

MÉTODO PARA EL SEGUIMIENTO DE BASTONES  
CONCURRENTES UTILIZANDO EL FILTRO DE KALMAN

RUBÉN VARGAS PALMA



Maestría en Ciencias de la Computación  
Facultad de Matemáticas  
Universidad Autónoma de Yucatán



## RESUMEN

---

En el presente trabajo se desarrolla una aplicación para el seguimiento tridimensional de bastones con el fin de demostrar las propiedades de operaciones vectoriales utilizando realidad aumentada. En un trabajo anterior realizado por Enrique Solis se efectuaba la localización tridimensional de bastones en cada cuadro, esto sin tomar en cuenta un modelo de movimiento. Disponer de información sobre el movimiento de los bastones permite realizar el seguimiento aun cuando no sea posible su localización mediante el uso de las cámaras estéreo.

Se trabajó bajo el supuesto de que siempre existen solo dos bastones en la escena y son cada uno de diferente color. Para detectar y localizar los bastones se generó un modelo de color por bastón, luego para cada imagen capturada se detectó la presencia de bastones en la escena utilizando el modelo de color y un vector de rasgos. Con ayuda de el sistema estereoscopio se hizo la localización 3D, esta recta es utilizada por un del filtro de Kalman para estimar las velocidades lineal y angular. Cuando no es posible la localización 3D de la recta los parámetros de velocidad nos permiten realizar predicciones.

En este trabajo se realizaron experimentos utilizando dos representaciones de la recta en el espacio 3D, una infinita utilizando coordenadas de Plücker y otra mediante el punto medio del segmento de recta formado por el bastón y la orientación. Esta ultima representación también se empleo para dar seguimiento a un plano formado por los dos bastones, cambiando el punto medio por el punto de intersección de ambas y la orientación por la normal del plano, asumiendo que ambos bastones forman parte de un solo cuerpo rígido.

Quedó demostrado que utilizar un filtro de Kalman para estimar los parámetros de movimiento cinemático permite realizar un seguimiento mas robusto, puesto que se pueden tratar casos de especiales de aparente colinealidad.



## DECLARACIÓN

---

Por este medio, declaro que yo escribí esta tesis, y que describe el trabajo de mi tesis de Maestría en Ciencias de la Computación.

Rubén Vargas Palma  
*Mérida, Yucatán,*  
*México*  
*Septiembre 2016*



## AGRADECIMIENTOS

---

Agradezco a las siguientes personas, quienes contribuyeron de una forma u otra al desarrollo de la presente tesis:

A mis padres, José Rubén Vargas y Zeidy Irasema Palma y a mi hermana Pamela, por su incondicional apoyo para superarme cada día.

Al Dr. Arturo Espinosa Romero, por su paciencia, apoyo y consejos que me ha brindado a lo largo de mi formación profesional

A mis profesores por todos los conocimientos que con dedicación me transmitieron.

A mis amigos y compañeros, por acompañarme en esta etapa.

Al Consejo Nacional de Ciencias y Tecnología (CONACYT), que me otorgó la beca con la cual pude sustentarme económicamente durante el estudio de la maestría.



# ÍNDICE GENERAL

---

1	INTRODUCCIÓN	1
1.1	Introducción	1
1.2	Objetivos	4
1.2.1	Objetivo General	4
1.2.2	Objetivos particulares	5
1.3	Estructura de la tesis	5
2	MARCO TEÓRICO	7
2.1	Introducción	7
2.2	Coordenadas Cartesianas	8
2.3	Coordenadas Homogeneas	8
2.3.1	Lineas	9
2.3.2	Puntos	10
2.3.3	Coordenadas de Plücker	11
2.4	Espacios de color	11
2.4.1	Espacio CIELab	12
2.5	Distancia de Mahalanobis	15
2.6	Movimiento de un cuerpo rígido	17
2.6.1	Movimiento Rotacional	19
2.6.2	Movimiento general de un cuerpo rígido	21
2.7	Filtro de Kalman	23
2.7.1	Etapa de predicción	24
2.7.2	Etapa de actualización	26
2.7.3	Algoritmo	30
2.7.4	Filtro de Kalman Extendido	30
3	METODOLOGÍA	33
3.1	Introducción	33
3.2	Segmentación Robusta	34
3.3	Reconstrucción 3D de la recta	35
3.4	Estimación de Parámetros de Movimiento	36
3.4.1	Coordenadas de Plücker	37
3.4.2	Plano y punto de intersección	38
3.5	Resumen	40

4	EXPERIMENTOS	43
4.1	Introducción	43
4.2	Simulaciones	44
4.2.1	Generación de Movimiento	44
4.2.2	Aproximación Inicial	46
4.2.3	Representación de Plücker	46
4.2.4	Representación con punto medio y orientación	53
4.3	Experimentos con cámaras	59
5	CONCLUSIÓN	61
5.1	Contribuciones del proyecto	62
5.2	Trabajo futuro	62
i	APÉNDICE	63
A	MATRIZ JACOBIANA DEL MODELO CON REPRESENTACIÓN PLÜCKER	65
A.1	Matriz Jacobiana con respecto a la medición	65
A.2	Matriz Jacobiana con respecto a la estado	67
	BIBLIOGRAFÍA	69

## ÍNDICE DE FIGURAS

---

Figura 1	En este trabajo los referentes físicos utilizados fueron las manos, y se empleo un dispositivo Kinect para la localización y seguimiento de dichos referentes. 2
Figura 2	Sistema de realidad aumentada utilizando cámaras estéreo y referentes físicos para manipular los vectores. 3
Figura 3	Capturas de pares de imágenes del sistema funcionando de forma inestable, en la cámara de la derecha aparece una condición de colinealidad, bajo este caso ya no es posible realizar la localización de la recta. 3
Figura 4	Coordenadas Cartesianas. 8
Figura 5	Respuesta de los conos según la longitud de onda. 12
Figura 6	Color matching functions. 13
Figura 7	Color matching functions sin valores negativos para el modelo CIExyz. 13
Figura 8	Dos datos en marcados en rojo, $d_1$ y $d_2$ aparentemente $d_1$ esta mas cerca de la media, esto sin tomar en cuenta la distribución de los datos. 15
Figura 9	Normalización y desnormalización de datos 16
Figura 10	Movimiento general de un cuerpo rígido asociado a un sistema coordenado C con respecto al sistema coordenado del mundo W. 22

- Figura 11 En esta figura se pueden observar las dos etapas del filtro de Kalman, la  $P_{t-1}$  y la  $X_{t-1}$  corresponden a la estimación del estado  $X$  con una matriz de covarianza  $P$  que representa la incertidumbre, la  $X_t$  y la  $P_t$  son la predicción realizada en base a la matriz de predicción  $F_t$ , luego en la etapa de actualización se obtiene una nueva medición y tomando en cuenta el ruido en la misma se obtienen nuevas estimaciones. 25
- Figura 12 Predicción de la distribución Gaussiana de  $x_{t-1}$ . 26
- Figura 13 Transformación del espacio de estados al espacio de las mediciones. 27
- Figura 14 Distribución Gaussiana de la medición. 28
- Figura 15 Distribución Gaussiana evaluada en una función lineal. 31
- Figura 16 Distribución Gaussiana evaluada en una función no lineal. 31
- Figura 17 Diagrama Metodología 33
- Figura 18 Sistema de camaras estereo. 36
- Figura 19 Recta en el tiempo  $t-1$  y en el tiempo  $t$ . 37
- Figura 20 Efecto de la velocidad lineal sobre el vector normal. 38
- Figura 21 Movimiento simulado de un bastón, la curva verde representa la trayectoria. 45
- Figura 22 Tres mediciones que se tomaron para hacer la estimación inicial. 47
- Figura 23 Gráficas de error de los componentes del vector de velocidad, utilizando la representación de recta de Plücker. 48
- Figura 24 Gráfica de la diferencia entre la velocidad angular real de la simulación y la estimada por el filtro de Kalman, el área verde representa  $3\sigma$  desviaciones estándar, se observa que la mayoría de los errores caen en ese rango. Para la generación de esta gráfica se empleó la representación de recta de Plücker. 49

- Figura 25 Gráfica de error en la magnitud de la velocidad de rotación para la representación de recta utilizando las coordenadas de Plücker. 50
- Figura 26 Gráfica del error en la estimación del eje de rotación, ángulo entre el vector original de rotación y el estimado, utilizando la representación de recta de Plücker.. 51
- Figura 27 Gráfica de la diferencia entre la velocidad lineal real de la simulación y la estimada por el filtro de Kalman, el área verde representa  $3\sigma$  desviaciones estándar, se observa que la mayoría de los errores caen en ese rango. Para la generación de esta gráfica se empleó la representación de recta de Plücker. 52
- Figura 28 Gráfica del error en la estimación del eje de rotación, ángulo entre el vector original de rotación y el estimado, utilizando el punto medio y la orientación de un segmento de recta. 54
- Figura 29 Gráfica de error en la magnitud de la velocidad de rotación para la representación de recta utilizando la representación de segmento de recta, con la orientación y el punto medio. 55
- Figura 30 Gráficas de error, arriba el error en la pose en el momento de realizar la predicción, abajo el error respecto al estado real, utilizando el punto medio y la orientación de un segmento de recta. 56
- Figura 31 Gráficas de error, arriba el error en la pose en el momento de realizar la predicción, abajo el error respecto al estado real, utilizando el punto medio y la orientación de un segmento de recta. 57
- Figura 32 Gráficas de error, arriba el error en la pose en el momento de realizar la predicción, abajo el error respecto al estado real, utilizando el punto medio y la orientación de un segmento de recta. 58

- Figura 33 Sistema funcionando en condiciones de paralelismo en la cámara izquierda, arriba el sistema sin filtro de Kalman, abajo el sistema funcionando con la estimación del movimiento. 59
- Figura 34 Sistema de cámaras estéreo utilizado para los experimentos. 60

## INTRODUCCIÓN

---

### 1.1 INTRODUCCIÓN

La realidad aumentada permite a los usuarios ver el mundo real con objetos virtuales superpuestos a éste, no reemplaza la realidad por completo, sino que añade mas información a la misma. Esta área está en un constante desarrollo, se han encontrado múltiples aplicaciones en áreas de la medicina, la ingeniería y el entretenimiento [2].

Una aplicación que se le puede dar a la realidad aumentada es en el área de educación, particularmente en la enseñanza de las matemáticas. Mediante esta técnica es posible mostrar visualmente conceptos abstractos como los vectores y sus diferentes operaciones, con el fin de facilitar la comprensión del estudiante en estos temas. Se han realizado dos trabajos al respecto en la Facultad de Matemáticas de la Universidad Autónoma de Yucatán, uno realizado por Chi Pot[4] y el otro por Enrique Solis[13].

En el trabajo realizado por Chi Pot[4] se utilizó un sensor Kinect de Microsoft® el cual es capaz de realizar seguimiento el movimiento del cuerpo humano. En [4] se utilizaron las manos como referentes físicos de los vectores y el torso como punto común, ese sistema es capaz de representar operaciones de adición, sustracción y producto cruz, una vez calculado el vector resultante este se muestra en una pantalla superpuesto a la realidad.

El trabajo de Enrique Solis[13], involucra el uso de visión computacional para llevar acabo el seguimiento de dos referentes físicos de los vectores, en este caso fueron usados dos bastones de diferentes colores como se observa en la Figura 2. El sistema consiste de un par de cámaras estéreo que junto con la aplicación de algoritmos de

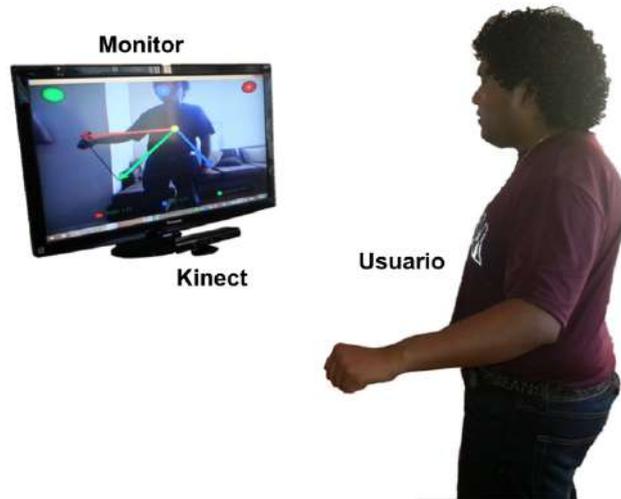


Figura 1: En este trabajo los referentes físicos utilizados fueron las manos, y se empleó un dispositivo Kinect para la localización y seguimiento de dichos referentes.

visión computacional permiten la localización tridimensional de los referentes. El sistema presenta ciertas debilidades como la que se observa en la Figura 3, bajo ciertas condiciones no es posible realizar una localización de los referentes.

En este trabajo se retoma lo realizado por Enrique Solis[13] y se propone un método para robustecer el seguimiento de los referentes físicos, utilizando para ello la información de movimiento para poder predecir la pose final del objeto en los siguientes cuadros, con ello nos permitirá seguir el objeto bajo condiciones de oclusión o aparentes condiciones de paralelismo debido a la perspectiva de la cámara.

Se realizarán las fases de segmentación y etiquetación similares a las fases efectuadas en el trabajo previo de Enrique Solis[13]. Las principales diferencias son que será utilizada la distancia normalizada de Mahalanobis[14], también se utilizará un espacio de color diferente, con estos dos cambios se pretende evitar los problemas segmentar en base a color que tenía el trabajo previo los cuales se describirán más adelante.

Para robustecer el seguimiento de los referentes se realizarán estimaciones de las velocidades de un cuerpo rígido, lo cual será llevado a cabo utilizando el filtro de Kalman extendido. La ventaja de utilizar

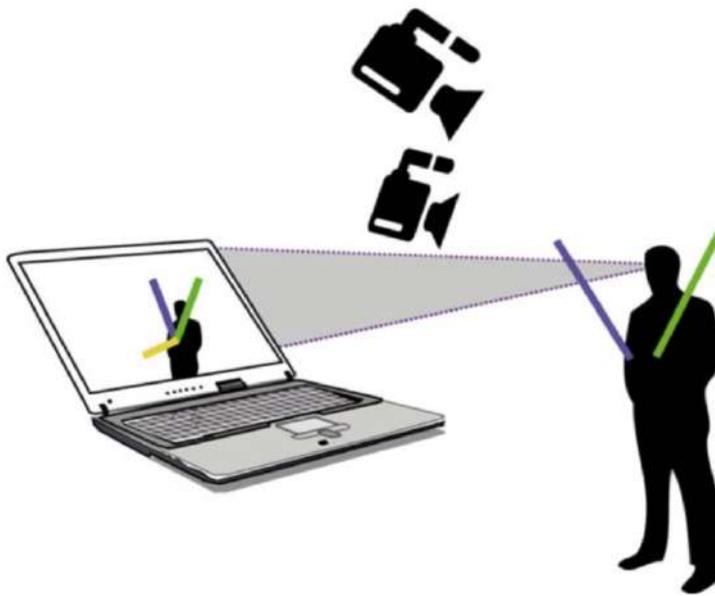


Figura 2: Sistema de realidad aumentada utilizando cámaras estéreo y referentes físicos para manipular los vectores.

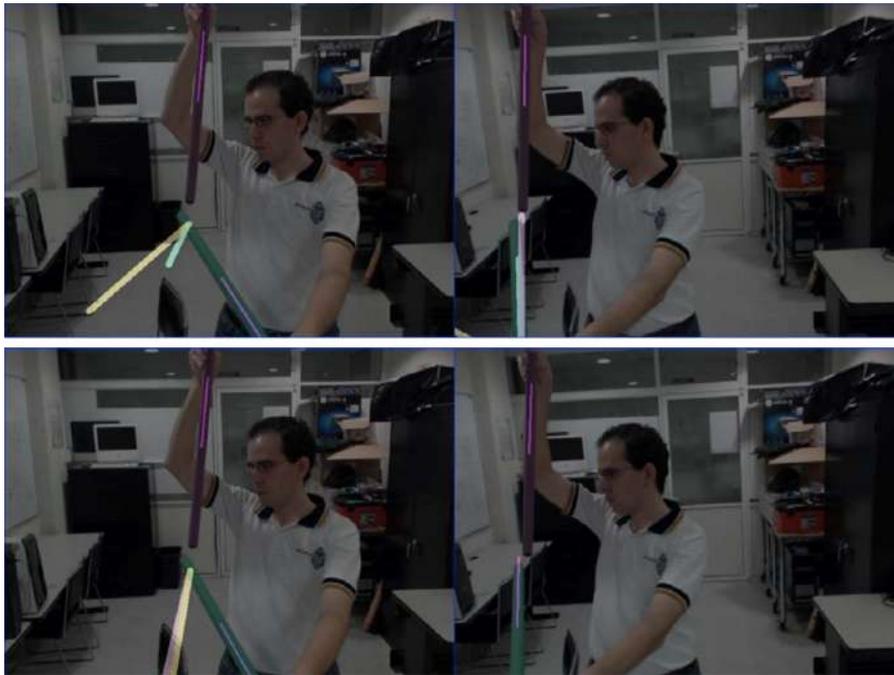


Figura 3: Capturas de pares de imágenes del sistema funcionando de forma inestable, en la cámara de la derecha aparece una condición de colinealidad, bajo este caso ya no es posible realizar la localización de la recta.

este filtro es que toma en cuenta el ruido en las mediciones introducido por la cámara y los algoritmos de visión de las fases anteriores, así como la incertidumbre ocasionada por no tomar en cuenta todas las variables físicas en el modelo.

Un trabajo similar para estimación de las velocidades lineal y angular para líneas rectas utilizando filtros de Kalman fue realizado por Zhengyou Zhang y Olivier D. Faugeras [15, 6].

En ambos trabajos se trata el problema de estimar los movimientos de un objeto tridimensional mediante de una secuencia de imágenes estéreo, se estableció una relación entre el movimiento tridimensional y el movimiento observado en la imagen. Para realizar esta estimación se utilizan las líneas de los objetos, las mediciones entregadas al sistema son pares de líneas 3D. En [15] se utilizó una representación de la línea 3D mediante el punto medio del segmento de recta y un vector de dirección, por otra parte [6] utilizó una representación de cinco parámetros, donde tres de ellos son el punto medio y dos representan la orientación en ángulos de Euler. Se utilizó un modelo de movimiento cinemática clásico, que divide el movimiento en dos partes: movimiento de rotación alrededor de un punto y movimiento de traslación del centro de rotación, en este trabajo para simplificación del modelo se asume una velocidad angular constante.

## 1.2 OBJETIVOS

Se presenta el objetivo de la presente tesis, el cual es la meta del presente trabajo, y los objetivos particulares a completar durante el desarrollo de la metodología.

### 1.2.1 *Objetivo General*

Desarrollar un método robusto para la detección y seguimiento tridimensional de bastones haciendo uso de modelos probabilísticos.

### 1.2.2 *Objetivos particulares*

- Generación del modelo de bastones.
- Localización y seguimiento de bastones en la imagen.
- Desarrollo del modelo probabilístico para el seguimiento de los bastones en el espacio tridimensional a partir de una secuencia de imágenes.
- Implementación del método propuesto.
- Realizar la evaluación de la metodología propuesta.

## 1.3 ESTRUCTURA DE LA TESIS

El presente trabajo de tesis se encuentra dividido en cinco capítulos los cuales se detallan a continuación.

- El primer capítulo que es el actual, presenta una breve introducción al trabajo de tesis y da un panorama general del problema y los objetivos.
- En el capítulo dos se presenta el marco teórico, conceptos en los cuales se apoya este trabajo para su desarrollo.
- En el capítulo tres se describe la metodología utilizada para cumplir con los objetivos.
- El capítulo cuatro describe el diseño de los experimentos y muestra los resultados.
- Por último el capítulo cinco se presentan las conclusiones de este trabajo así como también posible trabajo futuro.



## MARCO TEÓRICO

---

### 2.1 INTRODUCCIÓN

En este capítulo se describen los conceptos y fundamentos sobre los cuales se apoya la metodología presentada en este trabajo.

En sección 2.3 presentan los conceptos matemáticos utilizados que nos permitirán entender la relación entre los objetos y su proyección en el plano de la imagen. Enseguida se presentan los conceptos utilizados para la localización de líneas rectas en el espacio tridimensional así como su representación utilizando coordenadas de Plücker.

En la sección 2.4 se aborda el tema de la representación de colores en las imágenes, se describe con mayor detalle el espacio de color CIE Lab, el cual es utilizado para la segmentación de los bastones. Posteriormente en la sección 2.5 se introduce el tema de las distancias utilizadas para el proceso de segmentación.

En la sección 2.6 se presenta el modelo cinemático que describe el movimiento de un cuerpo rígido, entender este modelo nos permitirá comprender mejor el planteamiento realizado en la metodología para predecir de la pose de los bastones cuando la localización no es posible. Para realizar dicha estimación se utiliza el filtro de Kalman, sus fundamentos y su funcionamiento son explicados en la sección 2.7.

Finalmente se presenta una modificación al filtro de Kalman para tratar con sistemas no lineales, llamada filtro de Kalman extendido.

## 2.2 COORDENADAS CARTESIANAS

Un punto en el espacio euclidiano de dos dimensiones puede ser representado mediante un sistema de coordenadas cartesianas, en este sistema se definen dos ejes ortogonales que se interceptan en un punto llamado origen.

Cada punto en espacio esta representado como un par de numero reales  $(x, y)$  cada numero indica la distancia mínima de dicho punto al origen en cada uno de los ejes.

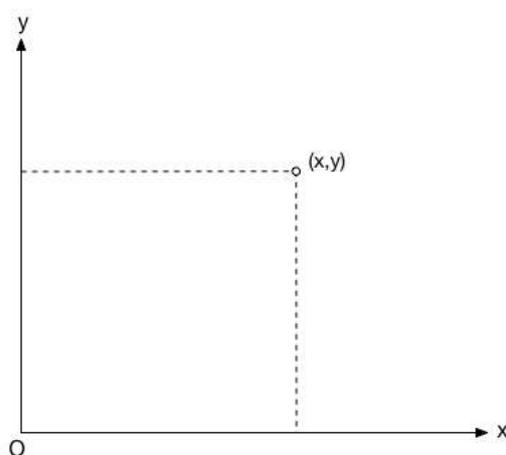


Figura 4: Coordenadas Cartesianas.

## 2.3 COORDENADAS HOMOGENEAS

Una manera alterna de representar puntos en el espacio de dos dimensiones es mediante la tripleta  $(x, y, w)$  en donde cada punto es representado por la razón entre sus componentes.

Una propiedad que posee esta representación es que el punto puede estar multiplicado por un escalar  $\lambda \neq 0$  y seguir representando el mismo punto, es decir es independiente de escala, esta forma de representación es conocida como coordenadas homogéneas.

La representación en coordenadas cartesianas de un punto en coordenadas homogéneas se obtiene dividiendo los tres componentes entre el último.

Sea  $p = (x, y, w)$  un punto en coordenadas homogéneas, la representación cartesiana está dada como sigue:

$$p' = \begin{bmatrix} \frac{x}{w} \\ \frac{y}{w} \end{bmatrix}$$

Las coordenadas homogéneas forman clases de equivalencia en las cuales un punto  $p$  tiene una relación de equivalencia  $\lambda p$  con  $\lambda \neq 0$  y cumple cada una de las propiedades de relación de equivalencia:

- Reflexividad:  $p \mathcal{R} p$
- Simetría: si  $p \mathcal{R} \lambda p$  entonces  $\lambda p \mathcal{R} p$  para cualquier escalar  $\lambda \neq 0$ .
- Transitividad: si  $p \mathcal{R} \lambda p \wedge \lambda p \mathcal{R} g p''$  entonces  $\lambda p' \mathcal{R} g p''$  para cualquier escalar  $g \neq 0$  y  $\lambda \neq 0$ .

El conjunto de clases de equivalencia forman lo que se conoce como el espacio proyectivo  $\mathbb{P}^2$ .

### 2.3.1 Líneas

Una línea en un espacio Euclidiano de dos dimensiones puede ser representada por la ecuación  $ax + by + c = 0$ , los valores de  $a$ ,  $b$  y  $c$  definen a recta, por tanto esta puede ser representada por un vector  $l = [a, b, c]^T$ .

La correspondencia entre líneas y vectores no es uno a uno, puesto que si multiplicamos  $l$  por un escalar  $\lambda \neq 0$ ,  $l$  seguirá representando la misma recta.

Los vectores que representan esas rectas  $l$  y  $\lambda l$  son considerados equivalentes, y son conocidos como vectores homogéneos.

Los puntos pertenecientes a la línea  $l$  cumplen con la propiedad de que  $ax + by + c = 0$ , esto puede ser escrito en términos de un producto punto de vectores como:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = p \cdot l = ax + by + c = 0$$

Para encontrar la intersección de dos líneas  $l_1$  y  $l_2$  debemos encontrar el punto  $p$  que pertenezca a ambas rectas, es decir que satisfaga  $l_1 \cdot p = 0$  y  $l_2 \cdot p = 0$ .

El producto triple escalar de las líneas definido como

$$l_1 \cdot (l_1 \times l_2) = l_2 \cdot (l_1 \times l_2') = 0 \quad (1)$$

podemos ver que si hacemos  $x = l_1 \times l_2'$  entonces  $l_1 x^\top = l_2 x^\top = 0$ , esto quiere decir que  $x$  está en ambas rectas, por tanto la intersección de dos líneas está dada por:

$$l_1 \times l_2.$$

Para el caso de líneas paralelas si consideramos  $l_1 = (a_1, b_1, c_1)^\top$  y  $l_2 = (a_2, b_2, c_2)^\top$  para las cuales los primeros dos componentes son iguales, entonces podemos calcular la intersección como  $l_1 \times l_2 = (c_2 - c_1)(b_1, -a_1, 0)$  e ignorando el factor de escala  $(c_2 - c_1)$  obtenemos el punto  $(b, -a, 0)^\top$ . Si intentamos obtener la coordenada no homogénea de este punto obtendríamos  $(b/0, -a/0, 0)^\top$  lo cual nos indica que el punto de intersección tiene componentes en el infinito y no corresponden con  $\mathbb{R}^2$ , estos puntos con el tercer componente igual a cero, se conocen como puntos en el infinito o puntos ideales.

### 2.3.2 Puntos

Todas las propiedades que se cumplen para las rectas se cumplen también para los puntos, de hecho observamos una dualidad entre rectas y puntos con coordenadas homogéneas, los cuales podemos intercambiar indistintamente, de hecho la ecuación de la recta es simétrica puesto que  $l^\top x = 0$  implica que  $x^\top l = 0$ .

### 2.3.3 Coordenadas de Plücker

Las líneas en un espacio 3D, pueden ser representadas por dos puntos, sean estos puntos  $a = (a_x, a_y, a_z)$  y  $b_1 = (b_x, b_y, b_z)$ ,

Las coordenadas de Plücker de una línea  $l$  que pasa por los puntos  $a$  y  $b$  se definen como  $l = (n, u)$ , donde  $u = b - a$  define la orientación de la recta, y  $n = x_2 \times x_1$  es conocido como el vector momento.

Un punto  $p$  está en la recta, si y solo si se cumple que  $p \times u = n$ . Cuando  $u$  es un vector unitario  $\|n\|$  define la distancia mínima del origen a la línea. En general dicha distancia está dada por la razón entre la norma del vector de momento y el vector de dirección, esto es:

$$d = \frac{\|n\|}{\|u\|}.$$

Las seis coordenadas de Plücker tienen dos redundancias. Puesto que la razón entre  $\|n\|$  y  $\|u\|$  es lo que determina la distancia de la recta al origen, las coordenadas de Plücker  $(n, u)$  y  $(\lambda n, \lambda u)$  con  $\lambda \neq 0$  describen la misma recta. Podemos escoger  $u$  como un vector unitario. Las restricciones  $\|u\| = 1$  y  $u \cdot n = 0$  remueven dos variables de  $l$ , reflejando los cuatro grados de libertad de  $l$ .

Cuando dividimos  $u$  y  $n$  entre  $\|u\|$  obtenemos las coordenadas lo que se conoce como las coordenadas normales de Plücker en donde  $\underline{u} = \frac{\vec{u}}{\|u\|}$  y  $\underline{n} = \frac{\vec{n}}{\|u\|}$ .

## 2.4 ESPACIOS DE COLOR

Un espacio de color es una forma de describir el color de la luz, usualmente con tres valores numéricos. Existen diferentes esquemas bajo los cuales se pueden definir estos tres valores, un esquema importante es el esquema aditivo que trabaja añadiendo cantidades específicas de tres tipos de luces con una cromancia específica. Los espacios de color aditivos pueden ser llamados espacios de color tricromático, ya que funcionan mediante la definición tres estímulos proporcionados a ojo.

## 2.4.1 Espacio CIELab

El espacio de color tridimensional CIExyz, es la base para varios sistemas de manejo de colores, este espacio de color contiene todos los colores percibidos por el ser humano.

El ojo humano tiene tres tipos de conos llamados “L”, “M” y “S”, cada cono tiene una respuesta espectral diferente, es decir una curva que nos indica que tanta “salida” tiene un cono ante una luz de potencia fija en una longitud de onda del rango visible, las letras L, M, S, se refieren a los picos que tienen los tres diferentes conos [8], dicha respuesta se observa en la Figura 5.

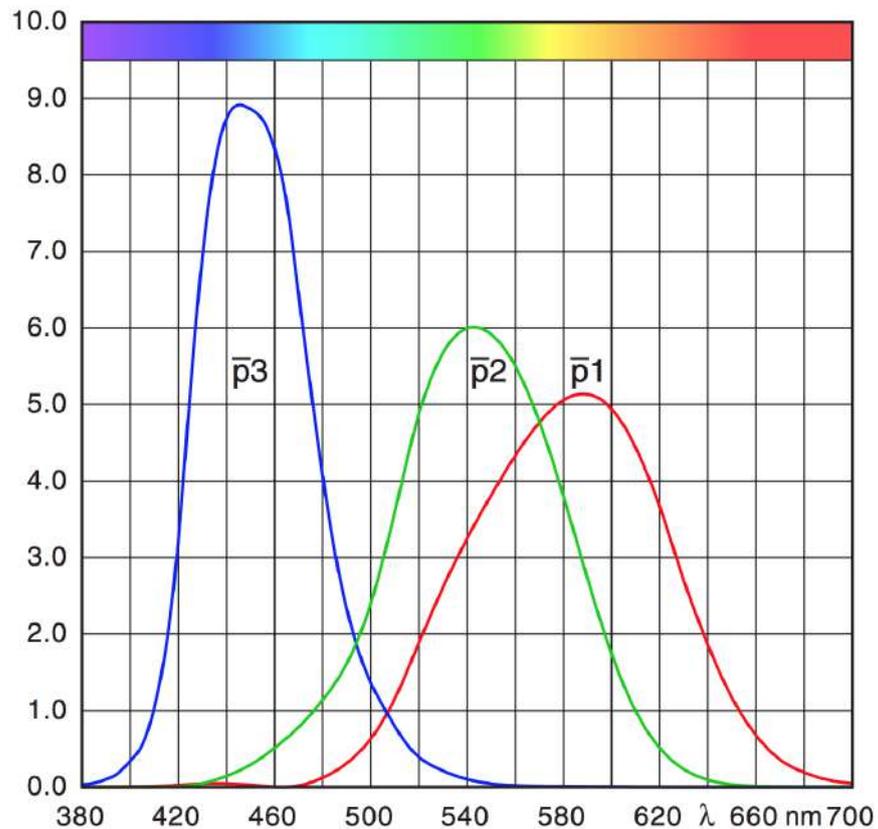


Figura 5: Respuesta de los conos según la longitud de onda.

Como menciona A. Kerr[8], estas curvas fueron determinadas mediante experimentos visuales, en el que se mostraba a una persona un color de una determinada longitud de onda y se trataba de igualar con tres diferentes componentes los cuales producían longitudes de onda de rojo, verde y azul, en ocasiones no se podía igualar el color

aun suprimiendo el rojo, lo que hicieron fue agregar color rojo al otro objeto lo que se interpreta como un valor negativo en el componente rojo, se analizaron los diferentes datos y se obtuvieron unas funciones que se conocen como “color matching functions”, en la Figura 6 se observa este comportamiento.

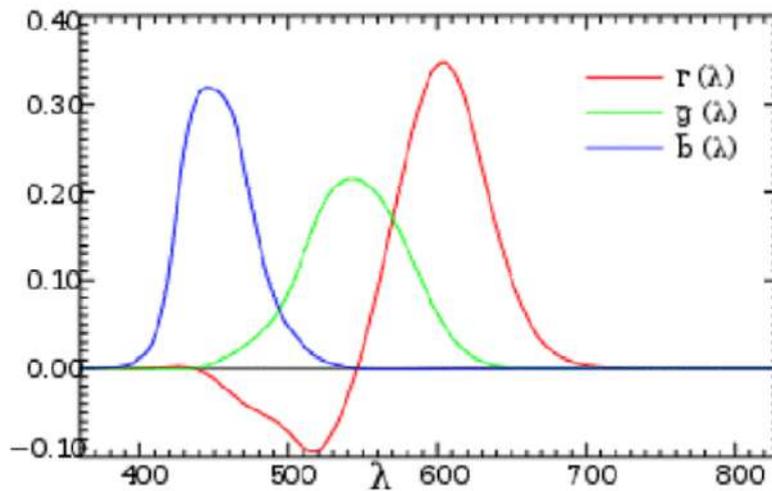


Figura 6: Color matching functions.

Esta forma de representar el color es llamada CI<sub>Ergb</sub>, una desventaja de este espacio de color es que la función de matching del rojo presenta valores negativos en una parte del espectro de longitud de onda, para solucionar esto se hicieron una serie de manipulaciones matemáticas para determinar unas nuevas funciones de matching y con ello tres nuevos componentes X, Y, Z.

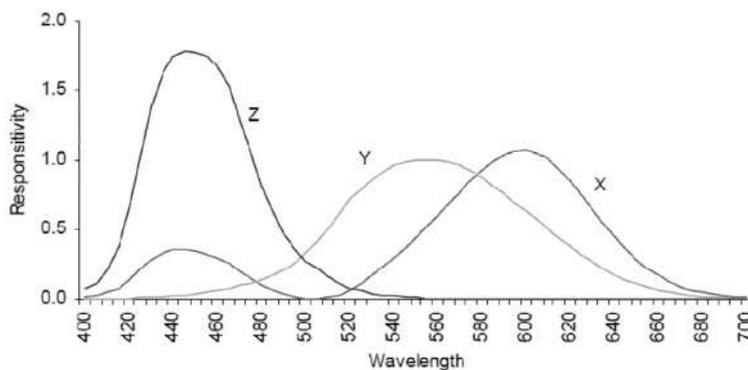


Figura 7: Color matching functions sin valores negativos para el modelo CIExyz.

El modelo RGB se basa en colores primarios aditivos rojo, verde y azul. Por el contrario, el CIExyz se basa en 3 primarios imaginarios con caracterización espectral ( $X, Y, Z$ ), que son los que representan el color. Éstos se combinan para formar todos los colores visibles. En el modelo CIExyz, la  $Y$  define la luminosidad, mientras la  $Z$  es casi igual a la estimulación producida por el azul o el cono "S", y  $X$  es una combinación lineal de las respuestas de los conos escogido para ser no negativo.

Existe un problema con el espacios CIErgb y CIExyz cuando es importante tomar en cuenta la forma en que el ser humano percibe los colores, son espacios no uniformes. Se considera un espacio perceptiblemente uniforme cuando la distancia entre dos puntos de color en cualquier parte del espacio corresponde a la diferencia perceptual entre esos dos colores para la visión humana. En el caso del espacio RGB dos colores distintos separados por una distancia fija en diversas partes del espacio no tienen la misma diferencia perceptual, es decir, dos colores que para el ser humano son muy parecidos en el espacio RGB podrían estar juntos o muy separados, también puede ocurrir que dos colores que no se parezcan puedan estar muy próximos.

El espacio de color CIELab es derivado del espacio CIExyz. El propósito de este es producir un espacio de color que sea más perceptualmente lineal que otros espacios de color. Los componentes principales del espacio de color son  $L$  que representa la luminosidad,  $a$  y  $b$  son componentes de color basados en el espacio de color CIExyz. el eje  $a$  representa una posición entre el rojo/magenta y el eje  $b$  representa una posición entre azul/amarillo.

Las relaciones no lineales para  $L, a$  y  $b$  pretenden emular la respuesta no lineal del ojo. Además, los cambios uniformes de los componentes en el espacio de color Lab tienen por objeto corresponder a cambios uniformes en el color percibido, por lo que las diferencias relativas de percepción entre dos colores en el espacio Lab se pueden aproximar tratando cada color como un punto en un espacio tridimensional, con tres componentes:  $L, a, b$  y tomando la distancia entre ellos[12].

## 2.5 DISTANCIA DE MAHALANOBIS

La distancia de Mahalanobis es una medida de distancia introducida por P. C. Mahalanobis[10] en 1936. Es una forma de determinar la similitud entre dos variables aleatorias multidimensionales. Esta distancia toma en cuenta la correlación entre las variables aleatorias [11].

Mediante esta distancia es posible determinar que tan similar es una muestra a cierta distribución de datos Gaussiana. Para el caso de una sola variable podemos decir que esta similitud puede ser dada en términos de la desviación estándar, la cual determina que tan cerca estamos de la media. Así, podemos decir que una observación que se encuentra a tres desviaciones estándar de la media esta mas lejos que una observación con dos desviaciones estándar. Para una distribución normal esto también significa que es mas probable observar el dato que esta mas cerca de la media, esto se debe a que la densidad de probabilidad es mas alta en ese punto.

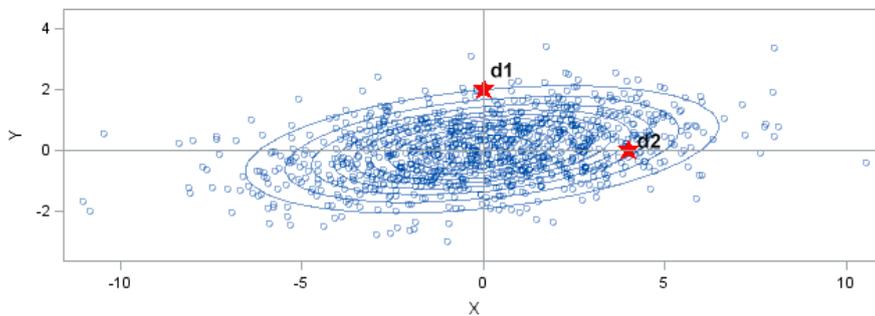


Figura 8: Dos datos en marcados en rojo,  $d_1$  y  $d_2$  aparentemente  $d_1$  esta mas cerca de la media, esto sin tomar en cuenta la distribución de los datos.

En la Figura 8 se observa una distribución normal bi-variada y dos observaciones las cuales son mostradas en rojo, en este caso aparenta que  $d_1$  esta mas cerca, y calculamos la distancia euclidiana pudiera ser que  $d_1$  este a menor distancia que  $d_2$ . Sin embargo la varianza sobre el Eje Y es menor que la varianza sobre el Eje X, lo cual no es tomado en cuenta por la distancia Euclidiana, por lo tanto  $d_1$  es

menos probable de ser observado bajo la distribución mostrada en la Figura 8 que  $d_2$ .

Para poder tomar en cuenta la correlación de los datos y la variación en cada Eje, se deben primero normalizar los datos para luego calcular la distancia Euclidiana. Empecemos por mostrar el caso para dos dimensiones.

Si se tienen una matriz de covarianza  $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_p^2)$ , La raíz cuadrada de  $\Sigma$  consiste en las desviaciones estándares  $\Sigma = \mathbf{D}\mathbf{D}^T$  donde  $\mathbf{D} = \text{diag}(\sigma_1, \dots, \sigma_p)$ .

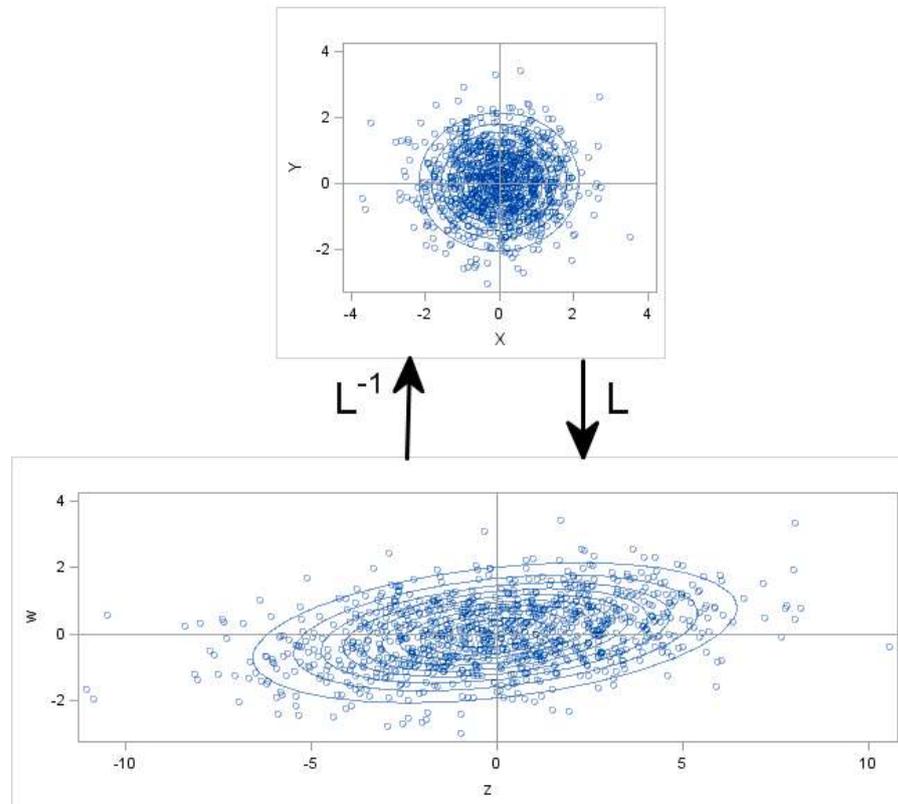


Figura 9: Normalización y desnormalización de datos

Se puede observar geoméricamente en la Figura 9 que la matriz  $L^{-1}$  escala cada coordenada en dirección independiente. En general la matriz de covarianza contiene elementos no diagonales, la interpretación geométrica es similar a la de la matriz diagonal que representa el caso en el cual solo se toma en cuenta la varianza en los ejes X y Y, es decir es solo un escalamiento, pero la transformación también

involucra otros tipos de transformaciones como rotaciones y estiramientos.

Para calcular la “raíz cuadrada” de la matriz de covarianza se utiliza la transformación de Cholesky[7]  $\sigma = L^T L$ , Luego se utiliza la inversa de dicha matriz para de-correlacionar los datos.

Entonces se puede reescribir el cuadrado de la distancia  $zz^T$  como:

$$\begin{aligned} d &= z^T z \\ d &= (L^{-1}(x - \mu))^T (L^{-1}(x - \mu)) \\ d &= (x - \mu)^T (LL^T)^{-1} (x - \mu) \\ d &= (x - \mu)^T \Sigma^{-1} (x - \mu) \end{aligned}$$

La ecuación anterior es el cuadrado de la distancia de Mahalanobis, la derivación utiliza diferentes identidades de matrices, geoméricamente transforma los datos correlacionados a datos no correlacionados y calcula la distancia euclidiana. Note que para el caso cuando  $\Sigma$  es la matriz identidad equivale a calcular la distancia euclidiana.

## 2.6 MOVIMIENTO DE UN CUERPO RÍGIDO

Como explica [9], para describir el movimiento de un objeto debemos definir la trayectoria de todos y cada uno de los puntos en dicho objeto. Podemos especificar las coordenadas de un punto como una función  $X(t)$  que depende del tiempo.

El movimiento de un cuerpo rígido se puede ver como una familia de transformaciones que describen como las coordenadas de cada punto del objeto se mueven en el tiempo.

$$g(t) : \mathbb{R}^3 \rightarrow \mathbb{R}^3, X \rightarrow g(t)X$$

Ademas de transformar puntos  $g$  también incluye transformaciones a vectores, si tenemos un vector  $v$  definido por dos puntos  $p$  y  $q$

con coordenadas  $v = Y - X$ , entonces después de la transformación  $g$  obtendremos un nuevo vector.

$$u = g_*(v) = g(Y) - g(X).$$

La transformación  $g$  para que represente las transformaciones de movimiento de un cuerpo rígido se deben preservar las distancias entre puntos. Luego entonces  $\|g_*(v)\| = \|v\|$ . Estas transformaciones que preservan las distancias se les conoce como Transformación Euclidiana y son denotadas como  $E(3)$ .

Para describir el movimiento de un cuerpo rígido no es suficiente con que la transformación preserve distancias, hay transformaciones que no son físicamente posibles para un cuerpo rígido, un ejemplo de este tipo de transformaciones imposibles es la reflexión de los puntos del objeto en el plano  $XY$ . Por lo tanto además de requerir que las estas transformaciones preserven la distancia entre sus puntos también se requiere preservar la orientación. Este tipo de transformaciones se conocen como Transformación Euclidiana Especial. Estas cumplen las siguientes dos condiciones:

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

El conjunto de todas las transformaciones especiales que cumplen las dos condiciones anteriores esta denotado por  $SE(3)$ .

Con estas propiedades podemos describir el movimiento de un cuerpo rígido escogiendo un punto en el objeto, asociando a este un sistema coordenado y siguiendo el movimiento del sistema coordenado relativo un sistema coordenado de referencia conocido también como sistema coordenado del mundo.

Sea un sistema coordenado con sus ejes dados por tres vectores ortonormales  $e_1, e_2, e_3 \in \mathbb{R}^3$ , estos vectores están ordenados en forma de un sistema de mano derecha tal que  $e_1 \times e_2 = e_3$ , después de aplicar la transformación que describe el movimiento del cuerpo rígido tendremos que:

$$g_*(e_1) \times g_*(e_2) = g_*(e_3)$$

es decir estos vectores aun preservan la ortonormalidad y forman un sistema de mano derecha el cual estará siempre asociado al cuerpo rígido y definirá su movimiento.

### 2.6.1 Movimiento Rotacional

El movimiento rotacional respecto a un sistema coordenado de referencia  $W$  es determinada por tres vectores ortonormales que conforman el sistema coordenado  $C$  los cuales son:  $r_1 = g_*(e_1), r_2 = g_*(e_2), r_3 = g_*(e_3)$ , en donde los vectores  $r_1, r_2, r_3$  forman los tres ejes principales  $x, y, z$  del sistema coordenado  $C$ ,  $e_1, e_2$  y  $e_3$  forman los tres ejes principales de  $W$  [9].

Este sistema determina la rotación del objeto y puede ser representado como una matriz de  $3 \times 3$ :

$$R_{wc} = [r_1, r_2, r_3] \in \mathbb{R}^{3 \times 3} \quad (2)$$

Esta matriz cumple que:

$$R_{wc}^T R_{wc} = R_{wc} R_{wc}^T = I \quad (3)$$

Cualquier matriz que cumpla la condición anterior es llamada matriz ortogonal y cumple que  $R_{wc}^T = R_{wc}^{-1}$ . Además  $r_1, r_2, r_3$  forman un sistema de mano derecha y cumple que  $\det(R_{wc}) = +1$ .  $R_{wc}$  es una matriz ortogonal especial que cumple el preservar la orientación, el conjunto de estas matrices  $\in \mathbb{R}^3$  se denota por  $SO(3)$ . Este conjunto de matrices es también conocido como matrices de rotación, cualquier elemento en  $SO(3)$  representa una posible configuración de un objeto rotado alrededor de un punto  $o$ , además  $R_{wc}$  puede representar la transformación del sistema coordenado  $C$  al sistema coordenado  $W$ .

La continua rotación de un objeto puede ser descrita con una trayectoria  $R(t) : t \rightarrow SO(3)$ , el movimiento relativo entre el tiempo  $t_1$  y  $t_2$  se denota como  $R(t_1, t_2)$  y por la composición de rotaciones obtenemos que:

$$R(t_2, t_0) = R(t_2, t_1)R(t_1, t_0) \quad (4)$$

Ahora bien, esta matriz de  $3 \times 3$  tiene 9 entradas, sin embargo estas entradas no son parámetros libres, puesto que deben satisfacer que  $R^T R = I$ , esto impone seis restricciones a las nueve entradas, en realidad la dimensión del espacio de matrices de rotación  $SO(3)$  debe ser de tres, seis parámetros de los nueve son redundantes. Dada una trayectoria  $R(t) : \mathbb{R} \rightarrow SO(3)$  que describe un movimiento rotacional continuo, debe satisfacer que:

$$R(t)R^T(t) = I \quad (5)$$

Calculando la derivada de la ecuación con respecto al tiempo  $t$  obtenemos lo siguiente:

$$\dot{R}(t)R^T(t) + R(t)\dot{R}(t)^T \implies \dot{R}(t)R^T(t) = -R(t)\dot{R}(t)^T \quad (6)$$

Por definición de matriz anti-simétrica y observando la ecuación anterior vemos que  $\dot{R}(t)R^T(t)$  es una matriz anti-simétrica, luego entonces existe un vector  $w(t) \in \mathbb{R}^3$  tal que :

$$\dot{R}(t)R^T(t) = \hat{w}(t) \quad (7)$$

Multiplicando ambos lados de la ecuación por  $R(t)$  obtenemos:

$$\dot{R}(t) = R(t)\hat{w}(t) \quad (8)$$

El conjunto de todas las matrices anti-simétricas esta denotado por:

$$so(3) = \{\hat{w} \in \mathbb{R}^{3 \times 3} | w \in \mathbb{R}^3\} \quad (9)$$

Observamos que la ecuación (8) es una ecuación diferencial ordinaria, y si asumimos que  $R(t)$  es una matriz de transición de estado, resolviendo la ecuación diferencial nos queda que:

$$x(t) = e^{\hat{w}t}x(0) \quad (10)$$

donde  $e^{\hat{w}t}$  es la matriz exponencial y puede ser expresado como la suma infinita de la serie que genera la constante  $e$ , Asumiendo que  $R(0) = I$  obtenemos la ecuación.

$$R(t) = e^{\hat{w}t} \quad (11)$$

La ecuación anterior podemos expresarla mejor si absorbemos el parámetro  $t$  dentro del vector  $w$ :

$$R(t) = e^{\hat{w}} \quad (12)$$

La interpretación física de esta ecuación es una rotación con respecto al eje  $w \in \mathbb{R}^3$  con un ángulo de  $\|w\|$ , por tanto esta ecuación define un mapeo de  $so(3)$  a  $SO(3)$  llamado mapeo exponencial. Para calcular la exponencial de la matriz usamos la formula de Rodrigues [9].

### 2.6.2 Movimiento general de un cuerpo rígido

El movimiento de un cuerpo rígido puede ser representado mediante dos componentes, uno de rotación  $R$  y uno de traslación  $T$ .  $g = (R, T)$  [9].

En la Figura 10 se muestra el movimiento de un cuerpo rígido con el sistema coordenado  $C$  adjunto a el. Para describir el movimiento de un punto  $p$  con respecto al sistema coordenado del mundo  $W$

Es claro que  $X_w$  es la suma de la traslación  $T_{wc}$  del origen del sistema  $C$  relativo al sistema  $W$ , y el vector  $X_c$  expresado en el sistema coordenado  $W$ . Como  $X_c$  esta expresado en el sistema  $C$  entonces  $R_{wc}X_c$  es la rotación entre el marco de referencia del mundo y el  $C$ .

$$X_w = R_{wc}X_c + T_{wc} \quad (13)$$

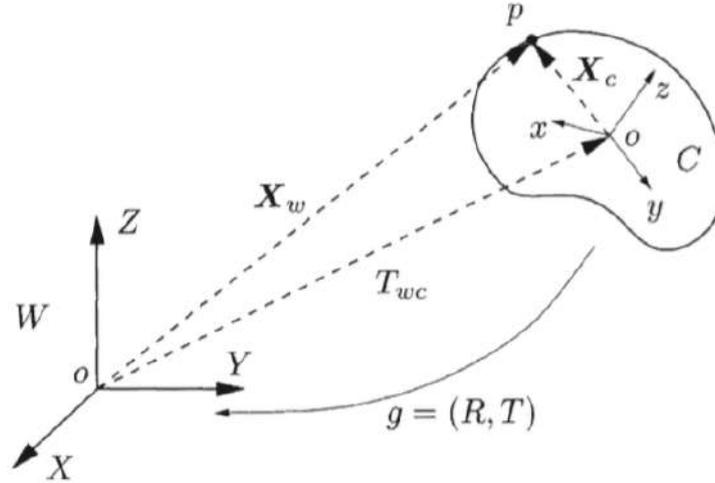


Figura 10: Movimiento general de un cuerpo rígido asociado a un sistema coordenado C con respecto al sistema coordenado del mundo W.

Para obtener una representación de  $g$  matricial se introducen las coordenadas homogéneas, la representación matricial de  $g$  esta dada de la siguiente forma:

$$g(t) = \begin{bmatrix} \mathbf{R}(t) & \mathbf{T}(t) \\ 0 & 1 \end{bmatrix}$$

Similar al caso de rotaciones puras vistas anteriormente, podemos obtener la derivada como se ve en la ecuación 14

$$g'(t)g^{-1}(t) = \begin{bmatrix} \mathbf{R}'(t)\mathbf{R}^\top(t) & \mathbf{T}'(t) - \mathbf{R}'(t)\mathbf{R}^\top(t)\mathbf{T}(t) \\ 0 & 0 \end{bmatrix} \quad (14)$$

Si hacemos  $\mathbf{w} = \mathbf{R}'(t)\mathbf{R}^\top(t)$  y  $\mathbf{v} = \mathbf{T}'(t) - \mathbf{w}(t)\mathbf{T}(t)$  entonces la representación queda expresada en la siguiente matriz.

$$g'(t)g^{-1}(t) = \begin{bmatrix} \mathbf{w}(t) & \mathbf{v}(t) \\ 0 & 0 \end{bmatrix}$$

Luego entonces definimos la matriz  $\hat{\xi}$  como:

$$\hat{\xi} = \begin{bmatrix} \mathbf{w}(t) & \mathbf{v}(t) \\ 0 & 0 \end{bmatrix}$$

Entonces:

$$g'(t) = (g'(t)g^{-1}(t))g(t) = \hat{\xi}(t)g(t)$$

Donde  $\hat{\xi}$  puede ser usado para aproximar  $g(t)$  localmente como se muestra en la ecuación 15.

$$g(t + dt)g(t) + \hat{\xi}(t)g(t)dt = (I + \hat{\xi}g(t)dt)g(t) \quad (15)$$

Por tanto la velocidad en un punto esta dada por la ecuación 16.

$$p' = w(t)p + v(t) \quad (16)$$

## 2.7 FILTRO DE KALMAN

El filtro de Kalman es un algoritmo utilizado cuando se desea estimar el estado de un sistema dinámico, un sistema cuyo estado evoluciona con el tiempo. El estado de el sistema no es directamente observable, pero se dispone de otro tipo de información producida por el sistema, esta información si puede ser medida. Este algoritmo también toma en cuenta la incertidumbre en la información disponible para la estimación del estado.

Algunos usos comunes del filtro de Kalman incluyen la reducción de ruido en datos y estimación de parámetros de interés.

Como se explica en [5], el filtro de Kalman asume que el estado del sistema evoluciona desde un estado anterior  $t - 1$  a un estado en  $t$  de acuerdo con la siguiente ecuación.

$$x_t = F_t x_{t-1} + w_t \quad (17)$$

Donde:

- $x_t$  es el vector de estado que contiene información sobre el sistema en el instante  $t$  (e.g. velocidad, posición, temperatura etc..).

- $w_t$  Representa la incertidumbre del sistema o ruido en el instante  $t$  en términos de cada parámetro del vector de estados, se asume  $\sim \mathcal{N}(0, Q_t)$
- $F_t$  es la matriz de transición de estados.

Las mediciones hechas al sistema de pueden modelar mediante la siguiente ecuación:

$$\hat{z}_t = \mathbf{H}_t x_{t-1} + v_t \quad (18)$$

Donde:

- $z_t$  es el vector de mediciones.
- $H_t$  es una matriz de transformación que relaciona el dominio estados con el dominio de mediciones.
- $v_t$  representa el ruido en la medición, se asume una distribución normal con media 0 y varianza  $R_t \sim \mathcal{N}(0, R_t)$ .

Como se puede observar en la Figura 11, el filtro de Kalman involucra dos etapas: predicción y actualización.

En la etapa de predicción se calcula el siguiente estado del sistema utilizando la ecuación 17. En etapa de actualización se combina la medición obtenida en el tiempo  $t$  con la predicción realizada en la etapa anterior, utilizando estas dos fuentes de información se provee una mejor estimación el estado (CITA REQUERIDA).

### 2.7.1 Etapa de predicción

En esta etapa se predice el estado que tendrá el sistema utilizando el estado en el tiempo  $t - 1$ . Se asume que se tiene una estimación del estado del sistema en  $t - 1$ , llámese  $x_{t-1}$ . Este estado es una variable aleatoria con distribución Gaussianas, con media  $\hat{x}_{t-1}$  que es el centro de la distribución, el estado mas probable, y varianza  $Q_t$ , que representa la incertidumbre.

Como el estado esta representado por una variable aleatoria con distribución Gaussiana, se debe realizar la predicción sobre todos los

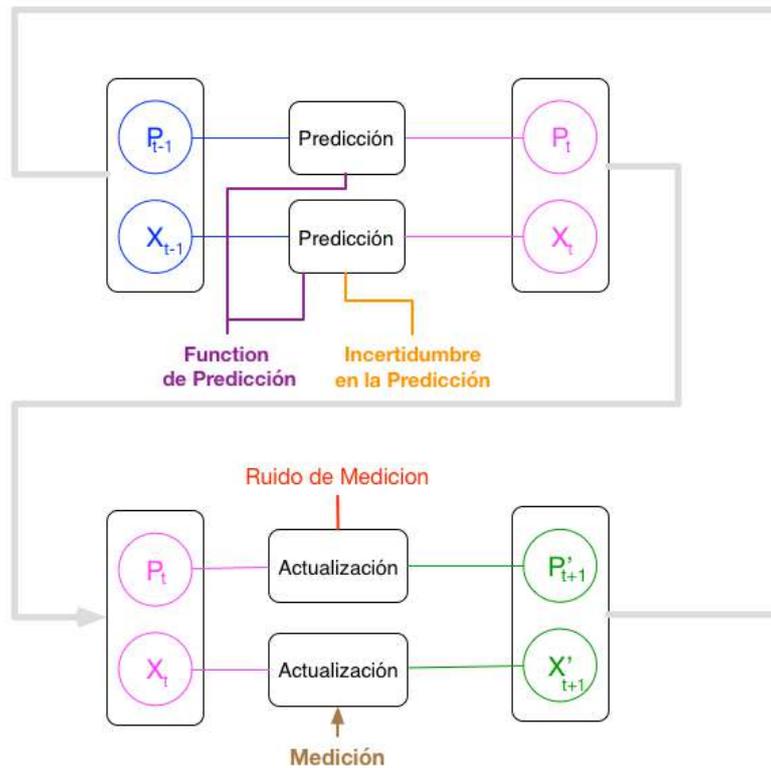


Figura 11: En esta figura se pueden observar las dos etapas del filtro de Kalman, la  $P_{t-1}$  y la  $X_{t-1}$  corresponden a la estimación del estado  $X$  con una matriz de covarianza  $P$  que representa la incertidumbre, la  $X_t$  y la  $P_t$  son la predicción realizada en base a la matriz de predicción  $F_t$ , luego en la etapa de actualización se obtiene una nueva medición y tomando en cuenta el ruido en la misma se obtienen nuevas estimaciones.

puntos que contiene la distribución utilizando la ecuación 17, para esto se aplica la transformación a cada parámetro de la distribución. De esta fase se obtendrá otra distribución Gaussiana que contendrá dichas predicciones

Para la distribución Gaussiana nueva se calculan la media y la matriz de covarianza. La media se calcula utilizando la ecuación 17, con lo cual tenemos que:

$$\hat{x}_t = \mathbf{F}_t \hat{x}_{t-1} \quad (19)$$

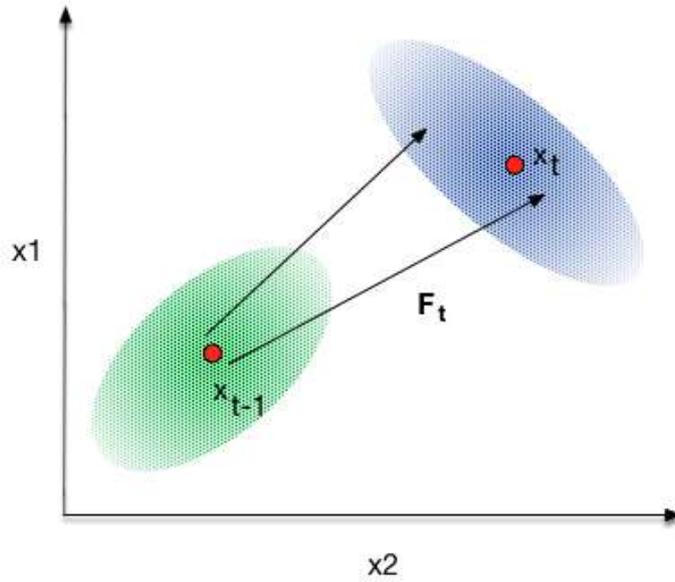


Figura 12: Predicción de la distribución Gaussiana de  $x_{t-1}$ .

Para calcular la nueva matriz de covarianza se emplea la siguiente propiedad

$$\begin{aligned} \text{Cov}(x) &= \Sigma \\ \text{Cov}(Fx) &= F\Sigma F^T. \end{aligned}$$

Entonces la nueva matriz de covarianza se calcula de la siguiente forma

$$P_t = F_t P_{t-1} F_t^T.$$

Así las ecuaciones para esta etapa para cálculo de la nueva Gaussiana con media  $x_t$  y matriz de covarianza  $P_t$ :

$$\hat{x}_t = F_t \hat{x}_{t-1} + w_t \quad (20)$$

$$P_t = F_t P_{t-1} F_t^T \quad (21)$$

### 2.7.2 Etapa de actualización

En esta etapa se utilizan las mediciones proporcionadas por los sensores, estas mediciones contienen información indirecta sobre el estado del sistema. Utilizando la matriz  $H_t$  de la ecuación 18 se puede establecer una relación entre la medición y el estado. Con esta

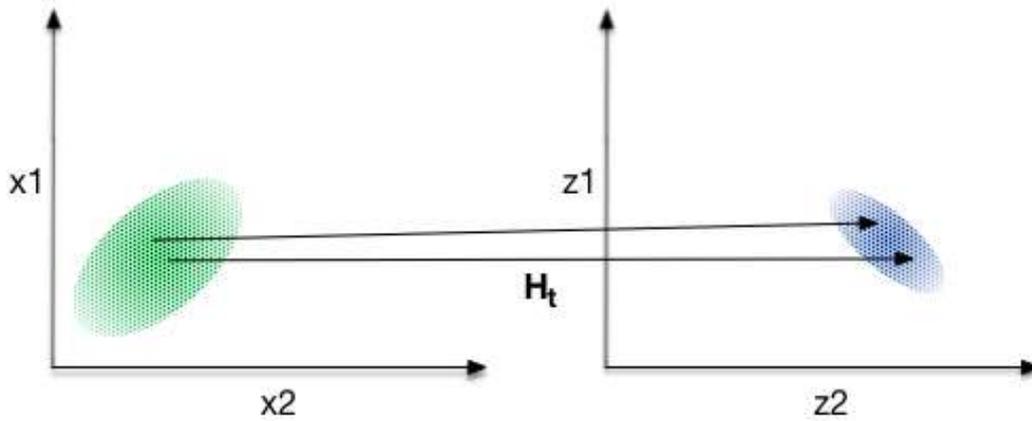


Figura 13: Transformación del espacio de estados al espacio de las mediciones.

relación se puede transformar la predicción de la etapa anterior al dominio de las mediciones.

Como la predicción  $\hat{x}_t$  es una variable aleatoria con distribución Gaussiana, deberá aplicarse una transformación para obtener otra distribución Gaussiana, esto se observa en la Figura 13, para realizar esta transformación se utilizan las ecuaciones 22 y 23.

$$\vec{\mu}_{\text{expected}} = \mathbf{H}_t \hat{\mathbf{x}}_t \quad (22)$$

$$\Sigma_{\text{expected}} = \mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T \quad (23)$$

En esta etapa también se dispone de la medición observada por los sensores, esta medición puede ser vista como una variable aleatoria, esto se debe a que el filtro de Kalman considera las mediciones como ruidosas.

La incertidumbre debida al ruido de los sensores puede ser representada con una matriz de covarianza  $\mathbf{R}_t$ . Se puede ver la medición como una variable aleatoria con distribución Gaussiana  $\mathcal{N}(z_t, \mathbf{R}_t)$  donde  $z_t$  es la medición obtenida con los sensores, esto se puede observar en la Figura 14. Ahora se tienen dos distribuciones Gaussianas que proporcionan información sobre el sistema, por un lado se tiene una distribución que indica la probabilidad de que la medición sea una

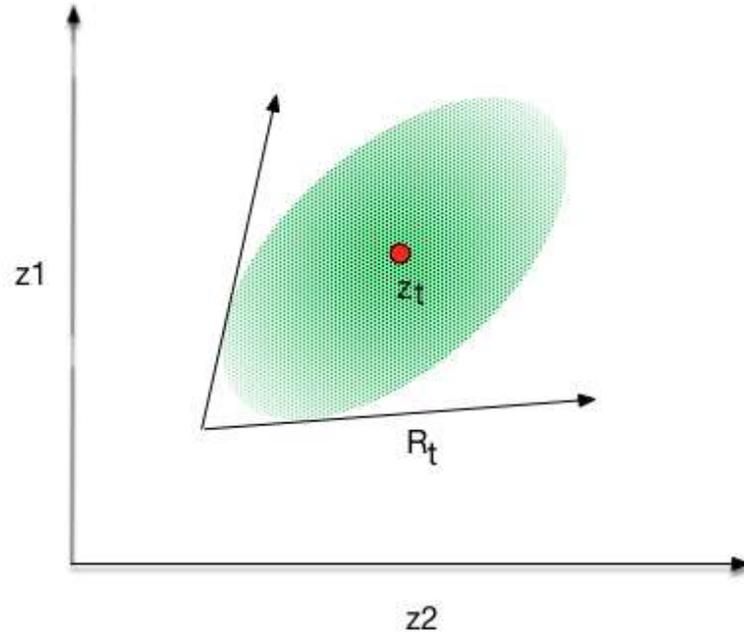


Figura 14: Distribución Gaussiana de la medición.

medición correcta y por otro lado la probabilidad del estado estimado pero en el dominio de las mediciones.

Con estas dos distribuciones de probabilidad se puede obtener una nueva distribución que indique la probabilidad de que ambas suposiciones son correctas. esto equivale a multiplicar ambas probabilidades. Esta multiplicación da como resultado otra distribución Gaussiana con su propia media y covarianza, la cual representa la mejor estimación.

Una distribución Gaussiana esta dada por la formula que se muestra en 24:

$$\mathcal{N}(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (24)$$

Si se tienen dos distribuciones Gaussianas  $G1 = (\mu_1, \sigma_1)$  y  $G2 = (\mu_2, \sigma_2)$  mediante el producto de ambas distribuciones de probabilidad se obtiene otra función Gaussiana, normalizando esa función Gaussiana para que la probabilidad total sea uno se obtiene una nueva distribución  $G' = (\mu', \sigma')$ .

Haciendo un poco de álgebra los parámetros de la nueva distribución quedan expresados como se muestra en la ecuación 25.

$$\begin{aligned}\mu' &= \mu_0 + \frac{\sigma_0^2(\mu_1 - \mu_0)}{\sigma_0^2 + \sigma_1^2} \\ \sigma'^2 &= \sigma_0^2 - \frac{\sigma_0^4}{\sigma_0^2 + \sigma_1^2}\end{aligned}\quad (25)$$

Para simplificar las ecuaciones, se introduce una variable llamada ganancia de Kalman, denominada  $\mathbf{k}$ , el cual es una factorización de algunos términos de la ecuación 25 y queda como se ve en 26.

$$\mathbf{k} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2}.\quad (26)$$

Luego entonces las ecuaciones en 25 se simplifican y obtienen las ecuaciones mostradas en 27.

$$\begin{aligned}\mu' &= \mu_0 + \mathbf{k}(\mu_1 - \mu_0) \\ \sigma'^2 &= \sigma_0^2 - \mathbf{k}\sigma_0^2\end{aligned}\quad (27)$$

Para el caso de las distribuciones Gaussianas en versión matricial se puede sustituir  $\Sigma$  y  $\mu$  en las ecuaciones mostradas en 27, las ecuaciones quedan como se muestra en 28.

$$\begin{aligned}\mathbf{K} &= \Sigma_0(\Sigma_0 + \Sigma_1)^{-1} \\ \bar{\mu}' &= \mu_0 + \mathbf{K}(\mu_1 - \bar{\mu}_0) \\ \Sigma' &= \Sigma_0 - \mathbf{K}\Sigma_0\end{aligned}\quad (28)$$

En el caso de las dos distribuciones Gaussianas para las cuales se quiere calcular el producto, se tienen la obtenida en la fase de predicción, transformada al dominio de las mediciones, por lo cual la distribución es:  $(\mu_0, \Sigma_0) = (\mathbf{H}_t \hat{\mathbf{x}}_t, \mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T)$  y la obtenida por los sensores, es decir la medición propiamente dicha, la cual esta representada por la distribución Gaussiana  $(\mu_1, \Sigma_1) = (\mathbf{z}_t, \mathbf{R}_t)$ . Usando las ecuaciones 28 y 26 se obtiene

$$\begin{aligned}\mathbf{H}_t \hat{\mathbf{x}}_t &= \mathbf{H}_t \hat{\mathbf{x}}_t + \mathbf{K}(\mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_t) \\ \mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T &= \mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T - \mathbf{K} \mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T \\ \mathbf{K} &= \mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1}\end{aligned}\quad (29)$$

Se puede eliminar el término  $H_t$  de las ecuaciones 29, las ecuaciones 30 son las ecuaciones finales de la etapa de actualización.

$$\begin{aligned}\hat{\mathbf{x}}_t &= \hat{\mathbf{x}}_t + \mathbf{K}(\mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_t) \\ \mathbf{P}_t &= \mathbf{P}_t - \mathbf{K} \mathbf{H}_t \mathbf{P}_t \\ \mathbf{K}_t &= \mathbf{P}_t \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1}\end{aligned}\tag{30}$$

### 2.7.3 Algoritmo

A continuación se presenta el algoritmo general del filtro de Kalman, utilizando las ecuaciones previamente descritas.

Paso 1	Inicialización	Inicializar $\hat{\mathbf{x}}$ y $\mathbf{P}$
Paso 2	Predicción	$\hat{\mathbf{x}}_t = \mathbf{F} \hat{\mathbf{x}}_{t-1}$ $\mathbf{P}_t = \mathbf{F} \mathbf{P}_{t-1} \mathbf{F}^T + \mathbf{Q}_t$
Paso 3	Matriz de ganancia de Kalman	$\mathbf{K}_t = \mathbf{P}_t \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_t \mathbf{H}_t^T + \mathbf{R}_t)^{-1}$
Paso 4	Actualización	$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_t)$ $\mathbf{P}'_k = \mathbf{P}_t - \mathbf{K}_t \mathbf{H}_t \mathbf{P}_t$
Paso 5	Volver al paso 2	

### 2.7.4 Filtro de Kalman Extendido

El filtro de Kalman asume que las ecuaciones que describen el sistema son lineales, pero en el mundo real muchos problemas no pueden ser modelados con sistemas lineales. Las dos fuentes de no linealidad del sistema se pueden presentar en el modelo de transición de estados o puede provenir de las mediciones.

Para desarrollar el filtro de Kalman, se considero que aplicar una función lineal a una Gaussiana produce otra Gaussiana, esto se ve en forma gráfica en la Figura 15. Esto no siempre ocurre en el caso de las funciones no lineales arbitrarias, Esto puede ser observado en la Figura 16

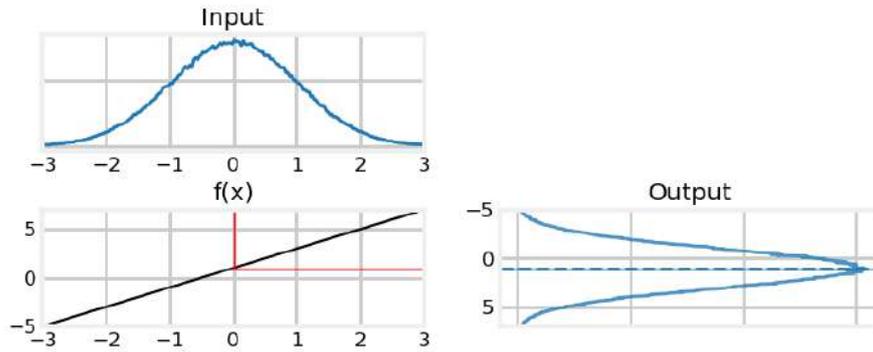


Figura 15: Distribución Gaussiana evaluada en una función lineal.

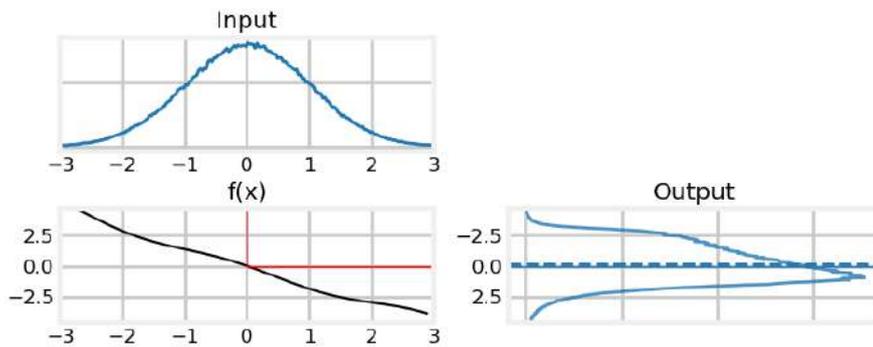


Figura 16: Distribución Gaussiana evaluada en una función no lineal.

Para solucionar este problema se desarrolló el filtro de Kalman Extendido, este algoritmo realiza una aproximación lineal utilizando series de Taylor en el punto de la estimación actual. Esta aproximación lineal es utilizada para por el filtro de Kalman visto en la sección anterior.

En el caso del filtro de Kalman como fue desarrollado originalmente, se utilizan dos ecuaciones para modelar el sistema las cuales son para modelar el proceso de evolución del sistema, Ecuación 17, y para el modelo de medición, el cual se puede observar en la ecuación 18.

Para el filtro de Kalman extendido la función de evolución del sistema  $F_t x_{k-1}$  será reemplazada por la función arbitraria  $f(z_t, x_t)$  y la función  $H_t x_{k-1}$  por  $h(z_t, x_t)$ , ambas pueden ser no lineales. Estas funciones se aproximan mediante una función lineal, un requisito para esto es que ambas funciones sean diferenciables. La aproximación lineal se realiza utilizando series de Taylor, por lo cual se requiere calcular Jacobiano.

Expandimos  $f(z'_t, x_t)$  utilizando series de Taylor, alrededor del punto  $(z'_i, x_{i|i-1})$ .

$$f(z'_i, x_i) = f(z'_i, x_i) + \frac{\partial f(z_i, x_{i|i-1})}{\partial z'_i} (z'_i - z_i) + \frac{\partial f(z_i, x_{i|i-1})}{\partial x_i} (x_i - x_{i|i-1}) + \dots$$

Solo tomaremos los términos de primer orden de la serie de Taylor, con esto obtendremos una nueva ecuación de medición.

$$y_i = M_i x_i + \xi_i$$

Donde:

$$M_i = \frac{\partial f_i(z_i, x_{i|i-1})}{\partial x_i}$$

$$y_i = -f_i(z_i, x_{i|i-1}) + \frac{\partial f(z_i, x_{i|i-1})}{\partial x_i} x_{i|i-1}$$

$$\xi_i = \frac{\partial f(z_i, x_{i|i-1})}{\partial z'_i} (z'_i - z_i)$$

Usando las ecuaciones anteriores se puede ver que:

$$E[\xi_i] = 0$$

$$E[\xi_i \xi_i^\top] = \frac{\partial f(z_i, x_{i|i-1})}{\partial z'_i} \Lambda_{\eta_i} \frac{\partial f(z_i, x_{i|i-1})}{\partial z'_i}^\top$$

El algoritmo Extendido de Kalman queda de la siguiente forma:

Paso 1	Inicialización	Inicializar $\hat{x}_0$ y $P_0$ $\hat{x}_t = f(x_{t-1})$
Paso 2	Predicción del Estado	$P_t = \frac{\partial f_t}{\partial x_t} P \frac{\partial f_t}{\partial x_t}^\top + Q_t$
Paso 3	Matriz de ganancia de Kalman	$K = P_t M_t^\top (M_t P_t M_t^\top + R_t)^{-1}$ $\hat{x}_t = \hat{x}_t + K'(y_t - H_t \hat{x}_t)$
Paso 4	Actualización	$\hat{x}_t = \hat{x}_t + K' f_t(z_t, x_t)$ $P_t = P_t - K_t M_t P_t$
Paso 5	Volver al paso 2	

## METODOLOGÍA

---

### 3.1 INTRODUCCIÓN

En este capítulo se describe la metodología utilizada en este trabajo para el seguimiento tridimensional de bastones. Como se observa en la figura 17, el proceso consta de seis etapas.

La primera etapa es de captura, mediante el sistema de cámaras estéreo, previamente calibrado, se obtienen dos imágenes de los bastones.

La segunda etapa consiste en la segmentación, en esta etapa se distinguen los posibles objetos de interés del fondo de la imagen, esta etapa se efectúa mediante modelos previamente calculados de los colores de los bastones. En esta etapa también se realiza la etiquetación de cada conjunto conectado de píxeles.

En la tercera etapa se procede a la detección de líneas rectas en la imagen, la entrada de esta etapa son los conjuntos de píxeles conectados, para detectar cuales conjuntos son potenciales bastones se emplea un vector de características, a los conjuntos que califiquen como líneas rectas se les calcula su ecuación de recta.

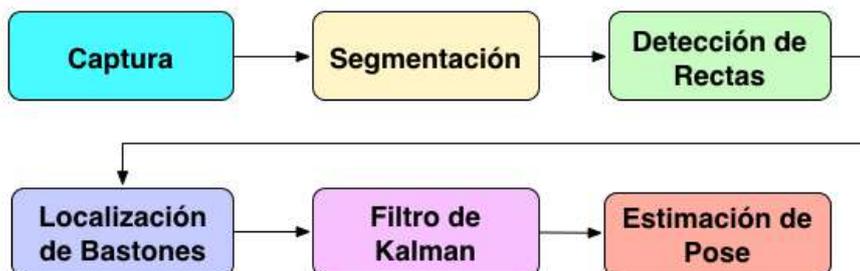


Figura 17: Diagrama general de la metodología.

En una siguiente etapa se efectúa la localización de bastones, ya se disponen de las ecuaciones de las proyecciones en 2D de la recta, utilizando ambas imágenes y las matrices del sistema de cámaras calibradas se realiza una localización 3D de los dos bastones respecto al sistema de cámaras, esto se realiza calculando la intersección de los dos planos formados por la proyección de las rectas en cada cámara.

Cuando ya se tienen dos o más capturas y se ha efectuado la localización tridimensional de los bastones en cada una, estas son usadas para formar la medición. Esta medición consiste en la representación 3D de la recta en el tiempo  $t$  y en el tiempo  $t - 1$ .

La medición obtenida se utiliza junto con en el filtro de Kalman para estimar los parámetros de movimiento de velocidad lineal y angular, mismos que servirán en la estimación de la pose.

### 3.2 SEGMENTACIÓN ROBUSTA

Para poder efectuar la segmentación se construye un modelo de color para cada tipo de bastón. Para construir el modelo se tomaron muestras de los bastones en diferentes ambientes de iluminación, para cada una de estas muestras se convierten sus píxeles al espacio CIELab. Se descarta el componente L el cual representa el brillo, se calculan la media y la matriz de covarianza sobre los otros dos componentes (a y b) con lo cual se obtiene una distribución del color. Este proceso se realiza fuera de línea.

Durante el funcionamiento del sistema cada par de imágenes capturadas por el sistema de cámaras estéreo se convierten al espacio de color CIELab. Se aplica un proceso de segmentación que consiste en calcular la distancia de Mahalanobis (véase sección 2.5) de cada píxel de la imagen a los modelos de color calculados previamente, aquellos píxeles que no pasan cierto umbral se consideran fondo. El resultado de esta fase es la separación de píxeles en grupos de píxeles que pertenecen a una figura y los píxeles que pertenecen al fondo.

Una vez segmentada la imagen se efectúa a un proceso de etiquetaron para identificar cada grupo de píxeles que se conectan entre sí.

Cada grupo es potencialmente un bastón, para discriminarlos se calcula por cada grupo contiguo unos rasgos para determinar si estos presentan características de bastón o no.

Con los grupos de pixeles pertenecientes a los bastones se procede a calcular la ecuación homogénea de la recta para cada uno de los que fueron considerados bastones.

### 3.3 RECONSTRUCCIÓN 3D DE LA RECTA

En esta sección se describe el proceso para la obtención de la línea recta en el espacio tridimensional. Para este trabajo se utiliza una representación de la línea recta tridimensional en coordenadas de Plücker, para obtener las coordenadas de Plücker se hace uso de un sistema de cámaras estéreo.

Se coloca un par de cámaras en una base, con una rotación y separadas simulando el sistema de visión humano. Estas cámaras se calibraron para obtener las matrices de proyección  $P_1$  y  $P_2$ , para calibrarlas se utiliza Matlab Calibration ToolBox [3].

Cada cámara proporciona una imagen a la cual, mediante el proceso previamente descrito, se le calculan las ecuación de las rectas homogéneas por cada bastón. Con estas ecuaciones se calculan los planos que pasan por las líneas y el origen en cada cámara. Con la intersección de estos planos se obtiene la línea recta en coordenadas de Plücker.

Para obtener la ecuación del plano se utiliza la matriz de proyección de la cámara, por cada cámara se tiene que:  $n_1 = P_1 l_1$  y  $n_2 = P_2 l_2$  donde los tres primeros componentes representan la normal del plano y el cuarto componente la distancia al origen. Calculando la intersección de estos planos se obtiene la recta en el espacio 3D con coordenadas de Plücker.

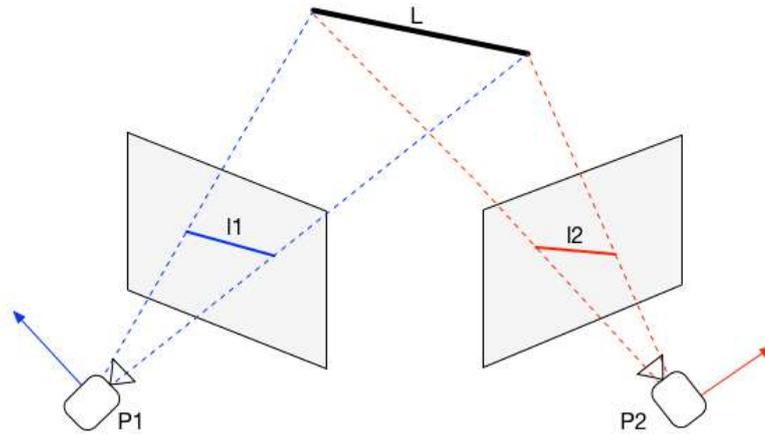


Figura 18: Sistema de cámaras estereo.

### 3.4 ESTIMACIÓN DE PARÁMETROS DE MOVIMIENTO

En este trabajo se consideran a las líneas rectas como cuerpos rígidos, por lo cual se puede describir su movimiento mediante la cinemática de cuerpos rígidos, con las ecuaciones vistas en la sección 2.6. Para la estimación de estos parámetros y asumiendo que nuestras mediciones provenientes de la reconstrucción tridimensional son ruidosas, se utiliza el filtro de Kalman extendido visto en la sección 2.7.

Para aplicar el filtro de Kalman se debe definir el estado que se desea estimar, las mediciones que se obtienen con la cámara, así como una función que relacione dichas mediciones con el estado.

En este trabajo se define el estado como las velocidades lineal y angular definidas por  $v$  y  $w$  respectivamente, esto con el fin de poder efectuar una predicción de la pose en el tiempo  $t + 1$ , por tanto el estado se  $x$  definió como  $x = (w, v)$ .

Se evalúan dos representaciones de la recta, la primera utilizando coordenadas de Plücker, y otra utilizando el plano definido por los dos bastones y el punto de intersección.

3.4.1 *Coordenadas de Plücker*

Para el caso de las coordenadas de Plücker se definen las mediciones como el par de rectas en el espacio tridimensional en el tiempo  $t$  y en el tiempo  $t - 1$ , para representar la recta se utilizan las coordenadas Bi-normalizadas de Plücker[1] con lo cual una medición en el tiempo  $t$  queda de la siguiente manera

$$z = (h_t, u_t, m_t, h_{t-1}, u_{t-1}, m_{t-1}) \quad (31)$$

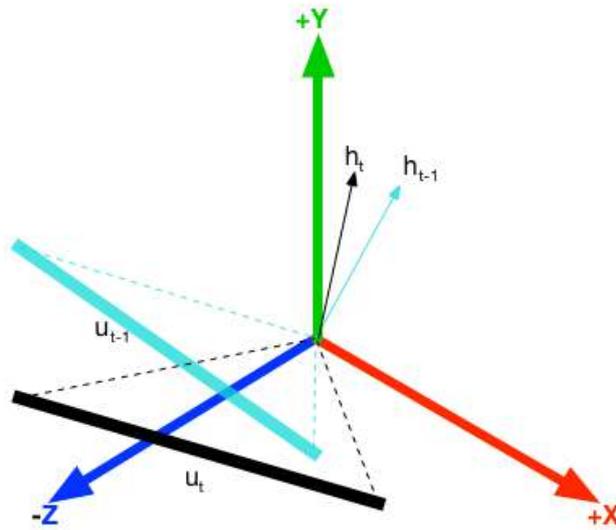


Figura 19: Recta en el tiempo  $t - 1$  y en el tiempo  $t$ .

Donde  $h_t$  y  $h_{t-1}$  son los vectores normales al plano formado por la recta y el origen,  $u_t$  y  $u_{t-1}$  definen la orientación de la recta así como  $m_t$  y  $m_{t-1}$  definen la distancia de la recta al origen.

Sabemos que en una representación de la línea recta en coordenadas de Plücker normalizadas, la  $u$  representa la orientación y la  $h$  la normal al plano. Por lo tanto, cuando la recta se mueve con una velocidad angular dicha orientación cambia, con lo cual la  $u$  se ve afectada por una matriz de rotación. Siendo  $\omega$  la velocidad angular instantánea, podemos calcular la matriz de rotación integrando con respecto a  $\Delta t$ , utilizando la fórmula de Rodrigues [9].

Si aplicamos la matriz de rotación resultante a  $u$  en el tiempo  $t$ , obtendremos  $u$  en el tiempo  $t + 1$ , por tanto  $Wu_t \times u_{t+1} = 0$ . Como menciona [1] la derivada de  $m$  con respecto al tiempo es:

$$\dot{m} = V^T(u \times h) \quad (32)$$

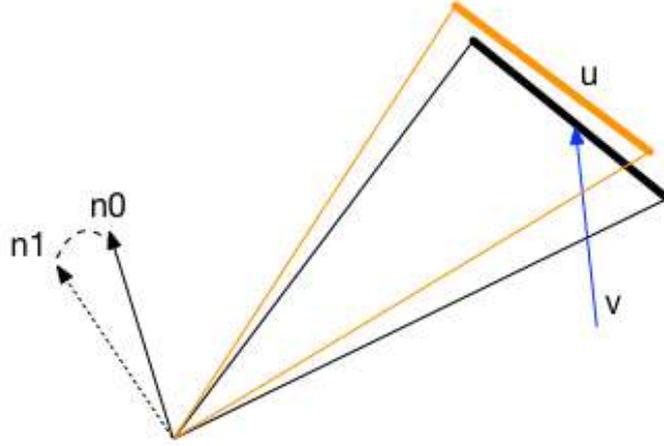


Figura 20: Efecto de la velocidad lineal sobre el vector normal.

Por tanto la función que relaciona la medición con el estado esta dada por:

$$f(z, s) = \begin{bmatrix} Wu_1 \times u_2 \\ m_2 - m_1 - (V^T(u_1 \times h_1))\Delta t \end{bmatrix} = 0 \quad (33)$$

Donde  $W$  es la matriz de rotación generada por la velocidad angular  $\omega\Delta t$  y  $R$  es la matriz de rotación al rededor del eje  $\omega - \frac{Vh}{m}u$ . Para poder utilizar esta función en el filtro de Kalman Extendido, necesitaremos calcular la matriz Jacobiana con respecto a la medición y con respecto al estado. El desarrollo de la matriz Jacobiana puede verse en el Apéndice A.

#### 3.4.2 Plano y punto de intersección

Para este caso se definen las mediciones como la normal del plano formado por las dos rectas de los bastones y el punto de intersección de los mismos.

Para obtener la norma del plano formado por ambas rectas se utilizan las coordenadas de Plücker obtenidas de cada recta en la etapa de localización, se toman los tres componentes de las coordenadas de Plücker que indica la dirección de la recta, y realizamos el producto cruz de de ambos, de esta operación obtendremos la normal al plano formado por ambos bastones.

Para obtener un punto en dicho plano, asumiendo que las proyecciones de las rectas se intersectan en cada cámara, se usa el punto de intersección de cada proyección y se lleva acabo una triangulación para obtener el punto 3D correspondiente.

Con esto se obtiene una medición que involucra a la normal del plano formado por las rectas y un punto del mismo, se optó por utilizar un punto y no la distancia, puesto que el punto proporciona mas información sobre la velocidad lineal en sus 3 ejes. La medición queda como se ve en la ecuación 34.

$$z = (n_t, p_t, n_{t-1}, p_{t-1}) \quad (34)$$

Para obtener la función de relación entre la medición y el estado compuesto por los parámetros de movimiento, se utilizó el modelo cinemático de un cuerpo rígido, como se menciona en el capítulo del marco teórico, el modelo de la velocidad del movimiento esta dado por la ecuación 14. De acuerdo con Z. Zhang y D. Faugeras [15, 6], la solución de esta ecuación para una velocidad angular constante y una velocidad lineal polinomial esta dada por la ecuación 35.

$$P_t = W_{p_o} + Vv \quad (35)$$

Donde:

- $W = I_3 + \frac{\sin(\theta\delta)}{\theta} \omega + \frac{1-\cos(\theta\delta)}{\theta^2}$
- $V = I_3\Delta t + \frac{1-\cos(\theta\Delta t)}{\theta^2} \omega + \frac{\theta\Delta t - \sin(\theta\Delta t)}{\theta^3} \omega^2$

Mediante este modelo podemos definir la función que relaciona la medición con el estado esta dada por:

$$f(z, s) = \begin{bmatrix} Wn_{t-1} \times n_t \\ Wp_{t-1} + Vv - p_t \end{bmatrix} = 0 \quad (36)$$

Una vez definidas estas funciones, se procede a la implementación del filtro de Kalman, utilizando para ello el filtro de Kalman extendido debido a que el modelo planteado no es lineales. Esto requiere el calculo de las matrices Jacobianas con respecto a la medición y al estado, para posteriormente realizar la aproximación lineal con series de Taylor.

### 3.5 RESUMEN

En este capítulo se describió el proceso para realizar el seguimiento de los dos bastones en el espacio tridimensional a partir de un sistema de cámaras estéreo calibradas, sobreponiendo en cada imagen el resultado de el producto cruz de los dos vectores formados por los bastones. Esto se logró mediante la segmentación y el etiquetado de cada par de imágenes, utilizando segmentación en base modelo de color. Esto dio como resultado un par de imágenes con conjuntos de pixeles conectados. Utilizando un vector de rasgos se determinó cuales conjuntos representan a bastones, se asumió que habrían únicamente dos bastones en la escena, ambos de diferente color.

Con los bastones identificados se utilizó la información proporcionada por el par de cámaras calibradas para localizar los bastones en la escena en coordenadas de Plücker, así como el punto de intersección, utilizando el punto de intersección en 2D en cada imagen, este punto de intersección junto con la orientación que forma parte de las coordenadas de Plücker sirvieron como vector de medición para el filtro de Kalman extendido, el cual nos ayudó a estimar los parámetros de movimiento cinemática de los bastones.

Cuando la localización de los bastones no era posible, se procedió a utilizar los parámetros estimados de movimiento para predecir la posición que debería tener el sistema de bastones.

Finalmente una vez teniendo la pose de los bastones, se realizó el producto cruz entre ambos bastones y se re-proyector el resultado encima de las imágenes capturadas.

En el próximo capítulo se describen los experimentos realizados basándose en la metodología presentada, en una primera etapa se utilizaron simulaciones del movimiento de los bastones para evaluar el filtro de Kalman, posteriormente se integro el filtro de Kalman a los algoritmos de visión descritos.



## EXPERIMENTOS

---

### 4.1 INTRODUCCIÓN

En este capítulo se describen los experimentos realizados utilizando la metodología presentada en el capítulo anterior. La realización de los experimentos se dividió en dos fases, en la primera fase se realizaron simulaciones con datos artificiales, en la segunda fase se utilizó el sistema con datos reales capturados con un sistema de cámaras estéreo.

Para la primera fase se compararon dos representaciones del sistema de bastones. Esto con el fin de determinar cual de ellas es mas conveniente a aplicar, es decir con cual se obtienen mejores estimaciones al aplicar el filtro de Kalman.

En la primera representación cada bastón individual, es representado mediante una coordenada de Plücker. En la segunda representación se toma el punto de intersección donde convergen los bastones y el vector normal a la superficie formada por ambos, esta representación es similar a la utilizada en la sección 3.4.2 para la estimación de movimiento de segmentos de rectas, la cual consiste en el punto medio del segmento y su orientación, es por esto que para la simulación se toma solo un segmento de recta.

La segunda fase se procedió a integrar dicha representación al sistema completo, utilizando como entrada las etapas de procesamiento de imágenes previamente descritas en la metodología, para su captura se empleó un sistema estereoscopio formado por dos cámaras.

Al final de este capítulo se explican los resultados obtenidos tanto en los experimentos artificiales así como los resultados que se obtu-

vieron utilizando el sistema de cámaras estéreo y se presenta una discusión.

## 4.2 SIMULACIONES

Los experimentos con datos artificiales fueron programados en lenguaje Python versión 2.7, para realizar los cálculos y operaciones con matrices se utilizó la biblioteca de computo científico NumPy y para la generación de gráficos se utilizó la biblioteca Matplotlib.

El entorno utilizado fue una distribución de Linux, Fedora 25.

Para generar los datos artificiales se realizó una simulación del movimiento de un segmento de recta, tomando las coordenadas de los extremos de un segmento el cual representa un bastón o referente físico.

### 4.2.1 Generación de Movimiento

Se simuló el movimiento de los bastones aplicando las ecuaciones de movimiento de cuerpo rígido a cada extremo del segmento modelado.

Las ecuaciones de movimiento describen la velocidad de un cuerpo rígido. Para obtener coordenadas de posición en cada instante se realizó una integración numérica que se puede observar en la ecuación 37 utilizando  $\Delta = 1/60$ . La trayectoria generada por la simulación puede observarse en la figura 21.

$$p_{t+1} = p_t + p'_t \Delta t \quad (37)$$

Donde  $p'_t$  esta dada por la ecuación 16 vista en la sección 2.6.2 del marco teórico.

El experimento recibe cuatro parámetros: el vector del eje de rotación, la velocidad angular (radianes/segundo), un vector unitario

de orientación del movimiento lineal y la magnitud de la velocidad lineal. Se consideró que tanto la orientación como el eje de rotación son unitarios, de no ser así se procede a normalizar el vector. Con estos cuatro parámetros se calcula  $w$  y  $v$  reales.

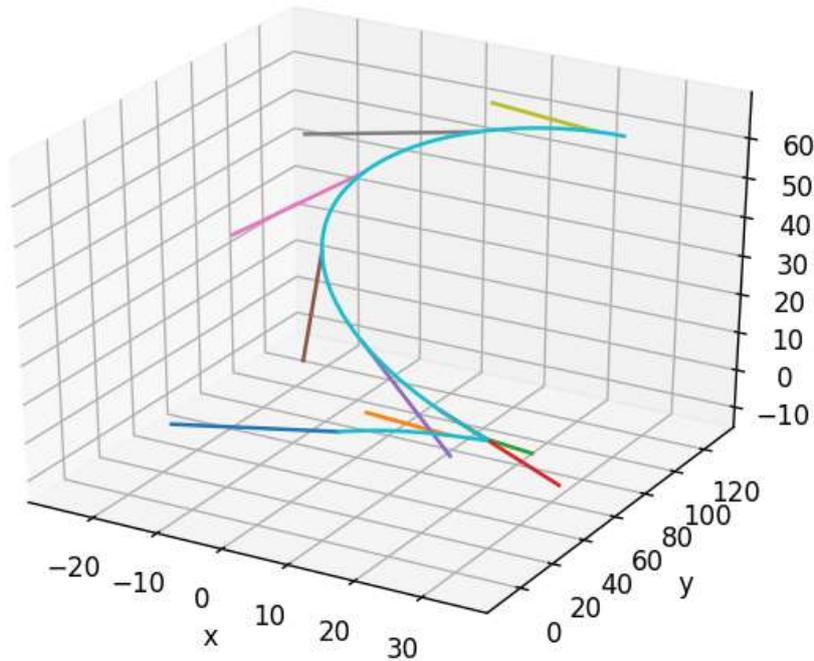


Figura 21: Movimiento simulado de un bastón, la curva verde representa la trayectoria.

Esta simulación genera una serie de segmentos de recta calculando la trayectoria de sus dos extremos, a estos extremos se les aplicó un ruido Gaussiano con el objetivo de simular ruido en las mediciones.

Como ya se mencionó, se hicieron experimentos utilizando dos representaciones de la recta, una propuesta por Faugeras, la cual representa una recta utilizando el punto medio y un vector de dirección y la representación en coordenadas de Plücker mencionada en capítulo 2.

En cada caso se tomaron los segmentos de recta y se convirtieron a sus representaciones correspondientes utilizando las ecuaciones ?? y ?? respectivamente.

Para cada experimento se generó una secuencia de 20 segundos, con el fin de observar el comportamiento un tiempo suficiente para determinar que el filtro de Kalman converge. El tiempo de 20s se obtuvo de manera empírica al observar que el error no decrecía.

Para medir el error en las simulaciones, se comparan los parámetros originales con las estimaciones hechas por el filtro de Kalman. Para medir el error en la velocidad lineal obtiene la diferencia absoluta entre los elementos del vector, en el caso de la velocidad angular se utilizó el ángulo formado por el eje de rotación estimado y el real, así como la magnitud.

#### 4.2.2 *Aproximación Inicial*

Para aproximar la velocidad lineal en el caso de la representación de Plücker no se encontró una clara forma de hacerlo, una aproximación fue la de buscar el punto más cercano al origen que pertenece a la recta, esta aproximación no funciona muy bien, puesto que el movimiento podría ser en la dirección de la recta y el punto no se movería, produciendo una estimación de velocidad cero. Para el caso donde contamos con el punto medio de un segmento, se puede estimar simplemente por la diferencia entre sus puntos medios.

Para la aproximación de la velocidad angular, en ambos casos se cuenta con el vector que indica la orientación de la recta. Se puede calcular la magnitud de la velocidad angular utilizando el ángulo formado entre dos vectores. Para calcular el eje de rotación se necesitan tres mediciones. En la figura 22, se observan  $u_1$ ,  $u_2$  y  $u_3$ , podemos estimar  $\sigma$  como  $(u_1 - u_2) \times (u_2 - u_3) = \omega$ .

#### 4.2.3 *Representación de Plücker*

Para el caso de la representación con coordenadas de Plücker, como se explicó en el capítulo anterior, se generaron datos artificiales para la normal y la dirección de la recta afectados por una velocidad lineal

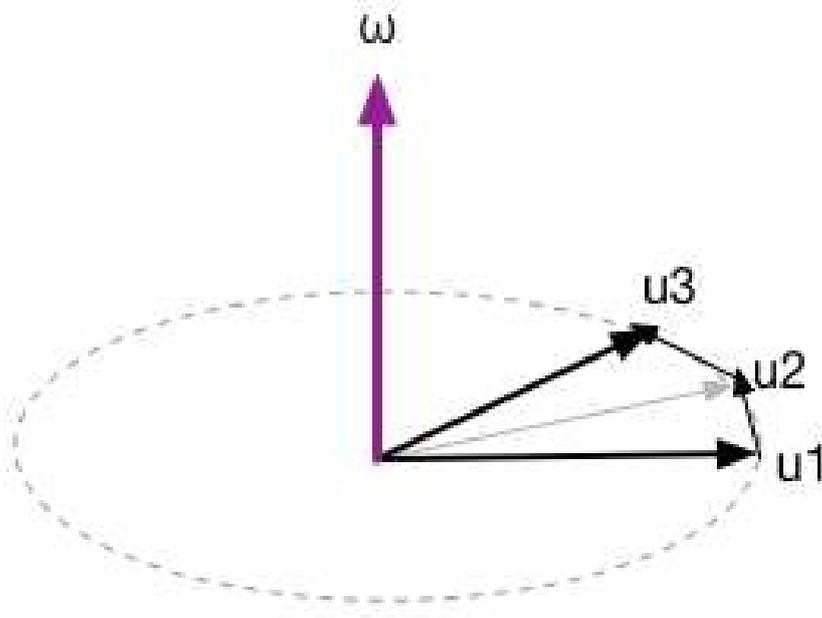


Figura 22: Tres mediciones que se tomaron para hacer la estimación inicial.

y angular constantes, a estos datos se les aplico un ruido Gaussiano aleatorio.

La gráfica en la Figura 23 se observa el error en cada componente del vector de velocidad lineal, esto indica físicamente como cambia la posición en cada componente con respecto al tiempo. En esta se observa como el error en la estimación inicial es muy alto, posteriormente decrece para estabilizarse al rededor de los 5.0 segundos.

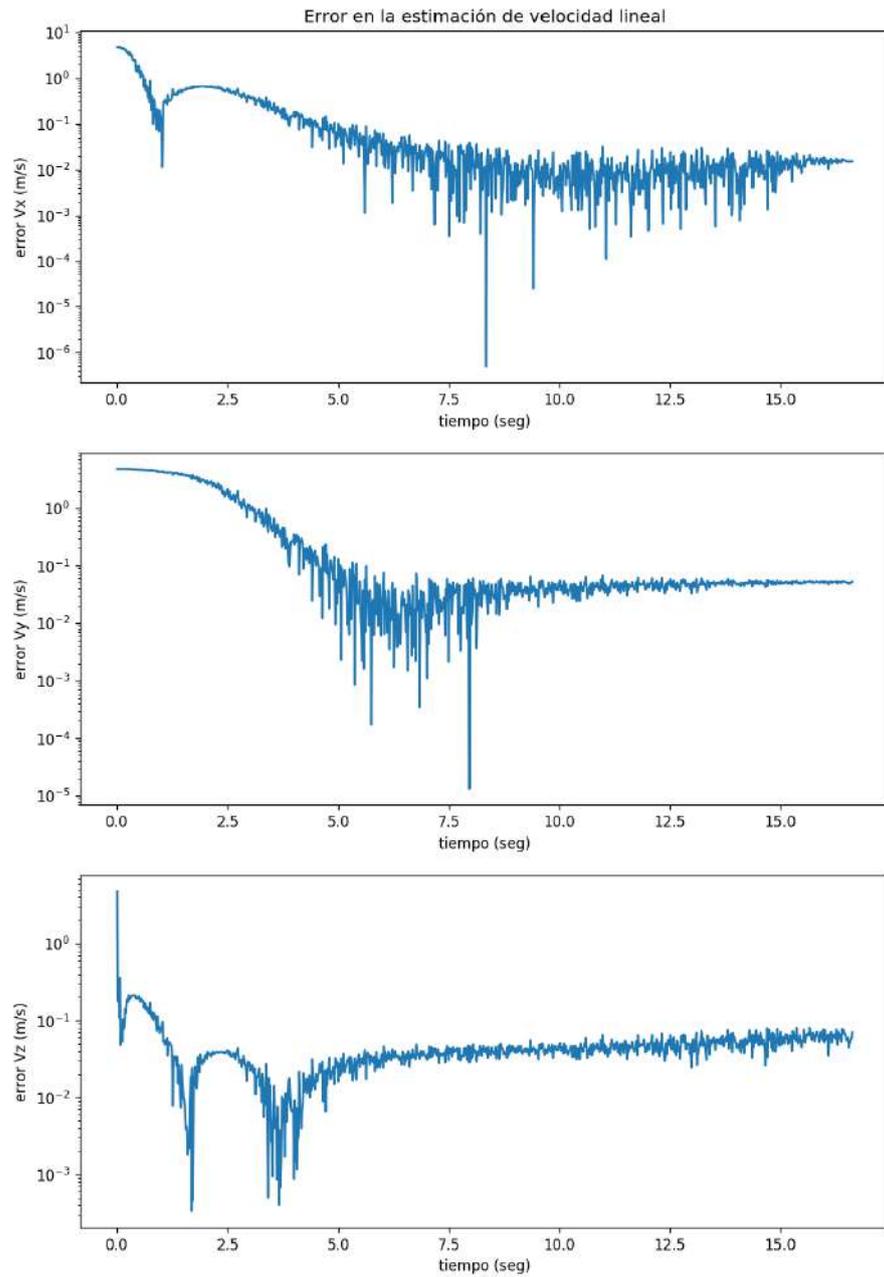


Figura 23: Gráficas de error de los componentes del vector de velocidad, utilizando la representación de recta de Plücker.

En la Figura a 24 se graficó el intervalo de  $-3\sigma$  a  $3\sigma$ , observamos como la mayoría de los errores caen en este intervalo, la varianza tiene un decremento hasta cierto nivel, luego del cual no ocurre un decremento mayor. Para la generación de esta gráfica se empleó la representación de recta de Plücker.

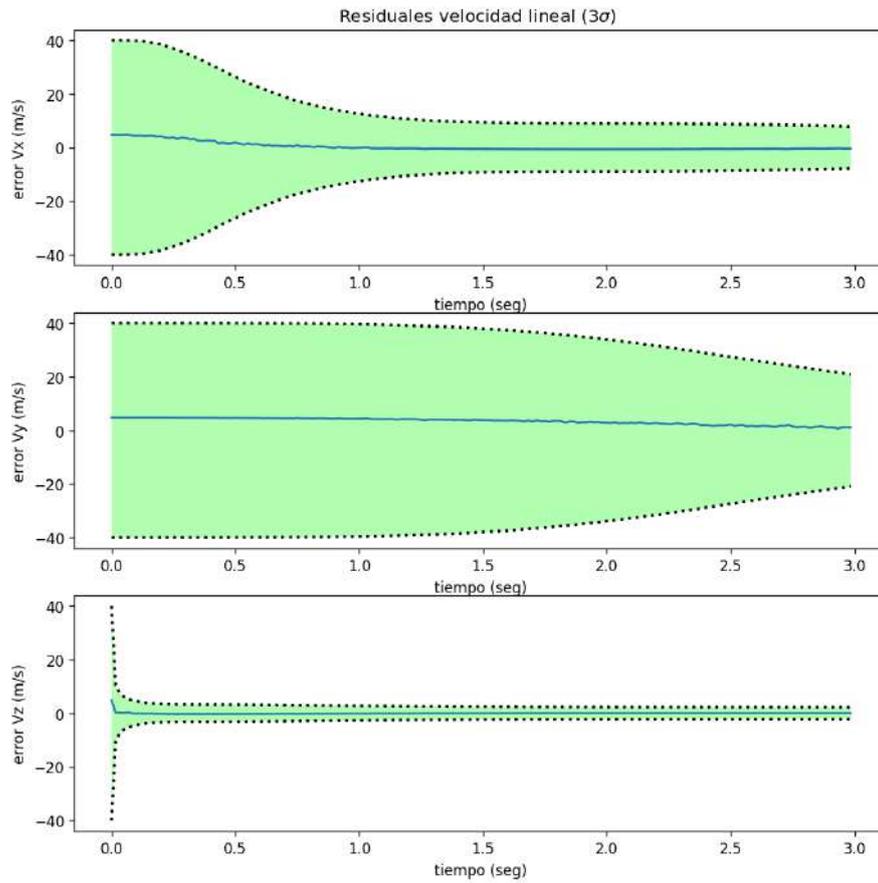


Figura 24: Gráfica de la diferencia entre la velocidad angular real de la simulación y la estimada por el filtro de Kalman, el área verde representa  $3\sigma$  desviaciones estándar, se observa que la mayoría de los errores caen en ese rango. Para la generación de esta gráfica se empleó la representación de recta de Plücker.

Lo mismo ocurre para la estimación de la velocidad angular, estos resultados pueden observarse en las Figura 25 y 26, donde se muestran los resultados para la estimación del eje de rotación así como la magnitud en radianes por segundo.

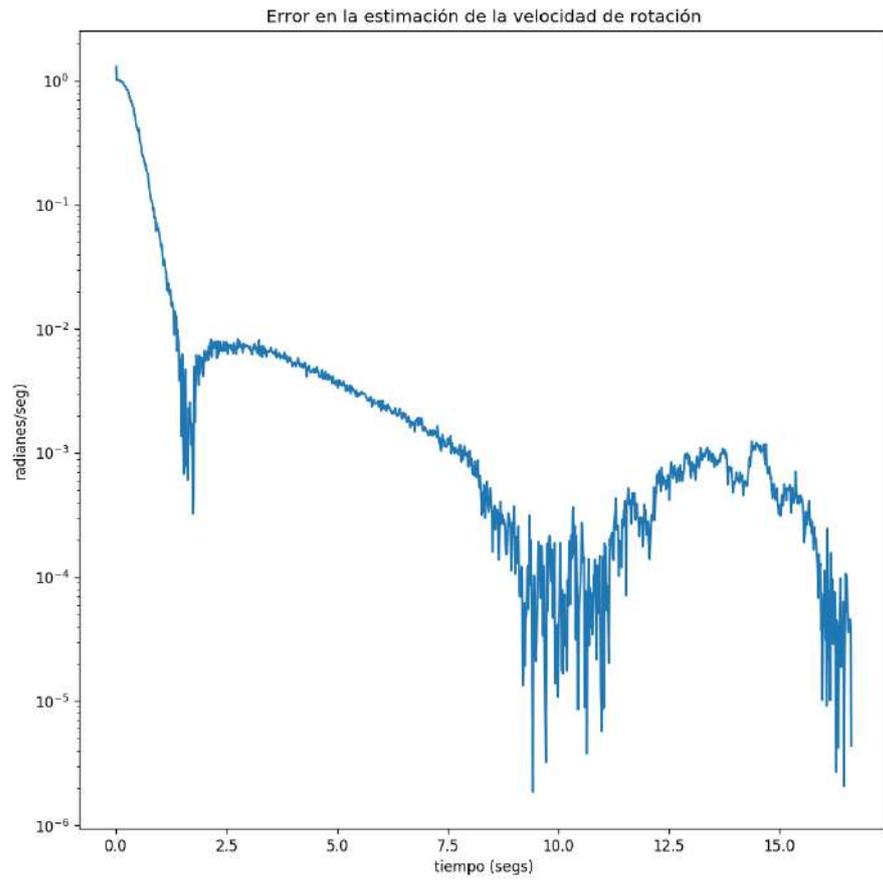


Figura 25: Gráfica de error en la magnitud de la velocidad de rotación para la representación de recta utilizando las coordenadas de Plücker.

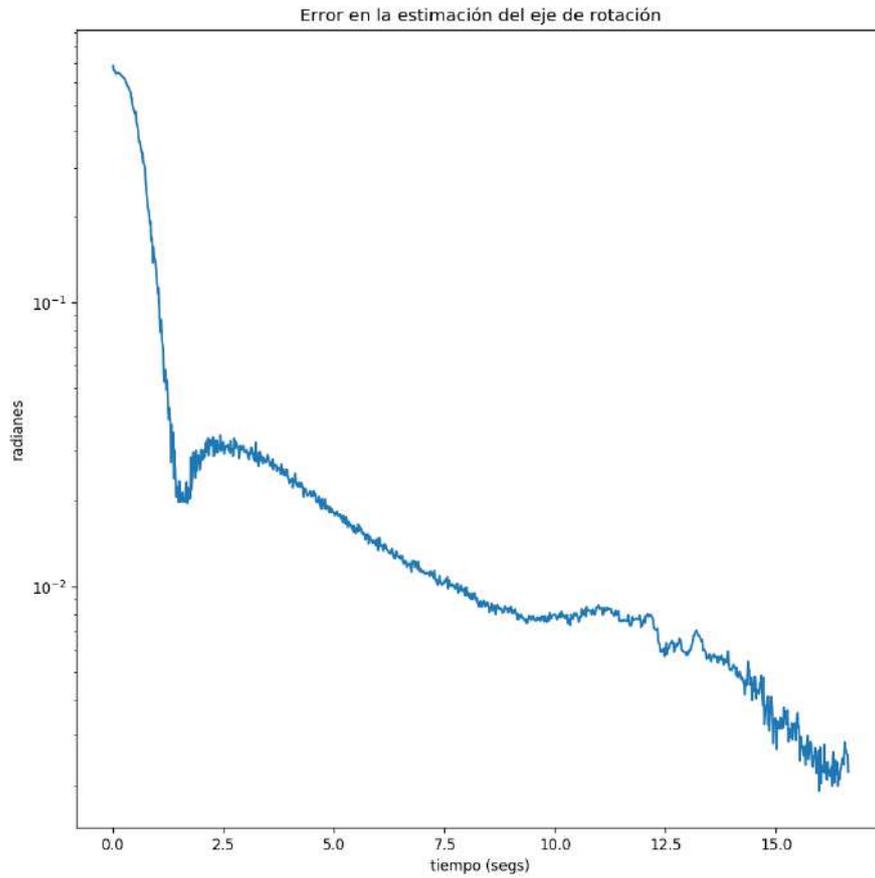


Figura 26: Gráfica del error en la estimación del eje de rotación, ángulo entre el vector original de rotación y el estimado, utilizando la representación de recta de Plücker..

En la figura 27 también se muestran los resultados de los errores con respecto a desviación estándar, la cual fue tomada de la traza de la matriz de covarianza.

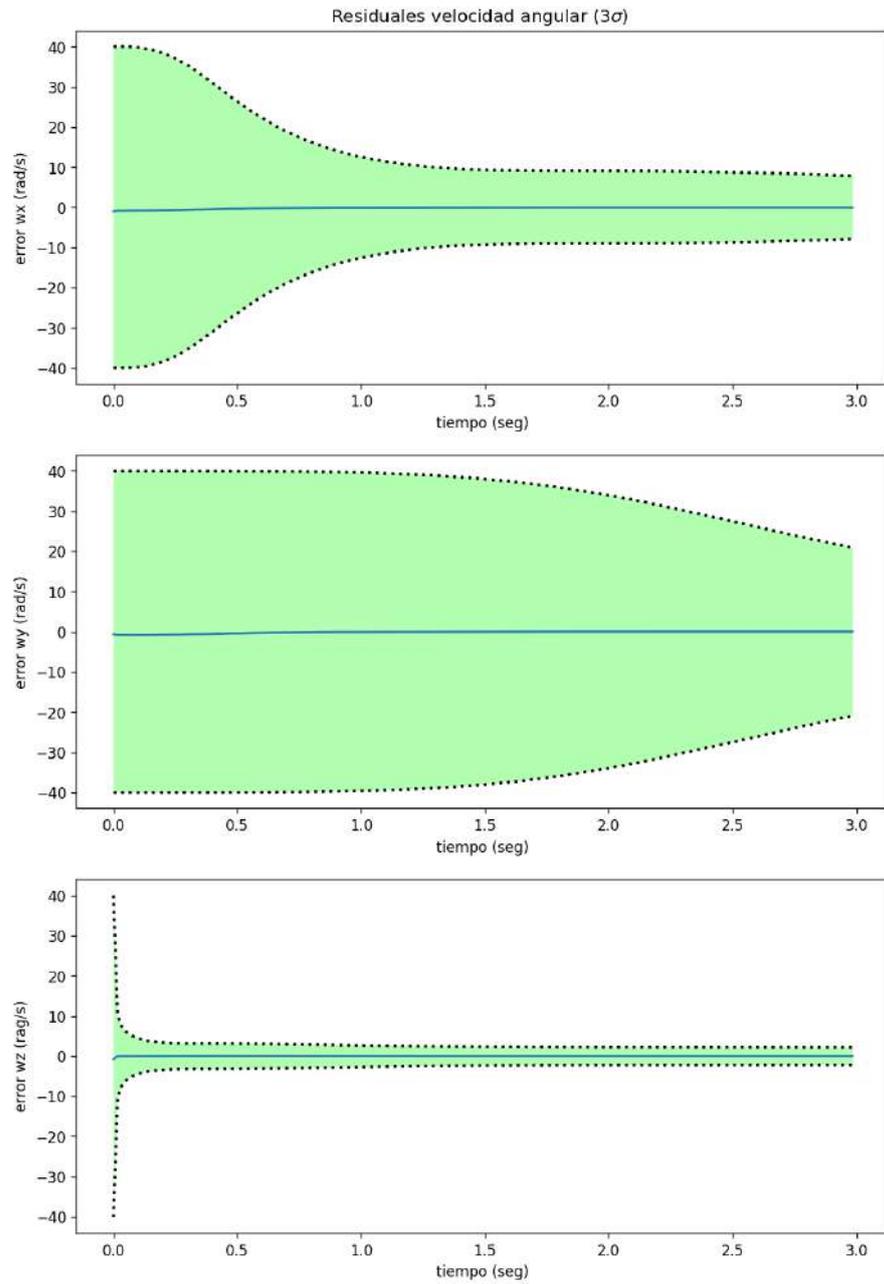


Figura 27: Gráfica de la diferencia entre la velocidad lineal real de la simulación y la estimada por el filtro de Kalman, el área verde representa  $3\sigma$  desviaciones estándar, se observa que la mayoría de los errores caen en ese rango. Para la generación de esta gráfica se empleó la representación de recta de Plücker.

#### 4.2.4 *Representación con punto medio y orientación*

Se utilizó el punto medio y la orientación como representación de un segmento de recta, mediante esta representación se planteó una función para relacionar la medición con el estado, como se explicó en el capítulo anterior.

Para este experimento se programó en Python una simulación del movimiento de los bastones utilizando la representación previamente descrita, pero se calculó el punto medio y la orientación de la recta para cada muestra y esto se utilizó como medición para el filtro de Kalman.

Los resultados de cómo va cambiando el error angular en el tiempo se pueden observar en las figuras 28 y 29, para el error lineal se muestra en gráfica en la figura 30.

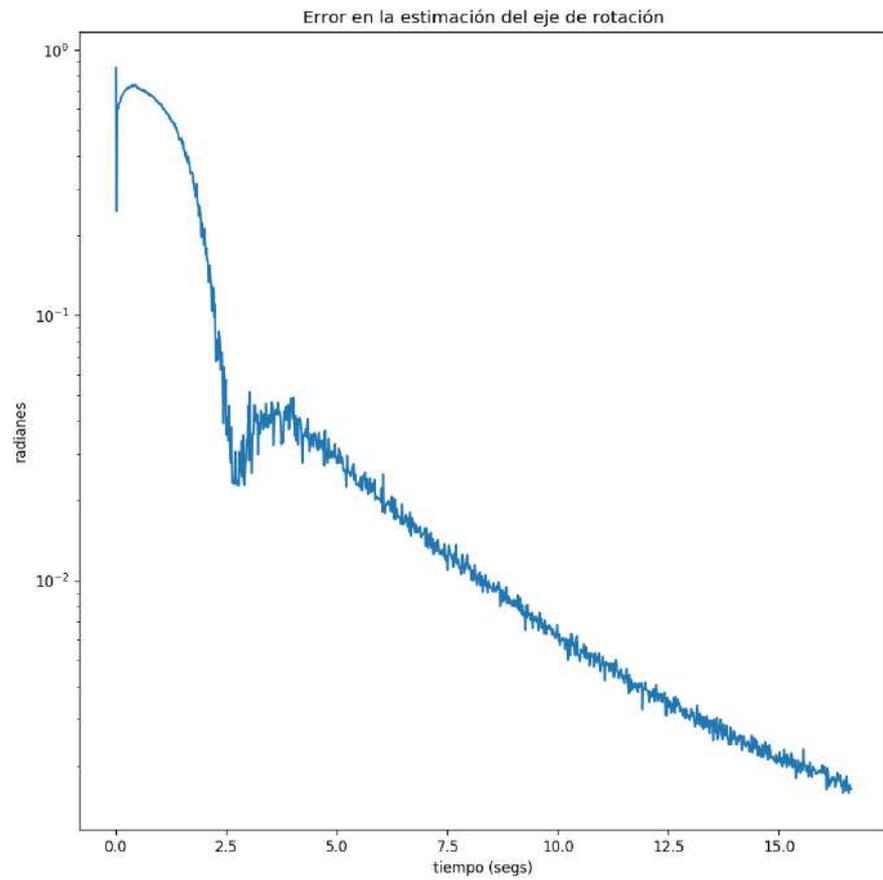


Figura 28: Gráfica del error en la estimación del eje de rotación, ángulo entre el vector original de rotación y el estimado, utilizando el punto medio y la orientación de un segmento de recta.

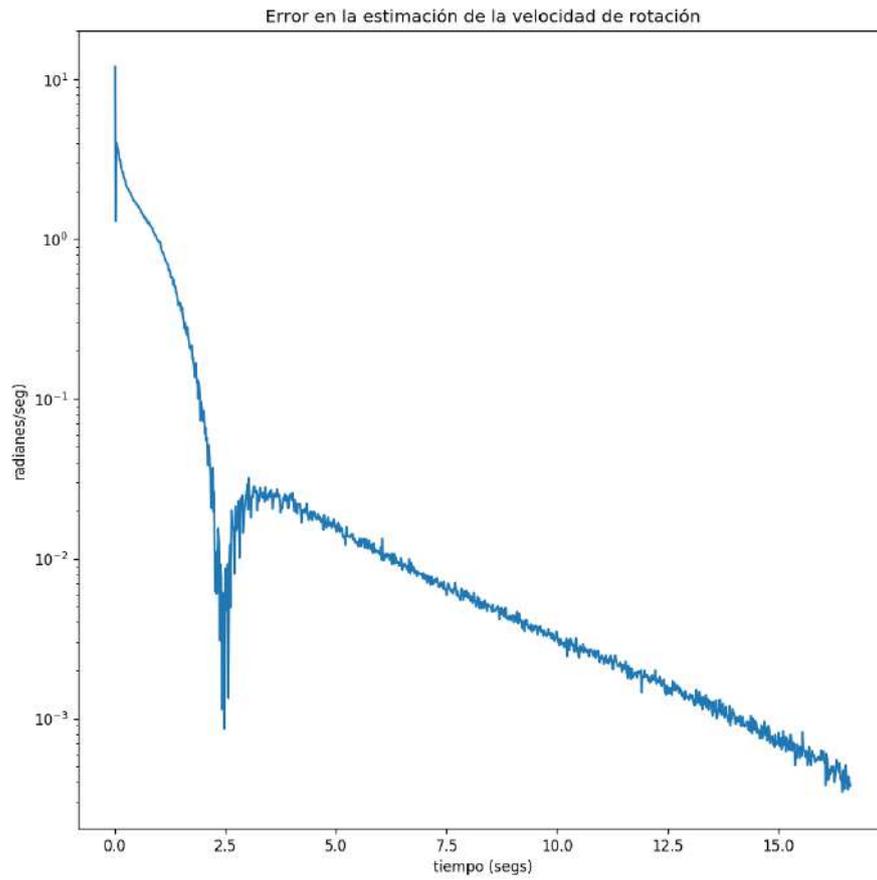


Figura 29: Gráfica de error en la magnitud de la velocidad de rotación para la representación de recta utilizando la representación de segmento de recta, con la orientación y el punto medio.

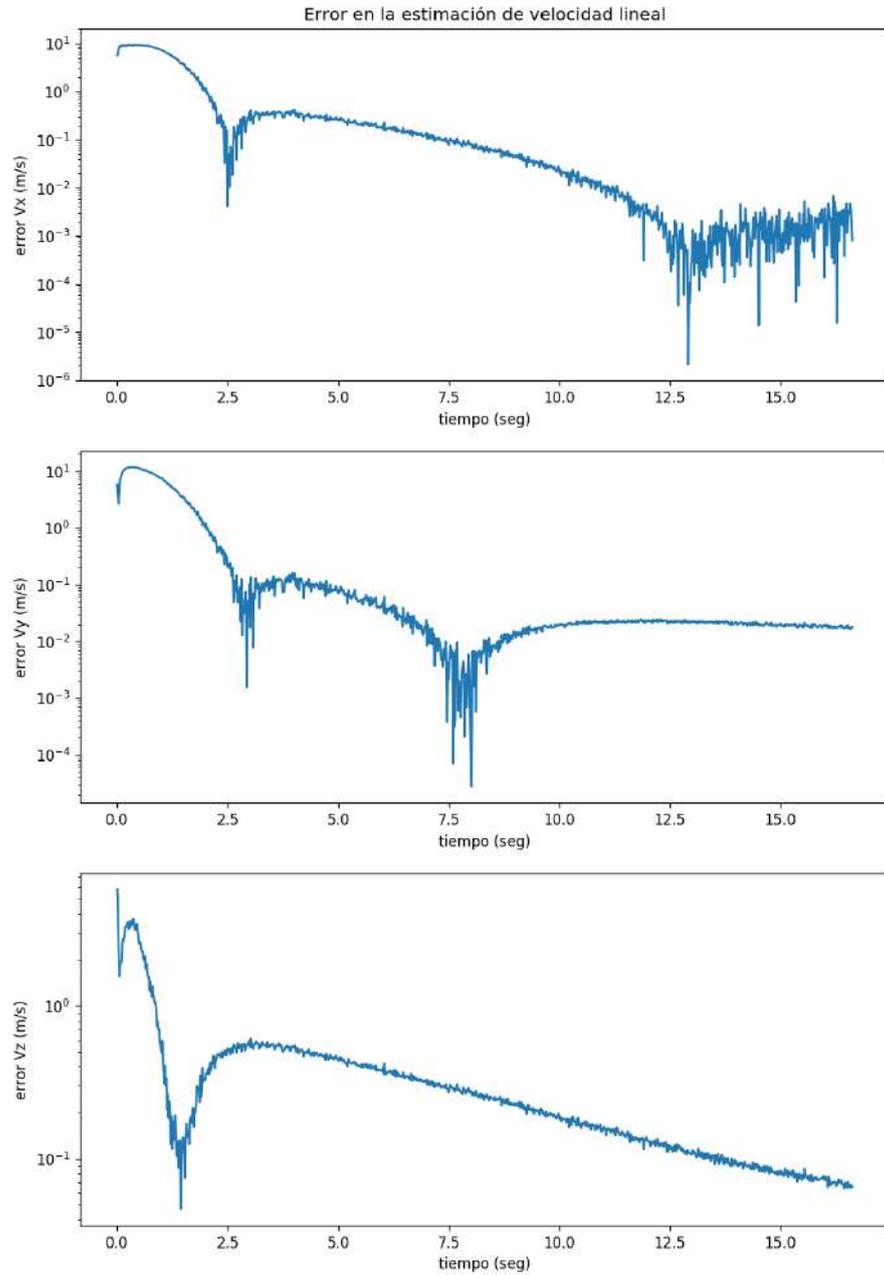


Figura 30: Gráficas de error, arriba el error en la pose en el momento de realizar la predicción, abajo el error respecto al estado real, utilizando el punto medio y la orientación de un segmento de recta.

Adicionalmente, como se hizo en el caso de la representación de Plücker, se generó una gráfica del rango de la desviación estándar, de  $-3\sigma$  a  $3\sigma$  y los resultados del error.

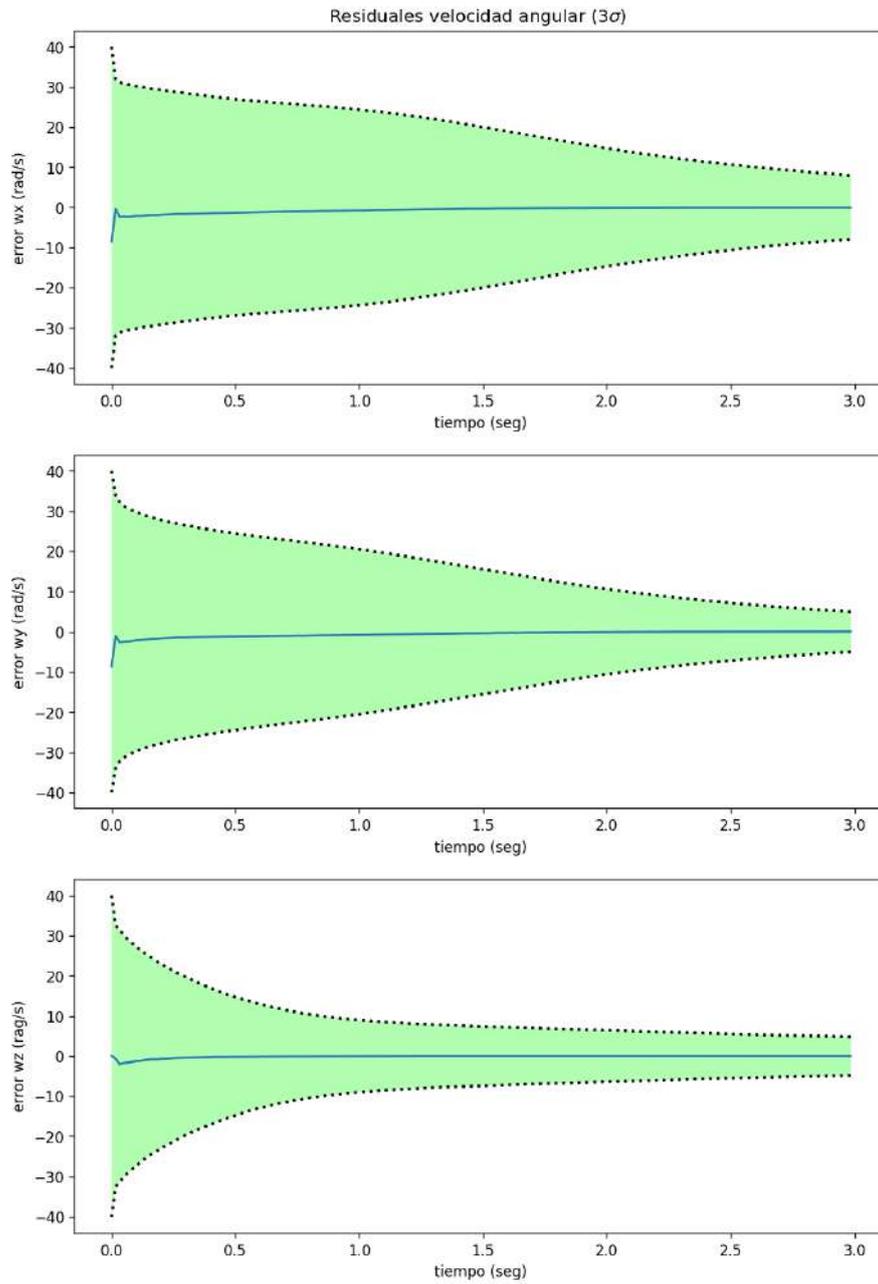


Figura 31: Gráficas de error, arriba el error en la pose en el momento de realizar la predicción, abajo el error respecto al estado real, utilizando el punto medio y la orientación de un segmento de recta.

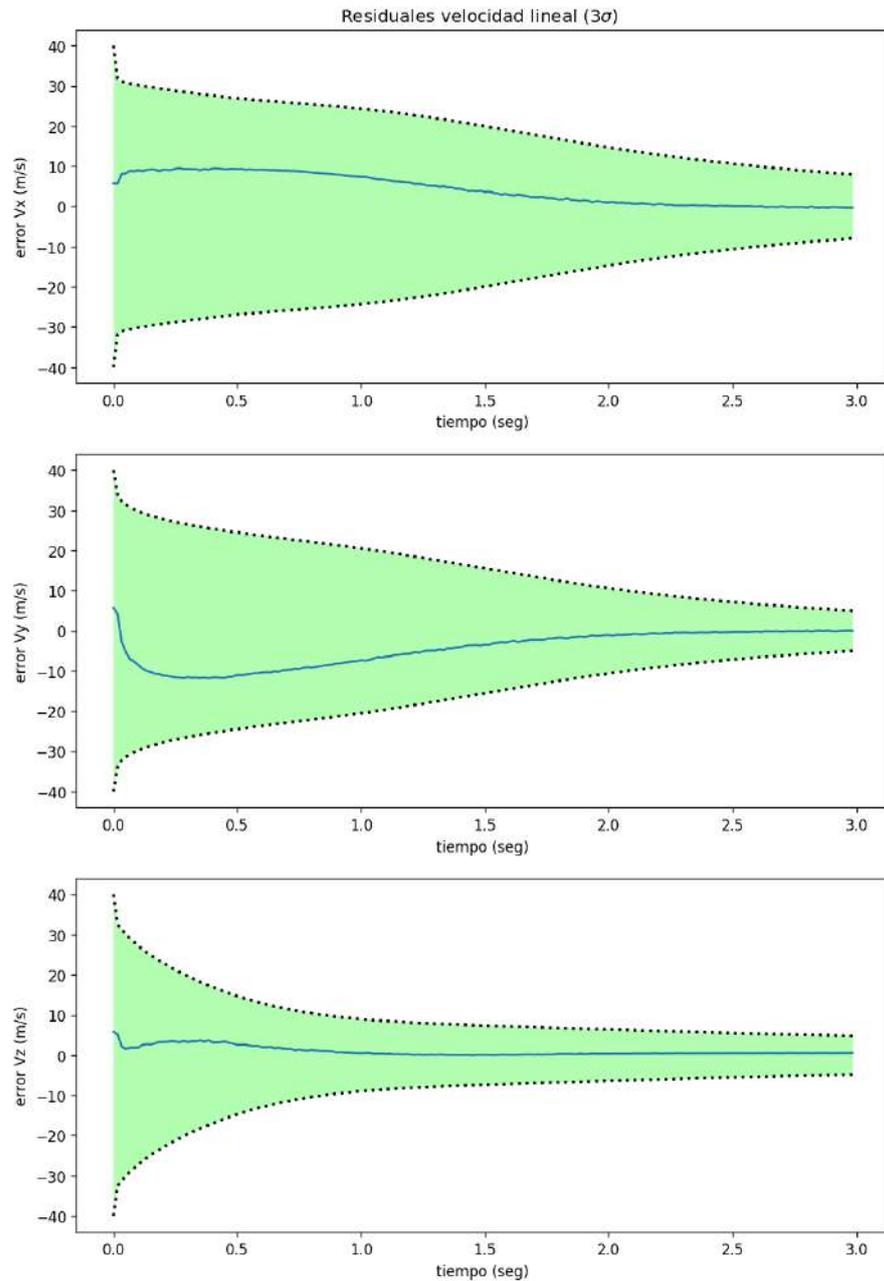


Figura 32: Gráficas de error, arriba el error en la pose en el momento de realizar la predicción, abajo el error respecto al estado real, utilizando el punto medio y la orientación de un segmento de recta.

En el caso de esta representación, la estimación de la velocidad lineal se comportó mucho mejor. Esto se debe a que un punto en el espacio proporciona mejor información al filtro, y por tanto se traduce en una mejor estimación y menor incertidumbre.

## 4.3 EXPERIMENTOS CON CÁMARAS

Después de probar con diferentes representaciones en simulaciones se procedió a implementar la representación que mejor resultados dio en las simulaciones, se calibró la cámara utilizando Matlab Calibration Toolbox [3], y se realizaron los procesos de segmentación y etiquetación explicados en la metodología.



Figura 33: Sistema funcionando en condiciones de paralelismo en la cámara izquierda, arriba el sistema sin filtro de Kalman, abajo el sistema funcionando con la estimación del movimiento.

Para los experimentos realizados con imágenes empleó un sistema de cámaras estéreo formado por dos cámaras web, las cuales están sujetas a una barra metálica que esta fija en un tripie, este sistema se muestra en la figura 34



Figura 34: Sistema de cámaras estéreo utilizado para los experimentos.

Se programo una herramienta que captura la secuencias de imágenes, el cual nos permitió sincronizar la captura para que ambas cámaras capturen simultáneamente, es decir en el mismo instante de tiempo.

Mediante la herramienta programada fueron capturados varios cuadros consecutivos de la escena con los cuales se realizaron los experimentos.

En el experimento a cada cuadro se le extrajeron las rectas tridimensionales que fueron utilizadas como mediciones del filtro de Kalman para estimar los movimientos. Con estos parámetros de movimientos estimamos la siguiente posición de la línea recta.

Los resultados de este experimento no se pudieron comprobar de manera cuantitativa, por no disponer de instrumentos para mover el bastón a unas velocidades conocidas. Pero mediante la re-proyección pudimos observar de forma cualitativa que las estimaciones fueron cercanas.

Por ultimo como se ve en la figura 33, se pudo estimar una pose para los casos donde el trabajo de Enrique Solis presentaba problemas, por lo se consiguió el objetivo de este trabajo el cual es robustecer el seguimiento de los bastones.

## CONCLUSIÓN

---

En este trabajo de tesis se desarrolló un algoritmo robusto para el seguimiento tridimensional de bastones utilizando un filtro de Kalman, esto con el propósito de elaborar una aplicación para visualizar las operaciones vectoriales para fines didácticos.

Tomando esto en cuenta, el sistema recibe una secuencia de imágenes en las que se da por supuesto que contienen 2 bastones de colores diferentes y conocidos, los cuales fueron localizados mediante algoritmos de visión.

Se generó un modelo para cada color de bastón, utilizando la misma metodología que Enrique Solis, pero en este trabajo se utilizó el espacio de color CIELab. Con este modelo y un vector de características se identificaron los bastones en cada par cuadros capturados por el sistema de cámaras estéreo.

Los bastones identificados en la escena fueron la entrada para las mediciones de un filtro de Kalman, el cual estimaba los parámetros de movimiento cinemático de velocidad lineal y angular, esto se repite con cada imagen capturada. Cuando los bastones no pueden ser localizados en la escena, se procede a estimar la pose en base a los parámetros de velocidad estimados por el filtro de Kalman. Finalmente se calcula el producto cruz de las orientaciones de los bastones para generar el resultado del producto cruz y los tres bastones se re-proyectan en la escena.

Para el desarrollo del filtro de Kalman se probó con dos representaciones diferentes, a fin de ver si el modelo de recta utilizado podía mejorar las estimaciones.

De acuerdo a los resultados obtenidos, quedo demostrado que la implementación de un filtro de Kalman para la estimación de velo-

tidades ayuda a mejorar el seguimiento de los bastones. Estos resultados permiten realizar una predicción de la pose de los bastones, con lo cual se consiguió mostrar la pose de los bastones bajo condiciones de colinealidad debido a la proyección en las cámaras, en las cuales no eran posibles la localización 3D, otorgándole más robustez al sistema.

También se observó que la representación para bastones utilizando un punto, el punto de intersección, y un vector de orientación logró mejores resultados que representar cada recta con coordenadas de Plücker.

### 5.1 CONTRIBUCIONES DEL PROYECTO

La contribución principal de este proyecto es proporcionar un algoritmo más robusto para el seguimiento de líneas rectas utilizado en una aplicación para mostrar propiedades de los vectores, utilizando realidad aumentada, para fines didácticos.

### 5.2 TRABAJO FUTURO

Aunque el filtro de Kalman demostró poder manejar casos donde la reconstrucción 3D no es posible, las pruebas de error solo pudieron ser realizadas mediante simulaciones. Con datos reales solo pudo ser observada mediante la re-proyección en las imágenes, sería deseable poder realizar pruebas con imágenes reales moviendo los bastones con unos parámetros de velocidad conocidos.

Además se debe hacer un mejor análisis de los parámetros con los cuales se inicializa el filtro de Kalman, en especial la matriz de ruido en la medición y las matrices de predicción.

También se puede mejorar el modelo cinemático tomando en cuenta las aceleraciones.

Parte I

APÉNDICE





## MATRIZ JACOBIANA DEL MODELO CON REPRESENTACIÓN PLÜCKER

---

Sea  $z = [h_1, u_1, m_1, h_2, u_2, m_2]$  la medición, donde  $h_1$  y  $h_2$  son las normales de las coordenadas de Plücker en el tiempo  $t$  y  $t + 1$ , respectivamente, y sean  $u_1$  y  $u_2$  los vectores de orientación de las rectas y  $m_1$  y  $m_2$  las distancias del origen a la recta.

El estado esta dado por  $x = [v, \omega]$ , donde  $v$  es el vector de velocidad lineal y  $\omega$  el el vector que representa la velocidad angular.

### A.1 MATRIZ JACOBIANA CON RESPECTO A LA MEDICIÓN

Se tiene la función que relaciona la medición con el estado, la cual se puede observar en 38.

$$f(z, x) = \begin{bmatrix} Wu_1 \times u_2 \\ m_2 - m_1 - (v^T(u_1 \times h_1))\Delta t \end{bmatrix} = 0 \quad (38)$$

Donde:

$$W = e^{\omega\Delta t} \\ = I_3 + \sin(\|\omega\|\Delta t)\hat{\omega} + (1 - \cos(\|\omega\|\Delta t))\hat{\omega}^2$$

Se obtendrá primero el Jacobiano con respecto a la medición, para eso se requieren calcular las derivadas parciales con respecto a  $f_1 = Wu_1 \times u_2$  y  $f_2 = m_2 - m_1 - (V^T(u_1 \times h_1))\Delta t$  evaluadas en  $t$ , la matriz Jacobiana queda como se ve en la ecuación

39

$$\frac{\partial f(z, s)}{\partial z} = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & 0 & 0 & \frac{\partial f_1}{\partial u_2} & 0 & 0 \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial h_1} & -1 & 0 & 0 & 1 \end{bmatrix}. \quad (39)$$

De la ecuación 39 se puede observar que solo se requiere calcular de las siguientes derivadas parciales  $\frac{\partial f_1}{\partial \mathbf{u}_1}$ ,  $\frac{\partial f_1}{\partial \mathbf{u}_2}$ ,  $\frac{\partial f_2}{\partial \mathbf{u}_1}$  y  $\frac{\partial f_2}{\partial \mathbf{h}_1}$ .

$$\begin{aligned}\frac{\partial f_1}{\partial \mathbf{u}_1} &= \widehat{W\mathbf{u}_1} \frac{\partial \mathbf{u}_2}{\partial \mathbf{u}_1} - \frac{\partial W\mathbf{u}_1}{\partial \mathbf{u}_1} \hat{\mathbf{u}}_2 \\ &= \widehat{W\mathbf{u}_1} \mathbf{0} - \frac{\partial W\mathbf{u}_1}{\partial \mathbf{u}_1} \hat{\mathbf{u}}_2 \\ &= -W\hat{\mathbf{u}}_2\end{aligned}$$

$$\begin{aligned}\frac{\partial f_1}{\partial \mathbf{u}_2} &= \widehat{W\mathbf{u}_1} \frac{\partial \mathbf{u}_2}{\partial \mathbf{u}_2} - \frac{\partial W\mathbf{u}_1}{\partial \mathbf{u}_2} \hat{\mathbf{u}}_2 \\ &= \widehat{W\mathbf{u}_1} \frac{\partial \mathbf{u}_2}{\partial \mathbf{u}_2} - \mathbf{0} \hat{\mathbf{u}}_2 \\ &= \widehat{W\mathbf{u}_1}\end{aligned}$$

$$\begin{aligned}\frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\mathbf{u}_1} &= \hat{\mathbf{u}}_1 \frac{\partial \mathbf{h}_1}{\partial \mathbf{u}_1} - \hat{\mathbf{h}}_1 \frac{\partial \mathbf{u}_1}{\partial \mathbf{u}_1} \\ &= -\hat{\mathbf{h}}_1 \\ \frac{\partial f_2}{\partial \mathbf{u}_1} &= -\frac{\partial(\mathbf{v}^\top(\mathbf{u}_1 \times \mathbf{h}_1))}{\partial \mathbf{u}_1} \Delta t \\ &= -\left[\frac{\partial \mathbf{v}^\top}{\partial \mathbf{u}_1}(\mathbf{u}_1 \times \mathbf{h}_1) + \mathbf{v}^\top \frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\partial \mathbf{u}_1}\right] \Delta t \\ &= -\left[\mathbf{0}(\mathbf{u}_1 \times \mathbf{h}_1) + \mathbf{v}^\top \frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\partial \mathbf{u}_1}\right] \Delta t \\ &= -\mathbf{v}^\top \frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\partial \mathbf{u}_1} \Delta t \\ &= \mathbf{v}^\top \hat{\mathbf{h}}_1 \Delta t\end{aligned}$$

$$\begin{aligned}
\frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\mathbf{h}_1} &= \hat{\mathbf{u}}_1 \frac{\partial \mathbf{h}_1}{\partial \mathbf{h}_1} - \hat{\mathbf{h}}_1 \frac{\partial \mathbf{u}_1}{\partial \mathbf{h}_1} \\
&= -\hat{\mathbf{u}}_1 \\
\frac{\partial f_2}{\partial \mathbf{h}_1} &= -\frac{\partial(\mathbf{v}^\top (\mathbf{u}_1 \times \mathbf{h}_1))}{\partial \mathbf{h}_1} \Delta t \\
&= -\left[ \frac{\partial \mathbf{v}^\top}{\partial \mathbf{h}_1} (\mathbf{u}_1 \times \mathbf{h}_1) + \mathbf{v}^\top \frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\partial \mathbf{h}_1} \right] \Delta t \\
&= -\left[ 0(\mathbf{u}_1 \times \mathbf{h}_1) + \mathbf{v}^\top \frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\partial \mathbf{h}_1} \right] \Delta t \\
&= -\mathbf{v}^\top \frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)}{\partial \mathbf{h}_1} \Delta t \\
&= -\mathbf{v}^\top \hat{\mathbf{u}}_1 \Delta t
\end{aligned}$$

## A.2 MATRIZ JACOBIANA CON RESPECTO A LA ESTADO

El calculo de la matriz Jacobiana con respecto al estado esta dado por la ecuación 40

$$\frac{\partial f(z, s)}{\partial \mathbf{s}} = \begin{bmatrix} \frac{\partial f_1}{\partial \omega} & 0 \\ 0 & \frac{\partial f_2}{\partial \mathbf{v}} \end{bmatrix}. \quad (40)$$

Para ello como se ve en la ecuación 40, debemos calcular las derivadas parciales de  $\frac{\partial f_1}{\omega}$  y  $\frac{\partial f_2}{\mathbf{v}}$ .

$$\begin{aligned}
\frac{\partial f_1}{\omega} &= (W\mathbf{u}_1)^\times \frac{\partial \mathbf{u}_2}{\partial \omega} - \mathbf{u}_2^\times \frac{\partial W\mathbf{u}_1}{\partial \omega} \\
&= (W\mathbf{u}_1)^\times 0 - \mathbf{u}_2^\times \frac{\partial W\mathbf{u}_1}{\partial \omega} \\
&= -\mathbf{u}_2^\times \frac{\partial W\mathbf{u}_1}{\partial \omega}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial W\mathbf{u}}{\partial \omega} &= -\frac{\sin(\theta \Delta t)}{\theta} \hat{\mathbf{u}} \\
&+ \frac{\theta \Delta t \cos(\theta \Delta t) - \sin(\theta \Delta t)}{\theta^3} (\hat{\omega} \mathbf{u}) \omega^\top \\
&+ \frac{\theta \Delta t \sin(\theta \Delta t) - 2(1 - \cos(\theta \Delta t))}{\theta^4} (\hat{\omega}(\hat{\omega} \mathbf{u})) \omega^\top \\
&+ \frac{1 - \cos(\theta \Delta t)}{\theta^2} [-\widehat{\omega \dot{\mathbf{u}}} + (\omega \dot{\mathbf{u}}) \mathbf{I}_3 - \mathbf{u} \omega^\top]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial f_2}{\mathbf{v}} &= \frac{\partial m_2}{\partial \mathbf{v}} - \frac{\partial m_1}{\partial \mathbf{v}} - \frac{\partial(\mathbf{v}^\top(\mathbf{u}_1 \times \mathbf{h}_1))\Delta t}{\partial \mathbf{v}} \\
&= -\frac{\partial(\mathbf{v}^\top(\mathbf{u}_1 \times \mathbf{h}_1))\Delta t}{\partial \mathbf{v}} \\
&= -\frac{\partial \mathbf{v}^\top}{\partial \mathbf{v}}(\mathbf{u}_1 \times \mathbf{h}_1)\Delta t - \frac{\partial(\mathbf{u}_1 \times \mathbf{h}_1)\Delta t}{\partial \mathbf{v}}\mathbf{v} \\
&= -(\mathbf{u}_1 \times \mathbf{h}_1)\Delta t
\end{aligned}$$

## BIBLIOGRAFÍA

---

- [1] Nicolas Andre and Bernard Espiau. Revisiting Plücker Coordinates in Vision-Based Control. 2002.
- [2] Gaurav Bhorkar. A survey of augmented reality navigation. *CoRR*, abs/1708.05006, 2017. URL <http://arxiv.org/abs/1708.05006>.
- [3] Jean-Yves Bouguet. Camera calibration toolbox for Matlab. [http://www.vision.caltech.edu/bouguetj/calib\\_doc](http://www.vision.caltech.edu/bouguetj/calib_doc).
- [4] Angel Chi-Poot, Anabel Martín-González, Víctor Hugo Menéndez-Domínguez, and Arturo Espinosa-Romero. Aprendizaje de vectores euclidianos utilizando un sistema de realidad aumentada. *Research in Computing Science*, 89:9–16, 2015.
- [5] Ramsey Faragher. Understanding the Basis of the Kalman Filter Via a Simple and Intuitive Derivation [Lecture Notes]. *IEEE Signal Processing Magazine*, 29(5):128–132, September 2012. ISSN 1053-5888. doi: 10.1109/msp.2012.2203621. URL <http://dx.doi.org/10.1109/msp.2012.2203621>.
- [6] Olivier D. Faugeras, Nassir Navab, and Rachid Deriche. Information contained in the motion field of lines and the cooperation between motion and stereo. *International Journal of Imaging Systems and Technology*, 2(4):356–370, 1990. ISSN 10981098. doi: 10.1002/ima.1850020410.
- [7] J.E. Gentle. *Numerical Linear Algebra for Applications in Statistics*. Statistics and Computing. Springer New York, 1998. ISBN 9780387985428. URL [https://books.google.com.mx/books?id=2J0ndF\\_LmqoC](https://books.google.com.mx/books?id=2J0ndF_LmqoC).
- [8] DA Kerr. The CIE XYZ and xyY Color Spaces. [http://graphics.stanford.edu/courses/cs148-10-summer/docs/2010--kerr--cie\\_xyz.pdf](http://graphics.stanford.edu/courses/cs148-10-summer/docs/2010--kerr--cie_xyz.pdf), 2010. URL [http://graphics.stanford.edu/courses/cs148-10-summer/docs/2010--kerr--cie\\_xyz.pdf](http://graphics.stanford.edu/courses/cs148-10-summer/docs/2010--kerr--cie_xyz.pdf).

[//graphics.stanford.edu/courses/cs148-10-summer/docs/2010--kerr--cie\\_xyz.pdf](http://graphics.stanford.edu/courses/cs148-10-summer/docs/2010--kerr--cie_xyz.pdf).

- [9] Yi Ma, Stefano Soatto, Jana Kosecka, and S. Shankar Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*. SpringerVerlag, 2003. ISBN 0387008934.
- [10] Prasanta Chandra Mahalanobis. On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2(1):49—55, 1936. URL <http://ir.isical.ac.in/dspace/handle/1/1268>.
- [11] María Teresa Escobedo Portillo and Jorge A. Salas Plata Mendonza. P. ch. mahalanobis y las aplicaciones de su distancia estadística. *Culcyt*, (27):13–20, 2008.
- [12] Janos Schanda. *Colorimetry: Understanding the CIE system*. John Wiley & Sons, 2007. ISBN 9780470049044.
- [13] Enrique Javier Solís-Pomar. Representación virtual e interactiva del producto cruz de dos vectores a partir del seguimiento tridimensional de referentes físicos. Master's thesis, Universidad Autónoma de Yucatán, 2014.
- [14] K. K. Sung and T. Poggio. Example-based learning for view-based human face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):39–51, Jan 1998. ISSN 0162-8828. doi: 10.1109/34.655648.
- [15] Zhengyou Zhang and Olivier D. Faugeras. Three-dimensional motion computation and object segmentation in a long sequence of stereo frames. *International Journal of Computer Vision*, 7(3):211–241, 1992. ISSN 09205691. doi: 10.1007/BF00126394.