

# EVALUACIÓN DE NUEVOS MODOS DE EMPLEO DE LOS DESCRIPTORES DE APARIENCIA GLOBAL EN TAREAS DE LOCALIZACIÓN

Vicente Román, Sergio Cebollada, Luis Payá, María Flores, Arturo Gil, Óscar Reinoso  
 Departamento de Ingeniería de Sistemas y Automática  
 Universidad Miguel Hernández, Avenida de la Universidad s/n 03202 Elche (Alicante), España  
 {vicente.roman, sergio.cebollada, lpaya, m.flores, arturo.gil, o.reinoso}@umh.es

## Resumen

*Los robots autónomos deben ser competentes en tareas de localización y creación de mapas. Una de las posibles soluciones para estos problemas es emplear descriptores de apariencia global que obtienen un único vector que describe globalmente una imagen panorámica previamente adquirida. Comparando los descriptores HOG y Gist, el propósito de este trabajo es el estudio de un nuevo modo de utilizar estos descriptores con la finalidad de sustituir el modo de empleo actual o lo que sería más interesante combinar ambos modos para obtener un método de mayor calidad. Este trabajo se realiza con imágenes reales tomadas en un campo de trabajo heterogéneo en el que simultáneamente conviven personas y robots, por ello en las imágenes hay cambios de iluminación, de mobiliario y personas que ocuyen parte de la escena. Como conclusiones se obtiene que el descriptor HOG tiene un resultado similar empleando ambos modos mientras que con el descriptor Gist se obtiene un peor resultado al utilizar el nuevo modo de ventanas verticales.*

**Palabras clave:** Localización, robótica móvil, descriptores de apariencia global, imágenes omnidireccionales.

## 1. INTRODUCCIÓN

Durante los últimos años ha habido una gran revolución en el mundo de la robótica. Por ello la presencia de los robots en diferentes ambientes se ha visto incrementada, lo que ha provocado un continuo progreso en sus habilidades de percibir la realidad. Sin embargo, la revolución no será total hasta que los robots no puedan operar autónomamente en lugares grandes y heterogéneos. El objetivo final de un robot autónomo es doble ya que el robot debe ser capaz de navegar por estos entornos cambiantes creando un mapa de la nueva área (tarea de mapping) y estimando correctamente su posición en él (tarea de localización).

A lo largo de los años ha aparecido diferentes métodos para resolver estos problemas de mapping

y localización. Una de las opciones es utilizar sensores que captan información como puede ser encoders, sonars [5] o lasers [7], pero esos sensores pueden resultar caros, necesitan un mantenimiento y consumen relativamente bastante energía. Otra opción es la utilización de GPS, pero esta solamente es útil en ambientes de exterior. Como tercer método para recoger información podemos encontrar utilización de información visual. Los sensores visuales ofrecen gran cantidad de datos a un precio relativamente bajo. Entre diferentes configuraciones aparece las cámaras de visión omnidireccional que ofrecen un campo de visión de 360 grados [21]. Esta forma de captar imágenes está siendo estudiada durante los últimos años debido a que el robot recoge una mayor información del escenario lo que le hace más fácil las tareas de navegación. En los trabajos [8], [19] y [16] se puede encontrar ejemplos de base de datos contraídas con estos sistemas de cámaras omnidireccionales.

Las imágenes ofrecen una gran cantidad de datos, por tanto, para resolver las tareas necesarias para la navegación es necesario describir las imágenes pudiendo ser reconocidas más fácilmente. A día de hoy el método de describir puntos característicos de la imagen es muy utilizado. Entre otros, se puede encontrar algoritmos de descripción de características locales como SIFT [9], SURF [2], BRIEF [4]. Esta es una opción muy madurada y con buenos resultados en tareas de mapping y localización. Por ejemplo, Angeli et al.[1] y Valgren and Lilienthal [22] utilizan estos descriptores para tareas de localización topológica o Murillo et al. [11] donde ese presenta un estudio de estos descriptores en tareas de interior. Por otro lado, otro modo de tratar la información visual es los métodos de apariencia global, estos métodos describen la imagen con un único vector que será más invariante antes cambios en la escena global de la imagen. Además, al tener un único descriptor por imagen la tarea de mapping y localización se simplifica a una simple comparación entre vectores permitiendo una posible reducción en tiempo de computación. Se han estudiado diferentes métodos de descripción global entre los que se puede destacar el Histograma de Orientación del Gradiente (HOG), usado por ejemplo en [15] para solventar

problemas de localización de robots móviles. Otra opción es el descriptor Gist, descrito por Oliva y Torralba [12] y utilizado en ambientes de exterior [20]. Otros descriptores de apariencia global conocidos son la Transformada de Fourier [10] o la transformada de Radon[17]. Estas alternativas han sido utilizadas en trabajos como el de Paya et al. [14] para la construcción de mapas o como el de Berenguer Y. et al. [3] que utilizan la Transformada de Radon para la búsqueda de vecinos cercanos. Por último, Román et al. [18] realizaron un estudio de estos métodos de apariencia global en entornos reales para tareas de localización bajo diferentes tareas de iluminación.

Como veremos en el siguiente apartado, estos descriptores de apariencia local son invariantes a la rotación del robot si se utiliza cámaras omnidireccionales. Las cámaras recogen información 360 y los descriptores HOG y Gist recogen la información en ventanas horizontales lo que hace obtener la misma información sin importa la orientación del robot. Este artículo comparará realizar la localización con ese método o hacerlo con ventanas verticales. En resumen, una vez los descriptores han sido utilizados en tareas de localización con relativamente buenos resultados, la finalidad de sustituir el modo de empleo clásico con ventanas horizontales o lo que sería mas interesante combinar ambos modos para obtener un método de mayor calidad. Este trabajo servirá como estudio preliminar a esta posible combinación de modos realizado una comparativa entre ambos modos de empleo.

El trabajo esta estructurado de la siguiente manera. En la siguiente sección los dos descriptores utilizados (HOG y Gist) son presentados. En la sección 3, se presenta la información importante sobre el data set junto al conjunto de imágenes y rutas. A continuación, se describe el modelo y los resultados de los experimentos. Por último las conclusiones son presentadas.

## 2. DESCRIPTORES DE APARIENCIA GLOBAL

Para comparar los resultados se ha utilizado los descriptores HOG y Gist. En ambos casos el punto inicial es la imagen panorámica  $i(x,y) \in R^{N_x \times N_y}$  y tras utilizar los métodos de descripción se obtiene como resultado un vector descriptor por imagen  $\vec{d} \in R^{l \times 1}$ .

Para obtener el descriptor la imagen se debe dividir en ventanas. Se va a estudiar dos modos de construcción. Un primer modo en el que las ventanas son horizontales, al ser las imágenes panorámicas y evaluar los resultados por ventas horizontales

el resultado por cada ventana siempre será el mismo dando igual la orientación del robot. En un segundo modo las ventanas son verticales, en este caso el método consta de dos etapas; en una primera se construye el descriptor de este nuevo modo y en una segunda etapa el propio algoritmo rota uno de los descriptores para adecuarlo una misma orientación antes de comparar dos vectores. De esta forma en este segundo modo de uso, aunque el descriptor depende de la orientación de la imagen se puede comparar vectores ya que relativamente siempre se tendrá la misma orientación. En la Figura 1 se puede observar los dos modos posibles de construir el descriptor con la condición de ventanas horizontales y verticales.

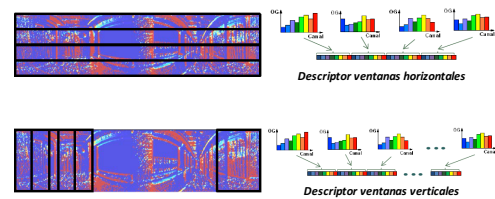


Figura 1: Modos de construir el descriptor

### 2.1. HOG

El método fue descrito por Dalal [6] y en esencia calcula el gradiente de la imagen y a continuación evalúa el módulo orientación del gradiente de cada píxel. Después almacena los resultados por áreas y construye el descriptor, en un primer modo estas áreas surgen de dividir la imagen en ventanas horizontales mientras que en el segundo modo las ventanas son verticales. En ambos casos el descriptor es invariante a la rotación del robot por lo que pueden ser utilizados para la localización.

Por último, como se ha dicho a imagen es reducida a un vector. La longitud de este vector depende de dos factores, el número de bins o posibles intervalos de orientación del gradiente  $b_1$  y el número de ventanas en las que la imagen ha sido dividida  $k_1$ . En este trabajo estos valores son variables para ver como afectan al resultado obteniendo el siguiente descriptor  $\vec{d} \in R^{b_1 \cdot k_1 \times 1}$ .

### 2.2. Gist

Descrito inicialmente por Oliva et al. [13], y desarrollado por Siagian et al. [20]. Este método evalúa la intensidad media de diferentes zonas tras exponer la imagen a diferentes orientaciones de filtros Gabor y a diferentes niveles de resolución. Como con HOG, los descriptores son invariantes

a la orientación del robot ya que el resultado se almacena por las conocidas ventanas horizontales y verticales.

En este caso el tamaño del descriptor depende de el número de orientaciones del filtro Gabor  $m_1$  y el número de ventanas en las que se divide la imagen  $k_2$ . El tamaño también va a depender del número de resoluciones diferentes de la imagen  $r_1$ . Por tanto, en estos estudios la imagen es reducida por el método Gist a un descriptor  $\vec{d} \in R^{r_1 \cdot m_1 \cdot k_2 \times 1}$ .

### 3. CONJUNTO DE IMÁGENES

Para llevar a cabo los experimentos es necesario un entorno dinámico y heterogéneo. La base de datos COLD [16] proporciona estas características ya que el robot recorre una ruta de un laboratorio en la que coetaneamente trabajan hombre y robot. De este modo el robot navega en un entorno cambiante enfrentándose, entre otras cosas, a cambios introducidos por humanos, variaciones en la iluminación y oclusiones de cámara. Esto provoca que las imágenes obtenidas en la ruta creen una base de datos ideal para nuestro estudio. En la Figura 2 se puede observar los fenómenos de heterogeneidad comentados.



Figura 2: Cambios dependiendo de la ruta.

Las imágenes trabajadas han sido adquiridas por un robot móvil el cual está equipado con una cámara omnidireccional. Este equipo de visión está constituido por una cámara montada en un espejo hiperbólico. Este montaje ofrece imágenes omnidireccionales, previamente a aplicar los descriptores de apariencia global la máquina transforma estas imágenes a panorámicas. Adicionalmente, el robot viene equipado con encoders y lasers. Estos lasers ofrecen la posición real del robot que utilizaremos para comprobar si el robot acierta o no posición y con que error, nunca para hacer la

tarea de localización, realizada únicamente con información visual.

Por último, la base de datos COLD ha sido tomada en diferentes edificios. En cada uno de ellos ha seguido diferentes rutas, tomando cada ruta en días diferentes y bajo iluminaciones diferentes. Para realizar nuestros experimentos se ha elegido la ruta más larga, Friburgo Part A, Path 2, size 3 [16]. Esta ruta se puede observar en la Figura 3. Se ha elegido esta ruta debido a la alta complejidad que ofrece realizar la tarea de localización en este entorno, ya que, además de ser la ruta más larga, las estancias de Friburgo contienen muchas ventanas así como paredes de cristal, lo que supone un reto debido a los cambios en las condiciones de luz.

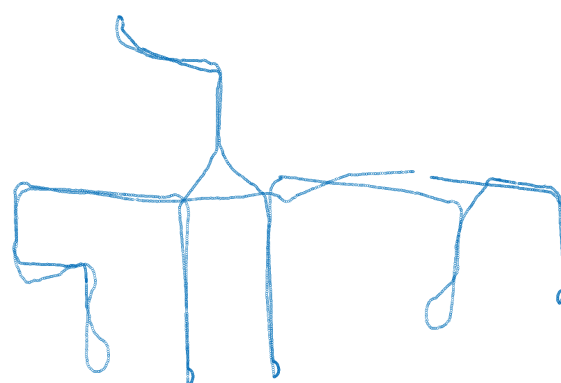


Figura 3: Ruta Friburgo.

### 4. EXPERIMENTOS

En esta sección se presentan de los experimentos que dan validez a los dos modos de uso de los descriptores de apariencia global. En un primer apartado se muestra como se crea el modelo en el que el robot se va a localizar mientras que en el siguiente se puede observar el error que cometido por cada método tras realizar la tarea de localización. Por último, también es importante que el coste computacional no sea muy elevado, por ello se ha hecho un tercer apartado en el que se estudia el tiempo empleado para la descripción y comparación y cálculo de distancia entre dos imágenes.

#### 4.1. CREACIÓN DEL MODELO

Para llevar a cabo los experimentos de localización el robot debe tener un mapa del entorno con anterioridad. Como se ha mencionado en la sección tres, la base de datos COLD ofrece varias rutas en diferentes edificios y con diferentes grados de iluminación. Para la creación del mapa preliminar se ha elegido una ruta en la que hay una imagen

Cuadro 1: Distance

$$\text{Euclidean } d_2(\vec{a}, \vec{b}) = \sqrt{\sum_{i=1}^l (a_i \cdot b_i)^2}$$

capturada cada 0.30m teniendo como modelo una ruta de 556 imágenes. Este modelo está creado por una ruta captada durante un día nublado.

Una vez construido el modelo con las imágenes conocidas. El método para resolver la tarea es comparar cada la imagen captada con las imágenes del modelo creado. Para comparar los descriptores es necesario que ambos tengan el mismo tamaño, para ello es necesario hacer una selección de parámetros. En el trabajo de Román V. et al. [18] se puede observar un estudio de los parámetros ante cambios de iluminación obteniendo como conclusión la utilización de los siguientes parámetros. Por un lado, al utilizar el descriptor HOG se sugiere utilizar  $b_1=16$  y  $k_1=16$ ; mientras que al utilizar el descriptor Gist se detecta los parámetros  $m_1=4$ ,  $k_2=32$  y  $r_1=2$  como ideales. Ambos descriptores tendrán un tamaño  $\vec{d} \in R^{256 \times 1}$ , de este modo los tiempos empleados serán comparables.

#### 4.2. LOCALIZACIÓN DEL ROBOT

La 556 imágenes de la ruta de referencia separadas aproximadamente 0.30m crean un modelo donde el robot va a tratar de localizarse. Para tratar de realizar esta tarea de localización se han empleado tres rutas distintas del mismo entorno pero con diferente grado de iluminación. Todas las rutas pasan por las mismas estancias pero fueron captadas durante día nublado, soleado y de noche. Estas rutas alternativas están formadas por 2778 imágenes, 2231 y 2889 respectivamente. Estas imágenes tienen una separación aproximada de 0.06m

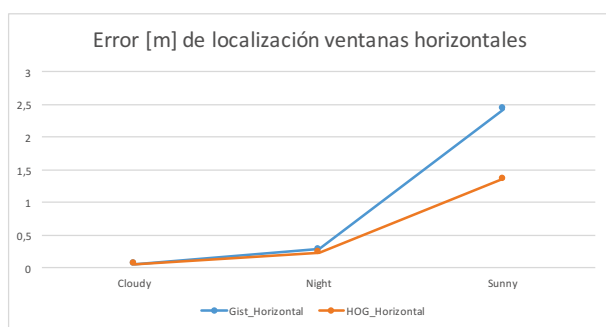


Figura 4: Error de localización al utilizar las ventanas en posición horizontal. Error calculado en metros.

Durante el proceso de localización el método describe la nueva imagen adquirida y la compara con las imágenes descritas de la ruta de referencia. Una vez evaluada la distancia entre descriptores

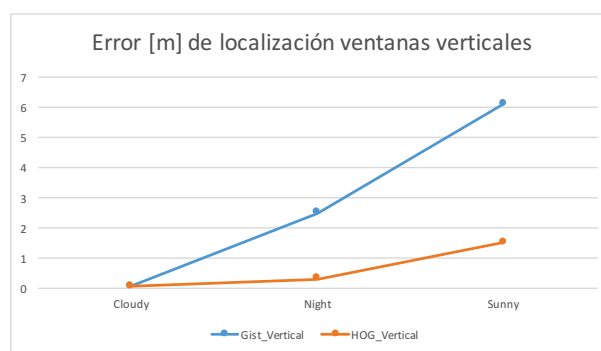


Figura 5: Error de localización al utilizar las ventanas en posición vertical. Error calculado en metros.

se selecciona, entre las imágenes de la ruta de referencia, aquella con menor distancia entre descriptores. Está será elegida como posición del robot. Para calcular el error se calcula la distancia *Euclidean*, tabla 1, entre la posición real del robot y la calculada. Este proceso es posible porque disponemos de la posición real del robot en cada momento, pero esta posición solo se utiliza como 'ground truth' para detectar el error.

Una vez realizado el proceso de localización las Figuras 4 y 5 muestran en el error de localización medio calculado en metros. Estos datos muestran el error en las tres condiciones comentadas durante el trabajo Cloudy (nublado), Night (noche) y Sunny (soleado). Con HOG y ventanas horizontales se obtiene un error sobre nublado de 0.0599m, sobre noche de 0.02377m y sobre soleado de 1.3715m mientras que al utilizar las ventanas verticales los resultados ascienden a un error medio de 0.0462m, 0.3058m, 1.5163m respectivamente. Por otro lado, al utilizar Gist y el modo clásico de ventanas horizontales el error medio ante situación nublado es de 0.0599m, ante noche es de 0.2914m y ante soleado se obtiene un error de 2.435m. Al utilizar el modo de ventanas verticales el error medio asciende en un mayor grado a valores de 0.0576m, 0.2377m y 6.0955m respectivamente.

Los resultados muestran un mejor comportamiento al utilizar el método HOG incrementándose el error al utilizar las ventanas verticales. Asimismo es observable un alto error al utilizar como base test las imágenes captadas en soleado, esto puede ser debido a que las condiciones estructurales de Friburgo son muy extremas para los cambios de luz ya que la estancia tiene muchas ventanas y paredes de cristal.

En la Figura 6 se puede ver una comparativa del error al utilizar los dos tipos de descripción de la imagen. La Figura 6 muestra en un diagrama de barras el error al utilizar los descriptores (Gist y

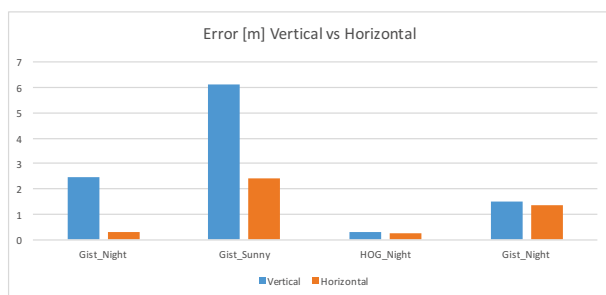


Figura 6: Comparativa del error al utilizar las ventanas verticales u horizontales.

HOG) únicamente en las condiciones de Night (noche) y Sunny (soleado). Las barras azules corresponden con los errores medios obtenidos al calcular la localización con la nueva disposición de ventanas verticales y las barraras naranjas muestran el error con una disposición horizontal de estas. El error está calculado en metros.

En general, y una vez estudiado los datos mostrados en la Figura 6, se observa que hay una diferencia muy grande al utilizar Gist como descriptor de localización con modo en el que las ventanas están verticales. Sin embargo, en el método de HOG la diferencia de error no es tan grande, aunque sigue siendo menor el error medio con las ventanas en posición horizontal.

### 4.3. TIEMPOS

Tener un error de localización bajo es importante para la calidad del descriptor pero también es importante coste computacional del proceso razonable. Por ello se ha estudiado el tiempo empleado por el proceso en describir dos imágenes y calcular la distancia entre ellas. La tabla 2 muestra el valor en segundos del proceso. Recordamos que para calcular los descriptores se ha elegido como parámetros de HOG  $b_1=16$  y  $k_1=16$ . Por tanto el descriptor HOG tiene un tamaño de vector:  $\vec{d} \in R^{256 \times 1}$ . A su vez, para el descriptor Gist se ha elegido  $m_1=4$ ,  $k_2=32$  y  $r_1=2$ . Por tanto, en estos estudios la imagen es reducida por el método Gist a un descriptor que también tendrá un tamaño  $\vec{d} \in R^{256 \times 1}$ .

Cuadro 2: Tiempo empleado descriptores.

Descriptor	Tiempo
<i>HOG ventanas horizontales</i>	0.0170s
<i>Gist ventanas horizontales</i>	0.0352s
<i>HOG ventanas verticales</i>	0.1689s
<i>Gist ventanas verticales</i>	0.0531s

Se puede observar en la tabla 2 que el proceso para obtener la distancia entre dos imágenes es más

rápido con las ventanas en posición horizontal. Esto es debido a que el método solo tiene que transformar las imágenes en vector y comparar ambos con la distancia *Euclidea*. Sin embargo, cuando las ventanas son verticales se describen las imágenes y se rota uno de los dos vectores para tener la misma orientación relativa, una vez se tiene la misma orientación se calcula la distancia entre ambos. Este proceso de rotación hace incrementar el tiempo de procesamiento cuando utilizamos ventanas verticales.

Los experimentos han sido realizados con un procesador 8-Core Intel Xeon E5® a 3GHz y utilizando la herramienta matemática Matlab®.

Los datos de tiempo no son absolutos ya que dependerán del procesador, pero pueden ser comparables ya que todos han sido realizados con el mismo equipo.

## 5. CONCLUSIONES

El presente trabajo se ha centrado en el estudio y comparativa de dos modos diferentes de describir las imágenes para una futura tarea de localización y/o navegación. Este estudio se ha realizado en un entorno donde los cambios de iluminación y la actividad humana influían negativamente en la tarea de localización. Haciendo uso únicamente de información visual se ha comparado dos algoritmos de apariencia global cada uno de ellos con los dos modos de uso, ventanas horizontales y verticales. Una vez realizada la descripción de las imágenes y calculado el error de localización se ha realizado también una evaluación del tiempo empleado en describir dos imágenes y calcular la distancia entre ambas.

Por un lado, tanto Gist como HOG ofrecen un error de localización adecuado con entorno nublado y de noche, teniendo errores medio de HOG y el nuevo modo de ventanas verticales de 0.0462m en entorno nublado y 0.3058m en entorno noche. Sin embargo este nuevo modo ante un entorno soleado ofrece unos errores medios de localización altos alcanzando el error de 1.5163m al utilizar HOG o incluso 6.0955m al usar Gist. En lo referente a la utilización de ventanas horizontales o verticales se observa que utilizando el descriptor HOG el resultando no varía en exceso y es posible la utilización de este segundo modo de ventanas verticales. Por otro lado, al utilizar ventanas verticales y el descriptor Gist el error de localización aumenta notoriamente y no sería adecuado su uso. En lo referente a los tiempo, utilizar ventanas verticales es más pesado computacionalmente debido a que hay que describir las imágenes, girarlas para tener la misma orientación relativa y a continua-

ción calcular la distancia.

Una vez estudiado los resultados se puede concluir que utilizar el modo de ventanas verticales únicamente no es una solución válida para realizar la tarea de localización. Pero este trabajo puede ser el primer paso en el estudio de este modo llegando a un punto en el que también sea utilizado en la tarea de navegación, no como condición suficiente pero sí como parte del proceso dándole un porcentaje de influencia en el resultado final. De este forma en una localización real no solo será importante la distancia de posición calculada con el modo de ventanas horizontales sino que también influirá otro tipo de errores como el calculado con un modo de ventanas verticales.

### Agradecimientos

Este trabajo está cofinanciado por la Unión Europea a través del Programa Operativo del Fondo Social Europeo (FSE) de la Comunitat Valenciana 2014-2020 a través de su beca predocctoral ACIF/2018/146 y ACIF/2018/224 y gracias al Ministerio de Ciencia e Innovación a través del proyecto DPI 2016-78361-R (AEI/FEDER, UE): 'Creación de mapas mediante métodos de apariencia visual para la navegación de robots.'

### English summary

## EVALUATING NEW GLOBAL-APPEARANCE USING MODELS IN A VISUAL LOCALIZATION TASKS

### Abstract

*Map building and localization are two important tasks that autonomous mobile robots have to deal with. In the last decades it has appeared many studies and techniques to approach these problems. This work briefly describes some possible solutions and it recommends global appearance descriptors to carry our localization tasks. The aim is to obtain a unique vector that describes globally the panoramic image. The mobile robot is able to estimate its position and to create maps using these methods, the result will be independent of the robot orientation. This paper goal is to use a brand new way to use the global appearance descriptors. The final result could be to use this new way or, a more interesting result, use a combination between both*

*ways taken as a result a third better way. This work will consist on a comparison between both global appearance method ways. The study has been carried out using real images taken in an heterogeneous atmosphere where humans and robots work together, for that reason light condition changes and humans can appear on the scene.*

**Keywords:** Localization, mobile robots, global appearance descriptors, omnidirectional images.

### Referencias

- [1] ANGELI, A., DONCIEUX, S., MEYER, J.-A., AND FILLIAT, D. Visual topological slam and global localization. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on* (2009), IEEE, pp. 4300–4305.
- [2] BAY, H., ESS, A., TUYTELAARS, T., AND VAN GOOL, L. Speeded-up robust features (surf). *Computer vision and image understanding* 110, 3 (2008), 346–359.
- [3] BERENQUER, Y., PAYÁ, L., PEIDRÓ, A., GIL, A., AND REINOSO, O. Nearest position estimation using omnidirectional images and global appearance descriptors. In *Robot 2015: Second Iberian Robotics Conference* (2016), Springer, pp. 517–529.
- [4] CALONDER, M., LEPETIT, V., STRECHA, C., AND FUA, P. Brief: Binary robust independent elementary features. In *European conference on computer vision* (2010), Springer, pp. 778–792.
- [5] CHA, Y., AND KIM, D. Omni-directional image matching for homing navigation based on optical flow algorithm. pp. 1446–1451. cited By 1.
- [6] DALAL, N., AND TRIGGS, B. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (June 2005), vol. 1, pp. 886–893 vol. 1.
- [7] HATA, A., AND WOLF, D. Outdoor mapping using mobile robots and laser range finders. pp. 209–214. cited By 6.
- [8] KRÖSE, B. J., VLASSIS, N., AND BUNSCHOTEN, R. Omnidirectional vision for appearance-based robot localization. In *Sensor based intelligent robots*. Springer, 2002, pp. 39–50.
- [9] LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision* 60, 2 (2004), 91–110.
- [10] MENEGATTI, E., MAEDA, T., AND ISHIGURO, H. Image-based memory for robot navigation using properties of omnidirectional images. *Robotics and Autonomous Systems* 47, 4 (2004), 251 – 267.

- [11] MURILLO, A. C., GUERRERO, J. J., AND SAGUES, C. Surf features for efficient robot localization with omnidirectional images. In *Robotics and Automation, 2007 IEEE International Conference on* (2007), IEEE, pp. 3901–3907.
- [12] OLIVA, A., AND TORRALBA, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision* 42, 3 (2001), 145–175.
- [13] OLIVA, A., AND TORRALBA, A. Building the gist of a scene: The role of global image features in recognition. *Progress in brain research* 155 (2006), 23–36.
- [14] PAYÁ, L., FERNÁNDEZ, L., REINOSO, Ó., GIL, A., AND ÚBEDA, D. Appearance-based dense maps creation-comparison of compression techniques with panoramic images. In *ICINCO-RA* (2009), pp. 250–255.
- [15] PAYÁ, L., REINOSO, O., BERENQUER, Y., AND ÚBEDA, D. Using omnidirectional vision to create a model of the environment: A comparative evaluation of global-appearance descriptors. *Journal of Sensors* 2016 (2016).
- [16] PRONOBIS, A., AND CAPUTO, B. COLD: COsy Localization Database. *The International Journal of Robotics Research (IJRR)* 28, 5 (May 2009), 588–594.
- [17] RADON, J. 1.1 über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. *Classic papers in modern diagnostic radiology* 5 (2005), 21.
- [18] ROMÁN, V., PAYÁ, L., AND REINOSO, Ó. Evaluating the robustness of global appearance descriptors in a visual localization task, under changing lighting conditions. In *ICINCO-RA* (2018), pp. 258–265.
- [19] SAITO, M., AND KITAGUCHI, K. Appearance based robot localization using regression models. *IFAC Proceedings Volumes* 39, 16 (2006), 584–589.
- [20] SIAGIAN, C., AND ITTI, L. Biologically inspired mobile robot vision localization. *IEEE Transactions on Robotics* 25, 4 (2009), 861–873.
- [21] STURM, P., RAMALINGAM, S., TARDIF, J.-P., GASPARINI, S., BARRETO, J., ET AL. Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends® in Computer Graphics and Vision* 6, 1–2 (2011), 1–183.
- [22] VALGREN, C., AND LILIENTHAL, A. J. Sift, surf & seasons: Appearance-based long-term localization in outdoor environments. *Robotics and Autonomous Systems* 58, 2 (2010), 149–156.

under the terms and conditions of the Creative Commons Attribution CC BY-NC-SA 4.0 license (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>).



© 2019 by the authors.  
Submitted for possible  
open access publication