



**MÁSTER UNIVERSITARIO  
EN VISION ARTIFICIAL**

Curso Académico 2016/2017

**Trabajo de Fin de Máster**

# Estado del Arte de VisualSLAM

**Autor:** Elías Barcia Mejias

**Tutores:** José María Cañas Plaza , Eduardo Perdices García



# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Aplicaciones</b>	<b>3</b>
2.1. Proyecto Tango . . . . .	3
2.2. Magic Plan . . . . .	4
2.3. Photo Tourism . . . . .	5
2.4. Aplicaciones en Robótica Móvil . . . . .	6
<b>3. Problemas de Visual Slam</b>	<b>10</b>
<b>4. Estado del Arte</b>	<b>11</b>
<b>5. Técnicas VisualSLAM</b>	<b>12</b>
<b>6. Conclusiones</b>	<b>13</b>
<b>Bibliografía</b>	<b>14</b>

# Índice de figuras

2.1. El primer smartphone compatible con Tango de Lenovo(a). El primer Smartphone compatible con Tango y DayDream de ASUS (b). . . . .	4
2.2. La pantalla de un smartphone utilizando Magic Plan. . . . .	5
2.3. Recreación del Coliseo de Roma. . . . .	6
2.4. El Robot Gita siguiendo a su dueño (a) El cinturón con cámaras estéreo (b) La capacidad de carga del robot Gita (c) . . . . .	7
2.5. Robot Dyson 360 Eye (a) Robot Roomba 966 (b) Robot Hombot de LG (c). . . . .	9
2.6. Dron equipado con dispositivo compatible con Tango . . . . .	9

# Capítulo 1

## Introducción

Actualmente la investigación y desarrollo en robótica móvil está en pleno auge. Uno de los objetivos más importantes que se persigue, es que los robots móviles puedan desplazarse por su entorno y navegar desde la posición A a la posición B de forma autónoma. Esta tarea no resulta muy complicada en entornos estructurados, donde el robot conoce a priori el terreno por el que se mueve o sabe de la existencia de alguna baliza que le de pistas de su posición. Pero en entornos no estructurados, donde el robot desconoce por completo el terreno, carece de mapas y no existe ningún tipo de marca o baliza que pueda guiar al robot, la navegación resulta mucho más compleja. En exteriores, podríamos guiar al robot mediante GPS, pero la señal GPS no llega con la suficiente potencia a todas partes, por ejemplo en interiores de edificios o en zonas subterráneas, o mejor aún imaginemos que enviásemos a nuestro robot a explorar la superficie del planeta Marte donde la señal GPS es inexistente. ¿Cómo se las arreglaría nuestro robot para desplazarse por el terreno de forma autónoma sin perderse? Actualmente existe una técnica que permite al robot navegar de manera autónoma por zonas desconocidas para él, esta técnica se llama VisualSLAM.

Visual SLAM (Simultaneous Localization and Mapping) es una técnica utilizada principalmente con robots móviles y que aporta al robot la capacidad de autolocalizarse y generar mapas del entorno que le rodea en tiempo real. Básicamente se comporta como una caja negra que procesa las imágenes en secuencia captadas por una o varias cámaras y a partir de esas imágenes, el robot es capaz de obtener su posición 3D en el mundo que le rodea, de esta forma el robot podrá desplazarse en su entorno de forma autónoma sin perderse. El robot además debe contar con una capacidad de cálculo suficientemente potente que le permita ejecutar un software de visión artificial para procesamiento de imágenes y al mismo tiempo realizar la generación de mapas. Estas tareas requieren ser ejecutadas con cierta velocidad, unos 30 frames por segundo. Dependiendo del tipo de cámaras con

las que esté equipado el robot, tendrá mayor o menor capacidad de ejecutar VSLAM. Como mínimo debe tener una cámara RGB , muy común en los drones, aunque también pueden tener 2 camaras estereo que le ayudarán a representar el entorno en 3d con mayor fiabilidad . Otras cámaras como las utilizadas en el proyecto Tango se ayudan de un sensor de profundidad que tambien capacita al robot para representar en tres dimensiones el mundo que les rodea con mayor robusted y precisión. Esta técnica es posible utilizarla hoy en día en pequeños dispositivos gracias al aumento de su potencia de computación.

# Capítulo 2

## Aplicaciones

Hoy en día Visual SLAM ya tiene muchas aplicaciones y aún más que están por llegar en un futuro próximo, a continuación se expondrán varios ejemplos de aplicaciones desde teléfonos móviles hasta robots aspiradora.

### 2.1. Proyecto Tango

Este proyecto trata de equipar a los smartphones y Tablets con sistema operativo Android de 2 cámaras, 1 RGB y otra que captura la profundidad , es decir la distancia a la que se encuentra cada pixel, así el smartphone es capaz de construir un mapa en 3d del entorno y aplicar realidad aumentada en la pantalla del movil.

Al ser una tecnología nueva aún no hay un elevado número de dispositivos que lo soporten . De momento existen 2 móviles compatibles con Tango , el Lenovo Phab 2 pro y el Asus Zenfone AR. En el caso del Zenfone AR estará equipado con 3 cámaras traseras , una para seguir objetos (motion tracking) , otra para detectar profundidad y otra de alta resolución de 23 MP . Con estas 3 cámaras el smartphone podrá crear una modelo tridimensional del entorno y seguir su movimiento. La cámara de localización permitirá al ZenFone conocer su posición 3D en todo momento mientras se mueve por el entorno. La cámara de profundidad está equipada con un proyector de Infrarrojos que le permite medir distancias hasta los objetos en el mundo real.



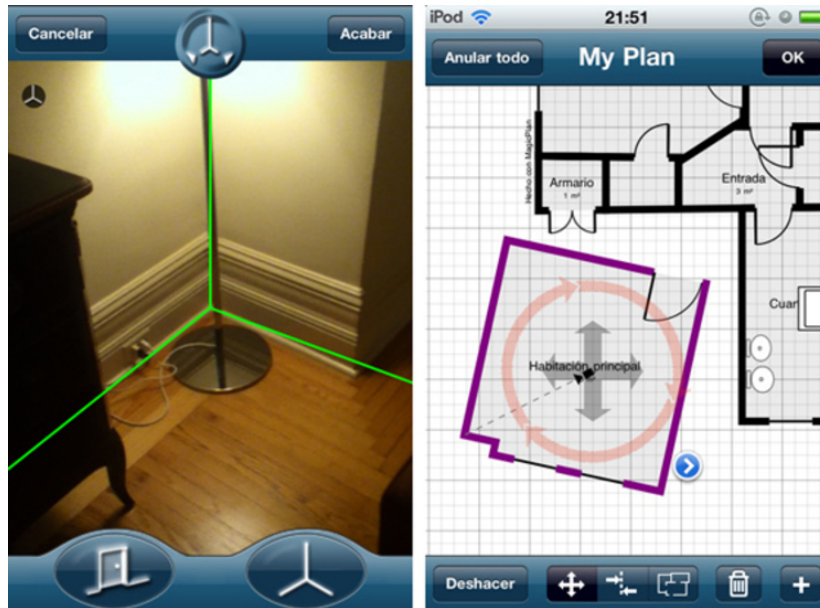
Figura 2.1: El primer smartphone compatible con Tango de Lenovo(a). El primer Smartphone compatible con Tango y DayDream de ASUS (b).

## 2.2. Magic Plan

Magic Plan es una aplicación que permite de forma interactiva obtener planos de habitaciones o del interior de un edificio, utilizando para ello la cámara de nuestra tablet o smartphone, sólo es necesario sacar fotos. Esta aplicación es gratuita, aunque si se desea obtener el plano en formato digital (pdf, jpg, csv y otros) será necesario pagar una pequeña cantidad de dinero. Es muy sencilla de utilizar y en cuestión de minutos se obtiene un plano fiable sin necesidad de medir, dibujar, mover muebles y sin necesidad de ser un experto. La aplicación utiliza técnicas de Visual SLAM y se apoya también en la información de los giroscopios de los dispositivos. Es compatible con Android y dispositivos Apple.

En el caso de Android, actualmente la última versión es compatible con el sistema Tango, por tanto el procedimiento de captura es mucho más sencillo, robusto y preciso ya que permite detectar con mayor exactitud todas las paredes de la habitación, visualizarlas en 3D y aplicar realidad aumentada.



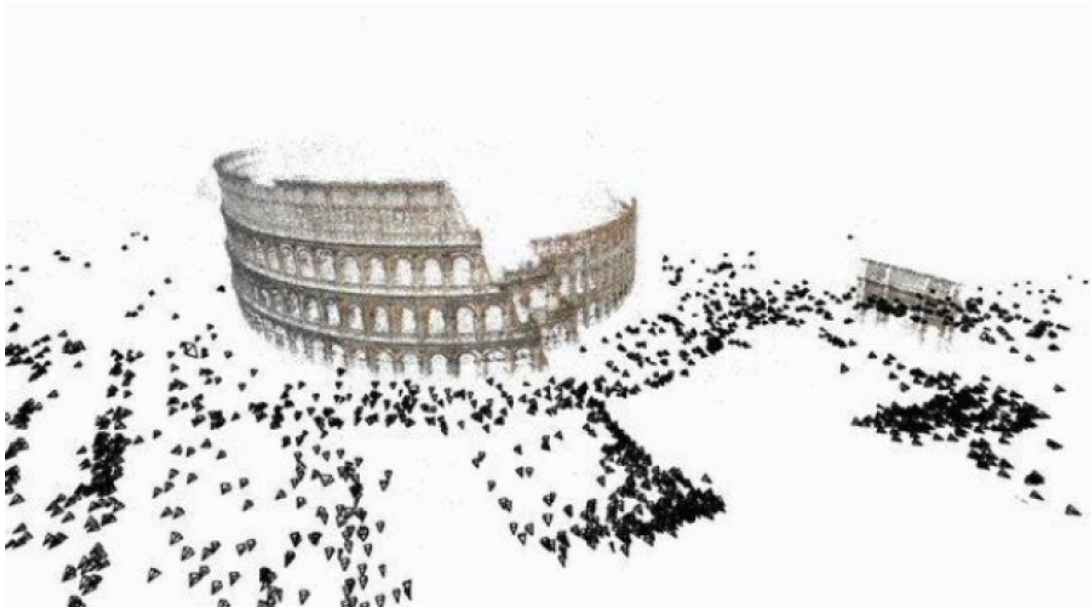


(a)

Figura 2.2: La pantalla de un smartphone utilizando Magic Plan.

### 2.3. Photo Tourism

PhotoTourism o Photo Synth es un software inicialmente creado por la universidad de Whashington en colaboración con Microsoft. Es un sistema que toma grupos de conjunto de fotografías disponibles online sobre un lugar en concreto ,normalmente sobre un monumento turístico mundialmente conocido (como NotreDame, el Coliseo, La Fontana de Trevi) y es capaz de reconstruir puntos 3D de los monumentos y tambien calcular la posición de la cámara desde donde se tomaron las fotografías. Proporciona una nueva forma de navegar a traves de fotografías de un destino turístico y una nueva forma de hacer visitas virtuales a monumentos. Este sistema utiliza la técnica de Structure From Motion SFM. SFM encuentra coincidencias de puntos característicos entre distintas fotografías de un mismo lugar y que han sido tomadas desde distintos puntos de vista y así es capaz de calcular la localización 3D de dichos puntos característicos y tambien la localización 3D desde donde se tomaron las fotografías. A diferencia de VisualSLAM , el procesamiento de estas fotografía es offline , por lo que pueden ser ejecutadas desde un PC que por lo general tiene una capacidad de computación mucho mayor que una tablet o teléfono movil.



(a)

Figura 2.3: Recreación del Coliseo de Roma.

## 2.4. Aplicaciones en Robótica Movil

Visual SLAM tiene aplicaciones directas en robótica . Un ejemplo podría ser el Robot Gita de Piaggio . El robot Gita tiene incorporadas varios pares de cámaras estereo, en la parte trasera y delantera. Con las imágenes captadas por estas cámaras es puede realizar VisualSLAM, además es capaz de seguir a a su dueño siempre y cuando el humano lleve un cinturón con otras 2 cámaras estereo. El robot dispone de un compartimento interior o maletero y tiene suficiente potencia como para poder transportar hasta 20 Kg. De esta forma el robot podría sustituir al conocido carrito de la compra Muy util como carrito de la compra, ya que el robot nos seguirá transportando la compra en su interior, ya no necesitaremos el típico carrito , incluso nos permitiría ir a hacer la compra en bici . Otra utilidad sería en el interior de un hotel podría ser camarero y hacer servicio de habitaciones transportando la comida directamente a las habitaciones del hotel. Tambien podría ser un estupendo ayudante para un mecánico, ya que podría transportar la pesadas herramientas o piezas.

Otra utilidad de SLAM es que mejoran la capacidad de los robots móviles a la hora de reconocer objetos. Los sistemas de reconocimiento de objetos utilizarán la información proporcionada por SLAM para mejorar su capacidad de reconocimiento. La capacidad de reconocimiento será muy util para aquellos robots que tengan que manipular objetos en su



(a)



(b)



(c)

Figura 2.4: El Robot Gita siguiendo a su dueño (a) El cinturón con cámaras estéreo (b)  
La capacidad de carga del robot Gita (c)

entorno. Con SLAM , los sistemas de reconocimiento pueden tomar como entradas varias imágenes desde distintos puntos de vista, por lo tanto el reconocimiento resulta más sencillo que si tuviesen tan sólo una imagen estática.

Recientemente ha entrado en los hogares el uso de VisualSLAM gracias a los últimos modelos de aspiradora equipados con cámaras . Estos aspiradores robótizados disponen de cámaras que le permiten obtener un mapa de la habitación o planta del edificio y gracias a este mapa son capaces de aspirar toda la superficie de manera eficiente, sin dejar ninguna zona de la planta sin limpiar. Además están equipados con sensores de proximidad , que les permiten esquivar obstáculos , aunque tengan que modificar su recorrido momentaneamente son capaces de seguir limpiando ya que pueden utilizar el mapa para continuar su ruta.

Entre los distintos aspiradores estarían:

- Aspirador Dyson 360 Eye.
- Aspirador Roomba 966.
- Aspirador LG-Hombot.

Tanto el modelo de Dyson como Roomba utilizan una cámara de 360 grados, en cambio el modelo de LG utiliza una doble cámara, y es capaz de aspirar la casa incluso en la oscuridad.

Por último no podemos olvidar los drones, robots voladores equipados con cámara que tambien pueden obtener mapas de su entorno con VSLAM. Existen tambien proyectos para equipar a drones con dispositivos compatibles con Tango para que sean capaz de obtener mapas de interiores con mayor precisión , robusted y velocidad.



(a)



(b)



(c)

Figura 2.5: Robot Dyson 360 Eye (a) Robot Roomba 966 (b) Robot Hombot de LG (c).



(a)

Figura 2.6: Dron equipado con dispositivo compatible con Tango

## Capítulo 3

### Problemas de Visual Slam

## Capítulo 4

### Estado del Arte

## Capítulo 5

### Técnicas VisualSLAM



## Capítulo 6

## Conclusiones

# Bibliografía

- [Agüero, 2010] Carlos Agüero. Técnicas de percepción compartida aplicadas a la asignación dinámica de roles en equipos de robots móviles. *Tesis doctoral - Universidad Rey Juan Carlos*, 2010.
- [Barrera, 2008] Pablo Barrera. Aplicación de los métodos secuenciales de monte carlo al seguimiento visual 3d de múltiples objetos. *Tesis doctoral - Universidad Rey Juan Carlos*, 2008.
- [Crespo, 2003] María Ángeles Crespo. Localización probabilística en un robot con visión local. *Proyecto Fin de Carrera. ITIS - Universidad Rey Juan Carlos*, 2003.
- [Domínguez, 2009] José Manuel Domínguez. Autolocalización probabilística visual de un robot en el simulador gazebo. *Proyecto Fin de Carrera. Ing. Informática - Universidad Rey Juan Carlos*, 2009.
- [García, 2007] Iván García. Reconstrucción 3d visual con algoritmos evolutivos. *Proyecto Fin de Carrera. ITIS - Universidad Rey Juan Carlos*, 2007.
- [Hidalgo Blázquez, 2006] Víctor Manuel Hidalgo Blázquez. Comportamiento de persecución de un congénere con el robot pioneer. *Proyecto Fin de Carrera. ITIS - Universidad Rey Juan Carlos*, 2006.
- [Kachach, 2005] Redouane Kachach. Localización del robot pioneer basada en láser. *Proyecto Fin de Carrera. ITIS - Universidad Rey Juan Carlos*, 2005.
- [Kachach, 2008] Redouane Kachach. Calibración automática de cámaras en la plataforma jdec. *Proyecto Fin de Carrera. Ing. Informática - Universidad Rey Juan Carlos*, 2008.
- [López, 2005] Alberto López. Localización visual del robot pioneer. *Proyecto Fin de Carrera. ITIS - Universidad Rey Juan Carlos*, 2005.

- [López, 2010] Luis Miguel López. Autolocalización en tiempo real mediante seguimiento visual monocular. *Proyecto Fin de Carrera. Ing. de Telecomunicación - Universidad Rey Juan Carlos*, 2010.
- [Martín *et al.*, 2010] Francisco Martín, Carlos Agüero, José M. Cañas, and Eduardo Perdices. Humanoid soccer player design. *Robot Soccer*, ed. Vladan Papic. *IN-TECH*, pp 67-100, 2010. ISBN: 978-953-307-036-0, 2010.
- [Martín, 2008] Francisco Martín. Aportaciones a la auto-localización visual de robots autónomos con patas. *Tesis doctoral - Universidad Rey Juan Carlos*, 2008.
- [Martínez, 2003] Juan José Martínez. Equipo de fútbol con jde para la liga simulada robocup. *Proyecto Fin de Carrera. ITIS - Universidad Rey Juan Carlos*, 2003.
- [Marugán, 2010] Sara Marugán. Seguimiento visual de personas mediante evolución de primitivas volumétricas. *Proyecto Fin de Carrera. Ing. Informática - Universidad Rey Juan Carlos*, 2010.
- [Perdices, 2009] Eduardo Perdices. Autolocalización visual en la robocup basada en detección de porterías 3d. *Proyecto Fin de Carrera. Ing. Informática - Universidad Rey Juan Carlos*, 2009.
- [RoboCup, 2009] RoboCup. Robocup standard platform league (nao) rule book. <http://www.tzi.de/spl/pub/Website/Downloads/Rules2009.pdf>, Mayo 2009.
- [Smith *et al.*, 2006] P. Smith, I. Reid, and Andrew J. Davison. Real-time monocular slam with straight lines. *British Machine Vision Conference*, Septiembre 2006.
- [Vega, 2008] Julio Vega. Navegación y autolocalización de un robot guía de visitantes. *Proyecto Fin de Carrera. Ing. Informática - Universidad Rey Juan Carlos*, 2008.
- [Álvarez, 2001] José María Álvarez. Implementación basada en lógica borrosa de jugadores para la robocup. *Proyecto Fin de Carrera. ITIS - Universidad Rey Juan Carlos*, 2001.