

Sistema multicámaras para detección de posición de un robot móvil

Christofer Alcaraz¹

¹Facultad de Ingeniería Electromecánica, Universidad de Colima, Carretera Manzanillo-Cihuatlán Km. 20, C.P. 28860, Manzanillo, Colima, México.

14 de agosto de 2024

Resumen: La orientación y posición de un vehículo es un problema fundamental en la robótica móvil aplicada. Ingenieros e investigadores han propuesto muchos métodos para resolverlo. Sin embargo, la mayoría de las propuestas permanecen en estado experimental y ninguna se ha establecido como el estándar de referencia. Además, existen sistemas especializados en posicionamiento que son costosos y poco accesibles. En este trabajo, proponemos un sistema de posicionamiento con un conjunto de cámaras convencionales (webcams). Evaluamos la calidad de los resultados usando equipamiento accesible para un robot móvil omnidireccional.

Palabras clave: Visión por computadora, Robótica móvil, triangulación, Posicionamiento, distorted pinhole model, segmentación por color.

Índice

| | |
|--|--|
| 1. Introducción | |
| 2. Metodología | |
| 2.1. Corrección de perspectiva | |
| 2.2. Coordenadas Homogéneas | |
| 2.3. Procesamiento de imagen | |
| 2.4. Modelado de cámara | |
| 3. Resultados | |
| 3.1. Calibración de cámaras | |

| | |
|---|---|
| 3.2. Detección de objetos por color | 8 |
| 3.3. Robot Omnidireccional | 8 |
| 4. Conclusiones | 9 |

1. Introducción

El conocimiento de la posición y orientación de los vehículos es fundamental en la robótica móvil aplicada. Existen muchas técnicas y sensores disponibles actualmente. Sin embargo, la estimación de posición en robótica móvil sigue siendo un problema abierto en la comunidad científica. Se están investigando alternativas para abordar este problema. La diversidad de métodos y requerimientos dificulta la comparación objetiva entre metodologías diferentes [1]. La odometría es una técnica de estimación de posición relativa, ofrece un bajo costo y una precisión aceptable en períodos cortos. Sin embargo, la odometría presenta altos niveles de error en períodos largos. Los sistemas de visión especializados, como OptiTrack, ofrecen alta precisión, también garantizan robustez en cualquier período de operación. Desafortunadamente, estos sistemas son poco accesibles por su alto costo. En este proyecto de verano, construimos un sistema de visión de bajo costo, utilizamos cuatro cámaras convencionales (webcams) para el posicionamiento de un robot móvil.

2. Metodología

2.1. Corrección de perspectiva

Un aspecto clave de los sistemas de visión es la corrección de perspectiva. Esto implica transformar imágenes de píxeles desde el plano interno de la cámara al plano de referencia en el mundo tridimensional donde se usarán. Para corregir la perspectiva, se generan homografías desde el plano imagen.

Para crear el algoritmo que genera matrices homografía, se deben entender conceptos de visión por computadora. La Figura 1 muestra un diagrama de un sistema de visión. El algoritmo busca encontrar la relación entre los puntos del plano imagen y los del plano de referencia [2].

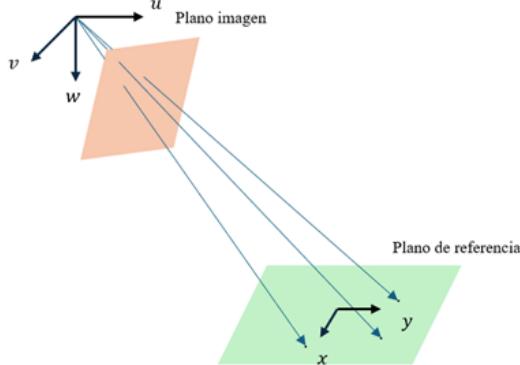


Figura 1: Diagrama de un sistema de visión con 1 cámara.

2.2. Coordenadas Homogéneas

Hay varias formas de representar puntos en el espacio 3D, como las coordenadas cartesianas, polares y ángulos directores. En los sistemas de visión, las coordenadas homogéneas son útiles. Esto se debe a que las transformaciones en este plano son lineales, a diferencia de las coordenadas cartesianas.

En coordenadas cartesianas, los puntos en 1 dimensión se representan con un vector de 1 componente. Lo mismo ocurre para 2 y 3 dimensiones, donde cada

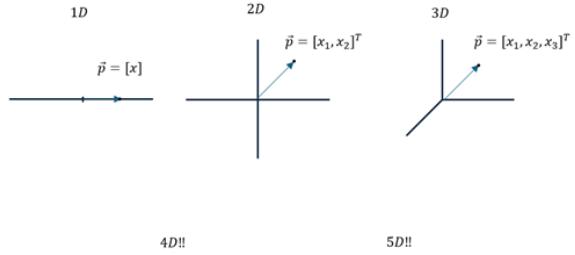


Figura 2: Representación de un punto en el espacio de n dimensiones con coordenadas cartesianas.

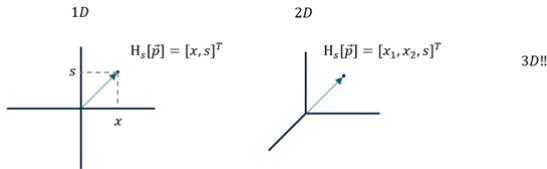


Figura 3: Representación de un punto en el espacio de n dimensiones con coordenadas homogéneas.

componente indica una dimensión, como se muestra en la figura 2.

En las coordenadas homogéneas, se añade una dimensión extra al vector. Esta dimensión se representa con un valor al final del vector, como se muestra en la figura 3. La distancia del punto al origen en esta nueva dimensión se llama escala. La escala indica cómo está representada la coordenada.

Se puede definir el operador H_s . Este operador convierte un punto P a coordenadas homogéneas. Lo hace añadiendo una dimensión extra con valor s . El valor s puede ser cualquier número real.

$$P = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, H_s[P] = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ s \end{bmatrix} \quad (1)$$

Este operador considera dos casos particulares. El primer caso ocurre cuando la escala es igual a uno. En este caso, se puede omitir la escala en el operador

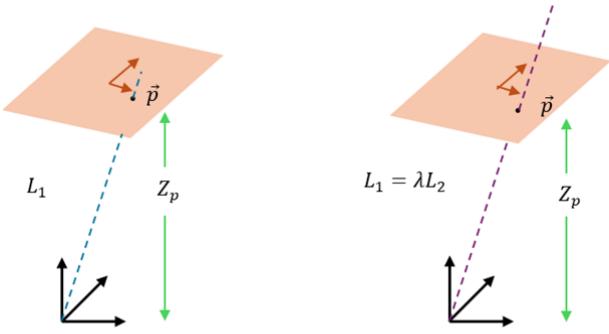


Figura 4: Representación gráfica de una propiedad fundamental de las coordenadas homogéneas, donde la recta que intercepta al punto en el plano se puede multiplicar por cualquier escalar λ

H . El segundo caso sucede cuando la escala es cero. Aquí, el resultado no afecta la dirección ni la norma del vector P . Esto permite describir la ubicación de puntos en el infinito.

$$H[P] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix}, H_0[P] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 0 \end{bmatrix} \quad (2)$$

Una propiedad interesante del operador H es que una recta generada es invariante al producto de un escalar λ . Supongamos que un punto $p = [x_1, x_2]^T$ representa la intersección de la recta L con el plano ubicado en z_p , como se muestra en la figura 4. Solo nos interesa la intersección de la recta con el plano. Por lo tanto, esta recta puede multiplicarse por un factor arbitrario λ sin afectar el punto P . Gracias a esta propiedad, podemos concluir que en el espacio de coordenadas homogéneas existe una infinidad de vectores que representan el mismo punto en el espacio [3].

$$Y = \lambda H_s[P] \quad (3)$$

El operador de coordenadas homogéneas inverso tiene como objetivo recuperar el punto dado a partir

de sus coordenadas homogéneas. Este se define con la notación H_s^{-1} , donde s puede ser representado por cualquier número real. Por ejemplo, si consideramos un vector de dimensiones $(n + 1)$, las cuales son las coordenadas homogéneas de un punto en espacio de n dimensiones.

$$Y = [y_1 \ \cdots \ y_n \ y_{n+1}]^T \quad (4)$$

Al realizar el proceso inverso, el operador $H_0^{-1}[Y]$ recupera los valores del vector excepto el último, es decir, y_{n+1} .

$$H_0^{-1}[Y] = [y_1 \ \cdots \ y_n \ y_{n+1}]^T \quad (5)$$

Este operador solo recupera el punto en la escala 0, sin embargo, cuando la escala corresponde a otro número real arbitrario s , es necesario considerar que el vector puede estar multiplicado por un escalar λ . Esto significa que:

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} \lambda H_s[x_1] \\ \vdots \\ \lambda H_s[x_n] \\ \lambda s \end{bmatrix} = \lambda H_s[X] \quad (6)$$

En este caso, X representa el punto de coordenadas cartesianas que queremos recuperar. Al observar la ecuación anterior, podemos ver que es necesario eliminar el valor escalar λ para obtener las componentes del punto $H_s[X]$. Este valor se puede determinar al observar el último elemento de las coordenadas homogéneas del punto. Este valor es importante para el operador de coordenadas homogéneas inversas, por lo que se define como S . Por lo tanto, se tiene que:

$$S[Y] = y_{n+1} = \lambda s \quad (7)$$

Si se conoce la escala S y la última entrada de un vector $(n+1)$ -dimensional, es posible eliminar el factor λ a partir de la siguiente ecuación:

$$H_s[X] = \frac{s}{S[Y]} = \frac{s}{\lambda s} \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \\ \lambda s \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ s \end{bmatrix} \quad (8)$$

Finalmente, para recuperar las coordenadas cartesianas queda eliminar el último elemento del vector, esto se puede realizar a través del operador de coordenadas homogéneas inverso (H_0^{-1}) [3], por lo tanto, el operador generalizado de coordenadas homogéneas inverso para cualquier escala queda definido de la siguiente manera:

$$H_s^{-1}[Y] = \frac{s}{S[Y]} H_0^{-1}[Y] \quad (9)$$

2.3. Procesamiento de imagen

2.3.1. Segmentación de imagen

La segmentación de imágenes implica dividir una imagen en sus regiones u objetos componentes. La extensión de esta división varía según el problema y la segmentación se completa cuando se aísla el objeto de interés. Este proceso es uno de los mayores desafíos en visión por computadora. Actualmente, hay muchos enfoques para abordarlo [4].

Aunque la segmentación de imágenes en escala de grises ha sido ampliamente explorada, la segmentación basada en el color ha ganado relevancia. Esto se debe a su capacidad para extraer más información que las imágenes en blanco y negro. Además, los avances en la capacidad computacional actual permiten implementar más algoritmos en este contexto [5].

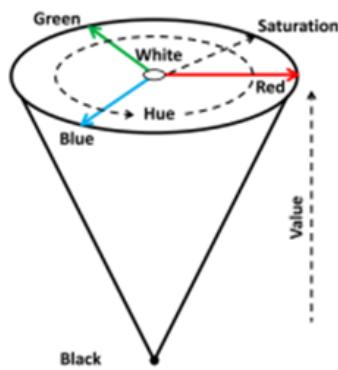


Figura 5: Espacio de color HSV.

Existen diferentes propuestas para la segmentación

por color. Estas pueden clasificarse de la siguiente manera: detección de bordes, umbral, umbral de histograma, región, agrupación de características, y métodos basados en redes neuronales. Después de analizar las diferentes técnicas, se concluyó que la segmentación por color era la opción más adecuada por su simplicidad y facilidad de implementación [4].

2.3.2. Modelos de color

Un modelo de color es un enfoque matemático para representar colores mediante listas ordenadas y finitas de elementos. Cada elemento corresponde a un componente diferente de color. La interpretación de estos componentes da como resultado un espacio de color. Existen varios espacios de color, como RGB, NTSC, YVbCr, HSV, CMY, entre otros.

El modelo HSV toma su nombre de las iniciales en inglés de Hue (matiz), Saturation (saturación) y Value (valor). Es un espacio de color que describe cómo los humanos perciben el color. El matiz se refiere al color dominante que ve el ojo humano, es decir, el tono. La saturación indica la cantidad de luz blanca mezclada con el matiz y representa la intensidad del color. El valor muestra la luminosidad o intensidad del color.

El espacio de color HSV se puede visualizar como un cono geométrico. La dimensión angular representa el matiz, comenzando con el rojo en 0 grados, pasando al verde a 120 grados, al azul a 240 grados, y volviendo al rojo a 360 grados. La distancia desde el eje central muestra la saturación, que varía de 0 a 1. El eje vertical representa el valor, con negro en la parte inferior (valor 0) y blanco en la parte superior (valor 1). [4]

2.3.3. Centroide de una figura

El centroide de una figura geométrica regular tiene una fórmula específica por figura. Sin embargo, en este trabajo emplearemos la fórmula generalizada, que permite trabajar con figuras irregulares que pueden incluso rotar, definida como [6]:

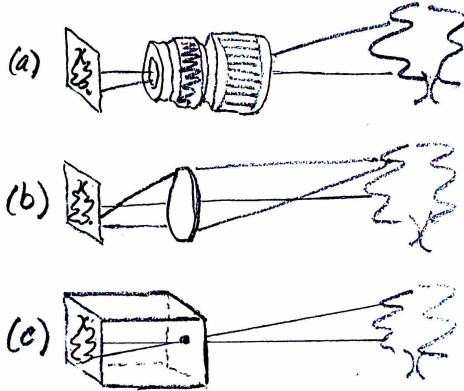


Figura 6: Diferentes modelos de cámaras. (a) Cámara de lente computesta. (b) Cámara de lente simple. (c) Cámara pinhole.

$$\begin{aligned} x &= \frac{\sum(X \times \text{PixelValues})}{\sum \text{PixelValues}} \\ y &= \frac{\sum(Y \times \text{PixelValues})}{\sum \text{PixelValues}} \end{aligned} \quad (10)$$

2.4. Modelado de cámara

Para analizar un sistema, proceso u objeto en la realidad, es necesario desarrollar una expresión matemática que lo describa, conocida como modelo. Para hacerlo, es crucial comprender cómo funciona el sistema. En el caso de las cámaras digitales, la imagen se crea a través de un complejo conjunto de lentes (a). Sin embargo, es posible aproximar este conjunto a una cámara con una sola lente simple (b). Este nuevo modelo comparte las mismas propiedades que el modelo más complejo. Puede enfocar múltiples rayos de luz de la realidad y mapearlos en el plano de la imagen (el sensor fotosensible) de manera idéntica.

La aproximación entre modelos se facilita gracias al propósito específico de la aplicación. No es esencial considerar la capacidad del sistema óptico para enfocar los rayos de luz de la escena observada. Lo importante es el mapeo de los rayos de luz entre la escena y el plano de la imagen. Por lo tanto, este modelo se puede simplificar aún más. Podemos con-

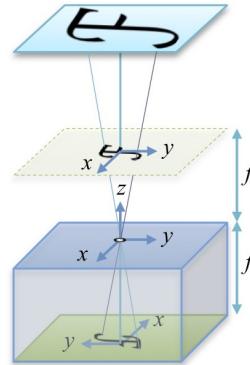


Figura 7: Proceso de formación de imagen a través del modelo Pinhole.

siderar el sistema óptico como una caja cerrada con un agujero por donde entran los rayos de luz (c).

Este modelo, conocido como *pinhole*, bloquea todos los rayos de luz excepto uno. Esto resulta en una correspondencia de uno a uno en el plano de la imagen. Los puntos están conectados por una línea que atraviesa el *pinhole*.

Al analizar el modelo de la figura 5, se observa que la proyección en el plano de la imagen está invertida en el eje X y el eje Y. Para corregir esto, se cambia la proyección de $z = -f$ a $z = +f$. En este punto, ningún eje está invertido y es equivalente a la proyección principal.

Otra característica importante de este modelo es que el sistema de coordenadas es estático y está ubicado en el *pinhole*. Aunque la mayoría de los rayos sufren una pequeña distorsión, el rayo que pasa por el origen es paralelo al eje z y no sufre distorsión. Este eje se conoce como el eje óptico [7].

2.4.1. Homografías

A menudo, en un sistema de visión, se busca encontrar las correspondencias entre un plano en el espacio tridimensional y el plano de imagen. Esto es un caso particular de la matriz de la cámara C , ya que los puntos en el espacio tridimensional deben ser coplanares. Esta nueva matriz, que realiza el mapeo entre dos sistemas de coordenadas coplanares, se llama *matriz homografía* y se denota con la letra G . Esta

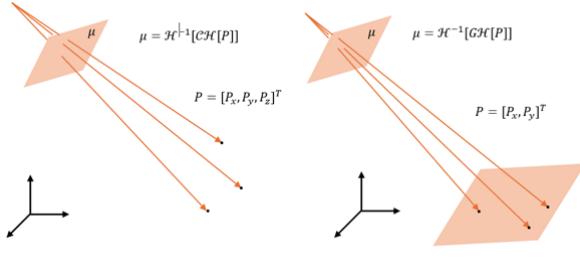


Figura 8: Comparación entre la correspondencia de puntos en el espacio y correspondencia de puntos coplanares.

matriz mapea los puntos ρ del plano de referencia con los puntos del plano de imagen μ .

Como se observa en la figura 8, la matriz de homografía es muy similar a la matriz de la cámara. La diferencia principal es que los puntos en el sistema de referencia de coordenadas están en $z = 0$. Debido a esto, la última fila de la matriz de la cámara es 0. Esto nos deja con una matriz no singular de tamaño 3×3 . Por lo tanto, podemos definir la matriz homografía a partir de la matriz de la cámara.

$$G = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} = \begin{bmatrix} \bar{g}_1^T \\ \bar{g}_2^T \\ \bar{g}_3^T \end{bmatrix} \quad (11)$$

De la expresión anterior se puede sustituir en la ecuación que une las correspondencias.

$$\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} = \begin{bmatrix} \bar{g}_1^T H[\rho] \\ \bar{g}_2^T H[\rho] \\ \bar{g}_3^T H[\rho] \end{bmatrix} \quad (12)$$

Al realizar la operación de las coordenadas homogéneas inversa, se obtiene

$$\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} = \frac{1}{\bar{g}_3^T H[\rho]} \begin{bmatrix} \bar{g}_1^T H[\rho] \\ \bar{g}_2^T H[\rho] \end{bmatrix} \quad (13)$$

Esta expresión se puede igualar a 0, donde se obtiene un sistema de ecuaciones de $2n$, donde n es el número de correspondencias en los dos sistemas de

referencia. Dado el número de incógnitas que se necesitan para calcular la matriz homografía G , se necesitan al menos 4 correspondencias.

$$\begin{bmatrix} \bar{g}_1^T H[\rho] \\ \bar{g}_2^T H[\rho] \end{bmatrix} - \begin{bmatrix} (\bar{g}_3^T H[\rho])\mu_x \\ (\bar{g}_3^T H[\rho])\mu_y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (14)$$

Al despejar la ecuación anterior se puede obtener el vector que corresponde a las filas de la matriz homografía.

$$\begin{bmatrix} H[\rho]^T & 0_3^T & -\mu_x H[\rho]^T \\ 0_3^T & H[\rho]^T & -\mu_y H[\rho]^T \end{bmatrix} \begin{bmatrix} \bar{g}_1^T \\ \bar{g}_2^T \\ \bar{g}_3^T \end{bmatrix} = 0_2 \quad (15)$$

Este sistema de ecuaciones se busca la solución mínima, sin embargo, tiene la particularidad de que esta igualada a 0, por lo que una de las soluciones para el vector de las columnas de la matriz G puede ser 0, por lo tanto, se utiliza la descomposición en valores singulares (SVD) de la matriz que acompaña al vector de incógnitas, es decir

$$A = U \Sigma V^T \quad (16)$$

Donde U y V son matrices ortogonales y σ es una matriz diagonal con los valores singulares de A en la diagonal. Dicha expresión se puede representar de la siguiente forma.

$$\begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} \sigma_1 v_1^T \\ \sigma_2 v_2^T \\ \sigma_3 v_3^T \end{bmatrix} \quad (17)$$

Donde $\sigma_1 \geq \sigma_2 \geq \sigma_3$. Debido a que se busca que la solución sea mínima se puede determinar que la solución del sistema de ecuaciones se encuentra en la última columna de la matriz V [7].

2.4.2. Triangulación

La mayoría de los algoritmos de triangulación en visión por computadora se basan en el modelo de la cámara pinhole. La forma generalizada de la ecuación que relaciona un punto en el espacio con el plano imagen se detemina

$$\mu = H^{-1}[CH[P]] \quad (18)$$

Al tratar de despejar el punto P de la ecuación anterior se observa que, por propiedades de las coordenadas homogéneas, aparece un número real λ , el cual representa todos los valores por los que atraviesa la proyección del punto desde el sistema de coordenadas de la cámara, es decir, resulta las siguientes ecuaciones.

$$\begin{aligned}\lambda_1 H[\mu_1] &= K_1 \begin{bmatrix} R_1^T & -R_1^T t_1 \end{bmatrix} \begin{bmatrix} P \\ 1 \end{bmatrix} \\ \lambda_2 H[\mu_2] &= K_2 \begin{bmatrix} R_2^T & -R_2^T t_2 \end{bmatrix} \begin{bmatrix} P \\ 1 \end{bmatrix}\end{aligned}\quad (19)$$

Donde $C_1 = K_1[R_1^T - R_1^T t_1]$. Al despejar P de las 2 ecuaciones se obtienen las siguientes expresiones

$$\begin{aligned}P &= t_1 + \lambda_1 R_1 K_1^{-1} H[\mu_1] \\ P &= t_2 + \lambda_2 R_2 K_2^{-1} H[\mu_2]\end{aligned}\quad (20)$$

Si sustituimos las expresiones $R_1 K_1^{-1} H[\mu_1] = d_1$ y $R_2 K_2^{-1} H[\mu_2] = d_2$ nos queda el siguientes sistema de ecuaciones

$$\begin{bmatrix} d_2 & -d_2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = t_2 - t_1 \quad (21)$$

En el cual podemos obtener los valores de λ_1 y λ_2 . Por último, la expresión que une los 2 puntos de coordenadas μ de las 2 cámaras con un punto en el espacio tridimensional se define como [8]

$$p = \frac{t_1 + t_2 + \lambda_1 d_1 + \lambda_2 d_2}{2} \quad (22)$$

2.4.3. Triangulación generalizada

Para realizar la triangulación con m número de cámaras, es necesario generalizar la expresión descrita anteriormente. Esta técnica se conoce como Media de puntos con mínima dispersión (Mean of Points with Minimum Dispersion). Para cada dispositivo, se observa un punto detectado por un rayo de luz. A menudo, estos rayos son líneas oblicuas debido a la resolución y los errores de detección.

$$p_i = t_i + \lambda_i b_i, \quad i = 1, 2, \dots, m. \quad (23)$$

Por lo tanto el punto observado puede ser definido como una media de puntos p_i como

$$P = \frac{1}{m} \sum_{i=1}^m p_i = \frac{1}{m} (Bl + T1_m) \quad (24)$$

, donde $1_m = [1, 1, \dots, 1]^T$ es un vector de dimensión $m \times 1$ y todos sus valores son 1, B y T son matrices de dimensión $3 \times m$ definidos como

$$\begin{aligned}B &= [b_1 \quad b_2 \quad \dots \quad b_m] \\ T &= [t_1 \quad t_2 \quad \dots \quad t_m]\end{aligned}\quad (25)$$

y l es un vector de dimensión $m \times 1$ desconocido dado por

$$l = [\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_m]^T \quad (26)$$

Para obtener los valores del vector l se puede determinar a través de la expresión

$$l = (N - \frac{1}{m} B^T B)^{-1} (\frac{1}{m} B^T T 1_m - h) \quad (27)$$

donde N es una matriz diagonal $m \times m$, y h es una matriz de dimensión $m \times 1$ definidos respectivamente como

$$\begin{aligned}N &= diag(b_1^T b_1, b_2^T b_2, \dots, b_m^T b_m), \\ h &= [t_1^T b_1 \quad t_2^T b_2 \quad \dots \quad t_m^T b_m]^T\end{aligned}\quad (28)$$

Al resolver la ecuación (27), el vector resultante l se substituye en la ecuación (24), para obtener las coordenadas del punto observado P en el espacio [8].

3. Resultados

3.1. Calibración de cámaras

Se realizó la calibración de cada una de las 4 cámaras. Se tomaron entre 13 y 17 fotos por cámara a una cuadrícula (patrón de calibración), como se muestra en la Figura 9(Izda). Después de calibrar la cámara, se capturó una imagen extra para calcular la posición



Figura 9: (Izda.) Imagen tomada para la obtención de los parámetros intrínseco y extrínsecos de la cámara. (Centro) Imagen de referencia utilizada para ubicar la cámara en el sistema de coordenadas de referencia. (Dcha.) Imagen procesada con la corrección de perspectiva.

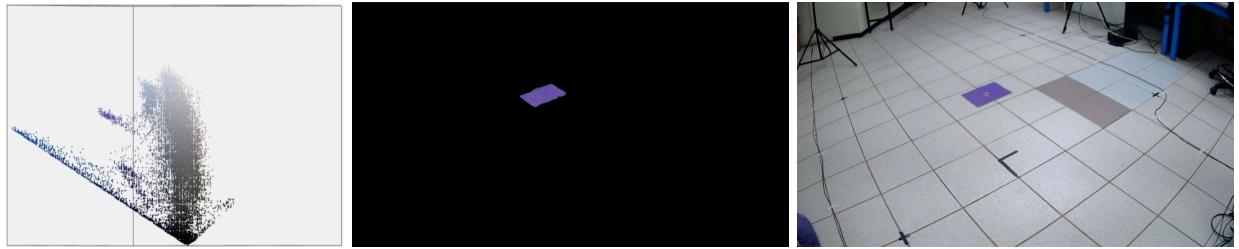


Figura 10: (Izda.) Cono característico del Espacio HSV. (Centro) Imagen segmentada por detección de color morado. (Dcha.) Centroide del objeto morado detectado en la imagen.

de la cámara respecto al sistema de referencia, como se muestra en la figura 9(Centro). Este último proceso se realizó usando cuatro puntos marcados en el suelo para obtener la homografía asociada y de ella calcular la ubicar y orientación de la cámara. Más aún, la homografía obtenida permite corregir la perspectiva de la imagen original, como se muestra en la Figura 9(Dcha). Este proceso de calibración se repitió para las cuatro cámaras del sistema propuesto.

3.2. Detección de objetos por color

Para detectar la posición del robot, se utilizaron distintivos de colores. Los colores tienden a variar con los cambios de luz exterior y la configuración de la cámara. Esto se debe a que el color percibido es subjetivo. Por lo tanto, en lugar de definir un color exacto, se establece un umbral para permitir cierta tolerancia ante estos cambios.

En la Figura 9(Centro) se muestra una imagen capturada por una cámara del sistema de visión. En la

Figura 10(Izda) se muestra el cono característico del modelo HSV. A esta imagen se le aplicó la segmentación al color morado, dando como resultado la figura 10(Centro).

Después el resultado se transforma en una nueva imagen, donde los píxeles detectados tienen el valor de 1, mientras que el resto de los píxeles toman el valor de 0, esto con la finalidad de ser capaz de calcular de forma más eficiente el centroide del objeto a través de la ecuación 10. La figura 10(Dcha.) muestra contiene una marca en las coordenadas píxel del centroide detectado.

3.3. Robot Omnidireccional

Para este sistema se utilizó un robot móvil omnidiereccional, al cual se le colocaron 2 distintivos, como se observa en la figura 12. El distintivo color naranja representa el punto de control del robot, mientras que el distintivo color rosa se utiliza para calcular la orientación del robot.

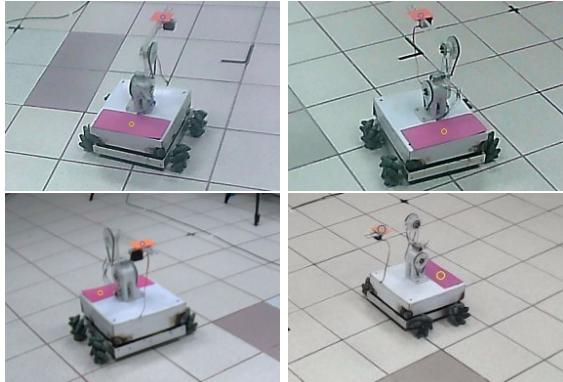


Figura 11: Detección del robot a través del sistema de cuatro cámaras propuesto usando distintivos de color.

Una vez detectado por las 4 cámaras, las coordenadas pixeles se convierten en rayos de dirección, seguido del algoritmo de triangulación, en la figura 12 (Arriba) se observa la altura real contra la altura calculada por el sistema de visión, visualizada en el servidor, la cual esta siendo enviada al control del robot, figura 12 (Abajo).

La creación de sistemas de posicionamiento mediante visión requiere varias etapas. Primero, se debe ubicar las cámaras. Es necesario saber cuál será el área de trabajo donde se detectará la posición. Luego, se realiza la calibración, que consiste en obtener fotografías útiles. A veces, las fotos tomadas no permiten calibrar las cámaras de manera adecuada.

Después, se comprueban las estimaciones. Es necesario verificar las posiciones calculadas por el sistema de calibración. Este proceso suele ser de prueba y error, especialmente para principiantes. Por último, se revisa la ubicación. Una vez completada la calibración, el sistema queda fijo.

Cuando el sistema de posicionamiento está listo, se puede usar en cualquier momento. Además, funciona para otras aplicaciones.

4. Conclusiones

Los sistemas de visión multicámara con cámaras convencionales son una solución de bajo costo pa-

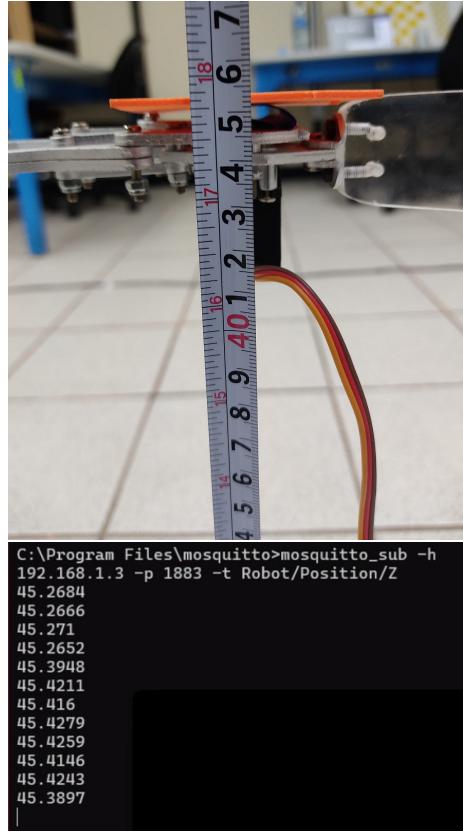


Figura 12: (Arriba) Distintivo naranja sobre el robot a una altura de 45.4 mm. (Abajo) Altura de 45.38 mm obtenida mediante el sistema multicámara propuesto.

ra el posicionamiento de robots móviles. Aunque la calibración puede ser laboriosa al principio, una vez calibrado y fijo, el sistema está listo para su uso. Un sistema de visión como el propuesto puede servir para otros propósitos, además de mejorar el conocimiento del entorno del robot.

Por ejemplo, se pueden detectar objetos para usar en algoritmos de evasión de obstáculos. También es posible generar un mapa del área de trabajo para realizar simulaciones y mejorar el entendimiento del entorno. Además, un mismo sistema de visión puede ser usado para el posicionamiento de cualquier tipo de sistema robótico, incluidos brazos robóticos y drones. Esto ofrece gran flexibilidad para su uso.

Referencias

- [1] J. Borenstein, H. R. Everett, L. Feng, and D. Wehe, “Mobile robot positioning: Sensors and techniques,” *Journal of robotic systems*, vol. 14, no. 4, pp. 231–249, 1997.
- [2] R. Juarez-Salazar, J. Zheng, and V. H. Diaz-Ramirez, “Distorted pinhole camera modeling and calibration,” *Applied Optics*, vol. 59, pp. 11310–11318, Dec 2020.
- [3] R. Juarez-Salazar and V. H. Díaz-Ramírez, “Operator-based homogeneous coordinates: application in camera document scanning,” *Optical Engineering*, vol. 56, no. 7, p. 070801, 2017.
- [4] D. Hema and D. S. Kannan, “Interactive color image segmentation using hsv color space,” *Sci. Technol. J*, vol. 7, no. 1, pp. 37–41, 2019.
- [5] F. Garcia-Lamont, J. Cervantes, A. López, and L. Rodriguez, “Segmentation of images by color features: A survey,” *Neurocomputing*, vol. 292, pp. 1–27, 2018.
- [6] Y. N. Naggar, A. H. Kassem, and M. S. Bayoumi, “A low cost indoor positioning system using computer vision,” *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, vol. 11, no. 4, pp. 8–25, 2019.
- [7] R. Juarez-Salazar, L. N. Gaxiola, and V. H. Diaz-Ramirez, “Single-shot camera position estimation by crossed grating imaging,” *Optics Communications*, vol. 382, pp. 585–594, 2017.
- [8] R. Juarez-Salazar, G. A. Rodriguez-Reveles, S. Esquivel-Hernandez, and V. H. Diaz-Ramirez, “Three-dimensional spatial point computation in fringe projection profilometry,” *Optics and Lasers in Engineering*, vol. 164, p. 107482, 2023.