

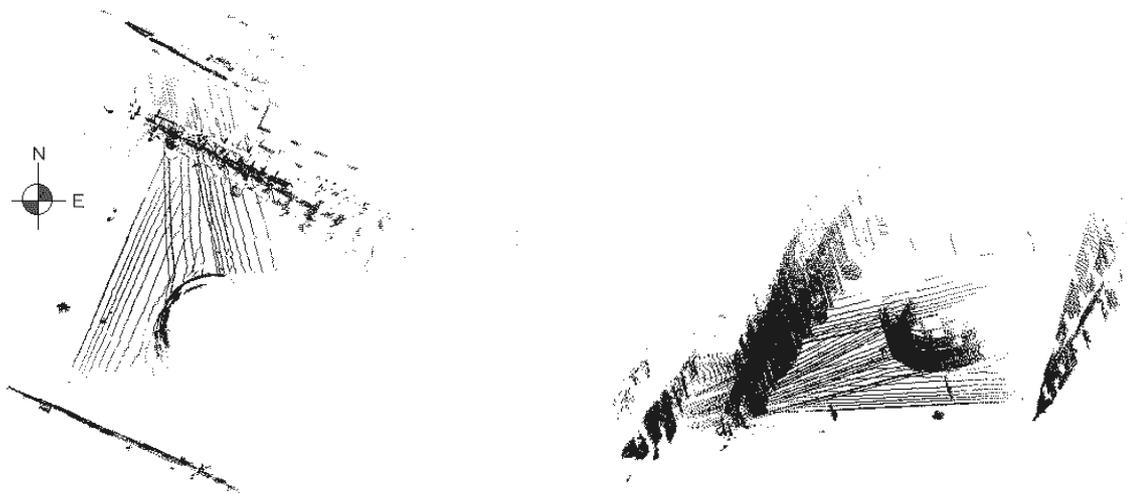


Fig. 20. Fusión de datos tridimensionales captados con el telémetro láser tridimensional.

Se han realizado los cambios de coordenadas de una base de datos de imágenes tomadas en la parte exterior de la Universidad Carlos III para poder llevar a cabo las pruebas. En las figuras 19 y 20 se muestra la fusión de tres imágenes utilizando la información de la brújula y del GPS. La imagen de la izquierda de la figura 20 representa la vista superior y la de la derecha, una vista lateral.

AGRADECIMIENTOS

Los autores quieren agradecer sinceramente la financiación proporcionada por el Gobierno Español a través del proyecto CICYT TAP 1997-0296 y el proyecto DPYT DPI 2000-0425.

BIBLIOGRAFÍA

- [1] D. Langer, J.K. Rosenblatt, and M. Hebert, “An integrated system for autonomous off-road navigation,” in *IEEE International Conference on Robotics and Automation*, 1994, vol. 1, pp. 414–19.
- [2] K. Mayora, I. Moreno, and G. Obieta, “Perception system for navigation in a 3d outdoor environment,” in *rd IFAC Symposium on Intelligent Autonomous Vehicles*, Madrid, Spain, March 1998.
- [3] In S. Kweon and T. Kanade, “High resolution terrain map from multiple sensor data,” in *IEEE International Workshop on intelligent robots and systems*, 1990.
- [4] A. de la Escalera, L. Moreno, J.M. Armingol, and M.A. Salichs, “Real time mobile robot localization by using a laser and a geometric map,” in *Intelligent autonomous Vehicles 95 (IAV95)*, 1995, pp. 197–201.
- [5] E. Krotkov and R. Hoffman, “Terrain mapping for a walking planetary rover,” *IEEE Transactions on Robotics and Automation*, vol. 10, no. 6, pp. 728–739, December 1994.
- [6] A. de la Escalera, M.A. Salichs, L. Moreno, and J.M. Armingol, “3d perception in non-structured environments using active vision,” in *IFAC Artificial Intelligence in Real-Time Control*, 1994, pp. 341–346.
- [7] L. Matthies and P. Grandjean, “Stochastic performance modeling and evaluation of obstacle detectability with imaging range sensors,” *Proceedings. 1993 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 657–8, 1993.
- [8] H.R. Everett, *Sensors for Mobile Robots, Theory and Application*, A K Peters, Ltd. Wellesley, Massachusetts, 1995.
- [9] J. Lobo, L. Marques, J. Dias, U. Nunes, and A. T. de Almeida, “*Sensors for Mobile Robot Navigation*”, *Autonomous Robotic Systems*, pp. 50–81, Lecture Notes in Control and Information Sciences 236. Springer-Verlag, February 1998.
- [10] Simon Lacroix, *Stratégies de Perception et de Déplacement Pour la Navigation D’un Robot Mobile Autonome Enenvironnement Naturel*, Ph.D. thesis, Université Paul Sabatier de Toulouse, Avril 1995.
- [11] M. Hebert, “Active and passive range sensing for robotics,” in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA ’00)*, April 2000, vol. 1, pp. 102 – 110.
- [12] F. Nashsashibi, P. Fillatreau, B. Dacre-Wright, and T. Simeon, “3-d autonomous navigation in a natural environment,” in *IEEE International conference on Robotics and Automation*, San Diego (USA), 1994, vol. 1, pp. 433–439.
- [13] D. Hähnel, W. Burgard, and S. Thrun, “Learning compact 3d models of indoor and outdoor environments with a mobile robot,” *Robotics and Autonomous Systems*, vol. 44, no. 1, pp. 15–27, July 2002.
- [14] C. Rocchimi, P. Cignoni, C. Montani, P. Pingi, and R. Scopigno, “A low cost 3d scanner based on structured light,” *Eurographics 2001*, vol. 20, no. 3, 2001.
- [15] A. de la Escalera, L. Moreno, M.A. Salichs, and J. M. Armingol, “Continuous mobile robot localization by using structured light and a geometric map,” *International Journal of Systems Science*, vol. 27, no. 8, pp. 771–782, 1996.
- [16] R. Klette, K. Schlüns, and A. Koschan, *Computer Vision Three-Dimensional Data from Images*, Springer, 1998.
- [17] D. Murray and J. Little, “Using real-time stereo vision for mobile robot navigation,” *Autonomous Robots*, vol. 8, no. 2, pp. 161 – 171, April 2000.
- [18] D. Murray and C. Jennings, “Stereo vision based mapping and navigation for mobile robots,” *IEEE International Conference on Robotics and Automation*, vol. 2, pp. 1694–9, 1994.
- [19] Patricia Lasserre, *Vision Pour la Robotique En Environnement Naturel*, Ph.D. thesis, Université Paul Sabatier de Toulouse, Laboratoire d’Analyse et d’Architecture des Systèmes du CNRS, 1996.
- [20] H. Al Haddad, *Contrôle Par Vision Du Mouvement D’un Robot Mobile En Environnement Naturel*, Ph.D. thesis, Université Paul Sabatier de Toulouse, Laboratoire d’Analyse et d’Architecture des Systèmes du CNRS, 1998.
- [21] E.E. Hemayed, A. Sandbek, and A. Wassal, “Investigation of stereo-based 3d surface reconstruction,” *Proc. SPIE*, vol. 3023, pp. 191–202, 1997.
- [22] E.E. Hemayed and Aly A. Farag, “A geometrical-based trinocular vision system for edges reconstruction,” in *IEEE Int. Conf. Image Processing*, October 1998, vol. 3, p. 162.

- [23] J. Majumdar and Seethalakshmy, “Efficient parallel processing for depth calculation using stereo,” *Robotics and Autonomous System*, pp. 1–13, 1997.
- [24] T. Kanade and M. Okutomi, “A stereo matching algorithm with an adaptative windows: Theory and experiment,” *International Conference on Robotics and Automation*, pp. 1088–1095, April 1991.
- [25] Vishvjit S. Nalwa, *A Guided Tour of Computer Vision*, Addison-Wesley, cop., 1993.
- [26] O. Faugeras and B. Hot, “Real time correlation-based stereo: Algorithm, implementations and applications,” *Tech. Rep. 2013*, Institut national de recherche en informatique et en automatique, août 1993.
- [27] C. Urmsen, V. Shamah, J.P. Teza, M. D. Wagner, and D. Apostolopoulos, “A sensor arm for robotic antarctic meteorite search,” in *rd International Conference on Field and Service Robotics FSR2001*, Helsinki, Finland, June 2001.
- [28] C. Castejón, L. Moreno, and M.A. Salichs, “Traversability modelling in 3d environments,” in *rd International Conference on Field and Service Robotics FSR2001*, Finland, June 2001.
- [29] Cristina Castejón, *Modelado de zonas cruzables en entornos exteriores para robots móviles*, Ph.D. thesis, Universidad Carlos III de Madrid, Julio 2002.

PERCEPTION SYSTEM FOR A MOBILE PLATFORM IN OUTDOOR ENVIRONMENT

Abstract – Autonomy is a technological goal to be achieved not only in mobile robotics systems, but also in personal and industrial vehicles. This issue (autonomy) involves the skills development, in order to operate and to react in environments that can be, a priori, unknown. For this reason, it is necessary to obtain an environment model from sensor data. The modelling process is essential, because it will be the base for future decision-taking, in order to fulfil the proposed task. To accomplish this functionality, it is necessary to obtain information about the outdoor workspace, using the perception system. In the present work the built-in sensorial system, onboard a mobile robot for outdoor environments, is presented. This sensorial system will provide useful three-dimensional information to the following modelling stages.

