

# **CORRELACIÓN DE FASE APLICADA A LA NAVEGACIÓN DE VEHÍCULOS AUTÓNOMOS EN ENTORNOS TECHADOS**

## *PHASE CORRELATION APPLIED TO AUTONOMOUS VEHICLE NAVIGATION IN ROOFED ENVIRONMENTS*

***Rosebet Miranda Luna***

Universidad Tecnológica de la Mixteca, México  
*rmiranda@mixteco.utm.mx*

***Raul Cruz Barbosa***

Universidad Tecnológica de la Mixteca, México  
*kruisebar@gmail.com*

***Antonio Orantes Molina***

Universidad Tecnológica de la Mixteca, México  
*tonito@mixteco.utm.mx*

***José Anibal Arias Aguilar***

Universidad Tecnológica de la Mixteca, México  
*anibal@mixteco.utm.mx*

***Alberto Elías Petrilli Barceló***

Tokyo University of Science, Japón  
*petrilliae@gmail.com*

**Recepción:** 6/diciembre/2022

**Aceptación:** 14/marzo/2023

### **Resumen**

En este artículo se considera el problema de la navegación autónoma del vehículo AutoNOMOS-mini-V2 en entornos techados. Este vehículo carece de un sistema de posicionamiento global GPS, por lo que se evalúa la técnica de correlación de fase para estimar la posición con respecto a una secuencia de imágenes indexada y adquirida con la cámara de video rgb del vehículo que apunta hacia arriba, y se determinan las condiciones bajo las cuales la técnica funciona. La metodología propuesta se probó con imágenes de 3 diferentes techos, obteniéndose errores máximos de 5.8 cm y 1° en la estimación de desplazamientos y orientación, respectivamente. Los resultados obtenidos indican que el método propuesto es eficiente para techos con alturas mayor o igual a 3 m.

**Palabras Clave:** Guiado visual, odometría visual, registro de imágenes, visión por computador.

## **Abstract**

*In this article we focus in the autonomous navigation of AutoNOMOS-mini-v2 vehicle in roofed environments. The vehicle does not have a GPS positioning system, so we evaluate the phase correlation technique to estimate its position using an indexed sequence of images acquired with the rgb camera pointing to the roof, we determine the conditions for the technique to work. Our proposed method was tested in images of 3 different roofs, getting maximum errors of 5.8 cm and 1° in the displacement and orientation respectively. Obtained results show that the proposed algorithm is efficient for roofs with heights equal or greater than 3 m.*

**Keywords:** *Computer vision, image registration, visual guidance, visual odometry.*

## **1. Introducción**

Actualmente, un número importante de compañías del ramo automotriz y otras empresas tecnológicas como BMW, Fiat, Chrysler Audi, GM, Ford, Honda, Toyota, Nissan, Suzuki, Mitsubichi, Mazda, Volvo, Renault, Volkswagen, Tesla, Google, Uber, Alphabet, Apple, Nvidia, intel, Microsoft, así como diversas universidades, están involucradas en la investigación y desarrollo de vehículos autónomos para pasajeros y/o transporte de mercancías [Bertozzi, 2011], [Brimbaw, 2015], [Funke, 2012], [Haslehurst, 2017]. El interés por el desarrollo de vehículos completamente autónomos incluye vehículos autoguiados en ambientes interiores, tales como entornos industriales, oficinas, entre otros [Girbés, 2011], [Blanc, 2005], [Xu, 2009]. Aunque algunas compañías y laboratorios ya han realizado pruebas con sus vehículos en autopistas y carreteras [Brimbaw, 2015], [Funke, 2012], [Krisada, 2012], a decir de algunos, los automóviles nivel 5 aún están lejos de estar listos [Moore, 2018]. Los vehículos autónomos incorporan un sistema GPS que les permite obtener información de su entorno, sin embargo, en ciertas circunstancias puede no ser operacional cerca de edificios [Suji, 2019], no estar disponible en los medios en los que se desenvuelve el vehículo [Stephan, 2011], o la precisión

ofrecida por éste puede no ser suficiente [Levinson, 2010]. En el caso de los vehículos autoguiados (AGV) de interior, se ha empleado la detección de marcas ya sea naturales o artificiales colocadas en posiciones estratégicas en pisos, paredes, techos, etc., [Cucchiara, 2007], [Saab, 2011], [Xing, 2014]. Los sistemas basados en marcas ubicadas en techos presentan la ventaja de que generalmente no habrá obstáculos entre el techo y el sistema de visión del vehículo [Xu, 2009].

En el contexto de vehículos y otros sistemas autónomos, la generación de mapas 2D o 3D mediante el registro de una secuencia de imágenes, al igual que el indexado de dicha secuencia son estrategias recurrentes para la localización y navegación del móvil cuando no se cuenta previamente con un mapa, o se quiere prescindir del GPS [Shuji, 2019], [Courbon, 2010], [Dryanovski, 2013], [Kazama, 2017]. Sin embargo, el registro de imágenes suele ser lento y más o menos complejo computacionalmente, dependiendo entre otras cosas de la medida de similitud empleada, del número de parámetros de transformación buscados, de la modalidad de las imágenes, de la información utilizada (marcas o niveles de gris) y del método de optimización implementado, siendo este último generalmente iterativo [Xu, 2009], [Miranda, 2008], [Courbon, 2010], [Huang, 2015]. En estas técnicas, la complejidad se encuentra en los algoritmos mismos, en su implementación y en el pre-procesamiento de las imágenes, además, cuando el registro se basa en marcas, los resultados dependen de la selección e identificación de las mismas en las imágenes. En contraste con estas técnicas, se tiene la correlación de fase que permite estimar desplazamientos entre imágenes 2D basada en el teorema de corrimiento, es de fácil implementación y con tiempos de ejecución pequeños, sin embargo, esta técnica es muy sensible a la información no común que pueda estar presente entre pares de imágenes.

Por otro lado, las Redes Neuronales Profundas están siendo muy utilizadas en la actualidad en muy diversas aplicaciones. En éstas, la complejidad se traslada a la adquisición de imágenes y/o la selección de la base de datos para el entrenamiento, al entrenamiento mismo y a la determinación de la arquitectura más adecuada para la red [Jianbo, 2019], [Lingyan, 2017], [Singandhupe, 2019]. Además, se requiere de GPUs para agilizar el entrenamiento.

En este trabajo se presenta la manera de aplicar la correlación de fase como técnica de registro de imágenes para estimar tanto la posición como la orientación de un vehículo AutoNOMOS mini V2 haciendo uso de imágenes adquiridas con una cámara dirigida hacia arriba. En la generalidad de los casos, los techos pueden presentar textura homogénea, patrones repetitivos o efecto de paralaje, que puede impedir el registro correcto de las imágenes. La principal contribución en este trabajo consiste en mostrar que es posible aplicar la técnica de correlación de fase para realizar el registro de imágenes de techos, adquiridas con la cámara del vehículo AutoNOMOS mini V2 dirigida hacia arriba, y con ello estimar sus parámetros de desplazamiento y rotación, así también se determinan las condiciones bajo las cuales la técnica funciona, sus alcances y sus limitaciones.

## 2. Métodos

En esta sección se presentan las etapas de la metodología, figura 1.

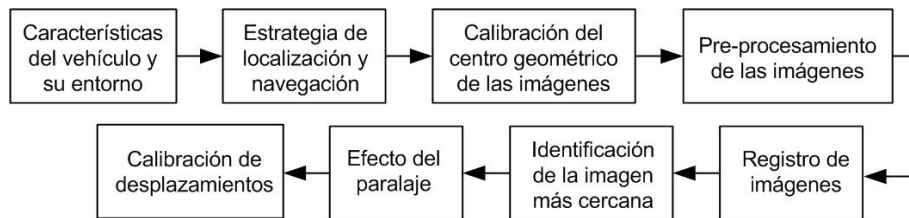


Figura 1 Etapas de la metodología.

### Vehículo AutoNOMOS mini V2 y entornos de navegación

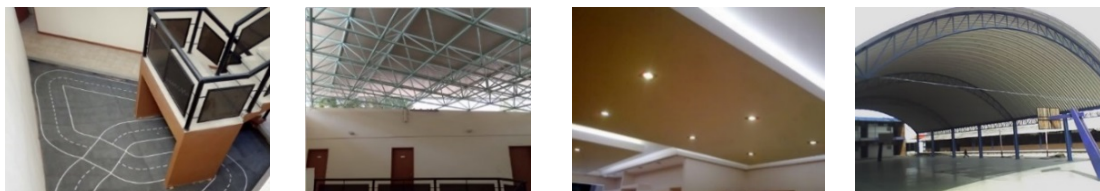
En este trabajo se utiliza un vehículo eléctrico autónomo no tripulado a escala 1:10, denominado AutoNOMOS mini V2. El vehículo cuenta entre otras cosas con una computadora principal Odroid (XU4 64GB), dos cámaras de video, una en la parte frontal (ELP-HD 720P) y otra en la parte superior (ELP 1080), un escáner giratorio, un sistema estereoscópico tipo Kinect (Intel RealSense SR300) y una unidad de medición inercial (IMPU6050). La cámara de video ELP 1080 apunta hacia arriba y tiene un ángulo de apertura de 170° (lente de ojo de pez). Esta cámara permite adquirir imágenes a color a una velocidad de 60 imágenes por segundo. El tamaño de las imágenes es de 640x480 píxeles después de la

corrección de la distorsión radial, figura 2. El vehículo no cuenta con algoritmos de navegación, ya que fue diseñado para promover el desarrollo de algoritmos de percepción, planeación y conducción autónoma en ambientes controlados.



Figura 2 Vehículo AutoNOMOS mini V2.

Los techos bajo los cuales se desplaza el vehículo pueden ser de una gran diversidad de formas y con características particulares, por ejemplo: superficies en diferentes niveles, cóncavas, planas, con o sin inclinación respecto al plano horizontal, con textura homogénea y diseños estructurales tridimensionales y repetitivos. Además, a lo largo de la ruta del vehículo pueden existir transiciones abruptas, por ejemplo, el paso por debajo de una escalera, etc, figura 3.



a) Alto y curvo.

b) Cuasi-homogéneo.

c) Estructura tridimensional.

d) Escalera.

Figura 3 Techos y estructuras de interiores.

### **Estrategia general para la localización y navegación**

La idea general de solución consiste en adquirir con la cámara superior una secuencia de imágenes  $I_n$ , con  $N$  imágenes del techo de un edificio, a lo largo del recorrido que se desea que el vehículo realice. La imagen  $I_1$  se adquiere en la posición inicial del recorrido, mientras que la imagen  $I_N$  es adquirida en la posición final, figura 4. Las imágenes de la secuencia  $I_n$  servirán de referencia durante la navegación autónoma del vehículo.

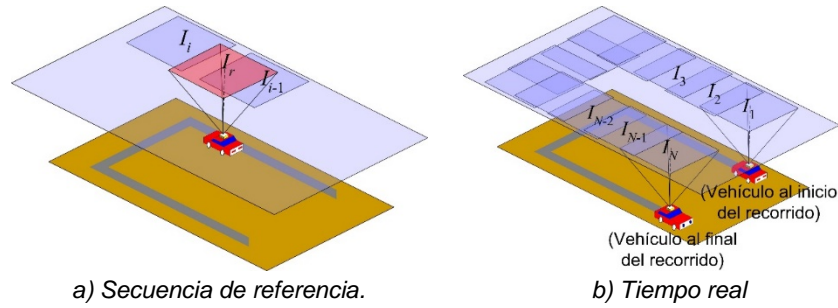


Figura 4 Adquisición de imágenes.

Al iniciar un recorrido de forma autónoma, se adquiere una imagen en tiempo real  $I_r$  y se busca dentro de la secuencia  $I_n$  la imagen más cercana a  $I_r$ , la cual es considerada como la imagen  $I_i$ . Enseguida se determinan las distancias y el ángulo de rotación que la relacionan con  $I_r$ . En caso de que esta imagen  $I_i$  se encuentre por detrás con respecto a la posición desde donde fue adquirida la imagen  $I_r$ , la imagen  $I_{i+1}$  pasa a ser la nueva imagen  $I_i$ . La imagen  $I_i$  es utilizada como referencia para guiar al vehículo a la ubicación desde donde ésta fue adquirida, y una vez ahí, la imagen  $I_{i+1}$  pasa a ser la nueva imagen  $I_i$ . El proceso de registro de imágenes entre  $I_i$  e  $I_r$  y el desplazamiento del vehículo hasta la posición desde donde fue adquirida la imagen  $I_i$  se debe repetir hasta que el vehículo ha llegado al final del recorrido.

El efecto de paralaje puede observarse en imágenes de una misma escena (techo) cuando estos tienen estructuras tridimensionales, o debido a la presencia de muros, cuando las imágenes son adquiridas desde posiciones diferentes. En general, el paralaje representa información útil para diferenciar las imágenes de la secuencia  $I_n$  unas de otras, sobre todo en aquellos casos en los que la textura de los techos es homogénea o estos presenten estructuras repetitivas. Sin embargo, cuando los contenidos de  $I_r$  e  $I_i$  no están relacionados mediante una transformación rígida, la técnica de correlación de fase no permite estimar correctamente los parámetros de desplazamiento que las relacionan. Para resolver este problema, se propone la adquisición de la secuencia de imágenes  $I_n$  de tal manera que las imágenes consecutivas  $I_i$  e  $I_{i+1}$  sean lo suficientemente cercanas. Si la condición anterior se cumple, el método de registro siempre encontrará una imagen  $I_i$  suficientemente

cercana a  $I_r$  con la cuál estimar correctamente los parámetros de desplazamiento y rotación que las relacionan.

### Calibración del centro geométrico de la imagen

Sean  $(O_R, X_R, Y_R, Z_R)$  el sistema de coordenadas de referencia fijo,  $(O_v, X_v, Y_v, Z_v)$  el sistema de coordenadas unido al vehículo justo por debajo de la cámara de video, con sus planos  $X_R Y_R$  y  $X_v Y_v$  paralelos y coincidiendo con el suelo sobre el que se desplaza el vehículo,  $(x_R, y_R)$  las coordenadas de un punto sobre el plano  $X_R Y_R$ , e  $I_a$  e  $I_b$  dos imágenes cualesquiera adquiridas con la cámara. Si  $I_a$  e  $I_b$  son adquiridas desde las mismas coordenadas  $(x_R, y_R)$ , pero con giro arbitrario  $\theta$  alrededor del eje  $Z_v$ , existirá un desplazamiento (en píxeles) entre ambas imágenes. Esto se debe a que el eje óptico de la cámara de video ELP 1080 no es paralelo al eje perpendicular al plano sobre el que se desplaza el vehículo. El desplazamiento estará relacionado con el ángulo de giro de la cámara y la inclinación de su eje óptico, figura 5a. Para que coincidan los centros geométricos de cualesquiera dos imágenes  $I_a$  e  $I_b$ , se siguen los siguientes pasos:

- Se adquieren dos imágenes de calibración  $I_A$  e  $I_B$  con la cámara ubicada en ambos casos en las mismas coordenadas  $(x_R, y_R)$ , y cuya diferencia en rotación alrededor del eje  $Z_v$  es de  $180^\circ$ , figura 5b.
- Se determinan los desplazamientos  $D_x$  y  $D_y$  (en píxeles) de la imagen  $I_B$  con respecto a la imagen  $I_A$ , en las direcciones  $X_i, Y_i$  del plano imagen.
- Los desplazamientos que corrigen el desplazamiento del centro geométrico de las imágenes son  $d_x = -D_x/2$  y  $d_y = -D_y/2$ .

Los desplazamientos  $d_x$  y  $d_y$  se deberán aplicar de manera sistemática a cada imagen adquirida con la cámara, de esta manera dos imágenes cualesquiera  $I_a$  e  $I_b$ , adquiridas con la cámara ubicada en las mismas coordenadas  $(x_R, y_R)$  pero con una rotación arbitraria  $\theta$  alrededor del eje  $Z_v$ , compartirán el mismo centro geométrico. Además, el área en común en píxeles entre las dos imágenes de calibración  $I_A$  e  $I_B$  será la misma sin importar la altura del techo, ya que depende únicamente de la inclinación y del ángulo de apertura de la cámara. Considerando

que la inclinación de la cámara no cambia y dado que el ángulo de apertura es fijo,  $D_x$  y  $D_y$  permanecen constantes, por lo que la calibración se realiza una sola vez.

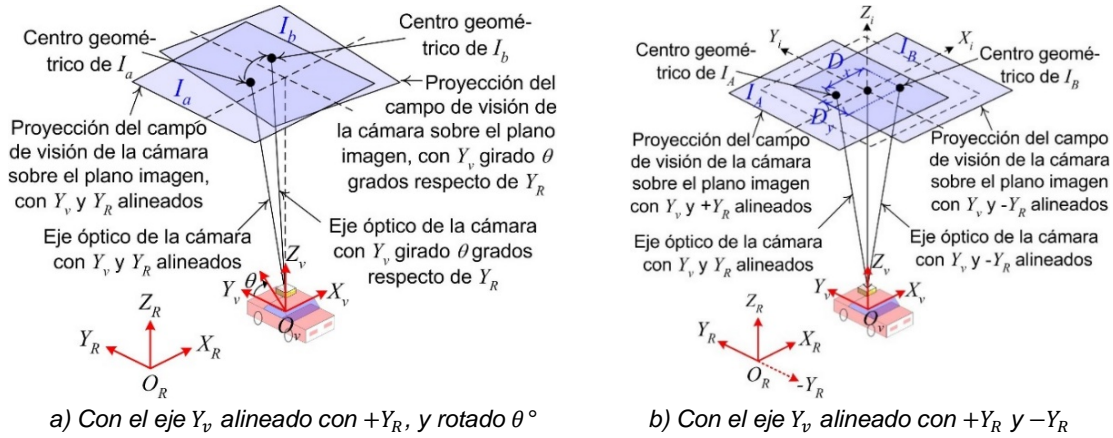


Figura 5 Adquisición de las imágenes.

### Pre-procesamiento de las imágenes

La cámara de video ELP 1080 que apunta hacia arriba genera imágenes con distorsión radial, los coeficientes para su corrección fueron estimados siguiendo el procedimiento y software propuestos en [Open source, 2017]. La corrección de la distorsión se realiza en tiempo real de manera que contamos con imágenes sin distorsión radial de 640x480 píxeles.

Sin embargo, las imágenes obtenidas después de dicha corrección presentan un alargamiento constante en la dirección del eje  $Y$ , de manera que una figura cuadrada con sus lados alineados con los ejes  $(X, Y)$  se verá como un rectángulo. Para corregir esta distorsión, se aplica a cada imagen un factor de escala (alargamiento) de 1.176 en la dirección del eje  $X$ .

Enseguida se aplican a cada imagen los desplazamientos  $d_x$ ,  $d_y$  (descritos en la sección anterior) para corregir su centro geométrico, y se recortan a un tamaño de 480x480 píxeles. El recorte es importante, ya que en imágenes rectangulares y con el contenido de una de ellas rotado respecto al de la otra, una componente de frecuencia en el espectro de la primera imagen y su equivalente en el espectro de la segunda imagen se encontrarán a diferente distancia con respecto a sus respectivos centros de espectro, impidiendo el registro de las imágenes.



## Registro de imágenes

La técnica de correlación de fase para el registro automático de imágenes 2D es capaz de manejar imágenes relacionadas únicamente por traslaciones, sin embargo, esta ha sido extendida para manejar imágenes relacionadas por traslaciones, rotación y escala [Srinivasa, 1996]. Cuando dos imágenes  $I_1(x, y)$  e  $I_2(x, y)$  están relacionadas mediante traslaciones y rotación, se corrige primero la rotación antes de poder estimar la traslación. Para determinar la rotación entre dos imágenes  $I_1(x, y)$  e  $I_2(x, y)$ , se calculan, a partir de las imágenes, los espectros de los logaritmos de sus espectros de magnitud en coordenadas polares. La figura 6 muestra el logaritmo de un espectro de amplitud de una imagen, en coordenadas rectangulares y en coordenadas polares respectivamente.

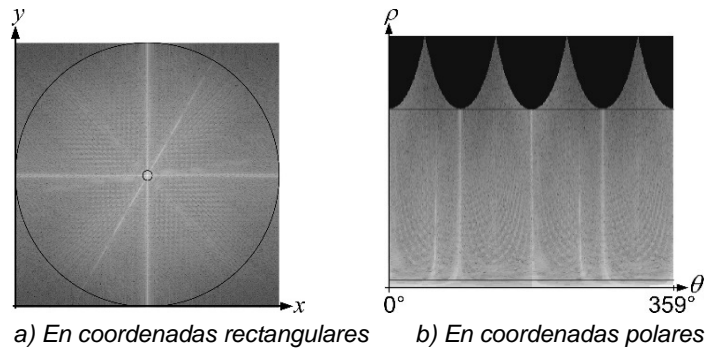


Figura 6 Logaritmo de un espectro de amplitud.

En la figura 6b se pueden observar dos patrones fijos, el primero es una banda brillante en la parte inferior de la imagen, y el segundo, en la parte superior, es un patrón negro en forma de arcos invertidos. Además, debido a que las imágenes  $I_1(x, y)$  e  $I_2(x, y)$  están rotadas entre sí, introducen información no común en las esquinas de sus espectros. Los dos patrones fijos y la información no común de las esquinas interfieren con el registro de las imágenes.

Para calcular la rotación y los desplazamientos, se debe retener únicamente el área comprendida entre las dos rectas horizontales, en el logaritmo del espectro de amplitud en coordenadas polares, figura 6b. Para retener dicha área en la imagen del logaritmo del espectro de amplitud en coordenadas polares se tienen dos opciones: a) se recorta la imagen, b) se mapea a coordenadas polares únicamente

la región de interés comprendida entre los dos círculos indicados en la figura 6a. Además, debido a la simetría del espectro, es suficiente retener una sección correspondiente a  $180^\circ$ . En este trabajo se realizó el mapeo de  $0^\circ$  a  $179^\circ$ .

Dado que entre los logaritmos de los espectros de amplitud en coordenadas polares solo existe desplazamiento en la dirección de  $\theta$  (eje horizontal), el valor máximo en la matriz resultante se busca en el primer renglón (eje  $X$ ), figura 7. La ubicación del valor máximo en la matriz resultante está directamente relacionada con los grados de rotación que relacionan las dos imágenes  $I_1(x, y)$  e  $I_2(x, y)$ . Una vez determinada la rotación, se corrige ésta en la imagen  $I_2(x, y)$ , para estar en condiciones de determinar los desplazamientos entre ambas imágenes.

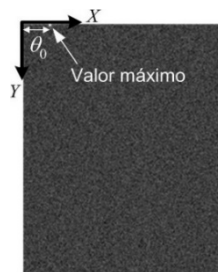


Figura 7 Resultado de la estimación de la rotación.

### Identificación de la imagen más cercana $I_i$

Como resultado del cálculo de la correlación de fase entre dos imágenes, se obtiene una matriz cuyo valor máximo representa la similitud entre ambas imágenes cuando están alineadas. Las coordenadas del valor máximo están directamente relacionadas con los desplazamientos entre las imágenes. Para identificar la imagen  $I_i$  (imagen de la secuencia  $I_n$  más cercana a  $I_r$ ) se calcula la correlación de fase entre la imagen  $I_r$  y cada una de las imágenes de la secuencia  $I_n$ . La hipótesis que se hace es que la matriz con el pico de mayor amplitud corresponde a la calculada entre las imágenes  $I_r$  e  $I_i$ . Una vez identificada  $I_i$ , se determina (únicamente para esta imagen) la ubicación del pico de mayor amplitud en la matriz correspondiente. Los parámetros de transformación que relacionan a ambas imágenes ayudan a determinar si la imagen  $I_i$  se encuentra por delante o por detrás con respecto a la posición desde donde se adquirió la imagen  $I_r$ .

## **Efectos de paralaje**

En la generalidad de casos, en imágenes que serán registradas, puede existir paralaje, además de desplazamientos en el plano  $(X_i, Y_i)$  y rotación alrededor del eje  $Z_i$ . El paralaje se presenta cuando los techos contienen estructuras tridimensionales, o en la periferia de las imágenes debido a los muros.

La técnica de correlación de fase es sensible al paralaje ya que los desplazamientos relativos de los contenidos en la imagen modifican su espectro. En tal situación, los desplazamientos calculados con el método de registro de imágenes no son confiables, lo que puede impedir identificar la imagen  $I_i$ .

El efecto de paralaje que se observa entre dos imágenes se acentúa al incrementarse la distancia desde donde son adquiridas las imágenes. Conforme el efecto de paralaje se incrementa, también lo hace la distorsión de sus espectros. La hipótesis que se hace es que en imágenes que han sido adquiridas desde posiciones muy cercanas, el paralaje no impedirá determinar los desplazamientos entre  $I_r$  e  $I_i$ . Por lo tanto, si las imágenes de referencia vecinas en la secuencia  $I_n$  son adquiridas desde posiciones muy cercanas, la imagen  $I_r$  tendrá siempre una imagen  $I_i$  que fue adquirida desde una posición muy cercana desde donde fue adquirida  $I_r$ . Empíricamente se verificó que, para desplazamientos del vehículo en el orden de los centímetros, el paralaje es despreciable en el sentido de que la técnica de registro de imágenes permite determinar los parámetros de transformación (desplazamientos y rotación) que alinean las imágenes.

## **Calibración de los desplazamientos**

La calibración de la cámara se realizó considerando el modelo simplificado Pin-hole, figura 8. De acuerdo con este modelo se tiene la ecuación 1.

$$f = \frac{Fd}{D} \quad (1)$$

Donde  $f$  es la distancia focal,  $F$  es la distancia del techo al punto focal,  $d$  es el tamaño de la zona sensible del sensor de la cámara en dirección del eje  $x$ , y  $D$  es la distancia máxima en la dirección del eje  $X$  que puede ser contenida en  $d$ .

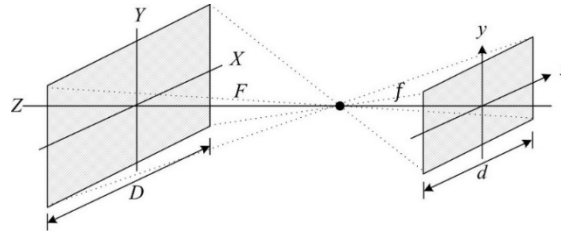


Figura 8 Modelo pin-hole.

Conocidos los valores teóricos de  $f$  y  $d$ , y el valor de  $F$ , se determina  $D$  a partir de (1). La relación pixeles/cm se determina igualando  $D$  con el número de pixeles  $N_p$  de la imagen contenidos en  $d$ . En este caso  $N_p$  es igual a 640 pixeles. La tabla 1 muestran las relaciones pixeles/cm para 3 alturas de techos (10, 3.5 y 1.5 m).

Tabla 1 Calibración para 3 alturas de techo.

Altura de techo (m)	Relación pix/cm
10	1/2.60
3.5	1/0.91
1.5	1/0.39

### 3. Resultados

Las imágenes utilizadas en las pruebas descritas en esta sección se adquirieron con la cámara de video ELP 1080, y se les aplicó el pre-procesamiento descrito en la sección Calibración del centro geométrico de la imagen.

#### Estimación del centro de rotación de la imagen

Se adquirieron las imágenes  $I_A$  e  $I_B$  del techo mostrado en la figura 9a, con la cámara ubicada en las mismas coordenadas  $(x_R, y_R)$ , pero con una diferencia de  $180^\circ$ . Desde el punto de vista de la cámara (viendo al techo), se obtuvieron  $D_x = 69$  y  $D_y = 23$  píxeles respectivamente entre las imágenes  $I_A$  e  $I_B$ . Este resultado implica que desde la perspectiva de la cámara, su eje óptico cruza el plano imagen en el primer cuadrante  $(x = 69/2, y = 23/2)$  cuando está en dirección hacia adelante, y lo cruza en el tercer cuadrante  $(x = -69/2, y = -23/2)$  cuando está girado  $180^\circ$ . De esta manera  $d_x = 34.5$  y  $d_y = 11.5$  píxeles respectivamente.

## Estimación de desplazamientos

Por cada techo mostrado en la figura 9 se adquirieron 3 secuencias de 20 imágenes cada una, de 480x480 píxeles, para ello se colocó el vehículo bajo el techo en turno en una posición inicial diferente para cada secuencia, y se desplazó manualmente en línea recta distancias (entre adquisiciones) de 15, 10 y 5 cm respectivamente en cada secuencia. Para cada secuencia se realizó el registro entre imágenes consecutivas.

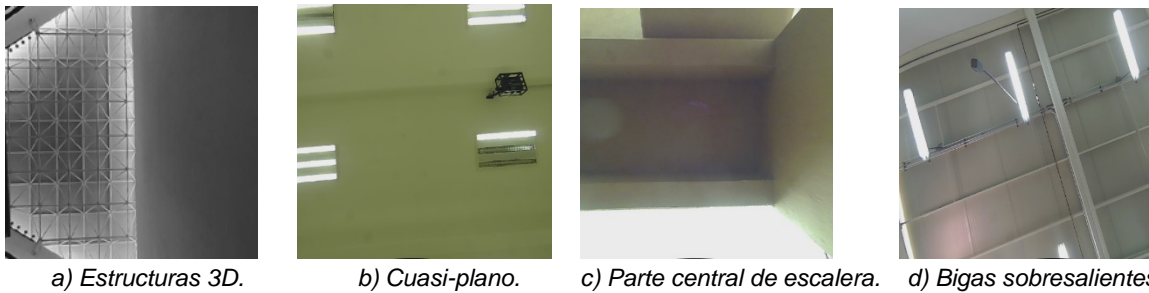


Figura 9 Techos con diferentes alturas.

Las tablas 2 a 4 muestran la media  $\mu$ , desviación estándar  $\sigma$ , y los errores promedio  $e_{prom}$  y máximo  $e_{max}$ , obtenidos para las tres secuencias de cada techo. En la tabla 5 se indican los porcentajes de imágenes para las cuales el algoritmo fue capaz de determinar desplazamientos y rotación.

Tabla 2 Resultados para el techo 1.

	Secuencia 1		Secuencia 2		Secuencia 3	
	pix	cm	Pix	cm	pix	cm
$\mu$	5.83	15.16	3.84	9.98	1.9	4.94
$\sigma$	1.2	3.12	0.98	2.55	0.94	2.44
$e_{prom}$	1.03	2.68	0.79	2.05	0.72	1.87
$e_{max}$	2.23	5.79	2.15	5.59	2.07	5.38

Tabla 3 Resultados para el techo 2.

	Secuencia 1		Secuencia 2		Secuencia 3	
	pix	cm	Pix	cm	pix	cm
$\mu$	16.58	15.1	11.67	10.62	5.13	4.66
$\sigma$	1.34	1.22	1.59	1.44	2.02	1.83
$e_{prom}$	1.28	1.16	1.00	0.91	1.48	1.34
$e_{max}$	1.51	1.37	4.00	3.64	4.49	4.08

Tabla 4 Resultados para el techo 3.

	Secuencia 1		Secuencia 2		Secuencia 3	
	pix	cm	Pix	cm	pix	cm
$\mu$	40.85	15.9	27.22	10.06	14.25	5.5
$\sigma$	1.46	0.57	1.14	0.42	2.19	0.85
$e_{prom}$	2.38	0.93	1.57	0.58	1.86	0.72
$e_{max}$	4.53	1.76	3.35	1.23	7.15	2.76

Tabla 5 Porcentaje de imágenes registradas correctamente.

Techo	Secuencia 1	Secuencia 2	Secuencia 3
1	100 %	100 %	100 %
2	67 %	73 %	95 %
3	39 %	47 %	60 %

### Rotación

Para cada techo mostrado en figuras 9a y 9b, se colocó el vehículo en las posiciones indicadas en la figura 10.  $A_0$  es la posición de referencia con el eje  $Y_v$  alineado con el eje  $Y_R$ . En las posiciones  $A_1$ - $A_4$ , el vehículo se colocó a diferentes ángulos entre  $-45^\circ$  y  $45^\circ$  a intervalos de  $22.5^\circ$  con respecto al eje  $Y_R$ . Se adquirieron 21 imágenes de cada techo, y se registraron las 20 imágenes adquiridas en las posiciones  $A_1$ - $A_4$  con la correspondiente imagen de referencia adquirida en la posición  $A_0$ . La tabla 6 muestra los resultados obtenidos. El cálculo del error promedio en la estimación de la rotación  $e_{promR}$  se realizó promediando el valor absoluto de las diferencias entre la estimación y la posición real.

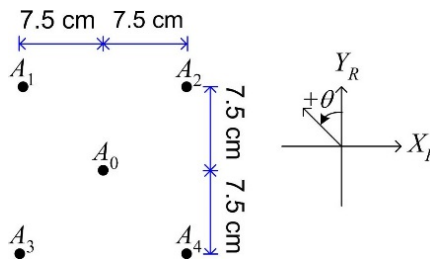


Figura 10 Colocación del vehículo para las pruebas de rotación.

Tabla 6 Resultados en la estimación de la rotación.

$\sigma_R$	$0.67^\circ$
$e_{promR}$	$0.43^\circ$
$e_{maxR}$	$1^\circ$

### Identificación de la imagen $I_i$

De cada techo mostrado en figuras 9a y 9b (techos 1 y 2 respectivamente), se adquirieron dos secuencias de 55 imágenes cada una, una secuencia a intervalos de desplazamiento de 10 cm (secuencia A) y la otra a intervalos de 5 cm (secuencia B), en trayectorias rectas. Cada imagen se registró con las restantes 54 imágenes de la respectiva secuencia, identificándose para cada caso la imagen con máximo valor en la matriz de la correlación de fase. La tabla 7 muestra el número de veces que la medida de similitud fue mayor entre imágenes vecinas para cada secuencia y para cada techo, y los porcentajes correspondientes. Solamente en dos ocasiones (3.6%) no fue posible identificar correctamente la imagen vecina (imagen  $I_i$ ).

Tabla 7 Resultado del registro de imágenes en la identificación de la imagen  $I_i$ .

Techo	Secuencia A		Secuencia B	
	Aciertos	Porcentaje	Aciertos	Porcentaje
1	53	96.4 %	54	98.2 %
2	52	94.5 %	53	96.4 %

### Pruebas en 3 escenarios

Se realizaron pruebas en 3 escenarios diferentes, y para cada uno se adquirió una secuencia de imágenes separadas por una distancia de 5 cm entre ellas, desplazando manualmente el vehículo a lo largo de cada recorrido deseado. Estas imágenes son las imágenes de referencia.

El primer escenario corresponde a una pista a escala de doble carril, cada uno de 40 cm de ancho que conforma un circuito cerrado, diseñada para el vehículo AutoNOMOS mini V2, figura 3d.

El techo tiene una estructura tridimensional repetitiva y su altura aproximada es de 10 m, figura 9a. La pista atraviesa por debajo de una escalera con altura de 1.5 m en su parte media, figura 9c. En este caso se adquirieron 215 imágenes de referencia a lo largo del circuito, enseguida se colocó el vehículo en la línea de salida y se le hizo avanzar de forma remota, a través del circuito mientras se iban adquiriendo imágenes de prueba  $I_r$ . En total se adquirieron 179 imágenes de prueba en la zona del techo con altura de 10 m, y para cada una de ellas el algoritmo fue capaz de determinar la ubicación del vehículo con respecto a la imagen de

referencia  $I_i$  más cercana. En la zona que atraviesa por debajo de la escalera el algoritmo no fue capaz de determinar la ubicación del vehículo con respecto a la imagen de referencia  $I_i$  más cercana.

El segundo y tercer escenario corresponden a recorridos por debajo de dos techos predominantemente planos con alturas de 3.5 m, figuras 9b y 9d, y en ambos casos se adquirieron 80 imágenes de referencia. Los desplazamientos y rotaciones estimados por el algoritmo permitieron en ambos casos conocer la ubicación del vehículo con respecto a las imágenes de referencia.

#### **4. Discusión**

Se pudo validar empíricamente la validez de la estrategia utilizada para reducir la influencia del paralaje en el registro de las imágenes, ya que fue posible determinar desplazamientos y rotación con una precisión aceptable para la aplicación. Igualmente se pudo verificar la hipótesis de que el valor máximo de la medida de similitud calculada con el método de correlación de fase es máximo para imágenes consecutivas, por lo que se puede utilizar para determinar la imagen más cercana. De las pruebas realizadas en los tres escenarios se observa lo siguiente: Para reducir la influencia del paralaje lo suficiente para que permita determinar los parámetros de desplazamiento y rotación entre dos imágenes, las imágenes deben contener una gran área en común, (cercana al 99%). Esto quiere decir por ejemplo que para techos con altura de 3.5 m, el desplazamiento del vehículo debe ser  $\leq 5$  cm entre la imagen adquirida en tiempo real y la imagen de la secuencia de referencia más cercana, y para techos con altura de 10 m, este desplazamiento puede ser de hasta 15 cm. Para techos con altura menor a 3.5 m se requiere que la separación entre las adquisiciones de las imágenes sea menor a 5 cm. El error máximo (de  $\pm 1$  grado) en la estimación de la rotación implica que al desplazarse entre dos imágenes de la secuencia de referencia, el vehículo se desviará a la izquierda o a la derecha menos de 0.5 cm con respecto a la posición de referencia (imagen  $I_i$ ) por lo que el error se considera igualmente despreciable para esta aplicación.



## 5. Conclusiones

El guiado visual de vehículos autónomos basado en una cámara de video dirigida hacia arriba constituye una alternativa viable en entornos techado.

A pesar de que el método de registro de imágenes es sensible al paralaje, es posible su utilización en el guiado de vehículos autónomos si se considera una secuencia de imágenes de referencia con poco desplazamiento entre imágenes consecutivas. La precisión alcanzada por el método de registro de imágenes en la estimación de la posición para techos con altura de 3.5 m o superior, y desplazamientos de 5 cm entre adquisiciones de las imágenes de la secuencia de referencia, es satisfactoria para la aplicación analizada ya que, de acuerdo a las pruebas realizadas, con un error máximo menor a 5.8 cm, el método permite determinar si el vehículo se encuentra o no en su carril.

Para techos con alturas inferiores a 3.5 m, la técnica de registro de imágenes propuesta no es apropiada, ya que el porcentaje de imágenes registradas correctamente es menor o igual a 60%.

El rango de rotación que permite determinar el método de registro implementado es apropiado para su utilización en rutas con curvas cerradas.

## 6. Bibliografía y Referencias

- [1] Bertozzi, M., et al. Viac: An out of ordinary experiment, 2011 IEEE Intelligent Vehicles Symposium (IV). Baden-Baden, Germany. Junio, 2011.
- [2] Bimraw, K. Autonomous Cars: Past, Present and Future - A Review of the Developments in the Last Century, the Present Scenario and the Expected Future of Autonomous Vehicle Technology. Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO-2015). Colmar, Francia. Enero, 2015.
- [3] Blanc, G., Mezouar, Y., & Martinet, P. Indoor Navigation of a Wheeled Mobile Robot along Visual Routes. Proceedings of the 2005 IEEE International Conference on Robotics and Automation. España. Abril, 2005.
- [4] Courbon, J., et al. Vision-based navigation of unmanned aerial vehicles. ELSEVIER, Control Engineering Practice, Vol. 18, No. 7, 789–799, 2010.

- [5] Cucchiara, R., Perini E., & Pistoni G. Efficient stereo vision for obstacle detection and AGV Navigation. 14<sup>th</sup> International conference on Image Analysis and Processing – ICIAP 2007.
- [6] Dryanovski, I., Valenti, R. G., & Xiao, J. An open-source navigation system for micro aerial vehicles. *Auton Robot*, Vol 34, No. 3, 177–188, 2013.
- [7] Funke, J., et al. Up to the limits: Autonomous Audi TTS. Comunicación presentada en 2012 IEEE Intelligent Vehicles Symposium (IV). Alcalá de Henares, España. Junio, 2012.
- [8] Girvés, V., Armesto, L., Tornero, J., & Solanes, J. E. Continuous-curvature kinematic control for path following problems. IEEE RSJ International Conference on Intelligent Robots and Systems. San Francisco CA. Septiembre, 2011.
- [9] Haslehurst, R., Lewis, A., & Moran, J. Mapping the Road to Autonomous Vehicles, L.E.K. Consulting / Executive Insights, Vol. XIX, No. 58, 2017.
- [10] Huang, Y., et al. Image-Based Localization for Indoor Environment Using Mobile Phone. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-4(W5), 211-215, 2015.
- [11] Jianbo, Z. J., et al. Semantic Segmentation based Dense RGB-D SLAM in Dynamic Environments. *Journal of Physics: Conference Series*, No. 1267, 1-7, 2019.
- [12] Kazama K., et al. Vehicle position estimation based on road area detection and image registration. *Transactions of the JSME*, Vol. 83, No. 849, 1-12, 2017.
- [13] Krisada, K., & Gerdes, J. C. Autonomous vehicle control at the limits of handling. *International Journal of Vehicle Autonomous Systems*, Vol. 10, No. 4, 271-296, 2012.
- [14] Levinson, J., & Thrun, S. Robust vehicle localization in urban environments using probabilistic maps, 2010 IEEE International Conference on Robotics and Automation. Anchorage, AK, USA. Mayo, 2010.
- [15] Moore, A. J., et al. Testing Enabling Technologies for Safe UAS Urban Operations. AIAA AVIATION Forum, Atlanta, GA. Junio, 2018.

- [16] Lingyan, R. L., et al. Convolutional Neural Network-Based Robot Navigation Using Uncalibrated Spherical Images. *Sensors*, Vol. 17, No. 1341, 1-18, 2017.
- [17] Miranda, L. R., et al. Mosaicing of Bladder Endoscopic Image Sequences: Distortion Calibration and Registration Algorithm. *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 2, 541-553, 2008.
- [18] Ojansivu, V., & Heikkilä, J. Image Registration Using Blur-Invariant Phase Correlation. *IEEE Signal Processing Letters*, Vol. 14, No. 7, 449-452, 2007.
- [19] Open source robotic foundation. *Camera calibration*. Recuperado el 05 de septiembre de 2017 de [http://wiki.ros.org/camera\\_calibration](http://wiki.ros.org/camera_calibration).
- [20] Saab, S. S., & Nakad, Z. S. A Standalone RFID Indoor Positioning System Using Passive Tags. *IEEE Transactions on Industrial Electronics*, Vol. 58, No. 5, 1961-1970, 2011.
- [21] Shuji, O. S., et al. SeqSLAM++: View-based robot localization and navigation. *Robotics and Autonomous Systems*, No. 112, 13–21, 2019.
- [22] Singandhupe, A., & La, H. M. A Review of SLAM Techniques and Security in Autonomous Driving. *IEEE International Conference on Robotic Computing (IRC)*. Naples, Italia. Febrero, 2019.
- [23] Srinivasa, R. B., & Chatterji, B. N. An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration. *IEEE Transactions on Image Processing*, Vol. 5, No. 8, 1266-1271, 1996.
- [24] Stephan, W. S., Scaramuzza, D., & Siegwar, R. Monocular-SLAM-based navigation for autonomous micro helicopters in GPS-denied environments. *Journal of field robotics*, Vol. 28, No. 6, 854-874, 2011.
- [25] Xing, W., Peihuang, L., Jun, Y., Xiaoming, Q., & Dunbing, T. Intersection Recognition and Guide-Path Selection for a Vision-based AGV in a Bidirectional Flow Network. *International Journal of advanced Robotic Systems*, Vol. 11, No. 3, 1-17, 2014.
- [26] Xu, D., Han, L., Tan, M., & Li, F. Ceiling-Based Visual Positioning for an indoor Mobile Robot With Monocular Vision. *IEEE Transactions on Industrial Electronics*, Vol. 56, No. 5, 1617-1628, mayo, 2009.