

Jornadas de Automática

Comparativa de SLAM RGB-D para manipulación robótica en entornos reducidos

Quim Serrats^a, Edmundo Guerra^{a,*}, Yolanda Bolea^a, Antoni Grau^a

^aDepartamento de Ingeniería de Sistemas, Automática e Informática Industrial (ESAI), Universitat Politècnica de Catalunya, Av. Diagonal 647, 08028 Barcelona, España.

To cite this article: Serrats, Q., Guerra, E., Bolea, Y., Grau, A. 2026. Comparison of RGB-D SLAM for robotic manipulation in reduced environments. Jornadas de Automática, 47. <https://doi.org/10.17979/ja-cea.2026.47.13865>

Resumen

Este trabajo compara experimentalmente tres sistemas SLAM —ORB-SLAM3, RTAB-Map y Cartographer— sobre datos RGB-D en un escenario de manipulación robótica a corta distancia. Las secuencias se capturaron con dos cámaras Intel RealSense, la D405 y la D435i, montadas sobre un brazo robótico UR3e, lo que permite ejecutar trayectorias repetibles y reducir la variabilidad propia de las capturas manuales. Se diseñaron cinco patrones de trayectoria y tres escenas de complejidad geométrica creciente. El procesamiento se realizó *offline*, homogeneizando las salidas para permitir una comparación común en términos de trayectoria estimada, mapa 3D, mapa 2D, interpretabilidad visual y coste computacional. Los resultados muestran que no existe un sistema universalmente superior: ORB-SLAM3 con la D405 ofrece el mejor compromiso para escenas próximas, RTAB-Map mantiene la mayor robustez frente al cambio de sensor, y Cartographer destaca en mapas 2D de ocupación. La información inercial aporta mejoras puntuales pero no generalizables.

Palabras clave: Robótica, SLAM, Sensores RGB-D, Manipulación robótica, Fusión de sensores, Visión por computador

Comparison of RGB-D -SLAM for robotic manipulation in reduced environments

Abstract

This paper experimentally compares three SLAM systems —ORB-SLAM3, RTAB-Map and Cartographer— on RGB-D data in a short-range robotic manipulation setting. Sequences were captured with two Intel RealSense cameras, the D405 and the D435i, mounted on a UR3e robotic arm, which enables repeatable trajectories and reduces the variability typical of hand-held captures. Five trajectory patterns and three scenes of increasing geometric complexity were designed. Processing was performed *offline*, homogenising the outputs to allow a common comparison in terms of estimated trajectory, 3D map, 2D map, visual interpretability and computational cost. Results show no universally superior system: ORB-SLAM3 with the D405 offers the best compromise for close scenes, RTAB-Map exhibits the highest robustness across sensors, and Cartographer stands out for 2D occupancy mapping. Inertial information yields situational improvements but no general benefit.

Keywords: Robotics, SLAM, RGB-D sensors, Robotic manipulation, Sensor fusion, Computer vision

1. Introducción

La localización y mapeado simultáneos (SLAM) es un problema central en robótica autónoma: permite que un robot estime su trayectoria mientras construye un mapa del entorno sin referencias externas (Durrant-Whyte and Bailey, 2006;

Bailey and Durrant-Whyte, 2006). Tras dos décadas de desarrollo, el problema se considera maduro en sus fundamentos, aunque persisten retos en entornos complejos, dinámicos y de operación prolongada (Cadena et al., 2016).

Existen familias de sistemas SLAM con filosofías de di-

*Autor para correspondencia: edmundo.guerra@upc.edu
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

seño distintas. Los métodos visuales basados en características estiman la pose a partir de puntos clave y producen mapas dispersos (Mur-Artal et al., 2015); los métodos directos operan sobre las intensidades de la imagen (Engel et al., 2014); las aproximaciones densas basadas en cámaras RGB-D reconstruyen superficies completas del entorno (Newcombe et al., 2011); y las formulaciones sobre grafo de poses optimizan globalmente la trayectoria y el mapa (Grisetti et al., 2010). La aparición de cámaras RGB-D de bajo coste ha facilitado especialmente el uso de SLAM visual denso en aplicaciones industriales y de servicio.

Sin embargo, la mayoría de comparativas experimentales publicadas se centran en entornos a escala humana o vehicular. La manipulación robótica a corta distancia —volumenes de trabajo del orden de 0,5–1 m, trayectorias condicionadas por la cinemática del brazo y escenas estructuradas— está menos representada en la literatura, pese a su interés práctico. La elección de la combinación cámara/sistema SLAM en este contexto no es evidente: los resultados publicados rara vez se obtienen sobre el mismo conjunto de datos y bajo condiciones de movimiento controladas.

Este trabajo aborda esa carencia. Las aportaciones son: (i) una metodología experimental para capturar secuencias RGB-D repetibles con un brazo UR3e y dos cámaras Intel RealSense (D405 y D435i); (ii) la comparación de ORB-SLAM3, RTAB-Map y Cartographer sobre las mismas trayectorias y escenas; y (iii) una síntesis por criterio —trayectoria estimada, mapa 3D, mapa 2D, interpretabilidad y coste computacional— que orienta la elección del sistema en función del requerimiento dominante.

El resto del artículo se organiza como sigue. La Sección 2 presenta los tres sistemas SLAM comparados. La Sección 3 describe la metodología experimental. La Sección 4 expone los resultados y la discusión. La Sección 5 recoge las conclusiones.

2. Sistemas SLAM comparados

Se han seleccionado tres sistemas SLAM representativos de filosofías distintas, todos compatibles con datos RGB-D y disponibles como software abierto. La Tabla 1 resume sus diferencias principales en el contexto de este trabajo.

ORB-SLAM3 (Campos et al., 2021) es un sistema visual basado en características, evolución de la familia ORB-SLAM (??), cuya precisión de localización en condiciones favorables se considera de referencia en SLAM visual. Su salida nativa es un mapa disperso de puntos característicos, y su rendimiento depende fuertemente de la textura disponible y de la coherencia geométrica de la profundidad, factor especialmente relevante en rango corto.

RTAB-Map (Labbé and Michaud, 2019) es un sistema sobre grafo de poses orientado al mapeado denso y a misiones largas. A partir del grafo y de la profundidad asociada a cada nodo genera nubes de puntos densas y mapas 2D de ocupación derivados, lo que lo convierte en el más versátil de los tres en cuanto a salidas cartográficas, a costa de una carga computacional mayor.

Cartographer (Hess et al., 2016), desarrollado originalmente para LiDAR, se utiliza aquí en su formulación 2D,

alimentado con escaneos planares derivados de la profundidad de la cámara RGB-D. Su interés en la comparación reside precisamente en su modelo de representación —planar geométrico— frente al visual basado en características y al visual sobre grafo de los otros dos sistemas.

La Tabla 1 resume las diferencias principales entre los tres sistemas en el contexto de este trabajo.

3. Metodología experimental

3.1. Equipamiento y montaje

Los experimentos se han realizado en un laboratorio estructurado. El equipamiento principal consta de un brazo robótico colaborativo Universal Robots UR3e (6 grados de libertad, alcance 500 mm) y dos cámaras Intel RealSense: la D405, orientada a rango corto (aproximadamente 7–50 cm), y la D435i, de rango medio con unidad de medida inercial integrada. En este trabajo ambas cámaras operan a distancia próxima, aproximadamente entre 30 y 50 cm, intervalo en el que las dos pueden registrar datos RGB-D pero no necesariamente con la misma calidad de profundidad: la D405 está especialmente adaptada a este rango, mientras que la D435i presenta más limitaciones al trabajar tan cerca de la escena. Las cámaras se montan alternativamente sobre el efector final mediante un soporte ajustable (Figura 1). El brazo ejecuta las trayectorias mediante movimientos *MoveP*, que permiten recorridos suaves e interpolados, repetibles entre series y entre sensores. La pose teórica del efector final se registra en paralelo a partir de la cinemática directa del UR3e y se utiliza como referencia geométrica.

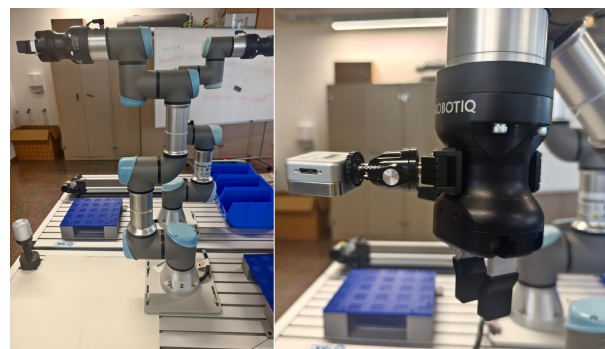


Figura 1: Montaje experimental: brazo UR3e y detalle del soporte de la cámara D405 sobre el efector final.

3.2. Escenas y objetos

Se han definido tres objetos con niveles crecientes de complejidad visual y geométrica: un cubo (geometría regular y textura uniforme), un conjunto de herramientas de laboratorio (geometrías heterogéneas y bordes definidos) y una planta artificial (geometría irregular y textura distribuida). Los tres objetos se sitúan sobre una zona de trabajo plana con elementos auxiliares estáticos que enriquecen el contexto visual. La escena se mantiene estructurada y de dimensiones acotadas, coherente con un caso de uso de manipulación a corta distancia.

Tabla 1: Comparación general entre los sistemas SLAM estudiados.

Aspecto	ORB-SLAM3	RTAB-Map	Cartographer
Filosofía	Visual, basado en características	Grafo de poses, mapeado denso	Geométrico planar
Entrada principal	RGB-D, mono, estéreo, visual-inercial	RGB-D, estéreo, LiDAR	LiDAR / escaneos 2D
Mapa nativo	Puntos dispersos	Nube densa + 2D derivado	Mapa 2D de ocupación
Cierre de lazo	Visual (BoW)	Apariencia visual	Geométrico (correlación)
Salida 3D	Sí (dispersa)	Sí (densa)	No (en este uso)
Salida 2D	No	Derivada	Nativa

3.3. Trayectorias experimentales

Se han diseñado cinco trayectorias (T1–T5) que cubren patrones de movimiento distintos (Figura 2):

- **T1** — órbita circular cerrada alrededor del objeto (longitud de referencia 0,97 m); caso base.
- **T2** — dobles órbitas a dos alturas con transición (2,14 m).
- **T3** — figura de infinito con doble lazo, en un plano horizontal (0,59 m).
- **T4** — recorrido rectangular con esquinas pronunciadas, con la cámara apuntando predominantemente hacia abajo (0,36 m, la más corta).
- **T5** — dobles órbitas con pausas intermedias, que retoma la geometría de T2 (2,14 m).

La longitud de referencia de cada trayectoria se obtiene de la cinemática directa del UR3e. El conjunto incorpora cierres de lazo, cambios de altura, geometrías no circulares y pausas, con el objetivo de excitar distintos modos de operación de los sistemas SLAM.

3.4. Captura y organización de los datasets

Cada combinación de trayectoria y objeto se ha registrado en formato *ROS bag* con los *topics* necesarios para cada sistema: imagen de color, imagen de profundidad, parámetros intrínsecos y, para la D435i, datos inerciales sincronizados. Los *datasets* se etiquetan de forma uniforme por trayectoria–sensor–objeto, lo que permite un procesamiento sistemático y reproducible.

3.5. Configuración y procesamiento

El procesamiento se ha realizado *offline* para garantizar la repetibilidad y desacoplar la comparación de las restricciones de tiempo real. Cada sistema se configura con los parámetros recomendados para RGB-D y se ajusta mínimamente para alinear las salidas: trayectoria estimada, nube de puntos 3D cuando procede, mapa 2D de ocupación cuando procede y registros temporales para estimar el coste computacional. La D435i se procesa tanto sin IMU como con IMU, esta última únicamente en RTAB-Map y Cartographer; la configuración visual-inercial de ORB-SLAM3 requería una excitación inercial superior a la disponible con los movimientos del UR3e y se ha excluido de la comparación principal.

Tabla 2: Configuración de los tres sistemas de SLAM.

Parámetro	ORB-SLAM3	RTAB-Map	Cartographer.
Modo	RGB-D	RGB-D (+IMU opc.)	2D (nubes)
Resol./frec.		640×480, 30 Hz	
Alin. prof.-color		Sí (píxel a píxel)	
Escala prof.		Fijada a metros	
Odometría	Visual (ORB)	RGB-D frame-to-frame	Correlativa 2D
Caract. ORB	1250 (D405); mayor (D435i)	—	—
Umbral FAST	Relaj. (D405); bajo (D435i)	—	—
Filtrado	—	Prof. máx., radio local	Plano de mesa
Cierre lazo	Visual (BoW)	Apariencia visual	Geom. (corr.)
Grafo global	Activo	Activo	Reduc./desact.
IMU	Excluida	Extrínsc. + inicializ.	Soporte (2D)
Reprod. offline	1,0×	0,5× (parcial)	1,0×

3.6. Criterios de evaluación

La comparación se articula sobre los siguientes criterios complementarios: *trayectoria estimada* (error RMS entre la trayectoria estimada y la referencia cinemática del UR3e, siguiendo prácticas de evaluación habituales en SLAM RGB-D (Sturm et al., 2012)); *mapa 3D* (evaluación cualitativa de la nube de puntos según coherencia geométrica global, densidad, presencia de ruido, continuidad de las estructuras e identificabilidad de los objetos, incorporando la interpretabilidad visual del color cuando es determinante para reconocer la forma); *mapa 2D* (evaluación cualitativa del mapa de ocupación según capacidad de diferenciar las estructuras de la escena, coherencia geométrica —ausencia de duplicaciones o deformaciones—, cobertura de la zona explorada, nitidez de los contornos y presencia de artefactos visuales); y *coste computacional* (tiempo equivalente por fotograma y carga relativa respecto al tiempo real).

Las trayectorias estimadas se normalizan respecto al origen y se alinean geoméricamente con la referencia antes de calcular el RMS, que se interpreta así como coherencia geométrica relativa y no como posición absoluta. En la D405, las trayectorias exportadas de ORB-SLAM3 y RTAB-Map incorporan una corrección de escala que compensa una discrepancia en la interpretación de la profundidad durante la exportación, sin alterar el comportamiento interno de los sistemas: solo las expresa en una escala coherente con la referencia cinemática y el resto de configuraciones. La misma discrepancia se compensó en los mapas 2D derivados mediante factores explícitos: 0,01 en las proyecciones de ORB-SLAM3 y RTAB-Map con D405 y 0,1 en las demás configuraciones afectadas.

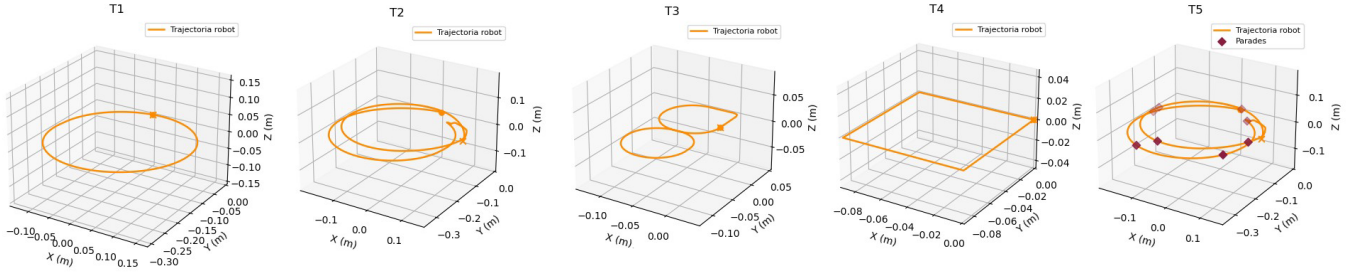


Figura 2: Las cinco trayectorias T1–T5 ejecutadas por el efector final del UR3e.

Tabla 3: Error RMS de la trayectoria estimada (mm) por trayectoria y configuración. Rangos por set de experimentos, min a max.

Trayectoria	Long. ref. (m)	D405			D435i		
		ORB	RTAB	Cartog.	ORB	RTAB	Cartog.
T1	0,97	40 a 41	12 a 60	22 a 103	24 a 100	14 a 18	39 a 222
T2	2,14	118 a 131	109 a 136	135 a 172	164 a 184	90 a 142	181 a 225
T3	0,59	62 a 70	65 a 68	55 a 65	49 a 57	65 a 72	71 a 77
T4	0,36	4 a 9	9 a 14	5 a 18	21 a 40	7 a 14	17 a 18
T5	2,14	170 a 189	174 a 195	137 a 150	155 a 181	142 a 208	157 a 282

4. Resultados y discusión

4.1. Trayectoria estimada

La Tabla 3 recoge los errores RMS por trayectoria y configuración; los rangos y casos que se discuten a continuación se apoyan en estos valores.

Con la cámara D405, ORB-SLAM3 obtiene los errores RMS más bajos en prácticamente todas las trayectorias, situándose de forma consistente entre 30 y 70 mm. Esta superioridad se atribuye a un seguimiento visual estable, basado en características fiables y una profundidad bien alineada con la imagen de color en el rango corto. RTAB-Map ocupa la segunda posición con la D405, con errores típicamente entre 50 y 100 mm. Cartographer presenta los mayores errores con D405 en general, pero destaca en T3, donde las revisitas repetidas activan el cierre de lazo geométrico y producen errores comparables a ORB-SLAM3.

Con la D435i sin IMU el orden se invierte. RTAB-Map mantiene errores moderados (70–150 mm) gracias a su odometría robusta, mientras que ORB-SLAM3 se degrada drásticamente hasta errores superiores al metro en varias configuraciones. Cartographer mantiene un comportamiento más estable que ORB-SLAM3 pero menos preciso que RTAB-Map.

La Figura 3 detalla este comportamiento sobre la órbita circular T1 para los tres sistemas y los dos sensores. Con la D405, ORB-SLAM3 cierra el círculo de forma limpia: el cierre de lazo visual se activa al completar la vuelta y redistribuye el error acumulado. RTAB-Map traza un círculo de longitud ligeramente sobreestimada, efecto de su selección de fotogramas clave por novedad visual, que reduce el número de nodos a costa de cierta deriva de escala. Cartographer es el que más se desvía: al carecer de reconocimiento visual de lugares, en una escena próxima y poco discriminativa geoméricamente su cierre de lazo no llega a activarse y el error crece sin corrección. Al pasar a la D435i el efecto es asimétrico.

Cartographer es el que más se desvía: al carecer de reconocimiento visual de lugares, en una escena próxima y poco discriminativa geoméricamente su cierre de lazo no llega a activarse y el error crece sin corrección. Al pasar a la D435i el efecto es asimétrico.

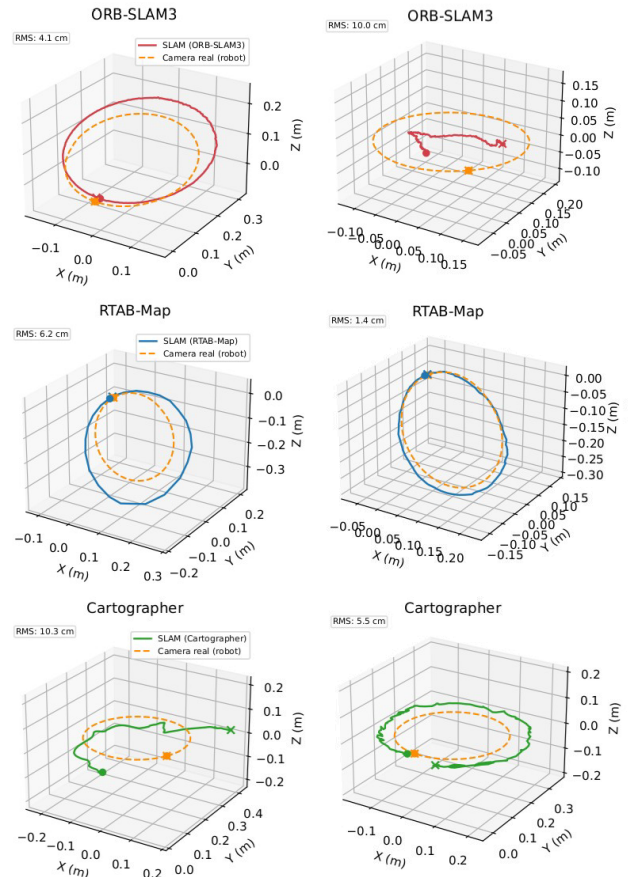


Figura 3: Trayectorias estimadas por ORB-SLAM3, RTAB-Map y Cartographer en la órbita circular T1, frente a la referencia: con la D405 y con la D435i.

ORB-SLAM3 se vuelve irregular sobre la superficie plana de la planta vista desde arriba, donde apenas hay características ORB estables entre fotogramas, lo que debilita el seguimiento e impide el cierre de lazo. RTAB-Map se mantiene

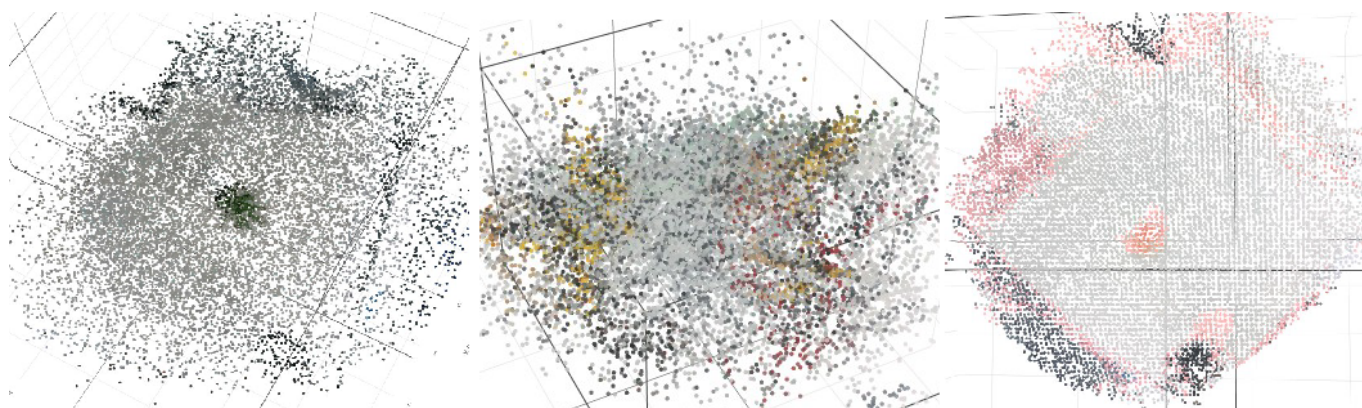


Figura 4: Comparación cualitativa del mapa 3D para una misma trayectoria: desde la izquierda ORB-SLAM3+D405, ORB-SLAM3+D435i y RTAB-Map +D435i.

robusto, beneficiado por la mayor resolución de color de la D435i.

Cartographer es el sistema cuyo comportamiento depende más de la escena al cambiar de sensor: en escenas geométricamente pobres el ruido de profundidad adicional de la D435i en rango corto produce traslaciones espurias —hasta 2,72 m de recorrido estimado frente a 0,97 m de referencia en el caso del cubo—, aunque en la escena de la planta, con más textura, llega a seguir la órbita incluso mejor que con la D405.

La aportación de la IMU es asimétrica; la integración de información inercial puede reforzar la robustez del SLAM visual (Mur-Artal and Tardós, 2017), pero aquí su efecto depende del sistema. En RTAB-Map mejora especialmente T2, donde la información inercial limita la deriva entre las dos órbitas y durante la transición de altura. En T3 y T5 el efecto es marginal. En Cartographer puede ayudar puntualmente en T1 y T3, pero también puede introducir problemas si la inicialización inercial es deficiente.

4.2. Mapa 3D

ORB-SLAM3 con D405 produce los mapas 3D más ricos e interpretables del conjunto experimental (Figura 4). La información de color de los fotogramas clave se incorpora a la nube de puntos, lo que facilita reconocer visualmente los objetos. La principal limitación es la confusión sistemática de las hojas de la planta con puntos del suelo, presente en los tres sistemas y no resuelta por ninguna configuración.

RTAB-Map ofrece los mapas 3D más consistentes entre sensores y trayectorias. La densidad de puntos es inferior a la de ORB-SLAM3 por el mecanismo conservador de selección de fotogramas clave, y se observa una contaminación cromática persistente de tonos rojizos relacionada con el empaquetado del color en los ficheros exportados; este artefacto no es propio del SLAM, pero afecta a la utilidad visual del mapa. Cartographer queda fuera de este criterio por haberse utilizado en formulación 2D.

4.3. Mapa 2D

Este criterio es el más dependiente de la aplicación porque los mapas 2D de los tres sistemas son conceptualmente distintos. Cartographer con D405 en T3 ofrece la nitidez de contornos más alta de toda la comparación (Figura 5 derecha).

El mapa de ocupación nativo en escala de grises delimita los obstáculos con una precisión que ningún otro sistema iguala en las mismas condiciones. Sus principales limitaciones son artefactos característicos: rayos radiales en forma de V cuando el alineamiento geométrico diverge con la D435i, y mapas duplicados puntuales en T2 y T5 cuando se acumulan errores de traslación.

ORB-SLAM3 con D405 produce mapas 2D en color que combinan geometría coherente e información textural; los objetos se reconocen con facilidad, pero con la D435i el mapa se degradó en bloques desconectados que lo hacen inservible. RTAB-Map produce mapas 2D derivados de su nube densa, más consistentes entre sensores pero menos nítidos que los nativos de Cartographer.

4.4. Coste computacional

ORB-SLAM3 es el sistema más eficiente cuando el sensor opera en condiciones favorables: con D405 el coste equivalente se sitúa en 37–41 ms por fotograma y la carga sobre el tiempo real entre el 67 y el 70 %. Es el único de los tres sistemas que puede operar holgadamente en tiempo real con margen para carga adicional. Sin embargo, con la D435i el coste se duplica o triplica y la carga supera el 100 % en la mayoría de trayectorias.

Cartographer presenta el comportamiento más estable entre sensores: el coste equivalente se mantiene en 113–138 ms y la carga sobre el tiempo real siempre por debajo del 100 % (67–86 %) independientemente del sensor utilizado. Es la opción más predecible computacionalmente, a costa de una frecuencia de salida menor (5–7 Hz).

RTAB-Map presenta el mayor coste absoluto: con D405 la carga se sitúa alrededor del 140 %, y con D435i supera el 300 %. La selección conservadora de fotogramas clave reduce el coste del nodo principal, pero no compensa el coste de la odometría RGB-D, que procesa todos los fotogramas de entrada. Este sobre coste es el precio de un flujo de procesamiento más completo (odometría, cierre de lazo, optimización del grafo, nube densa y mapa 2D).

4.5. Síntesis comparativa

La Tabla 4 resume el comportamiento por criterio: ninguno destaca de forma uniforme y la elección depende del cri-

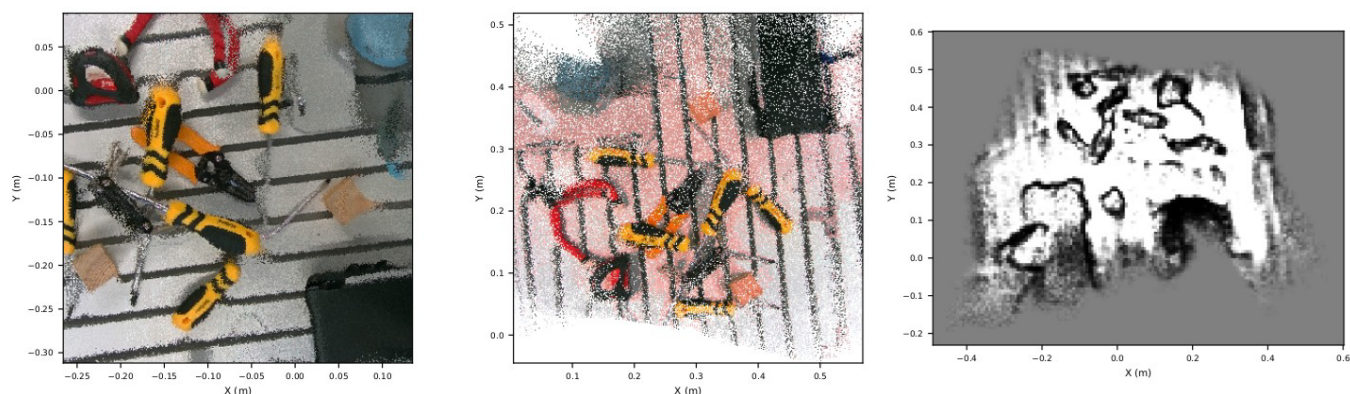


Figura 5: Comparación de Mapa 2D con D405 en secuencias T3 con giro, de izquierda a derecha: ORB-SLAM3, RTAB y Cartographer.

terio dominante y del sensor disponible. Para escenas próximas con la D405, ORB-SLAM3 ofrece el mejor compromiso en precisión de trayectoria, calidad de mapa 3D y coste; RTAB-Map es preferible cuando priman la robustez entre sensores o una salida cartográfica completa, a cambio de mayor coste; y Cartographer, cuando el objetivo es un mapa 2D de ocupación nítido y la trayectoria incluye revisitas geométricas suficientes.

Tabla 4: Síntesis comparativa por criterio. ●●● = mejor, ●● = intermedio, ● = menos favorable, — = no aplicable.

Criterio	ORB-SLAM3	RTAB-Map	Cartog.
Trayectoria (D405)	●●●	●●	●
Trayectoria (D435i)	●	●●●	●●
Mapa 3D	●●●	●●	—
Mapa 2D	●●	●●	●●●
Coste comp.	●●●	●	●●
Robustez sensores	●	●●●	●●

5. Conclusiones

Este trabajo propone una metodología experimental común para comparar sistemas SLAM con datos RGB-D en manipulación robótica a corta distancia, y la aplica a ORB-SLAM3, RTAB-Map y Cartographer con dos cámaras Intel RealSense montadas sobre un brazo UR3e, lo que permite ejecutar trayectorias repetibles y reducir la variabilidad propia de las capturas manuales. Ninguno de los tres sistemas resulta universalmente superior: la elección depende del sensor disponible, del tipo de escena y de la salida que se quiera priorizar (Tabla 4). La información inercial de la D435i no aporta mejoras generalizadas; su efecto depende del sistema, de la trayectoria y de la estabilidad de la inicialización, por lo que debe entenderse como un recurso situacional y no como una mejora automática. Las conclusiones son específicas del escenario evaluado —laboratorio estructurado, escenas próximas y trayectorias predefinidas— y no constituyen una validación general. Como continuación se proponen experimentos en escenarios más abiertos, con objetos dinámicos y trayectorias menos controladas, y una caracterización computacional con perfilado directo de CPU y memoria.

Referencias

- Bailey, T., Durrant-Whyte, H., 2006. Simultaneous localization and mapping (SLAM): part II. *IEEE Robotics & Automation Magazine* 13 (3), 108–117. DOI: 10.1109/MRA.2006.1678144
- Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J. J., 2016. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics* 32 (6), 1309–1332. DOI: 10.1109/TR0.2016.2624754
- Campos, C., Elvira, R., Gómez Rodríguez, J. J., Montiel, J. M. M., Tardós, J. D., 2021. ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM. *IEEE Transactions on Robotics* 37 (6), 1874–1890. DOI: 10.1109/TR0.2021.3075644
- Durrant-Whyte, H., Bailey, T., 2006. Simultaneous localization and mapping: part I. *IEEE Robotics & Automation Magazine* 13 (2), 99–110. DOI: 10.1109/MRA.2006.1638022
- Engel, J., Schöps, T., Cremers, D., 2014. LSD-SLAM: Large-scale direct monocular SLAM. In: *European Conference on Computer Vision (ECCV)*. Springer, pp. 834–849.
- Grisetti, G., Kümmerle, R., Stachniss, C., Burgard, W., 2010. A tutorial on graph-based SLAM. *IEEE Intelligent Transportation Systems Magazine* 2 (4), 31–43. DOI: 10.1109/MITS.2010.939925
- Hess, W., Kohler, D., Rapp, H., Andor, D., 2016. Real-time loop closure in 2D LIDAR SLAM. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 1271–1278. DOI: 10.1109/ICRA.2016.7487258
- Labbé, M., Michaud, F., 2019. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics* 36 (2), 416–446. DOI: 10.1002/rob.21831
- Mur-Artal, R., Montiel, J. M. M., Tardós, J. D., 2015. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics* 31 (5), 1147–1163. DOI: 10.1109/TR0.2015.2463671
- Mur-Artal, R., Tardós, J. D., 2017. Visual-inertial monocular SLAM with map reuse. *IEEE Robotics and Automation Letters* 2 (2), 796–803. DOI: 10.1109/LRA.2017.2653359
- Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J., Kohli, P., Shotton, J., Hodges, S., Fitzgibbon, A., 2011. KinectFusion: Real-time dense surface mapping and tracking. In: *2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. pp. 127–136. DOI: 10.1109/ISMAR.2011.6092378
- Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D., 2012. A benchmark for the evaluation of RGB-D SLAM systems. In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 573–580. DOI: 10.1109/IROS.2012.6385773