Universidad Nacional de San Juan
Facultad de Ingeniería
Instituto de Automática

Reconstrucción espacio-temporal de cultivos utilizando estrategias de sensado activo

Tesis Doctoral de
Sebastián Sansoni

Temas de Automática
111

Reconstrucción espacio-temporal de cultivos utilizando estrategias de sensado activo

Tesis doctoral de

Sebastian Sansoni

Aprobada por la
Facultad de Ingeniería de la Universidad Nacional de San Juan
para el otorgamiento del grado académico de

DOCTOR EN INGENIERÍA DE SISTEMAS DE CONTROL

Dirección de tesis: Dr. Ing. Santiago Ramón Tosetti Sanz
(Instituto de Automática, UNSJ)

Dr. Ing. Flavio Andrés Capraro Fuentes
(Instituto de Automática, UNSJ)

Dr. Lic. Javier Alejandro Gimenez Romero
(Instituto de Automática, UNSJ)

Jurado de tesis: Dr. Lic. Pablo De Cristoforis
(Instituto de Investigación en Ciencias de la Computación, UBA)

Dr. Ing. Juan Marcos Toibero
(Instituto de Automática, UNSJ)

San Juan, Argentina, 23 de Julio de 2025

Sansoni, Sebastian

Reconstrucción espacio-temporal de cultivos utilizando estrategias de sensado activo / Sansoni Sebastian. - 1.ª ed.

San Juan, Instituto de Automática, 2025.

Tesis (Doctoral) - Universidad Nacional de San Juan, 2025.

Formato	139 p. ; 21 x 15 cm.
ISBN	978-631-90962-2-4
Temas	1. Robótica. 2. Agricultura de Precisión. I. Título.
CDD	338.16



Repositorio Digital y Material Complementario

Escanee este código para acceder al repositorio en GitHub con el código fuente, los datos, el video de la defensa y material adicional de esta tesis.

https://github.com/Seba-san/thesis-phd-sansoni-2025



© 2025, Sebastian Sansoni
Hecho el depósito que marca la ley 11.723.
Impreso en Argentina.

Queda prohibida la reproducción total o parcial de esta tesis, su almacenamiento o transmisión por cualquier medio, ya sea electrónico, mecánico, fotocopiado o de cualquier otra índole, sin la autorización previa y por escrito del autor o de la institución a la que pertenece.

Índice general

Tabla de contenido III

Agradecimientos 1

Resumen 5

1 Introducción 7
- 1.1 Motivación . . . 7
- 1.2 Aportes y limitaciones del trabajo . . . 9
- 1.3 Difusión de los aportes . . . 10
  - 1.3.1 Revistas internacionales . . . 10
  - 1.3.2 Congresos Nacionales . . . 11
- 1.4 Organización de la Tesis . . . 12

2 Revisión Bibliográfica: Estado del arte y marco teórico 13
- 2.1 Introducción y Organización del Capítulo . . . 14
- 2.2 Simultaneous Localization and Mapping (SLAM) . . . 14
  - 2.2.1 Tipos de SLAM . . . 15
- 2.3 Reconstrucción 3D: Ejemplos de SLAM visual y con LiDAR . . . 18
  - 2.3.1 Introducción . . . 18
  - 2.3.2 ORB-SLAM . . . 18
  - 2.3.3 FAST-LIO . . . 22
- 2.4 Reconstrucción 3D: Structure From Motion . . . 27
  - 2.4.1 Introducción a SfM . . . 27
  - 2.4.2 COLMAP . . . 27
- 2.5 Reconstrucción 3D con Machine Learning . . . 31

2.5.1 Neural Radiance Fields (NeRF) . . . 32
2.6 Medición activa . . . 33
2.6.1 Active SfM o Next Best View Planning . . . 35
2.6.2 Active SLAM . . . 35
2.7 Reconstrucción 3D en la agricultura . . . 42
2.8 Conclusiones, oportunidades y desafíos . . . 50

3 Desarrollo Teórico 53
3.1 Organización del capítulo . . . 53
3.2 Modelado de la incertidumbre . . . 54
3.2.1 Probabilidades de dispersión e incertidumbre . . . 55
3.3 Proceso de Actualización . . . 61
3.3.1 Actualización mediante Media Ponderada . . . 62
3.3.2 Actualización mediante Filtro de Kalman . . . 65
3.4 Métrica de evaluación basada en Incertidumbre . . . 67
3.5 Fronteras de Incertidumbre: Objetivos de exploración . . . 72
3.6 Análisis del costo computacional . . . 73

4 Desarrollo experimental: Casos de estudio 75
4.1 Implementación en entornos de simulación . . . 78
4.1.1 Pruebas de las Fronteras de Incertidumbre . . . 79
4.1.2 Prueba del índice SiREn y de las FI . . . 81
4.1.3 Resultados y Discusión . . . 83
4.1.4 Conclusiones Preliminares de los Experimentos en Simulación 88
4.2 Implementación en entornos reales 3D . . . 90
4.2.1 Sistema de Localización 3D con Estimación de Incertidumbre para Entornos Agrícolas . . . 90
4.2.2 Estimación de la Posición Longitudinal mediante Landmarks Artificiales . . . 95
4.2.3 Metodología Experimental 3D . . . 101
4.2.4 Resultados y discusión . . . 104

5 Conclusiones y Trabajos Futuros 117
5.1 Síntesis del Trabajo . . . 117
5.2 Hallazgos Principales y Validación . . . 118
5.3 Contribuciones y Significado . . . 119
5.4 Limitaciones del Trabajo . . . 120
5.5 Trabajos Futuros . . . 121

5.6 Observaciones Finales 121

A Apéndice **123**

A.1 Apendice simulación 2D 123

A.1.1 Sistema de planificación de trayectorias 123

A.1.2 Sistema de SLAM 2D 125

A.2 Demostraciones y ejemplos numéricos 126

A.2.1 Desigualdad de Chebyshev 126

A.2.2 Prueba caso Gaussiano unidimensional 127

A.2.3 Ejemplo numérico para determinar β 127

A.2.4 Pruebas del Mapa de incertidumbre 129

A.2.5 Aproximación de D_{KL} 129

Agradecimientos

Creo que esta sección fue la más difícil de escribir de este trabajo. Principalmente porque no admite revisiones de pares.

Desde que tengo uso de razón, siempre pensé que la curiosidad es el motor por el cual hago las cosas, pero gracias al desarrollo de esta tesis doctoral y al doctorado en sí, me di cuenta que en mi caso, la motivación viene de la duda. Dudar me permite pensar en lo que asumimos conocido como si no lo fuera, como también, pensar en lo que imaginamos y dudar si eso ya no es parte de la realidad. En ese sentido, la duda me permitió desconfiar de lo que hoy asumimos como verdadero. Hay una frase extraída del GUM, que dice:

> "Una medida u observación solo es una aproximación de la realidad"

También hay otra frase sacada de "Microfísica del Poder" de Michel Foucault, que dice:

> "La verdad es aquella mentira que aún no se pudo desmentir"

Lo loco es que en la sección de agradecimientos de la tesis comienzo agradeciendo a la misma tesis. Este efecto autorreferencial genera paradojas (según Gödel), de las cuales no me voy a hacer cargo obviamente.

El Instituto de Automática es sin lugar a dudas el instituto de Robótica más grande e importante de Argentina. Con esto quiero remarcar la gran oportunidad que tuve de haber podido trabajar en este lugar. Hay gente trabajando en un montón de proyectos interesantes y de gran envergadura, lo cual para una persona curiosa como yo, es un paraíso. Por esto me gustaría agradecer inicialmente al Dr. Marcelo Moreyra por abrirme las puertas a este mundo, incentivarme a participar de las Jornadas Argentinas de Robótica del año 2017. Gracias a este congreso conocí al Dr. Ricardo Carelli, el cual posteriormente y gracias a Marcelo me invitó a trabajar

en el INAUT. Yo creo que sin esta serie de eventos no hubiera podido escribir esta tesis. Además, gracias al INAUT pude conocer a un montón de personas a nivel internacional, que de otra manera no hubiera sido posible.

Sin lugar a dudas, el Dr. Santiago Tosetti, mi director, es el responsable de que hoy pueda terminar mi doctorado. Cuando me postulé para el doctorado, no tenía ninguna referencia de este señor, era prácticamente un fantasma para mí. No sé si creer en el destino o no, pero claramente no era el director que yo quería pero sí el que necesitaba en esta etapa. Es un guía espiritual, un amigo, un compañero de trabajo que me ayudó a entender cómo funciona el mundo fuera y dentro del instituto. Es un excelente ser humano, con una paciencia infinita (testeada en reiteradas ocasiones). Por fuera de esto, me gustaría resaltar su extraordinaria capacidad de siempre exigirle a la vida un poco más de lo que le da. Aunque esto suene a que salió de un libro de autoayuda, a Santiago le funciona y es increíble verlo suceder.

El Dr. Flavio Capraro es el responsable en parte de que todo esto sea posible, gracias a su relación con el INTA y la generación del proyecto de unidad ejecutora (de donde salió mi beca). Siempre estuvo presente y atento.

El Dr. Javier Giménez, hoy uno de mis codirectores, además de sus aportes en lo académico, hizo como de salvavidas en momentos de crisis, tanto en la vida como en temas matemáticos. Ese salvavidas me animó a saltar más allá. En el aspecto personal, Javier es una referencia y, a la vez, un outlier (otra paradoja).

A mis compañeros de oficina, en particular a Estefanía y Humberto, les agradezco por escuchar mis teorías filosóficas y pido disculpas por los dolores de cabeza que les haya causado. Son y fueron piezas que ayudan a completar mi rompecabezas.

Moverse de ciudad no es fácil, siempre dije que me sentía y me siento un extranjero en San Juan. Mi permanencia se debe en gran parte a mis amigos, hoy separados por la distancia pero unidos electrónicamente: Cliver, Fran y Tania.

No puedo terminar esta sección de agradecimientos sin agradecerle a Ayelén, mi editora personal, mi compañera, mis otras dos patas de la mesa... si la defino la limito (frase robada).

Por último, me gustaría agradecer al lector de este texto. Ya sea por obligación, curiosidad o necesidad, te has tomado el tiempo para leer esto. Quizás no signifiquen nada estas palabras, pero para mí, en este momento mientras escribo esto, siento que concluí una etapa pero al mismo tiempo estoy abriendo otra (como un cierre de una prenda roto: por delante queda la parte que trato de cerrar y, por atrás, la que ya tendría que estar cerrada pero sigue abierta). El contenido del trabajo que estás a punto de leer es una idea loca dentro de un formato conocido para que sea digerible.

Como toda idea loca, no es una idea terminada, es un trabajo en proceso; sentite libre para enviarme un correo o contactarme por cualquier duda u observación.

Resumen

La gestión eficaz de la incertidumbre es crucial y desafiante en SLAM Activo (ASLAM), donde la toma de decisiones óptimas para la exploración y el mapeo autónomo es un problema NP-hard. Aplicaciones críticas como la robótica de servicios o la Búsqueda y Rescate (SAR) demandan mapas fiables, pero los enfoques tradicionales basados en mapas de ocupación y Fronteras Clásicas (FC) a menudo carecen de mecanismos robustos para representar, evaluar y explotar la incertidumbre del mapa y del agente.

Esta tesis introduce un marco teórico novedoso fundamentado en las Probabilidades de Dispersión (PD), una métrica probabilística ($PD \in [0, 1]$) flexible e interpretable propuesta para cuantificar la incertidumbre local de las estimaciones robóticas relativa a tolerancias definidas por el usuario. A partir de las PD, se construye un Mapa de Incertidumbre (MI) espacial, que a su vez permite definir dos herramientas originales: las Fronteras de Incertidumbre (FI) y el índice SiREn (Entropía Relativa con Signo). Las FI, detectadas por altos gradientes en el MI y condicionadas por un umbral de incertidumbre máxima tolerable ($\sigma_{\max}$), generalizan las FC para guiar no solo la exploración de lo desconocido, sino también la re-exploración de áreas mapeadas con alta incertidumbre, facilitando cierres de lazo. El índice SiREn, basado en la divergencia KL, proporciona una métrica global para evaluar cuantitativamente la calidad del mapa, balanceando cobertura y baja incertidumbre.

La viabilidad y utilidad del marco se validaron mediante simulaciones 2D, que demostraron la superioridad de planificadores basados en FI sobre los basados en FC en términos de calidad del mapa final (evaluado por SiREn), y mediante el análisis de datos reales 3D de un entorno agrícola complejo, utilizando un sistema SLAM adaptado que proporciona incertidumbre 6DOF (estimada vía PD). Este análisis 3D confirmó la capacidad de las FI para identificar regiones que requieren mejora y la

utilidad de SiREn para evaluar el impacto de la calidad de los datos y los criterios de incertidumbre ($\sigma_{\max}$) en la reconstrucción final.

En su totalidad, este trabajo aporta un conjunto coherente de herramientas para la toma de decisiones consciente de la incertidumbre en ASLAM, sentando las bases para sistemas de mapeo autónomo más eficientes, robustos y capaces de generar reconstrucciones con calidad controlada y evaluable.

Capítulo 1

Introducción

1.1 Motivación 7
1.2 Aportes y limitaciones del trabajo 9
1.3 Difusión de los aportes 10
1.3.1 Revistas internacionales 10
1.3.2 Congresos Nacionales 11
1.4 Organización de la Tesis 12

1.1. Motivación

La adquisición de grandes volúmenes de datos representa un desafío significativo en sistemas de reconstrucción espacial. Tradicionalmente, esta tarea se ha efectuado mediante un esquema de lazo abierto, utilizando protocolos de adquisición que no garantizan una posterior reconstrucción de calidad. Esta problemática se manifiesta en la recopilación de datos inadecuados, ya sea por exceso o por escasez, lo que genera una incertidumbre considerable en el modelo final. La insuficiencia de datos puede traducirse en la pérdida de información valiosa y en un derroche de recursos y tiempo, mientras que un exceso de datos no solo incrementa las demandas de almacenamiento y procesamiento, sino que también puede deteriorar la calidad de la reconstrucción al incorporar mediciones imprecisas o erróneas, conduciendo a posibles divergencias en el proceso de reconstrucción.

La rigidez inherente a los protocolos de adquisición convencionales compromete adicionalmente la robustez de la reconstrucción, especialmente

frente a perturbaciones no modeladas o cuando el objeto de interés es parcial o totalmente desconocido. Los errores resultantes, a menudo identificados solo tardíamente durante o después del procesamiento, magnifican la complejidad del problema. Ante este panorama, surge la necesidad imperante de sistemas de adquisición a lazo cerrado. Dichos sistemas deben facilitar la toma de decisiones en tiempo real y asegurar que la reconstrucción resultante cumpla con objetivos de calidad preestablecidos, como la completitud de una región determinada y la minimización de la incertidumbre del modelo final por debajo de un umbral aceptable. No obstante, una carencia notable en el estado del arte es la ausencia de criterios claros para determinar cuándo los datos adquiridos son suficientes y adecuados para satisfacer estos objetivos de calidad Placed and Castellanos (2022).

La reconstrucción 3D y, más específicamente, la adquisición activa de datos para la reconstrucción de entornos, son campos con una amplia trayectoria de investigación en robótica y visión por computadora, especialmente aplicados a la exploración. No obstante, la implementación sistemática y optimizada de estas técnicas en otros sectores, siendo el agrícola un ejemplo destacado, representa un área con un considerable potencial de desarrollo aún por explotar. Este último campo, en particular, podría transformar sus prácticas mediante la adopción de sistemas de adquisición de datos inteligentes y a lazo cerrado. De hecho, la agricultura de precisión, un paradigma en expansión que ya se fundamenta en la recopilación exhaustiva y oportuna de información, encontraría en estos sistemas avanzados la clave para una toma de decisiones agronómicas verdaderamente optimizada y una gestión de recursos más eficiente.

Una aplicación específica de este enfoque de adquisición de datos a lazo cerrado es la reconstrucción tridimensional de cultivos, la cual enfrenta desafíos singulares que van más allá de los problemas generales de la reconstrucción 3D. Por ejemplo, la integridad de la geolocalización puede verse comprometida por la pérdida de señales GNSS bajo canopias densas. A su vez, las técnicas de SLAM visual, si bien son versátiles, luchan con la sensibilidad a cambios abruptos de iluminación y con el bajo rendimiento en ambientes agrícolas que frecuentemente presentan texturas visuales pobres o altamente repetitivas. Incluso la reconstrucción con LiDAR, a menudo preferida por su robustez, no está exenta de dificultades, como la posible dispersión del haz láser al interactuar con vegetación fina, o la generación de nubes de puntos con densidad variable que complican la extracción de parámetros agronómicos consistentes. De manera transversal a cualquier modalidad sensorial, persiste la complejidad en la segmentación precisa de estructuras vegetales individuales. Todos estos factores, en conjunto, dificultan significativamente las tareas de reconstrucción en el ámbito agrícola.

Este desafío se agrava aún más al considerar la variabilidad temporal inherente a la estructura vegetal, la cual evoluciona constantemente a lo largo del ciclo de cultivo. Dicha variabilidad demanda reconstrucciones periódicas, que idealmente deberían realizarse de manera eficiente, enfocando los esfuerzos de adquisición en zonas de alto dinamismo vegetativo o en aquellas que presenten mayor incertidumbre o una relevancia particular para el manejo agronómico.

La relevancia de la reconstrucción 3D de los cultivos radica, fundamentalmente, en su capacidad para cuantificar parámetros biofísicos clave de las plantas y del dosel vegetal. Un ejemplo prominente es la estimación de la cantidad de luz interceptada por la canopia, como la Radiación Fotosintéticamente Activa interceptada (PARi, por sus siglas en inglés), un factor esencial para entender la fisiología de la planta y manejar el cultivo eficientemente. Además de este parámetro, la caracterización de la distribución espacial de la biomasa dentro de la canopia, así como la estimación del volumen o la densidad de la misma, son claves para mejorar la productividad del sector agrícola, tal como se discute en trabajos como el de Westling et al. (2020). Estos parámetros, derivados de modelos 3D, pueden informar y optimizar decisiones sobre poda, riego, fertilización y predicción de rendimiento.

En este contexto, el presente trabajo se enfoca en el desarrollo de herramientas metodológicas diseñadas para asistir en la toma de decisiones durante el proceso de la adquisición de datos. De esta manera, se puede garantizar que la adquisición de datos contribuya efectivamente a los esfuerzos de reconstrucción espacial y temporal en la agricultura y otros campos.

1.2. Aportes y limitaciones del trabajo

La labor investigativa que involucra cultivos a cielo abierto presenta retos notables. El objeto de estudio, es decir, el cultivo mismo, está expuesto a múltiples factores adversos e incontrolables, tales como condiciones climáticas severas (ej., sequías, inundaciones, heladas, granizo), la aparición de plagas y enfermedades, y la propia diversidad genética intra e intervarietal que puede influir en la respuesta a los tratamientos. Adicionalmente, la infraestructura experimental o de monitoreo puede sufrir fallas en sensores, en sistemas de irrigación y en las estaciones meteorológicas, entre otros elementos críticos para la recolección de datos y la gestión del ensayo.

Estos sistemas vegetales se caracterizan por ser dinámicos, evolucionando temporalmente con etapas de alta y baja variabilidad, lo que puede requerir la

aplicación de intervalos de muestreo no constantes. Sumado a esto, la ubicación de las plantaciones, frecuentemente distante de núcleos urbanos, complica el acceso a recursos esenciales como energía eléctrica y conectividad.

Estas circunstancias constituyen obstáculos significativos para el progreso y la aplicación de investigaciones en el ámbito de la agricultura de precisión. A ello se suma que los avances específicos utilizando determinados tipos de sensores o aplicados a un cultivo en particular, pueden enfrentarse a la obsolescencia tecnológica o presentar dificultades en su generalización a otros cultivos.

Por consiguiente, los aportes de este estudio tienen un enfoque teórico con aplicaciones prácticas extensibles más allá de la agricultura. No se restringe a un tipo de sensor, sistema de localización o cultivo en particular. Las contribuciones fundamentales de esta tesis son:

- Un indicador que evalúa tanto el área cubierta como la incertidumbre de la reconstrucción.
- Determinación de objetivos de exploración que puedan guiar la adquisición de datos. Estos objetivos de exploración extienden la funcionalidad de los algoritmos de exploración existentes.
- Un mapa de incertidumbre que permite la valoración de la calidad de los datos recolectados por zonas específicas.

Este enfoque integral enfatiza la importancia de superar las barreras existentes mediante soluciones escalables y adaptativas, capaces de enfrentar la variabilidad inherente al estudio de sistemas biológicos y sus entornos. Los aspectos teóricos desarrollados fueron probados y validados en un entorno de simulación y con datos reales en ambiente agrícola.

1.3. Difusión de los aportes

A continuación se enumeran las publicaciones realizadas durante el desarrollo de esta tesis:

1.3.1. Revistas internacionales

- Sansoni, S., Gimenez, J., Castro, G., Tosetti, S., & Capraro, F. (2025). Optimizing exploration with a new uncertainty framework for active SLAM systems. Robotics and Autonomous Systems, 193, 105059.

- Gimenez, J., Sansoni, S., Tosetti, S., Capraro, F., & Carelli, R. (2022). Trunk detection in tree crops using RGB-D images for structure-based ICM-SLAM. Computers and Electronics in Agriculture, 199, 107099.

1.3.2. Congresos Nacionales

- "Una Nueva Perspectiva para los Sistemas de SLAM Activo mediante Mapas de Incertidumbre", Sansoni, S., Gimenez, J., Castro, G.I, Tosetti, S., Capraro, F. (2024). XII Jornadas Argentias de Robótica (JAR), Ciudad de Buenos Aires, Buenos Aires, Argentina 2024.

- "ICM-SLAM: una propuesta de modelado y resolución versártil del problema SLAM", Gimenez, J., Sansoni, S., Tosetti, S., Carelli, R. (2022). XI Jornadas Argentinas de Robótica (JAR), San Carlos de Bariloche, Río Negro, Argentina 2022.

- "Alto Valle Dataset: colección de datos experimentales enfocados en el estudio y desarrollo de algoritmos de navegación mediante visión en ambientes frutícolas", Sansoni, S., Raverta Capua, F., Moreyra, M.L., Benitez Piccini, E. (2022). XI Jornadas Argentinas de Robótica (JAR), San Carlos de Bariloche, Río Negro, Argentina 2022.

- "Estimación de la posición lateral de un vehículo frutícola mediante visión estéreo", Sansoni, S., Moreyra, M.L. (2019). XVIII Reunión de Trabajo en Procesamiento de la Información y Control - RPIC 2019, Bahía Blanca, Buenos Aires, Argentina.

- "Comparative analysis of visual-SLAM algorithms applied to fruit environments", Raverta Capua, Sansoni, S.,F., Moreyra, M.L. (2018). 2018 Argentine Conference on Automatic Control (AADECA).

- "Estimación de la posición de una plataforma móvil en un ambiente frutícola basado en la detección de troncos con visión estéreo", Sansoni, S., Raverta Capua, F., Moreyra, M.L. (2017). IX Jornadas Argentinas de Robótica, Facultad Regional de Córdoba de la Universidad Tecnológica Nacional, Córdoba, Argentina.

1.4. Organización de la Tesis

El contenido de esta tesis doctoral está dividido en tres grandes capítulos: Revisión Bibliográfica, Desarrollo Teórico y Desarrollo Experimental.

En el primer capítulo, se realiza una Revisión Bibliográfica (Capítulo 2) de los distintos sistemas utilizados para la localización y mapeo simultáneo en $2D$ y en $3D$ (Sec. 2.2), seguido por sistemas de reconstrucción utilizando imágenes mediante *Structure from Motion* (*SfM*) (Sec. 2.4), luego sistemas basados en Aprendizaje Automático (*Machine Learning*) (Sec. 2.5) y técnicas de medición activa relacionadas con esas tres áreas (Sec. 2.6). Por último, se realiza una revisión de los distintos sistemas de adquisición de datos y reconstrucción aplicados en la agricultura (Sec. 2.7).

En el capítulo correspondiente al Desarrollo Teórico (Capítulo 3), se describe la metodología propuesta con sus distintos componentes, los cuales incluyen el modelado de la incertidumbre (Sec. 3.2), el proceso de actualización (Sec. 3.3), la umbralización (Sec. 3.5) y la métrica de evaluación basada en incertidumbre (Sec. 3.4).

En el capítulo correspondiente al Desarrollo Experimental (Capítulo 4), se describen los distintos experimentos realizados para validar la metodología propuesta. Se comienza con la implementación en entornos de simulación $2D$ (Sec. 4.1), luego se desarrolla un sistema *SLAM* $3D$ que provee una estima de la incertidumbre (Sec. 4.2.1) y, por último, se evalúa el sistema de reconstrucción con datos reales (Sec. 4.2.3).

Para finalizar, se presentan las Conclusiones y Trabajo Futuro (Capítulo 5). Seguido a esto se presenta el Anexo del trabajo (Capítulo A) en donde se desarrollan todos los temas que no se abordaron en los capítulos anteriores.

Capítulo 2

Revisión Bibliográfica: Estado del arte y marco teórico

2.1 Introducción y Organización del Capítulo . . . 14
2.2 Simultaneous Localization and Mapping (SLAM) . . . 14
2.2.1 Tipos de SLAM . . . 15
2.3 Reconstrucción 3D: Ejemplos de SLAM visual y con LiDAR . . . 18
2.3.1 Introducción . . . 18
2.3.2 ORB-SLAM . . . 18
2.3.3 FAST-LIO . . . 22
2.4 Reconstrucción 3D: Structure From Motion . . . 27
2.4.1 Introducción a SfM . . . 27
2.4.2 COLMAP . . . 27
2.5 Reconstrucción 3D con Machine Learning . . . 31
2.5.1 Neural Radiance Fields (NeRF) . . . 32
2.6 Medición activa . . . 33
2.6.1 Active SfM o Next Best View Planning . . . 35
2.6.2 Active SLAM . . . 35
2.7 Reconstrucción 3D en la agricultura . . . 42
2.8 Conclusiones, oportunidades y desafíos . . . 50

2.1. Introducción y Organización del Capítulo

Este capítulo aborda diversas metodologías y herramientas empleadas en la bibliografía referido a la reconstrucción de entornos. Además, se proporcionará una explicación de conceptos teóricos fundamentales para la comprensión de las metodologías discutidas, limitándose a lo estrictamente necesario. El objetivo de este análisis bibliográfico es ofrecer una visión general del estado del arte, así como identificar oportunidades y desafíos en el ámbito de la reconstrucción de entornos.

Un aspecto transversal a las primeras tres secciones, y que merece ser destacado, es la frecuente ausencia en la literatura de una estimación explícita de la incertidumbre asociada tanto a los sensores como a los algoritmos de reconstrucción. Aunque en esta tesis doctoral se asume la existencia de un sistema capaz de estimar la matriz de covarianza del agente y sus mediciones, es importante reconocer que hoy en día esta suposición representa un desafío considerable, cuya dificultad se analizará en el desarrollo de la sección. Precisamente, uno de los objetivos centrales de esta tesis es proponer una metodología que integre dicha incertidumbre en el proceso de reconstrucción de entornos.

Este capítulo se iniciará con el análisis de los sistemas que resuelven el problema de Localización y Mapeo Simultáneos (del inglés, *Simultaneous Localization and Mapping*, SLAM). En esta revisión, se diferenciará entre enfoques de localización y mapeo en 2D y 3D según el tipo de sensor principal empleado. Posteriormente, se examinarán sistemas de reconstrucción 3D mediante la técnica de *Structure from Motion* (SfM) y se presentarán ejemplos de sistemas de reconstrucción 3D basados en Aprendizaje Automático (*Machine Learning*, ML). Finalmente, se discutirán sistemas de medición activa y sus aplicaciones en el campo de la agricultura, conectando con los desafíos y oportunidades identificados previamente.

2.2. Simultaneous Localization and Mapping (SLAM)

El desafío de SLAM radica en estimar simultáneamente los estados de un agente móvil y la configuración espacial del entorno donde este se desplaza. Se define como agente a cualquier entidad capaz de moverse dentro de un entorno, como personas, sensores, robots autónomos o teleoperados. Los estados de un agente pueden incluir su posición en un plano (x, y) o en el espacio (x, y, z), así como su orientación representada por los ángulos de Tait-Bryan (*roll, pitch, yaw*). Un mapa, por otro lado, constituye una representación del entorno que puede ser bidimensional o

tridimensional y adoptar formas como grafos, grillas de ocupación, nubes de puntos, mallas de triángulos, entre otros.

Los algoritmos de SLAM son cruciales para la reconstrucción de entornos, ya que permiten asociar cada dato sensorial con una ubicación específica en el mapa, facilitando así la generación de una representación del entorno. Esta representación puede estar compuesta por *voxels*, mallas de triángulos o nubes de puntos. Para los propósitos de esta tesis, no se distinguirá entre los términos mapa y reconstrucción; sin embargo, es importante destacar que mientras que un mapa puede referirse a una abstracción como un grafo, una reconstrucción implica una representación espacial detallada.

En el panorama actual, proliferan sistemas dedicados a resolver la problemática de SLAM, cada uno caracterizándose por sus ventajas y limitaciones específicas. Dado el volumen significativo de revisiones bibliográficas existentes que exploran esta temática Gaia et al. (2023), se ha decidido prescindir de una revisión exhaustiva adicional sobre SLAM en el contexto de esta tesis. En cambio, se expondrán algunos ejemplos de sistemas SLAM en dos y tres dimensiones, acompañados de una descripción concisa de los principios esenciales subyacentes a SLAM. Entre las revisiones bibliográficas más relevantes se destacan Huang and Dissanayake (2016) y Cadena et al. (2016).

2.2.1. Tipos de SLAM

Basados en Filtros

Históricamente, el problema de SLAM se ha abordado mediante el uso de filtros de Kalman extendidos (*Extended Kalman Filter*, EKF) o de partículas (*Particle Filter*, PF). Estos métodos pueden ser altamente eficientes y se han utilizado en una gran cantidad de aplicaciones. En particular, los sistemas de SLAM basados en el filtro de Kalman, a medida que se adquiere la información de los sensores, el filtro actualiza la estimación de la posición del agente y del mapa de forma iterativa. El filtro requiere conocer la incertidumbre de medición y la del sistema, las cuales integra en cada iteración. Esto lo realiza de forma explícita, obteniendo una matriz de covarianza de todos los estados involucrados. Dado que dentro del filtro se requiere mantener una matriz de covarianza, esto hace que sea fácilmente accesible para trabajar con ella. Sin embargo, el cálculo de las matrices de covarianza está sujeto a errores de linealización, lo que puede llevar a una subestimación o sobreestimación de la incertidumbre (inconsistencia del filtro). Sumado a que se asume que las variables tienen distribuciones de probabilidad gaussiana, lo cual no se suele cumplir en la

práctica. Además, al integrar nueva información, luego no puede ser desechada, lo que puede llevar a una acumulación de errores y, eventualmente, a la divergencia del filtro. Esto significa que ante un fallo, el sistema no puede recuperarse. Por esto, es que un sistema de navegación autónoma crítico, no puede estar basado enteramente en este tipo de sistema sin atender esta problemática. En general, cuando se requiere implementar una solución de navegación autónoma en la agricultura, un fallo en el sistema puede ocasionar daños a la maquinaria, a los cultivos o a las personas. Para más detalles sobre este tipo de implementaciones se puede consultar Dissanayake et al. (2001). Entre los sistemas de SLAM 2D basados en EKF más conocidos se encuentra HECTOR-SLAM Kohlbrecher et al. (2011).

No obstante, hoy en día se siguen utilizando estos tipos de filtros para realizar fusión de sensores como GNSS (*Global Navigation Satellite System*), IMU (*Inertial Measurement Unit*), giróscopos y otros, permitiendo obtener una estimación *odométrica* de la posición del agente. Trabajos como FAST-LIO2 Xu et al. (2022) o MSCKF_VIO Sun et al. (2018) son ejemplos de cómo son utilizados para estimaciones odométricas con LiDAR y cámaras respectivamente. En general, una estimación *odométrica* es una estimación de la pose del agente con respecto a una pose anterior. No suelen ser precisas en el largo plazo, ya que acumulan errores de estimación. Además, se caracterizan por requerir un bajo costo computacional.

Por otro lado, los PF utilizan un conjunto de partículas (muestras) para representar la distribución de probabilidad posterior sobre las posibles soluciones al problema de SLAM, donde cada partícula consta de una estimación de la trayectoria del robot y un mapa del entorno. A diferencia de los filtros de Kalman, que asumen linealidad en el modelo de transición y medición (o linealizaciones locales para el EKF) y ruido gaussiano, los filtros de partículas pueden manejar eficazmente problemas no lineales y distribuciones de probabilidad multimodales. Gracias a su naturaleza de muestreo, los filtros de partículas pueden ser más robustos frente a incertezas significativas en la posición inicial del robot, proporcionando una buena convergencia incluso cuando la estimación inicial está lejos del verdadero estado. Como contrapartida, los PF tienen un costo computacional que crece con la cantidad de partículas. A medida que el mapa crece, se requiere un número mayor de partículas para converger a la solución correcta, lo cual puede ser computacionalmente costoso. La implementación más conocida de este tipo de filtros es *GMapping* Grisetti et al. (2007), una implementación de SLAM 2D muy utilizada mediante librerías de ROS.

Basados en Grafos

En los últimos años, los métodos basados en grafos han sido objeto de gran interés en la comunidad de SLAM. En estos grafos, los nodos representan las posiciones del agente y los bordes representan las mediciones, aunque existen variaciones en la forma de representación, como son los grafos de factores bipartitos. Con este modelado, el problema se traduce en minimizar un funcional de costo no lineal, el cual involucra las mediciones, las posiciones del agente y, en algunas ocasiones, puntos característicos del entorno (también conocidos como *landmarks*). Este funcional se minimiza mediante técnicas de optimización no lineales, como el método de Levenberg-Marquardt o Gauss-Newton. Tras la optimización de todo el grafo (conocida como *full SLAM* o *batch optimization*), es posible recuperar la matriz de covarianza marginal sobre todas las variables estimadas. Cuando la optimización se formula directamente sobre una variedad (ej., $SE(3)$ en el caso de las poses 3D) la covarianza obtenida conserva una interpretación coherente en cada grado de libertad. No obstante, su cálculo y tratamiento resultan más exigentes que en espacios estrictamente euclídeos, debido principalmente a la linealización necesaria para recuperarla a partir de la matriz de información, es decir, la inversa del Hessiano de la función de costo Barfoot and Furgale (2014).

Cuando los mapas crecen en tamaño y/o las trayectorias en longitud, el problema se vuelve computacionalmente intratable si no se toman medidas para reducir la complejidad. Para lograr la operación en tiempo real se recurre principalmente a dos estrategias: i) implementar una ventana deslizante sobre los nodos del grafo, optimizando solo la información reciente y marginalizando la información antigua, o ii) realizar una optimización local, dejando el resto de la información de los nodos fija. Ambas estrategias presentan problemas para recuperar las incertidumbres de estimación de forma global y consistente. En Castro (2021) se propone una metodología para obtener una aproximación de las incertidumbres de forma global realizando optimizaciones locales, sin embargo este tema sigue en desarrollo y no hay soluciones ampliamente adoptadas o triviales. En la práctica se utilizan librerías como *g2o* Kümmerle et al. (2011a), *Ceres-Solver* Agarwal et al. (2023) o GTSAM Dellaert (2012) para implementar la optimización no lineal.

Estos tipos de SLAM almacenan los datos de medición y con ello construyen un mapa. Si algún conjunto de mediciones es erróneo, este puede ser desechado eventualmente y el sistema puede recuperarse. En general, estos sistemas son más robustos que los basados en filtros, pero requieren un mayor costo computacional.

Algunas técnicas para reducir el costo computacional consisten en eliminar datos poco informativos, es decir, que tienen una alta incertidumbre o son redundantes. Esta selección de datos se suele realizar una vez que todos los datos ya fueron capturados. En esta tesis se busca implementar una metodología a lazo cerrado que permita realizar esta selección de datos antes que sean capturados, lo que reduce el costo computacional, entre otros beneficios.

2.3. Reconstrucción 3D: Ejemplos de SLAM visual y con LiDAR

2.3.1. Introducción

En esta sección se van a describir dos sistemas representativos de SLAM visual y con LiDAR. Con ello se busca familiarizar al lector con las metodologías y herramientas utilizadas en la reconstrucción de entornos. Como ya se mencionó en la Sec. 1.2, el aporte principal de esta tesis no depende intrínsecamente del tipo de SLAM implementado, por lo que una revisión bibliográfica exhaustiva sobre los sistemas de SLAM existentes escapa al alcance de este trabajo.

Como ejemplo de SLAM visual se presentará ORB-SLAM, un sistema que ha sido ampliamente adoptado en la comunidad científica y técnica. La elección de este sistema sobre otros se debe principalmente a su popularidad y a su eficacia como solución integral y de código abierto para SLAM monocular, estéreo y RGB-D.

Por otro lado, se presentará FAST-LIO, un sistema de odometría inercial-iDAR que ha demostrado ser eficiente en la reconstrucción de entornos urbanos y rurales, operando en tiempo real ($10 - 100$ Hz) en una computadora onboard[1]. Se menciona en particular este tipo de odometría ya que se utilizará a modo comparativo en la implementación de la metodología propuesta en esta tesis.

2.3.2. ORB-SLAM

Hasta la fecha, el paper de ORB-SLAM Mur-Artal et al. (2015) posee más de 7300 citas y, si se suman las citas de su versión 2 Mur-Artal and Tardós (2017) (más de 5900) y la versión 3 Campos et al. (2021) (más de 2100) se llega en total a más de 15300 citas. Indudablemente, ORB-SLAM es uno de los sistemas de SLAM visual más utilizados y citados en la literatura.

[1] https://www.dji.com/global/manifold-2/specs

Un sistema de SLAM visual es aquel que utiliza principalmente cámaras para estimar la posición del agente y el mapa del entorno. Estos sistemas se caracterizan por ser demandantes computacionalmente y relativamente poco robustos a los cambios de iluminación, a los ambientes con baja textura y a estructuras repetitivas con características visuales similares, como es el caso de la agricultura. En general, para obtener la posición de un punto en el espacio 3D, se requieren al menos dos imágenes. Sin embargo, hay posiciones relativas entre las imágenes y el punto a reconstruir que generan un sistema de ecuaciones mal condicionado, con errores de estimación prohibitivos. Por esto, los sistemas de SLAM visual poseen una lógica que intenta evitar este tipo de situaciones, aunque no siempre es posible.

ORB-SLAM es una colección de subsistemas armonizados para realizar SLAM visual en tiempo real, alcanzando errores de estimación de pose reportados del orden de 1 cm en ambientes interiores y de algunos metros en grandes espacios exteriores. Entre esos subsistemas se destaca la librería de optimización *g2o* Kümmerle et al. (2011b), la detección de lazos mediante *BagOfWords* Gálvez-López and Tardos (2012), el cierre de lazos locales en tiempo constante Strasdat et al. (2011) y, por último, la utilización de características ORB Rublee et al. (2011). De estos subsistemas, se describirá brevemente el algoritmo de detección de características ORB y el proceso de optimización conocido como *Bundle Adjustment* (BA), ya que son relevantes para el desarrollo de esta tesis.

Para localizar al agente en el ambiente mediante imágenes, lo que se hace es capturar una imagen, extraer puntos característicos[2] de la misma y compararlos con los puntos característicos de una imagen previa. Si estos puntos corresponden a la misma posición del espacio, se puede calcular la *pose*[3] del agente mediante técnicas como la triangulación (para inicializar nuevos puntos) o resolviendo el problema de Perspectiva-n-Puntos (PnP) (para localizar la cámara respecto a puntos conocidos). Dentro de los algoritmos más populares se encuentran los descriptores *Scale-Invariant Feature Transform* (SIFT) Lowe (2004) y *Speeded Up Robust Features* (SURF) Bay et al. (2006), entre muchos otros; sin embargo, en este sistema se implementó y modificó un algoritmo conocido por *Oriented FAST and Rotated BRIEF* (ORB) Rublee et al. (2011). Esta metodología combina dos descriptores como son FAST (*Features from Accelerated Segment Test*), Rosten and Drummond (2006) y BRIEF (*Binary Robust Independent Elementary Features*) Calonder et al. (2010). El algoritmo FAST es un algoritmo de detección de esquinas, que se basa en la intensidad de los píxeles de la imagen. El algoritmo BRIEF es un algoritmo de

[2]Un punto característico es aquel que es reconocible de forma sistemática y sin ambigüedad dentro de una imagen. Pueden ser invariantes ante los cambios de escala y rotaciones, entre otros.

[3]Posición y orientación.

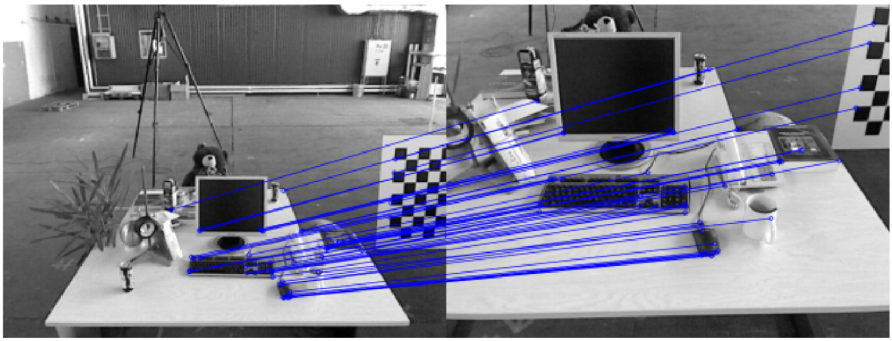


Figura 2.1: Imagen de ejemplo de cómo se utilizan los descriptores del tipo ORB para identificar los mismos puntos espaciales en dos imágenes distintas. En este caso se puede observar que el descriptor ORB identifica de forma eficaz los mismos puntos en una escena cuando interviene un cambio de escala. Imagen extraída de Mur-Artal et al. (2015).

extracción de características que se caracteriza por ser eficiente en el manejo de la memoria y de los recursos computacionales. Por otro lado, ORB combina ambos algoritmos, siendo un descriptor binario, invariante a la rotación y a los cambios de escala (hasta cierto punto) y menos vulnerable a los ruidos presentes en las imágenes obteniendo así un algoritmo más robusto. En la Fig. 2.1 se puede observar un ejemplo de cómo se utilizan los descriptores ORB para identificar los mismos puntos espaciales en dos imágenes distintas.

Por otro lado, ORB-SLAM utiliza la estrategia conocida como selección de *keyframes*. Un *keyframe* es una imagen que se selecciona de forma que sea representativa de un conjunto de imágenes. Esta imagen se utiliza para obtener la posición de la cámara y de los puntos en el mapa global. Todas las otras imágenes que no son consideradas *keyframe* son descartadas. En este sistema, se selecciona un *keyframe* cuando se cumplen una serie de criterios, entre ellos, está la cantidad de características ORB detectadas en la imagen. En otros sistemas de SLAM visual se suele agregar un nuevo *keyframe* cuando la cantidad de movimiento entre dos imágenes consecutivas supera un umbral, sin embargo, se utilizan también otros criterios. La idea de seleccionar *keyframes* es reducir la cantidad de imágenes que se deben procesar, lo que reduce el costo computacional del sistema.

Además, ORB-SLAM utiliza una técnica de optimización fundamental en

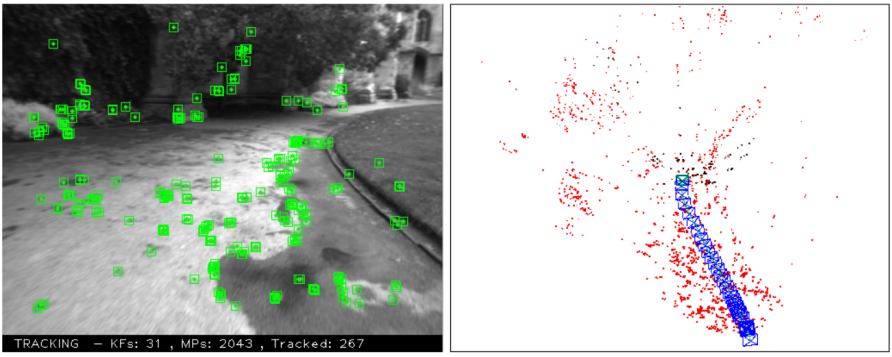


Figura 2.2: Ejemplo de reconstrucción 3D con ORB-SLAM. A la izquierda se ve una imagen en escala de grises con los puntos característicos detectados. A la derecha se observa la posición 3D de las cámaras y de los puntos detectados en las imágenes. Imagen extraída de Mur-Artal et al. (2015).

SLAM visual y SfM conocida como *Bundle Adjustment* (BA) Triggs et al. (2000). Dado un conjunto de imágenes (o más precisamente, poses de cámara estimadas correspondientes a los keyframes) y un conjunto de puntos característicos 3D que se ven en dos o más imágenes, BA busca refinar conjuntamente la estructura (posiciones 3D de los puntos) y el movimiento (poses de las cámaras, y opcionalmente, los parámetros intrínsecos de las cámaras [4]). Para esto, se parte de una estimación inicial de estas variables. Con esta estimación, se calcula la posición proyectada de cada punto 3D en cada imagen donde es visible y se compara con la posición observada del punto característico en dicha imagen. La diferencia entre la posición observada y la estimada (proyectada) es conocida como error de reproyección. La idea de BA es minimizar la suma de los cuadrados de estos errores de reproyección para todos los puntos y cámaras, lo que se traduce en una mejor estimación de la posición del agente y del mapa. Esta técnica de optimización es parte de los sistemas SfM que se explicarán en la Sec. 2.4.

A modo ilustrativo, en la imagen 2.2 se muestra un ejemplo de reconstrucción 3D con ORB-SLAM. Si bien ORB-SLAM Mur-Artal et al. (2015) es un sistema de SLAM visual monocular, en posteriores trabajos se integró la utilización de cámaras

[4]Los parámetros intrínsecos son: focal length, punto principal y coeficientes de distorsión, entre otros.

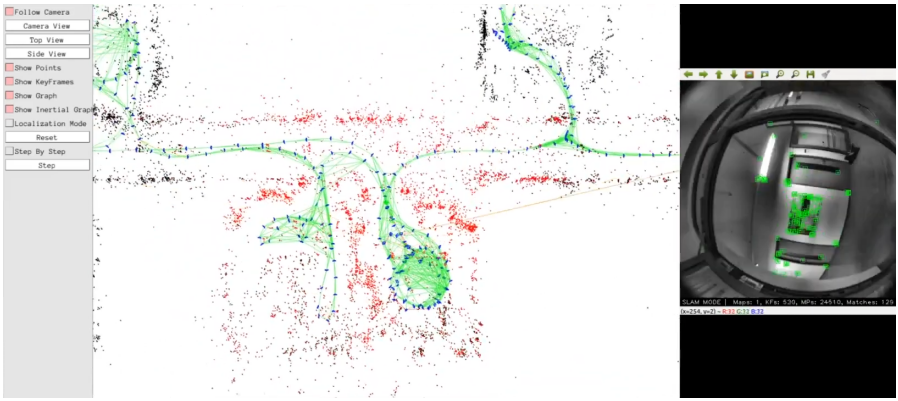


Figura 2.3: Ejemplo de reconstrucción 3D con ORB-SLAM3. A la izquierda se observa la trayectoria realizada (línea verde), los keyframes (objetos azul) y los puntos mapeados del entorno. A la derecha se observa la imagen procesada con las características ORB trackeadas. Imagen extraída del video https://www.youtube.com/watch?v=kZbpIbaBnZ0.

estéreo, IMU, entre otros Campos et al. (2021). En la Fig 2.3 se puede observar un ejemplo de reconstrucción 3D con ORB-SLAM3.

Por último, cabe remarcar que ORB-SLAM, en todas sus versiones publicadas, no proporciona explícitamente una matriz de covarianza completa y consistente de las estimaciones realizadas (poses de cámara y puntos del mapa) ni se analiza en detalle su consistencia. Esto se alinea con la discusión previa sobre los desafíos en la recuperación de incertidumbre global en sistemas de SLAM basados en grafos que emplean optimización local o por ventanas (Sec. 2.2.1), aunque la optimización por BA sobre el grafo completo podría, en principio, permitir la estimación de dicha covarianza a partir de la matriz de información.

2.3.3. FAST-LIO

Fast LiDAR Inertial Odometry o FAST-LIO Xu and Zhang (2021) y su segunda implementación Xu et al. (2022) es un sistema de odometría que utiliza datos

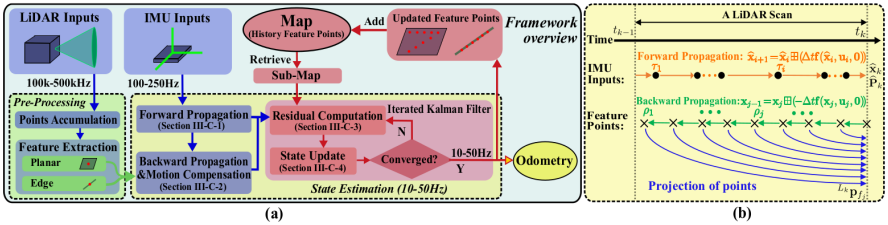


Figura 2.4: Esquema del sistema FAST-LIO Xu and Zhang (2021). (a) Diagrama de flujo principal, mostrando el preprocesamiento LiDAR con extracción de *features* planares y de borde, la estimación de estado mediante un iEKF que fusiona datos de IMU y LiDAR, y la gestión del mapa. (b) Ilustración de la propagación hacia adelante (con datos IMU) y la propagación hacia atrás para la compensación de movimiento de los puntos LiDAR. Imagen extraída de Xu and Zhang (2021).

de un LiDAR 3D y una IMU *fuertemente acoplados*[5] mediante un iEKF (*iterated Extended Kalman Filter*). Se caracteriza por su bajo costo computacional, obtenido gracias a una formulación particular que permite calcular la ganancia de Kalman eficientemente, actualizando un estado de dimensión fija (e.g., 18×18 para una pose de 6 DoF, velocidad, y sesgos de la IMU) sin importar la cantidad de puntos (features o no) de la nube de puntos utilizados en la etapa de medición. A modo descriptivo, en la Fig. 2.4 se muestra un esquema que explica el sistema completo de la primera versión.

Al igual que ORB-SLAM, FAST-LIO (en su primer versión) requiere la obtención de puntos característicos o features para funcionar. Estos features se obtienen a partir de una nube de puntos. Para obtenerlos, primero se retiran aquellos puntos mal condicionados (ángulos de incidencia cercanos a 0 o π rad, reflectividad del haz muy alta o muy baja, entre otros) Lin and Zhang (2020) y luego se extraen características de los puntos restantes. Estas características son principalmente planos y bordes, obtenidos mediante el cálculo de la *suavidad local* (del inglés *local smoothness*) Zhang and Singh (2014). Con estas características se genera una base de datos (en este caso, un *k-dimensional tree* o *k-d tree*) de features que se utiliza para comparar con los features de la nube de puntos actual (mediante el cálculo

[5]Todos los datos disponibles son procesados simultáneamente. A diferencia de los sistemas *débilmente acoplados* en donde los datos son procesados individualmente antes de fusionarse con el resto.

de distancia plano a plano o borde a borde) y así obtener la posición del agente y el mapa del entorno. Este último paso está influenciado por el patrón de escaneo propio del LiDAR (*prism-based*, *spinning*, *multi-line*, *solid-state*, *MEMS*, entre otros Raj et al. (2020))

Por otro lado, debido a los diferentes tiempos de muestreo entre los LiDARs y las IMUs, entre dos mediciones de LiDAR se pueden obtener varias mediciones de la IMU. Por esto, se realiza una propagación de los estados del sistema y de la matriz de covarianza a medida que hay nuevas mediciones de la IMU (propagación hacia adelante, *Forward Propagation*). Una vez que llega una nueva medición del LiDAR, esta medición está compuesta por muestras tomadas en diferentes tiempos, es decir, los puntos medidos son registrados uno después del otro (en lugar de hacerlo simultáneamente). Los puntos resultantes corresponden a posiciones del LiDAR distintas, por esto, se debe realizar una corrección de la distorsión de movimiento (*motion distortion compensation* o *deskewing*) para que todos los puntos medidos puedan ser referidos a una única pose del LiDAR (generalmente al final del barrido). Esta corrección se hace con las estimaciones de movimiento interpoladas a partir de las predicciones de la IMU durante la propagación hacia adelante y se la denomina propagación hacia atrás (*Backward Propagation*) en el contexto de este algoritmo, aunque el término más general es compensación de distorsión de movimiento. Dicha corrección resulta necesaria y evidente cuando los rangos de medición son muy grandes, ya que un pequeño cambio de orientación produce un gran cambio en la estimación de la posición de los puntos lejanos. En entornos abiertos, como en la agricultura, se obtienen muchas observaciones de puntos lejanos, que si bien son útiles para la estimación de la pose son muy sensibles a errores de orientación. Por esto es relevante realizar dicha corrección y tener una estimación confiable de las incertidumbres asociadas a las mediciones.

En el caso de FAST-LIO2 Xu et al. (2022), se evita calcular *features* explícitos de la nube de puntos, ganando robustez en ambientes con escasez de *features* geométricos distintivos e independencia del tipo de LiDAR que se utilice. En la Fig. 2.5 se muestra un esquema del sistema completo.

En esta versión, en lugar de generar un mapa de features, se van registrando las mediciones del LiDAR en un sistema de referencia global. Este mapa es organizado mediante un *k-d tree* incremental. A medida que se van registrando las mediciones, se van actualizando las posiciones de los puntos en el espacio global. Para obtener la posición del agente y el mapa del entorno, se compara la nube de puntos actual con la nube de puntos registrada en el sistema de referencia global. Para esto se utiliza un algoritmo de registro de nubes de puntos iterativo, similar en espíritu al *Generalized ICP* (GICP) Segal et al. (2009), calculando residuos entre los puntos actuales y las

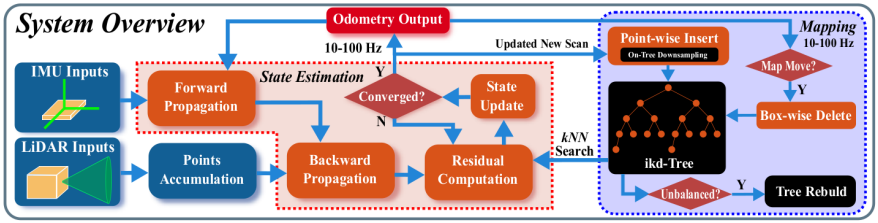


Figura 2.5: Visión general de la arquitectura de FAST-LIO2, que incluye los siguientes módulos: (i) preprocesamiento y *deskew* de los escaneos LiDAR, (ii) propagación del estado inercial con las lecturas de la IMU, (iii) registro directo punto-a-mapa y corrección en un filtro de Kalman extendido, y (iv) mantenimiento del mapa global en tiempo real mediante un *ikd-tree* incremental. Imagen extraída de Xu et al. (2022).

superficies locales (planos) inferidas del mapa global. Se calcula la distancia punto a plano en regiones pequeñas, asumiendo que los datos son lo suficientemente suaves como para ajustar un plano. Tanto con la propagación hacia adelante (mediante los datos de la IMU utilizando el modelo del sistema), como la distancia punto a plano, considerándolo como la entrada de medición, se implementa el iEKF para obtener la posición del agente y el mapa del entorno.

Los resultados con FAST-LIO2 son prometedores, ya que se obtiene una reconstrucción 3D del entorno en tiempo real, con un costo computacional bajo y una precisión aceptable. Sin embargo, al igual que ORB-SLAM, no se proporciona explícitamente una caracterización completa o un análisis de consistencia de la matriz de covarianza de las estimaciones realizadas, aunque debido al uso de iEKF esta matriz es un subproducto de la estimación, por lo que podría obtenerse. Debido a las largas distancias obtenidas en espacios exteriores, se requiere una estimación de las incertidumbres asociadas a las mediciones. Además, puede suceder que cuando se reconstruya un objeto, se tengan varias mediciones del mismo, realizadas a distintas distancias y por lo tanto con distintas incertidumbres. La asociación de una medición con un valor de incertidumbre es fundamental para fusionar datos de manera óptima o, cuanto menos, elegir la medición más confiable. En esta tesis se propone una forma de medir la incertidumbre tal que permita asociar variables con diferentes unidades, como son las mediciones de rango del LiDAR y las estimaciones angulares del sistema de posicionamiento.

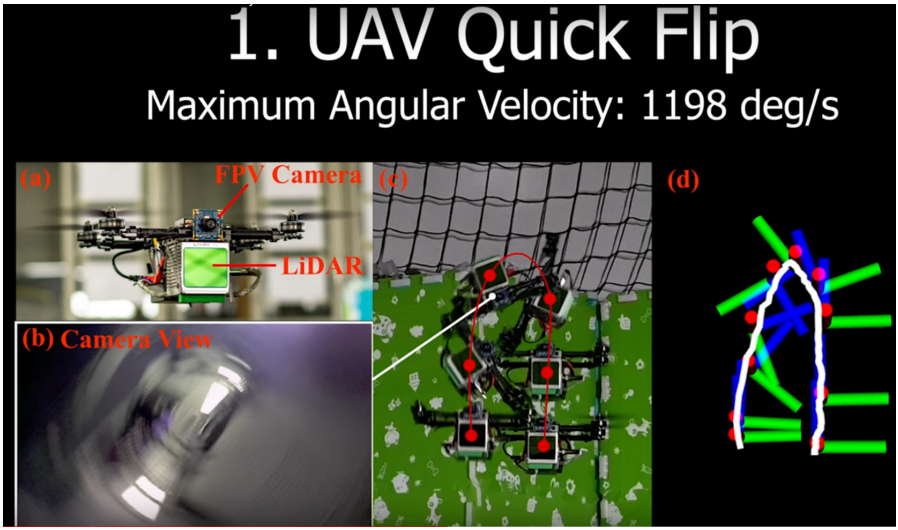


Figura 2.6: Demostración de la robustez de FAST-LIO2 Xu et al. (2022) durante una maniobra agresiva de un UAV ("Quick Flip") con una velocidad angular máxima de 1198 deg/s. (a) UAV equipado con LiDAR y cámara FPV. (b) Vista de la cámara FPV con severo desenfoque de movimiento. (c) Estimación de la trayectoria del UAV usando imágenes. (d) Poses estimadas por el sistema. Captura de pantalla del video oficial de FAST-LIO2 https://youtu.be/2OvjGnxszf8?t=234.

En la Fig. 2.6 se muestra un screenshot de un video de FAST-LIO2 en funcionamiento. En este video se puede observar la reconstrucción 3D del entorno en tiempo real sumado a un posicionamiento preciso y veloz del agente. En particular, el sistema en este caso funciona a 100 Hz en una computadora onboard[6]. Para probar esto, los autores montaron el sistema sobre un UAV y realizaron movimientos bruscos y rápidos. En particular en el screenshot se muestra un giro de $360°$ en $0.3s$.

[6] https://www.dji.com/global/manifold-2/specs

2.4. Reconstrucción 3D: Structure From Motion

2.4.1. Introducción a SfM

El término *Structure from Motion* (SfM) se refiere al problema de reconstruir la estructura tridimensional de un objeto o entorno a partir de una secuencia, a menudo desordenada y no necesariamente contigua temporalmente, de imágenes bidimensionales (2D). En esta técnica, a diferencia de los sistemas SLAM y de odometría visual[7], las imágenes pueden ser tomadas desde distintos dispositivos con diferentes parámetros de calibración.

COLMAP Schonberger and Frahm (2016) es un framework de código abierto que implementa SfM y *Multi-View Stereo* (MVS) con una interfaz gráfica de usuario. COLMAP es uno de los sistemas de SfM más utilizados en la comunidad científica y técnica. Hoy en día se lo utiliza como *gold standard*. En la Fig. 2.7 se muestra un ejemplo de reconstrucción 3D con COLMAP, en su versión de mapa disperso, como representación densa[8].

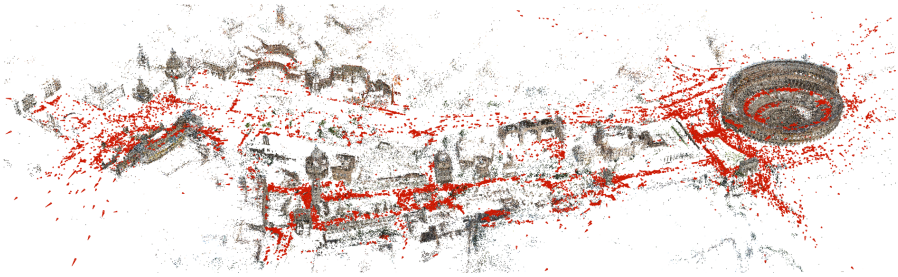
Por este motivo, se incluye COLMAP dentro de esta revisión bibliográfica y se usará como base para explicar algunos detalles de la técnica de SfM.

2.4.2. COLMAP

En la Fig. 2.8 se muestra el flujo de trabajo generalmente propuesto en COLMAP Schonberger and Frahm (2016), el cual ilustra los componentes típicos de muchos sistemas de SfM. En términos generales, el flujo de trabajo principal consta de varias etapas clave, que aquí agruparemos en tres fases principales: i) Búsqueda de Correspondencias (*Correspondence Search*), ii) Reconstrucción Incremental Dispersa (*Incremental Sparse Reconstruction*), y iii) Reconstrucción Densa (*Dense Reconstruction*, a menudo utilizando Estéreo Multi-Vista o *Multi-View Stereo*, MVS). En la primera fase, la Búsqueda de Correspondencias, se extraen los puntos característicos y se establecen correspondencias entre pares de imágenes. En la segunda fase, la Reconstrucción Incremental Disparza, se estima la pose de las cámaras y las coordenadas 3D de los puntos característicos, resultando en

[7] La odometría visual consiste en obtener las posiciones *relativas* entre dos frames utilizando la información de las imágenes, aunque también puede incorporar información de la IMU. En general, no utilizan algoritmos de optimización para minimizar los errores del mapa.

[8] Una reconstrucción densa tiene más puntos que una dispersa. En general, la reconstrucción dispersa se utiliza para localizar al agente y la densa en un procesamiento posterior, más detallado. El modelo denso depende del modelo disperso.



(a) Modelo disperso del centro de Roma usando 21 mil fotos. Este modelo se obtuvo con el flujo de trabajo SfM de COLMAP. Los puntos rojos corresponden a las posiciones de las cámaras.



(b) Modelo denso utilizando el flujo de trabajo MVS de COLMAP.

Figura 2.7: Ejemplo de reconstrucción densa 3D con COLMAP. Imágenes extraídas de https://colmap.github.io/.

un modelo disperso. En la tercera fase, se obtiene una reconstrucción 3D densa del entorno. Como el objetivo de esta tesis no se centra en esta última etapa (reconstrucción densa por MVS), no se profundizará en ella.

Extracción de características y emparejamiento (Correspondence Search)

En esta etapa inicial, se extraen puntos característicos (features) de cada imagen y, posteriormente, se buscan emparejamientos (matches) de estos features entre diferentes imágenes. Para extraer los puntos característicos, COLMAP comúnmente utiliza el algoritmo SIFT (*Scale-Invariant Feature Transform*) Lowe (2004), conocido por su robustez a cambios de escala, rotación e iluminación, aunque también puede implementar otros detectores y descriptores. Una vez que se han encontrado correspondencias tentativas basadas en la similitud de los descriptores, se realiza un proceso de verificación geométrica para cada par de imágenes con suficientes correspondencias. Este proceso típicamente implica estimar la matriz fundamental (para cámaras no calibradas) o la matriz esencial (para cámaras calibradas) entre el

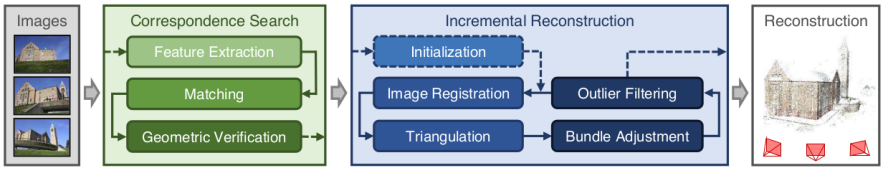


Figura 2.8: Flujo de trabajo propuesto en COLMAP. Imagen extraída de Schonberger and Frahm (2016).

par de imágenes usando un algoritmo robusto como RANSAC Fischler and Bolles (1981) para descartar outliers. El objetivo es garantizar que exista un solapamiento visual significativo entre las imágenes y que la cantidad de puntos emparejados consistentes geométricamente (*inliers*) sea suficiente para una estimación fiable de la pose relativa y, posteriormente, la triangulación de puntos 3D.

Agregado iterativo de imágenes al modelo, optimización y filtrado (Incremental Reconstruction)

En esta etapa se selecciona el par de imágenes desde donde se creará toda la reconstrucción. Este par debe ser cuidadosamente elegido ya que de esto depende la calidad de la reconstrucción (entendida como a la robustez, precisión y performance). Realizar el proceso de inicialización con un par de imágenes que presenten una buena base estereoscópica (baseline), suficientes correspondencias de *features* bien distribuidos y un solapamiento adecuado es crucial para obtener una buena reconstrucción. A partir de este par inicial, se estima la pose relativa y se triangulan los primeros puntos 3D.

Luego de esta inicialización, se van agregando nuevas imágenes al modelo una por una. Para cada nueva imagen candidata, primero se estima su pose 3D utilizando las correspondencias entre sus *features* 2D y los puntos 3D ya reconstruidos en el modelo existente. Este proceso se realiza típicamente mediante un algoritmo de Perspectiva-n-Puntos (PnP) Fischler and Bolles (1981) dentro de un esquema RANSAC para robustez a outliers. Una vez registrada la imagen (es decir, estimada su pose), se triangulan nuevos puntos 3D a partir de sus *features* que también son visibles en otras imágenes ya registradas.

Paralelamente a la integración de nuevas imágenes y puntos al modelo, se realizan optimizaciones periódicas mediante *Bundle Adjustment* (BA), como se describió en la sección 2.3.2. Esto es necesario para refinar conjuntamente todas las

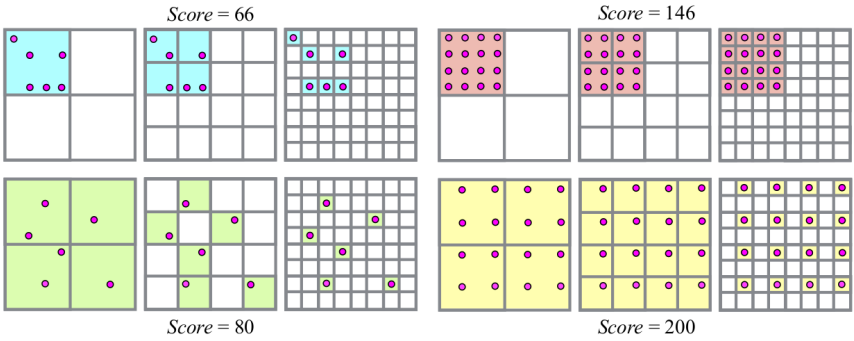


Figura 2.9: Puntajes para diferentes números de puntos y distintas distribuciones. Imagen extraída de Schonberger and Frahm (2016).

poses de cámara y puntos 3D estimados, minimizando los errores de reproyección globales y mitigando la acumulación de errores inherente al proceso incremental de triangulación y registro de poses.

En COLMAP, se utiliza una técnica de *Next Best View* (NBV) para decidir qué imagen no registrada se intentará agregar a continuación y, por ende, determinar el orden en que se van a agregar al modelo. En lugar de calcular la covarianza de las estimaciones como propone Haner and Heyden (2012) debido al costo computacional, se aproxima la incertidumbre de estimación mediante un análisis de múltiple resolución. Este análisis consiste en discretizar la imagen en celdas, si una celda posee al menos un punto característico se considera que está ocupada. Cada celda ocupada suma un puntaje asociado al área que ocupa. Este proceso se realiza para distintos tamaños de celda. Luego, la imagen con mayor puntaje será aquella que tenga muchos puntos característicos y uniformemente distribuidos. En la Fig. 2.9 se muestra un ejemplo de cómo se seleccionan las imágenes a agregar al modelo basado en este heurístico.

Esta estrategia de selección denominada *Pyramid* utiliza una aproximación para la incertidumbre de estimación. Según Schonberger and Frahm (2016) se seleccionan las mismas imágenes utilizando este método que si se utilizara otros métodos más costosos computacionalmente como son *Number* y *Ratio*. Además *Pyramid* obtiene una reconstrucción más precisa ya que obtiene un mejor orden de las imágenes a agregar al modelo. Este dato es relevante ya que el orden en que se procesan los

datos influye en la calidad de la reconstrucción. En esta tesis se propone una forma de detectar cuáles podrían ser los mejores siguientes datos a capturar, sin embargo no se especifica como adquirirlos, es decir, con qué trayectoria alcanzar ese conjunto de datos[9].

El proceso de filtrado de *outliers*[10] es continuo durante toda la reconstrucción incremental. Por ejemplo, durante el registro de nuevas imágenes con PnP, RANSAC se usa para identificar *inliers*. Similarmente, durante la triangulación de nuevos puntos, se aplican criterios para asegurar que los puntos estén bien condicionados (ej., ángulo de triangulación suficiente, error de reproyección bajo en las imágenes que lo ven, profundidad positiva). COLMAP utiliza algoritmos robustos para la triangulación, como el conocido por *Direct Linear Transform* (DLT) Hartley and Zisserman (2003) seguido a menudo por refinamiento no lineal para cada punto.

2.5. Reconstrucción 3D con Machine Learning

Esta revisión bibliográfica sobre los distintos métodos de reconstrucción 3D no puede estar completa sin mencionar algunas técnicas que utilicen Machine Learning. Dado el rápido avance en este campo, y riesgo de quedar desactualizado en el futuro, se ha seleccionado un método que ha tenido un impacto significativo en la comunidad científica y técnica. *Neural Radiance Fields* (NeRF) Mildenhall et al. (2020) es un método de reconstrucción 3D denso cuya innovación radica en su capacidad para producir imágenes fotorealistas de alta calidad de escenas 3D utilizando un modelo de red neuronal densamente conectado. Gracias a este enfoque se produjeron otros resultado interesantes como son *Gaussian Splatting* Kerbl et al. (2023), *Plenoptic Neural Radiance Fields* Fridovich-Keil et al. (2022), entre otros.

Es necesario señalar que, a pesar de los impresionantes resultados obtenidos con estas técnicas, ninguna ofrece una medida de incertidumbre de las estimaciones realizadas, por lo que no es posible comparar los métodos en términos de precisión de la representación sin conocer el ground truth. Sin embargo en el área de *active learning* y en la cuantificación de incertidumbre para redes neuronales profundas, se han propuesto métodos para medir la incertidumbre de estimación de los modelos basados en redes neuronales profundas Gawlikowski et al. (2023) utilizando técnicas como *Monte Carlo Dropout* Gal and Ghahramani (2016), *Deep*

[9] La planificación de trayectorias, en este contexto, es un tema complejo que requiere una mayor profundicación. En la Sec. 2.6.2 se trata este tema.

[10] Valores atípicos.

Ensembles Lakshminarayanan et al. (2017), entre otras. La aplicación de estas técnicas para obtener estimaciones de incertidumbre fiables en representaciones neuronales de escenas es un área de investigación activa.

2.5.1. Neural Radiance Fields (NeRF)

NeRF Mildenhall et al. (2020) es un método que aprende una representación continua y volumétrica de una escena 3D utilizando una red neuronal para aproximar la función de radiación de la luz. Esta función mapea una coordenada 3D (x, y, z) y una dirección de vista 2D (θ, φ) al color RGB y a la densidad volumétrica σ en ese punto. Si bien la síntesis de vistas puede ser eficiente una vez entrenado el modelo, el proceso de entrenamiento en sí mismo es computacionalmente costoso. En la Fig. 2.10 se muestra el flujo de trabajo de NeRF. Utiliza una red neuronal profunda totalmente conectada (*fully-connected*) del tipo MLP cuyas entradas son la posición 3D (x, y, z) (transformada mediante codificación posicional para capturar detalles de alta frecuencia) y la dirección de la vista (θ, φ), y las salidas son la densidad σ (dependiente solo de la posición) y el color RGB (dependiente de la posición y la dirección de vista). Una vez entrenada la red, se puede renderizar una imagen desde una pose de cámara arbitraria mediante la técnica de renderizado volumétrico a lo largo de los rayos de la cámara. Para el entrenamiento se requiere un conjunto de imágenes de la escena y sus respectivas poses de cámara precisas (generalmente obtenidas previamente mediante un método de SfM como COLMAP).

En la Fig. 2.10a) se muestra el proceso de entrenamiento de la red. Se toma un rayo de la cámara y se muestrean puntos a lo largo de este rayo. Para cada punto se obtiene la densidad y el color del volumen. Estos datos se ingresan a la MLP y se comparan con los datos reales. La red se entrena mediante la minimización de los errores entre los datos reales y los datos generados por la red. En la Fig. 2.10b) se muestra el proceso de inferencia. Se elige un punto de vista con sus 5 coordenadas $(x, y, z, \theta, \varphi)$ generando un rayo y se muestrean puntos a lo largo de él. Para cada punto se obtiene la densidad y el color del volumen. Luego mediante un proceso de integración Fig. 2.10c) se obtiene la imagen renderizada del modelo 3D desde ese punto de vista en particular. Por último en la Fig. 2.10d) se calculan los errores de la imagen generada con la imagen real.

Un resultado interesante de este proceso es que el modelo 3D está almacenado en la MLP y no en una nube de puntos, lo que permite obtener imágenes renderizadas de alta calidad desde puntos de vista no muestreados. Según Mildenhall et al. (2020) un modelo puede pesar 5 MB, incluso menos que su nube de puntos. En

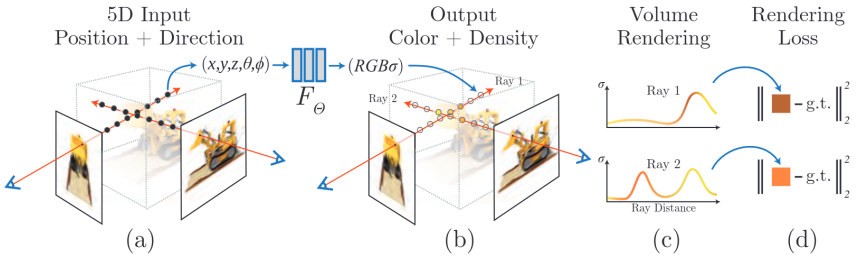


Figura 2.10: Flujo de trabajo de NeRF. (a) Muestreo de puntos a lo largo de un rayo de cámara y consulta a la MLP para obtener color y densidad. (b) Renderizado de una imagen completa mediante la integración volumétrica de los colores y densidades a lo largo de todos los rayos. (c) Detalle de la integración volumétrica. (d) Optimización de los pesos de la MLP minimizando el error fotométrico entre la imagen renderizada y la imagen real de entrenamiento. Imagen extraída de Mildenhall et al. (2020).

contrapartida, se requiere un proceso de entrenamiento de 12 hs[11] para una escena en particular. Sin embargo, una vez creada la red, es posible obtener el modelo 3D. A modo ilustrativo en la Fig. 2.11 se muestra un ejemplo de reconstrucción 3D con NeRF.

A modo informativo, en https://www.matthewtancik.com/nerf se pueden observar más ejemplos de reconstrucciones 3D con NeRF.

2.6. Medición activa

Teniendo en cuenta todos los métodos de reconstrucción 3D presentados en las secciones anteriores, se puede observar que todos ellos tienen en común que el procesamiento se realiza *offline*; es decir, sucede luego de que toda la adquisición de datos se haya completado. En otras palabras, se adquieren los datos y luego se procesan. En la medición activa (también conocida como sensado activo o percepción activa), en cambio, se busca guiar el proceso de adquisición de datos de forma inteligente, seleccionando acciones o puntos de vista que se espera aporten la

[11] Utilizando una idea parecida a NeRF, se puede reducir este tiempo notablemente Fridovich-Keil et al. (2022)

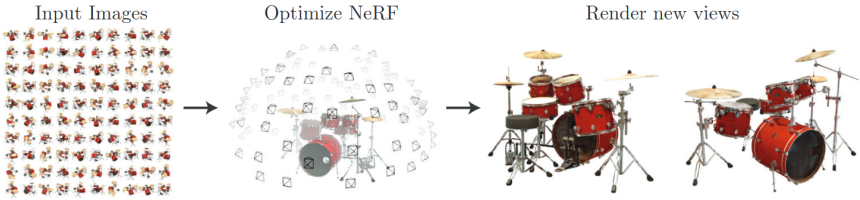


Figura 2.11: Ejemplo de reconstrucción 3D con NeRF. Se utilizaron 100 imágenes tomadas aleatoriamente sobre un casquete esférico. Luego del entrenamiento, se obtuvieron dos vistas nuevas. Imagen extraída de Mildenhall et al. (2020).

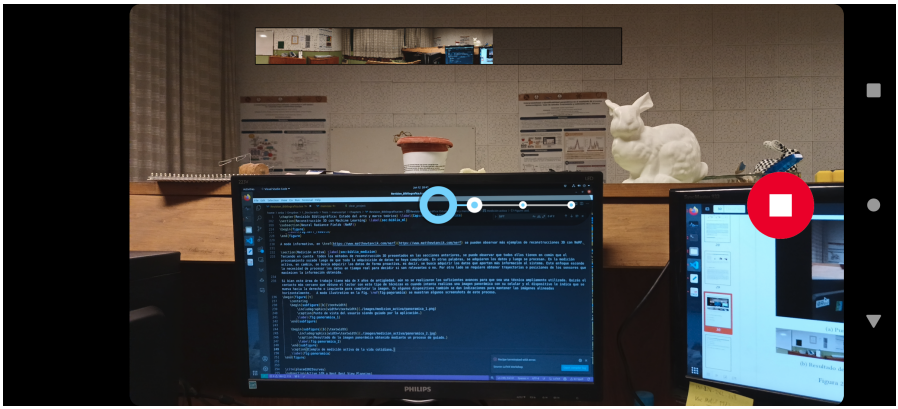
mayor cantidad de información relevante al sistema o que mejoren eficientemente la calidad del modelo. Este enfoque implica la necesidad de procesar los datos en tiempo real para decidir si son relevantes o no. Por otro lado, se requiere obtener trayectorias o posiciones de los sensores que maximicen la información obtenida. Hoy en día, el procesamiento en línea para la medición activa a menudo se realiza de forma simplificada debido a las restricciones computacionales, y aun así es un desafío implementarlo robustamente en simulación y, especialmente, en entornos reales.

Si bien esta área de trabajo tiene más de tres décadas de antigüedad[12], aún no se realizaron los suficientes avances para que sea una técnica ampliamente utilizada. Quizás el contacto más cercano que obtuvo el lector con este tipo de técnicas es cuando intenta realizar una imagen panorámica con su celular y el dispositivo le indica que se mueva hacia un lado u el otro para completar la imagen. En algunos dispositivos también se dan indicaciones para mantener las imágenes alineadas tanto como vertical como horizontalmente. Además, se informa al usuario cuando ocurrió un error y se requiere volver a tomar los datos. A modo ilustrativo, en la Fig. 2.12 se muestran algunos screenshots de este proceso.

Usualmente, este enfoque requiere de un sensor y de un agente que se desplace por el ambiente. En este segundo elemento es donde encontramos la mayor dificultad, ya que hoy en día solo hay algunos casos puntuales en donde un sistema pueda desplazarse de forma autónoma y segura en un ambiente complejo como el

[12] 27 años tomando como punto de partida el trabajo de Yamauchi (1997), ó 39 años considerando que la idea de visión activa apareció por primera vez en Connolly (1985), ó 22 años desde que se adoptó el término "Active SLAM" Davison and Murray (2002)

[13] Notar la autorreferencia, la panorámica involucra el mismo texto que usted está leyendo.



(a) Punto de vista del usuario siendo guiado por la aplicación.



(b) Resultado de la imagen panorámica obtenido mediante un proceso de guiado.

Figura 2.12: Ejemplo de medición activa de la vida cotidiana[13].

de la agricultura. En general, los sistemas de medición activa se utilizan en entornos controlados como laboratorios o en simulaciones.

A lo largo de esta sección se realizará una revisión de trabajos en los cuales se involucren alguna estrategia de sensado activo. Como fuente de información se utilizaron dos artículos que presentan un estudio exhaustivo del tema Lluvia et al. (2021); Placed et al. (2023).

2.6.1. Active SfM o Next Best View Planning

2.6.2. Active SLAM

En Lluvia et al. (2021) se realiza una revisión de los métodos de Active SLAM realizados en interior, con entorno estático y con un único robot. Se analizaron y se tabularon 23 trabajos publicados entre 2015 y 2020 en los que se aplica alguna técnica de Active SLAM. Las áreas segmentadas fueron: i) robot y sensor,

ii) representación del mundo, iii) técnica principal implementada, iv) objetivo de optimización y por último v) método de verificación: simulación o entorno real. En función de estos datos se realizó una estadística de los trabajos seleccionados:[14]

- Robot y sensor: El 52 % utilizó un robot terrestre, 39 % vehículo aéreo y 4 % vehículo submarino. Por otro lado 56 % de los sistemas empleó LiDAR (mayoritariamente 2D), 39 % cámaras (mayoritariamente RGB-D) y 13 % sonar.

- Representación del mundo: El 87 % utilizó mapa grillado como Octomap (mapa grillado en 3D) y 22 % utilizó otro tipo de mapa.

- Técnica Principal: Se observó una gran variedad de técnicas, pero predominaron enfoques de planificación basados en la selección NBV, a menudo utilizando *Rapidly-exploring Random Trees* (RRT) LaValle and Kuffner Jr (2001) para la generación de trayectorias hacia las fronteras entre el espacio explorado y el no explorado, consideradas como objetivos de exploración.

- Objetivo de optimización: El 96 % optimizan la posición deseada y 35 % optimizan la trayectoria.

- Método de verificación: El 78 % de los trabajos presentaron resultados en simulación y el 87 % en experimentos con robots en entornos reales.

A modo informativo, en Lluvia et al. (2021) se concluye que los criterios de finalización utilizados por los trabajos son muy variados y que no existe un criterio unificado. Por ejemplo se utiliza el número de escaneos, el tiempo transcurrido, la no existencia de las fronteras, la distancia recorrida, y otros que son muy dependientes del entorno como son el porcentaje del mapa explorado, los errores de estimación comparado con la posición real y más. En general todos los trabajos, de alguna forma u otra, utilizaban las fronteras entre lo explorado y lo desconocido como objetivos de exploración. Esta idea fue propuesta por primera vez por Yamauchi (1997) y se basa en la idea de que las fronteras son los lugares donde se puede obtener más información del entorno.

Por otro lado se considera especialmente a los trabajos realizados por Zhang (2020); Zhang and Scaramuzza (2020, 2019, 2018) por su innovación utilizando la técnica de *Receding Horizon* (RH) para la planificación de trayectorias en entornos

[14] La suma puede superar el 100 % ya que las categorías no son mutuamente excluyentes

3D con el objetivo de una localización precisa y evitando obstáculos usando estrategias de visión activa.

A modo ilustrativo en la Fig. 2.13 se observa un escenario de simulación propuesto por Zhang and Scaramuzza (2018). En este escenario, se busca que el *Micro Aerial Vehicle* (MAV) se traslade desde la esquina superior izquierda hasta la inferior derecha. Para esto, se presenta un escenario heterogéneo, con regiones de alta textura (generada por césped y rocas) y otras áreas con poca textura (representadas por un color oscuro). Las variables que el sistema de planificación utiliza para optimizar la trayectoria futura son: i) la probabilidad de colisión, ii) la calidad de la estimación del sistema de posicionamiento (incertidumbre de la pose), y iii) la distancia entre el MAV y el objetivo. A modo comparativo, se utilizan dos métodos: el llamado "activo" (azul), que incluye las tres variables a optimizar, y el denominado "reactivo" (rojo), el cual solo considera la probabilidad de colisión y la distancia al objetivo. Para este escenario, se observa que los errores en la estimación de posición son menores en el método "activo" que en el "reactivo". Además, con el método "activo", se prioriza navegar por zonas con abundante textura visual y así disminuir los errores de estimación de la localización.

Para calcular la calidad de las estimaciones del sistema de posicionamiento, Zhang et al. Zhang and Scaramuzza (2018) proponen lo siguiente: primero, para una trayectoria candidata (secuencia de poses), se identifican los *landmarks* del mapa que serían observables desde dichas poses. A continuación, se calcula (o aproxima) la matriz de covarianza de la pose, Σ, que resultaría de la observación de estos *landmarks* (e.g., mediante una linealización del proceso de triangulación o de la actualización del filtro). Por último, utiliza el criterio de optimalidad D (*D-optimality criterion*), referenciado como D-opt en el contexto de diseño óptimo de experimentos Carrillo et al. (2012), para obtener un escalar que resuma la incertidumbre, denominado R_{perc}. El criterio D-opt utilizado en este caso está dado por:

$$R_{\text{perc}} = k_{\text{perc}} \exp\left(-\frac{\log|\Sigma|}{6J}\right) \tag{2.1}$$

siendo k_{perc} una constante de ponderación y J la cantidad de poses a estimar en dicha trayectoria. Se menciona este criterio en particular ya que se utiliza comúnmente en la literatura de Active SLAM para cuantificar la información esperada de una acción.

Por otro lado en Placed et al. (2023), además de realizar una revisión bibliográfica, se realizan afirmaciones valorativas que ayudan a darle cohesión el conocimiento disperso en un solo trabajo. En este sentido, se extraerán una serie de

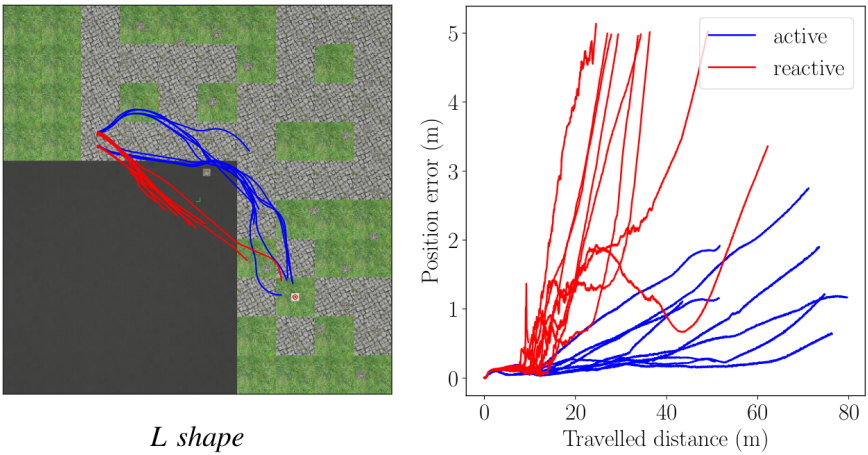


Figura 2.13: Resultados de simulación de Zhang and Scaramuzza (2018) comparando un planificador de MAV con reconocimiento de la percepción (activo, azul) frente a uno puramente reactivo (rojo). Izquierda: Vista cenital del entorno y trayectorias ejecutadas. El planificador activo dirige al MAV a través de regiones con mayor información visual. Derecha: Evolución del error de estimación de la posición, demostrando la superioridad del enfoque activo en este escenario. Imagen extraída de Zhang and Scaramuzza (2018).

declaraciones que fortalecerán los argumentos de esta tesis.

Funciones de utilidad

El trabajo de Yamauchi (1997) introdujo el concepto de "fronteras de exploración" que separan lo explorado de lo no explorado. Hoy en día, es una de las estrategias más utilizadas, especialmente para la exploración pura. Sin embargo, años más tarde, en Newman et al. (2003) y Stachniss et al. (2004), se observó que las regiones capaces de cerrar lazos contienen más información que las fronteras de exploración, especialmente cuando hay alta incertidumbre en la pose del agente. Esta observación ordena el objetivo de los sistemas de SLAM Activo, no solo a explorar todo el área lo más rápido posible, sino también mejorar la precisión de la localización y del mapeo. En sintonía con esto, en Stachniss et al. (2005) se demostró que para alcanzar estos objetivos de forma óptima, el funcional de costo debe tener una combinación de la distancia recorrida y de la información que se obtiene al realizar dicho recorrido. Sin embargo, debido a discrepancias entre los múltiples términos del funcional de costo, se requiere un ajuste manual en los factores de ponderación Carrillo et al. (2018).

La estrategia más común para calcular la cantidad de información del mapa es mediante el cálculo de su entropía. Sin embargo, a la hora de decidir qué trayectoria ejecutar, lo que interesa es seleccionar aquella que maximice la reducción de la entropía de todo el conjunto (mapa y agente). A este enfoque se lo conoce como *Mutual Information* (MI) que en este contexto puede interpretarse como la diferencia entre la entropía actual del sistema (mapa y pose) y la entropía esperada luego de ejecutar una acción y obtener una medición. Por otro lado, existe otra forma de medir este cambio de entropía, mediante la divergencia Kullback-Leibler, también conocida como *entropía relativa*.

Otro enfoque para computar la cantidad de información del mapa y del agente es mediante el cálculo de la matriz de covarianza (Σ) de las variables de interés. Estas variables pueden ser la posición del agente y de los landmarks del mapa. Con esta idea, para comparar el impacto de diferentes acciones futuras en la incertidumbre del estado, se proponen criterios de optimalidad que son funciones escalares de Σ, como su traza (A-optimality), determinante (D-optimality, relacionado con el volumen del elipsoide de incertidumbre), o el máximo/mínimo autovalor (E-optimality). Este enfoque es intuitivo, ya que se espera que las tareas de exploración inicialmente aumenten la incertidumbre, y que las acciones informativas la reduzcan. Sin embargo, la relación entre la reducción de la incertidumbre predicha y la mejora real del mapa no siempre es directa, y

depende de cómo el error y la incertidumbre son representados y propagados por el sistema SLAM subyacente Kim and Kim (2017); Rodríguez-Arévalo et al. (2018). Por esto, es fundamental prestar atención a la representación y consistencia de la incertidumbre en un sistema de SLAM activo, ya que, de no hacerlo, se pueden tomar decisiones erróneas. Además, los factores de ponderación utilizados en las funciones de utilidad son un punto vulnerable que puede comprometer la robustez del sistema.

Obtención de trayectorias

Una vez seleccionado el objetivo de exploración, ya sea para cerrar un lazo, realizar tareas de exploración o ambas, se requiere obtener la trayectoria que va desde el punto en donde se encuentra el agente hasta el objetivo en cuestión. Esta trayectoria debe ser aquella que optimice un funcional de costo. Para obtenerla, se pueden usar estrategias de muestreo como RRT LaValle and Kuffner Jr (2001), entre otros Karaman and Frazzoli (2011).

Sin embargo, este enfoque secuencial de "seleccionar objetivo y luego planificar trayectoria" es a menudo subóptimo, ya que quizás exista otro conjunto de objetivos de exploración y trayectorias asociadas que generen un mejor resultado global del sistema. Esto se debe a que se trata de forma desacoplada la selección del objetivo de exploración y la planificación de su trayectoria.

Una idea para alcanzar soluciones globalmente óptimas es modelar el problema como un Proceso de Decisión de Markov Parcialmente Observable (POMDP por sus siglas en inglés). Un POMDP es un modelo matemático que describe cómo tomar decisiones secuenciales en situaciones donde el estado del sistema no se puede observar de manera directa o completa. En un POMDP, el agente no tiene acceso directo al estado actual del entorno, sino que debe hacer inferencias basadas en observaciones parciales o ruidosas. El POMDP es una generalización de un Proceso de Decisión de Markov (MDP, por sus siglas en inglés), donde en un MDP el estado del sistema es completamente observable por el tomador de decisiones. En un POMDP, cada acción tomada por el agente no solo afecta el estado del sistema, sino que también contribuye a la información obtenida por el agente, lo cual afecta sus estimaciones sobre el estado actual del sistema. La solución a un POMDP es una política que mapea creencias a acciones, buscando maximizar la recompensa acumulada esperada a largo plazo.

Si bien los POMDP sobre espacios continuos, como los que se encuentran en robótica, no se pueden resolver de forma exacta (son computacionalmente PSPACE-completos o incluso indecidibles en el caso general), pueden ser resueltos

de forma aproximada utilizando diversas técnicas, incluyendo métodos basados en muestreo, optimización de políticas, o simplificaciones como el control óptimo estocástico bajo ciertos supuestos. Si la función de recompensa (o costo) a optimizar solo depende de la matriz de covarianza de los estados (y se asumen modelos lineales y ruido gaussiano), entonces el problema de planificación de la trayectoria bajo incertidumbre se puede abordar con herramientas del Controlador Lineal Cuadrático Gaussiano (LQG, por sus siglas en inglés) y sus extensiones, como se explora en Koga et al. (2021, 2022), entre otros.

Conclusiones y preguntas de investigación

Esta revisión preliminar de las soluciones propuestas pone de manifiesto que aún persisten numerosos desafíos por abordar. En Cadena et al. (2016); Placed et al. (2023); Placed and Castellanos (2022) se identifica como un problema abierto persistente la selección de un criterio de parada autónomo y significativo. En la bibliografía es usual encontrar criterios de tiempo, de área cubierta, energía, fronteras de exploración, entre otros. Sin embargo, estos enfoques a menudo no pueden ser usados de manera generalizable en espacios completamente desconocidos, tanto en interiores (entornos cerrados) como en exteriores (entornos abiertos), o no garantizan que se haya alcanzado un nivel de calidad deseado en el mapa o la localización.

Por otro lado, debido a la naturaleza inherentemente interactiva con el entorno de Active SLAM, no es sencillo utilizar bases de datos o *datasets* como benchmark para comparar distintos sistemas de forma justa y completa, ya que la secuencia de acciones del robot afecta las observaciones futuras. Sumado a esto, no hay establecida una métrica universal de contraste que permita cuantificar la mejora de un sistema sobre otro, considerando simultáneamente la calidad del mapa, la eficiencia de la exploración y el costo de las acciones. Además, sigue siendo un desafío asegurar que los resultados obtenidos en simulación sean directamente replicables y generalizables en un escenario real, dada la dificultad de modelar con alta fidelidad las complejidades sensoriales y de interacción del mundo real.

Todos estos desafíos se reflejan en una bibliografía a menudo muy dispersa y poco unificada, tanto en terminología como en metodología. Esto dificulta que los algoritmos maduren a través de comparaciones rigurosas y que se obtengan resultados con un mayor impacto en el mundo real.

2.7. Reconstrucción 3D en la agricultura

En la agricultura el fin último es incrementar la productividad y la eficiencia de los cultivos. Por esto, cualquier técnica que permita obtener información relevante alineada con este objetivo es bienvenida. La reconstrucción tridimensional (3D) de plantas y entornos agrícolas emerge como una herramienta prometedora en este contexto.

Las técnicas de reconstrucción 3D aplicadas en la agricultura difieren de las observadas en la literatura propia de la reconstrucción 3D. En general se tiene un mayor interés sobre el grado de precisión de las mediciones, la robustez de los algoritmos, la velocidad de adquisición de datos, la facilidad de uso, entre otros. Sin embargo, es frecuente que no se cuente con un análisis riguroso sobre las incertidumbres manejadas, en general se recurre a un cálculo estadístico en donde se obtiene un factor de correlación entre las mediciones y la variable a estimar. Este enfoque pragmático y orientado a la validación agronómica tiene sentido, ya que pioneros de la estadística moderna como Ronald Fisher[15] realizó gran cantidad de sus aportes cuando era parte de la estación agronómica de Rothamsted, sentando las bases para el diseño experimental y el análisis de datos en la agricultura.

Al igual que se observó en las secciones pasadas sobre SLAM y SfM, la reconstrucción 3D en la agricultura utiliza datos provenientes de dos tipos de sensores: cámaras (monoculares, estéreo, infrarrojas y térmicas) y LiDARs (2D y 3D). Sin embargo, los LiDARs son los preferidos ya que son sensores activos, que pueden operar en condiciones de iluminación adversas, condiciones climáticas poco favorables, además, son robustos y compactos. Por este motivo, en la bibliografía específica de aplicaciones agrícolas se cuenta con una mayor cantidad de trabajos que utilizan LiDARs en comparación con aquellos que dependen exclusivamente de cámaras para la reconstrucción 3D.

Conocer ciertos aspectos relacionados con la estructura tridimensional de un cultivo de árboles presenta una serie de ventajas para la toma de decisiones agronómicas. Los parámetros más comunes encontrados en la bibliografía son el volumen de la canopia, el índice del área de las hojas (del inglés, *Leaf Area Index* LAI), la estructura interna de la planta (discriminando por ramas principales y ramas secundarias), porosidad (espacio entre plantas), el área foliar total o proyectada, entre otros. Cabe destacar que si bien estos parámetros se pueden calcular a partir de una representación 3D del árbol, el objetivo final es obtener indicadores útiles para la gestión del cultivo. Por esto es importante resaltar que los requisitos de calidad de

[15] Uno de los estadísticos más influyentes del siglo XX.

la reconstrucción (entendidos como precisión y completitud de la representación) estarán fuertemente influenciados por la sensibilidad y el impacto de los errores de reconstrucción en dichos indicadores. Estas características de los indicadores dependen del tipo del cultivo y de la etapa del ciclo de vida en la que se encuentre.

Según Rosell and Sanz (2012) la medición de las características geométricas de una canopia permite ajustar mejor la dosis de productos aplicados, como pesticidas y fertilizantes, mejorando así tanto el impacto ambiental como económico de estas aplicaciones. Además, la caracterización geométrica de las canopias es crucial para establecer estimaciones de las necesidades de agua de los cultivos y para obtener información que puede usarse para cuantificar sus requerimientos nutricionales. Por otro lado, sirve para estimar la salud y la velocidad de crecimiento del cultivo. Un cambio drástico en el volumen puede señalar problemas como enfermedades, estrés hídrico o nutricional. Además, la cantidad de luz que absorbe la planta (relacionado a su geometría) está directamente asociada con su capacidad para realizar fotosíntesis y por ende para crecer.

En general, la reconstrucción 3D utilizando únicamente cámaras pasivas no cuenta con los mismos beneficios de robustez y precisión de rango que ofrecen los LiDARs en entornos agrícolas complejos. Sin embargo, se puede complementar la información obtenida de un LiDAR con una cámara para obtener información adicional, como el color o índices espectrales. Por ejemplo, en Underwood et al. (2016) se utilizan cámaras para realizar un mapeo de flores y frutos, además de obtener el volumen con un LiDAR.

Por otro lado, la reconstrucción 3D de árboles y otros cultivos complejos presenta una serie de desafíos particulares. En general, en casi todos los casos, salvo quizás en algunas aplicaciones comerciales bien establecidas como la mencionada en Underwood et al. (2016) (que derivó en la empresa Green Atlas), los resultados publicados en la literatura de investigación a menudo no son fácilmente generalizables por fuera del estudio en sí mismo. Estas limitaciones pueden deberse a diversos factores, tales como la falta de automatización en la toma de datos, las limitaciones de velocidad de los vehículos para que la medición sea lo suficientemente precisa (hasta $11Km/h$ según Rivera et al. (2023) para ciertas aplicaciones LiDAR), las limitaciones energéticas de los vehículos (en caso de usar UAV), la falta de robustez de los algoritmos frente a la variabilidad de las condiciones de campo, las necesidades de postprocesamiento de los datos, o incluso la necesidad de adaptar el cultivo para realizar la medición, entre otros.

Se ha observado que muchas estrategias que utilicen métodos de propósito general, como los vistos en las anteriores secciones, no se comportan adecuadamente en los entornos agrícolas Capua and Sansoni (2018). En este

trabajo, se realizó un análisis comparativo de distintos sistemas de visual SLAM en entornos agrícolas y se concluye que, debido a la estructura y textura repetitiva de los cultivos con árboles, los algoritmos de visual SLAM convencionales basados en características no pueden estimar la posición de la cámara de forma precisa. Además, el ambiente cambia significativamente en las cabeceras de las filas, lo cual requiere ajustes particulares del sistema dependiendo de la posición del agente. Esto, por lo tanto, dificulta el cierre de lazos y, consecuentemente, la reducción de los errores de estimación acumulados.

Como alternativa a esto, hay soluciones prometedoras que proponen utilizar la estructura del cultivo como fuente de información para estimar la localización, por ejemplo, usando una cámara RGB-D Gimenez and Sansoni (2022). Sin embargo, aún son trabajos en desarrollo y no cuentan con implementaciones en la industria.

A continuación, se describirán algunos trabajos recientes[16] en donde se aplican técnicas de reconstrucción 3D en la agricultura. Los trabajos seleccionados se caracterizan por ser elementos representativos de una técnica, tendencia, campo de aplicación, entre otros.

En general, se pueden identificar tres tendencias principales en la bibliografía respecto al grado de automatización de la adquisición de datos. Por un lado, se encuentran trabajos que utilizan un instrumento de medición portátil o manual, a menudo acarreado por personas para tomar mediciones. En otros casos, los trabajos se centran en la adquisición automática de datos mediante vehículos con cierto grado de autonomía o teleoperados. Por último, se encuentran menos trabajos en donde el proceso está totalmente automatizado e integrado en flujos de trabajo agronómicos.

Dentro del primer grupo, que corresponde a instrumentos de medición portátiles o manuales, se encuentra el trabajo de Westling et al. (2020). En este estudio se utiliza una técnica de reconstrucción muy llamativa llamada GeoSLAM Zebedee 1 [17] Bosse et al. (2012) desarrollada por *Commonwealth Scientific and Industrial Research Organisation* (CSIRO). En la Fig. 2.14 se muestra el sistema de adquisición utilizado en el campo. Este sistema consiste en un LiDAR 2D montado sobre un resorte que induce oscilación, permitiendo que el sensor capture datos en 3D mientras el operador camina por el entorno. El SLAM subyacente procesa los datos del LiDAR y de una IMU integrada para estimar la trayectoria del sensor y generar la nube de puntos.

[16] Es importante tener en cuenta que los tiempos de adopción y desarrollo de tecnología en la agricultura pueden ser mayores a los encontrados en otras áreas de rápida evolución, como puede ser el *Machine Learning*.

[17] Un video de su modo de uso se puede encontrar en el siguiente link https://youtu.be/DUEAz_naHHg.



Figura 2.14: Sistema de adquisición GeoSLAM Zebedee 1. Imagen extraída de Westling et al. (2020).

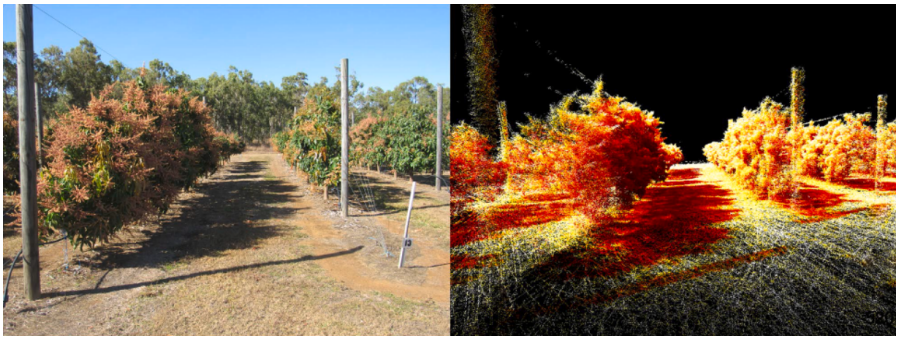


Figura 2.15: Ejemplo de aplicación de la reconstrucción 3D aplicado a la estimación de luz capturada por la canopia. A la izquierda se observa una fotografía del cultivo y a la derecha se encuentra una reconstrucción 3D del mismo en donde se simula la proyección de las sombras en la misma fecha del año y horario que la imagen de la izquierda. Imagen extraída de Westling et al. (2020).

Con esta estrategia se logra obtener una nube de puntos 3D del entorno. El objetivo del trabajo es determinar la intercepción de la radiación fotosintéticamente activa (PAR) por la canopia y así proponer una alternativa al uso de ceptómetros, que son instrumentos tradicionales para la medición de la cantidad de luz que llega a la planta. Estos instrumentos, si bien son precisos, son lentos (10 minutos por árbol) y requieren que las condiciones ambientales sean acordes para efectuar la medición (ej.: días soleados). Utilizando la reconstrucción 3D de la canopia, se puede realizar una simulación de trazado de rayos para conocer la distribución de la luz que absorbe la canopia. En la Fig. 2.15 se muestra un resultado cualitativo comparando las sombras de los árboles reales y las sombras simuladas por el modelo.

Con este enfoque Westling et al. (2020) reporta una correlación R de 0.899 entre las mediciones obtenidas con el ceptómetro y las simulaciones realizadas con la reconstrucción 3D. Además se obtiene una reducción del tiempo de medición de 10 minutos a 1 minuto por árbol. Se observa que los mayores errores se cometen cuando la densidad de la canopia es baja. Adicionalmente, se obtuvieron otros parámetros de los árboles como el volumen de copa, el cual reportó una correlación R de 0.8074 con estimaciones consideradas precisas como *partition hull* Cheein and Guivant (2014). Por último, el autor destaca que no es posible reconstruir con detalle la estructura interna de los árboles (ramas) con este sistema ya que la precisión y

densidad de la reconstrucción no son suficientes para tal fin.

La siguiente tendencia, que involucra la adquisición de datos mediante vehículos terrestres equipados con múltiples sensores, se observa en el trabajo de Underwood et al. (2016). Este sistema consiste en vehículo terrestre (tractor, cuatriciclo, etc) en el cual se montan distintos sensores como GNSS, LiDARs y cámaras. El objetivo de este trabajo es realizar un mapeo del volumen de la canopia, así como de la cantidad de flores y frutos en un cultivo de almendros. Este tipo de sistemas es relevante ya que permite la adquisición de datos a gran velocidad, lo cual solventa un cuello de botella existente entre los resultados obtenidos en investigaciones de pequeña escala o de laboratorio y su aplicación posterior a la industria. En base a este trabajo James Underwood en 2018 co-fundó la empresa *Green Atlas*[18] que se dedica a la recolección de datos en cultivos intensivos, teniendo presencia en Australia, Estados unidos, Chile, Brasil, Francia, Italia y Turquía.

Básicamente mediante el posicionamiento GNSS, asocia los datos capturados por los diferentes sensores a una posición geográfica en el campo. Las imágenes son capturadas en momentos estratégicos sincronizadas con la posición del vehículo. Luego mediante un proceso de segmentación de las imágenes se identifican y cuantifican las flores y los frutos permitiendo obtener un conteo por árbol. Además, con las mediciones del LiDAR, y luego de segmentar los puntos correspondientes a cada árbol, se calcula el volumen de la canopia. Los árboles se miden primero desde un lado y luego desde el otro; seguido a esto, los datos se alinean para obtener una nube de puntos completa. Para el cálculo del volumen, en este trabajo se discretiza la nube de puntos usando *voxels* de 10 cm de lado.

Para este trabajo en total se midieron 580 árboles en 3 temporadas (floración, cuajado de frutos y precosecha). Si bien se alcanzaron correlaciones muy altas comparadas con las mediciones manuales, el autor reporta que cuando las plantas tienen mucho follage, la señal del GNSS puede volverse inestable o imprecisa, provocando problemas en la sincronización de los datos y la posterior reconstrucción. Además, se observa que no se puede segmentar fácilmente las nubes de puntos de los árboles, ya que se mezclan con las canopias aledañas. En la Fig. 2.16 se muestra el sistema de adquisición utilizado, en conjunto con algunas muestras de los datos recolectados[19].

Por otro lado, además de los problemas generados con el GNSS, se observaron problemas en la segmentación de las imágenes debido a condiciones variables de iluminación y otros factores imprevistos. A la hora de realizar el análisis de los

[18] https://greenatlas.com/

[19] Ver el siguiente link para más detalles https://youtu.be/j-gJQXQoqX0

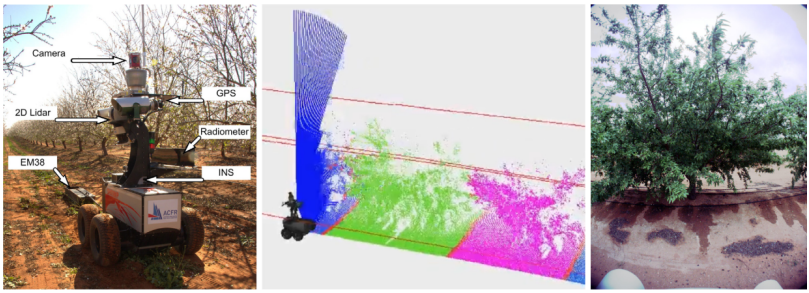


Figura 2.16: Plataforma de adquisición de datos y ejemplos de salida del sistema desarrollado en Underwood et al. (2016) para el mapeo de huertos de almendros. Izquierda: Vehículo terrestre equipado con múltiples sensores, incluyendo LiDAR 2D, cámara, GPS/INS y radiómetro. Centro: Visualización de una nube de puntos 3D parcial del cultivo, generada a partir de los datos del LiDAR. Derecha: Imagen representativa capturada por la cámara del sistema, mostrando un árbol individual. Imagen extraída de Underwood et al. (2016).

resultados, se detectó la presencia de valores atípicos o *outliers* en los datos por lo que se tuvo que realizar un proceso de limpieza de los mismos. Cabe destacar que todo el procesamiento se realiza offline; de aquí surge la necesidad de poder corregir o validar los datos adquiridos en tiempo real ya que muchas veces no se puede volver a realizar la medición. Además se menciona la necesidad de cambiar los puntos de vista de las imágenes de forma ad-hoc debido a la estructura compleja de las canopias para asegurar una cobertura adecuada.

La adquisición de datos de forma totalmente autónoma en agricultura sigue siendo, en muchos aspectos, una utopía, sin embargo, de la mano del sensado activo y la colaboración multi-robot, hay trabajos que se acercan a este objetivo. En particular, dentro del proyecto Flourish Pretto et al. (2020), se puede mencionar el trabajo de AgriColMap Potena et al. (2019) en el campo del mapeo colaborativo aire-tierra, y los trabajos de Popovic et al. (2016); Popović et al. (2020) en el área de clasificación activa para la detección de malezas.

En el caso de Potena et al. (2019), se busca realizar la reconstrucción del entorno agrícola con imágenes tomadas desde puntos de vista muy distintos: específicamente, imágenes tomadas desde un Vehículo Terrestre No Tripulado (UGV) y desde un Vehículo Aéreo No Tripulado (UAV). Este enfoque tiene el desafío

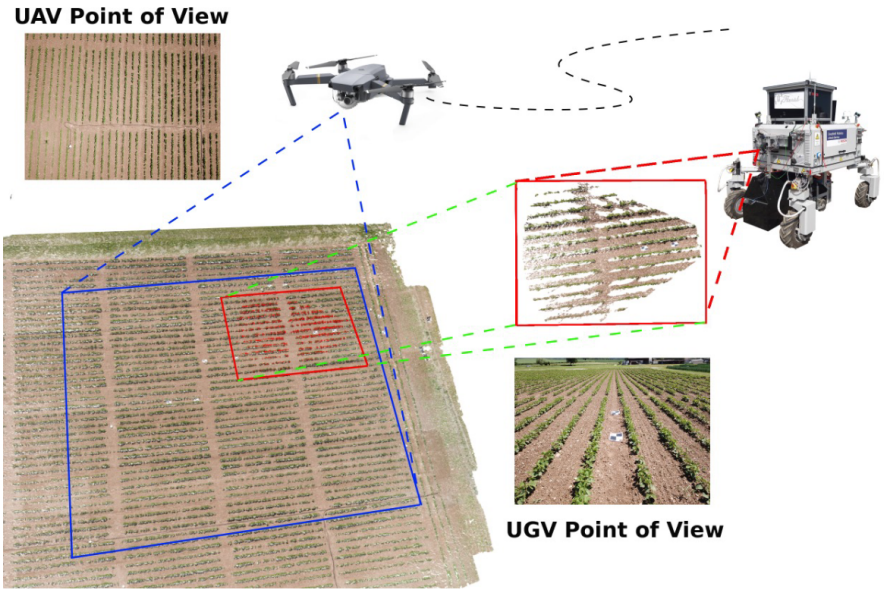


Figura 2.17: Esquema de la propuesta de AgriColMap. Imagen extraída de Potena et al. (2019).

de obtener las correspondencias y fusionar la información en un solo mapa. Para esto, se utiliza una técnica conocida como *Large displacement Dense Optical Flow* (LDOF) para encontrar correspondencias iniciales, combinada con información semántica del cultivo como el perfil de alturas de las plantas y el índice *Excess Green* (ExG)[20], en conjunto con información GNSS y técnicas de SfM para el refinamiento final del mapa. En la Fig. 2.17 se muestra un esquema de la propuesta.

Si bien Potena et al. (2019) está enfocado a cultivos extensivos (ej., remolacha azucarera), la problemática que aborda (fusión de datos de múltiples plataformas con diferentes perspectivas) es similar a la de los cultivos con árboles. Sin embargo, en un ambiente con árboles, la información que se obtiene desde un UGV es distinta a la que puede obtener un UAV desde una vista cenital u oblicua, siendo éstas muchas veces complementarias pero con regiones de no solapamiento que

[20]Obtenido como $ExG = 2G - R - B$, donde R, G, B son los valores de los canales de color del píxel.

dificultan la correspondencia directa.

Por otro lado, en Popovic et al. (2016) y en su trabajo posterior más generalizado Popović et al. (2020) se propone un sistema de planificación de trayectorias para UAVs tal que permita clasificar celdas de un mapa 2D del terreno entre dos clases: *con maleza* y *sin maleza*. Esta clasificación se basa en imágenes de la vegetación recogidas por el UAV, el cual planifica de forma adaptativa sus rutas para maximizar, en función del tiempo, la recolección de datos informativos sobre las áreas de interés. La planificación es del tipo de horizonte fijo y los objetivos de exploración son un conjunto finito de *waypoints* predefinidos o generados dinámicamente. La autora reporta una mejora del $50\,\%$ en la reducción de la entropía del mapa de clasificación de malezas en comparación con un enfoque de planificación de trayectorias estandard [21].

2.8. Conclusiones, oportunidades y desafíos

En esta extensa revisión bibliográfica, se han presentado diversos algoritmos y sistemas de reconstrucción 3D que, en muchos casos, poseen una gran precisión cuando se comparan con los valores reales en condiciones controladas o para métricas específicas. Sin embargo, un tema recurrente y llamativo a lo largo de las secciones es el limitado o nulo tratamiento sobre la propagación y cuantificación rigurosa de incertidumbres. Según la *Guide to the Expression of Uncertainty in Measurement* (GUM) Joint Committee for Guides in Metrology, (JCGM/WG 1) (1995), una medida solo es una aproximación del valor real, y esta es considerada completa solamente cuando su incertidumbre asociada es estimada. En este sentido, la incertidumbre es una medida de la calidad de la información y resulta fundamental para la toma de decisiones informadas y robustas. La carencia de un manejo explícito de la incertidumbre en muchos sistemas de reconstrucción limita su fiabilidad y la capacidad de evaluar la confianza en los resultados obtenidos.

Además, en el ámbito de la medición activa y el Active SLAM, se resalta la necesidad de aunar esfuerzos en la unificación de criterios de evaluación, el desarrollo de métricas estandarizadas, y en la utilización de *benchmarks* realistas y reproducibles para la comparación objetiva de algoritmos.

Por otro lado se observa que en la agricultura es fundamental asegurarse que los datos tomados son útiles para su propósito *en el momento en el que se están tomando*, ya que, de lo contrario, es altamente probable que no se pueda volver a medir o que

[21] En el trabajo se lo identifica como "lawnmower", pero hace referencia a las trayectorias típicas que se realizan en el corte de pasto y que son ampliamente utilizadas en las planificaciones de vuelo de UAVs.

la oportunidad de hacerlo en condiciones comparables se pierda. Esto sucede por varias razones: en primer lugar, los cultivos cambian dinámicamente con el tiempo; en segundo lugar, las condiciones ambientales pueden variar; en tercer lugar, dado que muchos flujos de trabajo actuales implican que los datos son procesados *offline* y, generalmente las zonas de medición están alejadas de los centros de cómputo, resulta costoso y poco práctico volver a medir. Se podrían seguir agregando motivos, pero, en general, se necesita la capacidad de evaluar la utilidad del dato medido en línea, sin necesidad de realizar todo el procesamiento posterior. Este requisito subraya la importancia de la medición activa y el procesamiento en el borde (*edge computing*) en aplicaciones agrícolas.

Por último se destacan los beneficios de la adquisición colaborativa de datos (ej., mediante múltiples robots o plataformas heterogéneas), ya que estas estrategias apuntan a resolver los desafíos de la recopilación de datos en grandes extensiones de terreno o en entornos complejos que requieren múltiples perspectivas. Esto se puede realizar ya sea desde distintos tipos de vehículos como con la utilización de enjambres o *swarms* de robots. En este sentido, se observa que la colaboración entre robots es una tendencia en alza en la literatura. Sin embargo, aún no se cuenta con un sistema ampliamente disponible y robusto que pueda realizar mediciones de forma plenamente autónoma ni colaborativa en un entorno agrícola desafiante y variable. Con estos desafíos identificados, se destaca que existe un punto de contacto y una sinergia potencial significativa entre la vigilancia de cultivos (monitorización continua y a gran escala) y las técnicas de sensado activo para optimizar dicha vigilancia. Esta intersección representa una oportunidad clave para futuras investigaciones, incluyendo la que se propone en esta tesis.

Capítulo 3

Desarrollo Teórico

3.1 Organización del capítulo 53
3.2 Modelado de la incertidumbre 54
3.2.1 Probabilidades de dispersión e incertidumbre 55
3.3 Proceso de Actualización 61
3.3.1 Actualización mediante Media Ponderada 62
3.3.2 Actualización mediante Filtro de Kalman 65
3.4 Métrica de evaluación basada en Incertidumbre 67
3.5 Fronteras de Incertidumbre: Objetivos de exploración 72
3.6 Análisis del costo computacional 73

3.1. Organización del capítulo

En este capítulo se desarrollan los conceptos teóricos que representan el núcleo del aporte de esta tesis. Se comienza con la definición de los conceptos de incertidumbre y su modelado. Luego se presenta el proceso de actualización de la incertidumbre utilizando estrategias de actualización Bayesiana y filtro de Kalman. Posteriormente se presenta una métrica de evaluación basada en la incertidumbre para la exploración de un entorno. Finalmente se presenta el concepto de umbralizar el mapa de incertidumbre y su relación con la exploración de un entorno. La información presentada en este capítulo también se encuentra en Sansoni and Gimenez (2025) y Sansoni and Gimenez (2024).

3.2. Modelado de la incertidumbre

Como se observó en la Sec. 2.6, más del $80\,\%$ de los trabajos analizados utilizan mapas grillados. En primer lugar, cuando se utiliza un mapa grillado la información contenida en él se suele obtener mediante la entropía de Shannon Stachniss et al. (2005); Stachniss (2006). En forma discreta esta entropía $H(\mathbf{X})$ se calcula según:

$$H(\mathbf{X}) = -r \sum_{k=1}^{n} \left(p_{t,k} \log p_{t,k} + (1 - p_{t,k}) \log(1 - p_{t,k})\right) \tag{3.1}$$

siendo n la cantidad de celdas en el mapa, $p_{t,k}$ la probabilidad de ocupación de la celda k en el instante t y r el área ocupada por cada celda. Notar que el valor obtenido con la Ec. 3.1 es dependiente del tamaño de las celdas del mapa y que además todas las celdas suman, sin importar si fueron exploradas o no Blanco et al. (2008). Este es un punto no menor ya que, de no tenerlo en cuenta, no es posible comparar dos métodos que reconstruyeron el mismo ambiente, pero con distinto tamaño de celda y distinto tamaño de mapa. Por otro lado, el uso de esta entropía requiere el cálculo y actualización de la probabilidad de ocupación de cada celda en el mapa. Esta variable se mantiene actualizada mediante la regla de Bayes utilizando los *log-odds*[1] Moravec (1989), dada por:

$$\ell_{t,k} = \ell_{t-1,k} + \ell_{\text{new},k} \tag{3.2}$$

con

$$\ell_{t,k} = \ln \frac{p_{t,k}}{1 - p_{t,k}} \;\Leftrightarrow\; p_{t,k} = 1 - \frac{1}{1 + e^{\ell_{t,k}}} \tag{3.3}$$

donde $\ell_{\text{new},k}$ es la nueva información obtenida sobre la celda k en el tiempo t. Para una explicación detallada de la relación entre las probabilidades y los *log-odds* se puede consultar Stachniss (2009); Thrun (2003); Moravec (1989).

Esta regla de actualización requiere asumir que las celdas son independientes y que la posición del robot es conocida al momento de realizar la medición. Es común, además, que la actualización de estas probabilidades se efectúe mediante dos valores fijos, l_{hit} o l_{miss}, lo cual impide una consideración real de la incertidumbre inherente a la medición. Este enfoque es el que utiliza Octomap Hornung et al. (2013) y su justificación está basada en la eficiencia computacional, sin embargo

[1] Es decir, a partir de la probabilidad $p_{t,k}$ se obtiene su correspondiente log-odd $\ell_{t,k}$, luego se realiza el proceso de actualización obteniendo así $\ell_{t+1,k}$ y por último se recupera con la función inversa el valor de $p_{t+1,k}$.

es una simplificación que atenta contra los algoritmos de planificación basados en incertidumbre. Estas observaciones son un lugar común, por lo que existen distintas alternativas propuestas. Una de ellas es O'Callaghan and Ramos (2012), en donde se utiliza un *Proceso Gaussiano* como clasificador de las celdas (entre ocupadas o libres), de esta manera se evita tener que asumir independencia de los elementos de la grilla. Un interesante resultado de este tipo de implementación es que se puede obtener una estima de la incertidumbre de la inferencia. Sin embargo, una crítica común que se le hace a este enfoque es que el costo computacional para el ajuste de los hiper parámetros mediante la maximización de la verosimilitud marginal es de $O(N^3)$ con N la cantidad de muestras cuando se invierte una matriz del kernel y de $O(kN^2)$ con k la cantidad de hiper parámetros cuando se calcula el gradiente [Murphy (2012), ch. 15.2.4]. Sin embargo, en O'Callaghan and Ramos (2012) se proponen algunas alternativas[2] para reducir el costo computacional. Aquí entra en juego el equilibrio que siempre aparece en la robótica: precisión vs costo computacional. A los fines prácticos se busca tener una representación del entorno que permita tomar decisiones lo suficientemente cercanas a la decisión óptima.

Obtener $\ell_{\text{new},k}$ (y por consiguiente $p_{\text{new},k}$) no es un proceso trivial y depende del sensor utilizado. En este sentido, como se está buscando obtener resultados generales, sin importar la tecnología utilizada y que además permita comparar distintos resultados de los sistemas, se propone utilizar un nuevo enfoque mediante las llamadas *Probabilidades de Dispersión*. Este enfoque no solo sortea los problemas identificados con la entropía de Shannon, sino que también es compatible con una amplia gama de distribuciones de probabilidad, no limitándose a la Bernoulli, incluyendo continuas, discretas y multivariables. Adicionalmente, establece una relación con el criterio D-opt (mencionado en la Sec. 2.6.2), lo que permite unificar los conceptos de incertidumbre del mapa, del agente y de las trayectorias.

3.2.1. Probabilidades de dispersión e incertidumbre

Se entiende por incertidumbre a la dispersión de una variable con respecto a su medida de tendencia central, como la media, la mediana, o moda, entre otros. Si las mediciones son modeladas con una variable aleatoria unimodal, entonces estas variables están alrededor de su media o moda con una probabilidad inversamente proporcional a la medida de dispersión. La desigualdad propuesta por Chebyshev y luego mejorada por Gauss Van Parys et al. (2016), provee una cota inferior para la probabilidad de que una variable aleatoria se encuentre a una cierta distancia de

[2] Como subsamplear los datos.

su media. Sea X una variable aleatoria (VA) con varianza σ^2 y su media igual a su moda μ, entonces esta desigualdad resulta[3]:

$$P\left(|X-\mu|<\frac{s}{2}\right)\geq\frac{s}{2\sigma\sqrt{3}},\quad\text{para}\quad 0\leq s\leq\frac{4\sigma}{\sqrt{3}}. \tag{3.4}$$

donde s es una constante positiva que representa la longitud del intervalo $(\mu - s/2, \mu + s/2)$ alrededor de la media μ. Este cuantificador de la concentración de X alrededor de μ se denomina *Probabilidad de Dispersión* (PD). A modo de ejemplo, cuando $X \sim \mathcal{N}(\mu, \sigma^2)$ se tiene que:

$$P\left(|X-\mu|<\frac{s}{2}\right)=2\Phi\left(\frac{s}{2\sigma}\right)-1 \tag{3.5}$$

donde Φ es la función de distribución acumulada de la normal estándar. En la Sec. A.2.2 se muestra con más detalle el procedimiento para obtener la expresión anterior. Cuando $\mathbf{X} \in \mathbb{R}^N$, $\mathbf{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$, y considerando el hiper rectángulo A con lados s_i dado por:

$$A=\{(x_1,x_2,\ldots,x_N)\in\mathbb{R}^N : |x_i-\mu_i|<s_i/2\ ,\forall i=1,\ldots,N\}, \tag{3.6}$$

entonces la probabilidad asociada a la incertidumbre de esta variable aleatoria es:

$$P(\mathbf{X}\in A)=\int_A \mathcal{N}(\mu,\mathbf{\Sigma})(w)dw. \tag{3.7}$$

Notar que cuando la función de distribución es conocida, la PD se puede calcular analíticamente y, si ese no fuera el caso, se puede obtener mediante simulaciones Monte Carlo (MC). Por fuera de esta observación, este enfoque no se limita a distribuciones normales, pueden ser distribuciones no paramétricas, en donde la probabilidad se define como la probabilidad de que el evento A suceda.

Una representación gráfica de este concepto se muestra en la Fig. 3.1. Los histogramas 2D de la distribución de la variable aleatoria gaussiana $(x_1, x_2) \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ se muestran tanto en el lado izquierdo como en el derecho de la figura. Para longitudes dadas (s_1, s_2), se define el intervalo A en ambos histogramas. En el lado izquierdo, la dispersión de las variables es menor que en el lado derecho, lo que hace más probable que (x_1, x_2) caiga en A en comparación con el histograma del lado derecho. Esta mayor probabilidad se debe a la relación inversa entre la dispersión de las variables y la probabilidad de que ocurra el evento A. Los parámetros

[3] En la Sec. A.2.1, se deriva esta desigualdad desde la desigualdad de Chebyshev.

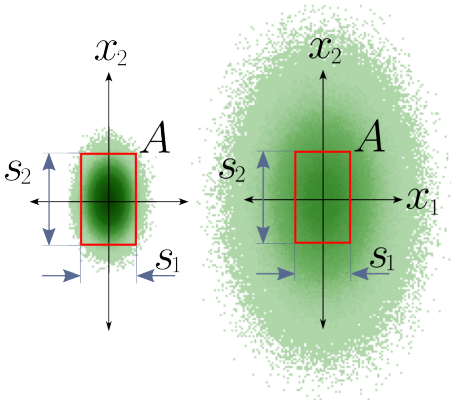


Figura 3.1: Se muestran dos histogramas de distintas variables gaussianas bidimensionales. Mientras más intenso sea el color verde, más observaciones se realizaron en ese bin. La varianza de la variable en el lado izquierdo es menor que la de la variable en el lado derecho. Sobre estos histogramas, se muestra la misma región de integración A.

(s_1, s_2) actúan como factores de normalización que permiten la comparación de la dispersión entre diferentes magnitudes. Para implementaciones prácticas, se puede elegir $s_i = \sigma_{i,\text{mín}}$ para asegurar que se cumpla la condición especificada en Ec. 3.4 para todo $\sigma_i \geq \sigma_{i,\text{mín}}$.

Existe una relación entre $\mathbf{\Sigma}_k$ y p_k de la celda k-ésima para un vector de variables aleatorias de dimensión N dado por[4]:

$$|\mathbf{\Sigma}_k|^{\frac{1}{2N}} = \tilde{\sigma}_k \geq \frac{a}{p_k^{1/N}} \tag{3.8}$$

donde a es una constante que depende de N y todos los parámetros s_i, y $\tilde{\sigma}_k$ es la media geométrica de las desviaciones estándar de la celda k-ésima. Esta relación establece una conexión entre la PD y el criterio D-opt, una medida estándar de incertidumbre en ASLAM Carrillo et al. (2012). Sin embargo, cada PD está acotado entre 0 y 1 con una clara interpretación probabilística. En contraste, el criterio D-opt no está acotado y solo se interpreta como el volumen del hiperelipsoide definido por $\mathbf{\Sigma}_k$ cuando las variables de estimación representan longitudes.

[4]Este resultado está demostrado en la Sec. A.2.4.

La metodología central propuesta en esta tesis se basa en la construcción y mantenimiento de una representación espacial de la incertidumbre utilizando las Probabilidades de Dispersión. Se emplea un mapa grillado, donde cada celda k almacena información sobre la certeza de la información asociada a la región espacial R_k que cubre dicha celda.

Es fundamental distinguir este mapa con incertidumbres de los mapas de ocupación tradicionales. Mientras un mapa de ocupación estima el *estado* de la celda (ej., probabilidad de estar ocupada), el mapa basado en PD cuantifica la *incertidumbre* o *confianza* sobre la información de esa celda, representada por su valor de Probabilidad de Dispersión p_k. En una implementación completa, este mapa con incertidumbre complementaría a un mapa de estado (ej., ocupación, nube de puntos), proporcionando información sobre la fiabilidad de este último.

Para facilitar ciertos análisis y la definición de las Fronteras de Incertidumbre (Sec. 3.5), es conveniente trabajar con un Mapa de Incertidumbre (MI) derivado, donde cada celda k almacena un valor escalar de incertidumbre σ_k^* obtenido directamente de p_k mediante la relación inversa $\sigma_k^* = a/p_k^{1/N}$ (Ec. 3.26). Un valor alto de p_k (alta confianza) corresponde a un valor bajo de σ_k^* (baja incertidumbre), y viceversa.

El proceso de actualización de este mapa es dinámico. Cuando una nueva medición proporciona información (ej., sobre la presencia de un objeto o sobre espacio libre), esta tiene asociada una incertidumbre, cuantificable mediante una PD (p_{new}). Dependiendo del modelo del sensor y la pose estimada del agente, esta nueva información afecta a un conjunto de celdas en el mapa. El valor p_k (o su log-odds ℓ_k) de cada celda k-ésima afectada se actualiza fusionando la información previa con la nueva p_{new}, utilizando reglas específicas que se detallarán en la Sec. 3.3. El resultado es un mapa que refleja la evolución espacial y temporal de la certeza sobre el entorno.

Además, como se explorará en la Sec. 3.4 y 3.5, la distribución espacial de los valores p_k (o σ_k^*) en el mapa facilita la evaluación global de la calidad de la reconstrucción y la identificación de regiones críticas para guiar futuras acciones de exploración o mapeo.

Con fines didácticos, se propone un ejemplo ilustrativo en el cual se asume que hay un agente cuya posición se conoce con cierto grado de incertidumbre, además se asume que este agente tiene un sensor de rango que mide la distancia entre el agente y otros objetos como puede ser una pared de ladrillos. Esta situación se muestra en la Fig. 3.2.

La posición del agente, obtenida mediante una simulación Monte Carlo

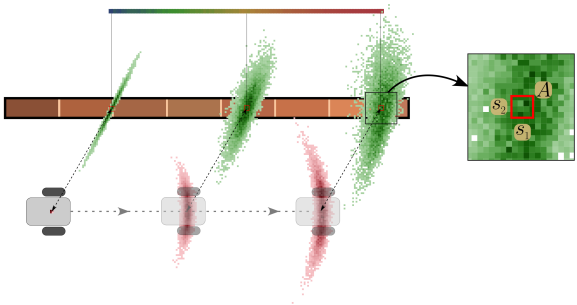


Figura 3.2: Se muestra la evolución de la incertidumbre de la pose del agente (representada por cuadrados rojos) y la medición de un sensor de rango, propagada a través de la incertidumbre del agente (representada por cuadrados verdes). La intensidad del color indica la probabilidad de la medición en esa región. La barra de color en la parte superior muestra los valores p_k de las celdas, donde valores p_k más altos corresponden a colores más azules.

asumiendo errores odométricos, se representa con diferentes tonalidades de rojo. Mientras más intenso el color, más probable es que el agente se encuentre en esa celda. En el primer paso de simulación, la incertidumbre en la posición del agente está confinada en una celda. Desde esta región, emerge una flecha que indica una medición realizada por el sensor, cuya dispersión se representa en diferentes tonalidades de verde. Nótese que la creciente incertidumbre en la posición del agente se propaga a las mediciones del sensor en el marco global. En el paso inicial, como la incertidumbre del agente es mínima, la incertidumbre de medición relativa al agente es similar a la incertidumbre de medición con respecto a un marco de referencia global externo. A medida que la simulación avanza, la incertidumbre del agente se incrementa, produciendo también un incremento en la incertidumbre de medición con respecto al sistema de referencia global. En particular, siendo la medición del sensor en coordenadas polares $[\rho, \theta]$ relativo al agente y la posición y orientación del mismo $[X_a, Y_a, \phi_a]$ con respecto al sistema de referencia global, entonces las mediciones se modelan según:

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} X_a + \rho\cos(\theta + \phi_a) \\ Y_a + \rho\sin(\theta + \phi_a) \end{bmatrix} \tag{3.9}$$

Entonces, dado un mapa grillado con celdas de lado d, se define:

$$A = \{(x_1, x_2) \in \mathbb{R}^2 : \|x_i - \mu_i\| < s_i/2,\ i = 1, 2\},$$

donde la PD resulta:

$$\begin{aligned} P(\mathbf{X} \in A) = P[&\mu_1 - s_1/2 < x_1 < \mu_1 + s_1/2, \\ &\mu_2 - s_2/2 < x_2 < \mu_2 + s_2/2]. \end{aligned}$$

en particular, si se asume que $\mathbf{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$, se tiene que:

$$\begin{aligned} P(\mathbf{X} \in A) = &F_X(\mu_1 + s_1/2, \mu_2 + s_2/2) \\ &-F_X(\mu_1 - s_1/2, \mu_2 + s_2/2) \\ &-F_X(\mu_1 + s_1/2, \mu_2 - s_2/2) \\ &+F_X(\mu_1 - s_1/2, \mu_2 - s_2/2) \end{aligned} \tag{3.10}$$

donde F_X es la función de distribución acumulada de la variable aleatoria bidimensional $\mathbf{X}$. Finalmente la probabilidad obtenida usando Ec. 3.10 se asigna a la celda k-ésima si $\mu \in R_k$. Notar que para obtener la PD a partir de una medición se requiere conocer μ y $\mathbf{\Sigma}$[5]. En este ejemplo la PD se calculó mediante la matriz de varianza-covarianza muestral y el vector de medias obtenidas con las muestras de cada paso de simulación. Cuando se implementa la metodología por fuera de este ejemplo se requiere que un subsistema de localización proporcione las estimaciones de μ y $\mathbf{\Sigma}$.

En la Fig. 3.3 se muestra la evolución de la PD en línea azul en conjunto con el determinante de la matriz de varianza-covarianza en línea naranja. Esta es una medida común para los sistemas de ASLAM tal como se explica en Carrillo et al. (2012) y se mencionó en Ec. 2.1. Para fines de visualización se muestra el logaritmo de estas variables. Notar que el criterio D-opt y la PD tienen un comportamiento similar y por lo tanto esta última puede ser utilizada para representar la incertidumbre de medición en un mapa de probabilidades de dispersión.

La ventaja que presenta este enfoque es que se mantiene la interpretabilidad de los datos, ya que se define una probabilidad de que un evento suceda. Mientras más probable resulte este evento, menor incertidumbre tendrán las variables analizadas. En cambio, cuando se utiliza el criterio D-opt por ejemplo, el valor obtenido no tiene una interpretación directa, es un número que resulta de una multiplicación

[5]También se puede obtener mediante la distribución conjunta de la variables bajo interés.

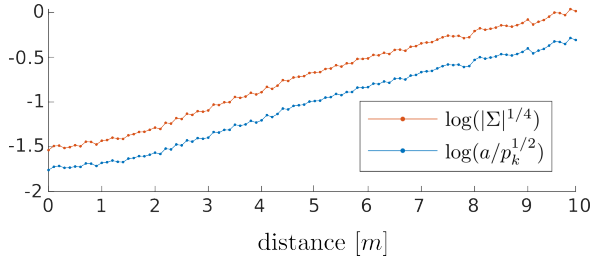


Figura 3.3: Se muestra la evolución de dos medidas de incertidumbre: p_k y Σ_k. Se grafican sus logaritmos para facilitar la visualización.

de varianzas *con distintas magnitudes*. Estas magnitudes tienen una dispersión asociada al proceso medido, por lo que un valor de D-opt determinado, no tiene una interpretación directa. Por ejemplo, en caso de que se esté utilizando para realizar tareas de exploración, las unidades de D-opt pueden ser $m^4 rad^2$, con valores que pueden variar entre 10^{-3} y 10^3, lo cual no es beneficioso para la interpretación de los resultados.

En cambio, cuando se utilizan las probabilidades de dispersión, es necesario definir un intervalo para cada variable bajo análisis. Con esto, la representación de la incertidumbre, se vuelve interpretable y permite relacionar cualquier tipo de variable con distintas unidades y escalas. Esta característica es relevante para: la toma de decisiones, la interpretación de los resultados, la comparación de distintos sistemas, la generalización de los resultados y, por último, la eliminación de factores de ponderación adimensionales que reducen la robustez del sistema [6], Carrillo et al. (2018); Placed et al. (2023).

3.3. Proceso de Actualización

En la práctica, a la misma celda se le asignan distintas probabilidades, por esto es necesario implementar un método para combinarlas. Una alternativa es quedarse con aquella que tiene mayor probabilidad Thrun (2002), otra opción es utilizar una media ponderada Sansoni and Gimenez (2025), sin embargo la opción más utilizada es la actualización Bayesiana, ya sea estimando la probabilidad de ocupación de cada celda dada la observación $p(m_k|z_t)$ Moravec (1989) o bien maximizando la

[6] Robustez entendida como la generalización de los resultados.

probabilidad a posteriori de observar una medición dada la ocupación de la celda $p(z_t|m_k)$, relajando la hipótesis de celdas independientes Thrun (2003).

Considerando la necesidad de combinar múltiples estimaciones de probabilidad para una misma celda, en las siguientes secciones se explorarán dos metodologías de actualización. Primero, se describirá en detalle la implementación de la actualización mediante una media ponderada, basada en el trabajo de Sansoni and Gimenez (2025). Esta propuesta busca abordar algunas de las deficiencias de la actualización Bayesiana estándar, particularmente su comportamiento con mediciones sucesivas bajo incertidumbre de pose y la saturación de la información, manteniendo al mismo tiempo una estructura que facilita su integración en los sistemas de mapeo actuales.

Segundo, se presentará una aproximación novedosa que utiliza el filtro de Kalman para la actualización de las celdas. Aunque esta segunda metodología puede requerir un mayor esfuerzo en su implementación, ofrece la capacidad fundamental de incorporar la evolución temporal de las celdas. Esto se logra mediante la formulación de un modelo dinámico para el estado de cada celda y la aplicación recursiva del filtro de Kalman, permitiendo así extender la reconstrucción de un dominio puramente espacial a uno espacio-temporal. Este enfoque abre la puerta a la predicción del estado de las celdas y a la planificación informada de mediciones futuras.

3.3.1. Actualización mediante Media Ponderada

Para poder entender en qué consiste la actualización mediante medias ponderadas, primero es necesario repasar el desarrollo de la actualización Bayesiana, también conocido como log-odds. Este tipo de representación requiere las siguientes hipótesis [Stachniss (2009), Sec. 2.2.1]:

1. Las celdas del mapa son independientes entre sí. Es decir:

$$p(\mathbf{m}) = \prod_{k=1}^{n} p(m_k) \tag{3.11}$$

   con $p(\mathbf{m})$ la probabilidad de que el mapa $\mathbf{m}$ tenga un estado y $p(m_k)$ es la probabilidad de que la celda k-ésima esté ocupada.

2. Se conocen perfectamente la posiciones del robot $x_{1:t}$ (sin incertidumbre).

3. Las mediciones actuales z_t son independientes de las mediciones anteriores $z_{1:t-1}$ y de las posiciones anteriores $x_{1:t-1}$.

4. El estado de una celda del mapa $p(m_k)$ no depende del estado del robot, es decir $p(m_k|x_t) = p(m_k)$.

5. El mapa es estático, es decir, no cambia en el tiempo.

Con estas hipótesis, la actualización Bayesiana de la celda k-ésima resulta:

$$\frac{p(m_k|x_{1:t}, z_{1:t})}{1 - p(m_k|x_{1:t}, z_{1:t})} = \frac{p(m_k|x_t, z_t)}{1 - p(m_k|x_t, z_t)} \frac{1 - p(m_k)}{p(m_k)} \frac{p(m_k|x_{1:t-1}, z_{1:t-1})}{1 - p(m_k|x_{1:t-1}, z_{1:t-1})} \tag{3.12}$$

luego si se define:

$$\ell(x) = \ln \frac{p(x)}{1 - p(x)} \tag{3.13}$$

la Ec. 3.12 se puede reescribir como:

$$\ell(m_k|x_{1:t}, z_{1:t}) = \ell(m_k|x_t, z_t) + \ell(m_k|x_{1:t-1}, z_{1:t-1}) - \ell(m_k) \tag{3.14}$$

$$\ell_{t,k} = \ell_{\text{new},k} + \ell_{t-1,k} - C \tag{3.15}$$

con $C = \ell(m_k)$, pero como normalmente se asume que $p(m_k) = 0.5$, resulta que $C = 0$ obteniéndose la Ec. 3.2. Además se asume que la posición del agente es conocida y sin incertidumbre, lo cual implica que toda la variabilidad está en la medición. Por tanto, si el agente está estático y mediciones sucesivas se realizan en la misma celda, es de esperar que algunas mediciones indiquen que la celda está ocupada y otras que está libre, dependiendo de la función de distribución de éstas generando que el valor $\ell_{t,k}$ converja. Sin embargo en la práctica la pose del agente no se conoce con exactitud y debido al estado de la tecnología actual, la incertidumbre en la medición del sensor suele ser más baja que la del agente [7]. Por lo tanto, en esta situación la mayoría de las mediciones indicarían el mismo estado de la celda, produciendo que $\ell_{t,k}$ crezca en módulo tanto como la cantidad de mediciones que se efectúen. Luego supongamos que el agente se mueve y vuelve a medir la misma celda desde otra perspectiva pero con menor incertidumbre y la medición resulta en un estado distinto al anterior. En este caso, para que el mapa refleje esta nueva medición se requiere que se efectúe una mayor cantidad de mediciones en la celda. Este comportamiento es indeseado ya que el estado de la

[7] Un punto medido por un lidar tiene una incertidumbre de $2cm$ y las estimas en la posición pueden alcanzar incertidumbres de $1m$

celda depende de la cantidad de mediciones efectuadas. Por esto en trabajos como Hornung et al. (2013), se aplica una función de saturación sobre los valores de $\ell_{t,k}$, sin embargo esto no resuelve el problema, solo lo atenúa.

En Sansoni and Gimenez (2025) se introduce una modificación al algoritmo de actualización basado en log-odds para incorporar la incertidumbre del agente mediante las PD. Esta modificación permite mantener la sencillez del método original y, al mismo tiempo, controlar la forma en que nuevas mediciones con mayor incertidumbre afectan a celdas previamente exploradas. Concretamente, se establece un umbral asociado a la probabilidad β, que representa el estado de ausencia de conocimiento previo sobre la celda.

Definido ℓ_β como el valor log-odds correspondiente a β, la regla de actualización propuesta tiene la siguiente forma:

$$\ell_{t,k} = \begin{cases} \ell_{t-1,k} + \kappa(\ell_{\text{new},k} - \ell_{t-1,k}), & \text{si } \ell_{\text{new},k} > \ell_{t-1,k} \vee \ell_{t-1,k} = \ell_\beta, \\ \ell_{t-1,k} & \text{en caso contrario} \end{cases} \tag{3.16}$$

donde $\kappa \in (0, 1)$ un coeficiente que controla la velocidad de convergencia. De este modo, si la celda ya poseía un valor log-odds superior a $\ell_{\text{new},k}$, se asume que no hay razón para "degradar" información confiable con mediciones más inciertas. Por tanto, solo se reajusta la celda cuando la nueva observación brinda un valor con menor incertidumbre que la "desconocida" (β). Este mecanismo evita situaciones en las cuales, al aumentar la incertidumbre, una celda que ya fue "explorada" se reclasifique erróneamente como "no explorada". En términos prácticos, eso significa que el log-odds de dicha celda no debe cruzar por debajo de ℓ_β salvo que la celda esté clasificada como "no explorada" y una nueva medición la clasifique como "explorada" pero con una incertidumbre mayor a β. Así, la PD no crece arbitrariamente en áreas ya analizadas con menor incertidumbre. Un ejemplo numérico ilustrativo para obtener este parámetro se encuentra en la Sec. A.2.3.

Por otro lado, cuando un objeto es detectado, la medición no solo nos informa sobre la distancia entre el objeto al sensor, sino que también nos dice que el espacio entre el sensor y el objeto está libre, con lo cual todas estas celdas observadas (espacio libre y ocupado) son actualizadas con la incertidumbre del sensor afectada por la del agente. Considerando que un sensor como un LiDAR o una cámara proporciona múltiples mediciones, entonces el mapa es actualizado en toda la región definida por su Campo de Visión (FOV, por sus siglas en inglés).

En resumen, esta modificación basada en PD garantiza que el mapa refleje de forma más realista la evolución de la incertidumbre en cada celda: celdas previamente exploradas y con menor dispersión no "retroceden" a estados

desconocidos ante mediciones con alta variabilidad, mientras que nuevas zonas se incorporan gradualmente con la incertidumbre correspondiente a las observaciones disponibles.

3.3.2. Actualización mediante Filtro de Kalman

Si bien la actualización mediante medias ponderadas es una alternativa sencilla y efectiva a los log-odds, no deja de ser una mejora sobre un método que tiene limitaciones. En particular, la principal limitación que se detecta a la hora de implementar este método en ambientes dinámicos es la hipótesis de que el mapa es estático. Por esto las celdas que no son observadas por el sensor, no se actualizan. Cuando se busca realizar tareas de planificación enfocadas a la reconstrucción espacio-temporal del ambiente, se requiere de un modelo de predicción que permita estimar el estado de las celdas no observadas. En este sentido, una alternativa que permite esto es aplicar un filtro de Kalman. Esta característica admite la utilización de los mapas de incertidumbre como herramienta para definir objetivos de exploración, planificar trayectorias y establecer sistemas de vigilancia de cultivos basados en la incertidumbre.

Si bien el enfoque de las PD fue desarrollado para el área de la reconstrucción de entornos, donde una celda puede o no estar ocupada, dado su enfoque general, permite la representación de la incertidumbre de cualquier conjunto de variables. Por ejemplo, si se realiza un mapa del nivel de humedad del suelo con su incertidumbre asociada, la evolución de esta humedad en el tiempo dependerá de varios factores, como las condiciones climáticas, el tipo de suelo, la superficie foliar de la vegetación cercana, entre otros. En base a un modelo de predicción, se puede estimar la evolución de la humedad en el tiempo en conjunto con una incertidumbre de estimación. Dicha incertidumbre se puede representar mediante las PD y en aquellas zonas que tengan una incertidumbre mayor a la esperada, se requerirá que se efectúe una nueva medición para actualizar el estado de la celda y así disminuir su incertidumbre. Este es solo un ejemplo de las aplicaciones que se pueden realizar con este enfoque.

Teniendo en cuenta que cada celda se puede modelar como una variable aleatoria, cuyo modelo está dado por:

$$\mathbf{x}_{t+1} = \mathbf{A_m}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t$$
$$\mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{v}_t$$

con $\mathbf{x}_t$ el vector de estados de la celda en el tiempo t, $\mathbf{u}_t$ el vector de entradas al sistema, $\mathbf{w}_t$ y $\mathbf{v}_t$ vectores de ruido blanco gaussiano con media cero y varianza $\mathbf{\Sigma}_w$

y $\mathbf{\Sigma}_v$ respectivamente. $\mathbf{A_m}$[8] y $\mathbf{B}$ son matrices de transición y control, $\mathbf{H}$ es la matriz de observación. Entonces el filtro de Kalman se puede expresar como:

$$\begin{aligned}
\hat{\mathbf{x}}_{t+1|t} &= \mathbf{A_m}\hat{\mathbf{x}}_{t|t} + \mathbf{B}\mathbf{u}_t \\
\mathbf{P}_{t+1|t} &= \mathbf{A_m}\mathbf{P}_{t|t}{\mathbf{A_m}}^T + \mathbf{\Sigma}_w \\
\mathbf{K}_{t+1} &= \mathbf{P}_{t+1|t}\mathbf{H}^T(\mathbf{H}\mathbf{P}_{t+1|t}\mathbf{H}^T + \mathbf{\Sigma}_v)^{-1} \\
\hat{\mathbf{x}}_{t+1|t+1} &= \hat{\mathbf{x}}_{t+1|t} + \mathbf{K}_{t+1}(\mathbf{z}_{t+1} - \mathbf{H}\hat{\mathbf{x}}_{t+1|t}) \\
\mathbf{P}_{t+1|t+1} &= (\mathbf{I} - \mathbf{K}_{t+1}\mathbf{H})\mathbf{P}_{t+1|t}
\end{aligned}$$

con $\hat{\mathbf{x}}_{t|t}$ la estimación a priori del estado de la celda en el tiempo t, $\mathbf{P}_{t|t}$ la matriz de covarianza a priori, $\hat{\mathbf{x}}_{t+1|t+1}$ la estimación a posteriori, $\mathbf{P}_{t+1|t+1}$ la matriz de covarianza a posteriori, $\mathbf{K}_{t+1}$ la ganancia de Kalman, $\mathbf{z}_{t+1}$ la medición en el tiempo $t+1$ y $\mathbf{I}$ la matriz identidad.

Si se toma el caso unidimensional y trivial en donde $\mathbf{A_m} = 1$, $\mathbf{B} = 0$, $\mathbf{H} = 1$, $u_t = 0$, $\mathbf{\Sigma}_w = \sigma_w^2$, $\mathbf{\Sigma}_v = \sigma_v^2$ y $\mathbf{P}_{t|t} = \sigma_{x_{t|t}}^2$, entonces el filtro de Kalman se reduce a:

$$\hat{x}_{t+1|t+1} = \frac{\hat{x}_{t|t}\sigma_v^2 + z_{t+1}(\sigma_{x_{t|t}}^2 + \sigma_w^2)}{\sigma_{x_{t|t}}^2 + \sigma_w^2 + \sigma_v^2} \tag{3.17}$$

$$\sigma_{x_{t+1|t+1}}^2 = \frac{\sigma_v^2(\sigma_{x_{t|t}}^2 + \sigma_w^2)}{\sigma_{x_{t|t}}^2 + \sigma_w^2 + \sigma_v^2} \tag{3.18}$$

luego como se asume que el sistema es lineal y las variables son gaussianas, entonces dado el intervalo A es posible obtener la PD asociada a la estimación para cada instante.

En conclusión, la actualización mediante filtro de Kalman permite incorporar una dinámica dentro del mapa mediante la utilización de modelos de predicción, sin la necesidad de medir continuamente. Además, se puede utilizar la PD para representar la incertidumbre de las variables y así poder planificar tareas de exploración, planificar trayectorias y establecer sistemas de vigilancia de cultivos basados en la incertidumbre. Incluso esta mecánica habilita la posibilidad de realizar tareas de prevención focalizadas. Por último, este enfoque permite incorporar la variable temporal dentro del mapa, transformando la reconstrucción espacial en espacio-temporal.

[8] Se agrega el sufijo *m* para diferenciarla del intervalo A

3.4. Métrica de evaluación basada en Incertidumbre

Como se mencionó en la Sec. 3.2, la entropía de Shannon se utiliza ampliamente como criterio de calidad de la información en mapas probabilísticos, especialmente en mapas de grillas de ocupación. Sin embargo, además de los problemas mencionados en Blanco et al. (2008), esta entropía no está bien definida para distribuciones continuas ya que pueden tomar valores negativos que carecen de sentido.

Un método alternativo que se puede utilizar para cuantificar la información o la diferencia entre estados de conocimiento es la divergencia de Kullback-Leibler (D_{KL}) Kullback and Leibler (1951)[9], también conocida como *entropía relativa*. Esta divergencia sirve para medir la diferencia o "distancia" informativa entre dos distribuciones de probabilidad, P y Q. Para el caso de un mapa de grillas de ocupación, se puede calcular la D_{KL} entre la distribución de probabilidad del estado de cada celda en un mapa M y la distribución correspondiente en otro mapa M', o con respecto a una distribución de referencia.

Si se asume independencia entre las celdas del mapa, la divergencia total entre dos mapas M (con distribuciones P_k para cada celda k) y M' (con distribuciones Q_k para cada celda k) se puede expresar como la suma de las divergencias individuales de cada celda:

$$\begin{aligned} D(M||M') &= \sum_k D_{KL}(P_k \parallel Q_k) \\ &= \sum_k \int_{-\infty}^{\infty} p_k(w) \ \ln\left(\frac{p_k(w)}{q_k(w)}\right) \ dw \end{aligned} \tag{3.19}$$

donde $p_k(w)$ y $q_k(w)$ son las funciones de densidad de probabilidad (PDF) de alguna variable continua w asociada a la celda k-ésima en los mapas M y M' respectivamente. Para obtener la D_{KL} se requiere conocer las distribuciones de probabilidad de cada celda, lo cual no es trivial. Además, como lo que se busca

[9] Es una divergencia porque no cumple con las propiedades de simetría y la desigualdad triangular, por lo que no es una métrica de distancia en el sentido estricto. Sin embargo, se han desarrollado extensiones, que la convierten en métrica, como la raíz cuadrada de la divergencia Jensen-Shannon Endres and Schindelin (2003): $d_{js}(P\|Q) = \sqrt{JS(P\|Q)} = \sqrt{\frac{1}{2}D_{\mathrm{KL}}(P\|M) + \frac{1}{2}D_{\mathrm{KL}}(Q\|M)}$, con $M = \frac{1}{2}(P+Q)$.

es medir la calidad del mapa, entendiéndose como la incertidumbre del mapa, se propone utilizar la D_{KL} entre dos distribuciones Gaussianas con la misma media pero distinta varianza para representar diferentes niveles de incertidumbre sobre una misma estimación puntual. Sin embargo, se podría utilizar cualquier distribución.

Basándose en la D_{KL}, se propone un índice de incertidumbre denominado Entropía Relativa con signo o *SiREn* por sus siglas en inglés de *Signed Relative Entropy*, el cual se define como:

$$D_s(M) = \sum_{k=1}^{K} C\Big(D_{KL}\big(P_k \parallel Q\big)\Big)\,\mathrm{sgn}\Big(|\mathbf{\Sigma}_Q| - |\mathbf{\Sigma}_{P_k}|\Big) \tag{3.20}$$

donde K es la cantidad de celdas del mapa, C es el espacio que ocupa la celda k-ésima, $P_k \sim \mathcal{N}(0, \mathbf{\Sigma}_{P_k})$ es la distribución de la celda k-ésima, $\mathrm{sgn}(x)$ es la función signo y $Q \sim \mathcal{N}(0, \mathbf{\Sigma}_Q)$ es la distribución correspondiente a la máxima incertidumbre considerada tolerable, definida por el usuario. Esta incertidumbre máxima puede representarse por su matriz de covarianza $\mathbf{\Sigma}_Q$ o, de forma equivalente en nuestro marco, por el valor escalar σ_{max} derivado de ella ($\sigma_{max} = |\mathbf{\Sigma}_Q|^{1/(2N)}$). Análogamente a cómo se definen los intervalos de referencia s_i (basados en la mínima incertidumbre esperada $\sigma_{i,min}$), el umbral σ_{max} puede establecerse considerando la máxima incertidumbre aceptable para la aplicación específica a partir del valor máximo tolerable i-ésimo $\sigma_{i,max}$. En la Sec. A.2.3 se puede acceder a un ejemplo numérico ilustrativo para ayudar a clarificar la obtención de este parámetro.

Este índice propuesto cuenta con las siguientes propiedades:

- Si todo el mapa está sin explorar, entonces $P_k = Q$, luego $D_s = 0$. Es decir, las celdas no exploradas no aportan en el índice.
- Cada celda aporta una cantidad proporcional al espacio que ocupa, esto permite incorporar mapas de resolución variable como Octomap Hornung et al. (2013).
- Se penalizan la celdas que tengan una incertidumbre mayor a la tolerable. El aporte será negativo si $|\mathbf{\Sigma}_Q| < |\mathbf{\Sigma}_{P_k}|$.
- Hay una relación monótona decreciente no lineal entre $D_{KL}(P_k||Q)\,\mathrm{sgn}(|\mathbf{\Sigma}_Q| - |\mathbf{\Sigma}_{P_k}|)$ y $|\mathbf{\Sigma}_{P_k}|$. Esta relación no lineal es más sensible a los cambios en la incertidumbre de medición cuando $|\mathbf{\Sigma}_Q| > |\mathbf{\Sigma}_{P_k}|$ que cuando $|\mathbf{\Sigma}_Q| < |\mathbf{\Sigma}_{P_k}|$.

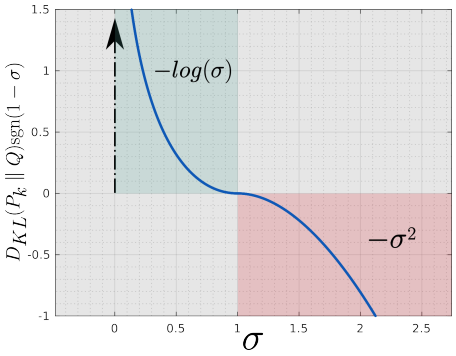


Figura 3.4: Comportamiento de un término de la sumatoria $D_s(M)$, para $\sigma_{\max} = 1$ en función de σ.

Para fines de simplificación, se define P_k y Q gaussianas ya que existe una expresión cerrada para calcular $D_{KL}\left(P_k \parallel Q\right)$ Hershey and Olsen (2007). Sin embargo, esta idea se puede modificar para cualquier tipo de distribución, incluso para aquellas no paramétricas. A modo ilustrativo en la Fig. 3.4 se muestra como se comporta el término $D_{KL}(P_k \parallel Q)\,\text{sgn}(|\mathbf{\Sigma}_Q| - |\mathbf{\Sigma}_{P_k}|)$ para el caso unidimensional tomando $Q \sim \mathcal{N}(0, \sigma^2_{\max})$ con $\sigma_{\max} = 1$ y $P_k \sim \mathcal{N}(0, \sigma^2)$, con σ variable, resultando en:

$$D_{KL}(P_k \parallel Q)\,\text{sgn}(1-\sigma^2) = \begin{cases} -\ln\sigma + \frac{1}{2}(\sigma^2 - 1) & \text{si } \sigma < 1 \\ \ln\sigma - \frac{1}{2}(\sigma^2 - 1) & \text{si } \sigma > 1 \end{cases} \tag{3.21}$$

Principalmente en la Fig. 3.4 se observa que la relación es monótona decreciente, además hay una asíntota vertical en $\sigma = 0$ y cuando $\sigma > \sigma_{\max}$ se comporta aproximadamente como un polinomio de grado 2. Teniendo en cuenta estas observaciones se concluye que $D_s(M)$ tiene una mayor sensibilidad (variación) cuando $\sigma < \sigma_{\max}$ comparando con el caso $\sigma > \sigma_{\max}$.

Por otro lado, cuando se utiliza este índice en un problema de exploración y, por ejemplo, el agente se encuentra navegando solamente con odometría[10] en línea recta, entonces al calcular $D_s(M)$ con $\sigma_{\max,1}$ y $\sigma_{\max,2}$ tal que $\sigma_{\max,1} > \sigma_{\max,2}$, resulta que el índice se volverá negativo con menor distancia recorrida cuando se calcule con $\sigma_{\max,2}$. Esta propiedad es útil cuando se realizan tareas de

[10]Es decir, la incertidumbre crece a medida que se mueve.

planificación de trayectorias, ya que el parámetro $\sigma_{\max}$ permite controlar la relación exploracion-explotación de forma interpretable y transparente para el usuario. Dependiendo de cual es la máxima incertidumbre tolerable, el agente puede explorar más o menos. Sin embargo, $\sigma_{\max}$ no es limite estricto para el sistema de planificación, ya que el agente puede explorar celdas con mayor incertidumbre, pero penalizando por el índice y esta penalización depende de la diferencia entre $\sigma_{\max}$ y σ.

Si se cuenta con un mapa grillado donde cada celda tiene asociada una PD notada como p_k, entonces una cota inferior para $D_s(M)$ resulta:

$$D_s(M) \geq \sum_k C\Big(\log \frac{p_k}{\beta} - \frac{N}{2} + \frac{N}{2}\big(\frac{\beta}{p_k}\big)^{2/N}\Big)\,\mathrm{sgn}(p_k - \beta) \tag{3.22}$$

siendo N la dimensión del vector de medidas. Cuando $p_k = \beta$ la cota coincide con el valor de $D_s(M)$, por esto se puede utilizar como aproximación para valores de p_k lo suficientemente próximos a β. La demostración de esta cota se puede encontrar en la Sec. A.2.5.

En general el índice SiREn no solo se puede utilizar para medir la incertidumbre de un mapa grillado, sino también para un mapa disperso, trayectorias de agentes, entre otros. En el caso del mapa disperso, cada elemento de la sumatoria se corresponde a un punto del mapa con su incertidumbre asociada. En el caso de las trayectorias de agentes, cada elemento de la sumatoria se corresponde a la incertidumbre de la posición del agente en un instante de tiempo. En este sentido, el índice SiREn se puede utilizar para medir la incertidumbre de cualquier tipo de variable con respecto a una incertidumbre máxima tolerable.

SiREn para mapa disperso de Landmarks

En la práctica, la estimación de la posición de un agente y su correspondiente incertidumbre puede obtenerse mediante un sistema de SLAM. Como se vió en la Sec. 2.2.1, algunos sistemas almacenan todas las mediciones pasadas, entonces cuando se realizan cierres de lazos es posible actualizar todas las celdas y sus incertidumbres del mapa grillado. Sin embargo, algunos sistemas de SLAM utilizan mapas dispersos con landmarks ($\mathbb{L}$), por ello el índice SiREn puede utilizarse como indicador de la incertidumbre de este otro tipo de mapas, en efecto para este caso se tiene:

$$D_s(\mathbb{L}) = \sum_{|\boldsymbol{\Sigma}_{l_i}| \leq |\boldsymbol{\Sigma}_Q|}^{N_l} D_{KL}(\mathbf{L}_i) - \sum_{|\boldsymbol{\Sigma}_{l_i}| > |\boldsymbol{\Sigma}_Q|}^{N_l} D_{KL}(\mathbf{L}_i) \tag{3.23}$$

donde si se asume que la función de distribución de cada i-ésimo landmark es gaussiana con covarianza $\Sigma_{l_i} \in M^{n_l \times n_l}$, entonces:

$$D_{KL}(\mathbf{L}_i) = \frac{1}{2}\left(\ln\left(\frac{|\mathbf{\Sigma}_Q|}{|\mathbf{\Sigma}_{l_i}|}\right) + \text{tr}\left(\mathbf{\Sigma}_Q^{-1}\mathbf{\Sigma}_{l_i}\right) - n_l\right)$$

SiREn para trayectorias

En tareas de planificación de trayectorias, el índice SiREn se puede utilizar como indicador de la incertidumbre de la estimación de sus estados en el intervalo $[0\ T]$. En efecto:

$$D_s(X_{0:T}) = \sum_{|\mathbf{\Sigma}_{x_t}| \leq |\mathbf{\Sigma}_Q|}^{T} D_{KL}(X_t) - \sum_{|\mathbf{\Sigma}_{x_t}| > |\mathbf{\Sigma}_Q|}^{T} D_{KL}(X_t) \tag{3.24}$$

donde $\mathbf{\Sigma}_{x_t} \in M^{n \times n}$ es la matriz de covarianza de la estimación de los estados del agente en el tiempo t. En caso de que los estados se modelen como variables aleatorias gaussianas, resulta:

$$D_{KL}(X_t) = \frac{1}{2}\left(\ln\left(\frac{|\mathbf{\Sigma}_Q|}{|\mathbf{\Sigma}_{x_t}|}\right) + \text{tr}\left(\mathbf{\Sigma}_Q^{-1}\mathbf{\Sigma}_{x_t}\right) - n\right) \tag{3.25}$$

Queda a disposición del usuario modificar estas expresiones para adaptarlas a su problema. Por ejemplo, se pueden agregar factores de ponderación que tengan en cuenta la longitud de la trayectoria o bien establecer otra relación con la variable temporal, entre otras alternativas.

En resumen el índice SiREn es un indicador de incertidumbre que permite comparar distintos mapas, trayectorias, etc., en función de una incertidumbre máxima tolerable. Este índice es interpretable y transparente para el usuario. Además se puede utilizar para cualquier tipo de variable con incertidumbre asociada, tanto para distribuciones paramétricas como no paramétricas. Por último se puede utilizar para comparar distintas reconstrucciones sin conocer el ground truth, ya que se basa en las incertidumbres estimadas por el sistema. Este es un factor relevante tanto para los sistemas de SLAM como para los de ASLAM.

3.5. Fronteras de Incertidumbre: Objetivos de exploración

Como se mostró en la Sec. 2.6, las fronteras propuestas por Yamauchi (1997) fueron y son la principal herramienta para identificar objetivos de exploración. Sin embargo, estos objetivos no integran naturalmente la posibilidad de cerrar lazos y así disminuir la incertidumbre del agente y del mapa. Además, como se usan generalmente en entornos indoor, también se utilizan como criterios de finalización para la exploración. Por otro lado, en entornos abiertos como la agricultura, la exploración no tendría fin, ya que siempre se puede explorar más debido a la falta de límites, como paredes. Por estas razones se propone redefinir el concepto de fronteras, llamándolas de ahora en más como *Fronteras Clásicas*, FC, y proponer unas nuevas fronteras denominadas *Fronteras de Incertidumbre*, FI. Estas fronteras se pueden utilizar para identificar objetivos de exploración, ya sea para explorar como para cerrar lazos tanto en espacios abiertos o cerrados. A su vez, pueden utilizarse como criterio de finalización y cuentan con un parámetro que ajusta la sensibilidad.

Al igual que las FC, las FI se obtienen a partir de calcular la norma del gradiente de un mapa. En este caso las FI son obtenidas a partir de un *Mapa de Incertidumbre* MI, el cual se genera con las PD de cada celda. Dadas las PD de cada celda, la celda de incertidumbre asociada σ_k^* del mapa MI se define según:

$$\sigma_k^* \triangleq \frac{\left(\prod_i^N s_i\right)^{1/N} (2\sqrt{3})^{-1}}{p_k^{1/N}} \tag{3.26}$$

$$= a p_k^{-1/N} \qquad \text{tomando } a = \left(\prod_i^N s_i\right)^{1/N} (2\sqrt{3})^{-1} \tag{3.27}$$

Para acceder a una demostración de esta expresión se puede consultar la Sec. A.2.4. Luego las FI se definen como:

$$FI_k \triangleq \{\sigma_k^* \in \text{MI} | T_h < ||\nabla \sigma_k^*||_2\} \tag{3.28}$$

siendo T_h un parámetro que determina el máximo salto en incertidumbre esperado. Mientras más pequeño sea este valor, más sensible será la detección de FI. El valor de T_h depende del sistema de medición y de σ^*_{max} seleccionado, en efecto si $T_h > \sigma^*_{\text{max}}$ no se generarán FI. Al igual que las FC, las FI se generan a partir de las

discontinuidades del FOV (por sus siglas en inglés de *Campo de Visión*[11]), ya sea por un acotado ángulo de visión, por el rango del sensor o por la oclusión generada por algún obstáculo. A diferencia de las FC, para eliminar la FI se requiere que el sensor vuelva a medir la celda con una incertidumbre menor, suavizando la discontinuidad. Sin embargo, no todas las FI son consideradas objetivos de exploración. Se descartan aquellas que pertenecen a un obstáculo, ya que no es posible eliminarlas, y también se descartan si $\sigma_k^* > \sigma_{\max}^*$. Esta última condición mantiene acotada la creación de FI en aquellas regiones en las que la incertidumbre está por debajo de la máxima tolerable.

Si bien es necesario un mayor análisis para determinar qué valor de T_h es adecuado para cierta aplicación, en la práctica se observa que eligiendo:

$$T_h = e^{\overline{l_{th}}} \tag{3.29}$$

$$\overline{l_{th}} = \frac{1}{N} \sum_{k=1}^{N} \log \left(||\nabla \sigma_k^*||_2 \right) \tag{3.30}$$

se obtienen resultados aceptables.

Es fundamental distinguir este Mapa de Incertidumbre (MI o el mapa de PD subyacente) de los mapas de ocupación grillados tradicionales. Mientras que un mapa de ocupación almacena una estimación del *estado* de cada celda (ej., $p(\text{ocupado})$), el mapa propuesto aquí almacena una medida de la *incertidumbre* o *certeza* sobre la información asociada a esa celda (sea su estado de ocupación, su reflectancia, la pose del robot que la observó, etc.), cuantificada por p_k o σ_k^*. Por lo tanto, en una implementación completa, este mapa de incertidumbre coexistiría o complementaría a un mapa que represente el estado estimado del entorno (como un mapa de ocupación o una nube de puntos etiquetada), proporcionando información crucial sobre la fiabilidad de dicha representación del estado.

3.6. Análisis del costo computacional

En esta sección se analiza el costo computacional de los métodos propuestos. Para calcular la medida PD en Ec. 3.7, o bien para cualquier otra distribución, se puede emplear una tabla precomputada (si las distribuciones son conocidas) o realizar una simulación Monte Carlo (MC). En el peor de los casos, computar una sola medida requiere $\mathcal{O}(k\, d^3)$ operaciones mediante MC, donde k es el número de muestras y d es la dimensión de la matriz de covarianza. A continuación, para

[11] Field Of View

cada celda afectada por dicha medida, se deben ejecutar los procedimientos de actualización de log-odds descritos en Ec. 3.3 y Ec. 3.16, incurriendo en un costo de $\mathcal{O}(N)$. Aquí, N es el número de celdas que se modifican a partir de la medición del sensor, valor que depende del campo de visión (*FoV*) del mismo. Si se calcula una medida de incertidumbre distinta para cada celda (por ejemplo, en el caso de cámaras), el costo computacional total pasa a ser $\mathcal{O}(N\,k\,d^3)$.

Por otra parte, el costo computacional del método *ray-casting* depende de la implementación específica. En una estructura basada en grafos, puede ser tan bajo como $\mathcal{O}(\log N)$, aunque esta complejidad varía según la forma en que se discretiza y consulte el entorno.

Para calcular la medida SiREn, las ecuaciones Ec. 3.3 y Ec. 3.22 se aplican a todas las celdas del mapa, denotadas por $\mathbb{N}$. En cuanto a la medida FI sobre una grilla de PD de $\mathbb{N}$ celdas, primero se debe obtener σ_k^* usando Ec. 3.3 (a partir de p_k o $l_{t,k}$ con Ec. 3.3). Finalmente, para determinar FI se requiere calcular el gradiente de σ_k^*. Todas estas etapas conllevan un costo de $\mathcal{O}(\mathbb{N})$.

Además, el costo computacional puede reducirse almacenando el mapa en una estructura de datos jerárquica, como un *octree*, limitando así el número de operaciones a aquellas celdas que pertenecen a las regiones actualizadas.

En conclusión, el peor caso de los métodos propuestos presenta un costo computacional de $\mathcal{O}(N)$, donde N depende del campo de visión del sensor. El cálculo de SiREn y FI sobre todo el mapa también posee una complejidad de $\mathcal{O}(\mathbb{N})$. Estos valores resultan comparables con los métodos clásicos utilizados en sistemas de SLAM activo.

Capítulo 4

Desarrollo experimental: Casos de estudio

El problema de Active SLAM puede interpretarse como un estimador de estados, cuyo objetivo es construir el mapa. Para lograrlo, se requieren diversas acciones de control sobre el sistema que permitan que ciertos estados se vuelvan observables.

Desafortunadamente, determinar las acciones de control óptimas es un problema NP-hard Tzoumas et al. (2017), lo que implica que encontrar la solución exacta es computacionalmente intratable para problemas de tamaño realista. Por lo tanto, se requiere de enfoques heurísticos para abordarlo. En el contexto de esta tesis, se propone y evalúa un nuevo marco basado en el modelado de la incertidumbre mediante las Probabilidades de Dispersión, el índice SiREn y las Fronteras de Incertidumbre, argumentando que este conjunto de herramientas facilita el diseño de heurísticas informadas para guiar el proceso ASLAM.

En este sentido, en la Sec. 4.1 se presenta una serie de ensayos en entornos de simulación con el fin de verificar la utilidad de este enfoque. Estas pruebas en 2D permiten una validación más rápida y reducen la complejidad inicial antes de extender el análisis a entornos 3D. Si bien los resultados son particulares para un entorno y para una configuración específica de robot y sensores, el método propuesto es independiente de los sistemas subyacentes, por lo que se espera que esta evidencia sea suficiente para extraer conclusiones de carácter general.

Por otro lado, en la Sec. 4.2, a partir de una trayectoria 3D cuya incertidumbre se estima en instantes de tiempo discretos, se implementa el marco de incertidumbre propuesto para distintos valores de σ_{max}[1]. Con ello se pretende mostrar que el

[1]A lo largo de la discusión experimental, se utilizará σ_{max} para referirse al umbral de máxima incertidumbre tolerable. Existe una relación directa entre $\mathbf{\Sigma}_Q$ (definida por σ_{max}) y σ^*_{max} a través de la Probabilidad de Dispersión β asociada: $\beta = P(\mathbf{X} \in A \,|\, \mathbf{\Sigma}_Q)$ y $\sigma^*_{max} = a/\beta^{1/N}$. Por

conocimiento adquirido en la etapa de simulación es aplicable a un entorno 3D real. No obstante, en esta sección no se aborda la planificación de trayectorias necesaria para un sistema de ASLAM completo, ya que, como se mencionó anteriormente, la planificación en 3D es más costosa que en 2D. Además, el seguimiento de dichas trayectorias requiere un agente capaz de ejecutarlas, lo cual queda fuera del alcance de esta tesis.

Los sistemas de SLAM empleados tanto en la simulación como en los experimentos con datos 3D se basan en el filtro de Kalman, dado que este proporciona la matriz de covarianza como subproducto de la estimación. En ambos casos, las soluciones al problema de SLAM fueron implementadas desde cero en el marco de esta tesis. Si bien en 2D el problema se considera mayormente resuelto y, por ende, no se aporta novedad a la literatura, la situación difiere en la reconstrucción 3D de entornos de agricultura intensiva, la cual continúa siendo un problema abierto y en constante evolución, tal como se evidenció en la revisión bibliográfica de la Sec. 2.7.

En este contexto, y específicamente para los experimentos 3D en entornos agrícolas presentados en la Sec. 4.2, se desarrolló una metodología propia para la estimación de poses 6DOF y su incertidumbre asociada. Esta propuesta combina la explotación de la estructura regular del cultivo (suelo y filas de árboles) para estimar robustamente cinco grados de libertad (pitch, roll, yaw, y, z) con el uso de landmarks artificiales (conos) para la estimación del grado de libertad longitudinal (x) mediante un KF-SLAM 2D. Aunque el sistema se basa en un Filtro de Kalman lineal para la parte longitudinal y prescinde del uso de una Unidad de Medición Inercial (IMU), lo cual simplifica la arquitectura sensorial, se mostrará que esta combinación metodológica alcanza resultados cualitativamente comparables o superiores a los obtenidos con FAST-LIO2 en este entorno específico, tanto en precisión visual de la reconstrucción como en la capacidad de proporcionar la estimación de incertidumbre necesaria para nuestro marco teórico.

Cabe destacar que, si bien el marco teórico (PD, FI, SiREn) y la calidad del SLAM subyacente son fundamentales, la efectividad de cualquier sistema ASLAM depende críticamente de la estrategia de planificación de trayectorias empleada. La elección de la próxima acción de control no solo afecta la nueva información adquirida, sino también la evolución de la incertidumbre del propio agente. Esta interdependencia se vuelve más compleja y computacionalmente costosa en entornos 3D, donde el espacio de estados y acciones es significativamente mayor que en 2D[2]. Los

simplicidad notacional, nos referiremos principalmente a $\sigma_{\max}$ como el parámetro que controla el nivel de incertidumbre aceptable aunque en algunos casos corresponde que sea σ^*_{max}.

[2]Para 2D, los estados posibles son (x, y, ϕ), mientras que para 3D se tienen

experimentos presentados en este capítulo explorarán aspectos de esta interacción, aunque sin abordar el diseño de planificadores 3D óptimos, lo cual queda fuera del alcance de este trabajo.

($x, y, z, roll, pitch, yaw$), ambos restringidos por el modelo de movimiento.

4.1 Implementación en entornos de simulación 78
4.1.1 Pruebas de las Fronteras de Incertidumbre 79
4.1.2 Prueba del índice SiREn y de las FI 81
4.1.3 Resultados y Discusión 83
4.1.4 Conclusiones Preliminares de los Experimentos en Simulación 88
4.2 Implementación en entornos reales 3D 90
4.2.1 Sistema de Localización 3D con Estimación de Incertidumbre para Entornos Agrícolas 90
4.2.2 Estimación de la Posición Longitudinal mediante Landmarks Artificiales 95
4.2.3 Metodología Experimental 3D 101
4.2.4 Resultados y discusión 104

4.1. Implementación en entornos de simulación

En esta sección se describen las simulaciones de un sistema SLAM 2D implementado en un entorno tridimensional sobre un vehículo que se mueve en el plano. Los resultados, publicados originalmente en Sansoni and Gimenez (2025), corresponden a la Sección 5 y subsecuentes de dicho artículo.

El objetivo principal consiste en analizar el comportamiento del modelado de mapa propuesto, la entropía relativa y las fronteras de incertidumbre. Para ello se diseñaron dos experimentos representativos, cada uno con propósitos bien delimitados. Asimismo, se evalúa el desempeño del conjunto de herramientas desarrolladas en distintos escenarios mediante la implementación de un sistema ASLAM completo: SLAM, planificador de trayectorias, controlador de bajo nivel y controlador de seguimiento de alto nivel, entre otros subsistemas. No obstante, dado que el interés se centra en un componente específico, los módulos restantes se mantienen deliberadamente simples para aislar las relaciones de causa-efecto introducidas por las modificaciones propuestas.

Cabe aclarar que el propósito de este capítulo no es desarrollar una implementación exhaustiva de Active SLAM ni optimizar el planeamiento en escenarios complejos, sino proponer y validar metodologías que mejoren aspectos específicos de dichos sistemas. En consecuencia, se utilizan entornos experimentales simplificados, lo que permite estudiar de forma controlada las

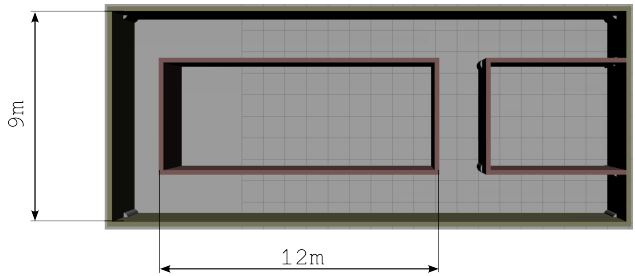


Figura 4.1: Entorno de simulación tridimensional utilizado para probar las FI. Los pasillos tienen un ancho de 2m y el largo total es superior a 12m.

interacciones entre variables clave sin las complicaciones geométricas ni las restricciones inherentes a escenarios reales.

El primer experimento ilustra el comportamiento de las FI, detalla su generación y analiza su dependencia respecto del parámetro T_h. El segundo experimento explora la relación entre entropía relativa, incertidumbre del mapa y cobertura: se comparan cuatro planificadores de trayectorias; dos emplean las FC de acuerdo con Yamauchi (1997) y los otros dos se basan en las FI.

Todos los experimentos se realizan con un robot diferencial simulado en *Gazebo*. El robot incorpora un LiDAR 2D montado sobre su estructura, capaz de medir distancias de hasta 5m con una cobertura de $360°$. La configuración del planificador consciente de la incertidumbre se presenta en la Sec. A.1.1, mientras que el sistema *KF-SLAM*, utilizado para obtener estimaciones de cada pose junto con su correspondiente matriz de varianza-covarianza, se detalla en el Sec. A.1.2.

4.1.1. Pruebas de las Fronteras de Incertidumbre

La primera simulación tiene por objeto demostrar las ventajas de las FI en comparación con el enfoque de FC. El robot se sitúa dentro de un edificio, tal como se aprecia en la Fig. 4.1, y se desplaza desde el punto P_i hasta el punto P_f mediante teleoperación, siguiendo la trayectoria indicada por la línea punteada en la Fig. 4.2.

Al iniciar el movimiento y al llegar a P_f, el FOV del LiDAR no alcanza a detectar las paredes del pasillo final ubicado a la derecha. Los colores en la figura representan los niveles de incertidumbre almacenados en el MI, donde las zonas azules indican menor incertidumbre que las zonas rojas (pasando por verde y luego por amarillo). Tanto las fronteras FC como FI que se encuentran a menos de 0.5m de un obstáculo

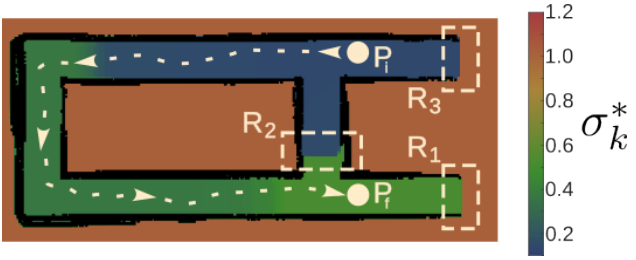


Figura 4.2: Entorno de simulación tridimensional utilizado en el primer experimento. El robot se desdeplaza desde P_i hasta P_f siguiendo la trayectoria de la línea punteada. Las regiones señaladas como R_1, R_2 y R_3 denotan incrementos abruptos del MI.

son descartadas.

Se establecen los siguientes parámetros: $\sigma_{\max} = 1.0$, $s_1 = s_2 = 0.1\,\text{m}$ y $T_h = 0.2$. Las tres regiones resaltadas, R_1, R_2 y R_3, son reconocidas como FI siguiendo el procedimiento descrito en la Sec. 3.5. Sus gradientes promedio son 0.28, 0.36 y 0.45, respectivamente; por lo tanto, un umbral T_h mayor que 0.28 excluiría R_1 de la clasificación FI. De acuerdo con la metodología de Yamauchi (1997), solo R_1 y R_3 se catalogan como FC.

Dependiendo de los objetivos detectados, el planificador selecciona una trayectoria. Cuando se utiliza el método de FC, R_1 y R_3 son designadas como objetivos de exploración. Sin embargo, al utilizar FI, R_2 también se convierte en un objetivo de exploración, lo que podría conducir a un cierre de lazo. El planificador debe determinar el camino que maximice la extracción de información, ya sea mediante exploración o explotación. Cuando el FOV del agente alcanza una FC, esta es eliminada, y potencialmente se generan nuevas fronteras si no hay obstáculos presentes. La situación es diferente en el caso de las FI: esta es eliminada si se cumple $||\nabla\sigma_k^*||_2 < T_h$, lo cual ocurre cuando la incertidumbre máxima en las celdas que la componen se reduce por debajo del umbral . En consecuencia, la planificación de trayectorias debe tener en cuenta la incertidumbre tanto del robot como de las mediciones al alcanzar una FI. La Sec. A.1.1 presenta un planificador que incluye un mecanismo de eliminación de FI. Se requiere además un sistema SLAM con cierre de lazos; de lo contrario, la incertidumbre del agente crecerá a medida que se aleje de la posición inicial.

En conclusión, a diferencia de las FC, las FI no solo sirven como metas

de exploración: su propósito principal es disminuir la incertidumbre del mapa mediante cierre de lazos y reexploración de zonas previamente visitadas con menor incertidumbre.

4.1.2. Prueba del índice SiREn y de las FI

Este segundo experimento tiene como objetivo resaltar los beneficios del uso de la entropía relativa (SiREn) y su relación con otras métricas clave, como la cobertura del mapa y la incertidumbre de la estimación, tanto para la construcción del mapa como para la localización de marcas de referencia (landmarks). Además, ofrece una comparación entre sistemas de planificación que utilizan FC y aquellos que emplean FI.

Para ello, el robot se ubica en un espacio cerrado de forma rectangular con marcas de referencia dispuestas a lo largo de su perímetro, como se muestra en la Fig. 4.3. El entorno presenta un espacio abierto en el sector derecho y pasillos en el sector izquierdo. Este escenario está diseñado para evaluar el rendimiento del sistema en un entorno que combina dos tipos de áreas (abierta y con pasillos), siendo lo suficientemente simple como para aislar la relación causa-efecto de las modificaciones propuestas.

Para la evaluación experimental, se definieron cuatro configuraciones de parámetros distintas, denominadas L_1, L_2, L_3 y L_4 (donde L hace referencia a *layout*). Estas configuraciones se diferencian en dos parámetros principales:

1. Posición inicial del agente: El agente comienza su recorrido desde la posición P_1 (definida como *superior* en la Fig. 4.3) o desde la posición P_2 (definida como *inferior* en la Fig. 4.3).

2. Disponibilidad de las marcas de referencia (*landmarks*): Se consideran dos escenarios:

   - Configuración *completa*: Todas las marcas de referencia mostradas en el perímetro del entorno (Fig. 4.3) están disponibles para el agente.
   - Configuración *incompleta*: Se simula la ausencia de una marca de referencia específica, concretamente aquella señalada con un círculo verde en la Fig. 4.3 (ubicada en la esquina inferior derecha).

Así, las cuatro configuraciones L_1 a L_4 representan las combinaciones posibles de estos dos parámetros, siendo: L_1 configuración incompleta inferior, L_2

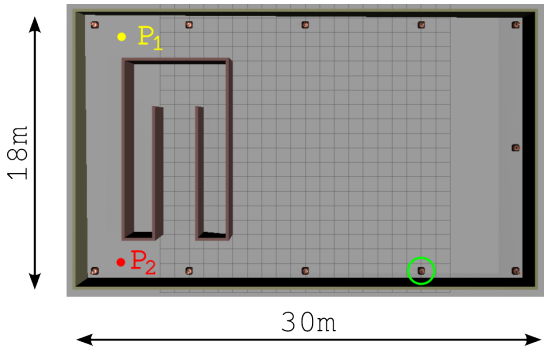


Figura 4.3: Escenario propuesto para evaluar el índice SiREn y las FI. Se identifican como P_1 (superior) y P_2 (inferior) las posiciones iniciales del agente. Además se muestran las marcas de referencia sobre el perímetro del entorno en color rojo y blanco (son conos de tránsito vistos desde arriba). El círculo verde resalta una marca que se excluye en la configuración *incompleta* del experimento, diseñada para probar el sistema bajo información parcial.

configuración incompleta superior, L_3 configuración completa superior y L_4 configuración completa inferior.

Por otro lado, se proponen cuatro sistemas de planificación de trayectorias (PPS, del inglés *Path Planning Systems*) con fines de comparación y análisis:

- PPS1: El sistema más simple. Utiliza las FC como objetivos de exploración y emplea un planificador basado en *Rapidly Exploring Random Trees* (RRT) para generar la trayectoria hacia la frontera alcanzable más cercana (estrategia codiciosa o greedy). Se selecciona el camino más corto encontrado por el RRT hacia esa frontera objetivo.

- PPS2: Similar a PPS1 en el uso de FC, pero implementa una versión modificada del planificador RRT* que considera la incertidumbre (asociada a la pose del robot y a las mediciones esperadas) dentro del proceso de planificación, como se describe en la Sec. A.1.1.

- PPS3: Idéntico a PPS2 en cuanto al planificador subyacente (RRT modificado)*, pero utiliza FI como objetivos de exploración con un valor de $\sigma_{\max} = 0.6$.

- PPS4: Similar a PPS3, pero utiliza un valor diferente para el parámetro clave de las FI: $\sigma_{\max} = 1.0$.

En todos los casos, la incertidumbre inicial del agente se fija en $\tilde{\sigma}_0 = 0.1$ y se utiliza un umbral $T_h = 0.2$ para la detección de FI. Dado que todos los PPS se basan en planificadores con componentes aleatorios (RRT/RRT*), cada condición experimental se repite un mínimo de cinco veces para mitigar los efectos del azar y obtener resultados estadísticamente más robustos. Como criterio de parada, todos los PPS se detienen cuando no quedan más fronteras objetivo (FI o FC, según corresponda) consideradas alcanzables por el planificador[3]. En total, se definen 16 condiciones experimentales distintas (2 posiciones iniciales $\times$ 2 configuraciones de landmarks $\times$ 4 estrategias de planificación), cada una repetida al menos cinco veces. Los resultados se presentan y analizan en la sección siguiente.

4.1.3. Resultados y Discusión

En la Fig. 4.4 se muestran los valores finales del índice *SiREn* obtenidos para cada configuración experimental mediante diagramas de caja (boxplot). En términos generales, se observa que *PPS1* y *PPS2* presentan una mayor dispersión (varianza) en los valores finales de *SiREn* en comparación con *PPS3* y *PPS4*. Además, los valores medianos de *SiREn* para *PPS3* y *PPS4* tienden a ser superiores a los de *PPS1* y *PPS2*, excepto en la configuración L_1 (incompleta, inicio en P_2, usando *PPS3* con $\sigma_{\max} = 0.6$). Un MI representativo de cada experimento se presenta en la Fig. 4.5 para discutir estos resultados.

Comenzando por *PPS1*, se observa que el mapa es explorado completamente en todas las configuraciones, pero también se evidencia una notable variabilidad en las incertidumbres de medición. Esta variabilidad proviene de la estrategia codiciosa de *PPS1* para alcanzar las FC, la cual está influenciada por la aleatoriedad del planificador RRT. También se observan discontinuidades significativas en el MI, indicadas por cambios bruscos en los colores. Estas variaciones se reflejan en los valores de *SiREn* obtenidos, ya que las áreas exploradas con mayor incertidumbre (cercanas y/o mayores a $\sigma_{\max}$) son más extensas que aquellas exploradas con baja incertidumbre (marcadas en azul).

En comparación con *PPS1*, *PPS2* también logra cubrir todo el mapa, pero presenta una menor variabilidad en la incertidumbre del mapa resultante. Esta mejora se debe a que *PPS2* emplea una versión modificada de *RRT** que considera

[3]Se considera alcanzable si existe un camino válido (transitable, sin colisiones) entre el agente y la frontera.

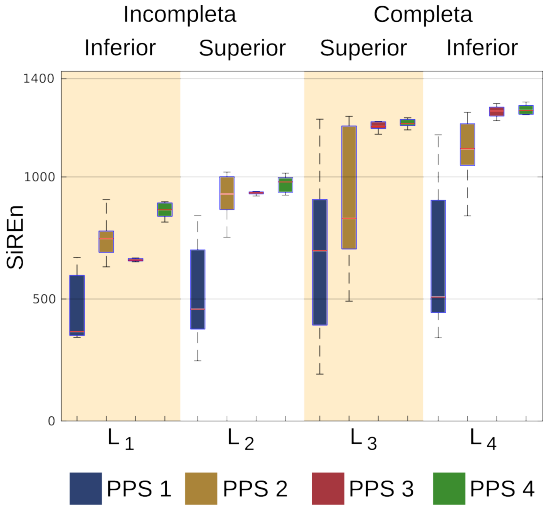


Figura 4.4: Diagramas de caja (boxplots) de los valores de *SiREn* obtenidos en todos los experimentos. Cada configuración, denotada por L_i (L de *layout*), varía según la disposición de las marcas de referencia y la posición inicial del agente: L_1 y L_2 corresponden a configuraciones incompletas, con posiciones iniciales P_2 y P_1, respectivamente. En cambio, L_3 y L_4 representan configuraciones completas, iniciando en P_1 y P_2, respectivamente.

la incertidumbre del agente en su función de costo o heurística (ver Sec. A.1.1), lo que permite al planificador elegir trayectorias a través de zonas con menor incertidumbre que contienen marcas de referencia bien conocidas. Además, si bien aún se observan discontinuidades en el MI, estas son menos abruptas y frecuentes que en *PPS1*. Este comportamiento se refleja en los valores de *SiREn* obtenidos, ya que las áreas exploradas con baja incertidumbre son más extensas en proporción (resultando en un área mayor de color azul en la Fig. 4.5), lo que conduce a valores medianos más altos de *SiREn*.

En el caso de *PPS3*, la cobertura del mapa no es total cuando se emplea una disposición incompleta de landmarks (L_1 y L_2). Esto se debe a que el sistema utiliza FI como objetivos de navegación con $\sigma_{\max} = 0.6$, lo cual provoca que el planificador considere que se ha completado el proceso de exploración y cierre de lazos en ciertas regiones una vez que la incertidumbre en las celdas del MI cae por debajo de

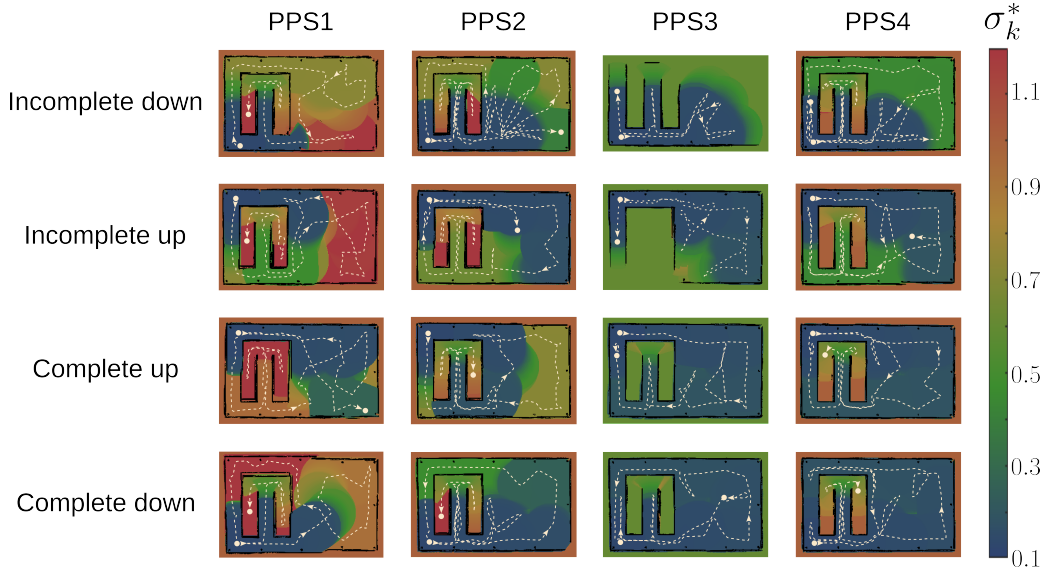


Figura 4.5: Una muestra representativa de cada configuración y sistema PPS se presenta junto con su trayectoria. El color de las celdas representa la incertidumbre estimada de la celda, de acuerdo con la barra de colores ubicada a la derecha. La línea punteada indica el recorrido seguido por el robot, empezando y terminando en un circulo.

dicho umbral. Por otro lado, visualmente, la evolución del nivel de incertidumbre en el mapa es más gradual (salvo en las zonas con obstáculos, donde se observan variaciones más abruptas en la escala de colores). *PPS3* utiliza el mismo planificador de trayectorias que *PPS2* (RRT* modificado), pero difiere en que emplea FI como objetivos de exploración, lo que explica los resultados observados. En general, *PPS3* explora zonas con menor incertidumbre que *PPS2*, alcanzando valores más altos de *SiREn* que PPS1/PPS2. Sin embargo, al operar con una disposición incompleta de landmarks, *PPS3* explora solo una parte del mapa, lo que resulta en una menor cantidad de celdas que contribuyen al cálculo de *SiREn*, en comparación con otros escenarios.

Finalmente, en el caso de *PPS4*, el mapa es ampliamente explorado en todas las configuraciones, con una variación homogénea en la incertidumbre de medición del mapa. El planificador, al igual que *PPS3*, agota los objetivos de exploración cuando la incertidumbre se aproxima a $\sigma_{\max} = 1.0$. *PPS4* cubre distancias mayores debido a que su máxima incertidumbre tolerable $\sigma_{\max}$ es mayor que la del *PPS3*. Este sistema explora un área más extensa con menor incertidumbre, lo que se traduce en valores mayores de *SiREn*. Además, aunque utiliza el mismo planificador que *PPS2*, la diferencia fundamental radica en que *PPS4* utiliza las FI como guía, las cuales, por diseño, dirigen al robot hacia regiones donde la reducción de incertidumbre es potencialmente mayor (altos gradientes de incertidumbre). Esta característica contribuye a alcanzar valores más altos de *SiREn* y menor variabilidad (ver Fig. 4.4), incluso cuando se emplea un planificador con componentes aleatorios.

Cabe destacar que la configuración y el sistema PPS que alcanzaron los valores más altos de *SiREn* corresponden a la disposición *completa* de landmarks, comenzando desde la posición P_2 (inferior) con *PPS4* (configuracuín L_4). Esto se debe a que *SiREn*, al ser una medida de entropía relativa no lineal, pondera más fuertemente la exploración con baja incertidumbre que aquella realizada en condiciones de alta incertidumbre, la cual es penalizada. Esta sensibilidad a la baja incertidumbre permite diferenciar entre posiciones iniciales[4].

En los sistemas PPS que utilizan FI (*PPS3, PPS4*), el planificador de trayectorias tiende a limitar la exploración a regiones donde la incertidumbre puede ser efectivamente reducida por debajo de los umbrales definidos, en lugar de intentar cubrir necesariamente todo el espacio. Esta estrategia resulta particularmente ventajosa en mapas de gran tamaño que no pueden ser explorados completamente en una sola sesión, o en condiciones adversas como una mala distribución de landmarks o la presencia de espacios abiertos extensos donde la localización puede

[4]Es decir, se obtienen valores distintos de SiREn cuando se cambia la posición inicial

degradarse rápidamente.

Por otro lado, con el objetivo de relacionar el MI con la incertidumbre de los landmarks, se realiza un análisis adicional sobre los mapas generados con PPS3 y PPS4. El proceso comienza con el cálculo de la mediana $\tilde{\sigma}_l$ de la incertidumbre para cada landmark a lo largo de las repeticiones de cada experimento. Esta mediana se obtiene a partir de la media geométrica de la incertidumbre asociada a cada marca, calculada como $\tilde{\sigma}_l = |\Sigma_l|^{1/(2N)}$, donde Σ_l representa la matriz de varianza-covarianza correspondiente al $l-$ésimo landmark. Una vez determinadas las medianas $\tilde{\sigma}_l$ individuales, se calcula la mediana global de estos valores entre todos los landmarks observados en una ejecución, denotada como $\bar{\sigma}_{\mathbb{L}}$. Esta mediana global sirve como una medida compuesta de la incertidumbre general de los landmarks observados al final de esa ejecución. Se encuentra que esta medida compuesta se correlaciona fuertemente (coeficiente de correlación de Pearson $\approx 97\,\%$) con la mediana de las celdas del MI final, denotada como $\bar{\sigma}_k^*$. La relación entre la mediana global $\bar{\sigma}_{\mathbb{L}}$ de los landmarks y la mediana $\bar{\sigma}_k^*$ del MI se ilustra a través de la siguiente regresión lineal:

$$\bar{\sigma}_k^* = 0.971\bar{\sigma}_{\mathbb{L}} + 0.00127. \tag{4.1}$$

Además, la pendiente es cercana a 1 y la ordenada al origen es cercana a 0, lo que indica que la incertidumbre mediana del MI refleja con precisión la incertidumbre mediana de los landmarks. En otras palabras, dentro del contexto de estos experimentos, la incertidumbre de reconstrucción asociada a $\bar{\sigma}_{\mathbb{L}}$ puede estimarse mediante $\bar{\sigma}_k^*$ con un error despreciable, y viceversa.

Basándose en los resultados presentados (principalmente Fig. 4.4), se puede intentar establecer un orden de impacto cualitativo de los parámetros experimentales sobre el valor final de *SiREn*. En primer lugar, el tipo de PPS utilizado (ya sea que tenga en cuenta la incertidumbre o no) tiene la mayor influencia, seguido por la disposición de los landmarks (completa vs. incompleta), luego por la posición inicial del agente (P_1 vs. P_2) y finalmente, por el valor de σ_{max}. Estos resultados son consistentes con el hecho ampliamente conocido de que la cantidad y distribución de landmarks afectan significativamente el proceso de reconstrucción SLAM. Además, la posición inicial del agente influye en el resultado del mapa, ya que una mala inicialización o una secuencia de exploración subóptima inicial puede generar mapas con mayor incertidumbre. Por último, el parámetro σ_{max} es el que menor impacto tiene sobre el índice SiREn en comparación con los otros parámetros.

En definitiva, si bien la cobertura del mapa influye en la entropía relativa, una mayor cobertura no implica necesariamente una mayor entropía relativa, ya que

esta depende fuertemente de la incertidumbre de las mediciones.

4.1.4. Conclusiones Preliminares de los Experimentos en Simulación

En esta sección experimental, se ha evaluado un conjunto de herramientas propuestas para sistemas ASLAM, incluyendo una metodología novedosa para modelar la incertidumbre espacial a través de las Probabilidades de Dispersión (PD). A partir de estas, se construye un Mapa de Incertidumbre (MI), del cual se derivan tanto las Fronteras de Incertidumbre (FI) como el índice de Entropía Relativa con Signo (SiREn). Este enfoque es independiente del sistema SLAM utilizado, ya que solo requiere que este último proporcione estimaciones de incertidumbre (ej. matrices de varianza-covarianza). De forma análoga, puede integrarse con cualquier tipo de sensor, incluyendo cámaras, LiDARs o sistemas de fusión sensorial, siempre que se pueda modelar adecuadamente la incertidumbre de sus mediciones y su propagación al MI.

La integración de este modelado de incertidumbre con un sistema de planificación basado en FI permite al agente explorar de forma autónoma espacios abiertos, un comportamiento no reportado previamente de manera explícita o con este enfoque en la literatura sobre Active SLAM. Asimismo, las FI han demostrado ser una herramienta eficaz para definir un criterio de parada, abordando un problema abierto identificado en trabajos previos. Por otro lado, cuando se utiliza un planificador consciente de la incertidumbre (como el RRT* modificado), se observó que el uso de FI (en PPS3 y PPS4) promueve la re-exploración de áreas previamente visitadas pero aún inciertas, facilitando así oportunidades para el cierre de trayectorias (loop closure) y mejorando potencialmente la calidad y consistencia del mapa en comparación con el mismo planificador utilizando FC.

Es importante destacar que el índice de entropía relativa, SiREn, mostró ser sensible tanto a la cobertura como a la incertidumbre de reconstrucción del mapa. Esta característica resulta esencial para comparar distintas reconstrucciones y seleccionar la trayectoria más adecuada, cumpliendo efectivamente con los objetivos fundamentales del Active SLAM. Los resultados sugieren que, a través del valor de *SiREn* asociado a cada reconstrucción, es posible establecer un orden de impacto cualitativo de los parámetros que afectan la calidad del resultado final. El orden observado en estos experimentos fue el siguiente (de mayor a menor influencia aparente): la disposición de los landmarks (completa vs. incompleta), seguido por la posición inicial del agente (P_1 vs P_2), luego por el tipo de PPS

utilizado (incluyendo la consideración de incertidumbre en el planificador) y, finalmente, el valor de $\sigma_{\max}$. Estos resultados son coherentes con la literatura general de SLAM, lo que representa otro indicio del potencial de SiREn como herramienta para ASLAM.

Cabe señalar que el objetivo de estos experimentos iniciales fue presentar y validar conceptualmente las herramientas propuestas. Por ello, los experimentos se diseñaron de forma deliberadamente simple, con el fin de aislar los efectos clave y facilitar la comprensión del concepto antes de extender el análisis a escenarios más complejos o realistas.

4.2. Implementación en entornos reales 3D

Tras haber validado los conceptos de modelado de incertidumbre, Fronteras de Incertidumbre (FI) y el índice SiREn en un entorno controlado de simulación 2D (Sec. 4.1), esta sección aborda el desafío de aplicar parte de este marco a datos reales obtenidos en un entorno 3D complejo, como es la agricultura intensiva.

El objetivo principal aquí no es implementar un sistema ASLAM completo con planificación activa de trayectorias, lo cual, como se mencionó, presenta desafíos computacionales y de actuación significativos en 3D, sino demostrar dos aspectos clave:

1. La viabilidad de generar estimaciones de pose 3D con incertidumbre asociada en un escenario agrícola real, utilizando un sistema SLAM adaptado (descrito en la Sec. 4.2.1).

2. La utilidad del marco de incertidumbre propuesto (Mapa de Incertidumbre, FI, SiREn) para analizar la calidad y completitud de la reconstrucción 3D resultante, sentando las bases para su futura integración en un sistema ASLAM 3D.

En la Sec. 4.2.1 se describen en detalle las partes del sistema de SLAM desarrollado. Luego, en la Sec. 4.2.3 se describe el hardware y la metodología de adquisición de datos. Finalmente, en la Sec. 4.2.4 se presentan los resultados utilizando esta metodología con datos reales, analizando cómo el Mapa de Incertidumbre, las FI y el índice SiREn permiten interpretar la reconstrucción bajo diferentes criterios de calidad (representados por variaciones de $\sigma_{\max}$).

4.2.1. Sistema de Localización 3D con Estimación de Incertidumbre para Entornos Agrícolas

La evaluación rigurosa del marco de incertidumbre propuesto (PD, MI, FI, SiREn) depende fundamentalmente de la disponibilidad de estimaciones de pose consistentes y accesibles, acompañada de una estima de la incertidumbre típicamente en forma de matrices de varianza-covarianza. Si bien existen sistemas SLAM 3D de última generación basados en LiDAR, como FAST-LIO2 Xu et al. (2022) o LOAM Zhang and Singh (2014) y sus variantes, su enfoque principal suele ser la precisión de la trayectoria estimada. La consistencia de sus estimaciones de incertidumbre puede ser problemática, especialmente en filtros no lineales Sola (2010), o bien la representación interna de la incertidumbre no siempre es

fácilmente integrable con nuestra metodología que requiere covarianzas explícitas para calcular las Probabilidades de Dispersión.

Aplicar SLAM en agricultura intensiva presenta desafíos adicionales que complican tanto la localización precisa como la estimación consistente de incertidumbre. La estructura altamente repetitiva de las hileras de cultivo Vougioukas (2019) y la escasez de rasgos geométricos distintivos y persistentes en el entorno natural Capua and Sansoni (2018) dificultan la robusta identificación de lugares (*place recognition*) y el cierre de lazos, problemas conocidos en entornos con baja texturización geométrica y alta ambigüedad perceptiva Cadena et al. (2016). Si bien existen trabajos que abordan SLAM específicamente en agricultura, como el enfocado en detección de troncos de Gimenez and Sansoni (2022), a menudo se centran en aspectos particulares de la localización y no priorizan la generación de la incertidumbre consistente y explícita que nuestro enfoque requiere para el análisis posterior.

Por lo tanto, para asegurar una fuente de datos 3D con incertidumbre adecuada para validar nuestro marco teórico en un entorno realista, se optó por desarrollar un sistema de localización 3D adaptado a las condiciones de la agricultura intensiva. Este sistema, que se describe a continuación, explota la estructura regular del suelo y las filas de cultivo para estimar 5 de los 6 Grados de Libertad (DoF, por sus siglas en inglés *Degrees of Freedom*) de la pose y utiliza landmarks artificiales para asegurar una estimación robusta del DoF longitudinal restante. Está diseñado para proporcionar las estimaciones de pose y covarianza necesarias para nuestro análisis, aunque reconocemos que el uso de landmarks artificiales implica que su aplicabilidad directa como sistema SLAM completo está sujeta a la presencia de dicha infraestructura o a la futura incorporación de métodos robustos de localización longitudinal basados en rasgos naturales.

El sistema desarrollado se compone de tres módulos principales que operan secuencialmente sobre cada nube de puntos adquirida, la cual puede provenir de diversos tipos de sensores capaces de generar datos 3D, como LiDAR o cámaras RGB-D: estimación del plano del suelo, detección de filas y, por último, estimación de la posición longitudinal utilizando landmarks artificiales. A continuación, se describen conceptualmente cada uno de estos componentes. Para acceder a los detalles técnicos de la implementación, se recomienda consultar el repositorio del proyecto asociado a esta tesis.[5]

[5] https://github.com/Seba-san/treeSLAM.

Estimación del plano del suelo

El primer módulo estima la orientación relativa del sensor respecto al plano del suelo π y la altura del sensor sobre este. Se asume que el suelo es localmente plano[6] y que una proporción significativa de los puntos capturados por el sensor inciden sobre él. En este sistema, se estableció heurísticamente que al menos el $30\,\%$ de la nube de puntos debe corresponder a muestras del suelo para una estimación robusta, condición que se busca satisfacer mediante la configuración de montaje y operación del sensor (ver más adelante).

Bajo esta premisa, se utiliza el algoritmo RANSAC (Random Sample Consensus) Fischler and Bolles (1981) para identificar el conjunto de puntos (*inliers*) que mejor se ajustan a un plano, presuntamente el suelo. La cantidad máxima de iteraciones de RANSAC (N) se determina para alcanzar una alta probabilidad ($p = 0.99$) de seleccionar una muestra libre de outliers, asumiendo una fracción de inliers ($\epsilon = 0.3$) y el mínimo número de puntos para definir un plano ($n = 3$):

$$N = \frac{\log(1-p)}{\log(1-\epsilon^n)} \tag{4.2}$$

resultando en un $N \approx 170$ iteraciones máximas.

En cada iteración de RANSAC, se seleccionan tres puntos de forma aleatoria ($A, B, C \in \mathbb{R}^3$) de la nube y se define el plano candidato calculando su normal $\mathbf{n} = (B-A) \times (C-A)$ y el parámetro independiente $d = -\mathbf{n} \cdot A$. A continuación, se calcula la distancia de todos los puntos de la nube a este plano candidato. Se consideran como inliers provisionales aquellos puntos cuya distancia al plano es inferior a un umbral $d_{\text{ransac}} = 0.2$ m [7]. El plano seleccionado por RANSAC es aquel que maximiza el número de inliers. Finalmente, se realiza un ajuste fino del plano utilizando únicamente el conjunto final de inliers mediante Mínimos Cuadrados, implementado eficientemente a través de la Descomposición en Valores Singulares (SVD) de la matriz de covarianza de los puntos inliers.

Una vez obtenido el plano π ajustado (con la normal $\mathbf{n}_{\text{opt}}$ normalizada y el parámetro independiente d_{opt}) se estima la orientación relativa del sensor respecto al suelo. Específicamente, se calculan los ángulos de *pitch* y *roll* que alinean $\mathbf{n}_{\text{opt}}$ con el eje vertical global $k = (0, 0, 1)$. Esto se logra calculando la matriz de rotación $\mathbf{R}$ mediante la fórmula de Rodrigues. Primero, se ajusta la orientación de $\mathbf{n}_{\text{opt}}$ para garantizar que su producto escalar con $\mathbf{k}$ no sea negativo. Luego se calcula el ángulo

[6] Como es habitual en muchos marcos de plantación de agricultura intensiva, donde el suelo se encuentra generalmente nivelado.

[7] Este es un valor empírico que tolera pequeñas irregularidades del suelo.

de rotación como $\theta = \arccos(\mathbf{n}_{\text{opt}} \cdot \mathbf{k})$ y el eje de rotación normalizado como $\mathbf{e} = \mathbf{n}_{\text{opt}} \times \mathbf{k}$. La matriz de rotación se construye como:

$$\mathbf{R} = \mathbf{I} + \sin(\theta)\mathbf{K} + (1 - \cos(\theta))\mathbf{K}^2 \tag{4.3}$$

donde $\mathbf{I}$ es la matriz identidad 3×3, y $\mathbf{K}$ es la matriz antisimétrica asociada al vector $\mathbf{e} = (e_1, e_2, e_3)$, dada por:

$$\mathbf{K} = \begin{pmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{pmatrix} \tag{4.4}$$

La altura del sensor sobre el plano ajustado, z_s se estima como la moda de las componentes z de todos los puntos de la nube después de aplicar la rotación $\mathbf{R}$.

La incertidumbre asociada a la estimación del plano π (que afecta tanto a la altura z_s como a los ángulos de pitch y roll) se cuantifica mediante el desvío estándar de los residuos de los *inliers* en el ajuste. En la práctica, estos residuos están artificialmente acotados por d_{ransac}; sin embargo, los desvíos obtenidos con esta metodología son mucho más pequeños que la distancia máxima.

Esta métrica σ_π no desglosa la incertidumbre entre altura y orientación, sino que sirve como un indicador pragmático de la calidad general del ajuste del plano en este entorno complejo. Se utilizará esta medida única en los análisis posteriores, reconociendo que una estimación separada de $\sigma^2_{pitch}, \sigma^2_{roll}, \sigma^2_{z_s}$ sería más rigurosa pero requiere un análisis más complejo.

Finalmente, todos los puntos que se encuentren a una distancia vertical (tras la rotación) mayor a $0.3\,\text{m}$ del plano ajustado se consideran parte de obstáculos (presumiblemente, el cultivo) y se utilizan como entrada para el siguiente módulo, la estimación de filas (descrita en la Sec. 4.2.1). Para facilitar la estimación robusta del plano (asegurando $\approx 30\,\%$ de puntos de suelo), se recomienda operar el sensor a una altura mínima de $0.5\,\text{m}$ y con una inclinación (*pitch*) de $\approx 20^\circ$ hacia el suelo.

En la Fig. 4.6 se muestra un ejemplo de la nube de puntos obtenida, el plano ajustado (morado), puntos pertenecientes al plano (azul) y los puntos identificados como obstáculos (verde).

Estimación de filas

Una vez identificados los puntos pertenecientes a obstáculos (presumiblemente el cultivo) en la etapa anterior, este módulo estima la posición lateral (y) y la orientación (*yaw*, Ψ) del sensor respecto a las filas de cultivo. Se asume que el cultivo

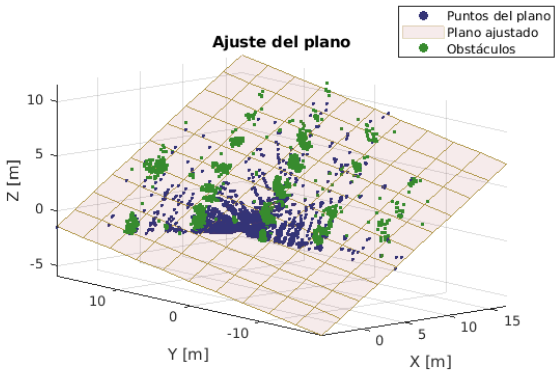


Figura 4.6: Ejemplo de la nube de puntos obtenida por el sensor, el plano ajustado y los puntos que se consideran como obstáculos.

está dispuesto en filas paralelas y que la distancia entre ellas, d_{filas}, es conocida (proveniente del marco de plantación).

El procedimiento es análogo al de la estimación del plano (Sec. 4.2.1), utilizando RANSAC para identificar los puntos que mejor se ajustan al modelo esperado, seguido de un ajuste fino por Mínimos Cuadrados. En este caso, el modelo geométrico consiste en un par de rectas paralelas en el plano xy (correspondiente a la vista superior tras la corrección de pitch/roll), separadas por la distancia d_{filas}.

Para la fase de RANSAC, se asume que al menos un $30\,\%$ de los puntos de obstáculos son inliers que pertenecen a las dos filas principales visibles por el sensor. En cada iteración, se seleccionan dos puntos aleatorios para definir una recta candidata. Luego se calcula la distancia de todos los puntos de obstáculos a esta recta y a su paralela a una distancia d_{filas}. Los puntos cuya distancia mínima a una de estas dos rectas es inferior a un umbral $d_{\text{row\_ransac}}$ se consideran inliers provisionales. Este umbral se fija en $d_{\text{row\_ransac}} = 0.8 \times (dA/2)$, donde dA es la distancia entre árboles en la fila (dato del marco de plantación). Este valor se justifica asumiendo que la copa de un árbol puede aproximarse por una esfera o un elipsoide, y que la distancia entre árboles adyacentes, d_A, es conocida a partir del marco de plantación. Así, el radio máximo de la copa se estima como $r_A = d_A/2$. Sin embargo, para dotar de mayor robustez al sistema, se adopta el valor $T_h = 0.8 d_A/2$. Dada la posible variabilidad

en la forma de los árboles, este umbral puede ser relativamente grande, por lo que su función principal es discriminar puntos muy alejados o pertenecientes a otras estructuras, más que limitar estrictamente la dispersión de los inliers. El par de rectas que maximiza el número de inliers es seleccionado por RANSAC.

Tras RANSAC, se realiza un ajuste fino de los parámetros del par de rectas (su orientación común Ψ y la posición lateral y) utilizando Mínimos Cuadrados sobre los inliers finales, minimizando la suma de las distancias al cuadrado a la recta más cercana del par.

Finalmente, la incertidumbre de este ajuste se cuantifica mediante la desviación estándar de los residuos (distancias mínimas) de los inliers finales respecto al par de rectas ajustado, σ_{row}, calculada como:

$$\sigma_{\text{row}} = \sqrt{\frac{1}{N_{\text{row\_inliers}} - 1} \sum_{i \in \text{Inliers}} \text{mín}(d_{1,i}, d_{2,i})^2} \tag{4.5}$$

donde $d_{1,i}$ y $d_{2,i}$ son las distancias del i-ésimo inlier a las dos rectas finales ajustadas, y $N_{\text{row\_inliers}}$ es el número de dichos inliers. Esta métrica σ_{row} sirve como un indicador de la calidad del ajuste sobre los puntos identificados como pertenecientes a las filas. Dado que el umbral RANSAC puede ser amplio, σ_{row} refleja principalmente la dispersión de los puntos considerados inliers. Al igual que σ_π para el plano, no desglosa la incertidumbre entre la orientación (Ψ) y la posición lateral (y). Por simplicidad, se utilizará σ_{row} como la medida de incertidumbre asociada a esta etapa en los análisis posteriores.

4.2.2. Estimación de la Posición Longitudinal mediante Landmarks Artificiales

En esta instancia del sistema, se tienen estimados cinco de los seis estados del agente: *roll*, *pitch*, *yaw*, y y z , junto con su respectiva incertidumbre de estimación. Estos estados se encuentran definidos con respecto al marco de plantación, y presentan varias características deseables: no dependen de estimaciones anteriores, son obtenidos mediante algoritmos robustos, incluyen una estimación de la incertidumbre no acumulativa (es decir, los errores de estimación no se acumulan en el tiempo), y no dependen del tipo de sensor utilizado (LiDAR o RGB-D). Como entrada, el sistema requiere únicamente una nube de puntos que cumpla ciertos requisitos en su composición.

La estimación precisa de la posición longitudinal (x), presenta un desafío

fundamental distinto. Una vez corregida la orientación y posición relativa a las filas, el problema se reduce a estimar el avance a lo largo de un corredor definido por el cultivo. En ausencia de características distintivas o cierres de lazo a gran escala, este escenario es propenso a la acumulación irrestricta de error longitudinal, análogo al problema de SLAM en entornos con alta simetría traslacional o características repetitivas Cadena et al. (2016), Gimenez and Sansoni (2022), Sansoni and Moreyra (2019), Capua and Sansoni (2018).

Si bien los árboles presentes en las filas podrían, en teoría, servir como landmarks naturales para mitigar esta deriva, nuestras pruebas preliminares indican dificultades prácticas significativas para lograr la robustez y, sobre todo, la consistencia en la estimación de incertidumbre requerida por esta tesis. Específicamente:

- Segmentación Ambigua: La proximidad y similitud entre árboles, sumada a la variabilidad del follaje y oclusiones, impide segmentar de forma fiable y consistente los puntos pertenecientes a un único árbol a lo largo de diferentes vistas.

- Puntos de Referencia Inestables: Definir un punto 3D geométricamente estable y observable de forma repetible sobre la estructura de un árbol (base del tronco, centroide de la copa visible, etc.) resulta altamente problemático, generando mediciones ruidosas o inconsistentes.

- Asociación de Datos: La similitud entre árboles consecutivos incrementa el riesgo de asociaciones incorrectas (confundir un árbol con el siguiente), un problema crítico que puede llevar a fallos en el SLAM.

Estas dificultades inherentes a la utilización de landmarks naturales en este entorno comprometen directamente la capacidad de obtener estimaciones de x con incertidumbre bien caracterizada y fiable. Para los objetivos de esta tesis, donde se busca evaluar rigurosamente el marco propuesto (MI, FI, SiREn) sobre una trayectoria 3D con incertidumbre conocida, era indispensable contar con una estimación de pose 6DOF lo más robusta y consistente posible. Por esta razón, se tomó la decisión metodológica de introducir landmarks artificiales (seis conos de tráfico, ver Fig. 4.7) espaciados ($\approx$ 7m) a lo largo del recorrido. Estos conos, fácilmente detectables y segmentables, proporcionan referencias espaciales inequívocas que permiten una estimación robusta y con incertidumbre caracterizable de la posición (x, y) del sensor mediante un filtro KF-SLAM 2D estándar.



Figura 4.7: Imagen del entorno real utilizado para probar el sistema de SLAM.

Si bien esta instrumentación del entorno simplifica el desafío de la localización longitudinal inherente a la agricultura intensiva, resulta crucial para poder aislar el fenómeno bajo estudio: la capacidad de nuestro marco teórico (MI, FI, SiREn) para modelar y analizar la incertidumbre de una reconstrucción 3D, alimentado por datos de pose 6DOF con la mejor calidad de incertidumbre posible bajo estas condiciones controladas. La evaluación de la aplicabilidad de nuestro marco en sistemas SLAM que utilicen exclusivamente landmarks naturales queda como un importante trabajo futuro.

El procedimiento implementado para estimar la posición (x, y) utilizando los conos como landmarks es el siguiente:

1. Segmentación de Puntos Candidatos a Cono: Para cada nube de puntos (ya rotada y alineada con el suelo y las filas), se identifican los puntos candidatos a pertenecer a los conos. Esto se realiza excluyendo aquellos puntos ya clasificados como suelo (inliers del plano π) y como filas (inliers de las rectas paralelas en la Sec. 4.2.1). Los puntos restantes son considerados candidatos. Adicionalmente, se aplica un filtro de altura para descartar puntos que esten por debajo de la altura esperada de los conos.

2. Clusterización: Sobre el conjunto de puntos candidatos, se aplica un algoritmo de clusterización basado en densidad, específicamente DBSCAN[8] Schubert

[8] *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN)

et al. (2017), con parámetros $\epsilon = 1.0\,\text{m}$ y MinPts $= 3$. Este paso agrupa los puntos que pertenecen a cada objeto discreto (presumiblemente, cada cono visible).

3. Extracción de Mediciones 2D y su Incertidumbre: Para cada clúster j resultante del DBSCAN, se calcula:

   - El centroide 2D $(\hat{x}_{c,j}, \hat{y}_{c,j})$ de los puntos del clúster, proyectados sobre el plano xy del sensor. Esta es la medición $\mathbf{z}_j$.
   - La matriz de covarianza 2×2, $\mathbf{R}_{c,j}$, asociada a esta medición. Esta se estima empíricamente a partir de la dispersión de los puntos (x, y) dentro del clúster j. Esta $\mathbf{R}_{c,j}$ representa la incertidumbre de la medición $\mathbf{z}_j$.

4. Procesamiento KF-SLAM 2D: Las mediciones válidas $\mathbf{z}_j$ y sus covarianzas estimadas $\mathbf{R}_{c,j}$ se introducen como entradas al filtro KF-SLAM 2D. Este filtro estima el estado $\mathbf{x} = [x_r, y_r, x_{L1}, y_{L1}, \dots]^T$ y su covarianza $\mathbf{P}$. Los detalles del filtro son:

   - Modelo de Predicción: Se asume un modelo de movimiento de caminata aleatoria para el robot (estado $[x_r, y_r]$), simplemente añadiendo la covarianza del ruido de proceso $\mathbf{Q_r} = \text{diag}([(0.3\,\text{m})^2, (0.3\,\text{m})^2])$ a la covarianza del robot en cada paso de predicción. Los landmarks (conos) se asumen estáticos.
   - Modelo de Medición: La medición esperada para el landmark k es $\mathbf{h}_k(\mathbf{x}) = [x_{Lk} - x_r, y_{Lk} - y_r]^T$. La matriz Jacobiana $\mathbf{H}_k$ utilizada en la actualización se deriva de este modelo.
   - Asociación y Gestión de Landmarks: Se utiliza una lógica de asociación de datos basada en distancia euclidiana y validación de consistencia para corresponder mediciones $\mathbf{z}_j$ con landmarks $\mathbf{L}_k$ existentes. Se inicializan nuevos landmarks para mediciones no asociadas, utilizando $\mathbf{R}_{c,j}$ como covarianza de medición para la inicialización de la covarianza del nuevo landmark.
   - Actualización: Se realiza la actualización estándar de Kalman, utilizando $\mathbf{R}_{c,j}$ como la covarianza de la medición $\mathbf{z}_j$ en el cálculo de la ganancia y la actualización de la covarianza del estado $\mathbf{P}$. Se incluye un test Chi-cuadrado para rechazar actualizaciones inconsistentes.

- Covarianza Inicial: La covarianza inicial del robot $\mathbf{P}_0(1:2, 1:2)$ se fijó como $\text{diag}([(0.05\,\text{m})^2, (0.05\,\text{m})^2])$.

La salida principal de este módulo para cada instante es la estimación de la pose 2D del robot (x_r, y_r) y su correspondiente submatriz de covarianza 2×2, $\mathbf{P}_r = \mathbf{P}(1:2, 1:2)$.

La pose 6DOF completa del sensor en el marco global se obtiene combinando los resultados de los tres módulos:

- La posición (x_r, y_r) proviene del estado estimado por el KF-SLAM 2D.
- La altura z_s proviene de la estimación del plano del suelo.
- La orientación (representada, por ejemplo, por ángulos de Euler θ_p, θ_r, ψ) se obtiene componiendo la rotación del plano (pitch/roll) y la rotación de las filas (yaw).

Combinación Final de Pose 6DOF y Cuantificación de Incertidumbre mediante PD

La pose 6DOF completa del sensor en el marco global, $\mathbf{x}_{6D} = [x_r, y_r, z_s, \theta_p, \theta_r, \psi]^T$, se obtiene en cada instante combinando las estimaciones de los tres módulos descritos: la posición (x_r, y_r) del KF-SLAM 2D (Sec. 4.2.2), la altura z_s del ajuste del plano (Sec. 4.2.1), y la orientación $(\theta_p, \theta_r, \psi)$ compuesta a partir de las rotaciones del plano y las filas. El siguiente paso crucial es cuantificar la incertidumbre asociada a esta pose $\mathbf{x}_{6D}$ utilizando el marco de Probabilidades de Dispersión (PD) desarrollado en esta tesis.

Desde una perspectiva teórica rigurosa, la cuantificación de la incertidumbre 6DOF se realizaría estimando la matriz de covarianza completa $\mathbf{\Sigma}_{6\times6}$ asociada a $\mathbf{x}_{6D}$. Esta matriz capturaría no solo las varianzas individuales de cada componente de la pose, sino también las correlaciones entre ellas, resultantes de la propagación formal de incertidumbres a través de los módulos de estimación (idealmente, obteniendo $\sigma^2_{z_s}, \sigma^2_{\theta_p}, \sigma^2_{\theta_r}, \sigma^2_{\psi}$ a partir del análisis estadístico de los ajustes geométricos y combinándolas adecuadamente con $\mathbf{P}_r$).

Con $\mathbf{\Sigma}_{6\times6}$ disponible, y asumiendo una distribución aproximadamente Gaussiana para la pose $\mathcal{N}(\mu, \mathbf{\Sigma}_{6\times6})$, la Probabilidad de Dispersión PD_{pose} se calcularía directamente a partir de su definición general (Sec. 3.2). Esto implica evaluar la probabilidad de que la pose se encuentre dentro de un hiper-rectángulo

A_{6D} de lados predefinidos por los s_i:

$$A_{6D} = \{\mathbf{x} \in \mathbb{R}^6 : |x_i - \mu_i| < s_i/2, \forall i = 1..6\} \tag{4.6}$$

$$PD_{\text{pose}} = \int_{A_{6D}} \mathcal{N}(\mu, \mathbf{\Sigma}_{6\times6})(\mathbf{w}) d\mathbf{w} \tag{4.7}$$

Sin embargo, la estimación rigurosa de todas las componentes (y correlaciones) de $\mathbf{\Sigma}_{6\times6}$ presenta desafíos prácticos significativos, especialmente dadas las heurísticas utilizadas para estimar la calidad de los ajustes del plano (σ_π) y las filas (σ_{row}) en nuestra implementación actual.

Ante las dificultades para obtener una $\mathbf{\Sigma}_{6\times6}$ formal completa, se adopta un enfoque pragmático que aprovecha la fortaleza clave del marco PD: su capacidad para integrar información de incertidumbre heterogénea proveniente de los distintos módulos de estimación. Se calcula la PD_{pose} combinando las PDs asociadas a los bloques de variables cuya incertidumbre sí se caracteriza:

1. PD para Posición 2D (PD_{xy}): Utilizando la covarianza $\mathbf{P}_r$ del KF-SLAM 2D y los parámetros $s_x = 0.05\,\text{m}$, $s_y = 0.05\,\text{m}$, se calcula utilizando la Ec. 3.10 en caso de que σ_x no sea independiente de σ_y. En caso contrario se puede utilizar la Ec. 3.5 haciendo:

$$PD_{xy} = [2\Phi(s_x/(2\sigma_x)) - 1] \times [2\Phi(s_y/(2\sigma_y)) - 1] \tag{4.8}$$

2. PD para DoF asociados al Plano (PD_π): La incertidumbre combinada de $(z_s, \theta_p, \theta_r)$ se representa mediante σ_π . Se calcula una PD asociada usando σ_π como medida de dispersión relevante y normalizándola mediante el parámetro $s_\pi = \text{mín}(\sigma_\pi)_{\text{exp}}$, el mínimo valor típico observado experimentalmente para σ_π:

$$PD_\pi = 2\Phi\left(\frac{s_\pi}{2\sigma_\pi}\right) - 1 = 2\Phi\left(\frac{\text{mín}(\sigma_\pi)_{\text{exp}}}{2\sigma_\pi}\right) - 1 \tag{4.9}$$

   Esta PD_π cuantifica la confianza en el ajuste del plano relativa a la mejor precisión observada.

3. PD para DoF asociado a las Filas (PD_{row}): Análogamente, usando σ_{row} (Ec. 4.5) y $s_{\text{row}} = \text{mín}(\sigma_{\text{row}})_{\text{exp}}$:

$$PD_{\text{row}} \approx 2\Phi\left(\frac{s_{\text{row}}}{2\sigma_{\text{row}}}\right) - 1 = 2\Phi\left(\frac{\text{mín}(\sigma_{\text{row}})_{\text{exp}}}{2\sigma_{\text{row}}}\right) - 1 \tag{4.10}$$

   Esta PD_{row} cuantifica la confianza relativa en el ajuste de las filas.

La Probabilidad de Dispersión global para la pose 6DOF, PD_{pose}, se obtiene combinando las PDs parciales. Se asume que los errores dominantes capturados por los ajustes del plano (PD_{π}), las filas (PD_{row}) y la estimación de posición XY (PD_{xy}) son estadísticamente independientes entre sí. Este supuesto de independencia *entre bloques* se sustenta en el análisis empírico de las propias medidas de incertidumbre: la matriz de coeficientes de correlación calculada entre σ_{π}, σ_{row}, $\sigma_x = \sqrt{\mathbf{P}_r(1,1)}$ y $\sigma_y = \sqrt{\mathbf{P}_r(2,2)}$ a lo largo de la trayectoria experimental muestra correlaciones cruzadas muy bajas (inferiores a 0.05) entre (σ_{π}, σ_{row}) y (σ_x, σ_y).

Bajo esta independencia empíricamente validada entre los bloques de estimación, la PD global se puede aproximar como el producto de las PDs parciales:

$$PD_{\text{pose}} \approx PD_{xy} \times PD_{\pi} \times PD_{\text{row}} \tag{4.11}$$

donde PD_{π} y PD_{row} se calculan con las Ecs. 4.9 y 4.10, y PD_{xy} se calcula mediante la Ec. 3.10 para tener en cuenta la alta correlación observada entre x_r e y_r (coeficiente de correlación ≈ 0.94) dentro del bloque estimado por el KF-SLAM 2D.

Este método de cálculo directo por bloques demuestra la flexibilidad inherente al concepto de Probabilidad de Dispersión. Permite fusionar, de manera probabilística y adimensional, información de incertidumbre proveniente de distintas fuentes y representaciones (una covarianza 2D como $\mathbf{P}_r$ y métricas de calidad escalares como $\sigma_{\pi}, \sigma_{row}$). La normalización realizada en el cálculo de PD_{π} y PD_{row} mediante los intervalos s_{π} y s_{row} (basados en los mínimos experimentales) es clave, pues convierte las medidas de dispersión heurísticas en probabilidades comparables y significativas dentro del contexto de la precisión alcanzable por el sistema.

Aunque este enfoque pragmático no captura las correlaciones entre todos los DoF como lo haría una $\mathbf{\Sigma}_{6\times6}$ completa, proporciona una métrica unificada, interpretable ($PD_{\text{pose}} \in [0,1]$) y coherente con el marco teórico para cuantificar la confianza en la estimación 6DOF con la información disponible. Esta PD_{pose}, calculada según la Ec. 4.11, es la que se utilizará en el resto de esta tesis para construir los Mapas de Incertidumbre (MI) y calcular el índice SiREn en los análisis de resultados experimentales.

4.2.3. Metodología Experimental 3D

Para capturar las nubes de puntos, se utilizó un LiDAR 3D marca Livox, modelo Mid-360, que cuenta con un alcance de 100 m y un patrón de escaneo no repetitivo. Este tipo de patrón permite obtener una mayor cobertura del entorno cuando el

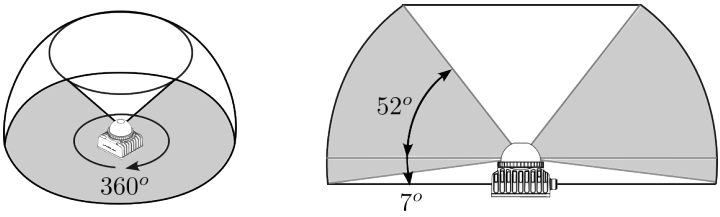


Figura 4.8: FOV del LiDAR Livox Mid-360, imagen recreada del manual de usuario.

LiDAR se encuentra estático en una posición. Por ejemplo, si se acumulan los puntos durante 0.1 s, la cobertura del FOV es equivalente a la de un LiDAR tradicional de 32 líneas; para 0.5 s, alcanza un $70\,\%$ de cobertura del FOV, lo cual es equivalente a un LiDAR tradicional de 64 líneas. Para esta prueba, se tomaron datos con una frecuencia de 10 Hz, lo que permite obtener una nube de puntos cada 0.1 s. El FOV de este LiDAR se muestra en la Fig. 4.8.

Este sensor fué montado en una estructura impresa en 3D para que pueda ser transladado y manipulado fácilmente por una persona. Por otro lado el sistema de adquisición consta de un computador portátil marca Intel modelo NUC8i3BEH con un procesador Intel Core i3-83109U @ 3Ghz $\times$ 4, y 8 GB de RAM corriendo Ubuntu 18.04.01 y ROS Noetic en un contenedor de Docker. Todo esto alimentado con una bateria de LiPO de 4 celdas de 6000 mAh 50C y 14.8V regulado a 12V mediante un conversor XL4016E1 DC/DC Buck de 8A.

El ensayo se realizó en la Estación Experimental Agropecuaria del INTA San Juan, ubicada en la ciudad de Pocito el día 8 de Enero de 2025. La plantación elegida fue un Olivar con un marco de plantación *intensiva* de $d_A \times d_{\text{filas}} = 3.5$ m $\times$ 7 m, conocida como *Parcela B*.

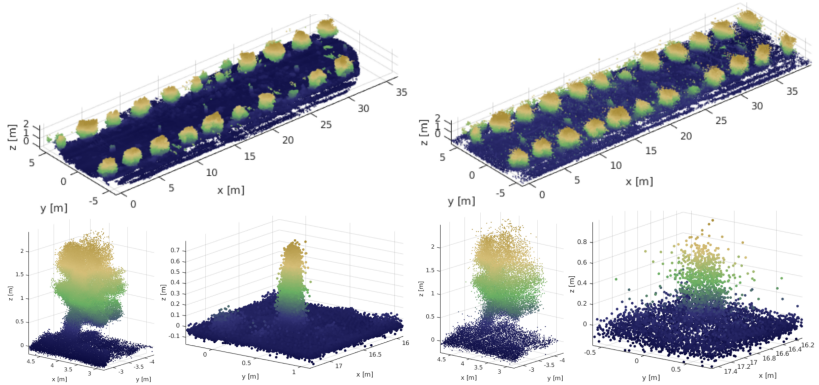
La captura de datos consistió en recorrer el ambiente mostrado en la Fig. 4.7, en un recorrido total de aproximadamente 70 m, considerando ida y vuelta, y procurando no realizar movimientos bruscos, tanto rotacionales como translacionales. Además, se tuvieron en cuenta los requisitos para que el sistema de SLAM funcione correctamente; es decir, que el sensor se encuentre a una altura mínima de 0.5 m del suelo, una inclinación mínima de $20°$, y se procuró que el FOV del LiDAR incluyera en todo momento tanto a los conos como a las dos filas de árboles simultáneamente. Hay que mencionar que no se contó con ningún tipo de realimentación que permitiera al operador saber si el sensor se encontraba en



Figura 4.9: Nube de puntos obtenida al dejar el LiDAR estático frente a un árbol.

la posición correcta o si el sistema de SLAM estaba estimando correctamente la posición.

Por otro lado, se eligieron al azar árboles y se dejo el lidar estático por más de 1 min a una distancia de 5 m de él aproximadamente. El objetivo de esto es obtener una nube de puntos con el menor ruido posible y usarlo para comparar con la nube de puntos obtenida en movimiento. En la Fig. 4.9 se muestra el árbol seleccionado.



(a) Sistema de SLAM propuesto.

(b) Sistema FAST-LIO2.

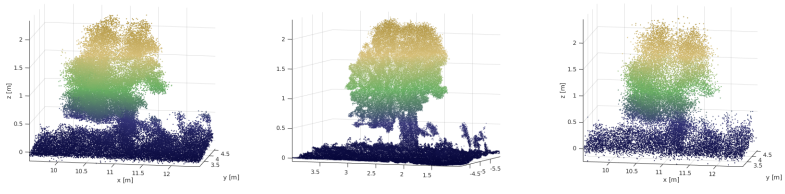
Figura 4.10: Comparación visual de reconstrucciones 3D. **Superior:** Nube de puntos completa del recorrido. **Inferior Izquierda:** Detalle de un olivo. **Inferior Derecha:** Detalle de un cono artificial. (Izquierda: SLAM propuesto, Derecha: FAST-LIO2).

4.2.4. Resultados y discusión

Con los datos capturados, se realizó la reconstrucción utilizando FAST-LIO2 Xu et al. (2022) y el sistema de SLAM propuesto en la Sec. 4.2.1, que proporciona estimaciones de pose con incertidumbre. La Fig. 4.10 presenta una comparación visual cualitativa de las nubes de puntos resultantes para todo el recorrido y detalles de un árbol y un cono. Se puede acceder a los datos mediante este link [9]. Visualmente, se observan diferencias notables: la reconstrucción obtenida con el sistema propuesto (Fig. 4.10a) presenta una mayor densidad de puntos aparente y una mejor definición de objetos pequeños como los conos, en comparación con la obtenida por FAST-LIO2 (Fig. 4.10b), donde el ruido parece más pronunciado y la estructura del cono es menos discernible.

La Fig. 4.11 compara las reconstrucciones del árbol seleccionado, tal como se mencionó anteriormente, obtenidas por ambos métodos SLAM con la nube de puntos de alta densidad generada por la captura estática del LiDAR (Fig. 4.11b). Como es de esperar, la captura estática presenta la mayor densidad (cubriendo aproximadamente el 90 % del FOV del sensor) y el menor ruido aparente, al no

[9] Para version impresa el link es: https://nextcloud.unsj.edu.ar/index.php/s/xaPtRAcxCTD3d29



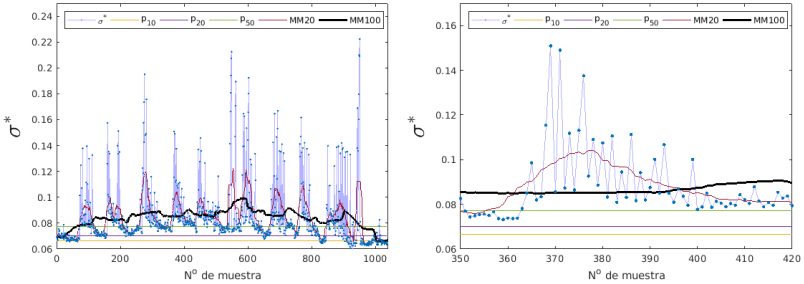
(a) SLAM propuesto. (b) LiDAR Estático. (c) Sistema FAST-LIO2.

Figura 4.11: Comparación cualitativa de la reconstrucción de un olivo específico.

sufrir de errores de estimación de pose. La comparación refuerza las observaciones anteriores: el SLAM propuesto (Fig. 4.11a) parece preservar mejor la estructura que FAST-LIO2 (Fig. 4.11c) en este caso particular, aunque ambos introducen artefactos y reducen la densidad respecto a la captura estática, ya que al estar el sensor en movimiento se cubre un mayor área con la misma cantidad de puntos. La diferencia principal radica en que el SLAM estático no acumula errores de pose, mientras que los sistemas SLAM en movimiento sí lo hacen, afectando la consistencia global de la nube de puntos fusionada. Además, factores como la incertidumbre intrínseca del sensor, la reflectancia variable de las superficies (hojas, tronco, suelo) y perturbaciones ambientales (viento) afectan a todas las capturas.

Una ventaja clave del sistema SLAM propuesto es la disponibilidad de una estimación de incertidumbre para cada pose 6DOF, cuantificada mediante la Probabilidad de Dispersión PD_{pose} (calculada según la Ec. 4.11). La Fig. 4.12 muestra la evolución de σ^* obtenida mediante la Ec. 3.26, tomando la PD calculada usando la Ec. 4.11 a lo largo del recorrido. En esta figura se observan dos comportamientos distintivos: tramos con incertidumbre relativamente baja y continua, correspondientes a instantes donde el KF-SLAM 2D pudo observar y utilizar eficazmente los landmarks (conos) para corregir la estimación; y saltos abruptos a niveles de incertidumbre más altos, que ocurren cuando no se detectan landmarks fiables y la estimación depende principalmente de la predicción basada en el modelo de movimiento (caminata aleatoria con ruido $\mathbf{Q}_r$), provocando un crecimiento rápido de la incertidumbre acumulada.

Además, en la Fig. 4.12 se muestran los umbrales que separan el $50\,\%$, $20\,\%$ y el $10\,\%$ de los datos con menor incertidumbre (denotados como p_{50}, p_{20} y p_{10} respectivamente), sumado a una media móvil con una ventana de 20 muestras (denotada como *MM20*) y otra de 100 muestras (*MM100*). Con estas medias móviles



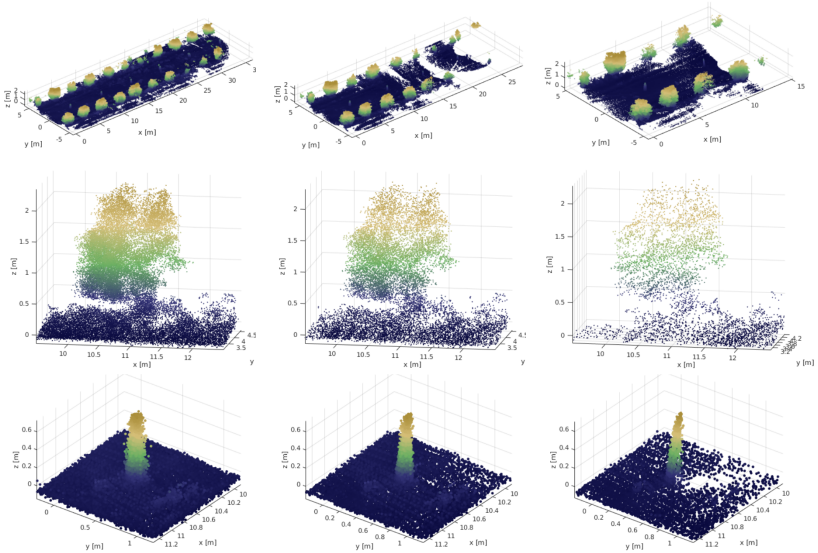
(a) Evolución temporal completa. (b) Detalle mostrando variaciones.

Figura 4.12: Evolución de la incertidumbre escalar estimada (σ^*) de la pose del sensor a lo largo del recorrido.

lo que se busca es remarcar el comportamiento del parámetro σ^* mientras se realizó el recorrido, filtrando la variabilidad de alta frecuencia (MM20) como también la de mediana frecuencia (MM100). Notar que con esta última media móvil se obtiene una forma similar a una letra "U" invertida, la cual es consistente con el movimiento realizado. En efecto, se partió de la cabecera de las filas con baja incertidumbre, se recorrieron aproximadamente 35 m hacia adentro de las filas, alcanzando la máxima incertidumbre del recorrido, y luego se volvió al punto inicial. De aquí se obtiene que mientras más lejos se esté del punto inicial, se espera que la incertidumbre tienda a aumentar.

Esta información de incertidumbre permite condicionar la reconstrucción 3D. En lugar de utilizar todas las nubes de puntos capturadas, podemos seleccionar solo aquellas correspondientes a poses estimadas con una confianza alta (baja incertidumbre). Esto es conceptualmente análogo a definir un umbral de incertidumbre máxima tolerable, $\sigma_{\max}$, en el marco teórico (aunque aquí se aplica post-procesado basado en percentiles de incertidumbre). La Fig. 4.13 ilustra este proceso: se muestran las reconstrucciones generadas utilizando solo el $50\,\%$, $20\,\%$ y $10\,\%$ de las poses con menor incertidumbre estimada (correspondientes a umbrales de $\sigma_k^* < \sigma_{\max,50}^* = 0.077183$, $\sigma_k^* < \sigma_{\max,20}^* = 0.070082$ y $\sigma_k^* < \sigma_{\max,10}^* = 0.066332$, respectivamente).

Se observa que, al reducir el porcentaje de datos (disminuyendo $\sigma_{\max}$), el ruido visual en la reconstrucción disminuye y la definición de las estructuras mejora, pero a costa de una reducción en la densidad y la cobertura espacial de la nube



(a) 50 % de los datos. (b) 20 % de los datos. (c) 10 % de los datos.

Figura 4.13: Reconstrucciones 3D condicionadas por la incertidumbre estimada de la pose. Solo se utilizan las nubes de puntos correspondientes a las poses con menor incertidumbre.

de puntos. Esta reducción no es homogénea: las zonas donde la incertidumbre fue consistentemente alta durante la captura (ej., tramos sin visibilidad de conos) quedan con muy pocos puntos o incluso vacías, como se evidencia en las Fig. 4.13c y Fig. 4.13b para las reconstrucciones con el 20 % y el 10 % de los datos (ver zonas $x \in [10, 25]$ de la Fig. 4.13b y en la reconstrucción del cono en la Fig. 4.13c). Identificar y analizar estas regiones de baja confianza o baja cobertura es precisamente donde las herramientas FI y SiREn pueden aportar valor, como se explorará a continuación.

Debido a la complejidad inherente a la visualización conjunta de nubes de puntos 3D con los diversos mapas (MI, FI y FC), se procederá a analizar estos elementos de forma separada o en vistas simplificadas antes de intentar una

integración visual completa.

Oclusiones

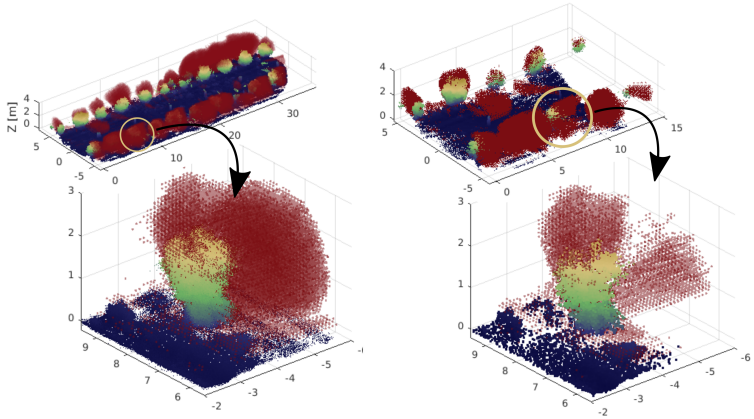
Un aspecto fundamental en la evaluación de la calidad de una reconstrucción 3D es la presencia de oclusiones, es decir, regiones del entorno que, aunque potencialmente dentro del área general de operación, no fueron observadas por el sensor debido a la interposición de otros objetos. En el contexto agrícola, esto ocurre comúnmente detrás de los troncos o en zonas densas del follaje de los árboles. Dado que la captura de datos se realizó en lazo abierto (como es práctica común actualmente) y sin planificación activa para minimizar estas oclusiones, su extensión y localización solo son evidentes tras el procesamiento offline. Determinar si un árbol o una zona de interés ha sido reconstruida completamente se vuelve, entonces, un desafío.

Estas regiones no observadas, a menudo denominadas "sombras" del sensor, representan información faltante en el mapa. Desde la perspectiva de los algoritmos de exploración basados en fronteras, cada una de estas oclusiones genera una interfaz entre el espacio conocido (los puntos reconstruidos alrededor de la oclusión) y el espacio desconocido (la región ocluida).

La Fig. 4.14 ilustra el efecto de las oclusiones en reconstrucciones generadas con diferentes niveles de confianza (filtrado por incertidumbre). Se compara la reconstrucción utilizando todas las poses estimadas ($100\,\%$ de los datos) con aquella que utiliza solo el $10\,\%$ de las poses con menor incertidumbre (correspondiente a $\sigma^* < \sigma^*_{\text{max},10}$). En la figura, las áreas detectadas como ocluidas (basadas en un análisis de visibilidad a posteriori) se resaltan en color morado.

Se observa que la extensión total de las regiones ocluidas aparenta ser mayor en la reconstrucción que utiliza el $100\,\%$ de los datos. Esto, aunque podría parecer contraintuitivo, tiene sentido: al utilizar más poses (incluyendo aquellas con mayor incertidumbre o desde puntos de vista menos informativos), se definen más claramente las superficies que *causan* la oclusión, delimitando así de forma más extensa las "sombras" detrás de ellas. Por el contrario, al usar solo las poses de alta confianza ($10\,\%$ de los datos), la reconstrucción es más dispersa y algunas de estas "sombras" pueden no quedar tan claramente definidas.

Independientemente del filtrado de incertidumbre, las fronteras internas debidas a oclusiones representan objetivos potenciales para una exploración activa destinada a completar la reconstrucción del entorno.



(a) Reconstrucción con el 100 % de datos. (b) Reconstrucción con el 10 % de datos.

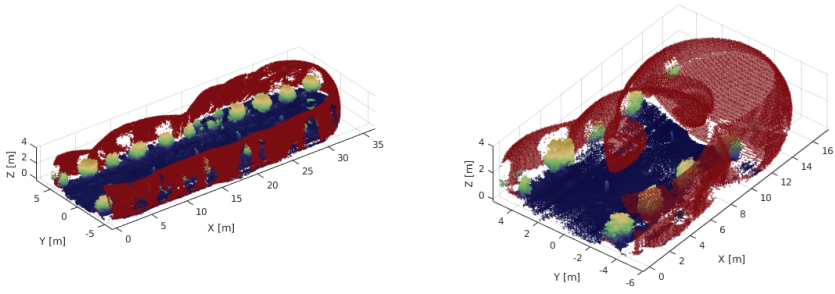
Figura 4.14: Visualización del efecto de las oclusiones en la reconstrucción 3D. Las figuras superiores comparan los resultados usando (a) el 100 % de los datos y (b) el 10 % de los datos con menor incertidumbre.

Límites de Exploración

Además de las oclusiones internas, la reconstrucción 3D está delimitada externamente por el alcance efectivo y el campo de visión (FOV) del sensor LiDAR. Aunque el Livox Mid-360 tiene un alcance nominal de 100 m, para el procesamiento y análisis en este trabajo, se consideró un rango máximo efectivo de 5 m, con el fin de concentrar el análisis en la zona cercana al sensor, simplificar la visualización y eliminar pequeños errores en las estimaciones que con las grandes distancias se amplifican.

La interfaz entre el volumen explorado (con puntos reconstruidos) y el espacio exterior no observado debido a las limitaciones del sensor (alcance o FOV) corresponde directamente al concepto de Fronteras Clásicas (FC) propuesto por Yamauchi (1997). Estas FC son un componente estándar en muchos sistemas de exploración robótica, señalando las direcciones hacia donde el robot podría moverse para adquirir nueva información sobre áreas desconocidas.

La Fig. 4.15 visualiza estas fronteras externas. Para aislar claramente los límites del sensor y evitar la confusión con las oclusiones internas discutidas previamente,



(a) 100 % de datos.

(b) Top 10 % de datos.

Figura 4.15: Visualización de los límites de exploración externos (Fronteras Clásicas), omitiendo las oclusiones internas. Se compara la extensión usando todos los datos (a) vs. datos filtrados por incertidumbre (b).

en esta visualización se han omitido los puntos clasificados como oclusión. Se comparan nuevamente las reconstrucciones obtenidas con el 100 % y el 10 % de los datos. Además, la Fig. 4.15b permite apreciar mejor la forma del FOV efectivo del sensor y cómo este define el perímetro superior e inferior de la zona explorada. Notablemente, en esta vista se aprecian vacíos en la parte superior (correspondientes a los 'conos ciegos' inherentes al patrón de escaneo del Livox Mid-360, como se ilustra en la Fig. 4.8), los cuales también constituyen límites de la información sensorial adquirida.

Observar estas FC es relevante porque representan el límite físico de la información adquirida en este experimento particular. Un sistema ASLAM completo utilizaría estas fronteras (o una combinación de ellas con otras métricas) para planificar trayectorias que expandan el área mapeada.

Análisis Direccional de Fronteras de Incertidumbre (FI)

A diferencia de las Fronteras Clásicas (FC) que únicamente marcan el límite entre lo explorado y lo desconocido, las Fronteras de Incertidumbre (FI), introducidas en la Sec. 3.5, identifican regiones del mapa donde existen altos gradientes de incertidumbre. Como se argumentó teóricamente, estas FI no solo engloban las FC y las fronteras internas por oclusión, sino que también señalan

áreas *dentro* del volumen ya explorado donde la incertidumbre es heterogénea. Esta característica es crucial para ASLAM, ya que las FI pueden guiar al agente no solo a explorar nuevas zonas, sino también a re-visitar áreas ya mapeadas para reducir su incertidumbre y facilitar cierres de lazo, mejorando así la consistencia global del mapa.

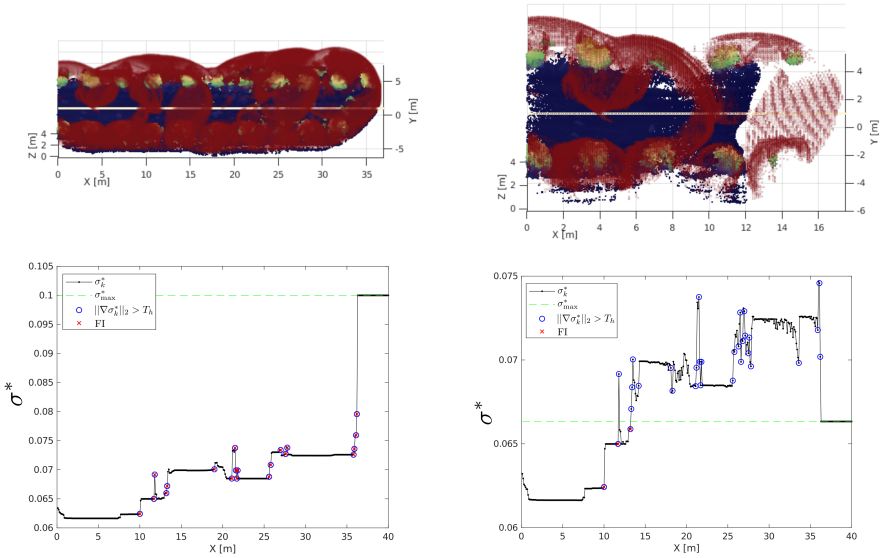
Para analizar el comportamiento de las FI en el entorno 3D real y superar los desafíos de visualización, se adopta un enfoque de análisis direccional. Se define una línea recta virtual l que atraviesa el entorno reconstruido (en este caso, $l : (x, y, z) = (d, 0, 1)$ m para $d \in [0, 35]$ m, paralela al eje x a una altura de 1 m sobre el origen $y = 0$). Luego, se extraen los valores de incertidumbre de las celdas del Mapa de Incertidumbre (MI) que intersecan esta línea l. Recordemos que el MI se construye asignando a cada celda k un valor de incertidumbre $\sigma_k^* = a/(p_k)^{1/N}$ (Ec. 3.26), donde p_k es la Probabilidad de Dispersión de la pose utilizada para generar los puntos en esa celda (Ec. 4.11), $N = 4$ y a es la constante asociada.

Las FI se detectan entonces como aquellas celdas k donde la magnitud del gradiente $||\nabla\sigma_k^*||_2$, supera un umbral T_h, y además cumplen $\sigma_k^* < \sigma_{\text{max}}^*$ (ver Sec. 3.5).

La Fig. 4.16 ilustra este análisis utilizando la reconstrucción con el $100\,\%$ y el $10\,\%$ de los datos para un $\sigma_{\text{max}}^* = 0.1$ y otro valor de σ_{max}^* (correspondiente al percentil 10, $\sigma_{\text{max},10}^* = 0.066332$), y un umbral de detección $T_h = e^{-6}$ seleccionado heurísticamente a partir de la Sec. 3.5. En la parte superior de la Fig. 4.16a se muestra la nube de puntos reconstruida, superpuesta con las celdas identificadas como FI en color morado y la línea l bajo análisis en un color claro. Se observa que las FI detectadas incluyen tanto los límites externos de exploración y la cáscara de las oclusiones internas (como se esperaba), como también aparecen en regiones intermedias del pasillo. En la parte inferior de la Fig. 4.16a se grafica el perfil de incertidumbre σ_k^* (puntos negros) a lo largo de la línea l. Se marcan los puntos donde existen saltos significativos de incertidumbre entre celdas adyacentes (círculos azules) y aquellos saltos que además son clasificados como FI (cruces rojas). En este caso, con σ_{max}^* alto, prácticamente todos los saltos significativos son FI. Se identifican FI claras cerca del final de la exploración ($t \approx 35$ m), y notablemente, en zonas intermedias ($d \in [10, 15] \cup [20, 25]$ m) donde la incertidumbre σ_k^* es localmente más alta o heterogénea, indicando potenciales oportunidades para mejorar la consistencia del mapa mediante re-observaciones o cierres de lazo.

Por otro lado, la Fig. 4.16b superior muestra la reconstrucción (con el $10\,\%$ de los datos) con las FI superpuestas. Se observan FI asociadas a los límites (incluyendo los "conos ciegos" del FOV), oclusiones, y también en regiones intermedias.

El perfil de incertidumbre de la Fig. 4.16b inferior revela una diferencia clave:



(a) Reconstrucción con el 100 % de los datos y $\sigma^*_{\max} = 0.1$.

(b) Reconstrucción con el 10 % de los datos y $\sigma^*_{\max,10}$.

Figura 4.16: Análisis direccional de Fronteras de Incertidumbre (FI) para distintos σ^*_{max}.

ahora, muchos de los saltos de incertidumbre detectados (círculos azules) NO se clasifican como FI (cruces rojas). Esto ocurre porque, aunque el gradiente supera T_h, el valor de incertidumbre σ_k^* en esas celdas supera el umbral $\sigma_{\max,10}$ seleccionado. El sistema considera que esas regiones ya son "demasiado inciertas" según el criterio actual y no las marca como objetivos prioritarios de FI. Las FI que sí se mantienen (cruces rojas) delimitan principalmente las regiones que aún están por debajo del umbral $\sigma_{\max,10}$ y presentan oportunidad de mejora, como las zonas entre $d \in [10, 15]$ m. Cabe aclarar que el perfil de incertidumbre final (parte inferior de las Figs. 4.16a y 4.16b) varía en función de las condiciones iniciales ($\sigma_{\max}$ utilizado). Esto se debe a que el salto de la actualización (ver Ec. 3.16) depende de la distancia entre $\ell_{\text{new},k}$ y $\ell_{t-1,k}$. Esta dependencia se manifiesta en las pequeñas pero observables

diferencias entre los perfiles de incertidumbre calculados con distintos valores de $\sigma_{\max}$.

Este análisis direccional en 3D, aunque simplificado, ilustra el significativo potencial informativo y práctico de las Fronteras de Incertidumbre. A diferencia de las FC, las FI no solo identifican los límites obvios de la exploración y las oclusiones. Crucialmente, mediante el ajuste de los umbrales $\sigma_{\max}$ (máxima incertidumbre tolerable) y T_h (mínimo gradiente relevante), permiten focalizar la atención del sistema ASLAM en regiones específicas donde la reducción de incertidumbre es más necesaria o factible según los objetivos de la misión.

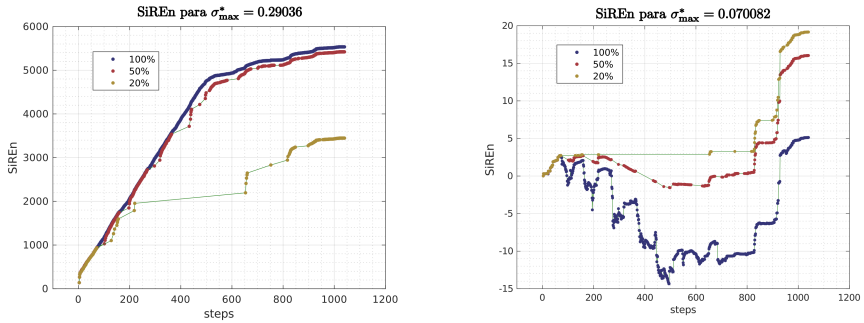
Este último resultado es el que potencialmente puede generar mayor impacto en la forma de adquirir datos de forma activa. Al procesar la información de incertidumbre (idealmente online), el sistema obtiene conocimiento explícito sobre el estado de la reconstrucción: sabe no solo *donde* faltan datos (oclusiones, límites), sino también *dónde* la información existente es poco fiable (alto σ_k^*) y, fundamentalmente, *dónde existen oportunidades claras para mejorarla* (alto $||\nabla\sigma_k^*||$ con $\sigma_k^* < \sigma_{\max}^*$).

En definitiva, las FI proporcionan la base informativa para implementar sistemas de mapeo 3D y adquisición de forma activa, capaces de tomar acciones correctivas en tiempo real para lograr una reconstrucción que cumpla con requisitos específicos de completitud e incertidumbre.

Evaluación de la Reconstrucción mediante el Índice SiREn

Complementando el análisis visual y geométrico anterior, el índice SiREn (Entropía Relativa con Signo, definido en la Sec. 3.4) proporciona una métrica cuantitativa para evaluar la calidad global de la reconstrucción 3D, considerando tanto la cobertura espacial como la incertidumbre asociada (PD_{pose}) a cada punto o celda del mapa. A medida que se adquieren los datos y se actualiza el Mapa de Incertidumbre (MI) (o directamente se asocia una PD_{pose} a cada pose), es posible estimar el valor acumulado de SiREn, utilizando la Ec. 3.20 o su aproximación basada en PD mediante la Ec. 3.22. Idealmente, en un sistema ASLAM online, este índice permitiría al agente cuantificar en tiempo real la "cantidad de información útil" que está adquiriendo, habilitando una toma de decisiones orientada a maximizar dicho índice.

La Fig. 4.17 muestra la evolución del índice SiREn acumulado a lo largo de la trayectoria en función del tiempo con el $100\,\%$, $50\,\%$ y el $20\,\%$ de los datos, calculado con un valor de $\sigma_{\max}$ correspondiente al umbral utilizado para filtrar el $100\,\%$ de los datos (es decir, el $\sigma_{\max}$ más permisivo que incluye todas las poses, o el



(a) $\sigma_{\max}$ equivalente al 100 % de los datos.

(b) $\sigma_{\max}$ equivalente al 20 % de los datos.

Figura 4.17: Evolución del índice SiREn acumulado a lo largo de la trayectoria en función del tiempo utilizando distinta cantidad de datos y $\sigma_{\max}$.

$\sigma_{\max}$ asociado al percentil 100). Para el caso de la reconstrucción con el 100 % de los datos se observa un comportamiento característico:

- Fase Inicial (0-500 pasos aprox.): Tras un salto inicial[10], SiREn crece de forma aproximadamente lineal y con pendiente pronunciada. Esto se correlaciona con la primera mitad del recorrido (ida), donde el sensor avanza desde la cabecera hacia el interior de la fila de árboles, adquiriendo información predominantemente nueva (exploración) con incertidumbre relativamente controlada (observación continua de conos).

- Fase Posterior (500 pasos en adelante): La pendiente de crecimiento de SiREn disminuye notablemente. Esto corresponde al trayecto de regreso, donde una parte significativa del entorno ya ha sido observada. Aunque se mapean zonas nuevas (oclusiones, bordes del FOV) y se refinan otras (potenciales cierres de lazo o re-observaciones), la ganancia de información *nueva y de alta calidad* por paso es menor.

- Comportamiento Local: Se aprecian "escalones" (aumentos abruptos) seguidos de "llanuras" (crecimiento lento o nulo). Los escalones corresponden a la incorporación de información significativa (nuevas áreas vistas, observaciones de zonas con menor incertidumbre que antes), mientras

[10] Antes de empezar el mapa está vacío, cuando se incorpora la primera medición se genera este salto.

que las llanuras indican fases donde las nuevas mediciones no reducen sustancialmente la incertidumbre global por debajo de σ^*_{max} o mapean áreas ya bien conocidas.

En la primera fase, la ganancia promedio es mayor a 9 unidades de SiREn por paso (calculado como $\approx 4800/500$), mientras que en la segunda fase desciende a ≈ 1.6 unidades por paso ($\approx (5600-4800)/500$). Esta reducción drástica sugiere que la trayectoria de ida y vuelta, sin planificación activa, podría mejorarse en términos de maximización de SiREn por unidad de tiempo, distancia recorrida o muestras. Un sistema ASLAM guiado por SiREn o FI probablemente elegiría acciones diferentes en la segunda mitad del recorrido.

Para investigar la sensibilidad de SiREn a la calidad de los datos, se recalculó el índice utilizando solo el $50\,\%$ de las poses con menor incertidumbre estimada, pero manteniendo el mismo valor alto de σ_{max} que en el caso anterior. Sorprendentemente, el valor final de SiREn es muy similar al obtenido con el $100\,\%$ de los datos. Esto sugiere que, para este valor de σ_{max}, la contribución principal al índice proviene de las regiones mapeadas con baja incertidumbre. Incluir las poses con alta incertidumbre (el $50\,\%$ restante) no aporta significativamente al valor final de SiREn (ya que las celdas mapeadas con $\sigma^*_k > \sigma^*_{max}$ aportan negativamente o poco, y las de $\sigma^*_k \approx \sigma^*_{max}$ aportan aproximadamente cero). Esto demuestra la capacidad de SiREn para valorar la calidad sobre la cantidad de los datos, alineándose con el objetivo de obtener reconstrucciones precisas. A modo cualitativo, compare la Fig. 4.13a correspondiente a la reconstrucción con el $50\,\%$ de los datos con la Fig. 4.10a correspondiente al $100\,\%$ de los datos. Si bien se observan diferencias, no se percibe que una solo tiene la mitad de los datos.

Continuando, se calcula el índice para el $20\,\%$ de los datos con menor incertidumbre, obteniendo aproximadamente la mitad de SiREn comparado con el $100\,\%$ de los datos. Aquí se observa que, aún priorizando la calidad de los datos, la cobertura también juega un rol significativo. Cualitativamente, se puede comparar la cobertura para las 3 reconstrucciones en la Fig. 4.13.

Finalmente, se analiza el comportamiento de SiREn cuando el umbral de incertidumbre máxima tolerable, σ_{max}, se ajusta para ser más exigente. Se utilizó un valor de σ_{max} equivalente al umbral del $20\,\%$ de los datos (es decir $\sigma^*_{\text{max},20}$) y se calculó SiREn utilizando nuevamente el $100\,\%$ $50\,\%$ y $20\,\%$ de los datos. El resultado se muestra en la Fig. 4.17b. Como era de esperar, el comportamiento es drásticamente diferente. El índice SiREn se mantiene en valores predominantemente negativos durante la mayor parte de la trayectoria en los tres casos. Esto se debe a que la gran mayoría ($80\,\%$) de las poses tienen una

incertidumbre asociada superior a este $\sigma_{\max}$ más estricto. Según la definición de SiREn (Ec. 3.20), las celdas mapeadas con una incertidumbre $\sigma_k^* > \sigma_{max}^*$ contribuyen negativamente al índice. Solo en la segunda mitad del recorrido se lograron re-observaciones con menor incertidumbre por debajo de σ_{max}, por lo que el índice comienza a crecer ligeramente.

Este resultado ilustra cómo $\sigma_{\max}$ actúa efectivamente como un parámetro que define el nivel de calidad deseado para la reconstrucción. Además, se observa que para valores grandes de $\sigma_{\max}$ el índice tiene una mayor sensibilidad a los cambios. Por otro lado, se observa que mientras más pequeño sea $\sigma_{\max}$, más se pondera la calidad sobre la cobertura. Esto se evidencia comparando la Fig. 4.17a con la Fig. 4.17b, ya que para el primer caso, ordenando de mayor a menor los valores de SiREn se obtiene primero la reconstrucción con el $100\,\%$ de los datos, seguido muy de cerca por la del $50\,\%$ y por último la del $20\,\%$, mientras que para el segundo caso el escenario es totalmente inverso: primero el conjunto del $20\,\%$, seguido por el del $50\,\%$ y por último el del $100\,\%$ de los datos.

Este análisis demuestra la aplicabilidad y utilidad del índice SiREn para evaluar cuantitativamente reconstrucciones 3D en entornos complejos. Permite comparar objetivamente resultados con diferentes conjuntos de datos (ej., filtrados por incertidumbre) o bajo distintos criterios de calidad (diferentes $\sigma_{\max}$). Su sensibilidad tanto a la cobertura como a la incertidumbre lo convierte en un candidato prometedor para guiar sistemas ASLAM. Se conjetura que la integración de SiREn en bucles de planificación activa, posiblemente combinada con modelos predictivos de la ganancia de información esperada, permitiría diseñar trayectorias óptimas para maximizar la calidad de la reconstrucción final.

Capítulo 5

Conclusiones y Trabajos Futuros

5.1 Síntesis del Trabajo 117
5.2 Hallazgos Principales y Validación 118
5.3 Contribuciones y Significado 119
5.4 Limitaciones del Trabajo 120
5.5 Trabajos Futuros 121
5.6 Observaciones Finales 121

5.1. Síntesis del Trabajo

El problema de SLAM Activo consiste en la construcción autónoma de mapas mediante la planificación secuencial y deliberada de acciones de exploración y mapeo. Este proceso representa un desafío fundamental en el campo de la robótica móvil, ya que busca optimizar la adquisición de información sobre el entorno.

La gestión explícita y eficiente de la incertidumbre, tanto del estado del agente como de la representación del entorno, es crucial para tomar decisiones informadas que conduzcan a reconstrucciones precisas y completas con recursos limitados. Sin embargo, los enfoques tradicionales basados en mapas de ocupación y Fronteras Clásicas (FC) presentan limitaciones en cuanto a la representación de la incertidumbre, la guía para la re-exploración y la evaluación cuantitativa de la calidad del mapa.

Esta tesis doctoral ha abordado estos desafíos mediante el desarrollo y la validación de un marco novedoso para el modelado, análisis y explotación de

la incertidumbre en sistemas ASLAM. La piedra angular de este marco son las Probabilidades de Dispersión (PD), una métrica probabilística ($PD \in [0, 1]$) propuesta para cuantificar la incertidumbre local de una estimación (ej., la pose del robot, el estado de una celda) relativa a una tolerancia definida por el usuario. Las PD ofrecen una representación interpretable y flexible, aplicable a distintas fuentes de incertidumbre y tipos de sensores.

A partir de las PD, se derivaron tres herramientas interconectadas:

- El Mapa de Incertidumbre (MI): Una representación espacial donde cada celda almacena un valor de incertidumbre σ_k^* derivado de la PD_k, proporcionando una visión global de la fiabilidad de la reconstrucción.

- Las Fronteras de Incertidumbre (FI): Detectadas como regiones con altos gradientes en el MI ($||\nabla\sigma_k^*||_2 > T_h$) y cuya incertidumbre está por debajo de un umbral máximo tolerable ($\sigma_k^* < \sigma_{\text{max}}^*$). Las FI generalizan las FC, identificando no solo límites de exploración sino también áreas internas con incertidumbre heterogénea, aptas para re-exploración y cierre de lazos.

- El Índice SiREn (Entropía Relativa con Signo): Una métrica global basada en la divergencia de Kullback-Leibler que cuantifica la calidad del mapa, ponderando positivamente las áreas mapeadas con baja incertidumbre (respecto a σ_{max}) y penalizando aquellas con alta incertidumbre.

5.2. Hallazgos Principales y Validación

La viabilidad y utilidad de este marco se investigaron mediante simulaciones en 2D y análisis sobre datos reales 3D de un entorno agrícola complejo:

- Validación del marco teórico mediante simulaciones 2D: Los experimentos comparativos demostraron que los planificadores basados en FI (utilizando RRT* consciente de la incertidumbre) podían generar mapas de mayor calidad (evaluada por el índice SiREn y con menor variabilidad) en comparación con enfoques basados en FC (RRT/RRT* greedy). Las FI guiaron eficazmente la re-exploración y condujeron a un mapa con una incertidumbre más homogénea. Se validó que el índice SiREn es sensible tanto a la cobertura como a la magnitud de la incertidumbre, correlacionándose con la precisión de los landmarks y permitiendo diferenciar cuantitativamente el rendimiento de distintas estrategias y configuraciones.

- Aplicabilidad y Análisis en 3D Real: Se demostró la aplicación del marco para analizar datos 3D de un olivar. Se desarrolló un sistema SLAM adaptado que proporcionó estimaciones de pose 6DOF con incertidumbre (PD). El análisis del MI resultante y la detección de FI (mediante un análisis direccional) ilustraron cómo estas herramientas identifican no solo oclusiones y límites de exploración (análogos a FC), sino también regiones internas con alta variabilidad de incertidumbre, indicando oportunidades concretas para mejorar la reconstrucción mediante acciones dirigidas. Finalmente, el análisis cuantitativo mediante el índice SiREn validó su eficacia como métrica de evaluación global en estos escenarios realistas. Su evolución temporal a lo largo de la trayectoria reflejó coherentemente las fases de exploración y re-exploración observadas. Asimismo, su demostrada sensibilidad al umbral $\sigma_{\max}$ y a la calidad intrínseca de los datos (evidenciada al comparar resultados con y sin filtrado por incertidumbre) confirmó su capacidad para cuantificar objetivamente la calidad de la reconstrucción 3D, valorando la precisión sobre la mera cantidad de información.

5.3. Contribuciones y Significado

Las principales contribuciones de esta tesis son:

1. Un nuevo formalismo (PD) para representar la incertidumbre local de forma interpretable, normalizada y flexible.

2. El desarrollo de las Fronteras de Incertidumbre (FI) como una herramienta de guía para ASLAM más informativa que las FC, capaz de dirigir el cierre de lazos, y parametrizable mediante T_h y σ_{max}.

3. La propuesta del índice SiREn como métrica cuantitativa para evaluar la calidad global de los mapas generados por SLAM/ASLAM, considerando conjuntamente cobertura e incertidumbre respecto a un nivel de calidad deseado (σ_{max}).

4. La demostración de la aplicabilidad y utilidad de este marco (PD, MI, FI, SiREn) mediante simulaciones y análisis sobre datos 3D reales, mostrando su potencial para mejorar la toma de decisiones en ASLAM.

El significado de este trabajo radica en proporcionar un conjunto de herramientas coherentes y fundamentadas para abordar la gestión de la incertidumbre en ASLAM. Aunque la validación experimental se centró en

simulaciones indoor y entornos agrícolas (predominantemente abiertos), el marco propuesto no se limita a estos escenarios. Su aplicabilidad se extiende a otros dominios exigentes donde la cuantificación y gestión de la incertidumbre son críticas. Por ejemplo, en operaciones de Búsqueda y Rescate (SAR, por sus siglas en inglés), tanto en entornos terrestres como subacuáticos, una estimación fiable de la incertidumbre asociada a la detección de objetivos y a la representación del mapa es crucial; esto permite minimizar el riesgo de falsos negativos (objetivos no detectados) y optimizar las estrategias de búsqueda para maximizar la probabilidad de éxito en condiciones de tiempo crítico. Asimismo, la flexibilidad del formalismo PD abre vías para su integración en sistemas de aprendizaje activo (Active Learning), donde la incertidumbre de la inferencia (estimada mediante técnicas como Monte Carlo Dropout o Deep Ensambles) podría representarse y utilizarse para guiar la adquisición de nuevos datos de entrenamiento. En resumen, este marco ofrece potencial para mejorar la toma de decisiones basada en incertidumbre en una amplia gama de aplicaciones robóticas, de aprendizaje automático y otras.

5.4. Limitaciones del Trabajo

Es importante reconocer las limitaciones de la investigación realizada:

- Estimación de Incertidumbre Base: La calidad de las FI y SiREn depende directamente de la calidad y consistencia de la incertidumbre proporcionada por el sistema SLAM subyacente. En los experimentos 3D, se utilizaron heurísticas (σ_π, σ_{row}) y aproximaciones para caracterizar la incertidumbre 6DOF.

- Implementación ASLAM 3D: No se implementó un sistema ASLAM 3D completo con planificación online basada en FI/SiREn. Los experimentos 3D se centraron en el análisis offline de la aplicabilidad del marco.

- Landmarks Artificiales: La validación 3D dependió de landmarks artificiales (conos) para robustecer la estimación longitudinal, limitando la generalización directa a entornos sin esta infraestructura.

- Profundización Teórica y Parametrización: Si bien se establecieron las bases conceptuales, el marco teórico requiere una mayor profundización en aspectos específicos. Esto incluye el desarrollo de métodos más sistemáticos para la determinación óptima del umbral T_h para las FI, así como la investigación de mecanismos de actualización de incertidumbre más

avanzados (como el propuesto basado en Filtro de Kalman) que los utilizados en la validación experimental actual (basado en log-odds modificado).

- Comparativa Limitada: No se realizaron comparaciones cuantitativas exhaustivas con otros enfoques ASLAM de última generación en los experimentos 3D.

5.5. Trabajos Futuros

Este trabajo abre diversas vías para investigaciones futuras:

1. Implementación ASLAM 3D Completa: Desarrollar e implementar un sistema ASLAM 3D online que utilice FI como objetivos de exploración/re-exploración y/o SiREn (o su ganancia esperada) en la función de utilidad del planificador de trayectorias.

2. Actualización con KF y Mapas Dinámicos: Implementar la actualización del MI basada en Filtro de Kalman (propuesta en la Sección 3.3) para manejar entornos dinámicos o variables temporales (ej., crecimiento de cultivos).

3. Planificación Basada en SiREn: Investigar algoritmos de planificación que busquen maximizar explícitamente la ganancia esperada de SiREn.

4. Adaptación de Parámetros: Desarrollar métodos para adaptar automáticamente los umbrales (T_h, σ_{max}) y los intervalos (s_i) según las características del entorno o los objetivos de la misión.

5. Validación Extensiva: Evaluar el sistema ASLAM completo en una mayor variedad de entornos (diferentes cultivos, escenarios indoor/outdoor) y compararlo cuantitativamente con otros enfoques ASLAM.

6. Explorar métricas para evaluar mapas, como la raíz cuadrada de la divergencia Jensen-Shannon Endres and Schindelin (2003).

5.6. Observaciones Finales

Esta tesis ha presentado un marco coherente y prometedor para la gestión de la incertidumbre en SLAM Activo. Las herramientas desarrolladas: Probabilidades de Dispersión, Mapa de Incertidumbre, Fronteras de Incertidumbre e Índice SiREn; ofrecen nuevas capacidades para analizar, evaluar y guiar el proceso de construcción

autónoma de mapas. Aunque restan desafíos para su implementación completa, validación exhaustiva y profundización teórica (particularmente en la selección de parámetros y mecanismos de actualización), los resultados sugieren un avance significativo. Se espera que este trabajo, concebido simultáneamente como punto de partida y como vector que oriente el rumbo de la investigación, dirija los esfuerzos de la comunidad hacia el desarrollo de sistemas de exploración y mapeo con mayores niveles de autonomía, capaces de optimizar la toma de decisiones durante la adquisición de datos y de garantizar, de manera objetiva, la calidad de las reconstrucciones.

Apéndice A

Apéndice

A.1 Apendice simulación 2D 123
A.1.1 Sistema de planificación de trayectorias 123
A.1.2 Sistema de SLAM 2D 125
A.2 Demostraciones y ejemplos numéricos 126
A.2.1 Desigualdad de Chebyshev 126
A.2.2 Prueba caso Gaussiano unidimensional 127
A.2.3 Ejemplo numérico para determinar β 127
A.2.4 Pruebas del Mapa de incertidumbre 129
A.2.5 Aproximación de D_{KL} 129

A.1. Apendice simulación 2D

A.1.1. Sistema de planificación de trayectorias

Como se indicó en la Sec. 3.5, una FI se considera eliminada cuando se cumple la condición $||\nabla\sigma_k^*||_2 < T_h$, lo cual implica que la incertidumbre máxima asociada a las celdas pertenecientes a la frontera ha sido reducida por debajo del umbral establecido. La reducción de la incertidumbre de una celda únicamente puede producirse si la medida de incertidumbre propagada por el agente es inferior a la incertidumbre presente en dicha celda. Por consiguiente, al planificar una trayectoria con el objetivo de eliminar una FI, es fundamental incorporar la incertidumbre estimada del agente en el proceso de decisión. Esta incertidumbre

es proporcionada por el sistema SLAM y depende de diversos factores, tales como la precisión de los sensores, las estimaciones odométricas y otras fuentes de error inherentes al sistema de localización.

Por esta razón, en este trabajo se propone una versión modificada del algoritmo RRT* Karaman et al. (2011) con el objetivo de planificar trayectorias que permitan la eliminación de una FI. El algoritmo RRT* extiende un árbol RRT tradicional, comenzando con su crecimiento estándar. No obstante, antes de establecer una conexión con el nodo más cercano, realiza una búsqueda dentro de un radio definido y selecciona la conexión que minimiza una función de costo. La función de costo C propuesta en este trabajo es:

$$C = d + d_{\text{odo}} + \frac{2\tilde{\sigma}_l}{c\tilde{Q}}, \tag{A.1}$$

donde d representa la distancia entre los nodos, d_{odo} es la distancia recorrida utilizando únicamente odometría (sin referencias externas), c corresponde al tamaño de celda del mapa, y $\tilde{Q}$ es un factor de conversión que relaciona la incertidumbre de los landmarks con una medida de distancia. Este valor puede interpretarse como el ruido del proceso en una implementación basada en filtros de Kalman, es decir, $|\mathbf{Q}|^{1/2N} = \tilde{Q}$. En este trabajo se ha fijado $\tilde{Q} = 0.01$. Cada landmark tiene asociada una desviación estándar media geométrica, $\tilde{\sigma}_l$, la cual se calcula como $\tilde{\sigma}_l = |\mathbf{\Sigma}_l|^{1/2N}$, donde Σ_l es la matriz de varianza-covarianza correspondiente al $l-$ésimo landmark.

Cuando el algoritmo inicia, tanto d como $\tilde{\sigma}_l$ se inicializan en cero. El valor inicial de d_{odo} corresponde a la distancia recorrida por el agente utilizando únicamente odometría (sin referencias externas) hasta alcanzar el nivel actual de incertidumbre. Posteriormente, cada vez que se agrega un nodo al árbol, d_{odo} se actualiza para reflejar la distancia recorrida sin referencias desde el nodo padre hasta el nuevo nodo. De manera simultánea, d se actualiza con la distancia entre dichos nodos. Si se observa un landmark, d_{odo} se reinicia a cero, lo cual indica que la estimación de la pose del agente pasa a estar influenciada principalmente por la incertidumbre en la medición. Al mismo tiempo, $\tilde{\sigma}_l$ se actualiza para igualar el valor de $\tilde{\sigma}_l$ del landmark observado. En caso de que se observen múltiples landmarks, se selecciona el último valor observado de $\tilde{\sigma}_l$.

A.1.2. Sistema de SLAM 2D

Dado que el objetivo principal de este trabajo es analizar el comportamiento del MI, el índice SiREn y las FI, se implementó un sistema SLAM óptimo. Es importante remarcar que el enfoque de este trabajo no se centra en el sistema SLAM en sí, sino en garantizar que las covarianzas tanto de los landmarks como del agente sean estimadas de manera consistente. Es bien sabido que todas las variantes no lineales del filtro de Kalman (KF-SLAM) tienden a subestimar las covarianzas Castellanos et al. (2004); Sola (2010). Este problema resulta crítico, dado que las decisiones dependen directamente de la incertidumbre representada en la matriz de varianza-covarianza del mapa. Por esta razón, se implementó un sistema lineal de KF-SLAM para la estimación de la posición, considerando que la orientación es conocida mediante sensores como magnetómetros, giróscopos, entre otros.

El robot utilizado en la simulación es un robot diferencial equipado con un escáner láser de $360°$ con un alcance de $5\,\text{m}$ y una resolución angular de $0.5°$. Sin embargo, se implementaron sistemas de control en un nivel inferior que permiten aproximar el comportamiento del robot no holonómico a uno holonómico, a fin de facilitar su integración con los módulos de planificación y SLAM. El control de bajo nivel fue diseñado para permitir un seguimiento preciso de trayectorias en un robot diferencial no holonómico. Para ello, se implementó un controlador no lineal basado en la retroalimentación de la posición estimada por el sistema SLAM. Este controlador se encarga de guiar al robot a lo largo de la trayectoria planificada, definida en términos de una secuencia de *waypoints* cercanos. La orientación del robot se considera conocida y libre de ruido, ya que es provista directamente por el simulador. Una característica particular de esta implementación es que las acciones de control entregadas al sistema SLAM no son velocidades angulares de las ruedas, sino posiciones relativas (x, y). Esta transformación permite abstraer el comportamiento no holonómico del robot y presentarlo al sistema SLAM como si fuera un robot holonómico. De este modo, se simplifica la integración con el sistema de planificación y posicionamiento. El algoritmo de seguimiento de trayectorias utilizado fue tomado de una tesis doctoral Secchi (2008); Secchi et al. (1999), y ha demostrado ser eficaz para mantener el error de seguimiento dentro de límites aceptables bajo condiciones simuladas.

Las ecuaciones discretas del sistema KF-SLAM son:

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{U}_k + \mathbf{w}_k, \tag{A.2}$$

$$\mathbf{z}_{k+1} = \mathbf{H}_k \mathbf{x}_{k+1} + \mathbf{v}_k, \tag{A.3}$$

donde $\mathbf{A}$ es una matriz identidad, $\mathbf{B} = [\mathbf{I}_{2\times 2},\ \mathbf{0}_{(n-2)\times 2}]$, $\mathbf{C} = [\mathbf{I}_{2\times 2},\ -\mathbf{I}_{2\times 2}]$ es el bloque de matriz de observación correspondiente a cada landmark, $\mathbf{U}$ representa la entrada de control, y $\mathbf{w}$ y $\mathbf{v}$ son los ruidos de proceso y de medición, respectivamente. Ambos ruidos se modelan como distribuciones gaussianas con media cero y matrices de covarianza $\mathbf{Q} = 0.01\mathbf{I}$ y $\mathbf{R} = 0.01\mathbf{I}$, respectivamente. La matriz de covarianza del estado se denota por $\mathbf{P}$, y su valor inicial es $\mathbf{P}_0 = 0.01\mathbf{I}$. No se asume conocimiento previo de ningún landmark al inicio del proceso.

Finalmente, en la implementación del sistema SLAM no se consideran errores de asociación entre observaciones y landmarks. Esta suposición garantiza la convergencia del filtro de Kalman. Para más detalles sobre la implementación y la convergencia del sistema SLAM, véase Dissanayake et al. (2001).

A.2. Demostraciones y ejemplos numéricos

A.2.1. Desigualdad de Chebyshev

La desigualdad de Chebyshev establece que para cualquier variable aleatoria X con media μ, y varianza σ^2, se cumple que:

$$P(|X-\mu| \geq k\sigma) \leq \begin{cases} \frac{1}{k^2} & \text{si } k > 1 \\ 1 & \text{si } k \leq 1 \end{cases} \tag{A.4}$$

donde k es un número real positivo. Cuando su moda es igual a su media, esta desigualdad se puede mejorar Van Parys et al. (2016):

$$P(|X-\mu| \geq k\sigma) \leq \begin{cases} \frac{4}{9k^2} & \text{si } k > \frac{2}{\sqrt{3}} \\ 1 - \frac{k}{\sqrt{3}} & \text{si } k \leq \frac{2}{\sqrt{3}} \end{cases} \tag{A.5}$$

luego:

$$P(|X-\mu| < k\sigma) = 1 - P(|X-\mu| \geq k\sigma) \geq \begin{cases} 1 - \frac{4}{9k^2} & \text{si } k > \frac{2}{\sqrt{3}} \\ \frac{k}{\sqrt{3}} & \text{si } k \leq \frac{2}{\sqrt{3}} \end{cases}$$

luego si se define $k\sigma = \frac{s}{2}$, la anterior desigualdad resulta:

$$P(|X-\mu| < s/2) = 1 - P(|X-\mu| \geq s/2) \geq \begin{cases} 1 - \left(\frac{4\sigma}{3s}\right)^2 & \text{si } s > \frac{4\sigma}{\sqrt{3}} \\ \frac{s}{2\sqrt{3}\sigma} & \text{si } s \leq \frac{4\sigma}{\sqrt{3}} \end{cases}$$

A.2.2. Prueba caso Gaussiano unidimensional

Dada una variable aleatoria $X \sim \mathcal{N}(\mu, \sigma^2)$, se desea conocer la PD de X en función de su media y varianza. Para ello se define la PD de X como:

$$P\left(|x-\mu| < \frac{s}{2}\right) = P\left(-\frac{s}{2} < x-\mu < \frac{s}{2}\right) = P\left(x-\mu < \frac{s}{2}\right) - P\left(x-\mu < -\frac{s}{2}\right) \tag{A.6}$$

$$= P\left(z < \frac{s}{2\sigma}\right) - P\left(z < -\frac{s}{2\sigma}\right) \quad \text{tomando } z = \frac{x-\mu}{\sigma} \tag{A.7}$$

$$= \Phi\left(\frac{s}{2\sigma}\right) - \Phi\left(-\frac{s}{2\sigma}\right) \tag{A.8}$$

$$= 2\left(\Phi\left(\frac{s}{2\sigma}\right) - \Phi(0)\right) \quad \text{por ser simétrica en } z = 0 \tag{A.9}$$

$$= 2\left(\Phi\left(\frac{s}{2\sigma}\right) - \frac{1}{2}\right) \tag{A.10}$$

$$= 2\Phi\left(\frac{s}{2\sigma}\right) - 1 \tag{A.11}$$

A modo ilustrativo en la Fig. A.1 se muestra la PD de una variable aleatoria gaussiana en función de su media y varianza.

A.2.3. Ejemplo numérico para determinar β

Supóngase que la incertidumbre en la *pose* de un robot diferencial se modela mediante una distribución gaussiana con matriz de varianza-covarianza

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_\theta^2 \end{bmatrix}, \tag{A.6}$$

donde σ_x, σ_y y σ_θ son las desviaciones típicas de la posición y la orientación del robot, respectivamente. Para los fines de diseño se fija la incertidumbre máxima en $\sigma_{x\max} = \sigma_{y_{\max}} = 2\,\text{m}$ y $\sigma_{\theta\max} = 0.02\,\text{rad}$. Por otra parte, el hiper-rectángulo de trabajo se define como $A = s_x \times s_y \times s_\theta = 0.1\,\text{m} \times 0.1\,\text{m} \times 0.002\,\text{rad}$.

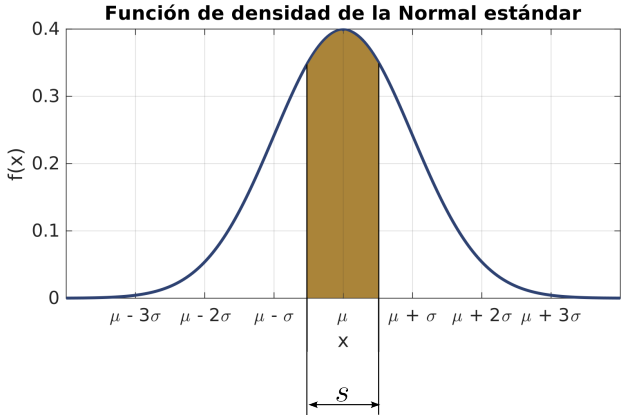


Figura A.1: El área pintada es la Probabilidad de Dispersión de una variable aleatoria gaussiana en función de su media y varianza.

Con estos parámetros, y aplicando las ecuaciones (3.7) y (3.3), se obtiene

$$\beta = P(|x| < s_x/2) \cdot P(|y| < s_y/2) \cdot P(|\theta| < s_\theta/2), \text{ por ser independientes} \tag{A.12}$$

$$= \left(2\Phi\left(\frac{s_x}{2\sigma_{x\max}}\right) - 1\right) \cdot \left(2\Phi\left(\frac{s_y}{2\sigma_{y\max}}\right) - 1\right) \cdot \left(2\Phi\left(\frac{s_\theta}{2\sigma_{\theta\max}}\right) - 1\right) \tag{A.13}$$

$$= 19.945 \times 10^{-3} \cdot 19.945 \times 10^{-3} \cdot 39.878 \times 10^{-3} \tag{A.14}$$

$$= 1.5863 \times 10^{-5} \tag{A.15}$$

$$\ell_\beta = -11.051, \qquad a = \frac{(s_x s_y s_\theta)^{1/3}}{2\sqrt{3}} = 7.8358 \times 10^{-3}. \tag{A.16}$$

Finalmente,

$$\sigma^*_\beta = \frac{a}{\beta^{1/3}} = 0.31186 < \left|\mathbf{\Sigma}_{\max}\right|^{1/6} = \tilde{\sigma}_{\max} = 0.43089. \tag{A.17}$$

Estos valores se emplean en los algoritmos de actualización de celdas para calcular la métrica *SiREn* y determinar el factor de actualización *FI* cuando la incertidumbre de la medición es despreciable.

A.2.4. Pruebas del Mapa de incertidumbre

Dada una medición con su respectiva matriz de varianza-covarianza $\mathbf{\Sigma} \in \mathcal{M}^{N\times N}$, entonces:

$$|\mathbf{\Sigma}| = \prod_{i=1}^{N} \lambda_i$$

con λ_i los autovalores de $\mathbf{\Sigma}$, estos autovalores son las varianzas σ_i^2 del vector de variables aleatorias del sistema. Entonces:

$$|\mathbf{\Sigma}| = \prod_{i=1}^{N} \sigma_i^2 \Longrightarrow |\mathbf{\Sigma}|^{\frac{1}{2N}} = \left(\prod_{i=1}^{N} \sigma_i\right)^{\frac{1}{N}} \triangleq \tilde{\sigma}$$

siendo $\tilde{\sigma}$ la media geométrica de los desvíos estándar de las mediciones el cual cuantifica la incertidumbre media de todas sus componentes. Luego $\tilde{\sigma}_k$ es la asociada a la celda k-th,

Por otro lado, considerando un vector de variables aleatorias independientes X_k con media cero, correspondiente a la celda k-th, entonces p_k resulta:

$$p_k = P(X \in A) = P\left(|x_1| < \frac{s_1}{2}\right) P\left(|x_2| < \frac{s_2}{2}\right) \dots P\left(|x_N| < \frac{s_N}{2}\right)$$

$$\geq \frac{s_1}{2\sqrt{3}\sigma_1} \frac{s_2}{2\sqrt{3}\sigma_2} \dots \frac{s_N}{2\sqrt{3}\sigma_N} = \frac{\prod_i^N s_i}{(2\sqrt{3})^N \tilde{\sigma}_k^N} \quad \text{si } 0 < s_i \leq \frac{4\sigma_i}{\sqrt{3}} \quad \text{, de la desigualdad 3.4}$$

definiendo $a^N = \frac{\prod_i^N s_i}{(2\sqrt{3})^N}$ se tiene que:

$$p_k \geq \left(\frac{a}{\tilde{\sigma}_k}\right)^N \tag{A.18}$$

luego el MI se obtiene según:

$$\sigma_k^* = \frac{a}{p_k^{1/N}} < \tilde{\sigma}_k \tag{A.19}$$

A.2.5. Aproximación de D_{KL}

En esta sección se describe un procedimiento para obtener una cota inferior para $D_s(M)$ a partir de un mapa de PDs.

La divergencia de Kullback-Leibler (KLD) entre dos distribuciones f y g gaussianas N dimensionales con matrices de varianza-covarianza $|\mathbf{\Sigma}_f|$ y $|\mathbf{\Sigma}_g|$ y vector de medias μ_f y μ_g, tiene una forma cerrada Hershey and Olsen (2007), dada por:

$$D_{KL}(f||g) = \frac{1}{2}\left(\ln\frac{|\mathbf{\Sigma}_g|}{|\mathbf{\Sigma}_f|} + \text{tr}\left(\mathbf{\Sigma}_g^{-1}\mathbf{\Sigma}_f\right) - N + (\mu_f - \mu_g)^T\mathbf{\Sigma}_g^{-1}(\mu_f - \mu_g)\right) \tag{A.20}$$

Luego tomando que $\mathbf{\Sigma}_f = I\tilde{\sigma}_k^2$, $\mathbf{\Sigma}_g = I\tilde{\sigma}_{\text{max}}^2$ y teniendo en cuenta los resultados del Ap. A.2.4, resulta que:

$$\begin{aligned}
\text{tr}\left(\mathbf{\Sigma}_g^{-1}\mathbf{\Sigma}_f\right) &= N\frac{\sigma_k^2}{\tilde{\sigma}_{\text{max}}^2} \geq N\left(\frac{p_k}{\beta}\right)^{-2/N} \\
|\mathbf{\Sigma}_g| &= \tilde{\sigma}_{\text{max}}^{2N} \geq \frac{a^{2N}}{\beta^2} \\
|\mathbf{\Sigma}_f| &= \tilde{\sigma}_k^{2N} \geq \frac{a^{2N}}{p_k^2}
\end{aligned}$$

además asumiendo que $\mu_f = \mu_g$, y reemplazando en Ec. A.20 resulta:

$$D_{KL}(f||g) \geq \frac{1}{2}\left(\ln\left(\frac{p_k}{\beta}\right)^2 - N + N\left(\frac{p_k}{\beta}\right)^{-2/N}\right) \tag{A.21}$$

$$\geq \ln\frac{p_k}{\beta} - \frac{N}{2} + \frac{N}{2}\left(\frac{\beta}{p_k}\right)^{2/N} \tag{A.22}$$

Por otro lado, si en lugar de reemplazar p_k y β se reemplaza con σ_k^* y σ_{max}^* se llega a que:

$$D_{KL}(f||g) = N\ln\frac{\sigma_{\text{max}}^*}{\sigma_k^*} - \frac{N}{2} + \frac{N}{2}\left(\frac{\sigma_k^*}{\sigma_{\text{max}}^*}\right)^2 \tag{A.23}$$

Bibliografía

Agarwal, S., Mierle, K., and Team, T. C. S. (2023). Ceres Solver. Apache-2.0.

Barfoot, T. D. and Furgale, P. T. (2014). Associating uncertainty with three-dimensional poses for use in estimation problems. *IEEE Transactions on Robotics*, 30(3):679–693.

Bay, H., Tuytelaars, T., and Van Gool, L. (2006). Surf: Speeded up robust features. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9*, pages 404–417. Springer.

Blanco, J.-L., Fernandez-Madrigal, J.-A., and González, J. (2008). A novel measure of uncertainty for mobile robot slam with rao—blackwellized particle filters. *The International Journal of Robotics Research*, 27(1):73–89.

Bosse, M., Zlot, R., and Flick, P. (2012). Zebedee: Design of a spring-mounted 3-d range sensor with application to mobile mapping. *IEEE Transactions on Robotics*, 28(5):1104–1119.

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., and 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.

Calonder, M., Lepetit, V., Strecha, C., and Fua, P. (2010). Brief: Binary robust independent elementary features. In *European conference on computer vision*, pages 778–792. Springer.

Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, J. M., and 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.

Carrillo, H., Dames, P., Kumar, V., and Castellanos, J. A. (2018). Autonomous robotic exploration using a utility function based on rényi's general theory of entropy. *Autonomous Robots*, 42(2):235–256.

Carrillo, H., Reid, I., and Castellanos, J. A. (2012). On the comparison of uncertainty criteria for active slam. In *2012 IEEE International Conference on Robotics and Automation*, pages 2080–2087. IEEE.

Castellanos, J. A., Neira, J., and Tardós, J. D. (2004). Limits to the consistency of ekf-based slam. *IFAC Proceedings Volumes*, 37(8):716–721.

Castro, G. I. (2021). *Mejoras de escalabilidad y consistencia en sistemas SLAM con aplicaciones en exploración activa multi-robot.* PhD thesis, FACULTAD DE CIENCIAS EXACTAS Y NATURALES DEPARTAMENTO DE COMPUTACIÓN Mejoras ....

Cheein, F. A. A. and Guivant, J. (2014). Slam-based incremental convex hull processing approach for treetop volume estimation. *Computers and Electronics in Agriculture*, 102:19–30.

Connolly, C. (1985). The determination of next best views. In *Proceedings. 1985 IEEE international conference on robotics and automation*, volume 2, pages 432–435. IEEE.

Davison, A. J. and Murray, D. W. (2002). Simultaneous localization and map-building using active vision. *IEEE transactions on pattern analysis and machine intelligence*, 24(7):865–880.

Dellaert, F. (2012). Factor graphs and gtsam: A hands-on introduction. *Georgia Institute of Technology, Tech. Rep*, 2:4.

Dissanayake, M. G., Newman, P., Clark, S., Durrant-Whyte, H. F., and Csorba, M. (2001). A solution to the simultaneous localization and map building (slam) problem. *IEEE Transactions on robotics and automation*, 17(3):229–241.

Endres, D. M. and Schindelin, J. E. (2003). A new metric for probability distributions. *IEEE Transactions on Information theory*, 49(7):1858–1860.

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395.

Fridovich-Keil, S., Yu, A., Tancik, M., Chen, Q., Recht, B., and Kanazawa, A. (2022). Plenoxels: Radiance fields without neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5501–5510.

Gaia, J., Orosco, E., Rossomando, F., and Soria, C. (2023). Mapping the landscape of slam research: A review. *IEEE Latin America Transactions*, 21(12):1313–1336.

Gal, Y. and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR.

Gálvez-López, D. and Tardos, J. D. (2012). Bags of binary words for fast place recognition in image sequences. *IEEE Transactions on Robotics*, 28(5):1188–1197.

Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., et al. (2023). A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 56(Suppl 1):1513–1589.

Grisetti, G., Stachniss, C., and Burgard, W. (2007). Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE transactions on Robotics*, 23(1):34–46.

Haner, S. and Heyden, A. (2012). Covariance propagation and next best view planning for 3d reconstruction. In *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part II 12*, pages 545–556. Springer.

Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press.

Hershey, J. R. and Olsen, P. A. (2007). Approximating the kullback leibler divergence between gaussian mixture models. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, volume 4, pages IV–317. IEEE.

Hornung, A., Wurm, K. M., Bennewitz, M., niss, C., and Burgard, W. (2013). Octomap: An efficient probabilistic 3d mapping framework based on octrees. *Autonomous robots*, 34(3):189–206.

Huang, S. and Dissanayake, G. (2016). A critique of current developments in simultaneous localization and mapping. *International Journal of Advanced Robotic Systems*, 13(5):1729881416669482.

Joint Committee for Guides in Metrology, (JCGM/WG 1) (1995). Guide to the expression of uncertainty in measurement. *Geneva, Switzerland*, 122:16–17.

Karaman, S. and Frazzoli, E. (2011). Sampling-based algorithms for optimal motion planning. *The international journal of robotics research*, 30(7):846–894.

Karaman, S., Walter, M. R., Perez, A., Frazzoli, E., and Teller, S. (2011). Anytime motion planning using the rrt. In *2011 IEEE international conference on robotics and automation*, pages 1478–1483. IEEE.

Kerbl, B., Kopanas, G., Leimkühler, T., and Drettakis, G. (2023). 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4).

Kim, Y. and Kim, A. (2017). On the uncertainty propagation: Why uncertainty on lie groups preserves monotonicity? In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3425–3432. IEEE.

Koga, S., Asgharivaskasi, A., and Atanasov, N. (2021). Active exploration and mapping via iterative covariance regulation over continuous se (3) trajectories. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2735–2741. IEEE.

Koga, S., Asgharivaskasi, A., and Atanasov, N. (2022). Active slam over continuous trajectory and control: A covariance-feedback approach. In *2022 American Control Conference (ACC)*, pages 5062–5068. IEEE.

Kohlbrecher, S., Von Stryk, O., Meyer, J., and Klingauf, U. (2011). A flexible and scalable slam system with full 3d motion estimation. In *2011 IEEE international symposium on safety, security, and rescue robotics*, pages 155–160. IEEE.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., and Burgard, W. (2011a). g 2 o: A general framework for graph optimization. In *2011 IEEE international conference on robotics and automation*, pages 3607–3613. IEEE.

Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., and Burgard, W. (2011b). g 2 o: A general framework for graph optimization. In *2011 IEEE International Conference on Robotics and Automation*, pages 3607–3613. IEEE.

Lakshminarayanan, B., Pritzel, A., and Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, 30.

LaValle, S. M. and Kuffner Jr, J. J. (2001). Randomized kinodynamic planning. *The international journal of robotics research*, 20(5):378–400.

Lin, J. and Zhang, F. (2020). Loam livox: A fast, robust, high-precision lidar odometry and mapping package for lidars of small fov. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3126–3131. IEEE.

Lluvia, I., Lazkano, E., and Ansuategi, A. (2021). Active mapping and robot exploration: A survey. *Sensors*, 21(7):2445.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60:91–110.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2020). Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*.

Moravec, H. P. (1989). Sensor fusion in certainty grids for mobile robots. In *Sensor devices and systems for robotics*, pages 253–276. Springer.

Mur-Artal, R., Montiel, J. M. M., and Tardos, J. D. (2015). Orb-slam: a versatile and accurate monocular slam system. *IEEE transactions on robotics*, 31(5):1147–1163.

Mur-Artal, R. and Tardós, J. D. (2017). Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE transactions on robotics*, 33(5):1255–1262.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Newman, P., Bosse, M., and Leonard, J. (2003). Autonomous feature-based exploration. In *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*, volume 1, pages 1234–1240. IEEE.

O'Callaghan, S. T. and Ramos, F. T. (2012). Gaussian process occupancy maps. *The International Journal of Robotics Research*, 31(1):42–62.

Placed, J. A. and Castellanos, J. A. (2022). Enough is enough: Towards autonomous uncertainty-driven stopping criteria. *IFAC-PapersOnLine*, 55(14):126–132.

Placed, J. A., Strader, J., Carrillo, H., Atanasov, N., Indelman, V., Carlone, L., and Castellanos, J. A. (2023). A survey on active simultaneous localization and mapping: State of the art and new frontiers. *IEEE Transactions on Robotics*.

Popovic, M., Hitz, G., Nieto, J., Sa, I., Siegwart, R., and Galceran, E. (2016). Online informative path planning for active classification using uavs. *arXiv preprint arXiv:1609.08446*.

Popović, M., Vidal-Calleja, T., Hitz, G., Chung, J. J., Sa, I., Siegwart, R., and Nieto, J. (2020). An informative path planning framework for uav-based terrain monitoring. *Autonomous Robots*, 44(6):889–911.

Potena, C., Khanna, R., Nieto, J., Siegwart, R., Nardi, D., and Pretto, A. (2019). Agricolmap: Aerial-ground collaborative 3d mapping for precision farming. *IEEE Robotics and Automation Letters*, 4(2):1085–1092.

Pretto, A., Aravecchia, S., Burgard, W., Chebrolu, N., Dornhege, C., Falck, T., Fleckenstein, F., Fontenla, A., Imperoli, M., Khanna, R., et al. (2020). Building an aerial–ground robotics system for precision farming: an adaptable solution. *IEEE Robotics & Automation Magazine*, 28(3):29–49.

Raj, T., Hanim Hashim, F., Baseri Huddin, A., Ibrahim, M. F., and Hussain, A. (2020). A survey on lidar scanning mechanisms. *Electronics*, 9(5):741.

Rivera, G., Porras, R., Florencia, R., and Sánchez-Solís, J. P. (2023). Lidar applications in precision agriculture for cultivating crops: A review of recent advances. *Computers and electronics in agriculture*, 207:107737.

Rodríguez-Arévalo, M. L., Neira, J., and Castellanos, J. A. (2018). On the importance of uncertainty representation in active slam. *IEEE Transactions on Robotics*, 34(3):829–834.

Rosell, J. and Sanz, R. (2012). A review of methods and applications of the geometric characterization of tree crops in agricultural activities. *Computers and electronics in agriculture*, 81:124–141.

Rosten, E. and Drummond, T. (2006). Machine learning for high-speed corner detection. In *European conference on computer vision*, pages 430–443. Springer.

Rublee, E., Rabaud, V., Konolige, K., and Bradski, G. (2011). Orb: An efficient alternative to sift or surf. In *2011 International conference on computer vision*, pages 2564–2571. Ieee.

Schonberger, J. L. and Frahm, J.-M. (2016). Structure-from-motion revisited. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4104–4113.

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., and Xu, X. (2017). Dbscan revisited, revisited: why and how you should (still) use dbscan. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21.

Secchi, H., Carelli, R., and Mut, V. (1999). Design of stable algorithms for mobile robot control with obstacle avoidance. *IFAC Proceedings Volumes*, 32(2):8456–8461.

Secchi, H. A. (2008). *Algoritmos para el control de robots basado en comportamientos*. PhD thesis, Universidad Nacional de San juan.

Segal, A., Haehnel, D., and Thrun, S. (2009). Generalized-icp. In *Robotics: science and systems*, volume 2, page 435. Seattle, WA.

Sola, J. (2010). Consistency of the monocular ekf-slam algorithm for three different landmark parametrizations. In *2010 IEEE International Conference on Robotics and Automation*, pages 3513–3518. IEEE.

Stachniss, C. (2006). *Exploration and Mapping with Mobile Robots*. PhD thesis, University of Freiburg, Freiburg, Germany.

Stachniss, C. (2009). *Robotic mapping and exploration*, volume 55. Springer.

Stachniss, C., Grisetti, G., and Burgard, W. (2005). Information gain-based exploration using rao-blackwellized particle filters. In *Robotics: Science and systems*, volume 2, pages 65–72.

Stachniss, C., Hahnel, D., and Burgard, W. (2004). Exploration with active loop-closing for fastslam. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, volume 2, pages 1505–1510. IEEE.

Strasdat, H., Davison, A. J., Montiel, J. M., and Konolige, K. (2011). Double window optimisation for constant time visual slam. In *2011 international conference on computer vision*, pages 2352–2359. IEEE.

Sun, K., Mohta, K., Pfrommer, B., Watterson, M., Liu, S., Mulgaonkar, Y., Taylor, C. J., and Kumar, V. (2018). Robust stereo visual inertial odometry for fast autonomous flight. *IEEE Robotics and Automation Letters*, 3(2):965–972.

Thrun, S. (2002). Probabilistic robotics. *Communications of the ACM*, 45(3):52–57.

Thrun, S. (2003). Learning occupancy grid maps with forward sensor models. *Autonomous robots*, 15(2):111–127.

Triggs, B., McLauchlan, P. F., Hartley, R. I., and Fitzgibbon, A. W. (2000). Bundle adjustment—a modern synthesis. In *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21–22, 1999 Proceedings*, pages 298–372. Springer.

Tzoumas, V., Atanasov, N. A., Jadbabaie, A., and Pappas, G. J. (2017). Scheduling nonlinear sensors for stochastic process estimation. In *2017 American Control Conference (ACC)*, pages 580–585. IEEE.

Underwood, J. P., Hung, C., Whelan, B., and Sukkarieh, S. (2016). Mapping almond orchard canopy volume, flowers, fruit and yield using lidar and vision sensors. *Computers and electronics in agriculture*, 130:83–96.

Van Parys, B. P., Goulart, P. J., and Kuhn, D. (2016). Generalized gauss inequalities via semidefinite programming. *Mathematical Programming*, 156:271–302.

Vougioukas, S. G. (2019). Agricultural robotics. *Annual review of control, robotics, and autonomous systems*, 2(1):365–392.

Westling, F., Mahmud, K., Underwood, J., and Bally, I. (2020). Replacing traditional light measurement with lidar based methods in orchards. *Computers and Electronics in Agriculture*, 179:105798.

Xu, W., Cai, Y., He, D., Lin, J., and Zhang, F. (2022). Fast-lio2: Fast direct lidar-inertial odometry. *IEEE Transactions on Robotics*, 38(4):2053–2073.

Xu, W. and Zhang, F. (2021). Fast-lio: A fast, robust lidar-inertial odometry package by tightly-coupled iterated kalman filter. *IEEE Robotics and Automation Letters*, 6(2):3317–3324.

Yamauchi, B. (1997). A frontier-based approach for autonomous exploration. In *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97.'Towards New Computational Principles for Robotics and Automation'*, pages 146–151. IEEE.

Zhang, J. and Singh, S. (2014). Loam: Lidar odometry and mapping in real-time. In *Robotics: Science and systems*, volume 2, pages 1–9. Berkeley, CA.

Zhang, Z. (2020). *Active robot vision: from state estimation to motion planning*. PhD thesis, Universität Zürich.

Zhang, Z. and Scaramuzza, D. (2018). Perception-aware receding horizon navigation for mavs. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2534–2541. IEEE.

Zhang, Z. and Scaramuzza, D. (2019). Beyond point clouds: Fisher information field for active visual localization. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 5986–5992. IEEE.

Zhang, Z. and Scaramuzza, D. (2020). Fisher information field: an efficient and differentiable map for perception-aware planning. *arXiv preprint arXiv:2008.03324*.

ISBN 978-631-90962-2-4



ISBN: 978-631-90-9622-4