

# Localización Simultánea a la Reconstrucción de Robots Móviles en Espacios Inteligentes mediante Múltiples Cámaras

Daniel Pizarro Perez, Enrique Santiso Gomez, Manuel Mazo Quintas  
Universidad de Alcalá/Departamento de Electrónica, Alcalá de Henares, España  
{pizarro,santiso,mazo}@depeca.uah.es

**Abstract**—El presente documento aborda el problema del uso de cámaras calibradas y fijas en el entorno para tareas de posicionamiento de robots móviles. El conjunto de cámaras forman un “Espacio Inteligente” con la capacidad de reconocer y controlar robots en su área de percepción visual. El objetivo de este trabajo es ofrecer una solución al problema de la obtención de la pose de robots móviles de estructura rígida mediante un sistema de múltiples vistas. Se propone una solución en ausencia de balizamiento activo a bordo del robot. El algoritmo propuesto se define como un proceso bayesiano de inferencia secuencial. Permite la obtención simultánea de la pose del objeto al mismo tiempo que se adquiere información tridimensional del mismo. Este trabajo conecta directamente el campo clásico de la localización y navegación de robots con una disciplina reciente en la que toda la inteligencia se traslada del robot al entorno equipado con sensores.

## I. INTRODUCCIÓN

El presente trabajo propone el uso de cámaras calibradas y fijas en un entorno para el posicionamiento de robots móviles dentro del mismo. La disposición de dichas cámaras junto con una infraestructura de procesamiento, forman un “espacio inteligente”. Entre los objetivos que ahora se otorgan al entorno se encuentra el posicionamiento y navegación de robots móviles mediante el uso de la percepción visual del mundo. Se pretende así trasladar la inteligencia al entorno, lo cual aporta ventajas sustanciales al caso del posicionamiento y navegación autónoma de vehículos en entornos interiores.

La utilización de cámaras como sensor de percepción es común en el esquema de espacio inteligente. La cantidad de información aportada y la posibilidad de incluir técnicas de procesamiento más costosas a nivel computacional se presentan como factores determinantes. A pesar de las ventajas que aporta la visión, las técnicas utilizadas para posicionamiento se reducen en la mayoría de los casos al balizamiento de aquellos objetos que el entorno debe localizar. El uso de este tipo de marcas artificiales reduce la complejidad de los algoritmos de obtención de pose, descartando sin embargo gran cantidad de información valiosa en aplicaciones reales.

El trabajo propuesto se centra en la utilización de técnicas novedosas en visión artificial, que permiten aprovechar en mayor medida el potencial de información del que se dispone. Estas técnicas extraen estructura tridimensional a partir de múltiples vistas, dando lugar a un mapa del entorno y los objetos que en él se encuentran. La reconstrucción de geometría a partir de

múltiples vistas es un campo de gran expansión y rápido desarrollo. La mayoría de los métodos se caracterizan por utilizar costosos algoritmos de optimización expocencial que no se ajustan a los requerimientos de actualización y complejidad que requiere el posicionamiento de robots móviles.

Se pretende diseñar un método de localización basado en información visual que permita obtener la posición de objetos rígidos contemplados por las múltiples cámaras del espacio inteligente. El sistema debe ser capaz de resolver la pose del robot en espacio métrico sin necesidad de un conocimiento previo de la estructura del mismo o sistema alguno de balizamiento. Como añadido, el sistema aporta el cómputo de la pose de manera que esté disponible en tiempo de ejecución.

El método de posicionamiento se define dentro de un marco probabilístico de inferencia bayesiana secuencial. La posición y orientación del robot se obtiene gracias a la relación existente entre la información visual y un conjunto de puntos o marcas naturales pertenecientes a su estructura tridimensional. Dicha relación se formula como un proceso de inferencia, en el que se busca la distribución ‘a posteriori’ que relaciona información obtenida por el conjunto de cámaras con la estructura y pose del robot.

Como será formalizado más tarde, se define un proceso de Markov que se compone de la pose del robot, así como un conjunto de marcas naturales cuyo número variará de manera dinámica. Este enfoque permite incrementar el conocimiento que el entorno tiene de los objetos que observa a medida que es capaz de obtener su posición. El algoritmo propuesto brinda por tanto un método de reconocimiento tridimensional de objetos en el que el modelo se actualiza dinámicamente.

El documento se organiza como sigue: En primer lugar se presenta la definición del problema como una estructura de estimación bayesiana. Se muestra el modelo de movimiento que se aplica al robot y el modelo de observación de marcas naturales. La sección III mostrará el proceso de observación de marcas naturales que será usado en el sistema, así como el proceso de correspondencia utilizado. El núcleo de inferencia se describe en las secciones IV y V, incluyendo menciones al proceso de inicialización para una o varias cámaras. En la sección VI se presentan resultados numéricos del sistema como validación del proceso completo. Finalmente las conclusiones y los trabajos futuros cierran el presente documento.

## II. LOCALIZACIÓN SIMULTÁNEA A LA RECONSTRUCCIÓN

El problema se define en una estructura bayesiana:

Sea un vector de estados no observables  $\{X_k; k \in \mathbb{N}\}$  que se modela como un *Proceso de Markov* de distribución de probabilidad inicial  $p(X_0)$  y un 'kernel' de transición  $p(X_k|X_{k-1})$ . El vector de observaciones  $\{Y_k; k \in \mathbb{N}\}$  se asume de variables independientes dado el vector de estados  $X_k$  y de distribución marginal  $p(Y_k|X_k)$ .

Todas las distribuciones corresponden y se relacionan con las transiciones del vector de estados y el de observación:

$$\begin{cases} X_k = f(X_{k-1}, W_k) \\ Y_k = h(X_k, V_k) \end{cases} \quad (1)$$

Donde  $W_k$  y  $V_k$  representan en este caso dos procesos Gaussianos incorrelados y mutuamente independientes, que definen respectivamente la incertidumbre en medidas y estimación.

El objetivo principal es la obtención de la distribución  $p(X_k|Y_{1:k})$  que representa la mejor estimación posible del proceso conocido el vector de observación.

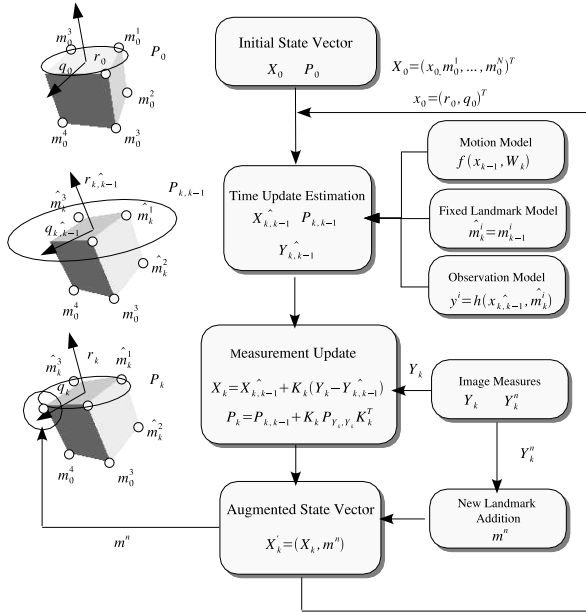


Fig. 1. Proceso de localización simultáneo a la reconstrucción

### A. Vector de estados aumentado

El vector de estados  $X_k$  se compone de dos partes. La primera,  $x_k$  representa la evolución temporal de la pose del robot. El segundo término consiste en la reconstrucción del robot conocida en el instante 'k'. Es decir un número 'M' de puntos  $m_k^i \in \mathbb{R}^3$  que pertenecen a la estructura del robot y se relacionan directamente con el proceso de observación.

$$X_k = (x_k \quad m_k^1 \quad m_k^2 \quad \cdots \quad m_k^M) \quad m_k^i \in \mathbb{R}^3 \quad (2)$$

El número de marcas puede ser dinámico; De esta manera se añaden nuevas marcas al vector durante el proceso 'k > 0'. El valor de las marcas es fijo a lo largo del tiempo puesto que representan valores de coordenadas de puntos que pertenecen a la estructura del robot referenciados a un origen fijo, independiente de la pose que tenga lugar en ese momento 'x\_k'.

De esta manera el vector de estados evoluciona a través del tiempo como se indica a continuación:

$$\begin{cases} x_k = f(x_{k-1}, u_k) \\ m_k^1 = m_k^1 + W_{m_k^1} \\ \vdots \\ m_k^N = m_k^N + W_{m_k^N} \end{cases} \quad (3)$$

La varianza del ruido aditivo  $W_{m_k^i}$  que describe las marcas inmóviles del vector de estados se fuerza generalmente a cero puesto que la incertidumbre de las marcas del vector de estado es constante con el tiempo.

### B. Modelo de Movimiento

En esta sección se presenta el modelo de movimiento usado como proceso de transición de estados. Como se describe en 3, el movimiento del robot se corresponde con un proceso de Markov Discreto 'x' que evoluciona a lo largo del tiempo  $k \in \mathbb{N}^+$  acorde a una función de transición  $f(x_k, u_k)$ .

$$x_k = f(x_{k-1}, u_k) \quad k \in \mathbb{N}^+ \quad (4)$$

El vector  $x_k$  se compone de la posición  $r_k$  y la orientación  $q_k$ :

$$x_k = \begin{pmatrix} r_k \\ q_k \end{pmatrix} \quad r_k \in \mathbb{R}^3, q_k \in \mathbb{R}^4 \quad (5)$$

El vector  $r_k$  representa la posición  $r_k = (r_{x,k} \quad r_{y,k} \quad r_{z,k})^T$  en el espacio euclídeo relativo a un origen de coordenadas que será establecido con posterioridad.

El vector de orientación  $q_k \in \mathbb{R}^4$  es representado por un vector cuaternión unitario. Existen ventajas numéricas en el uso de cuaterniones en lugar de matrices de rotación. Como contrapartida, añade una dimensión extra y necesita cumplir las condiciones de normalidad para representar una rotación.

$$q_k^t \cdot q_k = 1 \quad \forall k > 0 \quad (6)$$

La orientación se refiere a un eje del origen de coordenadas.

La entrada del sistema (4) se representa por 'u\_k'.

Se pueden añadir parámetros adicionales en el vector de estados, como puede ser la velocidad angular ' $\omega \in \mathbb{R}^3$ ' y la velocidad lineal ' $v \in \mathbb{R}^3$ '.

1) *Modelo de movimiento no holonómico*: Una plataforma móvil se dice que tiene un modelo de movimiento no holonómico si su movimiento y orientación están sujetos a restricciones. En general la dirección de la velocidad lineal depende de la orientación, por lo que se define el siguiente modelo de transición de estados.

$$\begin{pmatrix} r_k \\ q_k \\ \omega_k \\ v_k \end{pmatrix} = \begin{pmatrix} r_{k-1} + \Delta r_k \\ q_{k-1} \odot q((\omega_{k-1} + \Omega_k)\Delta T) \\ \omega_{k-1} + \Omega_k \\ v_{k-1} + V_k \end{pmatrix} \quad (7)$$

El valor  $v$  representa el módulo del vector velocidad y su dirección se describe mediante  $q_k$ .

$$\Delta r_k = v_k \Delta T \cdot (q_k)^{-1} \odot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \odot q_k \quad (q_k)^{-1} \odot q_k = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (8)$$

El modelo presentado consiste en 11 parámetros. Si se restringe el modelo de rotación a un plano, como es usual en un robot equipado con ruedas, el número de parámetros se reduce a 6. Este modelo clásico será usado en la sección de resultados y conclusiones.

### C. Modelo de Observación de Marcas Fijas

EL proceso de observación se compone de la proyección en la imagen de un conjunto reducido de marcas de las que existe una componente en correspondencia  $y$  del vector de observación  $Y_k$ .

$$\mathcal{C} = \{m_k^i \mid \exists y \in Y_k, y = h(m_k^i)\}$$

Sea  $m^i$  una marca fija que pertenece al vector de estados  $X_k$ . La posición de  $m^i$  en sistema de coordenadas global de la cámaras  $g_i$  se obtiene utilizando la pose actual que mantiene el robot.

$$\begin{pmatrix} 0 \\ g^i \end{pmatrix} = (q_k)^{-1} \odot \begin{pmatrix} 0 \\ m^i \end{pmatrix} \odot q_k + \begin{pmatrix} 0 \\ r_k \end{pmatrix} \quad (9)$$

El proceso de observación por tanto queda determinado por una función ' $h(x_k, m^i)$ ' que proyecta en el plano imagen cada una de las marcas que se han detectado en la estructura del robot.

$$Y_k = \bar{h}(X_k) + V_k \quad \bar{h} = \begin{pmatrix} h(x_k, m^1) \\ \dots \\ h(x_k, m^M) \end{pmatrix} \quad (10)$$

### III. PROCESO DE OBSERVACIÓN

Como se detalló previamente, varios puntos de la superficie del robot son usados para realizar la corrección del posicionamiento al mismo tiempo que dichos puntos son reconstruidos. Una de las condiciones de partida más importantes de este trabajo es evitar utilizar marcas artificiales y conocidas en el robot. La extracción de marcas naturales es un problema clásico en visión artificial. Existen numerosas técnicas que permiten no solo seguir una característica en sucesivas imágenes, sino que definen una medida de calidad en el proceso de seguimiento. El éxito de estas técnicas depende en gran medida de la deformación proyectiva y el grado de similitud entre imágenes consecutivas.

Se propone un detector de esquinas clásico para obtener un conjunto de puntos susceptibles de ser la correspondencia en la imagen de una marca natural existente en el vector de estados. Resuelto el problema de la extracción

de características, el proceso de correspondencia entre "frames" adyacentes se convierte en un proceso difícil de definir contando tan solo con la información en el plano imagen. Oclusiones y falsas detecciones pueden producir una correspondencia errónea en las observaciones. El proceso de seguimiento entre dos "frames" consecutivos se resuelve aprovechando la información estadística que el sistema posee de cada una de las marcas que sigue. De esta manera la correspondencia se obtiene aplicando un test de 'Mahalanobis' entre la estimación estadística de la posición de una marca ' $y_{k|k-1}$ ' y su observación correspondiente ' $y_k$ '.

$$d_m = (y_k - \hat{y}_{k|k-1})^T P_{y_k|y_k}^{-1} (y_k - \hat{y}_{k|k-1}) \quad (11)$$

Donde ' $P_{y_k|y_k}$ ' es la matriz de covarianza estimada de la proyección de la marca en el plano imagen. Esta distancia permite obtener la correspondencia, identificandola como aquella observación más cercana a la estimación bajo la norma ' $d_m$ '.

Con el objetivo de no despreciar completamente la información que ofrece el plano imagen, se define un factor de similitud entre regiones locales a una característica. Este factor se usará para descartar las marcas que no sean consideradas suficientemente estables durante el proceso. Al igual que en [ST94] la medida de similitud se define mediante el uso de un parche con información de apariencia en un entorno de cada característica. Para cada parche se obtiene un modelo afín que representa con suficiente fidelidad la transformación observada entre imágenes consecutivas. La evaluación de la calidad de esta representación puede ayudar a descartar correspondencias erróneas. Esta medida sin embargo puede significar un filtro demasiado restrictivo y descartar muchas correspondencias buenas en caso de transformaciones proyectivas severas debidas a desplazamientos rápidos.

### IV. NÚCLEO DE INFERENCIA BAYESIANA

Con el objetivo de encontrar una solución al proceso de inferencia Bayesiana, será necesario el uso de técnicas que permitan manejar procesos de transición de estados definidos con funciones no lineales. Dado que en general solo son interesantes los momentos de primer y segundo

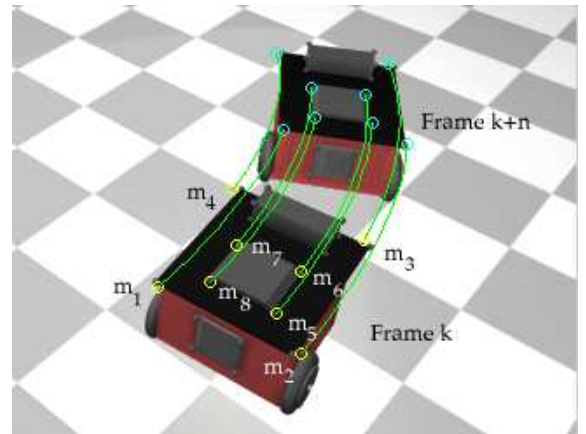


Fig. 2. Seguimiento de marcas en frames consecutivos

orden de la distribución *a posteriori*, es posible utilizar la mejor aproximación Gaussiana de cada uno de los procesos y propagarla a través del tiempo mediante un algoritmo de estimación-corrección como el filtro de Kalman.

Para la obtención del equivalente gaussiano de cada uno de los procesos, muchos autores [Dav02] [GN] hacen uso del famoso Filtro de Kalman Extendido EKF. Este algoritmo tiene un comportamiento bien conocido y una serie de inconvenientes debido a las aproximaciones de primer orden con las que opera, las cuales pueden producir estimaciones inconsistentes.

Con el objetivo de solucionar los problemas de inconsistencia en la estimación aparece una modificación del EKF denominada UKF 'Unscented Kalman Filter', el cual es capaz de conseguir una aproximación buena de los momentos de la distribución hasta segundo orden. A diferencia del EKF, el filtro UKF no aproxima el proceso mediante linealización; en su lugar obtiene una estimación de los estadísticos mediante muestreo de las funciones no lineales. El método de muestreo se denomina *Transformada Unscented* y su descripción detallada se halla en la literatura. [vdMDFW00]

La descripción general del núcleo de inferencia utilizando el filtro UKF se detalla a continuación.

- 1) *Inicialización*: Usando la definición descrita en (1) el punto de partida consiste en los estadísticos para  $k = 0$  de todos los procesos.

$$\begin{cases} \hat{X}_0 = E[X_0] \\ P_0 = E[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T] \\ Q = E[WW^T] \\ R = E[VV^T] \end{cases} \quad (12)$$

- 2) *Predicción en  $k$*

Cálculo de los puntos de muestreo óptimos (Sigma Points) y los pesos ' $\mathcal{W}$ ' para cada proceso a través de la dinámica del sistema y las observaciones.

$$\begin{aligned} \mathcal{X}_{i,k-1} & \mathcal{W}_i^{\mathcal{X}} & i = 1..N_{\mathcal{X}} \\ \mathcal{Q}_{i,k-1} & \mathcal{W}_i^{\mathcal{Q}} & i = 1..N_{\mathcal{Q}} \\ \mathcal{R}_{i,k-1} & \mathcal{W}_i^{\mathcal{R}} & i = 1..N_{\mathcal{R}} \end{aligned} \quad (13)$$

Estimación de la Actualización temporal:

$$\begin{aligned} \mathcal{X}_{k|k-1} &= f(\mathcal{X}_{k-1}, \mathcal{Q}_{k-1}) \\ \hat{X}_{k|k-1} &= \sum_{i=0}^{N_{\mathcal{X}}} \mathcal{W}_i^{\mathcal{X}} \mathcal{X}_{i,k|k-1} \\ P_{k|k-1} &= \mathcal{W}^{\mathcal{X}} [(\mathcal{X} - \hat{X})(\mathcal{X} - \hat{X})^T] \end{aligned}$$

$$\mathcal{Y}_{k|k-1} = h(\mathcal{X}_{k-1}, \mathcal{R}_{k-1}) \quad (14)$$

$$\hat{Y}_{k|k-1} = \sum_{i=0}^{N_{\mathcal{Y}}} \mathcal{W}_i^{\mathcal{Y}} \mathcal{Y}_{i,k|k-1} \quad (15)$$

- 3) *Ecuaciones de actualización de la medida*

$$\begin{aligned} P_{Y_k|Y_k} &= \mathcal{W}[(\mathcal{Y} - \hat{Y})(\mathcal{Y} - \hat{Y})^T] \\ P_{X_k|Y_k} &= \mathcal{W}[(\mathcal{X} - \hat{X})(\mathcal{Y} - \hat{Y})^T] \\ K_k &= P_{X_k|Y_k} P_{Y_k|Y_k}^{-1} \\ \hat{X}_k &= \hat{X}_{k|k-1} + K_k (Y_k - \hat{Y}_{k|k-1}) \\ P_k &= P_{k|k-1} + K_k P_{Y_k|Y_k} K_k^T \end{aligned} \quad (16)$$

Con motivo de una representación más sencilla, algunos subíndices se han omitido en la representación de las covarianzas.

Finalmente el proceso completo evoluciona como sigue y se ilustra en la figura 1:

- 1) En  $k - 1$  la pose que presenta el robot y algunas marcas se encuentran representadas por el vector de estados aumentado  $X_k$ .
- 2) El robot recibe comandos de movimiento desde el espacio inteligente que lo llevan a una nueva posición en  $k$ .
- 3) A través del proceso de estimación del filtro UKF, se obtiene la posición esperada de la imagen del conjunto de marcas que pertenecen en ese momento al robot. De este modo  $\hat{Y}_{k|k-1}$  y  $P_{Y_k|Y_k}$  se usan en (11) para obtener un nuevo vector de observaciones en  $Y_k$ .
- 4) Una vez que las medidas son accesibles, el filtro UKF corrige la estimación del vector de estados  $\hat{X}$  y su covarianza  $P_k$ .
- 5) Si nuevas marcas son detectadas, se incorporan al vector de estados mediante alguno de los métodos de inicialización que se describirán en la siguiente sección.

## V. PROCESOS DE INICIALIZACIÓN

En esta sección se discuten los procesos de inicialización que permiten obtener un conjunto inicial de marcas así como ingresar nuevas al vector de estados.

### A. Inicialización de nuevas marcas mediante una cámara

El proceso de localización descrito en este trabajo necesita de un proceso que añada marcas al vector de estados de una manera rápida y dinámica. Se presenta conveniente no romper la estructura de inferencia secuencial lograda en el proceso de localización. Del filtro UKF, se obtienen la posición y orientación del robot, que en añadido al cálculo de correspondencia se deberá llegar a la reconstrucción de la posición de la nueva marca.

La obtención de un enfoque totalmente bayesiano se realiza a partir de un proceso de triangulación usando un filtro de partículas similar al propuesto en [Dav03].

A partir del conjunto de correspondencias de una nueva marca detectada ' $y_k^n$ ' se obtiene una distribución de probabilidad mediante el filtro de partículas, que representa la posición de la marca con respecto al origen de coordenadas del robot.

La forma de la distribución será más afilada y unimodal con la aportación de nueva información, por lo que se

puede considerar que la incertidumbre en la estimación es menor. Cuando la incertidumbre llegue a un límite impuesto generalmente por la varianza de  $y_k$ , la marca puede ser introducida en el proceso usando la media y covarianza de la distribución resultante.

### B. Inicialización de nuevas marcas mediante múltiples cámaras

La inicialización de nuevas marcas con múltiples cámaras se realiza directamente gracias a (18) y los valores de la pose del robot

Una vez que el proceso de triangulación se termina para un punto ' $X_i$ ', las coordenadas de la nueva marca que será añadida al vector de estados se obtienen referenciando ' $X_i$ ' con respecto al origen de coordenadas del robot ' $r_k$ ' y la orientación ' $q_k$ '.

$$\begin{pmatrix} 0 \\ m^i \end{pmatrix} = (q_k) \odot \begin{pmatrix} 0 \\ X^i - r_k \end{pmatrix} \odot (q_k)^{-1} \quad (17)$$

Con el objetivo de obtener la incertidumbre asociada a la nueva marca, será necesario obtener los estadísticos conjuntos de ' $m_i$ ' en términos de incertidumbre en ' $r_k$ ', ' $q_k$ ' y la incertidumbre del proceso de triangulación. La nueva matriz de covarianza que será usada en el proceso de inferencia se obtiene fácilmente usando la transformada 'Unscented' en la función no lineal (17).

### C. Vector de Estados Inicial: Visión monocular

El proceso de inicialización intenta obtener conocimiento real acerca del movimiento realizado por el robot de acuerdo con un modelo a priori. Se puede considerar como un proceso de calibración del modelo del robot mediante visión y por tanto es de por sí un problema de considerable dificultad. En este documento se tratará de manera resumida por no ser el objetivo del trabajo.

El vector inicial ' $X_0$ ' se define mediante un modelo de movimiento escogido y supuesto preciso y un conjunto inicial de puntos de la estructura del robot a modo de marcas.

La solución propuesta se basa en la obtención de un algoritmo que obtenga ventaja de la información dada por la odometría. La idea es fusionar esta información con la proyección de la cámara para obtener una precisión suficiente en pocos frames y distancia a ciegas recorrida por el robot inferior a la que necesitaría un enfoque de recuperación de la estructura a partir de correspondencias. Dado que la incertidumbre de la posición crece con la distancia, sería recomendable el diseño de un algoritmo de minimización ponderado que tenga en cuenta este hecho.

Además de un algoritmo de optimización eficiente el camino que describe el robot durante el proceso de inicialización es también un aspecto sensible. No todas las trayectorias resultan en una buena o única solución. Las condiciones que permiten un valor estable de inicialización se pueden enumerar como sigue:

- 1) Se requiere un camino diseñado para ofrecer mínimo error de posicionamiento por odometría.

- 2) La trayectoria se deberá componer de una combinación de velocidad angular y lineal. Trayectorias puramente circulares o lineales no ofrecen suficiente información para la calibración.
- 3) Dependiendo de la aplicación el movimiento del robot puede sufrir de restricciones externas.

### D. Vector de Estados Inicial: Múltiples Cámaras

En este punto se discute el problema de la inicialización mediante múltiples cámaras. Mediante correspondencias de puntos y cámaras calibradas se pueden obtener directamente las coordenadas en tres dimensiones de un punto. De esta manera, con el objetivo de simplificar y compactar se define una función general de triangulación. Para  $N_c$  cámaras y un conjunto de correspondencias de puntos ' $(y_1, ..y_{N_c})$ ', las coordenadas euclídeas verdaderas se obtienen.

$$X = t(y_1, ..y_{N_c}) \quad (18)$$

Para asegurar un algoritmo independiente de la odometría, se propone una estrategia rápida y simple para obtener los parámetros de movimiento.

- 1) Haciendo que el robot gire en torno a su eje ' $r_0$ ', éste se determina fácilmente mediante identificación de todas las circunferencias en espacio tridimensional. Siendo Para un conjunto de  $N_p$  puntos, el algoritmo de optimización deberá obtener el conjunto de radios  $R_i$  para cada punto y el centro de rotación. La función de coste se construye como sigue:
- 2) Una vez que ' $r_0$ ' es obtenido, la orientación se calcula de manera trivial a partir de un movimiento en línea recta.

## VI. SIMULACIONES Y RESULTADOS NUMÉRICOS

Con el propósito de validar el algoritmo propuesto, se hace necesario un método que permita obtener parámetros de pose y estructura de un objeto que será proyectado en el plano imagen.

Se propone un enfoque de simulación para validar los algoritmos y obtener el error exacto que se introduce en los resultados del sistema completo. Mediante simulación es posible usar diferentes fuentes de luz, configuraciones de cámara y apariencia de los robots.

Mediante un software de generación de imágenes se realiza una simulación foto-realista de la escena completa y el conjunto de cámaras (Ver 3). Permite controlar las fuentes de luz y el tipo de modelo de cámara. Se puede considerar que los efectos derivados de oclusiones y la precisión obtenida son afines a los encontrados en imágenes reales. Además, el camino a lo largo del cual el robot es navegado puede ser definido por el usuario mediante curvas de tipo 'spline'.

Usando este tipo de simulación se presentan algunos experimentos para obtener la posición y orientación del robot. Ver figuras (4) y (5)

## VII. CONCLUSIONES

Este trabajo presenta un método novedoso para lograr la localización de un robot usando la información tomada por sensores de visión externos. Además del proceso de localización, se obtienen las coordenadas tridimensionales de puntos usados como marcas dinámicas y pertenecientes a la estructura del robot .

Para alcanzar los objetivos que propone el trabajo, se redefine el problema como uno de inferencia bayesiana, el cual permite el uso de potentes técnicas de inferencia ya usadas con éxito en casos de cámaras a bordo de robots o de movimiento libre.

La detección robusta de características y el proceso de seguimiento es crucial para la obtención de una corrección consistente de la pose del robot. El problema de correspondencia entre marcas y medidas es de complejidad exponencial. Se proponen dos métodos dependiendo de la información previa disponible:

- En primer lugar se propone un enfoque clásico a la búsqueda de marcas naturales para realizar su seguimiento a partir de sólo información visual.
- El segundo algoritmo parte de una representación estocástica de la proyección de marcas. La información extra permite el diseño de un test de 'Mahalanobis' que incluye correlación cruzada estadística. Se propone un test de Compatibilidad conjunta que hace uso de la relación estadística entre las diferentes marcas detectadas.

Se definen del mismo modo procesos de inicialización para una y varias cámaras, completando así las contribuciones de este documento.

Sin embargo algunos inconvenientes no son tratados con detalle y deberán ser el foco de atención en trabajos futuros.

Los más importantes se muestran a continuación:

- Un estudio minucioso de la detección y modelado de oclusiones.
- Evaluación de datos mínimos y diseño de un proceso robusto.
- Ampliación del proceso para múltiples cámaras añadiendo restricciones impuestas por la geometría

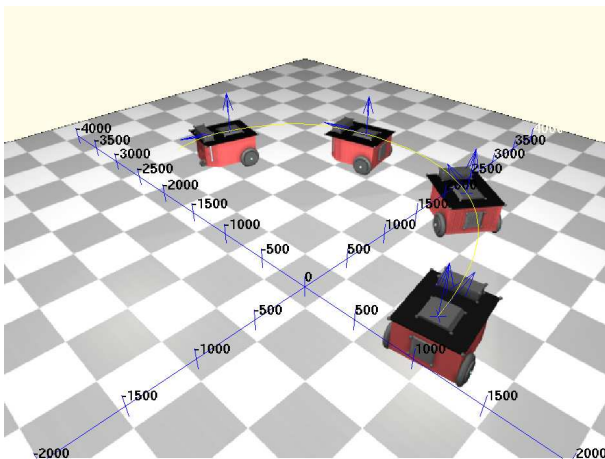


Fig. 3. Simulación 3D de la escena

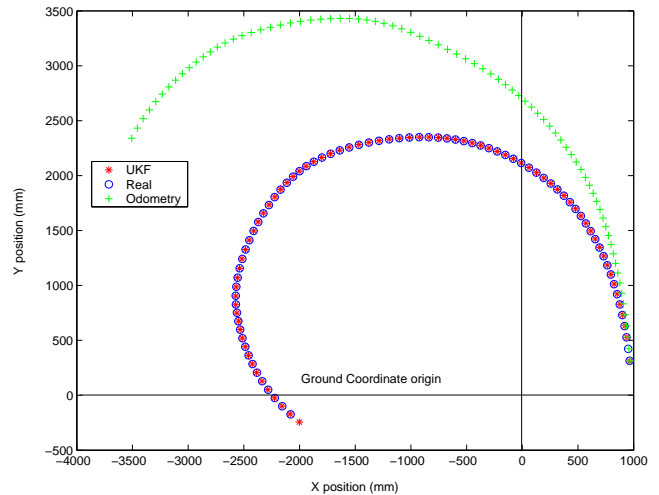


Fig. 4. Trayectoria del robot en el plano X-Y

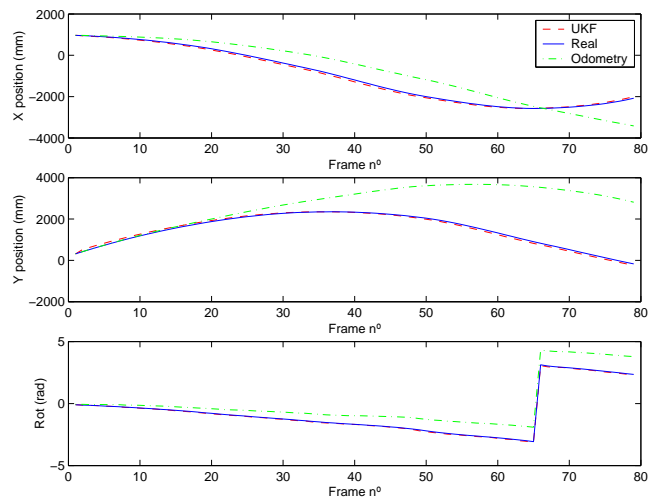


Fig. 5. Posición y rotación vs frame

de multiples cámaras.

- Extensión del algoritmo para múltiples robots.

## REFERENCES

- [Dav02] A. J. Davison. SLAM with a single camera. In *Workshop on Concurrent Mapping and Localization for Autonomous Mobile Robots, in conjunction with ICRA*, 2002.
- [Dav03] Andrew J. Davison. Real-time simultaneous localisation and mapping with a single camera. In *ICCV3*, pages 1403–1410, 2003.
- [GN] J. E. Guivant and E. M. Nebot. Optimization of the simultaneous localization and map building algorithm for real time implementation.
- [ST94] Jianbo Shi and Carlo Tomasi. Good features to track. In *1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pages 593 – 600, 1994.
- [vdMFDW00] Rudolph van der Merwe, Nando de Freitas, Arnaud Doucet, and Eric Wan. The unscented particle filter. Technical Report CUED/F-INFENG/TR380, Cambridge University Engineering Department, aug 2000.