Jornadas Nacionales de Robótica
Spanish Robotics Conference
14-15 Junio 2018 http://jnr2018.eii.uva.es

Uso de Descriptores basados en Apariencia Global para la Localización y Creación de Mapas

Yerai Berenguer*, Luis Payá, Adrián Peidró, Oscar Reinoso

Departamento de Ingeniería de Sistemas y Automática, Universidad Miguel Hernández de Elche, Avda. de la Universidad s/n, 03202, Elche, España.

Resumen

Este trabajo presenta las aportaciones más relevantes realizadas dentro de la tesis 'Uso de descriptores holísticos para la localización y creación de mapas: una aproximación al Graph-SLAM mediante apariencia visual', la cual será defendida el día 27 de abril de 2018 en la Universidad Miguel Hernández de Elche. Hoy en día, el campo de la robótica móvil está cada vez más presente en diferentes campos y aplicaciones como puede ser la conducción autónoma. Debido a esta creciente presencia, la comunidad científica trata de desarrollar nuevos algoritmos para proveer a los robots móviles de mayor autonomía para tomar mejores decisiones en sus tareas. Con este objetivo, el robot debe desarrollar diferentes habilidades esenciales a la hora de poder navegar por entornos desconocidos. Una de estas habilidades es la de crear un mapa del entorno a partir de información recogida con los sensores equipados en él. Otra de estas habilidades consiste en ser capaz de localizarse dentro del mapa creado anteriormente. En el caso de que ambas tareas (creación de mapas y localización) se ejecuten al mismo tiempo se conoce como SLAM (*Simultaneous Localization And Mapping*). En las investigaciones de la tesis de la que trata este trabajo se presentan diferentes algoritmos de creación de mapas y localización utilizando información obtenida a través de sensores visuales omnidireccionales. Asimismo, para la descripción de las imágenes se utilizan técnicas basadas en la apariencia global de las imágenes. Todos los algoritmos se testan con diversos conjuntos de imágenes, tanto imágenes generadas virtualmente como imágenes capturadas bajo condiciones de trabajo reales y teniendo en cuenta fenómenos habituales como cambios de iluminación y oclusiones. *Copyright © 2018 CEA.*

Palabras Clave:
Visión omnidireccional, descriptores holísticos, transformada de Radon, Localización, Creación de mapas, SLAM
Datos de la tesis: Uso de descriptores holísticos para la localización y creación de mapas: una aproximación al Graph-SLAM mediante apariencia visual.
Doctorando: Yerai Berenguer.
Directores: Óscar Reinoso y Luis Payá.
Fecha Inicio/Fecha estimada terminación: 01/05/2015 - 27/04/2018.
Programa de doctorado: Doctorado en tecnologías industriales y telecomunicación.
Universidad: Universidad Miguel Hernández de Elche

1. Introducción

Hoy en día, la presencia de robots en nuestra sociedad ha aumentado considerablemente. Inicialmente, se usaron para llevar a cabo algunas tareas que resultaban muy exigentes o peligrosas para las personas. Sin embargo, en la actualidad, se utilizan en otras innumerables tareas con diferentes propósitos, gracias a la evolución de los equipos y técnicas de percepción y computación empleadas. Actualmente, es posible diseñar robots más autónomos que no requieren la intervención humana para llevar a cabo sus tareas. Muchos de estos robots también son móviles y para cumplir con su tarea deben poder planificar una trayectoria para llegar a los puntos de destino y navegar hacia ellos mientras evitan los obstáculos presentes en el entorno. Para llevar a cabo la tarea de navegación de una manera eficiente, es necesario llevar a cabo dos tareas fundamentales. Por un lado, el robot debe crear una representación interna del entorno inicialmente desconocido (mapa) y, por otro, debe estimar su posición dentro del mapa. Por ello, el robot necesita uno o más sensores para extraer información del entorno. Hay muchos tipos de sensores que les brindan información útil, como sensores táctiles, encoders, sensores láser o de visión. Esta información puede usarse tanto para construir el modelo del entorno como para estimar la posición del robot.

Los sensores de visión se han convertido en una de las opciones más extendidas en robótica móvil gracias a la gran cantidad de información que proporcionan al robot. Garcia-Fidalgo y Ortiz (2015) presentan una recopilación de métodos de creación de mapas y localización utilizando sistemas de visión. Estos sistemas pueden tener diversas configuraciones, como cámaras individuales, cámaras estéreo, sistemas con arrays de cámaras, sistemas catadióptricos, etc. Los sistemas de visión catadióptricos

*Autor en correspondencia.
Correos electrónicos: yberenguer@umh.es (Yerai Berenguer), lpaya@umh.es (Luis Payá), apeidro@umh.es (Adrián Peidró), o.reinoso@umh.es (Oscar Reinoso)

consisten en una sola cámara que apunta a un espejo convexo. Esta configuración permite tomar imágenes con un campo de visión de 360 grados alrededor del eje del espejo. La riqueza de la información que capturan es la razón por la cual se ha elegido este tipo de sistema como dispositivo que proporciona la información que se tendrá en cuenta para reconocer el entorno por el que se desplaza el robot móvil. Hay muchos trabajos previos que usan sistemas de visión catadióptricos en tareas de navegación.

Sin embargo, las imágenes contienen mucha información redundante que puede cambiar en diversas circunstancias, como por ejemplo, el ruido y las oclusiones. Por esta razón, es necesario extraer información relevante de cada escena para crear el mapa. Esta información debe permitir estimar la posición del robot con robustez. Hay dos enfoques diferentes para llevar esto a cabo. Por un lado, la imagen se puede describir a través de la extracción y descripción de los puntos característicos locales de las escenas. Por ejemplo, Lowe (1999) llevan a cabo tareas de localización y creación de mapas utilizando SIFT (Scale-Invariant Feature Transform) y Bay et al. (2006) presentan otro detector y descriptor de puntos característicos, denominado SURF (Speeded-Up Robust Features). Por otro lado, algunas obras usan la apariencia global de las escenas para crear solo un descriptor por imagen. Payá et al. (2014) analizan y comparan varios métodos para obtener descriptores globales de escenas panorámicas para comprobar su validez en la creación y localización de mapas. La mayoría de estos descriptores holísticos se pueden usar en tiempo real, porque el tiempo de cálculo para generarlos y manejarlos es bajo, y generalmente conducen a algoritmos de localización y creación de mapas más directos que los algoritmos basados en la extracción de características. En los últimos años, algunos trabajos se han centrado en el uso de imágenes omnidireccionales como la única fuente de información para resolver las tareas de creación de mapas y localización utilizando descriptores basados en la apariencia global de las escenas, como Payá et al. (2016); Berenguer et al. (2015); Fernández et al. (2014). Payá et al. (2016) presentan una evaluación comparativa de diferentes técnicas de apariencia global en tareas de creación de mapas utilizando imágenes omnidireccionales. Berenguer et al. (2015) presentan diferentes métodos en tareas de localización y creación de mapas utilizando la apariencia global de imágenes omnidireccionales. Fernández et al. (2014) presentan un enfoque de apariencia global para llevar a cabo tareas de localización y creación de mapas simultáneamente utilizando mapas métrico-topológicos híbridos.

El mapa del entorno se puede crear utilizando dos enfoques diferentes: métrico o topológico. Por un lado, los mapas métricos representan el entorno que define las ubicaciones de algunas de las características relevantes con respecto a un sistema de coordenadas. Esta configuración permite estimar la posición del robot con precisión geométrica. Munguia et al. (2016) describen un sistema de localización y creación de mapas utilizando la información métrica obtenida por diferentes sensores; un sensor de orientación, un sensor de posición (GPS) y una cámara monocular. Hay más trabajos que utilizan este enfoque métrico en las tareas de localización, como Forster et al. (2013); Weiss et al. (2012). Forster et al. (2013) emplean localización y creación de mapas utilizando múltiples vehículos micro aéreos (MAV) en entornos desconocidos. Weiss et al. (2012) presenta un sistema para permitir vuelos autónomos de un micro vehículo aéreo (MAV) que solo dispone de mediciones lentas, ruidosas, retrasadas y posiblemente a escala arbitraria. Más recientemente, algunos investigadores han combinado los conceptos métricos y topológicos para generar mapas híbridos, donde la información se organiza en varias capas con diferentes niveles de detalle. Kostavelis et al. (2016) crean mapas compuestos por diversas capas utilizando el concepto de mapas híbridos para llevar a cabo tareas de navegación jerárquicas.

El resto de este trabajo se estructura de la siguiente manera. La sección 2 introduce el concepto de visión omnidireccional y presenta un programa de creación de imágenes virtuales. La sección 3 presenta un algoritmo de localización y creación de mapas locales. La sección 4 describe los descriptores holísticos empleados. La sección 5 muestra un algoritmo de estimación de altura relativa utilizando imágenes omnidireccionales. La sección 6 presenta un algoritmo de creación de odometría visual. Finalmente, la sección 7 esboza las conclusiones.

2. Creación de Imágenes Omnidireccionales Virtuales

A la hora de comprobar el funcionamiento de cualquier nuevo algoritmo se necesita una base de datos con la cual validar los resultados. En muchas ocasiones, los resultados de los experimentos con estas bases de datos capturadas muestran resultados poco satisfactorios. Para evitar este problema y poder realizar, de forma rápida, una comprobación inicial del desempeño de los algoritmos realizando además una sintonización gruesa de los parámetros de configuración del sistema, se ha creado una plataforma de adquisición de imágenes omnidireccionales simulando un entorno virtual. Esto proporcionará una base de datos virtual inicial con la cual comprobar el funcionamiento de los nuevos métodos y así elegir la correcta configuración a la hora de capturar la base de datos de imágenes reales acelerando el proceso de desarrollo de nuevos algoritmos de creación de mapas y localización.

Además, para probar los algoritmos de localización en un mapa previamente construido, es posible generar imágenes de prueba desde cualquier posición en el mapa simulando cualquier rotación u orientación de la plataforma robótica en el entorno (6 grados de libertad).

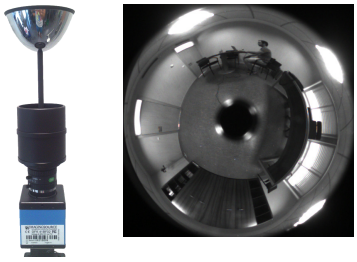
De esta manera, esperamos que el uso de esta plataforma ahorre tiempo y dinero durante el desarrollo de nuevos algoritmos de creación de mapas y localización de robots. Gracias a ello, los experimentos iniciales pueden ser llevados a cabo rápidamente, en una gran variedad de entornos y con precisión.

2.1. *Visión Omnidireccional*

Los sistemas de visión omnidireccional se caracterizan por obtener información del entorno para todas las orientaciones posibles, es decir, si este sistema estuviese montado en un robot móvil, éste sería capaz de obtener información de la habitación o entorno en el que estuviese operando para cualquiera de las posibles orientaciones alrededor suyo.

Para conseguir este propósito, estos sistemas deben de captar todos los rayos de luz procedentes de cualquier orientación alrededor de ellos. Por lo tanto, para los diferentes experimentos realizados a lo largo de la tesis, se ha utilizado un sistema de visión formado por una cámara CCD (Charge-Coupled Device), la cual capta la información visual y la transforma en información digital, y un espejo hiperbólico, el cual capta los rayos de

luz procedentes de todas las orientaciones y los proyecta hacia la lente de la cámara CCD. Dicha cámara y el espejo están anclados a través de mecanismos fijos (Figura 1(a)). En la Figura 1(b) se muestra un ejemplo de una imagen omnidireccional capturada con un sistema de visión catadióptrico.



(a) (b)

Figura 1: (a) Imagen omnidireccional real. (b) Sistema de adquisición de imágenes omnidireccionales. Está compuesto por una cámara CCD apuntando hacia un espejo hiperbólico.

2.2. *Plataforma de Creación de Imágenes*

Con el fin de poder crear imágenes omnidireccionales a partir de un entorno virtual, se ha modelado el sistema de visión catadióptrico mediante un algoritmo basado en la trayectoria que siguen los rayos de luz desde los objetos hasta que llegan al foco de la cámara, pasando por la reflexión del espejo hiperbólico. Dicho algoritmo ha sido desarrollado mediante el entorno de programación Matlab. En la Figura 2 se puede apreciar el esquema del sistema de visión omnidireccional y la trayectoria de los rayos de luz simulados.

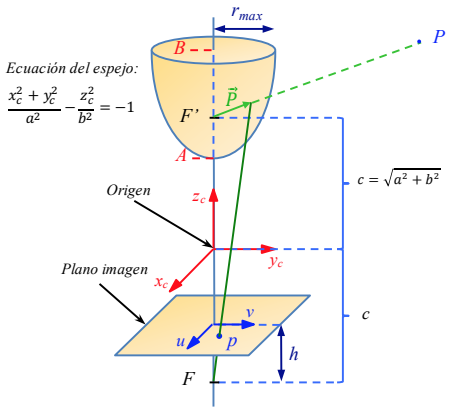


Figura 2: Modelo geométrico de proyección del sistema de visión omnidireccional con un espejo convexo.

La creación de cada entorno es muy fácil de realizar a través de líneas de comando definiendo las características de cada uno de los nuevos objetos definiendo su forma, posición, orientación y tamaño. Los entornos virtuales pueden ser definidos como entornos de interior o de exterior, dependiendo de la configuración de las paredes y de los objetos dentro de ellos.

3. Descriptores Holísticos

Los métodos basados en la apariencia global de escenas constituyen una alternativa robusta en comparación con los métodos basados en la extracción de puntos característicos. La clave es que los descriptores de apariencia global representan el entorno a través de características de alto nivel que se pueden interpretar y manejar fácilmente, y con un coste computacional razonablemente bajo.

Cualquier nuevo método de descripción de aspecto global debería satisfacer algunas propiedades: (1) debería ejercer un efecto de compresión en la información de la imagen; (2) debería existir una correspondencia entre la distancia entre dos descriptores y la distancia métrica entre las dos posiciones reales donde se capturaron las imágenes; (3) el coste computacional para calcularlos y compararlos debe ser bajo, de modo que puedan ser usados en tareas en tiempo real; (4) debería proporcionar robustez contra el ruido, cambios en las condiciones de iluminación, oclusiones y cambios en la posición de algunos objetos en el entorno; y, finalmente, (5) debería contener información de la orientación que tenía el robot cuando capturó la imagen.

En esta sección se muestran los descriptores de apariencia global utilizados como herramientas a la hora de desarrollar los algoritmos de las líneas de investigación.

3.1. *Firma de Fourier*

Si se parte de una imagen $im_j(x, y) \in \mathbb{R}^{N_x \times N_y}$, es posible extraer la información más relevante de la imagen a través de la transformada discreta de Fourier (DFT) de cada una de sus filas. Con este paso se obtiene una nueva matriz $d_j(u, v) \in (C)^{N_x \times N_y}$, del mismo tamaño que la imagen original. Dentro de esta matriz, la mayor parte de la información de la imagen se encuentra concentrada en los componentes de baja frecuencia de cada fila. De este modo, se puede trabajar únicamente con las k_1 primeras columnas de la matriz resultante. A esta nueva matriz de tamaño N_x filas y k_1 columnas se le llama *Firma de Fourier*.

3.2. *Transformada de Radon*

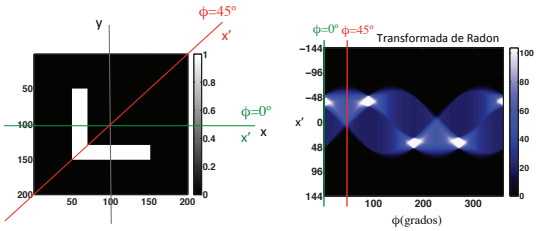
La transformada de Radon fue descrita inicialmente en Radon (1917) y consiste en la integral de una función 2D sobre líneas rectas (proyecciones integrales de línea). Esta transformada es invertible y la transformada inversa de Radon reconstruye una imagen a partir de sus proyecciones integrales de línea. Por este motivo, se utilizó inicialmente en imágenes médicas (como el TAC Yamazawa et al. (1995) y la resonancia magnética (IRM) Liang y Lauterbur (2000)).

La transformada de Radon de una función 2D $f(i, j)$ se puede definir matemáticamente como:

$$\mathcal{R}\{f(i,j)\} = \lambda_f(p,\phi) = \int\int_{-\infty}^{+\infty} f(i,j)\,\delta(p - \overrightarrow{r}\,\widehat{\overrightarrow{p}})\,di\,dj \quad (1)$$

donde δ es la función de la delta de Dirac ($\delta(x) = 1$ cuando $x = 0$, y $\delta(x) = 0$ en cualquier otro caso). La línea de integración está descrita por el vector $\overrightarrow{p}$ que se define como $\overrightarrow{p} = \widehat{\overrightarrow{p}} \cdot p$, donde $\widehat{\overrightarrow{p}}$ es un vector unitario en la dirección de $\overrightarrow{p}$. p es el módulo de $\overrightarrow{p}$. $\overrightarrow{r}$ es un conjunto de puntos perpendiculares a $\overrightarrow{p}$.

La Figura 3 muestra una imagen de muestra en blanco y negro, a la izquierda, y su transformada de Radon, a la derecha. Además, muestra gráficamente el proceso para calcular la transformada de Radon.



(a) (b)

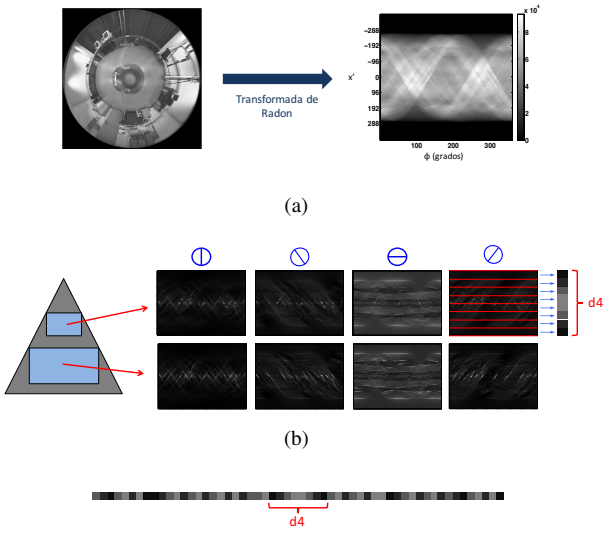
Figura 3: (a) Imagen ejemplo. (b) Transformada de Radon de la imagen ejemplo.

3.3. *Descriptor gist*

Los pasos que hemos implementado para obtener el descriptor *gist* de la transformada de Radon de una imagen omnidireccional son:

1. Construir una pirámide de resoluciones a partir de la transformada Radon original.
2. Filtrado de Gabor.
3. Agrupamiento en bloque.

La Figura 4 muestra el proceso que se propone para obtener el descriptor *gist* en este trabajo. En la Figura 4(a), podemos observar la transformada de Radon de una imagen omnidireccional. En la Figura 4(b) puede verse el uso de cada filtro de Gabor en la transformada de Radon y su agrupación en bloques. En este caso hemos usado $n = 2$ niveles y $b = 8$ bloques horizontales. Finalmente, en la Figura 4(c) se muestra la composición final del descriptor *gist*. El descriptor *gist* resultante estará formado por $4 \cdot n \cdot b$ componentes.



(a)

(b)

(c)

Figura 4: (**a**) Transformada de Radon de una imagen omnidireccional. (**b**) Matrices obtenidas aplicando cada uno de los cuatro filtros de Gabor a los dos niveles de la pirámide. (**c**) Composición final del descriptor *gist*.

3.4. *Histograma de Orientación del Gradiente (HOG)*

HOG se ha utilizado tradicionalmente como un método de descripción en el campo de la detección de objetos. Fue descrito inicialmente por Dalal y Triggs (2005) y lo usaron en tareas de detección de personas. Sin embargo, hay varias investigaciones en las que se ha mejorado este método de descripción, como Zhu et al. (2006), donde mejoran la precisión y el coste computacional.

La idea fundamental de los descriptores de Histograma de Orientación del Gradiente (HOG) es que la apariencia de un objeto y su forma dentro de imagen pueden ser descritas mediante la distribución de la intensidad de gradientes o dirección de bordes. La implementación básica consiste en dividir la imagen en pequeñas celdas conectadas y calcular el histograma de las orientaciones del gradiente de los píxeles en cada celda. Entonces, el descriptor se compone de estos histogramas dispuestos en un solo vector.

Fernández et al. (2016) analizan este tipo de descriptor en tareas de localización al aire libre. Además, hacen un análisis comparativo entre varios métodos para describir imágenes panorámicas.

La Figura 5 muestra el procedimiento de división de celdas horizontales para obtener el descriptor HOG de una imagen.

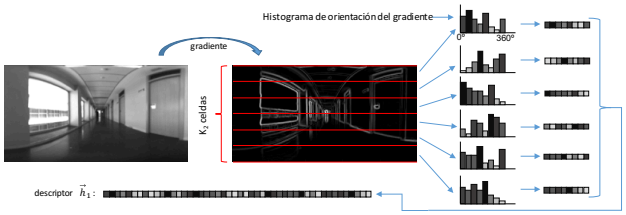


Figura 5: Procedimiento de obtención de un descriptor HOG de una imagen.

4. Localización con Mapas Densos

En esta sección, se propone una solución a los problemas de localización y creación de mapas utilizando sólo la información visual capturada por un sistema de visión omnidireccional montado en el robot. Este sistema se compone de una cámara que apunta a un espejo hiperbólico y captura imágenes omnidireccionales del entorno las cuales se describen con un solo descriptor de apariencia global.

Nuestro punto de partida es una base de datos de transformadas de Radon de imágenes omnidireccionales capturadas en una rejilla de puntos en el entorno donde el robot debe navegar. Se estudiarán dos posibilidades iniciales, en la primera de ellas el robot conoce las posiciones en las cuales han sido capturadas las imágenes de las transformadas de Radon del mapa y su objetivo es localizarse dentro de él. En la segunda de ellas no posee dicha información, por lo tanto deberá crear un mapa local seleccionando los vecinos más cercanos.

En la primera hipótesis el robot calcula cuál es la transformada de Radon que más se parece a la transformada de Radon de la imagen omnidireccional capturada en una posición aleatoria con el objetivo de localizarse dentro del mapa que posee. Para realizar esta comparación hace uso de la correlación en fase POC (*Phase Only Correlation*).

En la segunda hipótesis primero detectará los vecinos más cercanos (nodos de un mapa topológico) mediante el uso de la comparación usando POC. Una vez detectados cuales son los vecinos más cercanos, el robot crea un mapa local utilizando esas transformadas de Radon de los vecinos más cercanos y la transformada de Radon de la imagen capturada en la nueva posición. Para la creación del mapa local hace uso del algoritmo masa-muelle, utilizando como entrada del algoritmo las distancias entre los descriptores *gist* de estas transformadas de Radon.

Los experimentos se han llevado a cabo con varias bases de datos de imágenes virtuales diferentes. También usamos algunas bases de datos reales para probar la validez de los métodos propuestos, formadas por imágenes capturadas en áreas cuadradas de diferentes entornos de interior. La Figura 6 se muestra el ratio de acierto del algoritmo de localización basado en la transformada de Radon (R) utilizando bases de datos reales y añadiendo diferentes niveles de ruido y oclusiones. En esta figura también se pueden observar los resultados obtenidos con dos métodos alternativos basados en la Firma de Fourier (FS) y en la extracción de características locales SIFT (S).

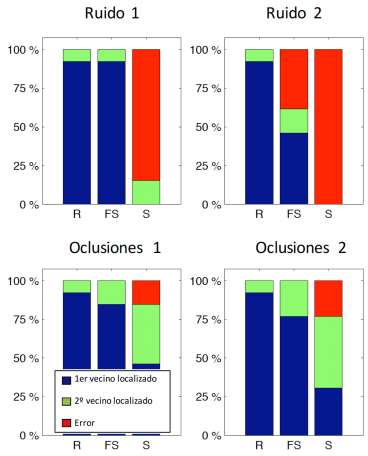


Figura 6: Resultados del algoritmo de localización con mapa denso utilizando bases de datos reales y añadiendo diferentes niveles de ruido (valor máximo de 20 % y 40 % de la intensidad máxima de la imagen) y oclusiones (10 % y 20 %).

5. Estimación de Altura Relativa

En esta sección se muestra el uso de un único sistema de visión catadióptrico para estimar la altitud del robot utilizando un enfoque de apariencia global directamente sobre la imagen omnidireccional adquirida por el sistema de visión. El robot opera en un entorno y utiliza un sistema de visión omnidireccional como única fuente de información. También consideramos que el mapa del entorno se construyó a partir de un conjunto de imágenes capturadas mientras el robot tenía un movimiento plano y usaba técnicas de apariencia global. Teniendo en cuenta este hecho, proponemos dar un paso más en esta línea. El objetivo es intentar estimar la altitud relativa donde se encuentra sin incorporar información adicional al mapa del entorno disponible y utilizando solo la información capturada por el sistema de visión omnidireccional.

Como en la sección anterior, cuyo objetivo era localizar el robot en el plano del suelo, en esta propuesta se utiliza la Transformada Radon como descriptor para estimar la diferencia de altura. La clave de este algoritmo es detectar la compresión sufrida por esta transformada al variar la altura en la imagen omnidireccional. En la Figura 7 se muestran los resultados probando con diferentes bases de datos reales.

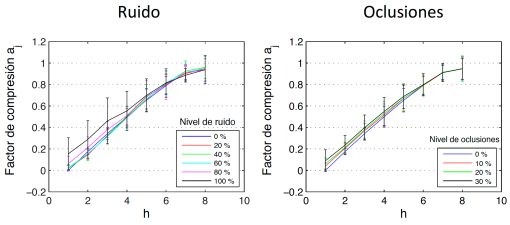


Figura 7: Resultados del algoritmo de estimación de altura utilizando bases de datos reales y añadiendo diferentes niveles de ruido y oclusiones.

6. Odometría Visual optimizada

En esta sección se presenta un algoritmo de odometría visual utilizando únicamente información visual de imágenes omnidireccionales. Básicamente el algoritmo consiste en ir calculando la distancia y la dirección del movimiento de imágenes consecutivas de una trayectoria seguida por el robot.

Para calcular la distancia entre imágenes consecutivas, el robot hace uso de la comparación POC entre transformadas de Radon, lo cual ha ofrecido muy buenos resultados en cuanto a cálculo de distancia entre transformadas de Radon.

Por otro lado para calcular la dirección de desplazamiento se ha desarrollado un algoritmo que consiste en extraer las características SURF de las imágenes omnidireccionales consecutivas y detectar las correspondencias entre ellas. Posteriormente se divide la circunferencia de la imagen mediante una recta que pase por el centro y orientada un determinado ángulo. Dicho ángulo se va variando, y el ángulo que maximice el número de correspondencias con un desplazamiento opuesto al desplazamiento de las correspondencias de la otra mitad de la circunferencia, será el ángulo de desplazamiento. Finalmente la odometría visual se optimiza calculando los cierres de bucle mediante las distancias entre descriptores HOG.

La Figura 8 muestra una ruta calculada mediante este algoritmo, en la cual el robot atraviesa diferentes estancias de un edificio. En azul se muestra la ruta calculada y en amarillo la ruta real seguida por el robot.

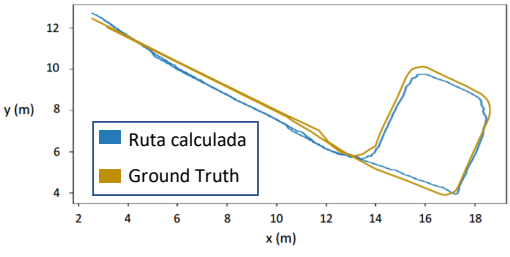


Figura 8: Resultados del algoritmo de creación de odometría visual. En azul se muestra la ruta calculada y en amarillo la ruta real seguida por el robot.

7. Conclusión

Durante los últimos años, los sensores de visión se han convertido en una opción muy robusta en el campo de la robótica a la hora de obtener información del entorno. A su vez, los algoritmos de descripción basados en la apariencia global de las imágenes están demostrando su validez y robustez a la hora de desempeñar tareas de localización y de creación de mapas.

Las investigaciones de esta tesis presentan diferentes algoritmos de localización y creación de mapas utilizando este tipo de descriptores, las cuales han demostrado su validez calculando la localización utilizando mapas densos, estimando la altura relativa entre dos posiciones distintas y calculando una odometría visual de una ruta seguida por el robot. Todo ello validado con experimentos utilizando bases de imágenes omnidireccionales capturadas en condiciones de trabajo reales y realizando comparaciones con diferentes tipos de algoritmos de localización y creación de mapas basados en otros tipos de descriptores.

English Summary

Use of Global Appearance Descriptors for Localization and Mapping

Abstract

This work presents the most relevant contributions made within the thesis 'Uso de descriptores holísticos para la localización y creación de mapas: una aproximación al Graph-SLAM mediante apariencia visual', which will be defended on April 27, 2018 at the University Miguel Hernández de Elche. Nowadays, the field of mobile robotics is increasingly present in different fields and applications such as autonomous driving. Due to this growing presence, the scientific community tries to develop new algorithms to provide mobile robots with greater autonomy to make better decisions in their tasks. With this objective, the robot must develop different essential skills when it comes to navigating unknown environments. One of these skills is to create a map of the environment from information collected with the sensors equipped on it. Another of these skills is to be able to calculate its location within the previously created map. In the case that both tasks are executed at the same time it is known as SLAM. In the investigations of the thesis dealt with in this work, different algorithms for mapping and localization are presented using information obtained through omnidirectional visual sensors. Also, for the description of the images, they use techniques based on the global appearance of the images. All algorithms are tested with different sets of images, both virtually generated images and images captured under real working conditions and taking into account common phenomena such as illumination changes and occlusions.

Keywords:

Omnidirectional vision, holistic descriptors, Radon transform, Localization, Mapping, SLAM.

Agradecimientos

Este trabajo ha sido apoyado por el gobierno español a través del proyecto DPI2016-78361-R (AEI/FEDER, UE) "*Creación de Mapas Mediante Métodos de Apariencia Visual para la Navegación de Robots*".

Referencias

Bay, H., Tuytelaars, T., Gool, L., 2006. Surf: Speeded up robust features. Computer Vision at ECCV 3951, 404–417.

Berenguer, Y., Payá, L., Ballesta, M., Reinoso, O., 2015. Position estimation and local mapping using omnidirectional images and global appearance descriptors. Sensors 15 (10), 26368.
DOI: 10.3390/s151026368

Dalal, N., Triggs, B., June 2005. Histograms of oriented gradients for human detection. En: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Vol. 1. pp. 886–893 vol. 1.
DOI: 10.1109/CVPR.2005.177

Fernández, L., Payá, L., Reinoso, O., Jiménez, L., Ballesta, M., 2016. A study of visual descriptors for outdoor navigation using google street view images. Journal of Sensors 2016.

Fernández, L., Payá, L., Reinoso, O., Jimenez, L. M., 2014. Appearance-based approach to hybrid metric-topological simultaneous localisation and mapping. IET Intelligent Transport Systems 8 (8), 688–699.
DOI: 10.1049/iet-its.2013.0086

Forster, C., Lynen, S., Kneip, L., Scaramuzza, D., Nov 2013. Collaborative monocular slam with multiple micro aerial vehicles. En: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 3962–3970.
DOI: 10.1109/IROS.2013.6696923

Garcia-Fidalgo, E., Ortiz, A., 2015. Vision-based topological mapping and localization methods: A survey. Robotics and Autonomous Systems 64, 1 – 20.
DOI: http://dx.doi.org/10.1016/j.robot.2014.11.009

Kostavelis, I., Charalampous, K., Gasteratos, A., Tsotsos, J. K., 2016. Robot navigation via spatial and temporal coherent semantic maps. Engineering Applications of Artificial Intelligence 48, 173 – 187.
DOI: http://dx.doi.org/10.1016/j.engappai.2015.11.004

Liang, Z.-P., Lauterbur, P. C., 2000. Principles of magnetic resonance imaging: a signal processing perspective. SPIE Optical Engineering Press.

Lowe, D., 1999. Object recognition from local scale-invariant features. En: ICCV 1999, International Conference on Computer Vision. Vol. 2. pp. 1150–1157.

Munguia, R., Urzua, S., Grau, A., 12 2016. Delayed monocular slam approach applied to unmanned aerial vehicles. PLOS ONE 11 (12), 1–24.
DOI: 10.1371/journal.pone.0167197

Payá, L., Amorós, F., Fernández, L., Reinoso, O., 2014. Performance of global-appearance descriptors in map building and localization using omnidirectional vision. Sensors 14 (2), 3033–3064.
DOI: 10.3390/s140203033

Payá, L., Reinoso, O., Berenguer, Y., Úbeda, D., 2016. Using omnidirectional vision to create a model of the environment: A comparative evaluation of global-appearance descriptors. Journal of Sensors 2016.
DOI: 10.1155/2016/1209507

Radon, J., 1917. Uber die bestimmung von funktionen durch ihre integralwerte langs gewisser mannigfaltigkeiten. Berichte Sachsische Akademie der Wissenschaften 69 (1), 262–277.

Weiss, S., Achtelik, M. W., Chli, M., Siegwart, R., May 2012. Versatile distributed pose estimation and sensor self-calibration for an autonomous mav. En: Robotics and Automation (ICRA), 2012 IEEE International Conference on. pp. 31–38.
DOI: 10.1109/ICRA.2012.6225002

Yamazawa, K., Yagi, Y., Yachida, M., 1995. Obstacle detection with omnidirectional image sensor hyperomni vision. En: Robotics and Automation, 1995. Proceedings., 1995 IEEE International Conference on. Vol. 1. IEEE, pp. 1062–1067.

Zhu, Q., Yeh, M.-C., Cheng, K.-T., Avidan, S., 2006. Fast human detection using a cascade of histograms of oriented gradients. En: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). Vol. 2. pp. 1491–1498.
DOI: 10.1109/CVPR.2006.119