

Coordinación de Robots Aéreo y Terrestre para la Exploración de Entornos Desconocidos

Diego Diaz Segovia¹, Evelyn Naomi Gallardo Licea¹, Miguel Ángel Rojas Andrade¹, Juan Pablo I. Ramírez Paredes^{1,*}

¹Departamento de Ingeniería Electrónica, Universidad de Guanajuato.
jpi.ramirez@ugto.mx

*Autor de correspondencia

Resumen

Este artículo implementa un sistema de Colaboración Heterogénea de Robots con un robot terrestre (Pioneer 3 DX) y un dron aéreo (cuadricóptero) en un entorno de interior, con el objetivo de generar un mapa ocupación compartido de la escena. Se integran detección de marcadores fiduciaros, mapeo mediante sensores ultrasónicos, sensores visuales y fusión de mapas. El sistema dota al dron de la capacidad de detectar y seguir un marcador fiduciario ubicado sobre el robot terrestre, permitiéndole sobrevolar u orbitar alrededor de este mientras registra información del entorno desde una perspectiva aérea. Simultáneamente, el robot terrestre realiza una navegación reactiva basada en sensores y genera un mapa de ocupación bidimensional. Los mapas construidos por ambas plataformas se fusionan posteriormente para obtener una representación espacial más completa del entorno. Se evaluaron distintas configuraciones experimentales: (1) mapeo únicamente con el robot terrestre; (2) dron estático sobre el robot terrestre a 12 m y 20 m de altura; y (3) dron orbitando al robot terrestre a 12 m y 20 m.

Palabras clave: robótica móvil, exploración robótica, robótica colaborativa, visión artificial, navegación reactiva.

Introducción

En la actualidad, la colaboración entre robots aéreos y terrestres representa una de las áreas más prometedoras dentro del campo de la robótica móvil. Este proyecto busca implementar un sistema de mapeo colaborativo entre un vehículo aéreo no tripulado (UAV, por sus siglas en inglés) y un vehículo terrestre no tripulado (UGV, por sus siglas en inglés), integrando sensores con diferentes enfoques y buscando sus capacidades para terminar complementándose durante la exploración desde diferentes perspectivas. A diferencia de otros enfoques donde cada agente opera de forma independiente, aquí se establece una arquitectura que permite el intercambio constante de información, orientado a generar un mapa global más completo y preciso del entorno.

La importancia de esta propuesta radica en su aplicabilidad directa en contextos reales como búsqueda y rescate, monitoreo de áreas inaccesibles o inspección de terrenos complejos, donde la combinación de movilidad aérea y análisis terrestre ofrece ventajas significativas en cobertura, adaptabilidad y eficiencia. Además, permite optimizar recursos al aprovechar las fortalezas de cada plataforma: el UAV para obtener vista panorámica y detección rápida, y el UGV para explorar con mayor detalle zonas cercanas al suelo.

Los robots UAV y UGV pueden aprovechar sus capacidades complementarias, con los UAV utilizando su movilidad rápida y los UGV permitiendo una interacción cercana con el entorno. Varios casos de integración de equipos con UAVs y UGVs se incluyen en la revisión de literatura de (Munasinghe *et al.*, 2024), donde también se consideran cuáles limitaciones importantes persisten en estas colaboraciones: restricciones en poder de cómputo, inestabilidad en redes de comunicación y adaptabilidad al entorno.

Cuando varios robots heterogéneos colaboran para explorar un entorno, pueden utilizar estrategias como la creación de mapas y localización simultáneas (SLAM, por sus siglas en inglés) para representar el escenario. Trabajos como el de (Fox *et al.*, 2006) describen cómo se realiza el proceso de SLAM con múltiples robots usando optimización de grafos que representan relaciones entre posiciones y orientaciones. Otras colaboraciones entre robots heterogéneos también han explorado el uso de diversas modalidades de locomoción, como drones y humanoides (Chatziparaschis *et al.*, 2020). La fusión de información de múltiples sensores durante las tareas de SLAM multi-robot puede realizarse de diversas maneras, varias de las cuales se exploran en la revisión de literatura de (Chen *et al.*, 2023).

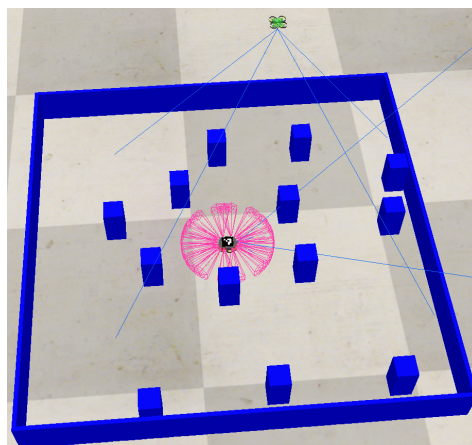


Figura 1. Escena global de la comunicación entre robots.

Existen antecedentes en trabajos que utilizan el sistema ROS (*Robot Operating System*, Sistema Operativo para Robots) y simuladores como CoppeliaSim para pruebas de coordinación multi-robot (Vigneron Olmos, 2025), sin embargo, son escasos los estudios que logren una integración efectiva de sensores visuales, estimación de poses mediante marcadores ArUco (Augmented Reality University of Cordoba), mapeo ultrasónico y fusión de mapas en un solo sistema funcional. Este proyecto retoma dichas bases y propone un enfoque unificado, aplicable a entornos simulados y transferible a escenarios reales. La solución que se propone en este trabajo combina un robot aéreo y un terrestre para que lleven a cabo la tarea de exploración de un escenario utilizando dos modalidades de censado de los obstáculos: una cámara monocular y una colección de sensores ultrasónicos de distancia.

Metodología

La creación de mapas de entornos desconocidos requiere que los robots móviles tengan acceso o conocimiento de su localización. Cuando la localización se lleva a cabo al mismo tiempo que la creación del mapa, se considera que se realiza SLAM (*simultaneous localization and mapping*). Consideremos por ahora que el problema de la localización se ha resuelto, por ejemplo, con el uso de un sistema de localización absoluta tal como los sistemas de navegación global por satélite (GNSS, por sus siglas en inglés) tales como GPS, Glonass, Galileo y BeiDou.

El problema por resolver es la creación de un mapa entre dos o más robots móviles heterogéneos, donde al menos uno sea terrestre y otro aéreo. Debido a que la plataforma terrestre es más estable, consume una menor cantidad de energía si permanece estático, además obtiene de manera más precisa su información de localización. En este proyecto, consideramos que el robot aéreo conoce su localización, pero basa su estrategia de exploración y control de movimiento en su posición relativa con respecto al robot terrestre, utilizando como referencia un marcador visual colocado en el robot terrestre que es captado por una cámara fija en el robot aéreo.

En este documento se representarán las conversiones entre marcos de referencia utilizando transformaciones homogéneas con la siguiente notación:

$$T_a^b = \begin{bmatrix} R_a^b & t_a^b \\ 0_{1 \times 3} & 1 \end{bmatrix}$$

donde R_a^b es la rotación entre dos marcos de referencia genéricos a y b , y t_a^b el vector de traslación entre los orígenes de a y b .

Los marcadores ArUco (Garrido-Jurado *et al.*, 2014), que están integrados a la biblioteca de código abierto OpenCV, son utilizados ampliamente en áreas de visión artificial, realidad aumentada, control de calidad, así como en la robótica, debido a que ofrecen una herramienta robusta, accesible y precisa para la detección, localización y estimación de pose en tiempo real. Se trata de un algoritmo de visión por computadora que utiliza patrones binarios 2D donde cada marcador consiste en un cuadrado con un borde negro ancho que facilita su detección, y un patrón interno de celdas blancas y negras que codifican un identificador único permitiendo distinguir entre diferentes marcadores dentro de una misma escena, incluso en condiciones desafiantes como iluminación variable, ruido visual o parcial oclusión.

Una de las principales ventajas de los marcadores ArUco es su capacidad para calcular con precisión la pose de un objeto en seis grados de libertad (posición en los ejes X, Y, Z y orientación en *yaw*, *pitch* y *roll*), lo que es especialmente útil para tareas de navegación autónoma, calibración de cámaras y seguimiento dinámico de objetos. En este trabajo, el marcador tiene una posición y orientación fijas con respecto al marco de referencia del robot terrestre, que se representa con la transformación de coordenadas T_m^r .

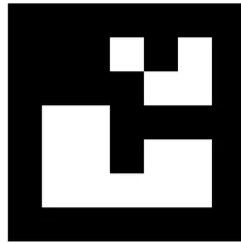


Figura 2. Ejemplo de marcador ArUco.

Una vez detectado el marcador ArUco desde la cámara aérea, se estima su posición relativa en el sistema de referencia de la cámara utilizando la rutina *solvePnP*, lo que permite obtener la transformación de coordenadas entre el marcador y la cámara. Llamemos a esta transformación T_c^m .

Fue implementado un controlador proporcional para guiar al dron como objetivo al marcador ArUco que fuese detectado en el entorno. Dado que la transformación T_c^a entre la cámara del dron y su cuerpo es constante y conocida, una vez que se ha obtenido la transformación entre el marcador y la cámara del dron T_c^m se puede obtener aquella entre el marcador y el cuerpo del dron, T_m^a . Utilizando esta información, y mediante el sistema de control, se calcula el error entre la posición actual del robot y la posición objetivo, aplicando una corrección proporcional a dicho error. La ganancia proporcional (k_p) determina la agresividad del movimiento: con valores altos produce respuestas rápidas, pero puede causar oscilaciones, mientras que con valores bajos garantiza estabilidad a costo de una convergencia lenta. Para tener una navegación adecuada se emplearon valores bajos tanto para el robot terrestre como en el aéreo, siendo en el caso de este último más baja para evitar oscilaciones. En el caso del robot terrestre, se empleó una navegación reactiva con evasión simple de obstáculos utilizando la información de los sensores ultrasónicos.

Una rejilla de ocupación, también llamado mapa de ocupación, es una representación discreta del entorno de un robot generado por sensores como láseres o cámaras que pueden detectar la distancia a los objetos del entorno. El espacio es dividido en celdas con un tamaño definido que contienen información sobre si está ocupada por un obstáculo, si no ha sido explorada o si está libre. El mapa se actualiza conforme el robot explora su entorno y obtiene nueva información desde sus sensores. La transformación de coordenadas entre el marco de referencia global y el mapa de rejilla de ocupación se representa como T_w^g .

Ahora bien, para poder fusionar la visión aérea con el mapa generado por el robot terrestre, es necesario traducir la imagen de la cámara aérea a un formato entendible para el sistema de mapeo, es decir, convertirla a una rejilla de celdas. Este proceso inicia con una segmentación de color basada en el espacio RGB (rojo-verde-azul, por sus siglas en inglés) que filtra zonas específicas del entorno, como marcadores o caminos, que aparecen en la imagen, conservando únicamente aquellas que se encuentre dentro del rango de color seleccionado.

La imagen resultante es binaria y rotada, para estar alineada con la perspectiva del mapa terrestre. Luego se convierte esta imagen aérea en un “submapa” que será fusionado con la cuadrícula global. Se realiza una conversión: se redimensiona la imagen aérea para que cada píxel corresponda a una celda en la cuadrícula de ocupación (suponiendo que cada celda representa 10cm). Esto se calcula considerando la altura del dron para estimar la resolución.

A continuación, se genera una máscara que simula la imagen aérea como si fuera una porción del mapa y se alinea con la cuadrícula global. Finalmente, esta se traduce en el mapa global según la posición del dron. Para traducir los píxeles de la imagen umbralizada y re-escalada del dron en celdas de la rejilla de ocupación e incorporar esta información en el mapa conjunto, se requiere transformar puntos en coordenadas de la cámara a puntos en el marco de referencia del mapa de rejilla. La transformación buscada se calcula a través de la siguiente operación:

$$T_c^g = T_w^g \cdot T_r^w \cdot T_m^r \cdot T_c^m$$

Con esto, cada imagen capturada por la cámara aérea no solo se interpreta como un conjunto de píxeles, sino como un bloque espacial con información útil para complementar el mapa global, esta fusión de datos aéreos y terrestres permite que ambos agentes trabajen como una única entidad exploradora, colaborando para generar un modelo ambiental más completo y coherente.

Pruebas y resultados

Se implementaron los enfoques de mapeo:

1. Mapeo únicamente con el robot terrestre Pioneer 3 DX
2. Mapeo dron estático sobrevolando el robot terrestre a 12 m y 20 m de altura
3. Mapeo dron orbitando al robot terrestre a 12 m y 20 m de altura

En los tres enfoques, el robot terrestre ejecutó una navegación reactiva de tipo vehículo de Braitenberg basada en los sensores ultrasónicos, evitando colisiones durante el recorrido.

El experimento se desarrolló en un escenario de entorno de interior con un muro perimetral de 10x10 m², trece obstáculos cúbicos distribuidos aleatoriamente en el escenario. Se integraron el robot Pioneer 3 DX y un dron como agentes de exploración. El Pioneer estaba equipado con 16 sensores ultrasónicos con un alcance máximo de 1 m, utilizados para la generación del mapa de ocupación. Además, se le colocó un marcador ArUco de 30 cm por lado en su parte superior.

El dron contaba con una cámara a bordo con modelo estenopeico sin distorsión radial, resolución de 640 × 480 píxeles y una longitud focal de 800 píxeles. La cámara se orientó en dirección nadir, observando verticalmente el escenario para la detección del marcador y realizar el proceso de mapeo.

Cada enfoque se desarrolló durante 120 s tomando como métricas el **Porcentaje de cobertura del mapa y Cobertura** por cada enfoque de mapeo.

La simulación fue implementada en el simulador CoppeliaSim versión 4.10, programado en Python 3.12 y la comunicación a través del protocolo ZeroMQ. El escenario puede apreciarse en la Figura 3.

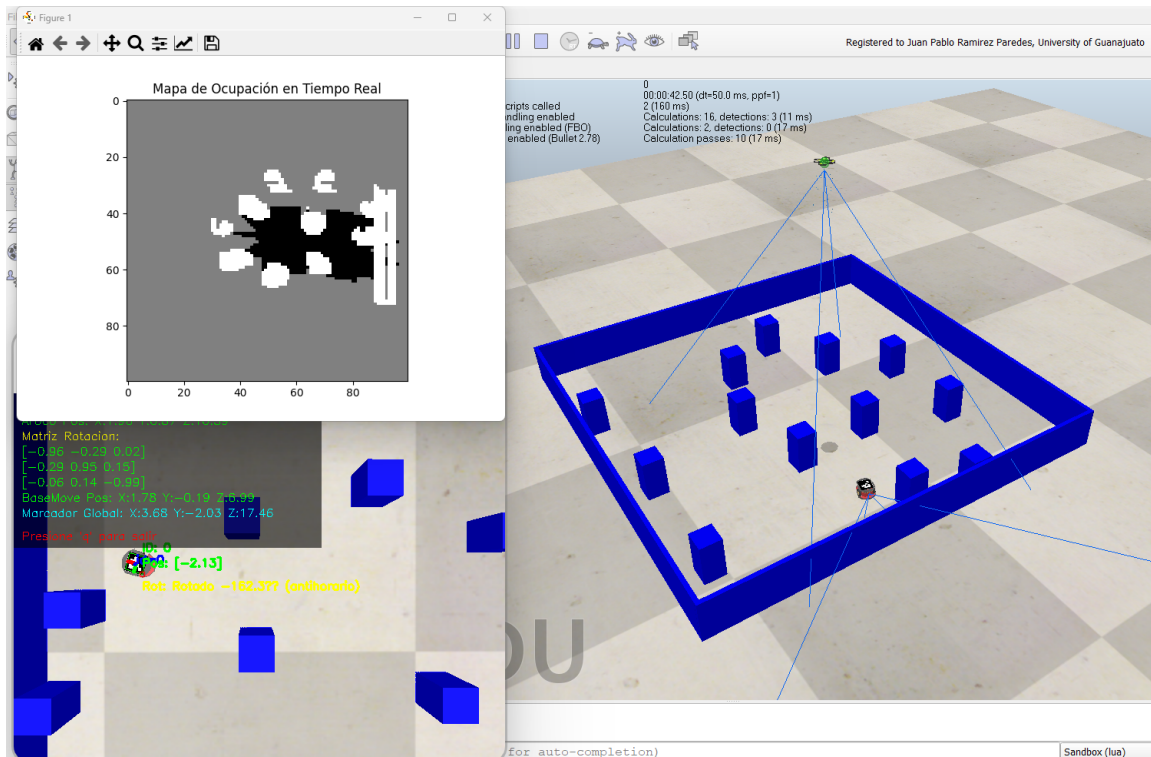


Figura 3. Escenario de simulación con robots terrestre y aéreo creando un mapa de los obstáculos.

La Figura 4 muestra la vista de la cámara simulada en el dron, mostrando el marcador ArUco detectado.



Figura 4. Vista desde la cámara aérea hacia el robot terrestre, detectando el marcador visual.

Mapeo robot terrestre Pioneer 3 DX

Se llevó a cabo una exploración del entorno únicamente con el robot terrestre. La exploración se llevó a cabo con una estrategia reactiva por parte del robot terrestre en todas las pruebas descritas en este trabajo, inspirado en los vehículos de Braitenberg. La Figura 5 se muestra el mapa de ocupación de esta exploración.

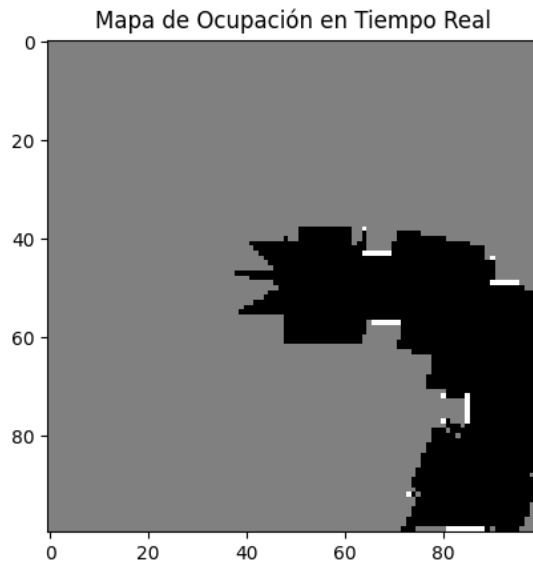


Figura 5. Mapa después de 120 segundos, usando solo los sensores del robot terrestre.

Mapeo dron sobrevolando el robot terrestre

En la Figura 6 se muestra un ejemplo de exploración que integra la información del robot terrestre y la del aéreo; la forma aproximada de los obstáculos se ve en tono gris para reflejar que la ocupación de esas celdas en la rejilla es menos confiable, mientras que las celdas que aparecen en color blanco reflejan una mayor confianza en las mediciones de los sensores ultrasónicos del robot terrestre.

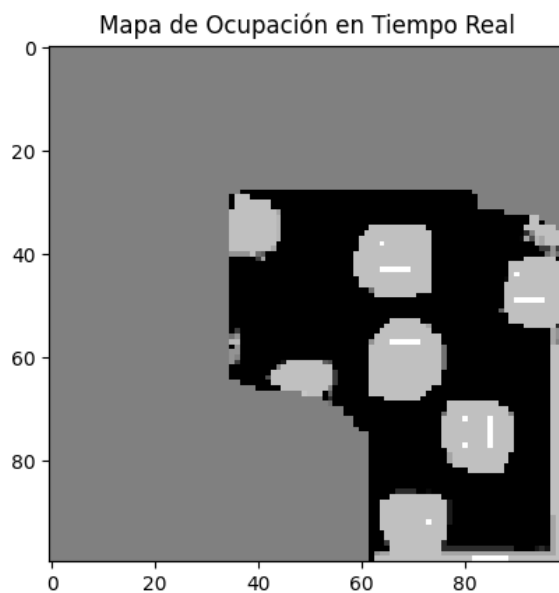


Figura 6. Mapa después de 120 segundos, con el dron sobrevolando a 12 m del robot terrestre

Los resultados se muestran en la Tabla 1.

Tabla 1. Resultados de las pruebas de cobertura.

Estrategia de Mapeo	Porcentaje de cobertura del mapa	Cobertura (celdas por segundo)
Robot terrestre	21.23 %	17.68
Robot aéreo sobre el terrestre (a 12m)	36.76 %	30.62
Robot aéreo orbitando al terrestre (a 12 m)	37.04 %	30.85
Robot aéreo sobre el terrestre (a 20 m)	56.78 %	47.30
Robot aéreo orbitando al terrestre (a 20 m)	58.29 %	48.55

Los resultados respaldan la noción de que, incluyendo información de sensores con mayor cobertura, se consigue mayor cantidad de celdas exploradas en un tiempo equivalente. Adicionalmente, incrementar la altura de sobrevuelo incrementa también el área cubierta por la cámara del robot aéreo y, en consecuencia, el porcentaje de celdas cubiertas en los 120 s. Sin embargo, esta simulación no considera la pérdida de resolución espacial que involucra el sobrevuelo a mayor altura con la misma cámara.

Conclusiones

Este trabajo demuestra que un sistema de colaboración heterogénea con un robot terrestre y un dron, que fusionan la información obtenida del entorno para crear el mapa de ocupación, logran un mayor porcentaje de mapeo del entorno. Los resultados de simulaciones de tareas de exploración muestran que agregar la información de la cámara del robot aéreo, el cual orbita al robot terrestre a las lecturas de los sensores ultrasónicos de este último, incrementa significativamente la cantidad de celdas cubiertas en el mapa de rejilla de ocupación. También se consideró cambiar la altura de sobrevuelo y la forma en que el robot aéreo se mantiene volando cerca del robot terrestre, y se mostró que es posible incrementar la cobertura de celdas de esta manera.

Bibliografía/Referencias

- Chatziparaschis, D., Lagoudakis, M. G., & Partsinevelos, P. (2020). Aerial and Ground Robot Collaboration for Autonomous Mapping in Search and Rescue Missions. *Drones*, 4(4), 4. <https://doi.org/10.3390/drones4040079>
- Chen, W., Wang, X., Gao, S., Shang, G., Zhou, C., Li, Z., Xu, C., & Hu, K. (2023). Overview of Multi-Robot Collaborative SLAM from the Perspective of Data Fusion. *Machines*, 11(6), 6. <https://doi.org/10.3390/machines11060653>
- Fox, D., Ko, J., Konolige, K., Limketkai, B., Schulz, D., & Stewart, B. (2006). Distributed Multirobot Exploration and Mapping. *Proceedings of the IEEE*, 94(7), 1325–1339. <https://doi.org/10.1109/JPROC.2006.876927>
- Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Marín-Jiménez, M. J. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6), 2280–2292. <https://doi.org/10.1016/j.patcog.2014.01.005>
- Munasinghe, I., Perera, A., & Deo, R. C. (2024). A Comprehensive Review of UAV-UGV Collaboration: Advancements and Challenges. *Journal of Sensor and Actuator Networks*, 13(6), 6. <https://doi.org/10.3390/jsan13060081>
- Vigneron Olmos, D. (2025, January). *Diseño e implementación de un protocolo de comunicación para entornos multi-robot* [Info:eu-repo/semantics/bachelorThesis]. E.T.S. de Ingenieros Informáticos (UPM). <https://oa.upm.es/87832/>