



*Estimación de pose de robots móviles
mediante mediciones de sensores inerciales
y una cámara externa*

Autor:

RUBEN IGNACIO COUOH KU

Directores:

Dr. Arturo Espinosa Romero

Dr. Carlos Brito Loeza

Programa de Maestría en Ciencias de la Computación

Facultad de Matemáticas

Universidad Autónoma de Yucatán

Enero de 2020

Este documento está preparado para ser imprimido a doble cara.

Resumen

Los robots manipuladores son usados de forma rutinaria en una amplia variedad de tareas en la industria manufacturera, en este contexto industrial los robots se especializan en una tarea y son diseñados para operar entornos totalmente controlados, por estas razones, se alcanza un alto nivel de exactitud. En contraste, los robots móviles operan en entornos no estructurados y dinámicos teniendo que lidiar con una gran diversidad de circunstancias no previstas, por lo tanto, la robótica móvil demanda cierta autonomía.

La navegación es una tarea esencial para alcanzar la autonomía en la robótica móvil, para navegar apropiadamente el robot tiene hallar la respuesta a tres preguntas fundamentales “¿dónde estoy?”, “¿dónde está todo lo demás?” y “¿qué hacer para alcanzar los objetivos o lugares particulares?”, es decir, ejecutar tareas de localización, mapeo, y planificación del movimiento. Estas tareas no pueden ser resultados de forma independiente, el primer paso para navegar consiste en determinar la localización del robot, así, la localización constituye los cimientos de la navegación, aunque existe una relación de dependencia mutua.

La localización consiste en el problema de estimar la pose (posición y orientación) del robot con respecto a su entorno. Diferentes métodos han sido propuestos para abordar este problema, entre ellos se encuentran la navegación inercial y los sistemas de localización basados en visión, ambos tratados en este trabajo.

Los sistemas de navegación inercial determinan su pose de forma incremental a partir de detectar el cambio de movimiento del vehículo, su solución es muy exacta en periodos de tiempo cortos, sin embargo, se degenera progresivamente con el tiempo. Una de sus principales ventajas es que no depende de ningún tipo de señal externa como sucede con los GPS, por ello, estos sistemas son útiles donde las señales externas se encuentren bloqueadas o no sea conveniente acceder a ellas.

En contraste, los sistemas de localización basados en visión no presentan divergencia en su solución, en consecuencia, son adecuados para mejorar la salida de los sistemas de navegación inercial. Estos sistemas determinan su pose a partir de las observaciones de un conjunto de características del entorno y sus correspondencias con el mundo.

La incertidumbre es el problema más crucial en la robótica móvil debido a que los robots operan en entornos impredecibles y el bajo costo de los sensores normalmente empleados dificulta establecer la correspondencia adecuada con las características del entorno para así poder determinar la pose del robot con un alto grado de exactitud.

En la práctica, todos estos errores, por muy pequeños que sean, contribuyen a la deriva casi inevitable a largo plazo, por ellos, diferentes autores han desarrollado varias técnicas para abordar este problema, entre ellas destacan la detección de cierre de bucle, relocalización y modelado de errores. Aunque las primeras dos tareas no pueden resolverse sin un mapa, el modelado de incertidumbre ha demostrado dar buenos resultados y puede o no utilizarse en conjunto con los métodos anteriores para aumentar la certeza de la estimación. El modelado de incertidumbre utilizará conceptos de la teoría de la probabilidad y es el enfoque para abordar el problema de deriva en este trabajo.

Específicamente, en este trabajo se usa la técnica denominada filtro de Kalman Unscented para estimar la pose del robot, esta técnica modela el sistema en términos de distribuciones de probabilidad para dar la estimación más probable a partir de información proveniente de diferentes fuentes, ponderadas de acuerdo al nivel de confianza que se tiene de cada una de ellas. La información necesaria es provista por el sistema de navegación inercial y el sistema de localización basado en visión, donde los sensores inerciales se utilizan para predecir el estado y una cámara externa para corregirlo.

Índice

Resumen	v
1. Introducción	1
1.1. La robótica	1
1.2. Visión artificial	2
1.3. La navegación	3
1.3.1. Navegación robótica en interiores	3
1.4. Descripción del problema	4
1.5. Estructura de la tesis	5
2. Sistema de Navegación Inercial	7
2.1. Principios de la navegación inercial	7
2.1.1. Tipos de sistema de navegación inercial	7
2.2. Marcos de referencia para la navegación inercial	8
2.3. Representación de la orientación	9
2.4. Sensores inerciales	10
2.4.1. Principales fuentes de error	10
2.4.2. Giroscopio	12
2.4.3. Acelerómetro	14
2.4.4. Magnetómetro	16
2.5. Orientación entre dos sistemas de referencia	17
2.6. Seguimiento de la orientación	18
2.6.1. Tasa de cambio de la orientación	18
2.6.2. Actualización de la orientación	20
2.7. Seguimiento de la posición	20
3. Localización visual	23
3.1. Visión robótica	23
3.2. Localización basada en visión	23
3.3. Movimiento de un cuerpo rígido	24
3.4. Formación de imágenes	24
3.5. Forma general de la proyección	27
3.5.1. Parámetros intrínsecos	27
3.6. Parámetros extrínsecos	28
3.7. Perspectiva desde N Puntos	28
3.8. Conclusiones	29
4. Filtro de Kalman	31
4.1. Filtros de Kalman	31
4.1.1. Proceso de estimación	31
4.1.2. Fase de corrección:	32
4.1.3. Limitaciones	32
4.2. El filtro de Kalman Extendido	33
4.2.1. Proceso de estimación	33

4.2.2. Limitaciones	34
4.3. El filtro de Kalman Unscented	34
4.3.1. Trasformada Unscented	35
4.3.2. Proceso de estimación	36
4.3.3. Fortalezas	37
5. Metodología	39
5.1. Descripción del sistema	39
5.2. Configuración de los marcos de referencia	40
5.3. Localización usando sensores inerciales	40
5.3.1. Seguimiento de la orientación	41
5.3.2. Cálculo de la orientación inicial	42
5.3.3. Seguimiento de la posición	42
5.4. Localización visual basada en marcas	42
5.4.1. Modelo 3D	43
5.4.2. Extracción y correspondencia de características	44
5.4.3. Recuperación de la pose	44
5.5. Modelo del sistema	45
5.5.1. Modelo del proceso	45
5.5.2. Modelos de observación	46
5.6. Estimación del estado	47
5.6.1. Puntos sigma	47
5.6.2. Cálculo de la media	47
5.6.3. Calculo de la matriz de covarianzas	48
6. Experimentación y Resultados	49
6.1. Entorno de experimentación	49
6.2. Metodología de experimentación	50
6.2.1. Evaluación	50
6.3. Resultados	51
6.3.1. Experimento 1	51
6.3.2. Experimento 2	53
6.3.3. Experimento 3	53
6.4. Conclusiones	55
Bibliografía	57

Índice de figuras

1.1. Esquema de interacción de un robot en su entorno.	2
2.1. Sistema de plataforma estabilizada.	8
2.2. Marcos de referencia (Kok et al., 2017).	9
2.3. Errores comunes en las mediciones de los sensores.	10
2.4. Diagrama esquemático de un giroscopio mecánico. Fuente: Titterton et al. (2004).	12
2.5. Principio de operación de un giroscopio de diapasón. Fuente: Titterton et al. (2004).	13
2.6. Acelerómetro mecánico. Fuente: Titterton et al. (2004).	14
2.7. Acelerómetro de superficie de onda acústica. Fuente: Titterton et al. (2004).	15
2.8. Líneas del campo magnético de la Tierra. Fuente: Kok et al. (2017).	16
2.9. Representaciones del campo magnético de la Tierra. Fuente: Wikipedia.	16
3.1. Representación de una imagen en la computadora (documentación oficial de OpenCV).	23
3.2. Movimiento de un cuerpo rígido relativo al marco del mundo W (Ma et al., 2012).	24
3.3. Esquema de una cámara pinhole.	25
3.4. Tipos de distorsión más comunes que afectan las lentes de las cámaras.	25
3.5. Modelo de la geometría de una cámara pinhole (Xu y Zhang, 2013).	26
3.6. Modelo equivalente de una cámara pinhole (Xu y Zhang, 2013).	26
3.7. Proyección de puntos en el sistema coordenado del mundo al plano de la imagen (Xu y Zhang, 2013).	28
4.1. Diagrama de bloques que ilustra el procedimiento de la UT (<i>Unscented Transformation</i>) (Wan y Van Der Merwe, 2001).	35
4.2. Ejemplo comparativo de la UT para calcular la media y la covarianza de un sistema bidimensional (Wan y Van Der Merwe, 2000).	38
5.1. Esquema general del seguimiento de la localización 3D del vehículo usando sensores inerciales, una cámara y el filtro de Kalman Unscented.	39
5.2. Principales marcos de referencia para la estimación de la pose: marco de navegación, marco del cuerpo y marco del patrón de puntos denotados con las letras \mathbf{n} , \mathbf{b} y \mathbf{p} , respectivamente.	40
5.3. Diagrama de bloques para la estimación de la localización usando navegación inercial.	41
5.4. Diagrama de bloques del procedimiento para la estimación de la localización del vehículo usando visión artificial.	43
5.5. Modelo 3D usado para el seguimiento del vehículo.	43
5.6. Marcas colocadas sobre los puntos característicos: círculo negro con radio igual a dos centímetros y círculo blanco con radio igual a un centímetro.	44
5.7. Resultado de la extracción de los puntos de interés representados en color verde.	44
5.8. Seguimiento del vehículo.	45
6.1. IMU (<i>Inertial Measurement Unit</i>) modelo BNO055 de 9-DOF (<i>Degrees-Of-Freedom</i>): empaquetada con tres acelerómetros, tres giroscopios y tres magnetómetros.	49
6.2. Cámara modelo Logitech.	50

6.3. Escenario usado para evaluar la solución del algoritmo de estimación de pose: las líneas punteadas representan la trayectoria de referencia; las flechas azules indican el sentido del movimiento; la marca de color verde representa la coordenada donde inicia y finaliza el movimiento.	51
6.4. Estimación de la pose del vehículo usando el algoritmo de navegación inercial en modo estacionario: los resultados muestran la degeneración de la solución en el tiempo debido a la acumulación de errores.	52
6.5. Escenario usado para la estimación de la pose basada en visión: El escenario real (izquierda) consta de una forma rectangular de 1.36×0.76 metros, dispuesta sobre una mesa de tal manera que quede dentro del campo visual de la cámara externa, a la izquierda su representación simbólica.	53
6.6. Estimación de la posición basada en visión: la velocidad promedio se calcula mediante la posición asumiendo que la velocidad es constante en las secciones rectas de la trayectoria, descartando el movimiento en las curvas donde ocurren las transiciones del cambio de rumbo.	54
6.7. Estimación de la orientación basada en visión: en la figura 6.7a se muestra la orientación en términos de los ángulos roll, pitch y yaw, y en la figura 6.7b se presentan sus correspondientes errores de estimación usando la raíz del error cuadrático medio.	54
6.8. Disposición del entorno usado para la estimación de la pose basada en el UKF (<i>Unscented Kalman Filter</i>): la figura 6.8a muestra la representación simbólica del escenario y la figura 6.8b ilustra el escenario real, el movimiento se efectuó siguiendo como guía el borde de la mesa con dimensiones 1.92×1.30 metros.	54
6.9. Estimación de posición basada en el UKF: la figura 6.9a muestra los resultados del UKF y la figura 6.9b muestra una comparativa de la solución entre el INS (<i>Inertial Navigation System</i>) y el UKF.	55
6.10. Estimación de la orientación basada en el UKF: la figura 6.10a muestra la orientación en términos de los ángulos roll, pitch y yaw, y en la figura 6.10b se presentan los errores para cada ángulo usando la métrica RMSE (<i>Root Mean Square Error</i>).	55

Índice de Tablas

6.1. Características de las señales de salida de la IMU.	49
--	----

Capítulo 1

Introducción

Este capítulo describe el problema de investigación que se desarrollará a lo largo de este documento. Antes de describir el problema de investigación se hace una breve introducción a la robótica, la visión artificial y la navegación, siendo estas las áreas en las que se circunscribe esta investigación. Después, se expresa lo que se desea indagar mediante la descripción del problema y la definición de los objetivos. Por último, se resume la estructura del resto del documento.

1.1. La robótica

Es evidente que a través de la historia el ser humano ha diseñado y construido desde simples artefactos hasta máquinas más complejas y sofisticadas con el fin de imitar el comportamiento humano para realizar principalmente tareas repetitivas, pesadas o peligrosas. Durante las últimas décadas ha aumentado la tendencia en transferir las habilidades humanas hacia entes electromecánicos conocidos como robots, cuyo fin es dotarlos con la autonomía suficiente para trabajar conjuntamente con los seres humanos en actividades rutinarias.

El concepto de un ente inteligente se remonta a nuestro pasado prehistórico, no obstante, la palabra robot fue introducida hasta 1920 por Karel Čapek en su obra de ciencia ficción *Rossum's Universal Robots*. La palabra robot proviene de la palabra checa *robot* que significa servidumbre o trabajador forzado, en esta obra se utiliza para nombrar a hombres artificiales con capacidad de pensar, aunque estas máquinas fueron creadas con el propósito de ayudar terminan destruyendo a la humanidad (Zamalloa et al., 2017).

La palabra robótica fue definida, en 1942, por Isaac Asimov como la ciencia que estudia los robots, en nuestro días, la robótica se ha convertido en un área multidisciplinaria que involucra disciplinas tales como la mecánica, la electrónica, la informática, la inteligencia artificial, la ingeniería de control, la física y otras; todas estas colaboran en el desarrollo del hardware y software del robot (Zamalloa et al., 2017).

A pesar que hace aproximadamente un siglo que se conoce el concepto de robot, fue hasta hace veinte años que los brazos robóticos alcanzaron las condiciones necesarias para ser utilizados en masa de forma rutinaria en cadenas de montaje industrial. Los brazos robóticos o robots manipuladores han superado en gran medida las habilidades humanas en tareas repetitivas, sin embargo, en la robótica móvil sigue siendo un reto imitar las capacidades innatas del ser humano que le permitan a los robots operar autónomamente en entornos complejos (Huat, 2006).

Un robot móvil es una máquina con capacidad de desplazarse en su entorno con cierta autonomía para reaccionar ante una gran variedad de circunstancias, para poder alcanzar dicha autonomía un robot deberá disponer tres habilidades fundamentales: percepción, decisión y acción.

El primer robot con estas capacidades fue desarrollado, en 1968, por el instituto de investigación de Stanford. Este robot llamado Shakey equipado con diversos sensores incluidos una cámara y sensores táctiles era capaz de navegar de forma autónoma sobre el suelo evitando obstáculos. Además, es considerado el primer robot inteligente de la historia y predecesor de los vehículos autónomos.

Actualmente, compañías como Tesla, Delphi, Uber y Google trabajan en el desarrollo de vehículos autónomos que han culminado en carros eficientes y seguros con funcionalidades avanzadas de autoconducción, pero para alcanzar vehículos totalmente autónomos se deberán incluir las capacidades pertinentes que les permitan a los vehículos comprender e interactuar naturalmente con los conductores humanos (Brown, 2017).

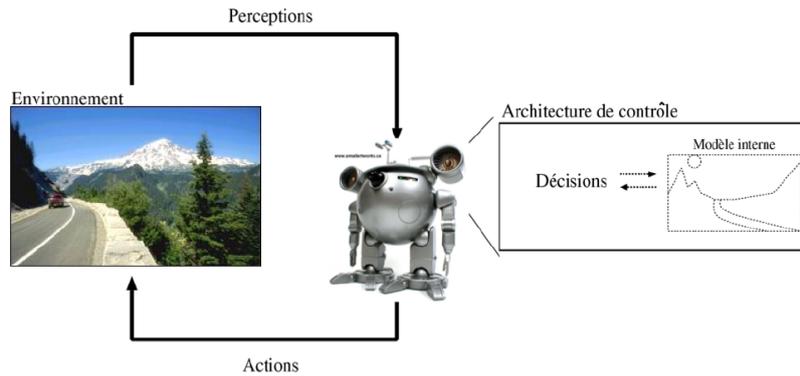


FIGURA 1.1: Esquema de interacción de un robot en su entorno.

Se estima que los avances en la inteligencia artificial y la robótica se verán reflejados en un período de gran prosperidad y rápido crecimiento, al grado de contar con robots médicos autónomos aproximadamente en veinte años, los cuales tendrán capacidades de proporcionar la medicación correcta, detectar irregularidades y actuar apropiadamente (Zamalloa et al., 2017).

Este trabajo se centra en robots móviles especializados en desplazarse sobre el suelo, como los ejemplos descritos anteriormente.

1.2. Visión artificial

La incertidumbre es el problema más crucial al que debe enfrentarse un robot móvil, por lo tanto, una capacidad clave de un robot es la interpretación sensorial cuyo propósito es relacionar las señales de salida de los sensores con las propiedades tridimensionales del entorno (Lozano-Perez, 2012).

La mayoría de las aplicaciones modernas utilizan una cámara como sensor principal para alcanzar dicho objetivo, ya que este es uno de los sensores que provee la información más valiosa del entorno. Para recuperar esta información se ha desarrollado una amplia variedad de técnicas en visión artificial y que pueden ser complementadas con otras técnicas utilizadas para el mismo propósito, entre las que se encuentran la odometría; detección de alcance por ultrasonidos, infrarrojos y láser; y sistemas inerciales.

En la percepción basada en visión, un robot usa una cámara con el apoyo de visión artificial como medio para percibir su entorno así como el ser humano usa los ojos en el proceso de visión para el mismo fin, por esta razón, un robot deberá ser dotado con capacidad visual.

La capacidad visual es uno de los sentidos más importantes que posee el ser humano, este sentido nos permite interactuar con el mundo en entornos altamente dinámicos realizando tareas complejas, engañosamente sin mucho esfuerzo. Esta simplicidad aparente llevó a pensar que sería fácil dotar a una máquina con capacidad visual.

Según la definición de Marr (1982), la visión es un proceso que produce a partir de las imágenes del mundo exterior una descripción que es útil para el observador y que no tiene información irrelevante.

El proceso que da como resultado la experiencia visual se basa en la siguiente idea: la luz que percibe el ojo pasa a través de la cornea para llegar a la retina donde se encuentran células especializadas (conos, bastones y demás) que transforman la luz en impulsos nerviosos formando una imagen en la retina, estos impulsos viajan a través del nervio óptico al cerebro para ser procesados por la corteza visual (Wandell, 1995).

La serie de transformaciones visuales que ocurren en el ojo para enfocar y recoger la luz están bien entendidas, sin embargo, no se tiene un conocimiento profundo de como el cerebro procesa las señales recibidas para reconocer la forma de los objetos, identificar distancias, detecta colores y movimiento, todo esto en una pequeña fracción de tiempo. El conocimiento parcial sobre los mecanismos biológicos involucrados en el proceso visual ha dificultado su transferencia de forma eficiente a los robots.

La visión artificial es un área de las ciencias que tiene como objetivo principal extraer información útil contenida en imágenes para su interpretación por una computadora. Esta es aplicable en casi todo tipo de tareas, algunas de las más relevantes relacionadas con la robótica que han sido posibles gracias a la visión artificial son: detección de objetos, evitar obstáculos, localización en el espacio; la última es la de mayor interés en este trabajo.

1.3. La navegación

La navegación es, en sentido amplio, el conjunto de técnicas y procedimientos que permiten conducir un vehículo desde un lugar origen hasta su lugar de destino de forma segura.

La navegación comenzó con balsas talladas en madera que posteriormente evolucionaron en embarcaciones propulsadas por velas y remos de apoyo que utilizaban recursos de la época; la ballestilla, la aguja imantada, la corredera de nudos, el reloj de arena, y las estrellas como mapa de referencia. Estas formas de navegación primitivas quedaban a merced de los cambios climáticos y de la experiencia de los navegantes, pero la competencia por la supremacía entre las civilizaciones aceleró el desarrollo de dispositivos muy precisos y nuevas técnicas de navegación que no dependen de las condiciones atmosféricas.

En la actualidad, la navegación ha sido estudiada extensamente y no guarda ningún misterio, es usada de forma convencional en aplicaciones marítimas, de aviación y espaciales. Independientemente del contexto en cual la navegación tenga lugar, las funciones de la navegación se pueden expresar mediante tres preguntas fundamentales (Jul y Furnas, 1997; Murphy, 2000):

1. ¿Dónde estoy?
2. ¿Dónde está todo lo demás?
3. ¿Qué hacer para alcanzar objetos o lugares particulares?

En la robótica, la primera pregunta se plantea como el problema de localización cuyo objetivo consiste en inferir la posición y orientación del vehículo con base en la información del entorno percibida mediante los sensores. La segunda pregunta se aborda como el problema de construir gradualmente, conforme se exploran nuevos lugares, una representación del entorno a partir de la información provista por los sensores, esta representación generalmente recibe el nombre de mapa. Finalmente, la tercera pregunta trata sobre el razonamiento de que acciones seguir para alcanzar el objetivo o lugares particulares.

Según la tecnología empleada para la navegación, esta se puede clasificar en (Iwaneczko et al., 2018): navegación por satélite; navegación por astronomía; navegación por piloto y radio; navegación costera; navegación por radar; y navegación inercial.

Uno de los temas centrales de este trabajo es la navegación inercial, la navegación inercial es una técnica cuyo objetivo es determinar progresivamente la posición y la orientación de un vehículo a partir de las mediciones obtenidas de sensores de movimiento y rotación. Es considerada una técnica de navegación autónoma al no depender de una infraestructura de localización ad hoc, sin embargo, el sistema pierde progresivamente su precisión con el tiempo a partir de que se alinea. Esta degeneración se debe a los errores acumulados en cada paso de la estimación, de manera que es muy adecuado combinarlo otro sistema que sirvan de referencia para acotar la tasa de crecimiento del error, específicamente en este trabajo se combina con un sistema de localización basado en visión.

La navegación es una de las habilidades fundamentales para lograr la autonomía en la robótica móvil, aunque en otros contextos se usa de forma rutinaria, en la robótica sigue siendo un reto debido a la incertidumbre inherente de los entornos no estructurados y dinámicos en los cuales tienen que operar los robots. Otro factor determinante es la calidad de los instrumentos empleados en la navegación robótica, los sensores comúnmente usados suele ser de bajo costo y en consecuencia de baja calidad si se comparan con los dispositivos de alta confiabilidad usados en otros contextos de navegación, por lo tanto, la información extraída de los sensores suele ser pobre para establecer la correspondencia adecuada con las características del entorno.

1.3.1. Navegación robótica en interiores

La navegación en interiores se puede dividir en tres categorías: navegación basada en mapas, navegación basada en construcción de mapas y navegación sin mapas (DeSouza y Kak, 2002).

Los sistemas de navegación basados en mapas requieren un modelo a priori de las características del entorno, el mapa es construido por el usuario y puede contener cualquier tipo de información que facilite la navegación, sin embargo, en diversos escenarios la construcción del modelo que describe el entorno puede llegar a ser una tarea difícil, especialmente cuando se incluye información métrica.

Esta dificultad motivó a los investigadores a idear métodos de navegación basados en la construcción de mapas, estos sistemas usan los sensores para construir gradualmente un modelo del entorno y luego usan este modelo para la navegación. El SLAM (*Simultaneous Localization And Mapping*) es una técnica representativa de esta categoría que ha tenido un desarrollo sorprendente en durante los últimos 30 años.

Esta técnica es computacionalmente costosa, por lo que diferentes autores han propuesto otras técnicas inspiradas en el SLAM que reducen substancialmente en tiempo de computo (Cadena et al., 2016).

El término SLAM se refiere a la habilidad que posee un robot para construir un mapa de su entorno, mientras que al mismo tiempo lo utiliza para determina su localización (Durrant-Whyte y Bailey, 2006). Una de sus características principales es que no depende de un infraestructura de localización, por ello se ha convertido gradualmente en la opción preferida para el posicionamiento de vehículos en entornos complejos como lo es la superficie lunar donde tecnologías como el sistema de posicionamiento GPS o el sensor equidistante de sonda no pueden aplicarse (Wang et al., 2017).

En contraste, la navegación sin mapas no requiere de una representación explicitan del entorno donde se quiere navegar, sino que recurren al reconocimiento y seguimiento de objetos con base a observaciones visuales, su desventaja es que la navegación solo se puede realizar con respecto a estos objetos (DeSouza y Kak, 2002).

Los sistemas modernos basados en visión han ganado popularidad gracias al aumento en las capacidades de computo aunado al desarrollo de algoritmos cada vez más eficientes en visión artificial. Estos sistemas suelen ser relativamente precisos en aplicaciones interiores donde se tiene más control sobre los factores que afectan su rendimiento.

Las imágenes borrosas, cambios de la iluminación, falta de textura, oclusiones, y movimientos rápidos son algunos de los problemas que tienen que enfrentar los sistemas visuales. El interés en superar estos inconvenientes ha aumentado la tendencia de las soluciones que aprovechan las características complementarias de los sensores visuales e inerciales para crear un sistema de estimación robusto (Corke et al., 2007).

La mayor parte de las investigación que trabajan con sistemas basados en sensores inerciales usan filtrado complementario y con mayor frecuencia las soluciones se basan en el filtro de Kalman para dar una estimación óptima o cercana a esta, dependiendo de las restricciones del problema Iwaneczko et al. (2018).

El filtro de Kalman permite inferir el estado no medible de un sistema a partir de observaciones que pueden contener ruido (Kalman, 1960). Los modelos basados en el filtro de Kalman caracterizan el sistema usando matrices de covarianza, sin embargo, las matrices de covarianza no capturan la física relevante, por lo tanto, es necesario implementar los mecanismos necesarios para evitar que mediciones erróneas tengan un impacto negativo en la estimación del estado.

En la práctica, la deriva a largo plazo es casi ineludible, actualmente existen varias métodos especializados abordar este problema que se pueden usar junto con técnicas de modelado de errores para una mejor solución, entre ellas podemos encontrar la detección de cierre de bucle y la relocalización, no obstante, estas tareas requieren un mapa del entorno. Otra forma de tratar con la deriva es incorporando información de las restricciones del dinámica para restablecer el sistema periódicamente como se hacer en el trabajo de Yun et al. (2007), ahí se aprovecha la naturaleza periódica de los movimiento durante una caminata, considerando durante periodos de tiempo cortos velocidad cero cuando el pie está en contacto con el suelo.

Una vez presentado el panorama general en el cual encuadra este trabajo, en la siguiente sección se describe el rumbo que tomará este trabajo.

1.4. Descripción del problema

El objetivo principal de este trabajo consiste en desarrollar un sistema visual-inercial para la estimación de pose de robots móviles en entornos interiores sin mapas.

El problema que se aborda es el de un robot moviéndose a lo largo de una trayectoria, mientras el robot ejecuta las maniobras necesarias para desplazarse habrá estimar el estado del robot, el estado deberá incluir la pose del robot con seis grados de libertad, tres para posición y tres para la orientación. Para determinar la pose se utilizará el filtro de Kalman Unscented utilizando sensores inerciales y una cámara externa, donde la IMU se use propagar el estado y la cámara para corregirlo.

Para indagar sobre el problema presentado en el párrafo anterior se han definido cuatro objetivos particulares:

- Desarrollar un sistema de navegación inercial.
- Desarrollar un sistema de localización basado en visión.

- Desarrollar un modelo que permita fusionar las mediciones de la cámara y los sensores inerciales utilizando el filtro del Kalman Unscented.
- Probar y validar el sistema de estimación de pose en un escenario real.

1.5. Estructura de la tesis

Después de entender el problema que se desarrolla en esta tesis, el resto de los capítulos detallan el trabajo de investigación, estos se estructuran de la siguiente manera:

- El capítulo 2 se centra en aspectos más importantes de los sistemas de navegación inercial, abarcando desde los principios de operación de los sensores hasta las ecuaciones que describen el movimiento del vehículo.
- El capítulo 3 describe de principio de funcionamiento de los sistemas de localización basados en visión, con este propósito, se describen los puntos clave del movimiento de un cuerpo rígido, las proyecciones en perspectiva, finalizando con el algoritmo usado en el seguimiento de la pose una cámara calibrada.
- El capítulo 4 aborda el problema de estimación de estado desde el enfoque del filtro de Kalman.
- En capítulo 5 describe el orden metodológico utilizado para alcanzar los objetivos planteados en el capítulo 1.
- El capítulo 6 presenta la metodología experimental y los resultados obtenidos siguiendo el procedimiento descrito en el capítulo 5.
- El capítulo 7 finaliza este trabajo, detallando las conclusiones que se desprenden del análisis de esta investigación. Además, se proponen los posibles trabajos futuros.

Capítulo 2

Sistema de Navegación Inercial

Este capítulo se centra en el estudio de los sistemas de navegación inercial. Primero se ofrece una breve descripción de los principios de funcionamiento de los sistemas de navegación inercial. Enseguida se definen los marcos de referencia indispensables para la navegación inercial y la manipulación de vectores entre dos marcos de referencia mediante matrices de rotación. Posteriormente, se describen las fuentes de error más comunes que afectan el rendimiento de los sistemas de navegación, se describen los principios de operación de los sensores inerciales y el magnetómetro, incluyendo sus modelos de medición probabilísticos. Finalmente, se derivan las ecuaciones para el seguimiento de la orientación y posición del sistema de navegación.

2.1. Principios de la navegación inercial

Las leyes del movimiento formuladas por Isaac Newton son los principios a partir de los cuales se desarrolla un INS.

Primera ley de Newton: todo cuerpo se preserva en su estado de reposo relativo o movimiento relativo, salvo que haya una fuerza que logre cambiar su estado.

Segunda ley de Newton: el cambio de movimiento es proporcional a la fuerza motriz que actúa sobre él, y se hace en la dirección de la línea recta en la que se aplica esa fuerza.

Estos dos principios constituyen los cimientos de la navegación inercial para hacer predicciones sobre el movimiento del vehículo a partir de la relación entre la razón de cambio de la posición y velocidad:

$$\text{posición : } \mathbf{p}(t), \quad \text{velocidad : } \mathbf{v} = \frac{d\mathbf{p}}{dt}, \quad \text{aceleración : } \mathbf{a} = \frac{d\mathbf{v}}{dt}.$$

La idea fundamental de la técnica de navegación inercial consiste en montar sensores inerciales en un punto del vehículo, por ejemplo su centro de masa, para proveer mediciones de los acelerómetros y los giroscopios en cada instante de tiempo y conociendo las condiciones iniciales, por medio de integrales numéricas, estimar la localización incremental del vehículo en movimiento sin la necesidad de referencias externas.

2.1.1. Tipos de sistema de navegación inercial

Un sistema de navegación inercial se pueden categorizar como sistema de navegación con plataforma estabilizada o sistema de navegación con sensores fijos a la estructura del vehículo según los mecanismos de operación de los sensores usados. La principal diferencia entre ambos es el marco de referencia en el que operan los sensores inerciales.

Sistema con plataforma estabilizada

Los sistemas con plataforma estabilizada son los predecesores de los sistemas con sensores fijos a la estructura del vehículo, en este tipo de sistemas de navegación los sensores inerciales se montan sobre una plataforma que es estabilizada usando principalmente cardanes y motores que reciben señales de corrección del controlador para mantener a la plataforma alineada en un marco de referencia global, como se observa en la figura 2.1.

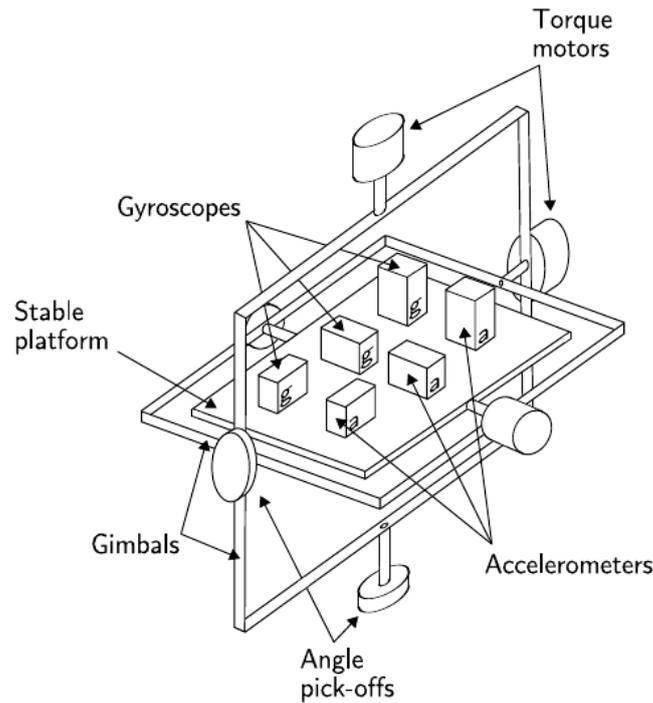


FIGURA 2.1: Sistema de plataforma estabilizada.

Sistemas de navegación con sensores fijos a la estructura del vehículo

Gracias a los avances tecnológicos se ha reducido la complejidad mecánica de los sistemas de plataforma estabilizada abriéndole paso a los sistemas de navegación con sensores fijos a la estructura del vehículo, siendo estos sistemas los que dominan las aplicaciones modernas.

Estos sistemas son conocidos también como sistemas *strapdown*, en estos las mediciones de los sensores se hacen en el marco del cuerpo con respecto al marco de referencia inercial por lo que se requieren hacer cálculos adicionales para seguir el movimiento del vehículo en comparación con los sistemas con plataforma estable.

2.2. Marcos de referencia para la navegación inercial

Es vital definir un conjunto de marcos de referencia antes de poder describir el movimiento relativo del vehículo.

Marco de referencia inercial - $\{i\}$: es un sistema de coordenadas donde se cumplen las leyes del movimiento de Newton. Su origen se localiza en el centro de masa de la Tierra y es estacionario con respecto a las estrellas. Las mediciones de los sensores inerciales se efectúan respecto a este marco de referencia.

Marco de referencia de la Tierra - $\{e\}$: al igual que el marco de referencia inercial tiene su origen en el centro de la Tierra y rota con esta alrededor del eje-z respecto al marco de referencia inercial. El eje-x pasa sobre la intersección del meridiano de *Greenwich* y el plano del Ecuador, el eje-z coincide con el eje polar, el eje-y es ortogonal al eje-x hacia el este.

Marco de referencia del cuerpo - $\{b\}$: también llamado marco de referencia del vehículo, ya que la IMU es alineada a la carcasa del este. Sus ejes corresponde los ejes de IMU con el eje-x apuntando hacia la dirección de avance del vehículo, el eje-y apuntando hacia la izquierda y el eje-z apuntando hacia arriba, conformado un sistema ortogonal.

Marco referencia de navegación - $\{n\}$: es el marco de referencia del área geográfica local en el cual la navegación tiene lugar, cuando el vehículo recorre una gran distancia hay que tener en cuenta el movimiento este marco de referencia sobre la superficie terrestre.

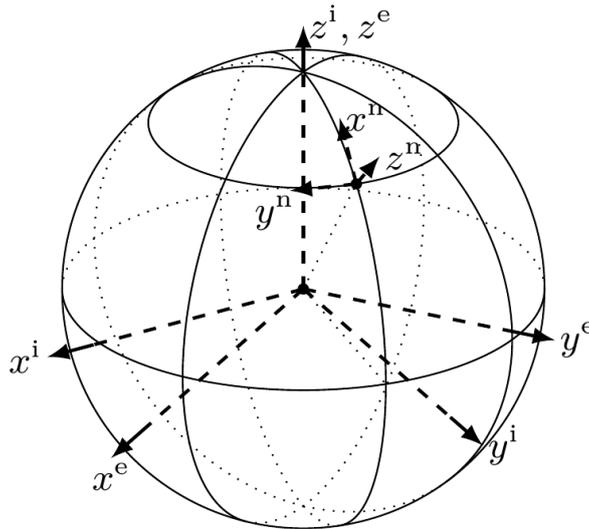


FIGURA 2.2: Marcos de referencia (Kok et al., 2017).

2.3. Representación de la orientación

La orientación de un marco de referencia A respecto a un marco de referencia B queda definida por una matriz R que pertenece a un grupo ortogonal especial (SO), en un espacio tridimensional este grupo se define como:

$$SO(3) = \{R \in \mathbb{R}^{3 \times 3} | R^T R = I, \det(R) = +1\}$$

A $SO(3)$ se le conoce como matrices de rotación. La razón es que al multiplicar un vector por una matriz de rotación cambia su dirección, pero preserva su magnitud.

Una rotación alrededor del eje-x por un ϕ se representa con la matriz de rotación:

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix}$$

De forma similar las matrices de rotación alrededor del eje-y y eje-z se representan como:

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$$

$$R_z(\psi) = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Toda rotación se puede expresar como el producto de tres rotaciones consecutivas sobre el eje-x, eje-y y eje-z. En general, el producto entre matrices no es conmutativo. Por lo tanto, es importante definir el orden en el que se aplican las rotaciones.

$$R = R_x(\phi)R_y(\theta)R_z(\psi).$$

Si conocemos la configuración (R_b^n) del marco-n respecto al marco-b, podemos transformar las coordenadas de un vector x^b expresadas en el marco-b a sus coordenadas x^n expresadas en el marco-n:

$$x^n = R_b^n x^b,$$

ya que R_b^n es una matriz de rotación la transformación inversa es su traspuesta:

$$x^b = (R_b^n)^{-1} x^n = (R_b^n)^T x^n = R_n^b x^n.$$

De manera general, para indicar en que marco de referencia está expresado un vector se utiliza un superíndice. Por ejemplo, x^i indica que el vector x se encuentra expresado en el marco de referencia inercial. Un subíndice y superíndice indica desde cual marco de referencia y hacia cual marco de referencia está definida la matriz de rotación. La notación R_b^n indica que la matriz R transforma las coordenadas de un vector expresado en el marco-b a sus coordenadas expresadas en el marco-n.

2.4. Sensores inerciales

Los acelerómetros y los giroscopios, conjuntamente conocidos como sensores inerciales, son los componentes sustanciales de los INS. Los principales factores que afectan la estabilidad del sistema son los errores presentes en las señales de salida de los sensores y su rendimiento depende de los mecanismos empleados reducirlos.

2.4.1. Principales fuentes de error

Son varios los factores que contribuyen en mayor o menor grado a la inestabilidad de un INS, por muy pequeño que este sea debería ser removido evitando así la degeneración de la solución. Entre los errores más relevantes podemos encontrar el sesgo, el factor de escala, el desalineamiento y el ruido aleatorio, todos ellos ligados con mediciones de los sensores y mediante la adecuada calibración de estos se puede reducir su impacto.

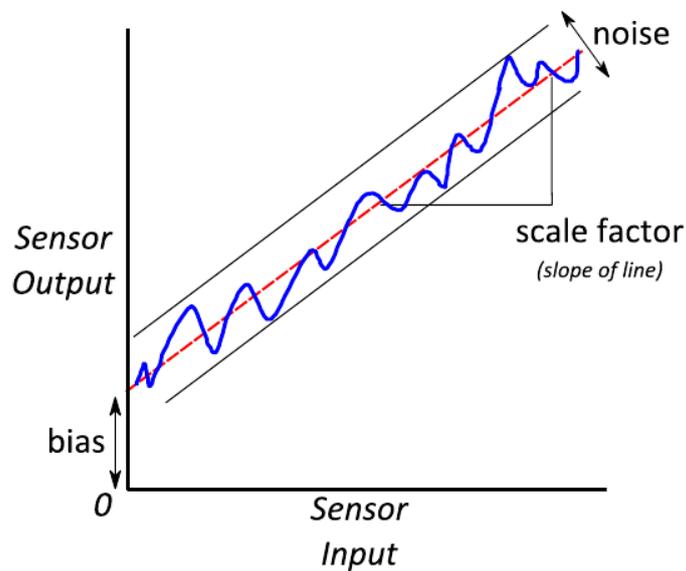


FIGURA 2.3: Errores comunes en las mediciones de los sensores.

Sesgo

Al medir una magnitud física con un sensor se espera que este entregue a la salida el valor real, sin embargo, existe una diferencia entre el valor observado y el valor real, a esta diferencia se le conoce como sesgo. El sesgo es independiente de la entrada, por lo que una estimación válida sería el promedio de las mediciones cuando el dispositivo se encuentra en estado de reposo.

La estabilidad del sesgo está estrechamente relacionado con la temperatura, el tiempo y el estrés mecánico del sistema. La pérdida de estabilidad afecta directamente el rendimiento de la salida, de manera que importante que el INS estime constantemente el sesgo y una vez conocido es necesario eliminarlo de las mediciones para compensar.

Factor de escala

El factor de escala es la relación que existe entre la salida del sensor y la cantidad medida. Idealmente (error de escala del 0%), si la entrada es del 100%, se espera una salida del 100%. El factor de escala esta compuesto por una parte lineal y otra no lineal, sin embargo, la última suele ser bastante pequeña y no tiene un impacto importante en la salida. Es común que en las especificaciones del dispositivo el factor de escala combine ambas partes en un solo valor. En términos matriciales los factores de escala se expresan como una matriz diagonal.

Desalineamiento

En un sensor ideal sus tres ejes son montados mutuamente perpendiculares, sin embargo, durante el proceso de fabricación esto no ocurre a la perfección provocando una correlación entre las mediciones. En condiciones ideales, un acelerómetro en reposo que esta perfectamente nivelado con un eje apuntando hacia arriba, el eje vertical debería medir la gravedad y los otros ejes no deberían medir ningún efecto de ella. Bajo los mismos supuestos, pero considerando que los ejes no son perfectamente ortogonales entre sí, la gravedad se proyecta en los tres ejes, provocando una correlación entre las mediciones. Así el error que se originan por la no ortogonalidad de los ejes del sensor tiene que ser compensado con una serie de factores que afectan la relación entre las medidas de los diferentes ejes. Estos factores se expresan en forma matricial donde los elementos de la diagonal son todos cero.

Ruido aleatorio

El ruido aleatorio $\epsilon(t)$ es un error inherente de las mediciones de la IMU que se describe como un proceso estocástico. Este se define como una secuencia de variables aleatorias independientes e idénticamente distribuidas con media cero y varianza σ^2 :

$$\mathbf{N}_i \sim \mathcal{N}(0, \sigma^2), \quad i.i.d. \quad (2.1)$$

Su impacto se pueden conocer ejecutando un procedimiento análogo al de la navegación inercial, pero únicamente operando sobre el ruido aleatorio (Woodman, 2007).

En la estimación de la orientación el ruido $\epsilon(t)$ es integrado con respecto al tiempo, esto se traduce en:

$$\int_0^t \epsilon(\tau) d\tau = \Delta t \sum_{i=1}^n \mathbf{N}_i, \quad (2.2)$$

donde n es la cantidad de muestras del sensor, Δt es el intervalo de tiempo entre dos muestras consecutivas y $t = n \cdot \Delta t$.

Aplicando las propiedades de la media y varianza de una combinación lineal de variables aleatorias se puede demostrar que el valor esperado y la varianza de (??) son:

$$E \left(\int_0^t \epsilon(\tau) d\tau \right) = 0, \quad Var \left(\int_0^t \epsilon(\tau) d\tau \right) = \Delta t \cdot t \cdot \sigma^2.$$

Estos resultados indican que el ruido aleatorio presente en la señal de giroscopio introduce un error en la estimación de la orientación que crece proporcionalmente a la raíz cuadrada del tiempo, \sqrt{t} .

Usando la misma idea, para estimar la posición el error $\epsilon(t)$ se integra dos veces con respecto al tiempo:

$$\begin{aligned} \int_0^t \int_0^t \epsilon(\tau) d\tau d\tau &= \Delta t \sum_{i=1}^n \mathbf{N}_i \Delta t \sum_{j=1}^i \mathbf{N}_j \\ &= \Delta t^2 \sum_{i=1}^n (n - i + 1) \mathbf{N}_i, \end{aligned}$$

el error esperado es:

$$E \left(\int_0^t \int_0^t \epsilon(\tau) d\tau d\tau \right) = 0,$$

suponiendo que Δt es pequeño se puede aproximar la varianza como:

$$Var \left(\int_0^t \int_0^t \epsilon(\tau) d\tau d\tau \right) \approx \frac{1}{3} \Delta t \cdot t^3 \cdot \sigma^2.$$

Estos resultados muestran que el ruido aleatorio presente en las mediciones del acelerómetro introduce un error en la estimación de la posición que crece proporcionalmente a la raíz cuadrada del tiempo elevado al cubo, $\sqrt{t^3}$.

2.4.2. Giroscopio

Un giroscopio es un sensor de rotación que mide velocidad angular. Existen diversas categorías en las que se puede clasificar este sensor, aquí se describen algunas, pero en trabajo de Titterton et al. (2004) se puede encontrar una lista bastante completa.

Mecánico

Un giroscopio mecánico está compuesto esencialmente por un rotor, cardanes y dispositivos de medición como se muestran en la figura 2.4.

- Un rotor es un elemento inercial que gira a altas velocidades.
- Un cardán es un sistema mecánico que permite unir dos ejes no coaxiales, sirve de soporte al rotor u otros cardanes con el fin de aislar al rotor de movimientos rotacionales.
- El dispositivo de medición mide los ángulos relativos entre el rotor y los cardanes.

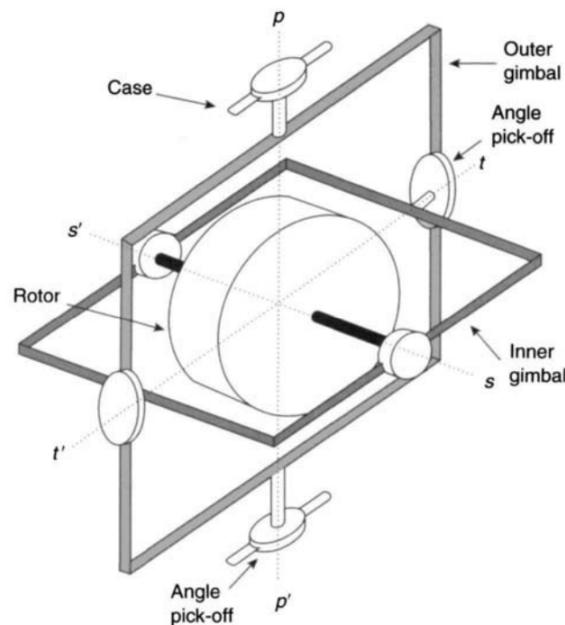


FIGURA 2.4: Diagrama esquemático de un giroscopio mecánico. Fuente: Titterton et al. (2004).

El modo de operación de un giroscopio mecánico se basa en la ley de conservación del momento angular. Este principio afirma que el momento angular permanece constante durante la evolución temporal de un sistema aislado. Así, los giroscopios mecánicos son diseñados como un cuerpo de simetría de rotación que usan cardanes para aislar el rotor de cualquier movimiento rotacional, es decir, mantener el momento angular constante.

El elemento inercial gira a velocidades altas generando un vector de momento angular en la dirección del eje de giro. El diseño del giroscopio le permite al rotor resistirse a cualquier cambio en su vector de momento angular relativo al marco de referencia inercial. Finalmente, el dispositivo de medición produce una señal eléctrica que indica el movimiento de rotación respecto a una referencia (Titterton et al., 2004).

Vibratorio

El efecto de Coriolis es el resultado que se observa cuando un cuerpo se encuentra en movimiento respecto a un sistema de referencia en rotación, descrito en 1836 por Gaspard Gustav Coriolis. Este principio asegura que todo cuerpo que se desplaza sobre un sistema de referencia rotatorio sufre una aceleración producida por una fuerza perpendicular a su movimiento y al eje de rotación del sistema, y que hace que todo cuerpo en movimiento se desvíe de su trayectoria. Esta fuerza es conocida como fuerza de Coriolis, definida por:

$$\mathbf{F}_c = -2m(\boldsymbol{\omega} \times \mathbf{v}).$$

Aprovechando este principio un giroscopio vibratorio es diseñado con partes vibrantes para generar una velocidad lineal oscilatoria. Cuando el dispositivo rota alrededor del eje perpendicular a esa velocidad se induce una fuerza de Coriolis que afectara el movimiento de los elementos vibrantes. A partir de los cambios detectados en el movimiento de los elementos vibrante se calcula la velocidad angular. Existen numerosas configuraciones en la geometría de los elementos vibrantes, un ejemplo es el giroscopio de diapasón que se muestra en la figura 2.5.

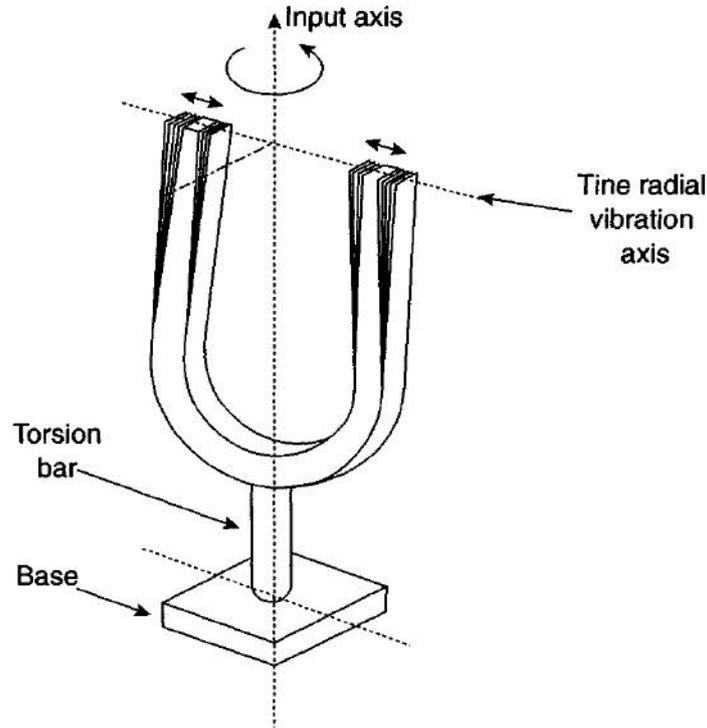


FIGURA 2.5: Principio de operación de un giroscopio de diapasón. Fuente: Titterton et al. (2004).

MEMS

El desarrollo de los sensores mecánicos han tenido un progreso significativo, sin embargo, el costo de estos se ha mantenido elevado. Esto se debe principalmente a la numerosa cantidad de piezas de alta fidelidad y tolerancia que exigen técnicas de montaje de alta precisión, caracterización y calibración.

Los MEMS (*Micromachined Electromechanical Systems*) se caracterizan por ser de bajo costo, alta confiabilidad, bajo consumo de potencia, corto periodo de inicialización, y robustez. Estas características son muy deseables para operar en entornos hostiles y son demandadas por las aplicaciones modernas. Usan el mismo principio de operación que los giroscopios vibrantes, es decir, utilizan el efecto Coriolis en una masa de prueba vibrante para detectar la velocidad angular inercial.

Modelo de medición

Para el propósito de navegación estamos interesados en medir la velocidad angular del vehículo respecto al sistema de referencia de navegación, denotada con el símbolo ω^b .

De forma general, la velocidad angular respecto al sistema de referencia inercial que nos provee el giroscopio se modela considerando los efectos de la rotación de la Tierra respecto al sistema de referencia inercial, denotada con símbolo ω_{ie}^n ; y la velocidad de transporte del sistema de referencia de navegación, denotada con el símbolo ω_{en}^n . Así, la señal de salida del giroscopio se modela como (Kok et al., 2017):

$$\omega_{ib}^b = {}^b_n\mathbf{R}(\omega_{ie}^n + \omega_{en}^n) + \omega_{nb}^b,$$

en el cual ${}^b_n\mathbf{R}$ es la matriz de rotación del sistema de referencia de navegación al sistema de referencia del cuerpo.

La Tierra tarda aproximadamente veinticuatro horas en completar una vuelta alrededor del eje-z, por lo que el término ω_{ie} se puede despreciar, además si se asume que sistema de referencia de navegación es estacionario, $\omega_{en} = 0$, por lo tanto, el modelo de la salida del giroscopio se puede simplificar a:

$$\omega^b(t) = \omega_{nb}^b + \delta_\omega^b + \mathbf{e}_\omega^b, \quad (2.3)$$

donde las variables δ_ω y $\mathbf{e}_\omega \sim \mathcal{N}(0, \Sigma_\omega)$ representan el sesgo y el ruido aleatorio en las mediciones del sensor, respectivamente.

2.4.3. Acelerómetro

En el trabajo de Titterton et al. (2004) se encuentra un estudio completo de la clasificación de los acelerómetros. Dependiendo de la forma de su construcción se pueden clasificar en alguna de dos categorías; mecánico o de estado sólido.

Mecánico

Un acelerómetro mecánico consiste en una masa de prueba suspendida en un resorte como se muestra en la figura 2.6. Su principio de funcionamiento se basa en la segunda ley de Newton:

$$\mathbf{F} = m \cdot \mathbf{a}.$$

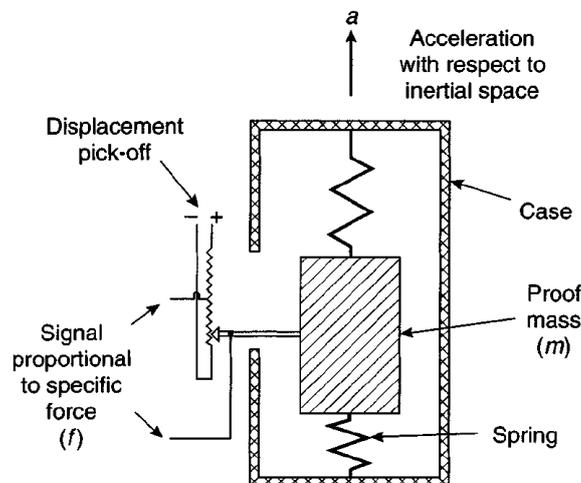


FIGURA 2.6: Acelerómetro mecánico. Fuente: Titterton et al. (2004).

Cuando una fuerza actúa sobre la masa a lo largo de su eje sensible esta tiende a resistirse al movimiento debido a su propia inercia, por lo que el dispositivo de medición mide la deflexión del resorte y entrega una salida proporcional a la fuerza específica aplicada a la masa.

Estado sólido

Los dispositivos de estado sólido se caracterizan por ser pequeños, robustos y confiables. En esta categoría podemos encontrar los dispositivos de superficie de onda acústica, vibratorios, silicio y cuarzo.

Una acelerómetro de superficie de onda acústica consta de un viga que tiene un extremo fijo a la estructura del dispositivo y el otro extremo puede moverse libremente, el extremo libre tiene acoplada una masa de prueba como se observa en la figura 2.7.

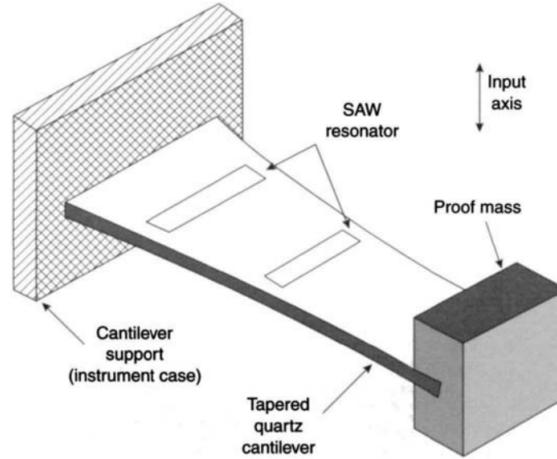


FIGURA 2.7: Acelerómetro de superficie de onda acústica. Fuente: Titterton et al. (2004).

El funcionamiento se basa en hacer resonar la viga por medio de un tren de pulsos superficiales a una frecuencia conocida. Cuando actúa una aceleración al plano perpendicular de la viga hace que ésta se deforme provocando que la frecuencia de la superficie de onda acústica cambie. A partir de los cambios en la frecuencia se calcula la aceleración aplicada al eje sensible.

MEMS

Existen dos clases principales de dispositivos MEMS que basan su operación en los dos tipos descritos anteriormente. La primera clase basa su funcionamiento en dispositivos mecánicos, es decir, miden la aceleración con base al desplazamiento de una masa de prueba. La segunda clase basa su operación en los dispositivos de estado sólido, de forma similar al acelerómetro de superficie de onda acústica, en este caso la aceleración se calcula utilizando los cambios en la frecuencia de elementos vibrantes causados por los cambios en la tensión cuando los elementos están sujetos a una aceleración.

Como sucede con giroscopios homólogos a esta categoría, estos dispositivos se caracterizan por ser de bajo costo, peso ligero, pequeños, robustos, bajo consumo de potencia y rápida inicialización. Estos dispositivos han ganado popularidad en los últimos años debido a las constantes mejoras en cuanto a precisión y otras características favorables que demandan las aplicaciones modernas.

Modelo de medición

El acelerómetro mide la fuerza específica externa que actúa en el sensor. La fuerza específica consiste tanto en la aceleración del sensor como en la gravedad de la Tierra, su valor se define como (Kok et al., 2017):

$$\mathbf{f}^b = {}^b_n\mathbf{R}(\mathbf{a}_{ii}^n - \mathbf{g}^n),$$

donde \mathbf{g}^n es el vector de gravedad y \mathbf{a}_{ii}^n representa la aceleración lineal del sensor expresada en el sistema de referencia de navegación.

Asumiendo que el sistema de referencia de navegación es fijo, se puede demostrar que la aceleración lineal del sensor es igual a:

$$\mathbf{a}_{ii}^n = \mathbf{a}_{nn}^n + \underbrace{2\omega_{ie}^n \times \mathbf{v}_n^n}_{\text{aceleración de Coriolis}} + \underbrace{\omega_{ie}^n \times \omega_{ie}^n \times \mathbf{p}^n}_{\text{aceleración centrífuga}},$$

donde \mathbf{a}_{nn}^n es la aceleración lineal que es de interés en la estimación de la posición, \mathbf{p}^n es la posición, los términos $(\omega_{ie}^n \times \omega_{ie}^n \times \mathbf{p}^n)$ y $(2\omega_{ie}^n \times \mathbf{v}_n^n)$ se conocen como aceleración centrífuga y aceleración de Coriolis, respectivamente.

La aceleración centrífuga normalmente la absorbe en el vector de gravedad y la magnitud de la aceleración de *Coriolis* es pequeña comparada con las mediciones del acelerómetro, despreciando estos términos y tomando en cuenta los errores que afectan la precisión del dispositivo, el modelo de medición se define como:

$$\mathbf{a}^b(t) = {}^b_n\mathbf{R}(\mathbf{a}_{nn}^n - \mathbf{g}^n) + \delta_a^b + \mathbf{e}_a^b, \quad (2.4)$$

donde $\mathbf{e}_a \sim \mathcal{N}(0, \Sigma_a)$ representa el ruido presente en las mediciones del acelerómetro y δ_a representa el sesgo. Para un estudio completo se recomienda revisar Kok et al. (2017).

2.4.4. Magnetómetro

El campo magnético de la tierra es parecido a una pequeña barra de imán localizada en el centro de la tierra. El eje a lo largo de la barra imantada se le conoce como eje magnético, atraviesa la superficie de la Tierra en dos puntos conocidos como polos magnéticos. Los polos magnéticos se localizan cerca de los polos geográficos de la Tierra como se observa en la figura 2.8.

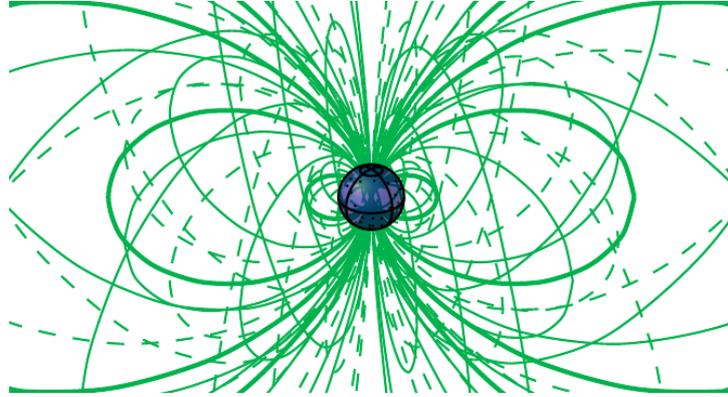


FIGURA 2.8: Líneas del campo magnético de la Tierra. Fuente: Kok et al. (2017).

El campo magnético de la Tierra se puede representar en términos de tres componentes rectangulares en coordenadas geográficas:

$$\mathbf{m} = [m_x, m_y, m_z]^T,$$

donde m_x es la componente horizontal hacia el Norte, m_y es la componente horizontal hacia el Este y m_z es la componente hacia abajo.

El campo magnético también se puede expresar en función de dos ángulos denominados declinación magnética e inclinación magnética y la magnitud del campo magnético. La declinación o variación magnética, se refiere al ángulo formado entre el Norte geográfico y el Norte magnético. La inclinación es el ángulo formado entre el plano horizontal y la dirección del campo magnético local. La relación entre estos parámetros se establece por medio de la siguiente ecuación:

$$\mathbf{m} = M [\cos \delta \cos \gamma \quad \cos \delta \sin \gamma \quad \sin \delta]^T, \quad (2.5)$$

donde γ es la declinación magnética, δ es la inclinación magnética y M es la intensidad del campo magnético local.

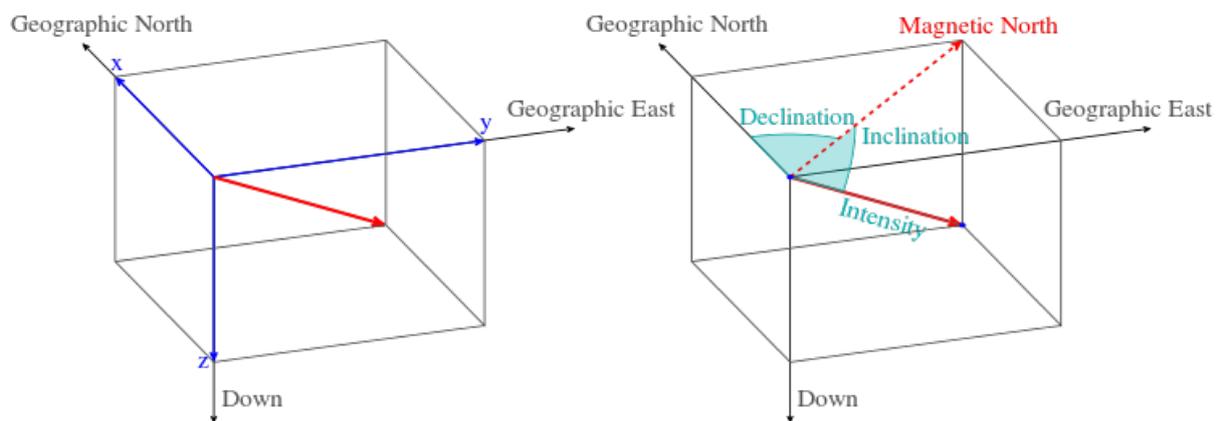


FIGURA 2.9: Representaciones del campo magnético de la Tierra. Fuente: Wikipedia.

La componente horizontal del campo magnético local siempre apunta al norte magnético, por esta razón es posible utilizar un magnetómetro para obtener una referencia global en aplicaciones de navegación sobre la superficie de la tierra.

MEMS

Un magnetómetro mide la intensidad y orientación del campo magnético de la Tierra. Este no es un sensor inercial, no obstante, es común encontrarlo empaquetado junto con sensores inerciales.

La mayoría de los magnetómetros MEMS operan detectando la fuerza de *Lorentz*. Esta fuerza se origina de la combinación de una fuerza eléctrica y otra magnética en una carga puntual debida a campos electromagnéticos, se expresa de acuerdo a la siguiente relación:

$$\mathbf{F}_L = q\mathbf{E} + q\mathbf{v} \times \mathbf{B},$$

donde las variables q , \mathbf{E} , \mathbf{v} , \mathbf{B} son: la carga del electrón, el campo eléctrico, la velocidad del electrón y el campo magnético, respectivamente.

Su funcionamiento se basa en la siguiente idea: las cargas dentro de una pieza conductora sometida a una fuente de voltaje constante estarán uniformemente distribuidas a lo largo del conductor, al colocar la pieza en un campo magnético se induce una la fuerza de *Lorentz*, \mathbf{F}_L . Esta fuerza provoca un cambio en la distribución de las cargas haciendo que cambie el voltaje a través del conductor. A partir de esos cambios se calcula el campo magnético, \mathbf{B} .

Esta tipo de dispositivos son pequeños y de bajo costo, tienen las mismas ventajas que los sensores inerciales de este tipo.

Modelo de medición

El campo magnético en un punto particular sobre la superficie de la Tierra sufre variaciones debido al material magnético presente en la vecindad de dicho punto, por lo tanto, el magnetómetro mide el campo magnético local de la Tierra y el campo magnético debido a la presencia de material magnético en la zona. Es importante calibrar este dispositivo en el entorno de trabajo para tener una buena estimación de la orientación. Por ejemplo, en caso de operar en interiores el material magnético presente en la estructura de las construcciones, electrónicos u otros se convertirán en una fuente de error si el dispositivo no se calibra adecuadamente.

En este trabajo, se utiliza un modelo que se deriva de la representación del campo magnético en función del ángulo de declinación y del ángulo de inclinación. Desafortunadamente, la declinación varía con el tiempo y del punto de observación de las mediciones, aunque es posible compensar ambos efectos en este trabajo se utiliza un modelo simplificado que depende solo del ángulo de inclinación. Este modelo proporciona la orientación respecto al norte magnético en lugar del norte geográfico. Para calcular la orientación solo es de interés la dirección del campo magnético, la magnitud es irrelevante, así se puede expresar el modelo del campo magnético de la Tierra como:

$$\mathbf{m} = [\cos \delta \quad 0 \quad \sin \delta]^\top, \quad \|\mathbf{m}\|_2 = 1. \quad (2.6)$$

Considerando los errores que afectan la precisión las mediciones del sensor, el modelo de medición del magnetómetro se puede definir como:

$$\mathbf{m}(t) = \frac{b}{n}\mathbf{R}\mathbf{m} + \delta_m^b + \mathbf{e}_m^b, \quad (2.7)$$

donde δ_m es el sesgo y $\mathbf{e}_m \sim \mathcal{N}(0, \Sigma_m)$ representa el ruido.

Para un estudio más completo se recomienda consultar Kok et al. (2017) y Titterton et al. (2004).

2.5. Orientación entre dos sistemas de referencia

Dado N vectores linealmente independientes observados en dos sistemas de referencia coordinados diferentes donde uno de ellos es conocido, entonces se puede determinar la rotación \mathbf{R} entre ambos sistemas de coordenadas.

En general, si $N \geq 2$, el sistema de ecuaciones está sobredeterminado, por lo tanto, no se cumple para los N vectores. Una forma de solucionar este problema es aplicando técnicas de optimización para estimar la rotación \mathbf{R} que minimice el error de proyección de los N vectores mediante la siguiente función de pérdida:

$$J(\mathbf{R}) = \frac{1}{2} \sum_{i=1}^N w_k |\mathbf{u}_i - \mathbf{R}\mathbf{v}_i|^2 \quad (2.8)$$

Este problema fue propuesto por Grace Wahba en 1965, en matemáticas aplicadas se le conoce como el problema de Wahba (Wahba, 1965).

Método-q

El método-q proporciona una solución exacta, en forma de cuaternión, al problema de Wahba. Esta técnica ajusta el indicador de rendimiento del problema en términos de un cuaternión unitario. Los extremos del nuevo indicador de rendimiento son los valores propios de una matriz comúnmente denotada con el símbolo \mathbf{K} , y la orientación óptima es el vector propio asociado al valor propio más grande Shuster y Oh (1981).

El procedimiento del método-q se puede resumir en las siguientes ideas esenciales: Se desarrolla la función de pérdida (??), y asumiendo que los vectores son unitarios se simplifica a:

$$J(\mathbf{R}) = \frac{1}{2} \sum_{i=1}^N w_k (1 - \mathbf{u}_i^T \mathbf{R} \mathbf{v}_i).$$

Se intercambia el problema de minimización a uno equivalente de maximización, el nuevo problema consiste en maximizar la función de ganancia:

$$G(\mathbf{R}) = \frac{1}{2} \sum_{i=1}^N w_k (\mathbf{u}_i^T \mathbf{R} \mathbf{v}_i).$$

Se reescribe el problema en términos de un cuaternión unitario, utilizando la relación entre una matriz de rotación y un cuaternión:

$$\mathbf{R} = (q_w^2 - \mathbf{q}^T \mathbf{q}) + 2\mathbf{q}\mathbf{q}^T - 2q_w \mathbf{q}^\times.$$

El resultado de dicha sustitución lleva a:

$$g(\mathbf{q}) = \mathbf{q}^T \mathbf{K} \mathbf{q},$$

donde \mathbf{K} es una matriz de 4×4 que se define como:

$$\mathbf{K} = \begin{bmatrix} \mathbf{S} - \sigma \mathbf{I} & \mathbf{Z} \\ \mathbf{Z}^T & \sigma \end{bmatrix},$$

$$\begin{aligned} \mathbf{B} &= \sum_{i=1}^N w_i (\mathbf{u}_i \mathbf{v}_i^T), \\ \mathbf{S} &= \mathbf{B} + \mathbf{B}^T, \\ \mathbf{Z} &= [B_{23} - B_{32} \quad B_{31} - B_{13} \quad B_{12} - B_{21}]^T, \\ \sigma &= \text{tr}(\mathbf{B}). \end{aligned}$$

El valor propio más grande de la matriz \mathbf{K} maximiza la función de ganancia $g(\mathbf{q})$ y el vector propio asociado a este valor propio es la estimación óptima de la orientación. Los detalles se pueden consultar en Shuster y Oh (1981).

2.6. Seguimiento de la orientación

Una rotación en un espacio tridimensional se puede definir mediante un vector $\omega = \theta \mathbf{u}$, en el cual el \mathbf{u} es un vector unitario que indica la dirección del eje de giro y θ el ángulo de rotación alrededor del eje. El cuaternión unitario equivalente a este vector de rotación se define como:

$$\mathbf{q} = \left[\cos\left(\frac{\theta}{2}\right), \mathbf{u} \sin\left(\frac{\theta}{2}\right) \right]^T. \quad (2.9)$$

2.6.1. Tasa de cambio de la orientación

Sean \mathbf{q} la orientación del vehículo hasta el tiempo t y \mathbf{p} es el responsable de los cambios en \mathbf{q} en el intervalo de tiempo Δt entonces la evolución de la orientación del vehículo hasta el tiempo $t + \Delta t$ se puede expresar como:

$$\mathbf{q}(t + \Delta t) = \mathbf{p} \odot \mathbf{q}, \quad (2.10)$$

con \mathbf{p} definido por:

$$\mathbf{p} = \left[\cos\left(\frac{\Delta\theta}{2}\right), \mathbf{u} \sin\left(\frac{\Delta\theta}{2}\right) \right] \quad (2.11)$$

Por definición, la derivada la orientación \mathbf{q} se expresa como:

$$\frac{d\mathbf{q}}{dt} = \lim_{\Delta t \rightarrow 0} \left[\frac{\mathbf{q}(t + \Delta t) - \mathbf{q}(t)}{\Delta t} \right] = \lim_{\Delta t \rightarrow 0} \left[\frac{\mathbf{p} \odot \mathbf{q}(t) - \mathbf{q}(t)}{\Delta t} \right], \quad (2.12)$$

factorizando el término \mathbf{q} en la ecuación (??) se llega a:

$$\frac{d\mathbf{q}}{dt} = \lim_{\Delta t \rightarrow 0} \left[\frac{\mathbf{p} - \mathbf{i}}{\Delta t} \right] \mathbf{q}(t) = \frac{d\mathbf{p}}{dt} \mathbf{q}(t). \quad (2.13)$$

Ahora, sustituyendo la ecuación (??) en la ecuación (??) se obtiene que:

$$\frac{d\mathbf{p}}{dt} = \lim_{\Delta t \rightarrow 0} \left[\frac{\left[\cos\frac{\Delta\theta}{2}, \mathbf{u} \sin\frac{\Delta\theta}{2} \right]^T - \mathbf{i}}{\Delta t} \right], \quad (2.14)$$

después de expandir el término coseno con series de Taylor, los términos $\Delta\theta^2 \rightarrow 0$ y $\sin\frac{\Delta\theta}{2} \rightarrow \frac{\Delta\theta}{2}$ cuando $\Delta t \rightarrow 0$, así:

$$\begin{aligned} \frac{d\mathbf{p}}{dt} &= \lim_{\Delta t \rightarrow 0} \left[\frac{\left[\left(1 - \frac{\Delta\theta^2}{4} + \dots\right), \mathbf{u} \sin\frac{\Delta\theta}{2} \right]^T - \mathbf{i}}{\Delta t} \right], \\ &= \lim_{\Delta t \rightarrow 0} \left[\frac{\left[0, \mathbf{u} \frac{\Delta\theta}{2} \right]^T}{\Delta t} \right], \\ &= \frac{1}{2} \lim_{\Delta t \rightarrow 0} \left[\frac{\Delta\theta}{\Delta t} \right] \begin{bmatrix} 0 \\ \mathbf{u} \end{bmatrix}. \end{aligned} \quad (2.15)$$

Por otra parte, por definición la velocidad angular es:

$$\omega^n = \lim_{\Delta t \rightarrow 0} \left[\frac{\Delta\theta}{\Delta t} \right] \mathbf{u}.$$

En consecuencia, la tasa cambio de la orientación $d\mathbf{q}$ se puede expresar en términos de la la velocidad angular como:

$$\frac{d\mathbf{q}}{dt} = \frac{1}{2} \omega_q^n(t) \odot \mathbf{q}, \quad (2.16)$$

donde $\omega_q^n = [0, \omega^n]^T$.

El giroscopio mide la velocidad angular expresada en el sistema de referencia del cuerpo ω^b , por lo tanto, es necesario proyectar la velocidad angular obtenida del giroscopio al sistema de referencia de navegación, esto se logra usando la orientación actual \mathbf{q} :

$$\omega^n = \mathbf{q} \odot \omega_q^b \odot \mathbf{q}^{-1}. \quad (2.17)$$

Finalmente, sustituyendo la ecuación (??) en la ecuación (??) se concluye que la razón de cambio de la orientación queda definida por:

$$\frac{d\mathbf{q}(t)}{dt} = \frac{1}{2} \mathbf{q} \odot \omega_q^b(t). \quad (2.18)$$

Los detalles del la demostración se encuentran en Kelly (2013).

2.6.2. Actualización de la orientación

La orientación del vehículo relativo al sistema de referencia de navegación se obtiene resolviendo la ecuación diferencial (??):

$$\mathbf{q}(t) = \frac{1}{2} \int_0^t \mathbf{q}(t) \odot \omega_q^b(t) dt \quad (2.19)$$

El producto entre cuaterniones se puede expresar como una multiplicación de matrices, esto permite escribir la ecuación (??) de forma equivalente como:

$$\mathbf{q}_{k+1} = \frac{1}{2} \int_{t_k}^{t_{k+1}} \times [\omega_q^b] dt \cdot \mathbf{q}_k \quad (2.20)$$

con

$$\times [\omega_q] = \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix}.$$

Usando la igualdad

$$\int_0^t \mathbf{A} d\tau = \exp(\mathbf{A}t),$$

se llega a la siguiente solución:

$$\mathbf{q}_{k+1} = \exp\left(\frac{1}{2} \times [\omega] \Delta t\right) \cdot \mathbf{q}_k, \quad (2.21)$$

donde Δt es el intervalo de tiempo comprendido entre t_k y t_{k+1} .

Una matriz antisimétrica tiene una matriz exponencial de forma cerrada, lo que permite escribir la solución como:

$$\mathbf{q}_{k+1} = \left[\cos \frac{\|\omega^b \Delta t\|}{2} I + \sin \frac{\|\omega^b \Delta t\|}{2} \left(\frac{\times [\omega^b]}{\|\omega^b\|} \right) \right] \cdot \mathbf{q}_k, \quad (2.22)$$

o equivalentemente en términos de cuaterniones:

$$\mathbf{q}_{t+1} = \mathbf{q}_t \odot \exp_q\left(\frac{1}{2} \omega^b \Delta t\right), \quad (2.23)$$

donde $\exp_q(\sigma)$ se define por:

$$\exp_q(\sigma) = \left[\cos |\sigma|, \frac{\sigma}{|\sigma|} \cdot \sin |\sigma| \right]^T. \quad (2.24)$$

Los detalles se pueden consultar en Titterton et al. (2004); Kelly (2013).

2.7. Seguimiento de la posición

La evolución temporal del vehículo se rige por las ecuaciones del movimiento. Dependiendo de las exigencias del tipo de aplicación se pueden usar modelos complejos que involucren todas las variables influyan en el movimiento. Normalmente se usan formulaciones matemáticas en función de la velocidad y aceleración en conjunto con las condiciones iniciales, en dicho caso la dinámica del vehículo respecto al sistema de referencia de navegación se expresa como:

$$\mathbf{v}(t) = \mathbf{v}(0) + \int_0^t \mathbf{a}(t) dt, \quad (2.25)$$

$$\mathbf{p}(t) = \mathbf{p}(0) + \int_0^t \mathbf{v}(t) dt, \quad (2.26)$$

donde $\mathbf{a}(t)$ es la aceleración lineal, $\mathbf{v}(0)$ es la velocidad inicial y $\mathbf{p}(0)$ es la posición inicial.

Para obtener la aceleración lineal se proyecta la aceleración $\mathbf{a}^b(t)$ obtenida del acelerómetro al sistema de referencia de navegación entonces se elimina la aceleración debida a la gravedad:

$$\mathbf{a}(t) = {}^n\mathbf{R}_b \mathbf{a}^b(t) - \mathbf{g}^n, \quad (2.27)$$

en el cual ${}^n\mathbf{R}_b$ es la rotación del sistema de referencia del cuerpo al sistema de referencia de navegación y \mathbf{g}^n es el vector de gravedad en el sistema de referencia de navegación.

Capítulo 3

Localización visual

Este capítulo presenta los fundamentos de los sistemas de localización basados en visión. Primero se brinda una breve descripción de las bases de operación de los sistemas basados en visión. En segundo lugar se dan a conocer las transformaciones que rigen el movimiento de un cuerpo rígido. Después se explican los mecanismos involucrados en el proceso de formación de imágenes, junto con la forma general de las proyecciones. Por último, se presentan el algoritmo para la estimación de la pose de un objeto.

3.1. Visión robótica

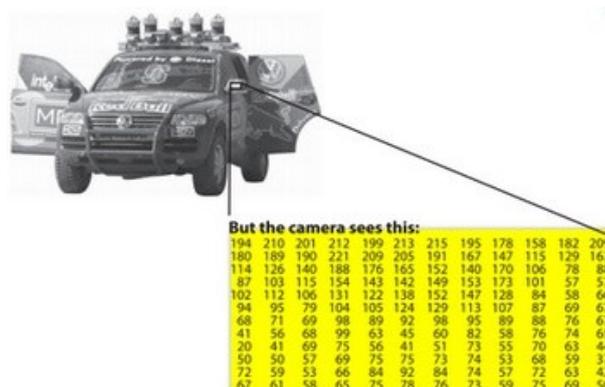


FIGURA 3.1: Representación de una imagen en la computadora (documentación oficial de OpenCV).

La visión robótica es la capacidad con la que cuenta un robot para procesar y analizar una secuencia de imágenes digitales que le permite recuperar información de interés de su entorno. Para este propósito, las cámaras sirven de interfaz entre el mundo y la computadora del robot, estas capturan la luz reflejada por los objetos en imágenes digitales. Una imagen se representa como una matriz de píxeles, donde cada píxel es vector de números que corresponde un color en particular como se muestra en la figura 3.1.

3.2. Localización basada en visión

Los sistemas de localización basados en visión basan su principio de funcionamiento en el reconocimiento de puntos de referencia provistos por las cámaras. Estos puntos de referencia o puntos característicos son identificados en las imágenes y asociados con propiedades del entorno con lo cual el robot puede determinar su localización. Esta procedimiento se puede dividir en cuatro tareas: adquisición y digitalización de imágenes; extracción de características; asociación de características; y cálculo de la localización (DeSouza y Kak, 2002).

Un punto característico identifica una localización en la imagen mediante un descriptor local que puede contener la información de la orientación, el gradiente, el brillo u otra información de los píxeles, dependiendo de la técnica empleada para su obtención. Es importante la selección adecuada de características ya que de esta depende la complejidad de la descripción, extracción y asociación de características.

Para estas tareas son varias las técnicas disponibles, entre ellas podemos encontrar Harris (Harris et al., 1988), SIFT (Lowe et al., 1999), SURF (Bay et al., 2008), BRIEF (Calonder et al., 2010), RANSAC (Fischler y Bolles, 1981).

Un punto de referencia pueden ser catalogado como natural o artificial. Las referencias naturales son los objetos o características que se encuentran en el entorno sin haber sido colocados intencionalmente para servir de apoyo en la navegación del robot, suele ser más dificultosa la extracción y asociación de este tipo de características por lo que debe procurarse un alto contraste relativo al fondo. Las referencias artificiales son los objetos o marcas específicamente diseñados para favorecer la navegación del robot y necesitan ser colocados en el entorno. Su detección es mucho más fácil ya que son diseñados con un alto contraste Borenstein et al. (1997).

3.3. Movimiento de un cuerpo rígido

La forma general para describir el movimiento de un objeto que se mueve enfrente de la cámara debería considerar la trayectoria de cada punto del objeto. Particularmente, cuando se trata de un cuerpo rígido en movimiento esto no es necesario debido a que la distancia entre cualquier par de puntos del objeto permanece constante en el tiempo. Por lo tanto, es suficiente considerar el movimiento de un marco de coordenadas unido a un punto particular, por ejemplo, el centro de masa del objeto.

Para describir el movimiento de un cuerpo rígido se usa una transformación conocida como SE(3) (*Special Euclidean transformation*), la cual preserva la distancia y también preserva la orientación, matemáticamente se expresa como (Ma et al., 2012):

$$g : \mathbb{R}^3 \rightarrow \mathbb{R}^3,$$

1. $\|g(\mathbf{v})\| = \|\mathbf{v}\|, \forall \mathbf{v} \in \mathbb{R}^3,$
2. $g(\mathbf{u}) \times g(\mathbf{v}) = g(\mathbf{u} \times \mathbf{v}), \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^3.$

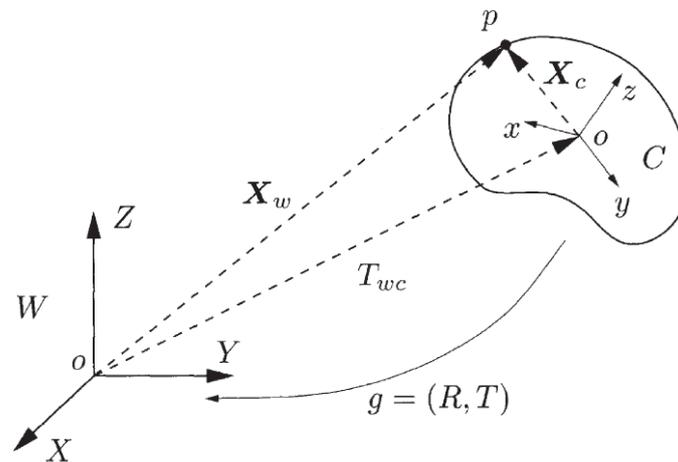


FIGURA 3.2: Movimiento de un cuerpo rígido relativo al marco del mundo W (Ma et al., 2012).

La transformación g se puede expresar en términos de dos parámetros denominados rotación y traslación, colectivamente conocidos como pose. La traslación es el nombre que recibe el cambio de movimiento del objeto entre su posición actual y su nueva posición, denotado por \mathbf{T} . La rotación es el cambio de orientación producida por un giro manteniendo un punto fijo, denotada por \mathbf{R} .

3.4. Formación de imágenes

La transformación un punto $\mathbf{M} = [X, Y, Z]^T$ en el mundo a un punto $\mathbf{m} = [x, y]^T$ en el plano de la imagen se denomina proyección en perspectiva, para establecer la relación entre ambos conjuntos de puntos se requiere de un modelo que capture con exactitud la geometría de la proyección en perspectiva, el modelo pinhole es el más utilizado por su simplicidad.

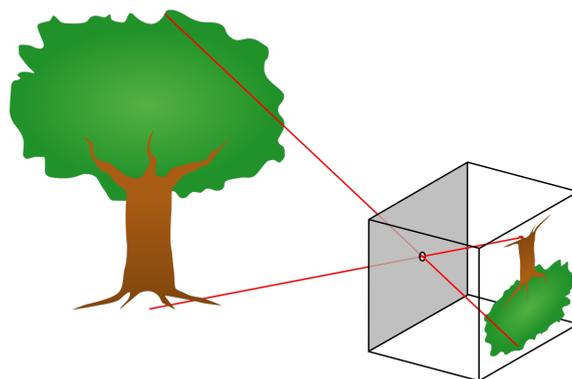


FIGURA 3.3: Esquema de una cámara pinhole.

El modelo de cámara pinhole representa el lente de una cámara ideal, y asume que los rayos de luz que reflejan los objetos viajan en línea recta atravesando un punto infinitesimal hasta el plano de la imagen, la mayoría de las cámaras actuales se pueden representar con bastante exactitud usando este modelo.

Desafortunadamente, los lentes de las cámaras son imperfectos por lo que sufren de algún tipo de distorsión, entre las distorsiones más comunes se encuentran la distorsión barril, distorsión cójín o la combinación de ambos, ilustrados en la figura 3.4. Por fortuna, estas distorsiones se pueden corregir digitalmente haciendo una rectificación previa con que se obtienen imágenes casi perfectas, permitiendo que el modelo de cámara pinhole siga siendo válido (Ojanen, 1999).

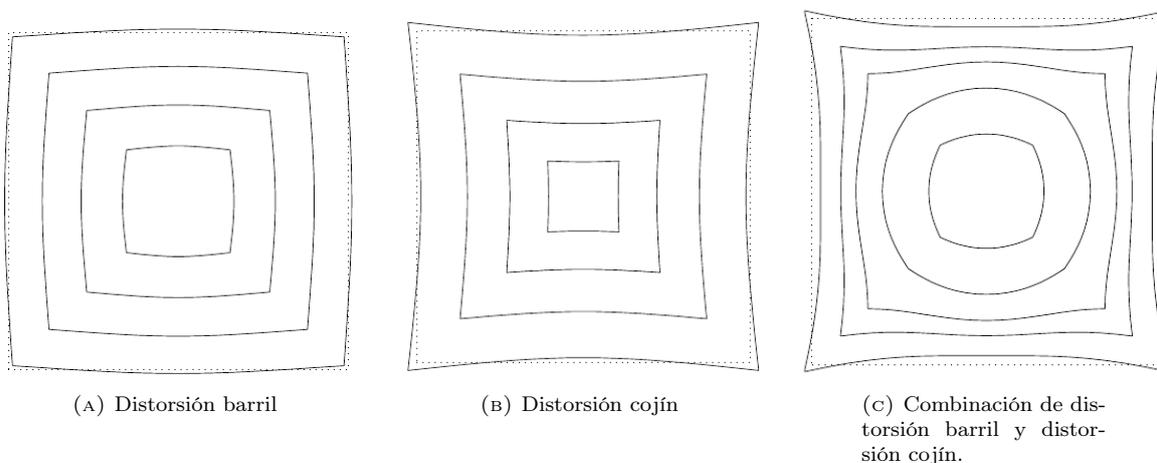


FIGURA 3.4: Tipos de distorsión más comunes que afectan las lentes de las cámaras.

El modelo geométrico de la cámara pinhole se muestra en la figura 3.5, como se puede observar está compuesto de (Xu y Zhang, 2013):

- C , centro óptico de la cámara,
- \mathcal{F} , plano focal que pasa por el centro óptico de la cámara,
- \mathcal{I} , plano de la imagen paralelo al plano \mathcal{F} ,
- f , distancia focal que separa los planos \mathcal{I} y \mathcal{F} ,
- eje óptico, línea que atraviesa el punto C y es perpendicular a ambos planos, y
- c , punto principal donde atraviesa el eje óptico al plano \mathcal{I} .

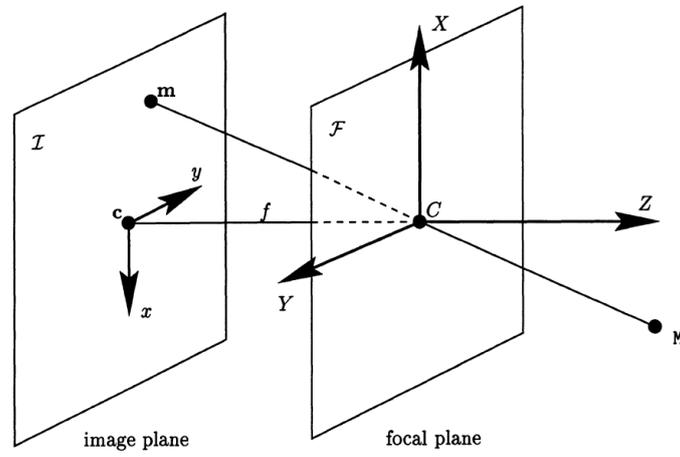


FIGURA 3.5: Modelo de la geometría de una cámara pinhole (Xu y Zhang, 2013).

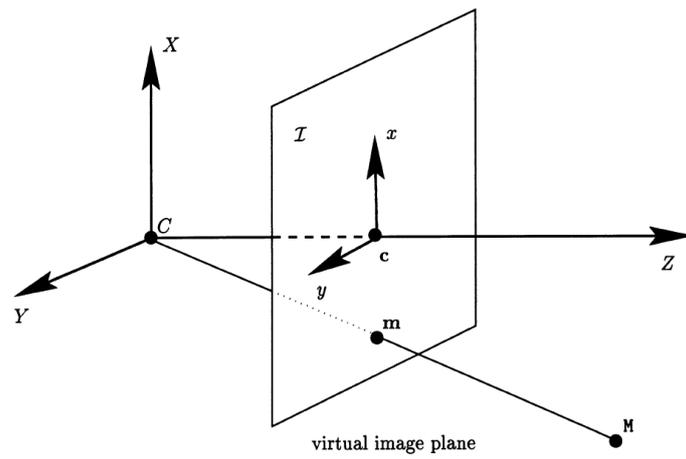


FIGURA 3.6: Modelo equivalente de una cámara pinhole (Xu y Zhang, 2013).

Un modelo más simple que reduce la complejidad matemática involucrada en la proyección se ilustra en la figura 3.6, en este modelo la proyección se puede expresar a partir de la razón de la proporción entre los lados de triángulos semejantes (Xu y Zhang, 2013):

$$[X, Y, Z]^T \rightarrow [x, y]^T = \left[\frac{fX}{Z}, \frac{fY}{Z} \right]^T. \quad (3.1)$$

La relación ?? describe una situación ideal, pero en escenarios reales existen varios factores que se tienen que tener en cuenta (Xu y Zhang, 2013):

- el origen de la imagen no coincide con la intersección del eje óptico y el plano de la imagen,
- debido a la electrónica de adquisición, los píxeles normalmente no son cuadrados, y
- las unidades de los ejes coordenados de la imagen no son necesariamente iguales.

3.5. Forma general de la proyección

El mapeo general que describe la proyección de un punto $\mathbf{M} = [X \ Y \ Z]^T$ en el mundo a un punto en el plano de la imagen $\mathbf{m} = [x \ y]^T$, expresado en cualquier sistema coordenado del mundo, se define por (Xu y Zhang, 2013):

$$\lambda \tilde{\mathbf{m}} = \mathbf{P} \tilde{\mathbf{M}}, \quad (3.2)$$

donde:

- λ es un factor de escala,
- $\tilde{\mathbf{m}}$ y $\tilde{\mathbf{M}}$ son las coordenadas homogéneas de los puntos \mathbf{m} y \mathbf{M} ,
- \mathbf{P} es una matriz de 3×4 con 11 grados de libertad, llamada matriz de proyección de perspectiva.

La matriz \mathbf{P} se puede descomponer en términos de los parámetros intrínsecos y los parámetros extrínsecos de la cámara, por lo tanto, la ecuación ?? se puede expresar de forma equivalente como (Xu y Zhang, 2013):

$$\lambda \tilde{\mathbf{m}} = \mathbf{K} [\mathbf{R} \ \mathbf{t}] \tilde{\mathbf{M}}. \quad (3.3)$$

3.5.1. Parámetros intrínsecos

Asumiendo que los puntos \mathbf{M} son expresados en el plano de coordenadas de la cámara ($\mathbf{M} = \mathbf{M}^c$) como se observa en la figura 3.6, la matriz de proyección se reduce a (Xu y Zhang, 2013):

$$\mathbf{P}_1 = \mathbf{K} [\mathbf{I}_{3 \times 3} \ \mathbf{0}_3] = \mathbf{K} \mathbf{P}_N,$$

donde \mathbf{P}_N es la matriz de proyección en perspectiva obtenida del sistema de coordenadas normalizado de la cámara, en este sistema se asume que el plano de la imagen se encuentra a una distancia unitaria del centro óptico. Los elementos de la matriz denotada con el símbolo \mathbf{K} son conocidos como parámetros intrínsecos de la cámara:

$$\mathbf{K} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix},$$

donde:

- (f_x, f_y) son los factores de escala definidos en la dirección x y y . Estos parámetros son proporcionales a la distancia focal $f_x = k_x f$ y $f_y = k_y f$. Los constantes k_x y k_y son el número total de píxeles por unidad en la dirección x y y , respectivamente.
- (c_x, c_y) es la coordenada del punto principal.
- s representa la inclinación de los píxeles debido a que los ejes no son exactamente ortogonales, asumiendo que los píxeles son cuadrados su valor llega a ser cero.

Los parámetros intrínsecos definen las condiciones de formación de imágenes, permiten relacionar información espacial con las imágenes al proyectar los puntos del sistema coordenado de la cámara al sistema coordenado de la imagen. Es importante tener una cámara correctamente calibrada, se dice que la cámara está calibrada cuando la matriz de parámetros intrínsecos es conocida, ya que de esto depende la exactitud de la localización estimada.

3.6. Parámetros extrínsecos

En la mayoría de aplicaciones es de interés expresar los puntos 3D en cualquier sistema de coordenadas del mundo como se ejemplifica en la figura 3.7, como se puede observar es posible ir del sistema coordenado de la cámara al sistema coordenado del mundo aplicando una rotación y una traslación, por lo tanto, los puntos del sistema coordenado del mundo \mathbf{M}^w se pueden expresar en el sistemas coordenado de la cámara utilizando la siguiente trasformación (Xu y Zhang, 2013):

$$\mathbf{M}^c = \mathbf{R}\mathbf{M}^w + \mathbf{t} \quad (3.4)$$

o, de forma equivalente como:

$$\tilde{\mathbf{M}}^c = \mathbf{D}\tilde{\mathbf{M}}^w, \quad \mathbf{D} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \quad (3.5)$$

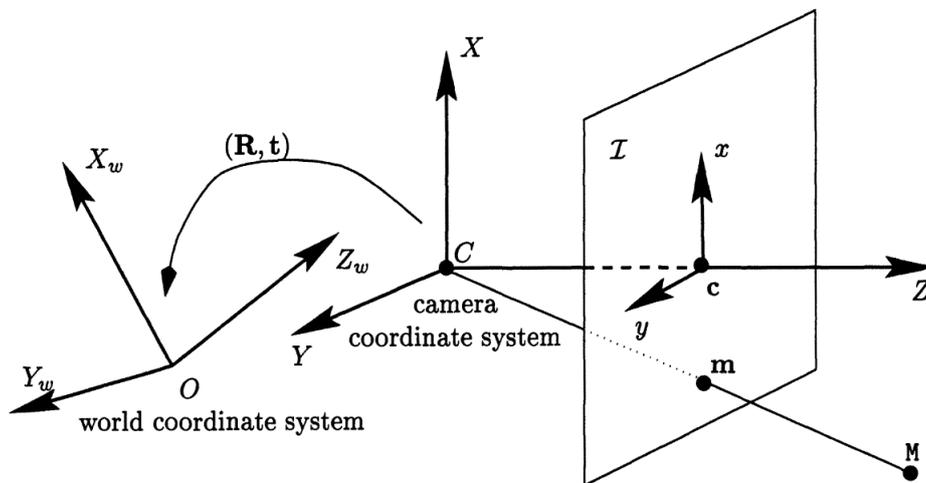


FIGURA 3.7: Proyección de puntos en el sistema coordenado del mundo al plano de la imagen (Xu y Zhang, 2013).

La matriz \mathbf{R} y el vector \mathbf{t} describen la orientación y la posición de la cámara con respecto al sistema coordenado del mundo, se les denomina parámetros extrínsecos de la cámara. Estos elementos forman la matriz \mathbf{D} que corresponde a un trasformación euclidiana del sistema coordenado del mundo al sistema coordenado de la cámara.

3.7. Perspectiva desde N Puntos

La recuperación de la pose de la cámara o de un objeto moviéndose enfrente de la cámara es un problema fundamental en visión artificial que ha sido abordado por los investigadores desde distintos enfoques. Este trabajo se centra en la técnica conocida como PnP (*Perspective-n-Point*). El problema PnP es el nombre que recibe la método cuyo propósito consiste en determinar la rotación y traslación de una cámara calibrada mediante n puntos 3D y sus correspondientes proyecciones 2D sobre la imagen, donde correspondencia entre puntos se rige por el siguiente modelo de proyección (Fischler y Bolles, 1981):

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (3.6)$$

Las soluciones del problema PnP pueden ser finitas o infinitas según la cantidad de correspondencias conocidas, el primer caso se manifiesta cuando $n \leq 2$, de otro modo, si $n \geq 3$ hay suficiente información para restringir el número de soluciones a una cantidad finita.

La pose definida en un espacio tridimensional consta de 6-DOF y cada correspondencia restringe 2-DOF, por lo tanto, para determinar la pose de un objeto con respecto a la cámara son necesarias al menos 3 correspondencias para restringir los 6-DOF.

Para soluciones finitas el numero soluciones está en función del número de correspondencias (Horaud et al., 1989):

- P3P: tiene 8 soluciones teóricas, no obstante, para cada solución positiva existe una solución geoméricamente correspondiente negativa, las soluciones negativas son irrelevantes y se pueden descartar y así el número máximo de soluciones posibles son 4.
- P4P o P5P: se llega a 2 soluciones, sin embargo, en el caso especial de puntos coplanarios la solución es única.
- A partir de P6P: siempre se produce una solución única.

El método numérico de mínimos cuadrados es una técnica clásica para resolver este problema, por ejemplo, Hesch y Roumeliotis (2011) presentan un método escalable basado en mínimos cuadrados para el caso general ($n \geq 3$). Por otro lado, existen técnicas no iterativas más eficientes, en el trabajo de Lepetit et al. (2008) se propone un método eficientes y precisos para el caso general ($n \geq 4$) y que maneja apropiadamente configuraciones de puntos coplanarios.

3.8. Conclusiones

Capítulo 4

Filtro de Kalman

4.1. Filtros de Kalman

«Cuando la gente te dice algo sobre algo que quieres saber y ya sabes un poco sobre eso, pero no estás seguro de lo que sabes, y tampoco están muy seguros de lo que te están diciendo, entonces es cuando necesitas un filtro de Kalman» (Orr, 1992).

El KF (*Kalman Filter*) es un algoritmo desarrollado por Rudolf Emil Kalman como una solución recursiva al problema de filtrado lineal de datos discretos (Kalman, 1960). Fue aplicado en la navegación del Proyecto Apolo para estimar la trayectoria de las naves espaciales tripuladas que iban a la Luna, revolucionó la teoría de control y es intensamente utilizado en sistemas de ingeniería; su impacto en la sociedad se refleja en una gran diversidad de productos de consumo, comerciales, salud y seguridad (Grewal y Andrews, 2010).

El KF es un conjunto de ecuaciones matemáticas que proporciona una solución óptima recursiva por el método de mínimos cuadrados al problema general de tratar de estimar el estado no medible de un proceso discreto a partir de mediciones que pueden contener ruido, se desarrolla suponiendo que el proceso de control se rige por un sistema lineal gaussiano (Welch et al., 1995).

El estado es el conjunto de aspectos del robot y del entorno que pueden impactar en el futuro, en otras palabras, el entorno se caracteriza por el estado, este se representa por los primeros dos momentos estadísticos de una función de distribución gaussiana. En el índice de tiempo k la estimación del estado es representado por la media denotada por \mathbf{x}_k y el error asociado a la estimación se representa por medio de la matriz de covarianza denotada por \mathbf{P}_k (Thrun et al., 2005).

Para asegurar que la distribución del estado sea siempre lineal gaussiano se deben cumplir las siguientes cuatro propiedades (Thrun et al., 2005):

- el estado es un resumen completo del pasado, es decir, se debe cumplir la propiedad de Markov; la suposición de Markov postula que los datos pasados y futuros son independientes si se conoce el estado actual \mathbf{x}_k ,
- la función de transición del estado debe ser una función lineal en sus argumentos con ruido gaussiano agregado.
- la función de medición también debe ser lineal en sus argumentos, con ruido gaussiano agregado.
- el estado inicial debe ser una distribución normal.

4.1.1. Proceso de estimación

Partiendo de la suposición que el sistema en cuestión satisface las propiedades de los sistemas lineales gaussianos, descritas anteriormente, el proceso de estimación del estado ejecuta de manera recursiva en dos partes comúnmente conocidas como fase de predicción seguida por la fase de corrección.

Fase de predicción:

La idea central de la fase de predicción es inferir el estado del sistema con cierto grado de confianza, con este propósito, se proyecta el estado del sistema hacia adelante en el tiempo resultando en una

estimación a priori del estado y su correspondiente error:

$$\begin{aligned}\mathbf{x}_k^- &= \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k, \\ \mathbf{P}_k^- &= \mathbf{A}_k \mathbf{P}_{k-1} \mathbf{A}_k^\top + \mathbf{Q}_k,\end{aligned}$$

donde:

- k : es el índice de tiempo discreto en la k -ésima iteración;
- \mathbf{x}_k^- : es la estimación a priori del estado;
- \mathbf{w}_k : representa el ruido del proceso, se asume media cero y covarianza \mathbf{Q} ;
- \mathbf{A}_k : es la matriz de transición de estado, relaciona el estado previo \mathbf{x}_{k-1} con el estado actual \mathbf{x}_k ;
- \mathbf{u}_k : es la entrada de control;
- \mathbf{B}_k : es la matriz de control, relaciona la entrada de control con el estado actual; y
- \mathbf{P}_k^- : es la matriz de covarianza, representa el error de la estimación a priori.

4.1.2. Fase de corrección:

El propósito de esta fase es refinar la estimación del estado resultante de la fase anterior, con este fin, se incorporan las nuevas mediciones de los sensores en la estimación a priori para obtener una nueva estimación a posteriori mejorada, la estimación se basa en cuánto se confía en los sensores y cuánto se confía en la estimación a priori, se calcula de la siguiente manera:

1. en función del estado se predice el vector de observación m -dimensional:

$$\hat{\mathbf{z}}_k = \mathbf{H}_k \mathbf{x}_k^- + \mathbf{v}_k,$$

2. se calcula la ganancia de retroalimentación del error en función de los ruidos que afectan el sistema:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^\top + \mathbf{R}_k)^{-1},$$

3. se corrige el estado predicho mediante la actualización proporcional del residuo entre la observación predicha y observada:

$$\begin{aligned}\mathbf{x}_k &= \mathbf{x}_k^- + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k), \\ \mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^-, \end{aligned}$$

donde:

- \mathbf{x}_k : es el estado a posteriori;
- \mathbf{z}_k y $\hat{\mathbf{z}}_k$: son las mediciones de los sensores y las observaciones estimadas a partir del estado, respectivamente;
- \mathbf{v}_k : es el ruido presente en las mediciones, se asume media cero y covarianza \mathbf{R}_k ;
- \mathbf{H}_k : es la matriz de medición, relaciona el estado con la medición actual; y
- \mathbf{K}_k : es la ganancia de Kalman.

Para profundizar sobre el filtro de Kalman, se puede consultar la referencia Welch et al. (1995).

4.1.3. Limitaciones

El filtro de Kalman estima la salida más probable con base a la certeza que se tiene de la evolución del sistema y las observaciones, sin embargo, para su adecuado funcionamiento el sistema debe cumplir con las propiedades de los sistemas lineales gaussianos, cuando alguna de esas propiedades no se cumplen se debe pensar en usar otras alternativas, entre las más conocidas se encuentran EKF (*Extended Kalman Filter*), el PF (*Particle Filter*) y el UKF (*Unscented Kalman Filter*).

4.2. El filtro de Kalman Extendido

En la práctica los modelos que gobiernan los sistemas no satisfacen los requerimientos impuestos por el KF, siendo su mayor limitación que necesita que el sistema sea lineal. La estimación en sistemas no lineales es de gran importancia debido a que la mayoría de aplicaciones tales como el seguimiento de objetos, navegación de vehículos y plantas de control de procesos químicos involucran no linealidades (Julier y Uhlmann, 2004).

El EKF es una variante del KF clásico que extiende su aplicación a sistemas no lineales gaussianos a costa de resultar menos óptimo, siendo menos óptimo cuanto menos lineal sea el modelo que rige el proceso de control del sistema.

EKF opera usando los mismos principios que el KF, pero parte de la premisa que el sistema es casi lineal para poder linealizar alrededor de la media y covarianza usando una herramienta matemática denominada series de Taylor (Jazwinski, 2007). Explotando esta suposición el EKF simplemente aproxima linealmente todas las transformaciones no lineales usando derivadas parciales de las funciones del proceso y observación, mismas que sustituyen las transformaciones en las ecuaciones del KF (Julier y Uhlmann, 2004).

4.2.1. Proceso de estimación

El resultado de sustituir las transformaciones no lineales en las ecuaciones del KF clásico con su correspondientes matriz jacobiana da como resultado el siguiente conjunto de ecuaciones.

Fase de predicción

La fase de predicción se define por Welch et al. (1995):

$$\begin{aligned}\mathbf{x}_k^- &= f(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{0}), \\ \mathbf{P}_k^- &= \mathbf{A}_k \mathbf{P}_{k-1} \mathbf{A}_k^\top + \mathbf{W}_k \mathbf{Q}_{k-1} \mathbf{W}_k^\top,\end{aligned}$$

donde:

- \mathbf{x}^- : es la estimación a priori del estado;
- \mathbf{u}_k : es la entrada de control;
- \mathbf{w} : representa el ruido del proceso, se asume media cero y covarianza \mathbf{Q} ;
- \mathbf{f} : es la función de transición de estado;
- \mathbf{A}_k : es la matriz jacobiana de derivadas parciales de \mathbf{f} con respecto a \mathbf{x} :

$$\mathbf{A}_k[i, j] = \frac{\partial \mathbf{f}^{[i]}}{\partial \mathbf{x}^{[j]}}(\mathbf{x}, \mathbf{u}_k, \mathbf{0});$$

- \mathbf{W}_k : es la matriz jacobiana de derivadas parciales de \mathbf{f} con respecto a \mathbf{w} :

$$\mathbf{W}_k[i, j] = \frac{\partial \mathbf{f}^{[i]}}{\partial \mathbf{w}^{[j]}}(\mathbf{x}, \mathbf{u}_k, \mathbf{0});$$

- \mathbf{P} : es la matriz de covarianza, representa el error de la estimación.

Fase de corrección

La fase de corrección se definen por Welch et al. (1995):

$$\begin{aligned}\hat{\mathbf{z}}_k &= h(\mathbf{x}_k^-, \mathbf{0}), \\ \mathbf{K}_k &= \mathbf{P}_k^- \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^\top + \mathbf{V}_k \mathbf{R}_k \mathbf{V}_k^\top)^{-1}, \\ \mathbf{x}_k &= \mathbf{x}_k^- + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k), \\ \mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^-, \end{aligned}$$

donde:

- \mathbf{x} : es el estado a posteriori;
- \mathbf{v} : es el ruido presente en las mediciones, se asume media cero y covarianza \mathbf{R} ;
- \mathbf{z} y \mathbf{z}^* : son las mediciones de los sensores y las observaciones estimadas a partir del estado, respectivamente;
- \mathbf{h} : es la función de observación que relaciona el estado con la medición;
- \mathbf{H} : es la matriz jacobiana de derivadas parciales de \mathbf{h} con respecto \mathbf{x} :

$$\mathbf{H}_k[i, j] = \frac{\partial \mathbf{h}^{[i]}}{\partial \mathbf{x}^{[j]}}(\mathbf{x}, \mathbf{u}_k, \mathbf{0});$$

- \mathbf{V}_k : es la matriz jacobiana de derivadas parciales de \mathbf{h} con respecto a \mathbf{v} :

$$\mathbf{V}_k[i, j] = \frac{\partial \mathbf{h}^{[i]}}{\partial \mathbf{v}^{[j]}}(\mathbf{x}, \mathbf{u}_k, \mathbf{0});$$

- \mathbf{K} : es la ganancia de Kalman.

4.2.2. Limitaciones

Las principales limitaciones del EKF son:

- la estimación es cercana al valor real solo si el sistema es casi lineal alrededor del punto de operación, de lo contrario puede producir alta inestabilidad (Julier y Uhlmann, 1997);
- el cálculo de las matrices jacobianas suele ser un trabajo complicado, lo que puede llevar a dificultades en la implementación (Julier y Uhlmann, 1997);
- la linealización solo se puede aplicar si existe la matriz jacobiana, en consecuencia, las discontinuidades en sistemas son un problema (Julier y Uhlmann, 2004); y
- el error de estimación puede llegar a ser divergente si el sistema es no lineal (Reif et al., 1999).

Estos inconvenientes han motivado el desarrollo de otras técnicas. Por ejemplo, el filtro de partículas es una técnica que no hace suposiciones del sistema, puede aproximar cualquier distribución multimodal, su funcionamiento se basa en métodos Monte Carlo de muestreo secuencial donde la idea clave consiste en calcular recursivamente las funciones de distribución de probabilidad usando un conjunto de muestras aleatorias con pesos asociados, aunque la alta carga computacional es su principal desventaja Gustafsson et al. (2002).

En contraste con el PF, el UKF no es una técnica general ya que solo aborda el problema de la no linealidad e intenta mejorar el rendimiento del EKF infiriendo el punto de operación estocásticamente a un costo computacional menor que el PF. En este trabajo es de mayor importancia el UKF, por lo que en la siguiente sección se describen los aspectos más relevantes de este método.

4.3. El filtro de Kalman Unscented

El UKF es un estimador que generaliza a sistemas no lineales propuesto por Julier y Uhlmann, tiene un rendimiento superior al EKF, mientras que la demanda computacional es comparable entre ambos cuando los jacobianos son calculados usando diferencias finitas (Kandepu et al., 2008).

El componente clave del UKF es la UT (*Unscented Transformation*). La idea fundamental de la UT consiste en predecir los primeros dos momentos estadísticos de sistemas no lineales a partir de un conjunto de puntos seleccionados estratégicamente.

4.3.1. Transformada Unscented

La UT es un método para calcular las estadísticas de una variable aleatoria la cual sufre de una transformación no lineal (Julier y Uhlmann, 1997). Para cualquier variable aleatoria no gaussiana, los resultados garantizan al menos la precisión de un filtro truncado de segundo orden para todas las no linealidades, pudiéndose aumentar el grado de precisión de las aproximaciones mediante la selección adecuada de los parámetros α y β , descritos posteriormente. Particularmente, para sistemas gaussianos brinda aproximaciones con precisión de tercer orden para cualquier transformación no lineal (Wan y Van Der Merwe, 2001).

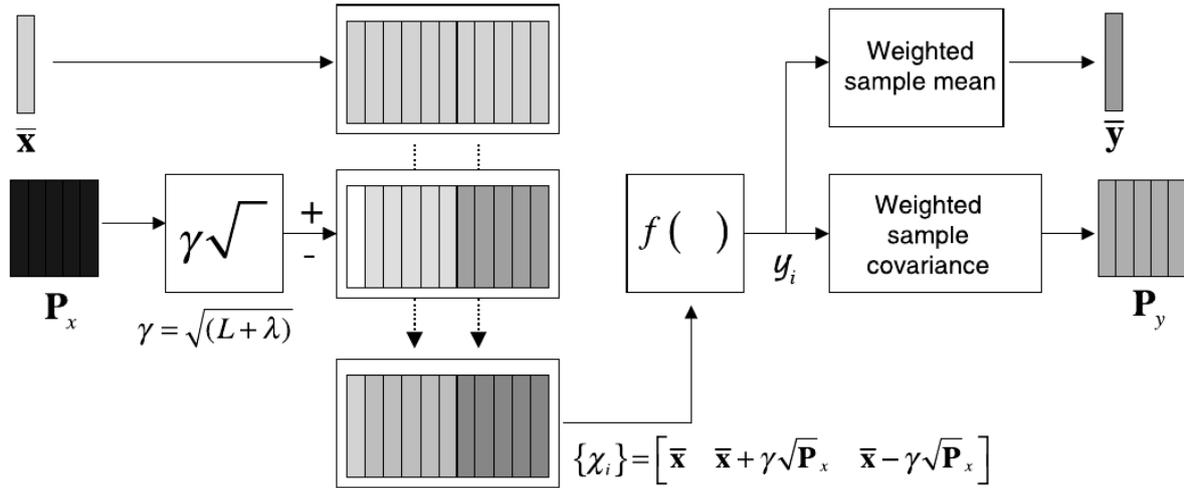


FIGURA 4.1: Diagrama de bloques que ilustra el procedimiento de la UT (Wan y Van Der Merwe, 2001).

Para una variable aleatoria \mathbf{x} de dimensión L con media $\bar{\mathbf{x}}$ y matriz de covarianza \mathbf{P}_x , el procedimiento para calcular la UT es de la siguiente manera (Julier y Uhlmann, 1997; Wan y Van Der Merwe, 2001):

1. Se selecciona el conjunto de puntos sigma χ de $2L + 1$ vectores:

$$\begin{aligned} \chi_0 &= \bar{\mathbf{x}} \\ \chi_i &= \bar{\mathbf{x}} + \left(\sqrt{(L+\lambda)\mathbf{P}_x} \right)_i & i = 1, \dots, L \\ \chi_i &= \bar{\mathbf{x}} - \left(\sqrt{(L+\lambda)\mathbf{P}_x} \right)_{i-L} & i = L+1, \dots, 2L, \end{aligned}$$

donde:

- $\left(\sqrt{(L+\lambda)\mathbf{P}_x} \right)_i$: es la i -ésima columna de la raíz cuadrada de la matriz $(L+\lambda)\mathbf{P}_x$. La raíz cuadrada \mathbf{S} de una matriz \mathbf{A} satisface $\mathbf{A} = \mathbf{S}\mathbf{S}^T$, donde \mathbf{S} es una matriz triangular inferior. La matriz \mathbf{S} se puede calcular de forma eficiente y estable utilizando la descomposición de Cholesky debido a que la matriz de covarianza es simétrica y definida positiva, propiedades que se pueden demostrar.
- λ : es un factor de escalamiento definido como:

$$\lambda = \alpha^2(L + \kappa) - L.$$

- α : es un parámetro que determina la dispersión de los puntos sigma alrededor de la media, normalmente toma valores positivos pequeños:

$$1e-4 \geq \alpha \leq 1.$$

- κ : es un parámetro de escalamiento, generalmente se establece a:

$$\kappa = 3 - L.$$

Esto se hace con la finalidad de eliminar los efectos de la dimensión del espacio de estados, ya que a medida que esta aumenta también lo hace la curtosis de los puntos sigma lo que implica que el muestreo pierda su efecto local. Este valor hace que la curtosis de la distribución coincida con la de una distribución Gaussiana y establece a tres el radio de la esfera que limita los puntos sigma, en consecuencia, los momentos de orden superior de los puntos sigma son de crecimiento más lento que los momentos de orden superior de una distribución Gaussiana Julier et al. (1995).

2. Se proyectan los puntos sigma a través de la función no lineal f para obtener los puntos sigma transformados:

$$\mathcal{Y}_i = f(\chi_i).$$

3. Se calcula la media y la matriz de covarianza en función de los puntos sigma transformados y sus respectivos pesos asociados:

$$\bar{\mathbf{y}} = \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Y}_i$$

$$\mathbf{P}_y = \sum_{i=0}^{2L} W_i^{(c)} [\mathcal{Y}_i - \bar{\mathbf{y}}] [\mathcal{Y}_i - \bar{\mathbf{y}}]^\top,$$

con los pesos W_i definidos como:

$$W_0^{(m)} = \frac{\lambda}{L + \lambda},$$

$$W_0^{(c)} = \frac{\lambda}{L + \lambda} + 1 - \alpha^2 + \beta,$$

$$W_i^{(m)} = W_i^{(c)} = \frac{1}{2(L + \lambda)}, \quad i = 1, \dots, 2L,$$

en el cual β es un parámetro utilizado para incorporar el conocimiento a priori de la distribución \mathbf{x} , $\beta = 2$ es un valor óptimo para una distribución es gaussiana.

Los pasos involucrados en el procedimiento para calcular la UT se resumen en el diagrama de bloques de la figura 4.1.

4.3.2. Proceso de estimación

El filtro de Kalman Unscented es un extensión de la UT para estimar el estado de un sistema recursivamente, a continuación se describe el algoritmo.

1. Inicialización:

$$\bar{\mathbf{x}} = \mathbb{E}[\mathbf{x}_0], \quad \mathbf{P}_0 = \mathbb{E}[(\mathbf{x}_0 - \bar{\mathbf{x}})(\mathbf{x}_0 - \bar{\mathbf{x}})^\top].$$

Para $k = 1$ hasta ∞ :

2. El conjunto de puntos sigma es calculado como:

$$\chi_{k-1} = \left[\bar{\mathbf{x}}_{k-1}, \quad \bar{\mathbf{x}}_{k-1} + \sqrt{(L + \lambda) \mathbf{P}_{k-1}}, \quad \bar{\mathbf{x}}_{k-1} - \sqrt{(L + \lambda) \mathbf{P}_{k-1}} \right].$$

3. Fase de predicción:

$$\begin{aligned}\chi_{k|k-1} &= \mathbf{F}(\chi_{k-1}, \mathbf{u}_{k-1}), \\ \bar{\mathbf{x}}_k^- &= \sum_{i=0}^{2L} W_i^{(m)} \chi_{k|k-1}^i, \\ \mathbf{P}_k^- &= \sum_{i=0}^{2L} W_i^{(c)} \left[\chi_{k|k-1}^i - \bar{\mathbf{x}}_k^- \right] \left[\chi_{k|k-1}^i - \bar{\mathbf{x}}_k^- \right]^\top + \mathbf{Q}, \\ \mathbf{Z}_{k|k-1} &= \mathbf{H}(\chi_{k|k-1}), \\ \bar{\mathbf{z}}_k^- &= \sum_{i=0}^{2L} W_i^{(m)} \mathbf{Z}_{k|k-1}^i,\end{aligned}$$

4. Fase de corrección:

$$\begin{aligned}\mathbf{P}_{z_k z_k} &= \sum_{i=0}^{2L} W_i^{(c)} \left[\mathbf{Z}_{k|k-1}^i - \bar{\mathbf{z}}_k^- \right] \left[\mathbf{Z}_{k|k-1}^i - \bar{\mathbf{z}}_k^- \right]^\top + \mathbf{R}, \\ \mathbf{P}_{x_k z_k} &= \sum_{i=0}^{2L} W_i^{(c)} \left[\chi_{k|k-1}^i - \bar{\mathbf{x}}_k^- \right] \left[\mathbf{Z}_{k|k-1}^i - \bar{\mathbf{z}}_k^- \right]^\top, \\ \mathbf{K}_k &= \mathbf{P}_{x_k z_k} (\mathbf{P}_{z_k z_k})^{-1}, \\ \bar{\mathbf{x}}_k &= \bar{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \bar{\mathbf{z}}_k^-), \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{P}_{z_k z_k} \mathbf{K}_k^\top,\end{aligned}$$

donde:

- \mathbf{Q} es la covarianza de ruido del proceso,
- \mathbf{R} es covarianza del ruido de las mediciones.

4.3.3. Fortalezas

La UT aborda el problema de la linealización del EKF resultando una serie de propiedades importantes en la práctica (Julier y Uhlmann, 2004):

- puede ser utilizada como una caja negra para predecir cantidades de cualquier transformación;
- la demanda computacional del algoritmo es equivalente a la del EKF;
- garantiza la precisión de segundo orden para cualquier no linealidad, pero sin la necesidad de calcular jacobianos y hessianos; y
- trabaja adecuadamente con transformaciones discontinuas.

En la figura 4.2 se ilustra un ejemplo comparativo del rendimiento de la UT, el método Monte Carlo de muestreo y un enfoque de linealización.

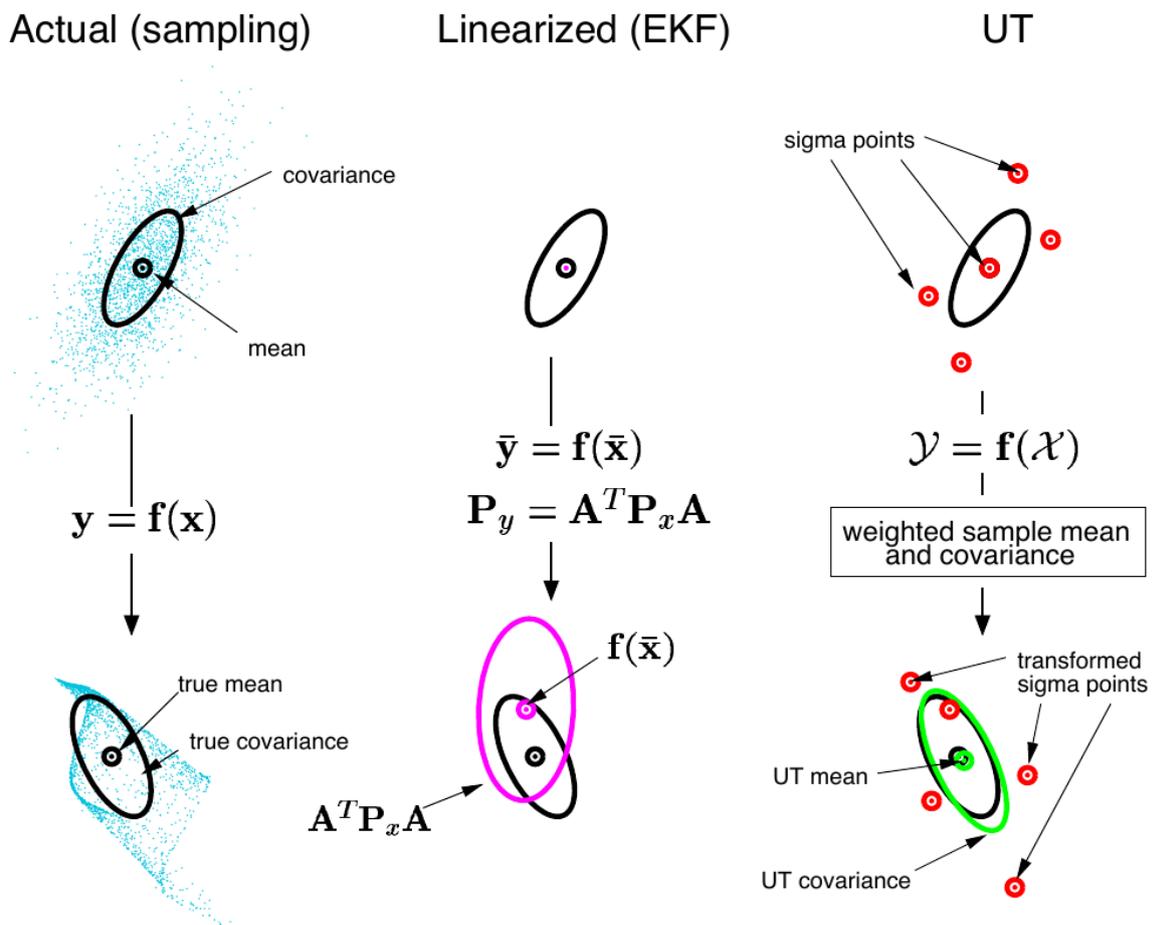


FIGURA 4.2: Ejemplo comparativo de la UT para calcular la media y la covarianza de un sistema bidimensional (Wan y Van Der Merwe, 2000).

Capítulo 5

Metodología

Después de haber definido el problema de investigación y los objetivos a alcanzar. En esta sección se describe el orden metodológico que se siguió para ejecutar la investigación y así poder responder al problema planteado. Primero se da un panorama general de las partes que componen el sistema de navegación. Después se describen los mecanismos fundamentales para estimar la pose del vehículo desde dos enfoque diferentes; por un lado usando sensores inerciales y por el otro usando una cámara externa. Finalmente, se plantea una solución para la estimación robusta de la pose fusionando ambos enfoques usando el filtro de Kalman Unscented.

5.1. Descripción del sistema

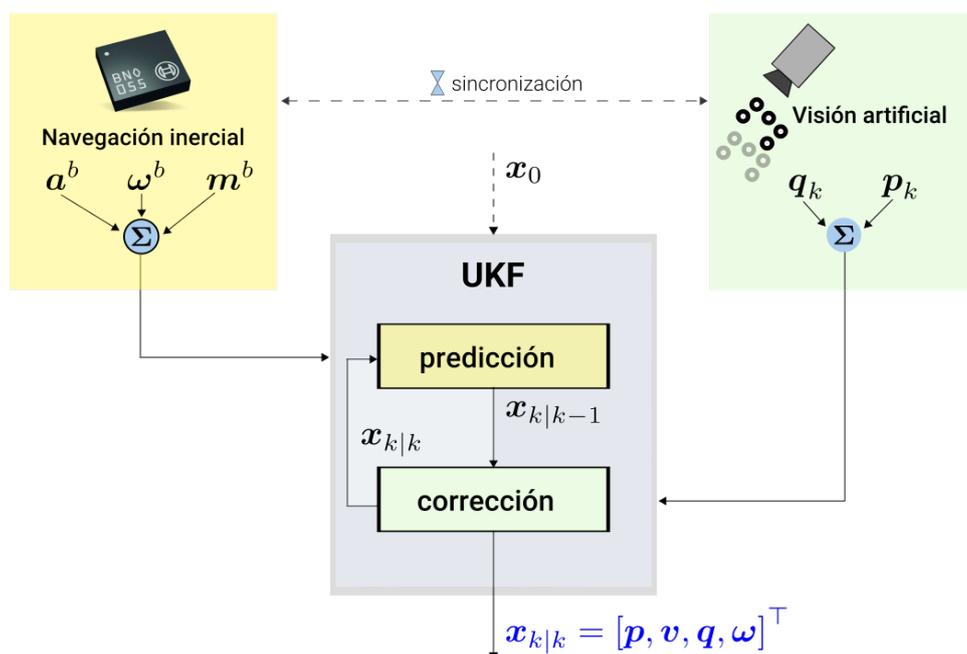


FIGURA 5.1: Esquema general del seguimiento de la localización 3D del vehículo usando sensores inerciales, una cámara y el filtro de Kalman Unscented.

El sistema está compuesto de dos partes que trabajan en sinergia con el filtro de Kalman Unscented para lograr un sistema de navegación más robusto como se ilustra en la figura 5.1.

La primera parte, denominada sistema de navegación inercial, comprende de un vehículo al cual se le ajustó rígidamente a su estructura una IMU de 9-DOF para medir aceleración, velocidad angular y campo magnético local. Estas mediciones permiten seguir el movimiento del vehículo aplicando la técnica de navegación inercial descrita en la sección 2.1.

La segunda parte, llamada sistema de navegación visual basada en marcas, se compone de una cámara externa que es utilizada para inferir la localización del vehículo aplicando una técnica de visión compu-

tacional conocida como *Perspective-n-Point*, se asume que el vehículo se mantiene en el campo visual de la cámara.

5.2. Configuración de los marcos de referencia

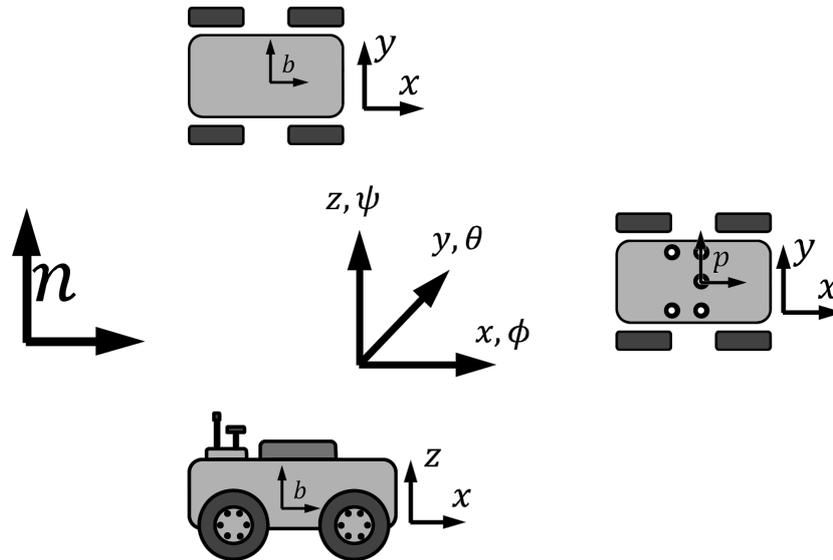


FIGURA 5.2: Principales marcos de referencia para la estimación de la pose: marco de navegación, marco del cuerpo y marco del patrón de puntos denotados con las letras \mathbf{n} , \mathbf{b} y \mathbf{p} , respectivamente.

El sentido de los vectores y las direcciones de los movimientos de rotación del vehículo se rigen por la regla de la mano derecha. Los marcos de referencia aquí usados son sistemas de coordenadas cartesianas que permiten seguir la localización en el espacio tridimensional, estos se ilustran en la figura 5.2.

La posición y la orientación del vehículo se calcula respecto al marco de navegación \mathbf{n} , el cual se define con el eje X apuntado hacia el Este, el eje Y hacia el Norte y el eje Z alineado con el vector de gravedad, según la regla de la mano derecha.

La IMU se montó a la estructura del vehículo con el eje X hacia adelante y el eje Z apuntando hacia arriba, su centro corresponde al marco del cuerpo \mathbf{b} .

El marco \mathbf{p} hace referencia al marco del patrón de puntos colocado sobre el vehículo, es usado por el sistema de visión basado en marcas para estimar la localización del vehículo, su centro coincide con el segundo punto del patrón que será descrito posteriormente. El patrón de puntos fue colocado sobre el vehículo de tal manera que el marco \mathbf{b} y el marco \mathbf{p} están sobrepuestos, idealmente son el mismo marco de referencia.

Los movimientos de rotación ocurren primero alrededor del eje X , después alrededor del eje Y y finalmente alrededor del eje Z . Estas rotaciones expresadas en términos de los ángulos llamados roll ϕ , pitch θ y yaw ψ se miden en sentido contrario de las manecillas del reloj, conforme a la regla de la mano derecha.

5.3. Localización usando sensores inerciales

La navegación inercial usa los datos proporcionados por la IMU, la estimación anterior y el tiempo para estimar la posición y orientación del vehículo que depende de las condiciones iniciales como se ilustra en la figura 5.3.

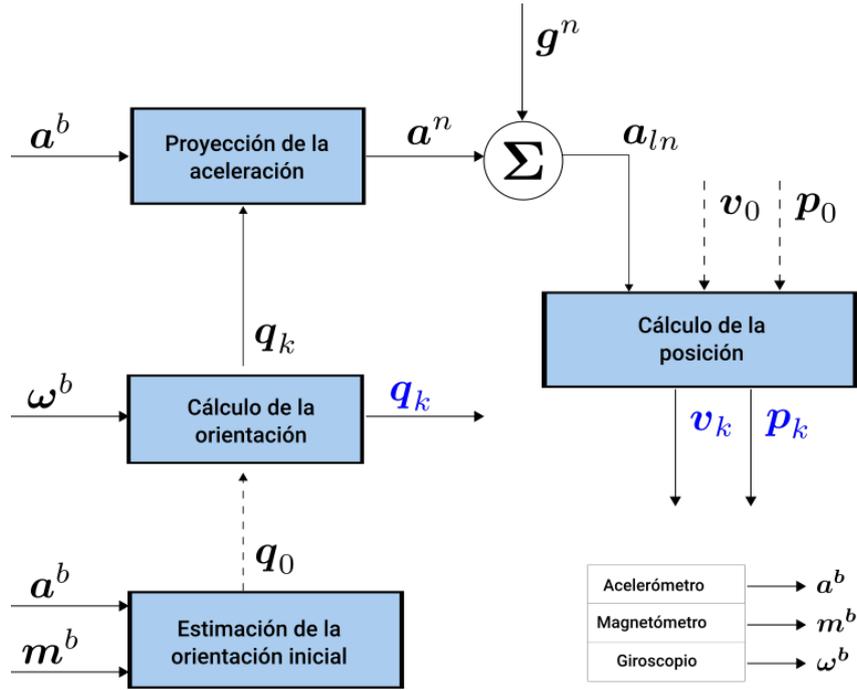


FIGURA 5.3: Diagrama de bloques para la estimación de la localización usando navegación inercial.

La IMU provee señales de velocidad angular $\omega^b(t)$, aceleración $\mathbf{a}^b(t)$ y campo magnético $\mathbf{m}^b(t)$ de los ejes del marco de cuerpo, en la sección 2.1 se demostró que estas señales de salidas están afectadas por sesgo y otro tipo de errores, por lo que se definió un modelo de medición para el giroscopio (??), el acelerómetro (??) y el magnetómetro (??). Después de haber calibrado correctamente estos sensores, se asume que las mediciones en los tres ejes son independientes, además que el sesgo δ y los errores \mathbf{e} han sido compensados, de tal forma que la salida de los sensores se encuentran libres de cualquier perturbación:

$$\omega^b = \omega - \delta_\omega - \mathbf{e}_\omega, \quad (5.1)$$

$$\mathbf{a}^b = {}^b_n\mathbf{R}(\mathbf{a}^n - \mathbf{g}^n) - \delta_a - \mathbf{e}_a, \quad (5.2)$$

$$\mathbf{m}^b = {}^b_n\mathbf{R}\mathbf{m}^n - \delta_m - \mathbf{e}_m. \quad (5.3)$$

5.3.1. Seguimiento de la orientación

La orientación del vehículo respecto al marco de navegación se calculó agregando continuamente los cambios detectados en la orientación usando la señal ω^b obtenida del giroscopio, de acuerdo a:

$$\mathbf{q}_{k+1} = \mathbf{q}_k \odot \mathbf{q}\Delta, \quad (5.4)$$

donde k es el índice del tiempo discreto, \mathbf{q}_k es la orientación estimada hasta el índice de tiempo k y $\mathbf{q}\Delta$ es el cuaternión que representa el cambio en la orientación en el intervalo de tiempo Δt , desde el índice de tiempo k hasta $k+1$.

Asumiendo que la velocidad angular ω^b es constante en el intervalo de tiempo Δt , el cuaternión correspondiente al cambio de orientación en este intervalo de tiempo está dado por:

$$\alpha = \|\omega^b\| \cdot \Delta t, \quad \mathbf{n} = \frac{\omega^b}{\|\omega^b\|},$$

$$\mathbf{q}\Delta = [\cos(\alpha), \mathbf{n} \cdot \sin(\alpha)]^T. \quad (5.5)$$

Así, la orientación evoluciona en pasos discretos de manera recursiva de acuerdo a (??), iniciando con $\mathbf{q}_0 = [1, 0, 0, 0]^T$.

5.3.2. Cálculo de la orientación inicial

En este trabajo es de interés navegar con respecto a una pose particular, por lo que \mathbf{q}_0 tiene que ser estimado en una localización determinada al momento de iniciar el INS. Aquí, se calculó de forma estadística usando el método-q. Esta técnica se fundamenta en el hecho que dado un conjunto de N vectores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ en un marco de referencia y sus correspondientes observaciones $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N\}$ desde otro marco de referencia, entonces se puede determinar la rotación que lleva al primer conjunto a la mejor coincidencia con el segundo:

$$\mathbf{u}_i = \mathbf{R}\mathbf{v}_i. \quad (5.6)$$

Seguindo esta idea, se usó el vector de gravedad y el campo magnético para determinar la orientación del marco del cuerpo respecto al marco de navegación:

$$\mathbf{g}^n = [0, 0, 1]^\top, \quad \mathbf{m}^n = [1, 0, 0]^\top. \quad (5.7)$$

El acelerómetro sirvió para proveer información de la inclinación ya que suele ser más precisa en comparación con el magnetómetro, que también provee información de la inclinación excepto en el Ecuador. El magnetómetro se usó únicamente para obtener información del rumbo, por lo cual fue necesario proyectar el campo magnético y las mediciones del magnetómetro en el plano horizontal:

$$\mathbf{g}^b = \frac{\mathbf{a}^b}{\|\mathbf{a}^b\|}, \quad \mathbf{m}^b = \mathbf{g}^b \times \left(\frac{\mathbf{m}^b}{\|\mathbf{m}^b\|} \times \mathbf{g}^b \right). \quad (5.8)$$

De modo que tomando muestras de ambos sensores usando (??) con el vehículo estático en el lugar de partida se calculó la orientación inicial \mathbf{q}_0 .

5.3.3. Seguimiento de la posición

El acelerómetro mide la aceleración debida a la gravedad de la Tierra y la aceleración debida al movimiento al que está sometido el sensor. La aceleración que se produce debido al movimiento es conocida como aceleración lineal, y es esta la que requiere el INS para estimar su posición.

Por lo tanto, para calcular la posición del vehículo la señal de aceleración \mathbf{a}^b es proyectada al marco de navegación, de acuerdo a:

$$\mathbf{a}^n = {}^n_b\mathbf{q} \odot \tilde{\mathbf{a}}^b \odot ({}^n_b\mathbf{q})^{-1}. \quad (5.9)$$

Inmediatamente después la aceleración debida a la gravedad es eliminada para obtener la aceleración lineal:

$$\mathbf{a}_{ln}^n = \mathbf{a}^n - \mathbf{g}^n, \quad (5.10)$$

donde $\tilde{\mathbf{a}}^b$ es la aceleración en forma de cuaternión y \mathbf{g}^n es el vector de gravedad.

Partiendo de la premisa que la aceleración se mantiene constante en el intervalo de tiempo T , la posición y la velocidad se calculan agregando continuamente los cambios en la posición y la velocidad correspondientes a este intervalo de tiempo usando las formulas de la cinemática:

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \mathbf{a}_k \cdot T, \quad (5.11)$$

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{v}_k \cdot T + \frac{1}{2} \cdot \mathbf{a}_k \cdot T^2, \quad (5.12)$$

en las ecuaciones de arriba k es el índice de tiempo discreto, \mathbf{a}_k es la aceleración lineal obtenida de (??) y T es intervalo de tiempo desde k hasta $k+1$. Cuando el vehículo parte del estado de reposo la velocidad inicial \mathbf{v}_0 y la posición inicial \mathbf{p}_0 son establecidas a cero.

5.4. Localización visual basada en marcas

La idea central de lo que aquí se entiende como localización visual consiste en usar una cámara para capturar una escena del mundo real en una imagen digital. Esta transformación que ocurre del mundo analógico al mundo digital conocida como transformación en perspectiva guarda información geométrica de la relación de correspondencia de $3D$ a $2D$, la cual es usada para determinar la pose de un objeto con la ayuda de un modelo que describe la correspondencia de dicha transformación.

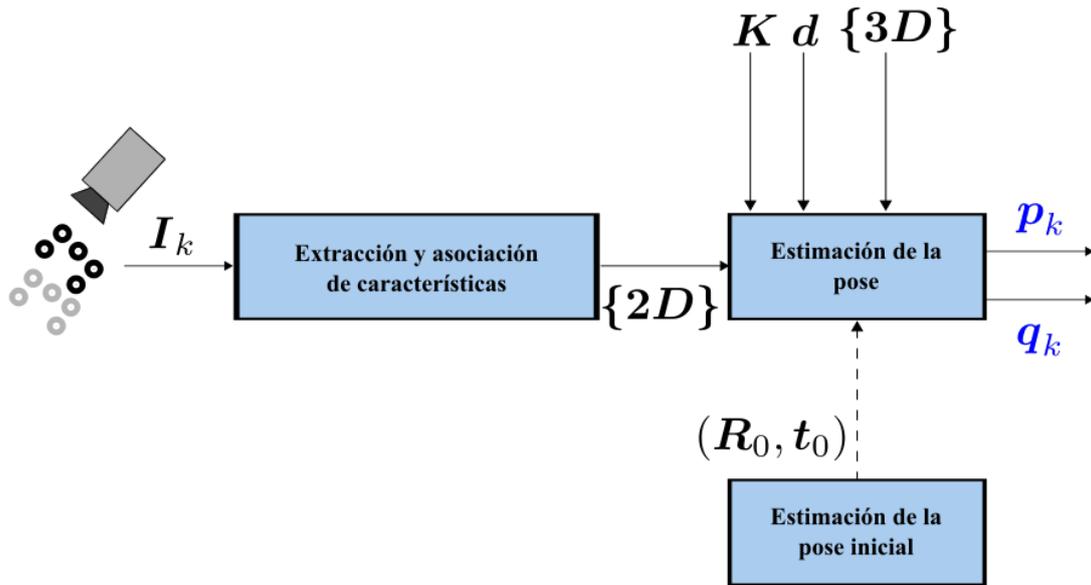


FIGURA 5.4: Diagrama de bloques del procedimiento para la estimación de la localización del vehículo usando visión artificial.

El seguimiento de un objetos consiste en calcular su posición y orientación en una secuencia de imágenes. El escenario común para ejecutar esta tarea requiere registrar previamente un modelo 2D de características naturales que identifican al objeto de otros objetos, por medio del cual es posible reconocer al objeto en la secuencia de imágenes. En este modelo cada punto característico tiene asociado una coordenada 3D, respecto al marco de referencia del objeto, que serán usadas para establecer las correspondencias necesarias para recuperar la pose del objeto.

El procedimiento anterior requiere calcular cientos de características en cada imagen que también tienen que ser asociados con sus respectivas coordenadas 3D, por lo que puede convertirse en un trabajo costoso. Aquí se aborda el problema desde un enfoque más simple que usa una cantidad mínima de puntos característicos, el procedimiento general se ilustra en la figura 5.4.

5.4.1. Modelo 3D

El modelo 3D usado aquí corresponde a un conjunto de puntos en el espacio coplanario. La solución al problema de la pose a partir de puntos característicos en un mismo plano es de bastante importancia en la práctica. Existen muchas aplicaciones donde el suelo no es completamente plano, pero la distribución de los puntos es mucho mayor que su elevación y por lo tanto pueden ser considerados como puntos coplanarios.

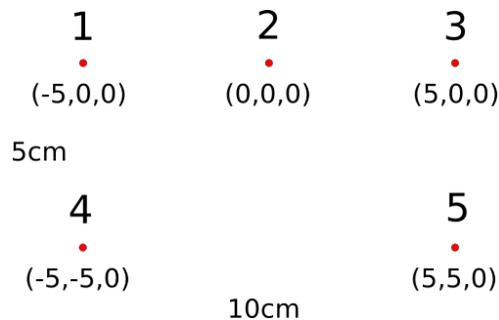


FIGURA 5.5: Modelo 3D usado para el seguimiento del vehículo.

Para el seguimiento de la pose del vehículo se usó un modelo 3D compuesto por cinco coordenadas tridimensionales en el que todas las coordenadas yacen en el plano XY como se muestra en la figura 5.5. La distancia entre un punto y su vecino es de cinco centímetros, excepto la distancia entre los dos últimos que es de diez centímetros.

De este modelo surge el marco del patrón de puntos \mathbf{p} , su centro corresponde con el segundo punto, el eje X apunta en dirección $P1 \rightarrow P3$ y el eje Y apunta en dirección $P4 \rightarrow P1$.

Debido a que se requiere detectar el modelo en la secuencia de imágenes, se colocaron círculos sobre los puntos de tal forma que el centro de los círculos coincide con las coordenadas del modelo como se ilustra en la figura 5.6.

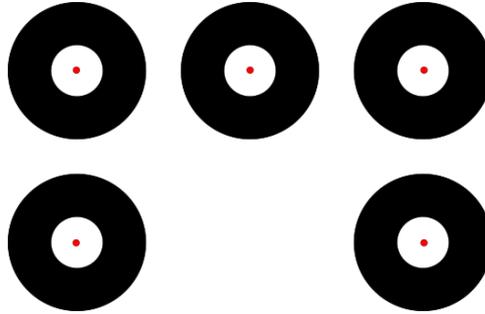


FIGURA 5.6: Marcas colocadas sobre los puntos característicos: círculo negro con radio igual a dos centímetros y círculo blanco con radio igual a un centímetro.

5.4.2. Extracción y correspondencia de características

Los puntos característicos representan una localización particular en la imagen, estos provienen del patrón de puntos descrito previamente. Los pasos necesarios para su extracción se describen a continuación:

- Se suaviza la imagen para reducir el ruido usando un filtro gaussiano con un kernel de tamaño 5×5 .
- Se obtiene una imagen binaria donde las áreas negras de la imagen original ahora son blancas y viceversa.
- Se recuperan todos los contornos de la imagen y se filtran estructuras circulares padre-hijo que compartan el mismo centro.
- Se encuentra la correspondencia de puntos ordenándolos del 1 al 5.
- Se valida que los puntos 1,3,4,5 sean los vértices de una forma rectangular para evitar reconocer puntos característicos atípicos.

La figura 5.7 muestra el resultado de la extracción de estos puntos característicos en una imagen.



FIGURA 5.7: Resultado de la extracción de los puntos de interés representados en color verde.

5.4.3. Recuperación de la pose

La pose de la cámara se calculó usando el enfoque PnP. El problema PnP consiste en estimar la pose de una cámara calibrada a partir de n correspondencias de puntos de $3D$ a $2D$ (Fischler y Bolles,

1981). Aquí se usó una variante caracterizada por su eficiencia y precisión conocida como EPnP (*Efficient Perspective-n-Point*), este método es aplicable para todo $n \geq 4$ con la ventaja que maneja adecuadamente configuraciones de puntos coplanarios (Lepetit et al., 2008).

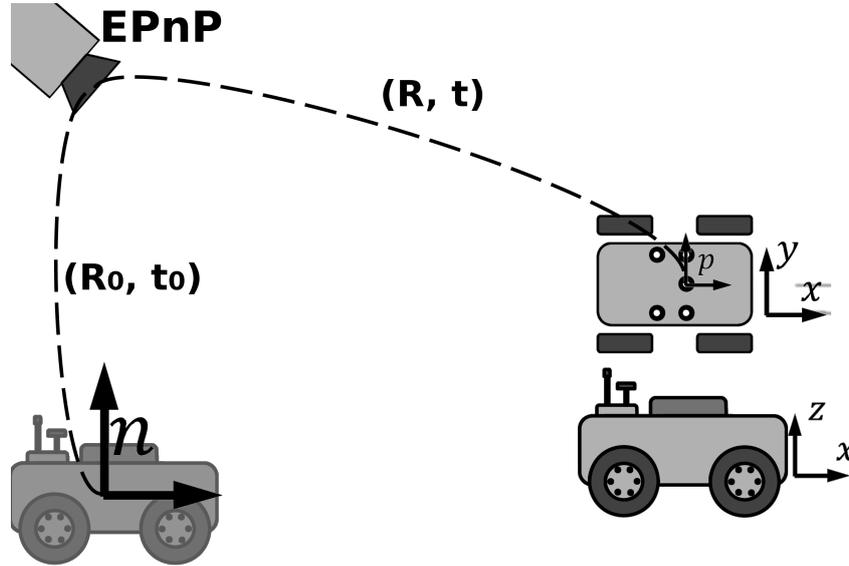


FIGURA 5.8: Seguimiento del vehículo.

Para cada imagen de la secuencia se hace la extracción y correspondencia de características, descrita en la sección 5.4.2. Esta correspondencia, el modelo 3D, la matriz de calibración y los coeficientes de distorsión son usados por el algoritmo EPnP para estimar la pose de la cámara, el esquema general de este procedimiento se muestra en la figura 5.4.

El seguimiento del vehículo se hace respecto a una localización inicial, esta localización $(\mathbf{R}_0, \mathbf{t}_0)$ se estima usando la primera imagen de la secuencia. El resto de estimaciones se hacen respecto a esta localización, por lo que las siguientes rotaciones y traslaciones (\mathbf{R}, \mathbf{t}) calculadas son proyectadas usando $(\mathbf{R}_0, \mathbf{t}_0)$ para obtener la estimación deseada. Las rotaciones se hacen aplicando operaciones entre cuaterniones.

5.5. Modelo del sistema

Siguiendo el enfoque del filtro de Kalman descrito en la sección 4.1, el estado del sistema está compuesto por la posición, la velocidad, la orientación y la velocidad angular, definidos en el espacio tridimensional.

$$\mathbf{x}_k = [\mathbf{p}_k, \quad \mathbf{v}_k, \quad \tilde{\mathbf{q}}_k, \quad \omega_k]^\top \in \mathbb{R}^{13}, \quad (5.13)$$

donde:

- $\mathbf{p}_k \in \mathbb{R}^3$ es la posición medida en $[m]$,
- $\mathbf{v}_k \in \mathbb{R}^3$ es la velocidad de traslación medida en $[m/s]$,
- $\tilde{\mathbf{q}}_k \in \mathbb{R}^4$ es el cuaternión de orientación medida en $[rad]$,
- $\omega_k \in \mathbb{R}^3$ es la velocidad de rotación medida en $[^\circ/s]$.

Los cuaterniones ofrecen una representación compacta para la orientación y evitan el problema de *gimbal lock*.

5.5.1. Modelo del proceso

De forma general, el estado del sistema evoluciona en pasos discretos de acuerdo a la siguiente relación:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k),$$

en la cual $f(\cdot)$ es la función de transición de estado, \mathbf{u}_k es la entrada de control y $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ es el ruido del proceso. La entrada de control es irrelevante en este trabajo.

El ruido del proceso se representa como un vector $\mathbf{w} \in \mathbb{R}^{12}$. Este se puede escribir en términos de cuatro vectores tridimensionales que causan disturbios en la posición, la velocidad, la orientación, y la velocidad angular, respectivamente:

$$\mathbf{w} = [\mathbf{w}_p, \mathbf{w}_v, \mathbf{w}_q, \mathbf{w}_\omega]. \quad (5.14)$$

Esta incertidumbre es tratada como ruido aditivo, excepto \mathbf{w}_q que es un vector tridimensional y no puede ser agregado al proceso de forma convencional.

El ruido \mathbf{w}_q es interpretado como un vector de rotación aleatorio, definido por:

$$\alpha_w = \|\mathbf{w}_q\|, \quad \mathbf{e}_w = \frac{\mathbf{w}_q}{\|\mathbf{w}_q\|}, \quad (5.15)$$

el cual permite obtener el cuaternión de ruido equivalente a \mathbf{w}_q como:

$$\mathbf{q}_w = \left[\cos\left(\frac{\alpha_w}{2}\right), \quad \mathbf{e}_w \cdot \sin\left(\frac{\alpha_w}{2}\right) \right]^\top. \quad (5.16)$$

Así, usando (??) el ruido \mathbf{w}_q es integrado al proceso mediante una multiplicación entre cuaterniones.

Finalmente, la dinámica del estado se describe con las ecuaciones (??), (??) y (??) aplicando los disturbios correspondientes, de acuerdo a:

$$\begin{bmatrix} \mathbf{p}_{k+1} \\ \mathbf{v}_{k+1} \\ \mathbf{q}_{k+1} \\ \omega_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{p}_k + \mathbf{v}_k T + \frac{1}{2} \mathbf{a}_k T^2 + \mathbf{w}_{p,k} \\ \mathbf{v}_k + \mathbf{a}_k T + \mathbf{w}_{v,k} \\ \mathbf{q}_k \odot \mathbf{q}_{w,k} \odot \mathbf{q}_\Delta \\ \omega_k + \mathbf{w}_{\omega,k} \end{bmatrix}, \quad (5.17)$$

la aceleración \mathbf{a} tratada como una entrada del modelo, esto permite que los cambios en la aceleración tengan un efecto más rápido en la evolución del estado.

La orientación expresada como cuaternión ofrece ventajas frente a otras alternativas, sin embargo, tiene dos inconvenientes. El primer inconveniente es que hace que el sistema sea no lineal; y el segundo inconveniente es que produce un desajuste en las dimensiones del estado-ruido, lo que impide que el ruido del proceso pueda ser modelado como ruido aditivo incumpliendo las restricciones del filtro de Kalman clásico. Por estas razones se seleccionó el filtro de Kalman Unscented, en contraste con otras alternativas de tiene un mejor rendimiento y no es tan demandante computacionalmente como se explicó en la sección 4.1.

5.5.2. Modelos de observación

El modelo de observación permite relacionar el estado con las mediciones en la fase de corrección, lo que da como resultado en una estimación de estado mejorado, dicho modelo se representa como:

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{v}_k),$$

en el cual $h(\cdot)$ es la función de observación y $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_z)$ representa ruido. Se parte de la premisa que los errores en las mediciones de los diferentes sensores son independientes, por lo que \mathbf{R}_z es una matriz diagonal.

La estimación del estado a priori se corrigió utilizando el acelerómetro, el giroscopio y la cámara, en consecuencia se emplearon tres modelos de observación.

El sistema de visión basado en marcas nos entrega la posición del vehículo \mathbf{p} , esta forma parte del estado, por lo cual para inferir las mediciones del sistema de visión se utiliza el siguiente modelo:

$$\mathbf{h}_c = \mathbf{p}_k + \mathbf{v}_p, \quad (5.18)$$

donde $\mathbf{v}_p \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_c)$ es el ruido en las mediciones de la posición del sistema de visión.

De forma análoga a lo que sucede con la posición, la velocidad angular que nos entrega el giroscopio se relaciona con el estado de acuerdo a:

$$\mathbf{h}_\omega = \omega_k + \mathbf{v}_\omega, \quad (5.19)$$

donde $\mathbf{v}_\omega \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_\omega)$ es el ruido en las mediciones del giroscopio.

Asumiendo que el vehículo se desplaza a una velocidad constante, las mediciones del acelerómetro se infieren utilizando el vector de gravedad, el cual es proyectado del marco de navegación al marco del cuerpo usando la última orientación estimada y añadiendo ruido:

$$\mathbf{h}_a = (\mathbf{q}_k)^{-1} \odot \tilde{\mathbf{g}} \odot \mathbf{q}_k + \mathbf{v}_a$$

donde $\mathbf{v}_a \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_a)$ es el ruido en las mediciones del acelerómetro y $\tilde{\mathbf{g}}$ es el vector de gravedad en forma de cuaternión.

5.6. Estimación del estado

El marco de trabajo del filtro de Kalman basa sus conceptos y criterios enfocados a operar sobre un espacio vectorial. Los cuaterniones no pertenecen a un espacio vectorial por lo que surgen problemas cuando se trabaja con esta representación de la orientación. Estas problemáticas se abordan en el trabajo de Kraft (2003), del cual fueron tomadas las ideas planteadas a continuación.

5.6.1. Puntos sigma

El problema de desajuste de la dimensión estado-ruido se solucionó agregando un paso extra para calcular los puntos sigma χ_i . El cual consiste en multiplicar el cuaternión de orientación por el cuaternión correspondiente al ruido:

$$\mathbf{q}_w = \left[\cos\left(\frac{\|\mathbf{w}_q\|}{2}\right), \frac{\mathbf{w}_q}{\|\mathbf{w}_q\|} \sin\left(\frac{\|\mathbf{w}_q\|}{2}\right) \right]^\top, \quad (5.20)$$

la variable \mathbf{w}_q es el error de la orientación respectivo a los ángulos roll, pitch y yaw.

El procedimiento para calcular los puntos sigma consiste en multiplicar cada fila de la matriz $\sqrt{\mathbf{P}_x}$ por $\pm\sqrt{2L}$ para formar el conjunto \mathbf{W} , siendo L la dimensión del estado disminuido en uno. Los elementos de este conjunto se pueden agrupar como:

$$\mathbf{W}_i = [\mathbf{w}_p, \mathbf{w}_v, \mathbf{w}_q, \mathbf{w}_\omega]^\top, \quad (5.21)$$

estas componentes representan los errores en la posición, la velocidad, la orientación y la velocidad angular, respectivamente.

Finalmente, los puntos sigma se generaron de acuerdo a:

$$\chi_i = \begin{bmatrix} \mathbf{p}_k + \mathbf{w}_p \\ \mathbf{v}_k + \mathbf{w}_v \\ \mathbf{q}_k \odot \mathbf{q}_w \\ \boldsymbol{\omega}_k + \mathbf{w}_\omega \end{bmatrix}. \quad (5.22)$$

5.6.2. Cálculo de la media

Los cuaterniones de orientación no pertenecen a un espacio vectorial. Por lo tanto, el cuaternión media no puede ser calculado de forma convencional, de no ser así podría originar graves problemas de inconsistencia en la orientación estimada Kraft (2003).

En este trabajo, el cuaternión media se calculó usando el enfoque propuesto por Markley et al. (2007), mismo que consiste en optimizar el siguiente indicador de rendimiento:

$$\bar{\mathbf{q}} = \underset{\tilde{\mathbf{q}} \in \mathcal{H}}{\operatorname{argmax}} \tilde{\mathbf{q}}^\top \mathbf{M} \tilde{\mathbf{q}}$$

con \mathbf{M} definida por:

$$\mathbf{M} = \sum_{i=1}^n w_i \tilde{\mathbf{q}}_i \tilde{\mathbf{q}}_i^\top,$$

cuya solución corresponde al vector propio de la matriz \mathbf{M} asociado al máximo valor propio.

5.6.3. Cálculo de la matriz de covarianzas

Dado dos cuaterniones \mathbf{q}^a y \mathbf{q}^b la rotación relativa que hace coincidir el primero con el segundo:

$$\mathbf{q}^b = {}^a_b\mathbf{q} \odot \mathbf{q}^a,$$

se define por (Kraft, 2003):

$${}^a_b\mathbf{q} = \mathbf{q}^b \odot (\mathbf{q}^a)^{-1}. \quad (5.23)$$

Usando un esquema similar, se puede escribir el cuaternión de error como la rotación relativa que lleva a los puntos sigma transformados \mathbf{q}_y al cuaternión media $\bar{\mathbf{q}}$:

$$\delta\mathbf{q} = \mathbf{q}_y \odot \bar{\mathbf{q}}^{-1}. \quad (5.24)$$

Así, el error en la estimación de la orientación se calcula como:

$$\mathbf{P}_\omega = \sum_{i=1}^{2L} \frac{1}{2L} \mathbf{w}_i \mathbf{w}_i^T + \mathbf{Q}_\omega, \quad (5.25)$$

en el cual \mathbf{w}_i es el vector de rotación equivalente al cuaternión de error $\delta\mathbf{q}$, definido por:

$$\begin{aligned} \theta &= 2 \cdot \arctan 2(|\tilde{\mathbf{q}}|, q_0), \\ \mathbf{u} &= \theta \cdot \frac{\tilde{\mathbf{q}}}{|\tilde{\mathbf{q}}|}. \end{aligned} \quad (5.26)$$

Capítulo 6

Experimentación y Resultados

El objetivo de este capítulo es presentar los experimentos y resultados con el propósito de validar el sistema mediante el desarrollo de la metodología descrita en el capítulo 5. Primero se brinda un breve descripción de las características del hardware y software utilizados en las pruebas. Posteriormente se presenta la metodología de experimentación que describe los escenarios de experimentación y la métrica usada para medir el nivel de error de la solución. Por último, se presentan los resultados del seguimiento de las trayectorias del vehículo.

6.1. Entorno de experimentación

Los algoritmos fueron implementados en Python3 y ejecutados sobre una computadora personal usando un conjunto de datos experimentales tomados con anterioridad, para su recolección se ha utilizado una placa Arduino equipada con una IMU modelo *BNO055*.

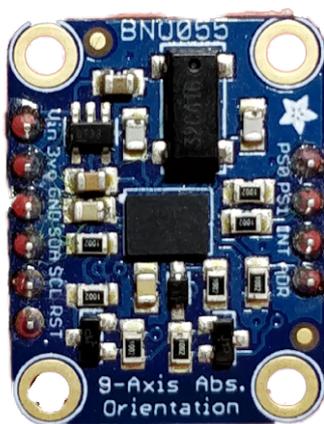


FIGURA 6.1: IMU modelo BNO055 de 9-DOF: empaquetada con tres acelerómetros, tres giroscopios y tres magnetómetros.

Esta IMU es adecuada para seguir el movimiento del vehículo en el espacio tridimensional ya que incluye tres acelerómetros, tres giroscopios y tres magnetómetros que permiten medir las magnitudes físicas que actúan sobre el vehículo a lo largo del eje-x, eje-y y eje-z, en la tabla 6.1 se describen las características de las señales de salida de los sensores.

TABLA 6.1: Características de las señales de salida de la IMU.

Tipo de medición	Unidades	Frecuencia de muestreo
Aceleración	$[m/s^2]$	100 Hz
Velocidad angular	$[^\circ/s]$	100 Hz
Campo magnético	$[\mu T]$	20 Hz

El dispositivo se configuró para operar en modo NDOF, el cual fusiona los tres sensores para alcanzar una alta robustez de las distorsiones del campo magnético, dando como resultado una rápida calibración del magnetómetro y alta precisión en los datos de salida (BOSCH, 2016).



FIGURA 6.2: Cámara modelo Logitech.

Por otro parte, en el sistema de visión se utilizó una cámara modelo Logitech como la que se ilustra en la figura 6.2. Esta cámara cuenta con un sensor de 2 Mega píxeles que fue configurada para trabajar con una resolución de 1600×1200 puntos a 10 FPS (*Frames Per Second*).

6.2. Metodología de experimentación

La idea principal de las pruebas consisten en capturar datos de la dinámica del vehículo y ejecutar el algoritmo en un momento posterior con los datos recopilados. Siguiendo este esquema, se recolectaron datos en modo estacionario y en modo dinámico para evaluar las cualidades de la solución.

Modo estacionario

Este escenario está diseñado para evaluar la tasa de crecimiento del error inherente a los sistemas de navegación inercial. La prueba consiste en mantener al vehículo en estado de reposo mientras se recolectan datos de la IMU durante un lapso de aproximadamente 60 segundos.

Modo dinámico

Este escenario está diseñado para evaluar la certeza del movimiento del vehículo sobre una trayectoria rectangular como se ilustra en la figura 6.3. El rectángulo gris simboliza la trayectoria de referencia la cual se describe por cuatro vértices \mathbf{A}_1 , \mathbf{A}_2 , \mathbf{A}_3 y \mathbf{A}_4 que yacen en el plano XY , cuyos valores serán definidos posteriormente. En cada prueba el movimiento inicial y finaliza en el punto denotado por $\mathbf{A}_0 = [0, 0]^T$, expresado en metros.

Para recolectar los datos experimentales, se posicionó el vehículo en \mathbf{A}_0 y permaneciendo en estado estacionario se tomaron 100 muestras de la aceleración y campo magnético local que fueron usadas para alinear el vehículo. Inmediatamente después, se inicio el movimiento recorriendo la trayectoria de referencia en el sentido contrario de las manecillas del reloj, siguiendo la dirección que indican las flechas azules, hasta llegar de nuevo a la posición \mathbf{A}_0 completando un circuito cerrado. Durante el desplazamiento las mediciones de los sensores en conjunto con el tiempo fueron trasferidos a la computadora para su almacenamiento.

6.2.1. Evaluación

Las pruebas se llevaron a cabo en entornos reales, una de las limitaciones fue que no se cuenta con una plataforma que proporcione información de la posición y la orientación del movimiento del vehículo

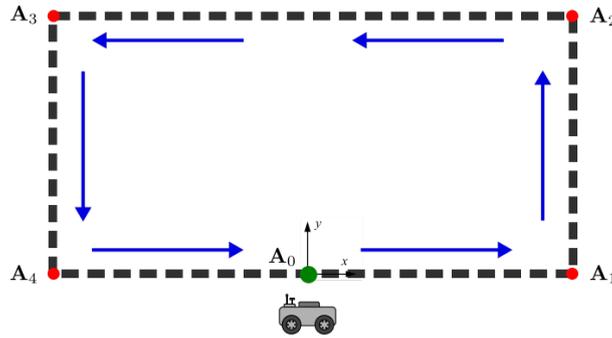


FIGURA 6.3: Escenario usado para evaluar la solución del algoritmo de estimación de pose: las líneas punteadas representan la trayectoria de referencia; las flechas azules indican el sentido del movimiento; la marca de color verde representa la coordenada donde inicia y finaliza el movimiento.

que sirva de guía o referencia para hacer un análisis cuantitativo del error. Por lo tanto, las evaluaciones se centran principalmente en la congruencia de los resultados y el movimiento que le dio origen.

Particularmente, en la orientación es posible hacer un análisis numérico del error ya que los ángulos roll(ϕ) y pitch(θ) mantienen una referencia a cero durante toda la trayectoria, el ángulo yaw(ψ) mantienen un valor constante por intervalos, salvo en las cuatro coordenadas (\mathbf{A}_1 , \mathbf{A}_2 , \mathbf{A}_3 y \mathbf{A}_4) donde ocurren los cambios rumbo ($0^\circ \rightarrow 90^\circ$, $90^\circ \rightarrow 180^\circ$, $180^\circ \rightarrow 270^\circ$ y $270^\circ \rightarrow 0^\circ$).

Así, para evaluar la orientación se ha determinado usar la RMSE, la métrica RMSE mide error promedio del grado de correspondencia entre pares de valores predichos por un estimador y sus correspondientes valores reales, definida como:

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(y_i - \hat{y}_i)^2}{N}}, \quad (6.1)$$

donde:

y es el valor real,

\hat{y} es el valor predicho, y

N es el número de valores analizados.

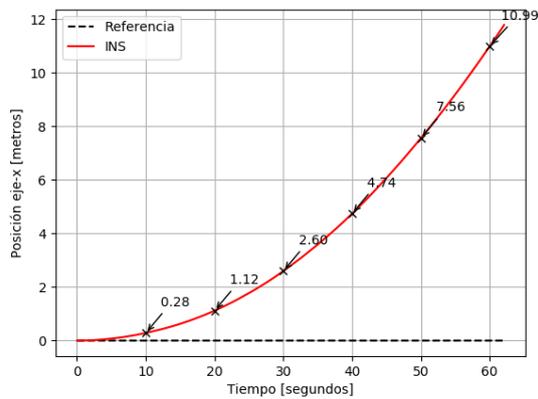
6.3. Resultados

En este apartado se presenta los resultados obtenidos en un entorno interior, con el fin de resaltar las características de cada una de las partes que componen el sistema se diseñaron tres experimentos para su validación. En el primero se valida el sistema de navegación inercial, en modo estacionario. El segundo tiene como propósito validar el sistema de estimación de pose basado en visión. Finalmente, el tercer experimento presenta los resultado de la validación del sistema basado en el UKF. En los últimos dos experimentos el sistema se evaluó usando el modo dinámico, descrito anteriormente.

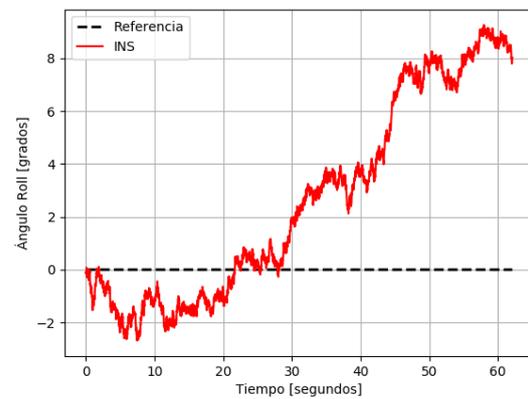
6.3.1. Experimento 1

En este experimento se valida el sistema de navegación inercial en modo estacionario. La figura 6.4 muestra los resultados de la posición y orientación para cada eje. En los resultados del experimento no se usa ningún tipo de corrección en la estimación ya que la intención es mostrar el comportamiento de este tipo de sistemas.

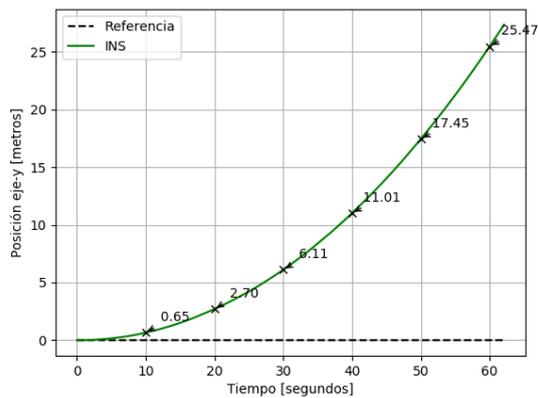
El aspecto más interesante de estas gráficas es que ilustran el comportamiento característicos de los sistemas de navegación inercial. Como muestran los resultados, la deriva incrementa rápidamente debido a los errores presentes en las mediciones de los sensores, descritos en la sección 2.4.1. Por lo tanto, un enfoque basado netamente en la navegación inercial no es aplicable para seguir el movimiento del robot con la precisión esperada.



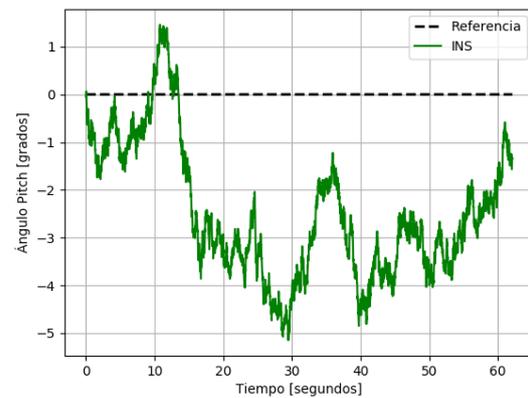
(A) Posición eje-x.



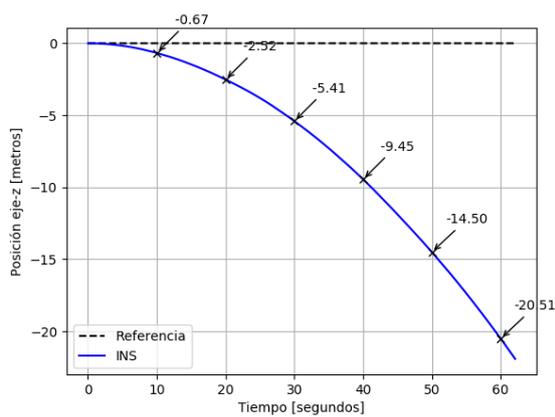
(B) Ángulo roll.



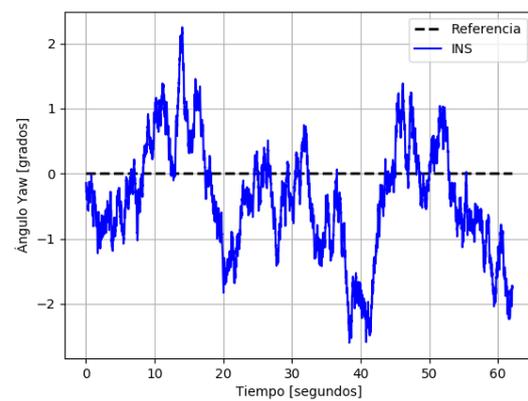
(C) Posición eje-y.



(D) Ángulo pitch.



(E) Posición eje-z.



(F) Ángulo yaw.

FIGURA 6.4: Estimación de la pose del vehículo usando el algoritmo de navegación inercial en modo estacionario: los resultados muestran la degeneración de la solución en el tiempo debido a la acumulación de errores.

6.3.2. Experimento 2

El propósito de este experimento es validar el sistema de estimación de pose basado en visión, para ello, se usa el modo dinámico siguiendo las pautas descritas anteriormente. Para este experimento se usó el escenario que se ilustra en la figura 6.5, donde las coordenadas cartesianas 2- D de los vértices tienen los siguientes valores, expresados en metros:

$$\begin{aligned}\mathbf{A}_1 &= [0.68, 0.0]^T, \\ \mathbf{A}_2 &= [0.68, 0.76]^T, \\ \mathbf{A}_3 &= [-0.68, 0.76]^T, \\ \mathbf{A}_4 &= [-0.68, 0.0]^T.\end{aligned}$$

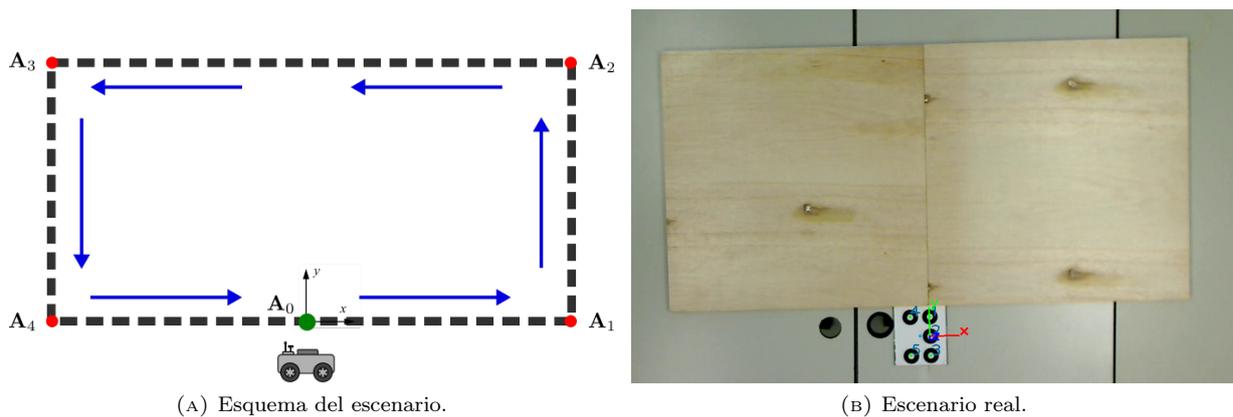


FIGURA 6.5: Escenario usado para la estimación de la pose basada en visión: El escenario real (izquierda) consta de una forma rectangular de 1.36×0.76 metros, dispuesta sobre una mesa de tal manera que quede dentro del campo visual de la cámara externa, a la izquierda su representación simbólica.

La figura 6.6 presenta los resultados de la posición, obtenidos a distintas velocidades (0.23 m/s y 2.36 m/s), las velocidades reportadas representan la velocidad promedio calculada a partir de los cambios en posición en los tramos rectos donde se asume que la velocidad es constante. En ambos casos, los resultados concuerdan con los movimientos del vehículo.

Los resultados muestran que el sistema no presenta divergencia en la estimación, sin embargo, la frecuencia de muestreo disminuye conforme aumenta la velocidad. Por otro lado, los resultados de la estimación de la orientación se muestran en la figura 6.7, tal y como se esperaba, los resultados tampoco presentan divergencia. En la figura 6.7b se muestran el error para cada ángulo, como se puede observar de los datos de la gráfica, los ángulos roll y pitch son los más afectados por el ruido, aunque en todos los ángulos el error es inferior a los 4 grados.

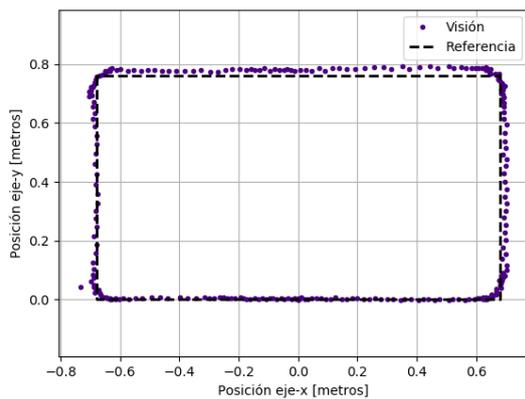
A pesar de los inconvenientes, el sistema basado en visión es adecuado para funcionar como guía del sistema del INS, consolidando así un sistema híbrido capaz de ofrecer una estimación de pose más robusta.

6.3.3. Experimento 3

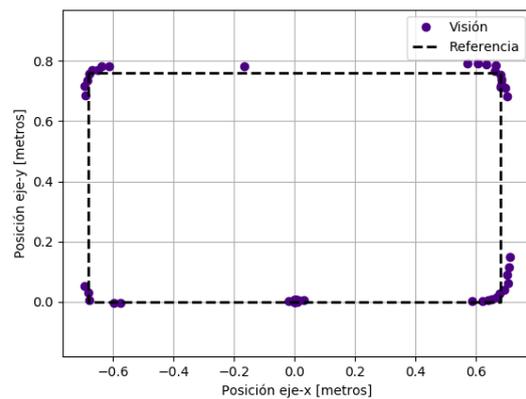
Para validar el sistema de estimación de pose basado en el UKF se siguió nuevamente el procedimiento descrito en el modo dinámico, para esto, se usó el escenario mostrado en la figura 6.8 con los siguientes valores de coordenadas cartesianas 2- D , expresados en metros:

$$\begin{aligned}\mathbf{A}_1 &= [1.72, 0.0]^T, \\ \mathbf{A}_2 &= [1.72, 1.30]^T, \\ \mathbf{A}_3 &= [-0.20, 1.30]^T, \\ \mathbf{A}_4 &= [-0.20, 0.0]^T.\end{aligned}$$

En la figura 6.9 se presentan los resultados para la estimación de la posición usando el INS y el UKF. Al comparar ambos métodos, se puede observar un mejor rendimiento por parte del UKF, aunque no

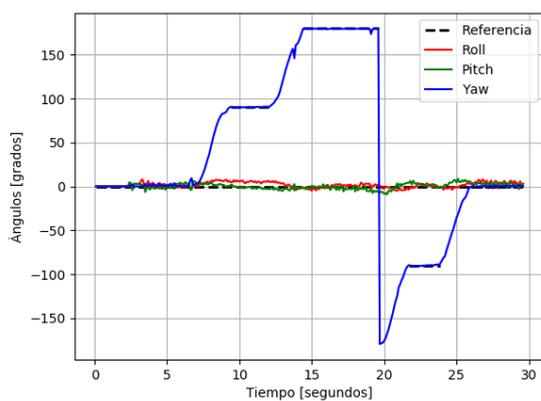


(A) Posición con velocidad promedio de 0.23 m/s.

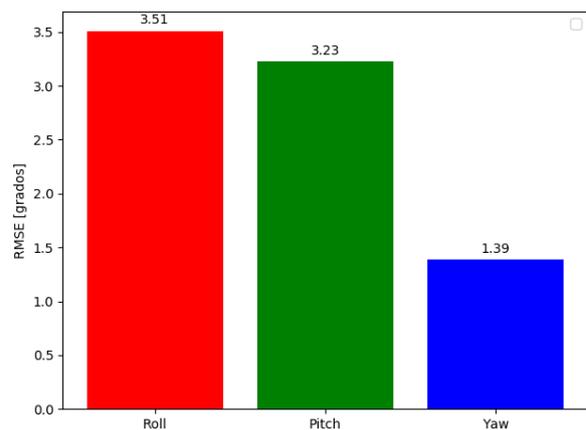


(B) Posición con velocidad promedio de 2.36 m/s.

FIGURA 6.6: Estimación de la posición basada en visión: la velocidad promedio se calcula mediante la posición asumiendo que la velocidad es constante en las secciones rectas de la trayectoria, descartando el movimiento en las curvas donde ocurren las transiciones del cambio de rumbo.

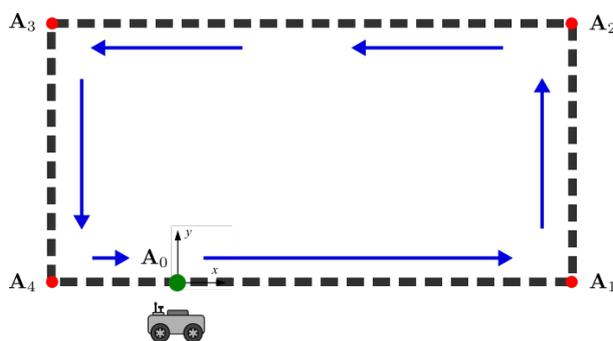


(A) Orientación.



(B) Error en la orientación.

FIGURA 6.7: Estimación de la orientación basada en visión: en la figura 6.7a se muestra la orientación en términos de los ángulos roll, pitch y yaw, y en la figura 6.7b se presentan sus correspondientes errores de estimación usando la raíz del error cuadrático medio.



(A) Esquema del escenario.



(B) Escenario real.

FIGURA 6.8: Disposición del entorno usado para la estimación de la pose basada en el UKF: la figura 6.8a muestra la representación simbólica del escenario y la figura 6.8b ilustra el escenario real, el movimiento se efectuó siguiendo como guía el borde de la mesa con dimensiones 1.92×1.30 metros.

es muy notable la corrección de la posición se aprecia claramente es esfuerzo en la corrección del rumbo del vehículo, esto se confirma con los datos de figura 6.10 donde se puede observar la convergencia de la orientación con un error inferior a los 3 grados para cada ángulo.

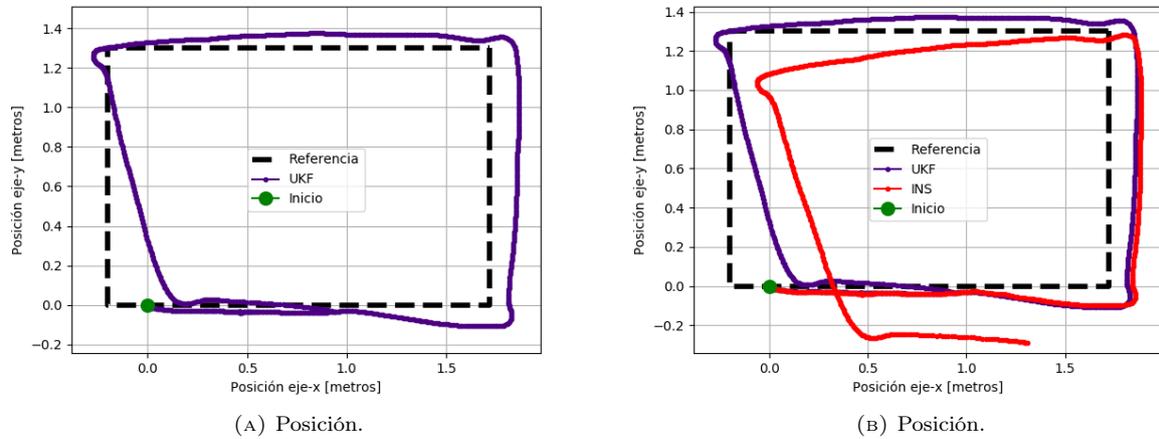


FIGURA 6.9: Estimación de posición basada en el UKF: la figura 6.9a muestra los resultados del UKF y la figura 6.9b muestra una comparativa de la solución entre el INS y el UKF.

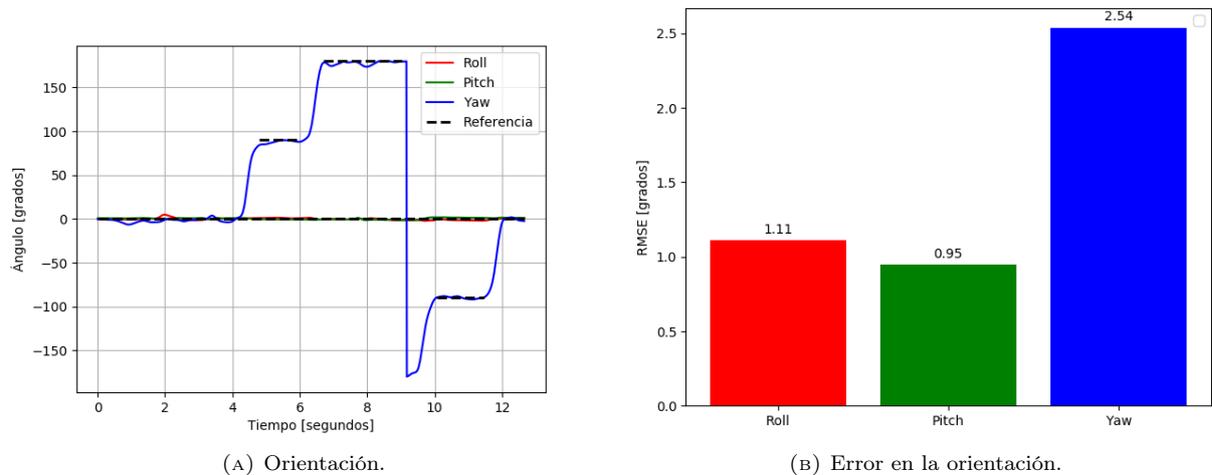


FIGURA 6.10: Estimación de la orientación basada en el UKF: la figura 6.10a muestra la orientación en términos de los ángulos roll, pitch y yaw, y en la figura 6.10b se presentan los errores para cada ángulo usando la métrica RMSE.

6.4. Conclusiones

A lo largo del desarrollo de este capítulo se han presentado los resultados obtenidos para la estimación de pose, de acuerdo a lo expuesto en los resultados podemos resaltar la importancia de un sistema híbrido visual-inercial para mejorar la robustez de la estimación, aprovechando sus características complementarias.

Los sistemas de navegación inercial son sistemas autocontenidos capaces de proveer información de la dinámica del vehículo a altas velocidades y son precisos por periodos de tiempo cortos, sin embargo, los errores numéricos involucrados el proceso de estimación y los errores característicos de los dispositivos de bajo costo provocan un comportamiento caótico, dejando prácticamente inutilizable el sistema. Por muy pequeños que sean los errores, debido a la integración, el error de deriva crece ilimitadamente, para

evitar esta situación es crucial utilizar un sistema de referencia externo o conocimiento de la dinámica para restablecer continuamente el sistema.

Por otro lado, los sistemas basados en visión pueden operar por un periodo de tiempo largo sin degenerar su rendimiento, a cambio de un costo computacional elevado. Aunque son muchos los factores que afectan su estabilidad, brindan mejores resultados siguiendo movimientos relativamente lentos con periodo de latencia superior en comparación con los sistemas de navegación inercial.

Los movimientos rápidos normalmente producen imágenes borrosas que hacen que el sistema prolongue el tiempo de respuesta hasta alcanzar las condiciones adecuadas para estimar la pose nuevamente. Este inconveniente se puede solventar usando el INS para rellenar los espacios entre muestras del sistema basado en visión.

Es evidente que ambos sistemas se pueden complementar para alcanzar una estimación robusta de la pose, en la práctica la deriva es casi inevitable en periodos de tiempo prolongados, por lo tanto, es crucial implementar otros mecanismo para incrementar el nivel de confianza de la estimación.

Específicamente, se ha utilizado el UKF que modela el sistema en términos de distribuciones de probabilidad para dar una estimación más certera a partir de información proveniente de diferentes fuentes, ponderadas de acuerdo al nivel de confianza que se tiene de cada una de ellas. Las mediciones de la IMU se han utilizado para propagar el estado y se ha propuesto el sistema de visión para corregirlo. Sin embargo, uno de los principales problemas fue sincronizar las mediciones que permitan corregir el estado de forma adecuada, así que para corregir se ha utilizado únicamente la IMU.

Los resultados muestran que el sistema basado en el UKF ofrece una solución aceptable, que no se podría alcanzar únicamente utilizando el INS ya que un error en la orientación de 1 mili-radian producirá un error de compensación de 1 mili-g en los acelerómetros, en consecuencia, la posición habrían derivado 4.5 metros después de solo 30 segundos (Welch y Foxlin, 2002).

Bibliografía

- BAY, H., ESS, A., TUYTELAARS, T. y VAN GOOL, L. Speeded-up robust features (surf). *Computer vision and image understanding*, vol. 110(3), páginas 346–359, 2008.
- BORENSTEIN, J., EVERETT, H. R., FENG, L. y WEHE, D. Mobile robot positioning: Sensors and techniques. *Journal of robotic systems*, vol. 14(4), páginas 231–249, 1997.
- BOSCH. *BNO055 data sheet*, 2016. Rev. 1.4.
- BROWN, B. The social life of autonomous cars. *Computer*, vol. 50(2), páginas 92–96, 2017.
- CADENA, C., CARLONE, L., CARRILLO, H., LATIF, Y., SCARAMUZZA, D., NEIRA, J., REID, I. y LEONARD, J. J. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics*, vol. 32(6), páginas 1309–1332, 2016. ISSN 1941-0468.
- CALONDER, M., LEPETIT, V., STRECHA, C. y FUA, P. Brief: Binary robust independent elementary features. En *European conference on computer vision*, páginas 778–792. Springer, 2010.
- CORKE, P., LOBO, J. y DIAS, J. An introduction to inertial and visual sensing. 2007.
- DESOUZA, G. N. y KAK, A. C. Vision for mobile robot navigation: A survey. *IEEE transactions on pattern analysis and machine intelligence*, vol. 24(2), páginas 237–267, 2002.
- DURRANT-WHYTE, H. y BAILEY, T. Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, vol. 13(2), páginas 99–110, 2006.
- FISCHLER, M. A. y BOLLES, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, vol. 24(6), páginas 381–395, 1981.
- GREWAL, M. S. y ANDREWS, A. P. Applications of kalman filtering in aerospace 1960 to the present [historical perspectives]. *IEEE Control Systems Magazine*, vol. 30(3), páginas 69–78, 2010. ISSN 1066-033X.
- GUSTAFSSON, F., GUNNARSSON, F., BERGMAN, N., FORSELL, U., JANSSON, J., KARLSSON, R. y NORDLUND, P. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on Signal Processing*, vol. 50(2), páginas 425–437, 2002. ISSN 1053-587X.
- HARRIS, C. G., STEPHENS, M. ET AL. A combined corner and edge detector. En *Alvey vision conference*, vol. 15, páginas 10–5244. Citeseer, 1988.
- HESCH, J. A. y ROUMELIOTIS, S. I. A direct least-squares (dls) method for pnp. En *2011 International Conference on Computer Vision*, páginas 383–390. 2011. ISSN 2380-7504.
- HORAUD, R., CONIO, B., LEBoulLEUX, O. y LACOLLE, B. An analytic solution for the perspective 4-point problem. En *Proceedings CVPR '89: IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, páginas 500–507. 1989. ISSN 1063-6919.
- HUAT, L. K. *Industrial robotics: programming, simulation and applications*. 2006.
- IWANECZKO, P., JĘDRASIAK, K. y NAWRAT, A. Indoor navigation with micro inertial navigation technology. En *Advanced Technologies in Practical Applications for National Security*, páginas 337–346. Springer, 2018.

- JAZWINSKI, A. H. *Stochastic processes and filtering theory*. Courier Corporation, 2007.
- JUL, S. y FURNAS, G. W. Navigation in electronic worlds: a chi 97 workshop. *SIGCHI bulletin*, vol. 29, páginas 44–49, 1997.
- JULIER, S. J. y UHLMANN, J. K. New extension of the kalman filter to nonlinear systems. En *Signal processing, sensor fusion, and target recognition VI*, vol. 3068, páginas 182–194. International Society for Optics and Photonics, 1997.
- JULIER, S. J. y UHLMANN, J. K. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, vol. 92(3), páginas 401–422, 2004.
- JULIER, S. J., UHLMANN, J. K. y DURRANT-WHYTE, H. F. A new approach for filtering nonlinear systems. En *Proceedings of 1995 American Control Conference - ACC'95*, vol. 3, páginas 1628–1632 vol.3. 1995.
- KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, vol. 82(1), páginas 35–45, 1960.
- KANDEPU, R., FOSS, B. y IMSLAND, L. Applying the unscented kalman filter for nonlinear state estimation. *Journal of process control*, vol. 18(7-8), páginas 753–768, 2008.
- KELLY, A. *Mobile robotics: mathematics, models, and methods*. Cambridge University Press, 2013.
- KOK, M., HOL, J. D. y SCHÖN, T. B. Using inertial sensors for position and orientation estimation. *arXiv preprint arXiv:1704.06053*, 2017.
- KRAFT, E. A quaternion-based unscented kalman filter for orientation tracking. En *Proceedings of the Sixth International Conference of Information Fusion*, vol. 1, páginas 47–54. 2003.
- LEPETIT, V., MORENO-NOGUER, F. y FUA, P. Epnnp: An accurate $o(n)$ solution to the pnp problem. *International Journal of Computer Vision*, vol. 81(2), página 55, 2008. ISSN 1573-1405.
- LOWE, D. G. ET AL. Object recognition from local scale-invariant features. En *iccv*, vol. 99, páginas 1150–1157. 1999.
- LOZANO-PEREZ, T. *Autonomous robot vehicles*. Springer Science & Business Media, 2012.
- MA, Y., SOATTO, S., KOSECKA, J. y SASTRY, S. S. *An invitation to 3-d vision: from images to geometric models*, vol. 26. Springer Science & Business Media, 2012.
- MARKLEY, F. L., CHENG, Y., CRASSIDIS, J. L. y OSHMAN, Y. Averaging quaternions. *Journal of Guidance, Control, and Dynamics*, vol. 30(4), páginas 1193–1197, 2007.
- MARR, D. *Vision*. san francisco: W. h. H. Freeman, 1982.
- MURPHY, R. R. *Introduction to AI Robotics*. MIT Press, Cambridge, MA, USA, 1st edición, 2000. ISBN 0262133830.
- OJANEN, H. Automatic correction of lens distortion by using digital image processing. *Rutgers University, Dept. of Mathematics technical report*, 1999.
- ORR, M. J. *A Mickey Mouse guide to Kalman filtering*. Edinburgh University, 1992.
- REIF, K., GUNTHER, S., YAZ, E. y UNBEHAUEN, R. Stochastic stability of the discrete-time extended kalman filter. *IEEE Transactions on Automatic control*, vol. 44(4), páginas 714–728, 1999.
- SHUSTER, M. D. y OH, S. D. Three-axis attitude determination from vector observations. *Journal of Guidance, Control, and Dynamics*, vol. 4(1), páginas 70–77, 1981.
- THRUN, S., BURGARD, W. y FOX, D. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005. ISBN 0262201623.
- TITTERTON, D., WESTON, J. L. y WESTON, J. *Strapdown inertial navigation technology*, vol. 17. IET, 2004.

- WAHBA, G. A least squares estimate of satellite attitude. *SIAM review*, vol. 7(3), páginas 409–409, 1965.
- WAN, E. A. y VAN DER MERWE, R. The unscented kalman filter for nonlinear estimation. En *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, páginas 153–158. Ieee, 2000.
- WAN, E. A. y VAN DER MERWE, R. The unscented kalman filter. *Kalman filtering and neural networks*, páginas 221–280, 2001.
- WANDELL, B. A. *Foundations of vision*, vol. 8. Sinauer Associates Sunderland, MA, 1995.
- WANG, Y., ZHANG, W. y AN, P. A survey of simultaneous localization and mapping on unstructured lunar complex environment. En *AIP Conference Proceedings*, vol. 1890, página 030010. AIP Publishing, 2017.
- WELCH, G., BISHOP, G. ET AL. An introduction to the kalman filter. 1995.
- WELCH, G. y FOXLIN, E. Motion tracking: no silver bullet, but a respectable arsenal. *IEEE Computer Graphics and Applications*, vol. 22(6), páginas 24–38, 2002. ISSN 1558-1756.
- WOODMAN, O. J. An introduction to inertial navigation. Informe técnico, University of Cambridge, Computer Laboratory, 2007.
- XU, G. y ZHANG, Z. *Epipolar geometry in stereo, motion and object recognition: a unified approach*, vol. 6. Springer Science & Business Media, 2013.
- YUN, X., BACHMANN, E. R., MOORE, H. y CALUSDIAN, J. Self-contained position tracking of human movement using small inertial/magnetic sensor modules. En *Proceedings 2007 IEEE International Conference on Robotics and Automation*, páginas 2526–2533. 2007. ISSN 1050-4729.
- ZAMALLOA, I., KOJCEV, R., HERNÁNDEZ, A., MUGURUZA, I., USATEGUI, L., BILBAO, A. y MAYORAL, V. Dissecting robotics-historical overview and future perspectives. *arXiv preprint arXiv:1704.08617*, 2017.