

SLAM Monocular Basado en UKF para la Localización de un Robot Móvil

Gonzalo Perez Paina Claudio Paz Martin Baudino Luis Canali
 Centro de Investigación en Informática para la Ingeniería, CIII
 Universidad Tecnológica Nacional, Facultad Regional Córdoba, UTN-FRC
 gperez@scdt.frc.utn.edu.ar

Resumen—En la actualidad el algoritmo de SLAM es una parte fundamental de cualquier esquema de navegación autónoma de robots móviles, siendo la localización una de sus aplicaciones más utilizadas. En este trabajo se presenta un esquema de SLAM visual monocular para la localización de un robot móvil con ruedas, aprovechando la información de odometría para ajustar la escala a unidades métricas. Mediante la simulación del esquema propuesto se muestra que el filtro UKF resulta más consistente en la estimación comparado con el tradicionalmente utilizado EKF. Además, se presentan resultados experimentales utilizando el filtro UKF en un robot móvil con ruedas.

I. INTRODUCCIÓN

Uno de los bloques fundamentales de cualquier esquema actual de navegación autónoma o semi-autónoma de robots móviles es el algoritmo de SLAM (Simultaneous Localization and Mapping). El algoritmo de SLAM permite a un robot que se sitúa en una localización desconocida, en un entorno también desconocido, construir incrementalmente un mapa de su entorno de operación y determinar su localización en el mismo [1]. Los algoritmos de SLAM pueden utilizarse para diferentes propósitos, los cuales pueden ser: para la generación de mapas en procesamiento offline, para localización en procesamiento online, o bien como algoritmo de operación permanente para actualización continua del mapa [2].

El enfoque clásico del algoritmo de SLAM es mediante el filtrado bayesiano, particularmente utilizando el filtro de Kalman y sus diferentes versiones (EKF, UKF, EIF, etc.) [3][4]. En este contexto el SLAM puede verse como un problema de estimación de estado, donde el estado a estimar incluye la posición y orientación (o pose) del robot, junto a las posiciones de los landmarks que componen el mapa. En los esquemas de SLAM con filtro de Kalman en la función de movimiento generalmente se incorpora información de sensores propioceptivos, mientras que en la función de observación se utiliza información de los sensores exteroceptivos del robot.

En los últimos años ha habido un creciente interés en la utilización de cámaras digitales como principal sensor exteroceptivo para percibir el entorno. Sin embargo, cuando se utiliza una única cámara (visión monocular) surge el problema de la inicialización de landmarks, debido al hecho de la imposibilidad de determinar la profundidad de un punto 3D a partir de una única vista. Además, los esquemas de SLAM puramente visuales tienen el inconveniente de que la estimación se realiza hasta un factor de escala desconocido. En este sentido, la aplicación del SLAM en un robot móvil con ruedas puede aprovechar la información de odometría como sensor propioceptivo, lo que permite ajustar la escala a unidades métricas.

Uno de los primeros trabajos de SLAM monocular se presenta en [5], el cual está basado en un filtro del Kalman extendido (EKF, Extended Kalman Filter) para una cámara que se mueve libremente con movimientos suaves. En dicho

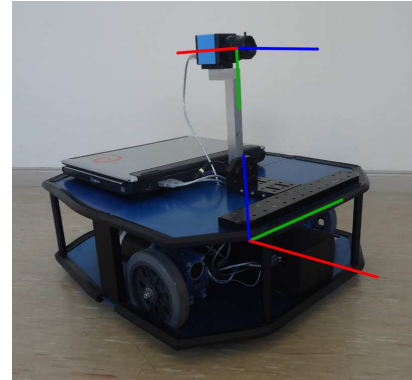


Figura 1: Robot móvil RoMAA-II y sistemas de coordenadas del robot y cámara

trabajo se presenta un esquema de inicialización de landmarks que opera en forma independiente del filtro de estimación del SLAM. Este proceso de inicialización permite estimar la profundidad de los nuevos landmarks (puntos 3D) los cuales son luego agregados al mapa del SLAM. La estimación de la profundidad se realiza utilizando un filtro de partículas sobre el rayo óptico de la nueva observación. Otro enfoque más recientemente permiten inicializar los landmarks de forma inmediata cuando son observados por primera vez [6]. Dicho esquema se basan en la parametrización con profundidad inversa para representar los puntos 3D que conforman el mapa del SLAM [7].

El actual trabajo es una continuación de [8] donde se presenta un esquema de SLAM monocular basado en filtro de Kalman Extendido; el cual aquí se reemplaza por un filtro de Kalman Unscented (UKF, Unscented Kalman Filter). Se muestra mediante simulaciones un mejor comportamiento del UKF comparado con el EKF. Además, se presentan resultados experimentales utilizando el esquema de SLAM propuesto en un robot móvil con ruedas, utilizando el robot móvil RoMAA-II [9], que se muestra en la Fig. 1.

En la literatura existen diferentes trabajos que utilizan el filtro UKF para la estimación del SLAM. Entre ellos en [10] se usa la transformada Unscented (UT, Unscented Transform), sobre la que se basa el filtro UKF, para la estimación a priori de la pose del robot (vector de estado y matriz de covarianza) únicamente, mientras que para obtener la covarianza cruzada entre el mapa (considerado estático) y el robot se utiliza la linealización como en un EKF. En [11] se presenta uno de los primeros trabajos de SLAM monocular con parametrización de profundidad inversa con UKF, aplicado a una cámara que se mueve con 6 grados de libertad, presentando sólo resultados de simulación. Mientras que en [12] se muestra la implementación de la versión Square Root del UKF (SRUKF), comparando los resultados entre EKF y UKF estándar, para SLAM monocular con modelo de movimiento de velocidad constante como el

utilizado en [5]. El principal aporte del presente trabajo es utilizar los modelos adecuados (tanto de movimiento como de observación) para el caso particular de un robot que se mueve sobre una superficie plana, incorporando también la información de la odometría del robot para que la estimación resulte en unidades métricas.

La Sección II presenta la formulación del SLAM basada en filtro de Kalman y el caso particular del filtro UKF. En la Sección III se describe el esquema de SLAM monocular propuesto junto a la parametrización de landmarks utilizada y los modelos aplicados al filtro. La Sección IV muestra resultados de simulación comparando el filtro EKF y UKF, y resultados experimentales utilizando un filtro UKF. Finalmente, la Sección V presenta las conclusiones y trabajo a futuro.

II. SLAM CON FILTRO DE KALMAN

En el contexto del filtrado de Kalman, una de las operaciones principales y de mayor importancia en cualquier implementación es la propagación de variables aleatorias gaussianas a través de funciones que describen el sistema, las cuales generalmente son no lineales.

En el caso del filtro de Kalman Extendido (EKF) la distribución de probabilidad del estado del sistema, al igual que las densidades de los ruidos involucrados, se representan por variables aleatorias gaussianas que se propagan de forma analítica mediante una aproximación lineal de primer orden de funciones no lineales. Sin embargo, dicha aproximación puede producir errores importantes tanto en la media como en la covarianza de la densidad de la variable aleatoria gaussiana transformada. Esto lleva al filtro a un comportamiento subóptimo o bien, bajo ciertas circunstancias, puede llegar a producir una divergencia en el resultado de la estimación.

Por otro lado, los filtros de Kalman de puntos sigma (SPKF) [13] abordan el problema del EKF antes mencionado mediante el uso de un mecanismo de muestreo determinista. De forma similar al EKF, en los SPKF el estado se aproxima mediante una variable aleatoria gaussiana, y se representa por un conjunto mínimo de muestras con sus pesos asociados, los cuales son seleccionados cuidadosamente, y reciben el nombre de *puntos sigma*. Estas muestras o puntos sigma representan fielmente la media y covarianza de la variable aleatoria gaussiana, y cuando se propagan a través de la función no lineal, representan la media y covarianza a posteriori hasta en un segundo orden (comparada con una aproximación en expansión en serie de Taylor). Además, la implementación de los SPKF resulta sustancialmente más simple comparada con el EKF, dado que no necesita el cálculo analítico de los jacobianos de este último.

El filtro de Kalman Unscented es un caso particular de filtro de Kalman de puntos sigma, donde la selección de los puntos sigma se realiza mediante la transformada Unscented (UT, Unscented Transform) [14].

II-A. Filtro de Kalman Unscented

Para la aplicación del filtro de Kalman se necesita describir el sistema mediante ecuaciones de estado estocásticas, las cuales vienen dadas por

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

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k \quad (2)$$

donde \mathbf{x}_k es el estado del sistema (no observable), \mathbf{u}_k es la acción de control, \mathbf{z}_k es la observación, $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ es el ruido de proceso, y $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ es el ruido de medición, todos en el paso temporal k .

El vector de estado $\mathbf{x} = [\mathbf{r}^T \ \mathbf{m}^T]^T$ esta compuesto por el estado del robot $\mathbf{r} = [x \ y \ \phi]^T$, y el estado del mapa $\mathbf{m} = [\mathbf{m}_1^T \ \dots \ \mathbf{m}_N^T]^T$, donde $\mathbf{m}_i, i = 1, \dots, N$ son los landmarks que conforman el mapa, el cual se asume estático $\mathbf{m}_k = \mathbf{m}_{k-1} = \mathbf{m}$.

El objetivo del filtro de Kalman es estimar el estado \mathbf{x}_k a partir de las observaciones \mathbf{z}_k hasta el paso temporal k . El filtro de Kalman opera en dos etapas: predicción y corrección. La etapa de predicción consiste en aplicar directamente la transformada Unscented para obtener el vector estado y la matriz de covarianza a priori utilizando la ecuación de proceso (1). En la etapa de corrección se aplica la transformada Unscented para determinar la predicción de la medición dada por la ecuación de observación (2), luego se corrige el estado y la matriz de covarianza mediante

$$\mathbf{K}_k = \mathbf{P}_{rz,k} (\mathbf{R}_k + \mathbf{R}_k^-)^{-1} \quad (3)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - h(\hat{\mathbf{x}}_k^-)) \quad (4)$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k (\mathbf{R}_k + \mathbf{R}_k^-) \mathbf{K}_k^T. \quad (5)$$

II-B. La transformada Unscented

La transformada Unscented es un método numérico que puede ser utilizado para aproximar una distribución de probabilidad conjunta de las variables aleatorias $\mathbf{x} \sim \mathcal{N}(\bar{\mathbf{x}}, \mathbf{P}_x)$ e $\mathbf{y} = g(\mathbf{x})$.

Para el caso de filtrado esta transformación se aplica en dos situaciones: la primera en la etapa de predicción donde \mathbf{x} es $\hat{\mathbf{x}}_{k-1}$, y es $\hat{\mathbf{x}}_k^-$ y $g(\cdot)$ es $f(\cdot)$, y la segunda en la predicción de la medición donde \mathbf{x} es $\hat{\mathbf{x}}_k^-$, y es \mathbf{z}_k^- , y $g(\cdot)$ es $h(\cdot)$. La transformada Unscented se aplica mediante los siguientes pasos:

1. Formar un conjunto con $2L + 1$ puntos sigma

$$\mathcal{X}_0 = \mathbf{x}$$

$$\mathcal{X}_i = \mathbf{x} + \left(\sqrt{(L + \lambda) \mathbf{P}_x} \right)_i \quad i = 1, \dots, L \quad (6)$$

$$\mathcal{X}_i = \mathbf{x} - \left(\sqrt{(L + \lambda) \mathbf{P}_x} \right)_{i-n} \quad i = L + 1, \dots, 2L$$

donde L es la dimensión del vector \mathbf{x} , $\lambda = \sigma^2(L + \kappa) - L$ es un parámetro de escalado. La constante α determina la dispersión de los puntos sigma alrededor de $\bar{\mathbf{x}}$ y generalmente se fija a un valor positivo pequeño (*p.e.*, $1 \leq \alpha \leq 1 \times 10^{-4}$). La constante κ es un segundo parámetro de escalado, y generalmente se fija a 0 o $3 - n$.

2. Propagar los puntos sigma a través de la función no lineal $g(\cdot)$ como

$$\mathcal{Y}_i = g(\mathcal{X}_i), \quad i = 0, \dots, 2L.$$

3. La estimación de la media y covarianza de la variable transformada se puede calcular a partir de los puntos sigma como

$$\mathbb{E}[g(\mathbf{x})] \approx \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Y}^{(i)} = \bar{\mathbf{y}} \quad (7)$$

$$\text{Cov}[g(\mathbf{x})] \approx \sum_{i=0}^{2L} W_i^{(c)} (\mathcal{Y}^{(i)} - \bar{\mathbf{y}})(\mathcal{Y}^{(i)} - \bar{\mathbf{y}})^T, \quad (8)$$

donde las constantes de peso $W_i^{(m)}$ y $W_i^{(c)}$ están dadas por

$$W_0^{(m)} = \lambda / (L + \lambda) \quad (9)$$

$$W_0^{(c)} = \lambda / (L + \lambda) + (1 - \alpha^2 + \beta) \quad (10)$$

$$W_i^{(m)} = W_i^{(c)} = 1 / (2(L + \lambda)), \quad i = 1, \dots, 2L \quad (11)$$

β se utiliza para incorporar información a priori de la distribución de \mathbf{x} (para una distribución gaussiana, opcionalmente puede ser $\beta = 2$). $\left(\sqrt{(L + \lambda)\mathbf{P}}\right)_i$ es la i -ésima columna de la raíz cuadrada de la matriz (la cual se obtiene como *p.e.* mediante la matriz triangular inferior de la factorización de Cholesky).

III. SLAM MONOCULAR EN ROBOT MÓVIL

El esquema de SLAM visual monocular propuesto se basa en el filtro de Kalman Unscented donde el mapa se corresponde a puntos 3D (nube de puntos no densa), los cuales se representan ya sea mediante la parametrización euclídea o en profundidad inversa para la inicialización. El modelo de movimiento utilizado se basa en la información odométrica del robot y se describe en [1] y en [4]. Además, los parámetros intrínsecos de la cámara al igual que la transformación rígida entre el robot y la cámara (o parámetros extrínsecos) se consideran conocidos. Dichos parámetros se obtienen previamente a la ejecución del SLAM mediante la calibración de los mismos utilizando por ejemplo [15] y [16] para determinar los parámetros intrínsecos y extrínsecos, respectivamente.

III-A. Esquema general

El esquema general se muestra en la Fig. 2, el cual está compuesto de las siguientes partes: la plataforma móvil y sus sensores (odometría y cámara), el filtro de estimación, los algoritmos de visión por computadoras para el procesamiento de las imágenes y la asociación de datos junto a la eliminación de outliers con RANSAC 1 punto [6]. La administración del mapa consta de una heurística simple de agregado y eliminación de landmarks del mapa. Los nuevos landmarks se agregan al mapa basados en un umbral de la mínima cantidad de características en la imagen, y la eliminación considera la relación entre la cantidad de landmarks del mapa proyectados al plano imagen (predicción) y la cantidad utilizados realmente en la etapa de corrección del filtro (asociación correcta).

III-B. Parametrización de landmarks

Un punto 3D, o landmark del mapa del SLAM, se representan mediante sus coordenadas cartesianas (x, y, z) , en la parametrización euclídea. Sin embargo, los puntos en coordenadas euclídeas presentan no linealidades importantes en las funciones de observación de sistemas bearing-only como el SLAM monocular, resultando inadecuado en el proceso de inicialización. Además de que no es posible invertir la función de observación, dado que no se conoce la profundidad de un punto 3D a partir de una única vista en un sistema monocular.

Una parametrización que permite representar puntos hasta distancia infinita, y que resultan adecuados para la inicialización de landmarks es la parametrización en profundidad inversa, la cual se presenta en [7]. Un punto 3D en dicha

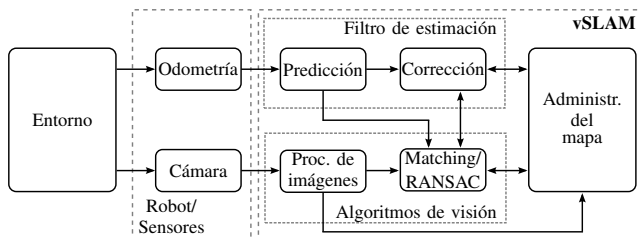


Figura 2: Esquema general del algoritmo de SLAM visual monocular propuesto

parametrización queda representado por $(x^0, y^0, z^0, \gamma, \psi, \rho)$; donde (x^0, y^0, z^0) (punto de anclaje) son las coordenadas de la posición de la cámara en el instante cuando el punto es observado por primera vez, (γ, ψ) son los ángulos de azimut y elevación del rayo óptico que une el punto de anclaje con el punto 3D, y $\rho = 1/d$ es la inversa de la distancia d al punto.

La parametrización en profundidad inversa tiene el inconveniente de requerir mas parámetros que la parametrización mínima euclídea. Por ello en SLAM se deben convertir aquellos puntos cuya profundidad ha sido ya estimada con cierta precisión, lo que reduce el tamaño del vector de estado del mapa. Un criterio de conversión de la parametrización en profundidad inversa a coordenadas euclídeas se presenta en [17].

III-C. Modelo de observación

En el SLAM visual presentado cada landmark del mapa $\mathbf{m}_i, i = 1, \dots, N$ se inicializa utilizando la parametrización en profundidad inversa dada por

$$\mathbf{m}_{\text{ID}}^i = [x_i^0 \ y_i^0 \ \gamma_i \ \psi_i \ \rho_i]^T \in \mathbb{R}^5, \quad (12)$$

donde se omite incluir la coordenadas z del punto de anclaje, dado que se conoce la transformación rígida robot/cámara. Luego, dichos puntos se convierten a la representación euclídea dada por

$$\mathbf{m}_{\text{EU}}^i = [x_i \ y_i \ z_i]^T \in \mathbb{R}^3. \quad (13)$$

La función de observación, que proyecta los landmarks del mapa del SLAM, al plano imagen de la cámara viene dada por

$$\mathbf{z} = \mathbf{K}\mathbf{R}_g^c(\mathbf{m}_{\text{EU}} - \mathbf{t}_c^g) \in \mathbb{P}^2, \quad (14)$$

para la parametrización en profundidad inversa de (12), y

$$\mathbf{z} = \mathbf{K}\mathbf{R}_g^c(\rho(\mathbf{t}_0 - \mathbf{t}_c^g) + \mathbf{d}(\gamma, \psi)) \in \mathbb{P}^2, \quad (15)$$

para la parametrización euclídea de (13).

Las entidades de (14) y (15) son:

- $\mathbf{z} = [u \ v \ w]^T$ es el punto imagen (observación) en coordenadas homogéneas.
- \mathbf{K} es la matriz de parámetros intrínsecos de la cámara dada por

$$\mathbf{K} = \begin{bmatrix} f/h_u & 0 & u_0 \\ 0 & f/h_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (16)$$

con f la distancia focal, y h_u y h_v el ancho y alto de los píxeles.

- $(\mathbf{t}_c^g, \mathbf{R}_g^c)$ es la transformación rígida de la posición y orientación de la cámara en el sistema de referencia global.
- $\mathbf{d}(\gamma, \psi) = [\cos \gamma \sin \psi \ -\sin \psi \ \cos \gamma \sin \psi]^T$ es el vector unitario dado por los ángulos de azimut (γ) y elevación (ψ).

El punto imagen \mathbf{z} convertido a coordenadas euclídeas como $\mathbf{z}_u = [u/w, v/w]^T$ es el punto imagen sin distorsión al que finalmente se le aplica el modelo de distorsión de la lente obteniendo la observación final utilizada en (2).

III-D. Asociación de datos

La asociación de datos necesaria para la etapa de corrección del filtro se realiza de igual forma que como se presenta en [8]. El proceso asociación de datos busca la correspondencia entre la predicción de las mediciones y el descriptor de la característica, almacenado en el proceso de extracción; lo

cual recibe el nombre de matching activo. El extractor de características utilizado es Harris y el descriptor es el patch en la vecindad del punto de interés detectado. Como medida de similitud entre descriptores (patches) se utiliza la correlación cruzada normalizada (NCC, Normalized Cross-Correlation).

El matching activo opera como sigue: dada la predicción de la medición ($\{z_i^-, R_i^-\}$) la cual determina una elipse de incertidumbre en el plano imagen, se determina una zona de interés donde buscar el matching en base a la caja o bounding box que resulta tangente a dicha elipse y orientada según los ejes del plano imagen. Luego, se le realiza una operación de warping al patch inicial del landmark en cuestión a partir de la pose estimada actual de la cámara, y se busca este nuevo patch, en la zona de interés mediante NCC.

IV. RESULTADOS

Los siguientes resultados muestran la comparación de la estimación del SLAM utilizando el filtro EKF y UKF en simulación. Los resultados de la comparación indican un mejor comportamiento del UKF frente al EKF en cuanto a la consistencia de la estimación.

Los resultados experimentales muestran el comportamiento del SLAM 3D estimado mediante el filtro UKF. El robot móvil utilizado es el RoMAA-II [9], que se muestra en la Fig. 1, el cual está equipado con una cámara ImagingSource modelo DFK21BU04 con una lente Computar modelo T3Z3510. El tamaño de las imágenes utilizadas es de 640x480 píxeles.

IV-A. Resultados de simulación (EKF/UKF)

Las simulaciones realizadas para comparar el comportamiento del filtro EKF y UKF en la estimación del SLAM se llevaron a cabo con diferentes longitudes del camino seguido por el robot, y diferentes cantidades y ubicaciones de los landmarks del mapa verdadero; aunque, por cuestiones de espacio se muestra un sólo caso.

En la simulación se generan los datos de odometría sin ruido (valor verdadero o ground truth) al cual se le agrega ruido. Se generan también una serie de puntos 3D (mapa verdadero), a los cuales se le agrega ruido de medición en el proceso de observación al proyectar dichos puntos al plano imagen.

La Fig. 3 muestra los datos de odometría, generados mientras el robot se mueve hacia adelante y produciendo una desviación en aprox. 0.5m de avance (en una curva tipo “S”) volviendo luego a la orientación inicial. Esto produce un cambio de vista de los landmarks 3D que mejora la estimación de profundidad. La Fig.3a muestra la pose verdadera del robot para cada paso temporal en línea continua (color rojo) y la pose dada por la odometría en línea de trazo (color azul), mientras que la Fig. 3b muestra la diferencia de estos valores o error de odometría. La Fig. 3c muestra el camino verdadero del robot y la aplicación del modelo de odometría.

La Fig. 4 muestra el sistema de coordenada del robot y de la cámara y los landmarks estimados por el SLAM para el inicio y fin de la simulación; las gráficas de la izquierda se corresponden al filtro EKF mientras que la derecha al UKF.

La Fig. 5 muestra el error de estimación de la profundidad de los landmarks del mapa del SLAM tanto para el filtro EKF en línea de trazo (color azul) como el UKF en línea continua (color rojo). Como se observa, el error de estimación de profundidad converge más rápidamente a cero y logra un menor error al final de la simulación para el caso del filtro UKF.

Finalmente, la Fig. 6 muestra la estimación de la pose del robot para cada paso temporal dada por el filtro EKF a la

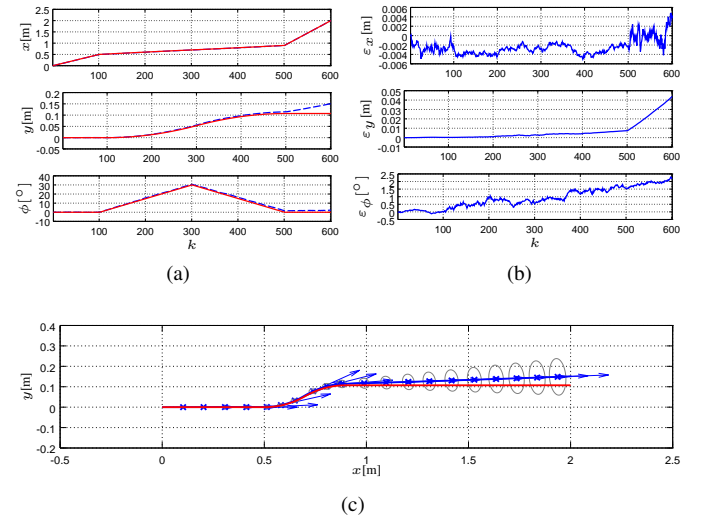


Figura 3: Datos de simulación de odometría. (a) Odometría con y sin ruido. (b) Error de odometría. (c) Camino con y sin ruido.

izquierda (color azul) y el UKF a la derecha (color rojo). Como se observa el filtro UKF presenta un comportamiento más consistente comparado con el filtro EKF, donde el error se encuentra siempre dentro de la cota de incertidumbre.

Todas las simulaciones realizadas muestran resultados similares en cuanto a la comparación del filtro EKF y UKF. La estimación de pose del robot resulta siempre más consistente en el UKF; y el error de profundidad de los landmarks converge más rápido a cero y logra un menor error al final de la simulación para el UKF.

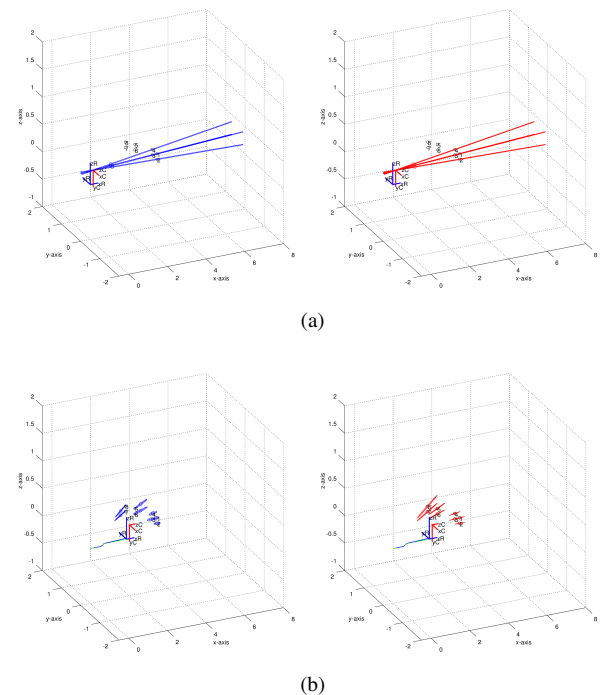


Figura 4: Estimación del SLAM, pose de robot/cámara y landmarks 3D. (a) En el primer paso temporal. (b) Al final de la simulación.

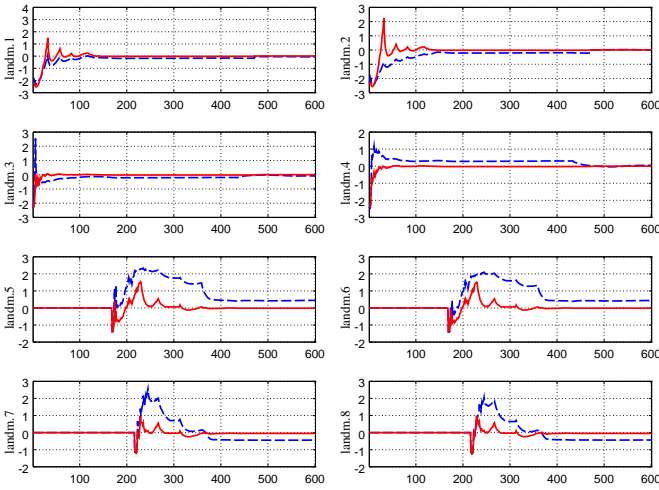


Figura 5: Error en la estimación de profundidad para cada landmark, para el filtro EKF (línea de trazo) y para el UKF (línea continua).

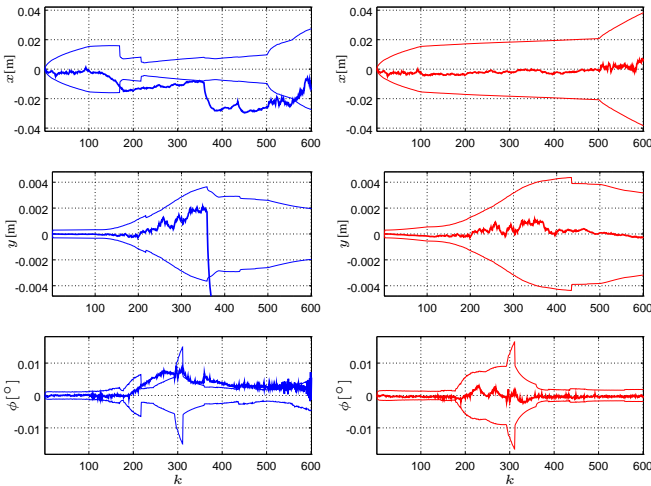


Figura 6: Estimación de la pose del robot y cota de incertidumbre de $\pm 3\sigma$, para el filtro EKF (izq.) y el UKF (der.)

IV-B. Resultados experimentales

En la Fig. 7 se muestra la trayectoria seguida por el robot, el cual se mueve hacia adelante en línea recta y luego retrocede, en un camino de longitud total de 2.50m. Se observa la posición (x, y) y orientación ϕ ; además de la velocidad lineal y angular de robot.

La Fig. 8 muestra el resultado en la estimación de la pose del robot para el camino completo. En la Fig. 8a se observa la posición (x, y) y orientación ϕ dados por la odometría en línea de trazo (color azul), y la pose estimada según la

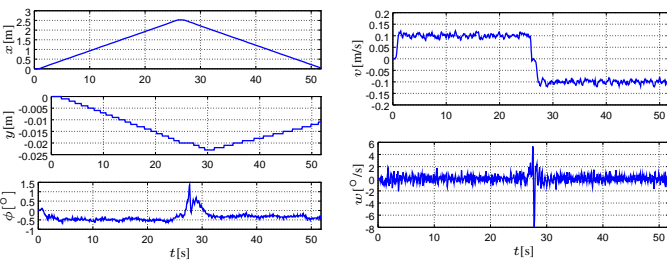
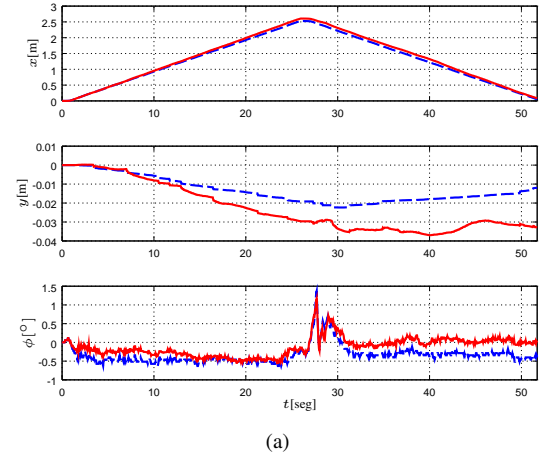
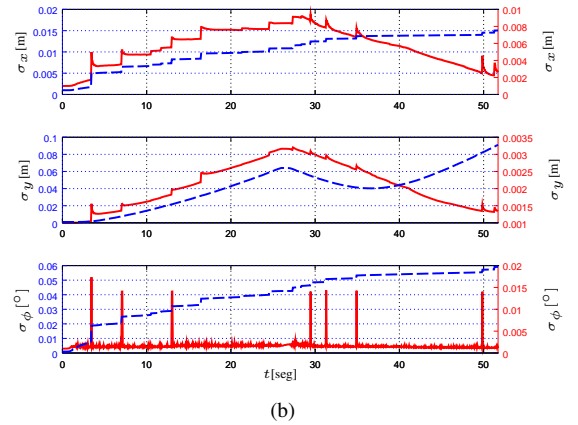


Figura 7: Datos de odometría obtenidas del robot (a) Pose, (b) velocidad lineal y angular.



(a)



(b)

Figura 8: Estimación de la localización del robot. (a) Pose estimada, e (b) incertidumbre de la pose.

localización del SLAM en línea continua (color rojo). En la Fig. 8b se muestra la incertidumbre en la pose del robot según el modelo de movimiento utilizado (sin la etapa de corrección del filtro) en línea de trazo (color azul) y el eje de ordenadas izquierdo, y la incertidumbre en la pose según la estimación del filtro en línea continua (color rojo) y eje de ordenadas derecho. Como se observa, la incertidumbre de la pose del robot sin la aplicación de la etapa de corrección del filtro del SLAM crece sin límites; por otro lado, al aplicar la corrección dicha incertidumbre se mantiene acotada. Se puede apreciar también que la incertidumbre en la orientación del SLAM monocular se mantiene a valores pequeños la mayor parte del tiempo, indicando que un sistema de visión monocular resulta adecuado para medir orientación.

La Fig. 9 muestra el resultado final de la estimación del SLAM. En la Fig. 9a se muestra la incertidumbre de cada landmark del mapa en función del tiempo. Como se observa, dichas incertidumbres disminuyen y convergen a un valor menor al inicial. En la Fig. 9b se muestra la matriz de correlación final del SLAM.

Por último, la Fig. 10 muestra el mapa final estimado por el SLAM y los puntos en el plano imagen que resultan inlier en el proceso de matching por RANSAC. Estos se indican con las elipses color verde que representan la incertidumbre de la predicción de las mediciones utilizadas para el proceso de matching activo. Como se observa, la incertidumbre de los landmarks disminuye en general, excepto aquellos que son observados en pocas imágenes al inicio de la secuencia, como los que se encuentra más a la izquierda en el plano imagen. Dichas observaciones tienen poco ángulo de paralaje a medida

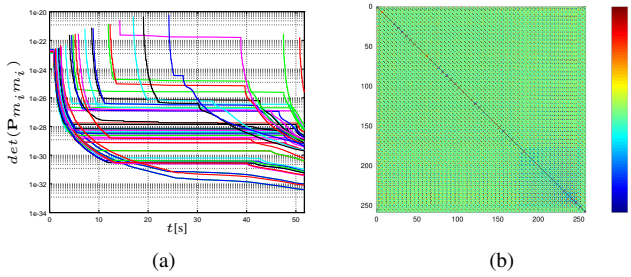


Figura 9: Resultado de la estimación del SLAM. (a) Incertidumbre de landmarks, y (b) matriz de correlación final.

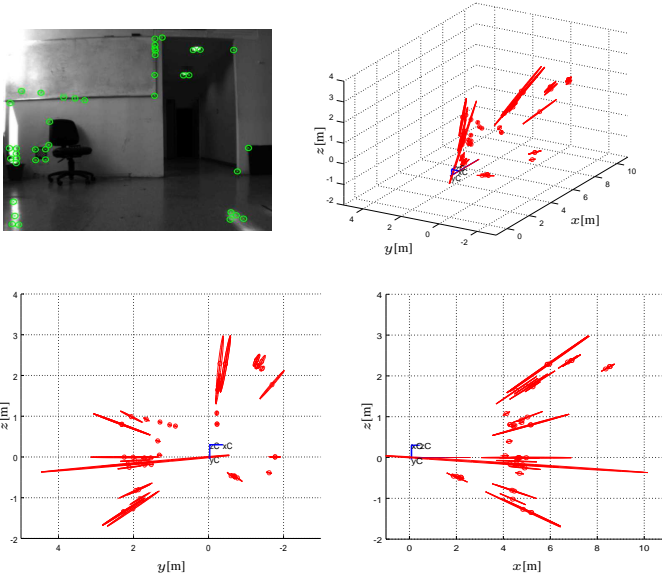


Figura 10: Imagen y mapa final para un camino en línea recta en avance y retroceso de 2.5m de longitud.

que el robot se mueve, lo que imposibilita poder estimar la profundidad de los landmarks correspondientes.

V. CONCLUSIONES Y TRABAJOS FUTUROS

El trabajo presenta un esquema de SLAM visual monocular basado en un filtro de Kalman Unscented para la estimación, aplicado a un robot móvil con ruedas, el cual utiliza la información odométrica del robot para poder ajustar la escala del SLAM a unidades métricas. En dicho esquema se utiliza la inicialización de landmarks mediante la parametrización con profundidad inversa, la cual fue simplificada al operar sobre un robot móvil que se mueve sobre una superficie plana. Además, se asumen conocidos todos los parámetros del sistema robot/cámara, los cuales tienen que ser calibrados previamente a la operación del SLAM.

Los resultados presentados se obtuvieron tanto de simulaciones como de experimentos. La simulación permitió determinar un mejor comportamiento del filtro UKF frente al filtro EKF, utilizado comúnmente en SLAM. Esta mejora se ve reflejada tanto en la consistencia de la estimación de la localización del robot, como en el error de estimación de la profundidad de los landmarks del mapa; los cuales convergen más rápidamente a cero y resultan con menor error final para el filtro UKF. Los resultados experimentales muestran que la estimación tanto de la localización como del mapa final obtenido del SLAM visual. La estimación del filtro de Kalman Unscented resulta suficientemente precisa para lograr un matching correcto de características en el plano imagen, al volver al punto inicial del camino del robot.

Los trabajos a futuro son: realizar una evaluación más exhaustiva de la estimación del algoritmo de SLAM presentado, en caminos de mayor longitud; y determinar de forma absoluta la precisión de la estimación de localización utilizando un sistema de referencia externo como información de ground truth. Otro aspecto importante es la calibración del sistema robot/cámara, la cual podría ser incluida como variables a estimar dentro del filtro, principalmente los parámetros extrínsecos que relacionan el sistema de coordenadas de la cámara con el del robot.

VI. AGRADECIMIENTOS

El presente trabajo se enmarca en el proyecto “Localización de Robots Móviles Utilizando Información Métrica y Semántica”, UTN-PID-2173.

REFERENCIAS

- [1] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.
- [2] U. Frese, R. Wagner, and T. Röfer, “A slam overview from a user’s perspective,” *KI-Künstliche Int.*, vol. 24, no. 3, pp. 191–198, 2010.
- [3] F. Auat Cheein, G. Steiner, G. Perez Paina, and R. Carelli, “Optimized eif-slam algorithm for precision agriculture mapping based on stems detection,” *Comput. Electron. Agric.*, vol. 78, pp. 195–207, Sep. 2011.
- [4] G. Perez Paina, C. Paz, M. Baudino, A. Delfino, and E. Destéfani, “Implementation and performance evaluation of ukf for simultaneous localization and mapping,” in *Proceedings of the VII Jornadas Argentinas de Robótica (JAR)*, 2012.
- [5] A. Davison, I. Reid, N. Molton, and O. Stasse, “Monoslam: Real-time single camera slam,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 6, pp. 1052–1067, June 2007.
- [6] J. Civera, O. Grasa, A. Davison, and J. M. M. Montiel, “1-point ransac for ekf-based structure from motion,” in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ Int. Conf. on*, Oct 2009, pp. 3498–3504.
- [7] J. Montiel, J. Civera, and A. Davison, “Unified inverse depth parametrization for monocular slam,” in *Proceedings of Robotics: Science and Systems*, Philadelphia, USA, August 2006.
- [8] G. F. Perez Paina and E. Destéfani, “Monocular simultaneous localization and mapping,” in *Proceedings of the XV Workshop on Information Processing and Control (RPIC)*, 2013, pp. 248–253.
- [9] G. Perez Paina, G. Araguas, D. Gaydou, G. Steiner, and L. Rafael Canali, “Romaa-ii, an open architecture mobile robot,” *Latin America Trans., IEEE*, vol. 12, no. 5, pp. 915–921, Aug 2014.
- [10] J. Andrade-Cetto, T. Vidal-Calleja, and A. Sanfeliu, “Unscented transformation of vehicle states in slam,” in *Robotics and Automation, ICRA 2005. Proceedings of the IEEE Int. Conf. on*, April 2005, pp. 323–328.
- [11] N. Sunderhauf, S. Lange, and P. Protzel, “Using the unscented kalman filter in mono-slam with inverse depth parametrization for autonomous airship control,” in *Safety, Security and Rescue Robotics, 2007. SSR 2007. IEEE International Workshop on*, Sept 2007, pp. 1–6.
- [12] S. Holmes, G. Klein, and D. Murray, “A square root unscented kalman filter for visual monoslam,” in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, May 2008, pp. 3710–3716.
- [13] R. van der Merwe and E. Wan, “Sigma-point kalman filters for integrated navigation,” in *Proceedings of the 60th Annual Meeting of The Institute of Navigation (ION)*, Dayton, OH, June 2004.
- [14] S. Julier, “The scaled unscented transformation,” in *American Control Conference, 2002. Proceedings of the 2002*, vol. 6, 2002, pp. 4555–4559 vol.6.
- [15] Z. Zhang, “A flexible new technique for camera calibration,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [16] G. Araguás, G. Perez Paina, G. Steiner, and L. Canali, “Extrinsic calibration of a camera-robot system under non-holonomic constraints,” in *Proceedings of the AST in 40th Argentine Conference on Informatics (IAII)*, 2011, pp. 157–167.
- [17] J. Civera, A. Davison, and J. M. M. Montiel, “Inverse depth to depth conversion for monocular slam,” in *Robotics and Automation, 2007 IEEE International Conference on*, 2007, pp. 2778–2783.