

Esquema de Navegación Reactiva con Sensores RGB-D

Reactive Navigation Scheme with RGB-D Sensors

ANDRÉS FELIPE SUÁREZ-SÁNCHEZ

*Ingeniero Electrónico y candidato a Magister en ingeniería con énfasis en automática
Universidad del Valle
andres.suarez@correounivalle.edu.co
Santiago de Cali, Colombia*

HUMBERTO LOAIZA-CORREA

*Doctor en Robótica y Visión Artificial, profesor titular de Escuela Ingeniería Eléctrica y Electrónica y Codirector Grupo
Percepción y Sistemas Inteligentes (PSI)
Universidad del Valle
humberto.loaiza@correounivalle.edu.co
Santiago de Cali, Colombia*

Fecha de recibido: 12/04/2014

Fecha de aceptado: 23/12/2014

Forma de citar: SUÁREZ, Andrés, LOAIZA, Humberto. Implementación de un esquema de navegación reactiva con sensores RGB-D . Rev. UIS Ingenierías,2015,vol.14,n1,p.p 7-19.

RESUMEN

En este artículo se presenta la integración de un robot móvil con un sistema de visión artificial, que permite al agente móvil detectar y evadir obstáculos presentes en su trayectoria. El sistema de visión tiene la capacidad de presentar un modelo 3D de la escena local como una representación de nubes de puntos. La integración hardware es conformada por una cámara RGB-D de bajo costo, una unidad de procesamiento y una plataforma robótica. Para el sistema de percepción visual se implementa un algoritmo de reducción de información 3D a 2D, similar a la información de suministrada por sensores de rango, esto con el propósito de reducir etapas adicionales de manipulación de datos y tiempos en la respuesta de los algoritmos de navegación. El sistema en conjunto se puso a prueba en un entorno controlado, superando diversas pruebas de navegación y evasión de obstáculos. Con este trabajo se inicia la exploración de las capacidades del sensor RGB-D en robótica móvil a nivel de la región del pacífico colombiano y es el primer paso de una propuesta de investigación que busca implementar una solución al problema de Localización y Mapeo Simultaneo (SLAM).

PALABRAS CLAVE: Visión Artificial, Robótica Móvil, Percepción, Navegación, Reconstrucción, Kinect, Nube de puntos, Campos potenciales.

ABSTRACT

This paper presents the integration of a mobile robot with a vision system, which allows to a mobile agent to detect and avoid obstacles in its path. The vision system has the ability to present a 3D model of the local scene as a point clouds representation. The hardware integration is formed by an RGB-D inexpensive camera, a processing unit and a robotic platform. For visual perception, an algorithm for reduction 3D to 2D information similar to the information provided by range sensors was implemented in order to reduce response time and manipulation data for navigation algorithms. The whole system was tested in a controlled environment, overcoming various tests of navigation and obstacle avoidance. This work exploring the capabilities of RGB-D sensors in mobile robotics at Colombian Pacific

region and is the first step of a research proposal aims to implement a solution to the problem of Simultaneous Localization and Mapping (SLAM).

KEYWORDS: Computer Vision, Mobile Robotics, Perception, Navigation, Reconstruction, Kinect, Point Cloud, Potential Fields.

1. INTRODUCCIÓN

Los robots móviles son sistemas que tienen la capacidad de navegar autónomamente en un espacio, lo que los hace muy atractivos para ser empleados como robots de servicio en aplicaciones de transporte como: automóviles totalmente autónomos para el transporte de personas dentro de la ciudad visión que se dará en un futuro cercano, transporte de material en entornos industriales. Otra de las aplicaciones son la exploración de terrenos hostiles y de difícil acceso, la interacción con personas como: guías turísticos en museos, personal de apoyo en centros hospitalarios, etc.

Para lograr total autonomía en un robot inicialmente se es necesario dotar de sistemas de percepción al agente móvil, esto con el fin de brindarle la información de la estructura del entorno; posteriormente se debe brindar un algoritmo que procese la información recibida de los sensores para proporcionarle información sobre la posición y orientación de los posibles obstáculos en su trayectoria. Como último nivel en la pirámide de control se encuentran los algoritmos que permiten la localización, el mapeo del entorno y la planificación de trayectorias “optimizadas” en un mapa global.

El problema de la percepción del entorno en la robótica móvil, es un campo de continua investigación. Uno de los propósitos de la percepción artificial es brindarle la mayor información del entorno a un agente móvil en el menor tiempo posible. Para un robot es importante percibir y entender la estructura de un ambiente para poder llevar a cabo una tarea asignada en lugares que suelen ser desconocidos. Es por eso que se han desarrollado una variedad de sistemas sensoriales que le permiten al robot captar esta información.

Para la percepción del entorno generalmente se han empleado sensores del tipo sonar y laser 2D (Bui, T., et al. 2009, Quian, Song, 2012), con este tipo de sensores se implementan estrategias de navegación que permiten realizar tareas complejas de localización y mapeo obteniendo buenos resultados. También se han desarrollado sistemas laser 3D que permiten capturar el entorno en su estructura básica de profundidad, altura y ancho, estos dispositivos brindan una mayor cantidad

de información que puede ser utilizada por el robot, igualmente con este tipo de sensores se logran realizar tareas complejas de navegación y mapeo, ofreciendo la capacidad de representar un mapa del entorno en un modelo de tres dimensiones generalmente representado por una nube de puntos (Nüchter, A., et al. 2007, Cole, Newman, P.2006)

Otro enfoque de los investigadores en percepción artificial y robótica ha hecho uso de sensores que brindan información visual del entorno. Los desarrollos realizados con este tipo de sensores han empleado cámaras monoculares como sensor principal o auxiliar para la navegación de robots (Davison, A., et al. 2007, Seo-Yeon, Jae-bok, 2011, Lemaire, T., et al. 2007), la dificultad de trabajar con este tipo de sensor es que no proporciona directamente la información 3D de la escena lo que hace complejo la estimación de la estructura y geometría del entorno, ante esta dificultad se han desarrollado sistemas de visión estéreo (Lemaire, T., et al. 2007, Viejo, Carloza, 2004, Cañas. et al. 2008), con este tipo de cámaras es posible la percepción de la profundidad de la escena capturada, permitiendo estimar con mayor precisión la estructura del entorno respecto a una cámara monocular, sin embargo para poder estimar la profundidad es necesario procesar las imágenes estéreo, en donde la dificultad principal reside en la correspondencia de las imágenes estéreo y la reconstrucción 3D de la escena capturada (Trucco, Verri, 1998). En general los desarrollos que emplean sistemas de visión se basan en enfoques donde se realiza una extracción de características o primitivas geométricas del entorno capturado y se hacen coincidir éstas en la secuencia de video permitiendo así, realizar la navegación y una reconstrucción 3D del entorno.

Otros desarrollos como el presentado en (Alenyà, G., et al. 2004) han puesto a trabajar en conjunto diversos sistemas sensoriales explotando y complementando las capacidades de cada sensor, la complejidad de trabajar con más de un sistema sensorial es la asociación de los datos entre estos.

En la actualidad desde que se dispone de sensores RGB-D de bajo costo, como el sensor de profundidad de Microsoft Kinect, se han realizado diversos trabajos

Esquema de Navegación Reactiva con Sensores RGB-D

para diferentes aplicaciones en el campo de la robótica móvil terrestre.

El trabajo aquí descrito muestra el inicio de la exploración de las capacidades y ventajas que traen las cámaras RGB-D de bajo costo como sistemas de percepción en robots móviles, en este trabajo se embarca una cámara RGB-D en un robot móvil con el objetivo de realizar una navegación local y reactiva, principalmente para la evasión de obstáculos presentes en la trayectoria del robot. Adicionalmente se presenta una visualización 3D de la escena local. Este trabajo es un inicio de una investigación que busca implementar un algoritmo de localización y mapeo simultáneo SLAM con cámaras RGB-D, con el objetivo de acercar los robots móviles a aplicaciones de servicio como publicidad, transporte industrial, etc. El artículo está distribuido de la siguiente forma: en la Sección 2 realiza un breve estado del arte de trabajos relacionados con la automatización del transporte industrial y trabajos que hacen uso del sensor Kinect en su aplicación; en la Sección 3 se detalla el sistema desarrollado; la descripción de la aplicación se detalla en la sección 4; en la Sección 5 se realiza la descripción de las pruebas y los resultados obtenidos y finalmente en Sección 6 se presenta las conclusiones y trabajos futuros.

2. TRABAJOS RELACIONADOS

El transporte de materia prima en la industria realizada por robots ha sido trabajado entre otros autores, por (Mora, A. et al. 2004). Este trabajo se enmarca en un proyecto general de automatización de la gestión y transporte de cargas. En él se presenta una solución al problema de navegación de un robot en un entorno industrial con la capacidad de realizar la localización y mapeo del entorno, mediante el uso de técnicas multifrecuenciales para la fusión sensorial, el algoritmo de navegación es determinado por campos de potencial y es solucionado el problema de mínimos locales con implementación de fuerzas ficticias. El trabajo presentado por (Alenyà, G., et al. 2004) describe una experiencia práctica para el transporte automatizado. En éste se dota de dos sistemas sensoriales a una plataforma industrial: el primer sistema es un láser con el que consiguen realizar la estimación de posición mediante un sistema dinámico de triangulación basado en un algoritmo geométrico fundamentado en intersecciones circulares y el segundo sistema es un cámara monocular con la cual logran estimar la posición mediante una técnica que relaciona la deformación de los contornos de los objetos presentes en una secuencia de video a un movimiento en el espacio 3D. Finalmente

la investigación concluye calificando que la técnica basada en visión es una alternativa prometedora para llevar a cabo tareas de transporte, mientras que el posicionamiento basado en laser podría ser empleada para operaciones donde se necesite mayor precisión como la carga y descarga de material.

Respecto a trabajos que emplean sistemas sensoriales basados en cámaras RGB-D de bajo costo, como el sensor de profundidad de Microsoft Kinect. Se puede citar el trabajo desarrollado en (Wu, Bainbridge-smith, 2012) en el cual se realiza una breve recopilación de trabajos donde se ha utilizado el Kinect, resaltando campos como UAV (Vehículos Aéreos No Tripulados), robótica móvil terrestre y la medicina, adicionalmente en su trabajo realizan una comparación entre el sensor y otros sensores de bajo costo, presentado las ventajas y desventajas de este sensor, igualmente resalta las principales razones para el uso de este sensor en investigaciones de robótica.

En el trabajo reportado en (Romero, Díaz, 2011) se desarrolla una plataforma robótica que logra navegar en un ambiente controlado, el algoritmo implementado se basa en el reconocimiento de obstáculos de un color definido, se realiza una segmentación por color para detectar el objeto, seguidamente se analiza la imagen de profundidad para obtener el valor de distancia promedio de la región segmentada de la imagen de color. En (Ruiz-Sarmiento, J., et al., 2011) se presenta la integración del sensor Kinect en un robot móvil para mejorar la realización de las tareas de detección de obstáculos y navegación reactiva, en este trabajo los autores hacen un estudio de la capacidad del sensor para la detección de obstáculos confirmando el buen desempeño de éste, para la navegación del robot toman la nube de puntos tridimensional y la simplifican a una nube de puntos bidimensional que es conformada por los obstáculos detectados en la escena, esta información es llevada al algoritmo de navegación reactivo.

En Huang y Bachrach (2011) los autores desarrollan un sistema de localización y reconstrucción 3D con aplicación al control de vuelo y navegación de un robot aéreo autónomo (UAV). En (Henry, P., et al. 2010) se desarrolla un sistema SLAM capaz de generar mapas 3D de la escena capturada, los autores presentan un algoritmo que combina la características visuales con la información de profundidad, la metodología seguida en este trabajo, inicialmente se extrae características tipo SIFT de las imágenes de color, seguidamente se busca una correspondencia de características entre dos *frames* consecutivos, las características correspondientes se

evalúan en el mapa de profundidad para obtener un conjunto de puntos 3D correspondientes entre las dos nubes de puntos, base a estas correspondencias se realiza la estimación la transformación relativa entre las nubes de puntos utilizando un algoritmo basado en RANSAC, posteriormente realizan un refinamiento de la posición inicial usando una variante del algoritmo ICP. Siguiendo con el mismo enfoque, se encuentra el trabajo de (Engelhard, N., et al., 2011) que a diferencia del trabajo anterior emplea características tipo SURF. Estos trabajos reportados realizan aportes significativos en el tema de la localización y mapeo simultáneo SLAM.

3. DESCRIPCIÓN DEL SISTEMA

El sistema implementado está conformado por tres componentes hardware: La cámara RGB-D (sensor Kinect), la unidad de procesamiento y la plataforma robótica Pioneer 3DX, en la figura 1 se presenta una ilustración que describe la integración de estos componentes.

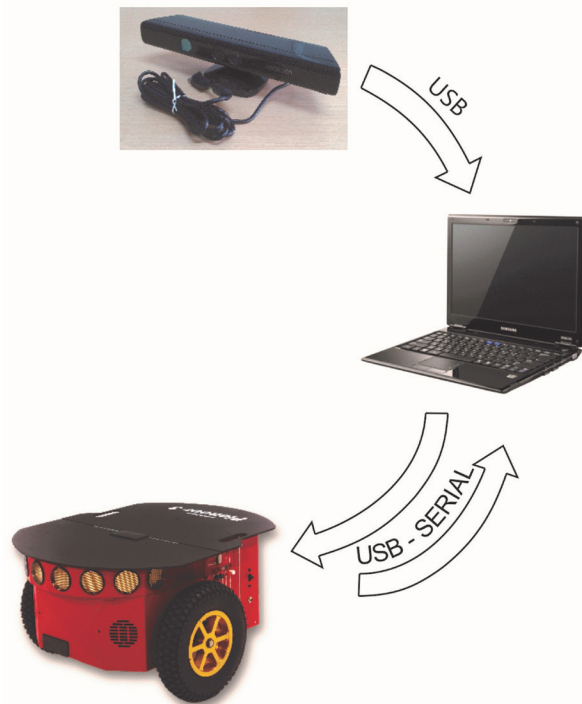


Figura 1. Ilustración de la integración de los componentes hardware del sistema

Como se aprecia en la Figura 1, el sensor Kinect se encuentra comunicado con la unidad de procesamiento o PC por una conexión USB, para adquirir la información

Andrés Felipe Suárez-Sánchez, Humberto Loaiza-Correa

suministrada por el sensor. El PC y el robot están comunicados por medio de una conexión USB-SERIAL por el cual se ordenan los comandos de velocidad y rotación al robot determinados por el algoritmo reactivo de navegación.

3.1 Sensor Kinect

El sensor de captura principal es el sensor Kinect de Microsoft, este es conformado por una cámara RGB que captura el entorno en una imagen a color con una frecuencia de 30 Hz en formato VGA con una resolución de 640 X 480 píxeles, un sensor de profundidad que cuenta con un emisor laser infrarrojo que proyecta una luz estructurada sobre la escena y una cámara infrarroja que captura el retorno de la luz proyectada con el cual se calcula la diferencia del tiempo de vuelo del haz de luz proyectada con una electrónica incorporada internamente, está codifica la información de profundidad en una imagen de rango, igualmente presenta una frecuencia de 30 Hz y una resolución de 640 X 480 píxeles (Smisek, J., et al. 2011).

Este dispositivo presenta un alcance de medición de 0.7 a 4 m, sin embargo experimentalmente se ha trabajado en un rango de 0.4 a 7 m, según reporta Huang y Bachrach (2011). En (Smisek, J., et al. 2011) se presenta un estudio más detallado de la resolución de las imágenes del sensor Kinect mediante un análisis experimental, en este trabajo se determina una fórmula que permite calcular la resolución de la profundidad, adicionalmente en el trabajo se realiza una comparación cuantitativa del sensor con una cámara 3D-TOF (Time Of Fly), proponen un modelo geométrico y un procedimiento para la calibración del sensor.

En la ecuación 1 se presenta el modelo para la resolución de la profundidad del sensor Kinect, descrito en (Smisek, J., et al. 2011).

$$q(z) = 2.73z^2 + 0.74z - 0.58 \quad (1)$$

donde:

$q(z)$: Resolución de la profundidad [mm].

z : Valor de distancia de prueba [m].

Considerando un rango de operación de 0,6 m a 3 m, la resolución asociada al sensor es entre 0,8468 mm y 26.21 mm. Los anteriores valores indican que el sensor presenta un error de medida aceptable para aplicaciones de detección de obstáculos en robótica.

3.2 Unidad de Procesamiento

La unidad de procesamiento es un computador portátil con un procesador Intel core i3 a 2.4 Ghz y memoria RAM de 3 GB con Sistema Operativo Linux (Distribución Ubuntu 10.04 de 32 Bits). La aplicación desarrollada emplea el driver OpenNI, el cual permite la comunicación del sensor Kinect y el computador. Se hace uso de tres librerías especializadas OpenCV (OpenCV), Point Cloud Library (PCL) y Player/Stage (Player/Stage), la primera de ellas se utiliza para la captura y procesamiento de imágenes, la segunda como motor para la representación y visualización de la escena 3D y la última para la programación del algoritmo de navegación reactivo.

3.3 Plataforma Robótica

La plataforma robótica es el robot móvil Pioneer 3DX, el cual tiene una configuración del tipo diferencial, cuenta con arco frontal de ocho sensores de ultrasonido que cubren 180° del robot.

4. DESCRIPCIÓN DE LA IMPLEMENTACIÓN

Para integrar los tres componentes descritos anteriormente, se implementó una aplicación software conformada por dos módulos, *percepción* y *navegación*. El primer módulo es el encargado de procesar la imagen de rango entregada por el sensor RGB-D. En éste se extrae la información de profundidad de la escena y se entrega una estructura de datos con información de distancia y ángulo para el algoritmo de navegación. En el segundo bloque se implementa el algoritmo de navegación local, el cual toma la información entregada por el bloque de percepción y determina los comandos de velocidad y rotación que deberá tener el robot para evadir los obstáculos. A continuación se describirán con mayor detalle los dos bloques mencionados.

4.1 Percepción Artificial

El proceso de percepción artificial se ilustra en la Figura 2.

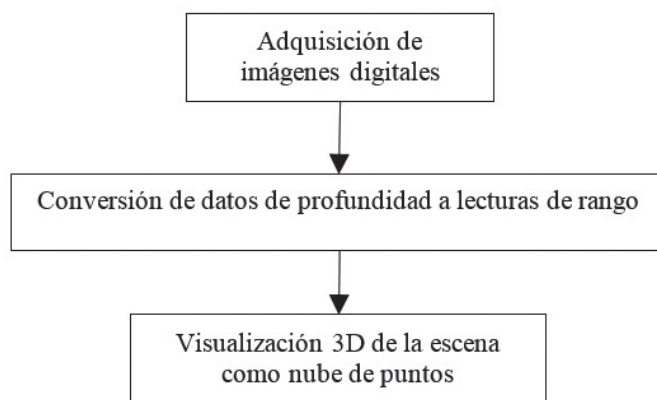
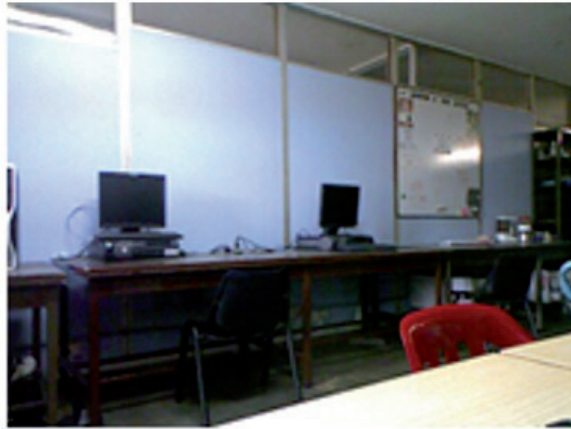


Figura 2. Diagrama de flujo para la percepción artificial

4.1.1 Adquisición de imágenes

En esta etapa se captura la imagen de color RGB, la imagen de profundidad DEPTH y la imagen con la información 3D POINTCLOUD. En la imagen DEPTH se codifican los valores de profundidad en representación de escala de grises, por otra parte la información 3D con los valores de la posición X, Y, Z

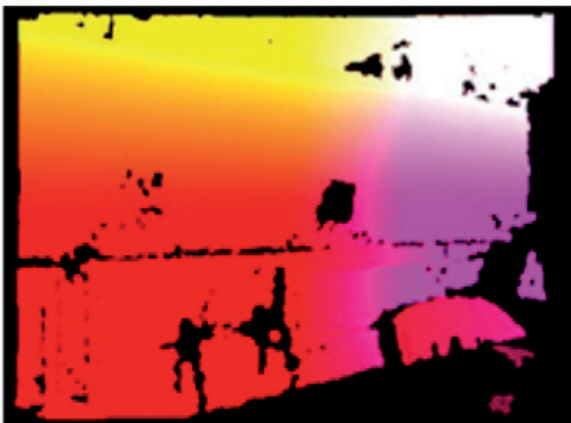
de cada punto es codificada en una imagen de color. Dado que el tamaño de las imágenes en el momento de la captura es de 640 x 480, se realizó una reducción de la información mediante un escalado a 320 X 240, con el fin de reducir carga computacional sin comprometer calidad de la imagen. En la figura 3 se presenta una muestra de tres imágenes que suministra el sensor Kinect para una escena capturada.



(A)



(B)



(C)

Figura 3. Resultado de la captura de imágenes. A) Imagen RGB. B) Mapa de Profundidad. C) Imagen de la nube de puntos

4.1.2 Conversión de datos de profundidad a lecturas de rango

En esta función se realiza un proceso de reducción de la información 3D de la cámara RGB-D a información 2D como si se tratase de un sensor de rango, es decir, lo que se busca es reducir la información 3D obtenida por el sensor y proyectarla sobre el plano 2D en el que se encuentra el robot como lecturas coplanares de distancia, de esta manera se podrán utilizar los algoritmos clásicos de navegación reactiva. La idea de este proceso es simular un sensor tipo rango que retornará una distancia medida y un ángulo asociado al valor de distancia capturado, para extraer la información necesaria para el algoritmo de navegación y no sobrecargar la unidad de procesamiento manejando una gran cantidad de datos 3D, con el propósito de no descuidar la tarea simultánea de evasión de obstáculos.

Esta técnica se puede apreciar en el trabajo pionero reportado por Murray y Little (1998), en donde obtienen un vector de lecturas de rango a partir de un sensor de estéreo visión. El vector de lecturas de rango tiene un tamaño igual al ancho de la imagen de disparidad, en donde el valor de profundidad de cada posición del vector se calcula tomando el valor máximo de disparidad de cada columna del mapa de disparidad. Otro trabajo que comparte la misma metodología es presentado en Castillo y Saéz (2003), en el cual se plantea una mejora al trabajo realizado por Murray y Little (1998) calculando la distancia a partir del valor de mayor frecuencia en cada una de las columnas de la imagen de disparidad; eliminando así, lo que ellos denominan un error estadístico, al tomar el máximo de un conjunto. En Hamzah (2010) se plantea la búsqueda de obstáculos en una región de interés horizontal del mapa de disparidad, la profundidad es calculada de la misma forma que en Murray y Little (1998) tomando el valor máximo de disparidad.

El proceso implementado es similar a un trabajo anterior presentado en Suárez y Loaiza (2012) y es mostrado como un diagrama de bloques en la figura 4. En la figura 5 se presenta una ilustración del concepto de reducción de datos 3D a 2D.

El proceso se inicia fijando una línea de escaneo horizontal sobre la cual se tomarán los datos de profundidad, esta línea tiene la propiedad de ser variable,

Esquema de Navegación Reactiva con Sensores RGB-D

esta fue configurada 30 pixeles por debajo de la línea media del mapa de profundidad. Igualmente se fijó la cantidad de lecturas de rango a obtener, equivalente a una *resolución angular*, la cual es calculada por la ecuación 4, esta sólo permite los valores [320, 160, 80, 40, 20, 10, 5] para el parámetro que fija la cantidad de sensores de rango. Se escogió una cantidad de 40 sensores de rango equivalente a una resolución angular $R_\theta = 1.425^\circ$.

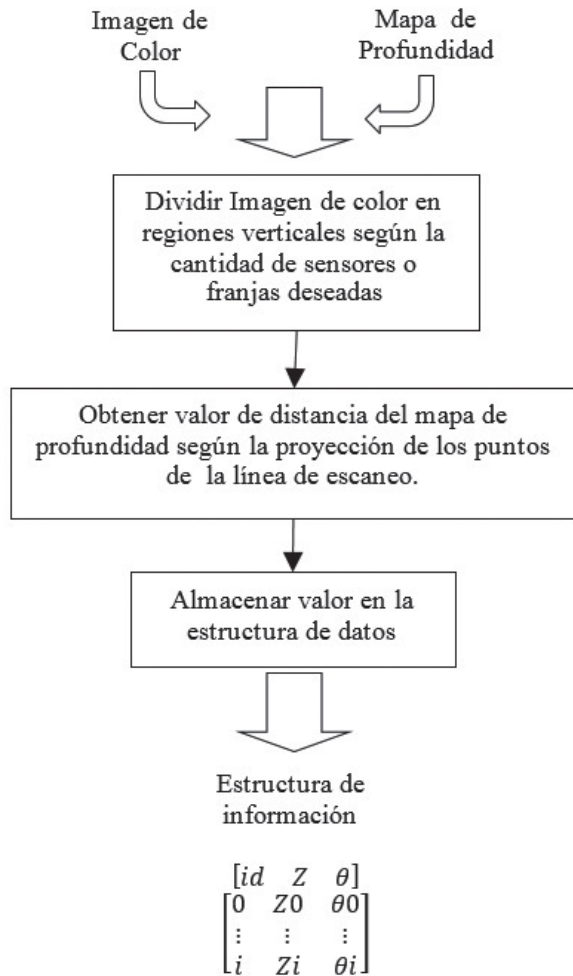


Figura 4. Diagrama de flujo del proceso de conversión de datos 3D a 2D

$$R_\theta = \frac{FOV_{xkinect}}{\#Sensoresrango} = \frac{57^\circ}{\#Sensoresrango} \quad (2)$$

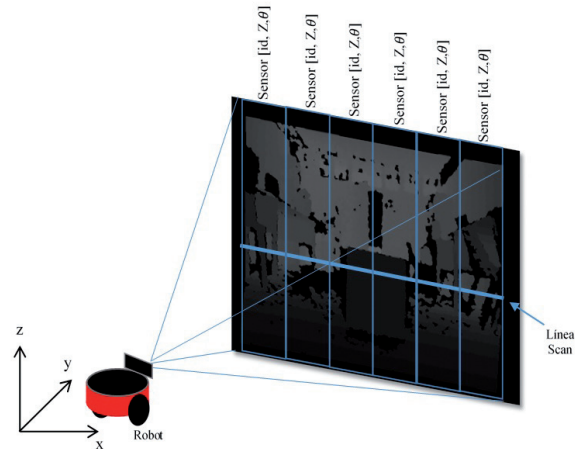


Figura 5. Ilustración del concepto de reducción de datos 3D a 2D

Inicializada la línea de escaneo horizontal y la cantidad de lecturas de rango que sean obtener (). Se toma la imagen de color y se fracciona en regiones verticales de interés, la cantidad de estas regiones es la misma cantidad de lecturas fijada en el paso anterior. Se Divide la imagen en las regiones de interés, se procede a fijar un punto medio a lo ancho de cada región. En la Figura 6, se muestra el resultado de este proceso inicial.



Figura 6. Puntos de escaneo para la adquisición de los valores de profundidad

Para la adquisición de los valores de profundidad de los puntos de interés proyectados en la imagen de color,

lo que se realiza es un proyección de los puntos sobre el mapa de profundidad para así extraer el valor de distancia para cada uno, de esta manera se colectará la información de distancia de la escena y se almacenará en una estructura de datos denominada $DATA_{\#sensores_rango \times 3}$, esta es una matriz que posee un tamaño igual a $\#Sensores_{rango} \times 3$, donde la primera columna lista los identificadores (id) de cada sensor o región de interés, en la segunda columna se agrupa los datos de distancia (Z) adquiridos y la última columna reúne la posición angular (θ) del centro de cada región de la imagen, donde la posición angular de cada dato es calculada por medio de la ecuación 5.

$$\theta(id_{sensor}) = 28.5^\circ - \left((2 * \frac{57^\circ}{\#Sens}) + (\frac{57^\circ}{\#Sens}) * id_{sensor} \right) \quad (3)$$

Los datos de distancia son representados en una imagen que representa el plano del robot. Cada lectura de distancia se presenta como un punto azul entre seis secciones horizontales que se utilizan como valores guía de distancia. Ver figura 7.

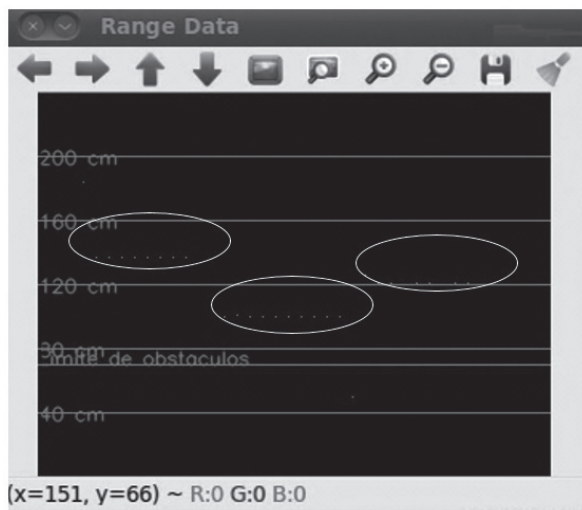


Figura 7. Representación de las lecturas de profundidad para los objetos de la figura 6

4.1.3 Visualización 3D de la escena como nube de puntos

En esta etapa se concibe la idea de representar en 3D el entorno que captura el robot, el propósito es generar una visualización que presente con detalles la escena¹. En la figura 8, se presenta el diagrama de bloques del proceso de reconstrucción 3D local con sus respectivas entradas y salida.

¹ En este trabajo no se resolverá el problema de mapeo 3D.

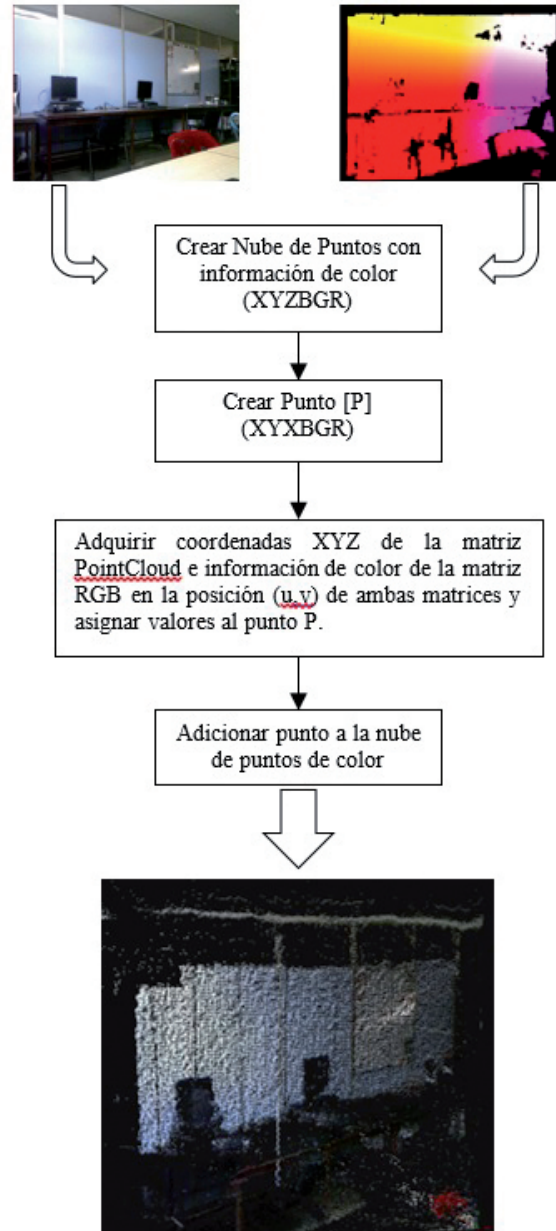


Figura 8. Representación del proceso de visualización 3D como diagrama de entradas y salidas

Este proceso tiene como entrada la matriz de la nube de puntos (POINTCLOUD image) que contiene la información espacial y la matriz de color (RGB image) que contiene la información de detalle y entrega como salida una nube de puntos a color que recrea la escena capturada. Este se inicia creando una nube de puntos con la característica que contenga información de color, seguidamente se procede a recorrer simultáneamente las dos matrices de entrada para extraer la información espacial o coordenadas XYZ y la información de color

Esquema de Navegación Reactiva con Sensores RGB-D

RGB para luego asignárselo a cada punto de la nube de puntos creada, terminado el proceso se procede a visualizarla en un motor de gráficos 3D.

Los datos de la matriz nube de puntos son las coordenadas X, Y, Z del mundo real de cada uno de los elementos presentes en la imagen de color, estas coordenadas son calculadas por medio de la transformación de perspectiva inversa. En las ecuaciones 4, 5 y 6² se presentan las expresiones que permiten obtener las coordenadas X, Y, Z a partir de la información de profundidad, para el sensor empleado.

$$X = \left(\frac{x_{proc}}{x_{res}} - \frac{1}{2} \right) * depth * fx \quad (4)$$

$$Y = \left(\frac{1}{2} - \frac{y_{proc}}{y_{res}} \right) * depth * fy \quad (5)$$

$$Z = depth \quad (6)$$

4.2 Navegación de Robot

Para que el robot móvil sea autónomo y pueda desplazarse de un punto A a un punto B, debe desarrollarse un software de control que se encargue de procesar los datos de los sensores para planificar y ejecutar las tareas encomendadas. En nuestra implementación el propósito del navegador es asegurar la integridad del robot para movimientos parciales o *navegación local*. Ésta se logra mediante un navegador reactivo, el cual solo tiene en cuenta los obstáculos existentes en cada instante de tiempo. En esta sección se resume el funcionamiento del navegador reactivo implementado.

Se implementó una estrategia de navegación de tipo local, basada en campos de potencial. Este algoritmo representa al agente móvil como una partícula bajo la influencia de un campo potencial artificial cuyas variaciones locales reflejan la estructura del espacio libre de obstáculos. La fuerza artificial que determina el movimiento del robot es el resultado de una fuerza atractiva (vector de meta) y una fuerza repulsiva (vector de obstáculos) (Dudek, Jenkin, 2010).

En la ecuación 7 se presenta el modelo que representa la fuerza potencial que experimenta el robot según la configuración del espacio.

$$U(q) = U_{meta}(q) + \sum U_{obstaculos}(q) \quad (7)$$

Donde,

$U(q)$: Potencias que experimenta el robot en el punto q

$U_{meta}(q)$: Potencial generado por la meta

$U_{obstaculos}(q)$: Potencial generado por un obstáculo

Se seleccionó como potencial meta $U_{meta}(q)$, un vector unitario que apunta siempre hacia adelante y como potencial repulsivo $U_{obstaculos}(q)$ el resultado de la sumatoria de los vectores donde la profundidad de los objetos fuera menor a la distancia de la zona de detección de obstáculos, esta zona tiene un rango de 40 a 70 cm desde el frente del agente móvil, el potencial $U(q)$, es un vector resultante de la operación de la ecuación 7, la magnitud y el *ángulo de este vector determina* los comandos de velocidad lineal y el ángulo que debe tomar el robot para evadir los obstáculos.

5. RESULTADOS

Se estableció un protocolo de pruebas conformado por dos fases, la primera consistió en determinar el desempeño del sensor para la medición de distancias de objetos presente en una escena y simultáneamente se verificaba el funcionamiento del enfoque de reducción de información de 3D a 2D. La siguiente fase consistió en poner a navegar un robot móvil con la información adquirida con el fin de determinar, si, el enfoque de reducción de datos suministra la información adecuada para navegar y simultáneamente se evaluaba el algoritmo implementado.

La primera fase de pruebas se realizó en el interior de un laboratorio, el cual presenta una iluminación artificial controlada y una entrada de luz natural. Para esta prueba se ubicó el sensor de tal forma que la luz natural no incidiera sobre éste. La prueba consistió en desplazar uno o varios obstáculos frente al sensor en un rango de 0,6 a 3 m.

En la figura 9 se presenta una recopilación de imágenes de la prueba de medición de obstáculos, se muestra tres obstáculos localizados en diferentes distancias y se presenta la imagen como lecturas de rango respectivamente. Se determinó que el sensor empieza a registrar mediciones a partir de 0,5 m y que el error de las mediciones realizadas por el sensor está por debajo de 2,5 cm, presentando un error promedio de medición de 0,636% del valor de distancia real. En la ecuación 8, se presenta el modelo obtenido para la profundidad

² Las ecuaciones presentadas son obtenidas de la librería OpenNI

Z considerando el error promedio registrado en las pruebas.

$$Z = (1 \pm 0.00636)depth_{sensor} \quad (8)$$

Respecto a la reconstrucción de la escena local, se puede decir que como se aprecia en las ecuaciones 4 y 5, existe una dependencia del valor de profundidad para la coordenadas X e Y. Esta dependencia implica que el error que se presente en la coordenada Z se propagaría para las coordenadas X e Y.

La segunda fase de las pruebas se realizó dentro de las instalaciones de uno de los edificios de la Universidad del Valle³, con iluminación artificial controlada. Para esta prueba se realizaron 3 escenarios con obstáculos. El objetivo de la prueba consistía en que el robot con el sistema de visión pudiese ir desde el punto de partida a un punto meta evadiendo los obstáculos presentes en la trayectoria.

En la figura 10, se muestra los diferentes escenarios donde se puso a navegar el robot. Para cada escenario se realizaron cuatro repeticiones para un total de 12 pruebas⁴, considerando como prueba exitosa cualquier recorrido donde el robot no colisionará con ningún obstáculo y como prueba fallada cuando colisionará con un obstáculo. En esta prueba se obtuvo 11 recorridos exitosos de 12 para un 91.66% de éxito. Igualmente se pudo determinar que el control del algoritmo de navegación para las velocidades lineales y rotacionales fue apropiado, evitando que se presentaran movimientos bruscos e inapropiados en el robot.

El funcionamiento en conjunto presenta un tiempo de respuesta promedio de 125ms, para capturar, procesar, generar el modelo 3D y planear el próximo movimiento. El mayor consumo de tiempo de ejecución se presenta en el proceso de generación del modelo del entorno con un tiempo promedio 85 ms. El tiempo de respuesta que presenta el sistema limita a que el robot pueda navegar a una velocidad menor a 1 m/s.



Figura 9. Imágenes de la prueba de medición y su respectiva representación a lecturas de rango

³ No se realizaron pruebas en entornos industriales.

⁴ Las pruebas realizadas se registraron en video, estos pueden verse en canal de youtube, <http://www.youtube.com/user/anfesusaable>

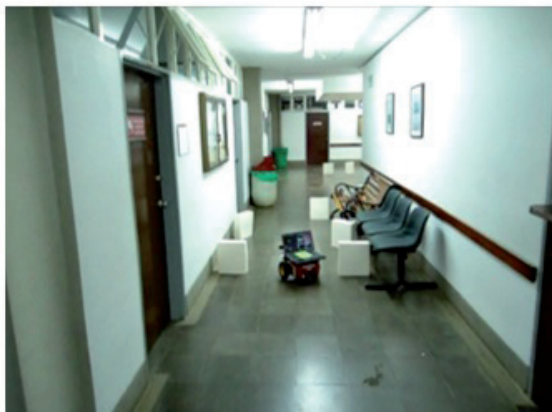


Figura 10. Escenarios de prueba para el algoritmo de navegación

6. CONCLUSIONES

En este trabajo se presentó la integración de un robot móvil con un sistema de visión artificial basado en una cámara RGB-D, que puede servir de base para lograr la automatización en el transporte de carga en entornos industriales.

El enfoque de reducción de datos de 3D a 2D presentado resultó ser una buena opción para brindar la información necesaria al planificador de trayectoria local (Algoritmo de Navegación). Por medio de este enfoque se pudo localizar y representar obstáculos presentes en la trayectoria del robot de una forma sencilla, rápida y con buena precisión. Este enfoque ofrece reducir tiempos para el cálculo del siguiente movimiento sin saturar la unidad de procesamiento, considerando el sistema de visión sin el proceso de reconstrucción de la escena local se obtuvo a una reducción en tiempo de 68%, con un tiempo de ejecución promedio de 45 ms.

El uso de la cámara RGB-D para el sistema de visión, fue buena opción para la aplicación desarrollada por su relación costo-beneficio. El sistema de visión presentó un buen desempeño frente a las condiciones del entorno como la forma geométrica de los objetos presentes en la escena, texturas y color de los objetos y frente a las irregularidades en la iluminación artificial del entorno, problemas que suelen presentarse en sistemas de visión conformados por cámaras mono o estéreo. Otra gran ventaja de usar la cámara RGB-D como sensor en robótica, es que ahorra tiempo de cómputo en el algoritmo de visión ya que entrega directamente un mapa de profundidad sin dedicar recursos de tiempo y computó, evitando procesos y problemas propios de la visión estero convencional como son la alineación y la correspondencia de dos imágenes RGB independientes.

En un futuro se plantea realizar un esquema de navegación reactiva que emplee toda la información densa del sensor y tenga la capacidad de detección de obstáculos en todo el rango de alturas del robot. Igualmente continuar con investigaciones que den solución al problema de SLAM, de manera que la información 3D de la cámara de rango también sea útil para la localización del robot en el entorno y construcción de mapas métricos o semánticos.

7. AGRADECIMIENTOS

Andrés F. Suárez agradece a COLCIENCIAS (Departamento Administrativo de Ciencia, Tecnología e Innovación) por la beca-pasantía p-2011-1833 otorgada del programa Jóvenes Investigadores e Innovadores, al grupo de investigación Percepción y Sistemas Inteligentes (PSI) de la Universidad del Valle por el apoyo brindado.

8. REFERENCIAS

- BUI, T.; WIDYOTRIATMO, A.; LE HOA NGUYEN; KEUM-SHIK HONG. “Sonar-based collision avoidance and probabilistic motion model for mobile robot navigation,” *Asian Control Conference*, 2009. ASCC 2009. 7th, vol., no., pp.1462-1467, 27-29 Aug. 2009.
- QIAN, K; SONG, A. “Autonomous navigation for mobile robot based on a sonar ring and its implementation”, *Instrumentation and Control Technology (ISICT), 2012 8th IEEE International Symposium on*, vol., no., pp.47-50, 11-13 July 2012.
- NÜCHTER, A; LINGEMANN, K; HERTZBERG, J; SURMANN, H. “6D SLAM—3D Mapping Outdoor Environments” *Journal of Field Robotics*, 2007.
- COLE, D.M.; NEWMAN, P.M. “Using laser range data for 3D SLAM in outdoor environments”, *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, vol., no., pp.1556-1563, 15-19 May 2006
- DAVISON, A.J.; REID, I.D.; MOLTON, N.D.; STASSE, O. “MonoSLAM: Real-Time Single Camera SLAM.” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol.29, no.6, pp.1052-1067, June 2007.
- SEO-YEON HWANG; JAE-BOK SONG. “Monocular Vision-Based SLAM in Indoor Environment Using Corner, Lamp, and Door Features From Upward-Looking Camera”, *Industrial Electronics, IEEE Transactions on*, vol.58, no.10, pp.4804-4812, Oct. 2011.
- LEMAIRE, T; BERGER, C; JUNG, I; LACROIX, S. “Vision-Based SLAM: Stereo and Monocular Approaches”. *International Journal of Computer Vision* 74(3), 343–364, 2007. DOI: 10.1007/s11263-007-0042-3
- VIEJO, D.; CARLOZA, M. “Construcción de Mapas 3D y Extracción de Primitivas Geométricas del Entorno”; *5th Workshop de Agentes Físicos*; 2004
- CAÑAS, J., M.; LEON, O.; CALVO, R. “Reconstrucción 3D visual atenta para la navegación de un robot móvil”. *IX Workshop en Agentes Físicos*, Vigo, pp 205-21, 2008.
- TRUCCO, E; VERRI, A. *Introductory Techniques for 3-D Computer Vision*. New Jersey. Prentice Hall, 1998.
- ALENYÀ, G.; ESCODA, J.; MARTÍNEZ, A. B., TORRAZ, C. “Using Laser and Vision to Locate a Robot in an Industrial Environment: A Practical Experience”, *Proceedings of the 2005 IEEE International Conference on Robotics and Automation Barcelona*, pp 3528-3533, 2005.
- MORA AGUILAR, MARTHA C; ARMESTO ÁNGEL, LEOPOLDO; TORNERO MONSERRAT, JOSEP. “Sistema de navegación de robots móviles en entornos industriales”, *XXV Jornada de Automática*, Ciudad Real. 2004.
- WU, H., BAINBRIDGE-SMITH, A. “Advantages of using a Kinect Camera in various applications”. [Documento en línea]: <[https://bitbucket.org/soko/quadtrack/src/d1f3b069ed9a/Reports/Research Papers/hhw26paper3rdPro.pdf](https://bitbucket.org/soko/quadtrack/src/d1f3b069ed9a/Reports/Research%20Papers/hhw26paper3rdPro.pdf)> [consulta: 15-3-2012].
- ROMERO MOLANO, C. A.; DÍAZ CELIS, C. A. “Navegación de robot móvil usando Kinect y OpenCV”, *3er Congreso Internacional de Ingeniería Mecatrónica*, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia, Vol 2, No 1, 2011.
- RUIZ-SARMIENTO, J.R.; GALINDO, C.; GONZALEZ-JIMENEZ, J., BLANCO, J.L. “Navegación Reactiva de un Robot Móvil usando Kinect”, [Documento en línea]: <<http://ingmec.ual.es/~jlblanco/papers/robot2011ruiz-sarmiento.pdf>> [consulta: 15-3-2012].
- HUANG, A; BACHRACH, A. “Visual odometry and mapping for autonomous flight using an RGB-D camera”, *International Symposium on Robotics Research (ISRR)*, 2011. [Documento en línea]: <<http://www.cs.washington.edu/research/projects/aiweb/media/papers/Huang-ISRR-2011.pdf>> [consulta: 15-3-2012].
- HENRY, P.; KRAININ, M.; HERBST, E.; REN, X.; FOX, D. “RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments”. *In Proc. of the Intl. Symp. on Experimental Robotics (ISER)*, Delhi, India, 2010. [Documento en línea]: <<http://www.cs.washington.edu/robotics/postscripts/3d-mapping-iser-10-final.pdf>> [consulta: 15-3-2012].
- ENGELHARD, N.; ENDRES, F.; HESS, J.; STURM, J.; BURGARD, W. “Real-time 3D visual SLAM with

a hand-held RGB-D camera”. 2011. [Documento en línea]: <https://ias.informatik.tu-muenchen.de/_media/events/rgbd2011/03_engelhard.pdf> [consulta: 15-3-2012].

SMISEK, J.; JANCOSSEK, M.; PAJDLA, T. “3D with Kinect,” *Computer Vision Workshops (ICCV Workshops)*, 2011 *IEEE International Conference on*, vol., no., pp.1154-1160, 6-13 Nov. 2011.

OpenCV: Open Source Computer Vision [Web en línea]. <<http://opencv.org>> [consulta: 15-3-2012].

PCL: Point Cloud Library [Web en línea]. <<http://pointclouds.org>> [consulta: 15-3-2012].

Playe/Stage: The Player Project. [Web en línea]. <<http://playerstage.sourceforge.net>> [consulta: 15-3-2012].

MURRAY, D.J., LITTLE, J. “Using Real-Time Stereo Vision for Mobile Robot Navigation” *Computer Vision*

and Pattern Recognition (CVPR’98). Santa Barbara CA. 1988

CASTILLO, M. A., SAÉZ, J. M. “Modelo de Sonar de Largo Alcance Basado en Tecnología Estéreo”. *4th Workshop de Agentes Físicos*. 2003.

HAMZAH, R.A.; RAHIM, R.A.; ROSLY, H.N. “Depth evaluation in selected region of disparity mapping for navigation of stereo vision mobile robot,” *Industrial Electronics & Applications (ISIEA)*, 2010 *IEEE Symposium on*, vol., no., pp.551-555, 3-5 Oct. 2010.

SUÁREZ, ANDRÉS; LOAIZA, HUMBERTO. “Navegación de un robot móvil por estereovisión”. *Revista entre ciencia e ingeniería, Colombia*, pp 9-23, 2012.

DUDEK, G.; JENKIN, M. “Computational Principles of Mobile Robotics, 2da Edition, Cambridge University Press, 2010.