



Universidad Nacional de Quilmes

Simulación de un sistema basado en cámaras Fisheye y PTZ para la toma de fotografías de objetos en movimiento.

Hernán Guimaraynz.

Director:
Dr. Damián Oliva.

Índice

1. Introducción.	3
2. Materiales utilizados.	3
2.1. Cámara Fisheye con visión de campo amplio (VCA).	3
2.2. Cámara Pan-Tilt-Zoom.	3
3. Simulación de un sistema de adquisición con una cámara (Pan-Tilt-Zoom) PTZ. Calibración y estimación de posición.	4
3.1. Modelado de la cámara PTZ.	4
3.1.1. Modelo de cámara pinhole.	5
3.2. Calibración de la cámara PTZ.	7
3.2.1. Calibración intrínseca de la cámara PTZ.	7
3.2.2. Simulación de los parámetros intrínsecos.	9
3.2.3. Incerteza de los parámetros intrínsecos.	9
3.2.4. Calibración extrínseca de la cámara PTZ.	9
3.2.5. Simulación de los parámetros extrínsecos.	10
3.2.6. Incerteza de los parámetros extrínsecos.	11
3.3. Simulación del nivel de zoom.	11
3.4. Simulación del proceso de calibración. Minimización del error de ajuste sobre el mundo.	13
3.4.1. Incerteza de los parámetros extrínsecos usando la técnica de bootstrap.	15
3.5. Simulación de cámara PTZ.	15
4. Simulación de un sistema de adquisición con una cámara fisheye (FE). Calibración y estimación de posición.	16
4.1. Modelado de la cámara fisheye.	16
4.1.1. Modelo de proyección.	16
4.2. Calibración de la cámara fisheye.	17
4.2.1. Calibración intrínseca de la cámara fisheye.	17
4.2.2. Simulación de los parámetros intrínsecos.	18
4.2.3. Incerteza de los parámetros intrínsecos.	19
4.2.4. Calibración extrínseca de la cámara fisheye.	19
4.2.5. Simulación de los parámetros extrínsecos.	19
4.2.6. Incerteza de los parámetros extrínsecos.	20
4.3. Simulación del proceso de calibración. Minimización del error de ajuste sobre el mundo.	20
4.3.1. Incerteza de los parámetros extrínsecos usando la técnica de bootstrap.	22
4.4. Simulación de cámara FE.	22
5. Simulación de los delays de movimiento y de comunicación para las cámaras PTZ y FE.	22
5.1. Retardos en la captura.	22
5.1.1. Retardo en la captura de la cámara PTZ.	22
5.1.2. Retardo en la captura de la cámara fisheye.	22
5.1.3. Diferencia de retardo en la captura de ambas cámaras.	22

5.1.4. Frecuencia de capturas de las cámaras.	22
5.2. Modelado del retardo de movimiento de la cámara PTZ.	22
6. Simulación del sistema PTZ+FE.	24
6.1. Cinemática de apunte de la cámara PTZ.	25
6.1.1. Calibración de la base de la cámara PTZ.	25
6.1.2. Cálculo de los ángulos de apunte.	26
6.1.3. Determinación del nivel de zoom.	27
7. Métodos para la estimación de velocidad y trayectorias de los vehículos con la cámara PTZ y FE.	29
7.1. Modelos del movimiento de los vehículos en el mundo.	29
7.2. Estimación de velocidad con Filtro Kalman, Cuadrados mínimos recursivos y Filtro de partículas.	29
8. Análisis de estrategias de posicionamiento de la cámara PTZ en función del posicionamiento inferido con la cámara FE.	29
8.1. Diseño de una estrategia de apunte de la PTZ intersectar el vehículo y para capturar una imagen de la patente.	29
8.2. Diseño de una estrategia para optimizar el tiempo en el que la cámara graba el video.	29
9. Bibliografía.	29

1. Introducción.

2. Materiales utilizados.

Debido a que este trabajo tiene como objetivo diseñar un simulador de un sistema de transporte real, las características y parámetros de las cámaras a simular fueron obtenidos a través de experimentos con cámaras reales, para dichos experimentos y obtención de datos se utilizaron las siguientes cámaras:

2.1. Cámara Fisheye con visión de campo amplio (VCA).

Cámara VCA de la marca VIVOTEK modelo FE8172, con las siguientes características:

- Lente de tipo ojo de pez con campo visual de $360^\circ \times 183^\circ$.
- Tasa máxima de adquisición de fotogramas por segundo de 12 fps con una resolución máxima de 1920×1920 píxeles.



Figura 1: Cámara VIVOTEK FE8172.

2.2. Cámara Pan-Tilt-Zoom.

Para apuntar a un objetivo específico y verlo en detalle se utilizó una cámara domo de la marca Hecker (representante de Hikvision en Argentina) modelo HE-Z9116, que tiene las siguientes características:

- Rango de alcance de 360° en el plano horizontal, 180° en el vertical y zoom óptico de 30x.
- Tasa máxima de adquisición de datos de 25 fps con una resolución de 1280×960 píxeles.
- Codificación de video con el protocolo H.264.
- Modo de enfoque automático, semiautomático y manual.

Cabe aclarar que el movimiento del equipo en el plano vertical no se produce en todo el rango de alcance, sino de 0° a 90° . Para llegar al otro hemisferio se necesita rotar 180° en el plano horizontal relativos a la posición actual, desde donde se permite mover los 90° restantes.



Figura 2: Cámara Hecker HE-Z9116.

3. Simulación de un sistema de adquisición con una cámara (Pan-Tilt-Zoom) PTZ. Calibración y estimación de posición.

Se realizará un programa en Matlab para implementar el modelo intrínseco y extrínseco de la cámara PTZ. El nivel de zoom se modelará como un cambio de distancia focal en el modelo pin-hole. Los resultados obtenidos se comparan con el toolbox de Corke [2] y con mediciones experimentales en la cámara real. Para lograr dicho objetivo, se utilizará el *RoboticsToolbox* de *MATLAB*, el cual contiene funciones y clases para representar la orientación y pose en 2D y 3D de una cámara, pudiendo realizar el modelado del funcionamiento de ambas cámaras simulando sus parámetros.

El método de simulación consiste en generar un conjunto de puntos que simulen una intersección de calles y uno o varios automóviles, con posiciones conocidas, para luego obtener una predicción de los píxeles en los que se va a encontrar en las imágenes capturadas por las cámaras FE y PTZ.

Una vez logrado esto, se desarrolla un sistema de apuntado para la PTZ tal que ésta realice los movimientos de pan tilt y zoom necesarios para que un punto marcado por un operario en la imagen de la cámara FE aparezca en el centro de la imagen de la PTZ. De esta manera se pretende obtener una imagen en alta resolución de un automóvil, que el operario considere en una situación sospechosa o en infracción interceptando dicho automóvil con la cámara PTZ. Para interceptar al automóvil con la cámara PTZ se deberá estimar y predecir la trayectoria y velocidad de éste, para lo cual se utilizará un filtro de Kalman.

3.1. Modelado de la cámara PTZ.

Los parámetros que describen el funcionamiento de una cámara se denominan intrínsecos, éstos parámetros convierten las coordenadas en el marco de referencia del dispositivo a las coordenadas en píxeles de la imagen, dichos parámetros son; la distancia focal y el punto principal.

A su vez, son los parámetros extrínsecos los que definen la posición y la orientación del marco de referencia de la cámara con respecto al mundo se denominan y están determinados por una rotación y una traslación en dicho sistema de coordenadas.

Para modelar el comportamiento de la cámara PTZ , se opta por el modelo pinhole , el cual supone ciertas simplificaciones en el proceso de formación de la imagen, y también se desprecian los efectos de distorsión y desenfoque [2].

3.1.1. Modelo de cámara pinhole.

El modelo simplificado de cámara pinhole describe la relación matemática entre las coordenadas de un punto (X, Y, Z) en el mundo y su proyección en el plano de la imagen (u, v) .

Las simplificaciones que se realizan son:

- Considera la apertura de la cámara como un punto infinitesimal, despreciando así la distorsión y el efecto del foco que puede ocasionar el lente en la imagen.
- El plano de la imagen está a distancia. f en frente del origen de la cámara, y se forma una imagen no invertida sobre el eje Z , el cual determina el centro del campo de visión, (en la geometría del sistema original la imagen en el plano queda invertida, pero esta simplificación es matemáticamente equivalente [2]).

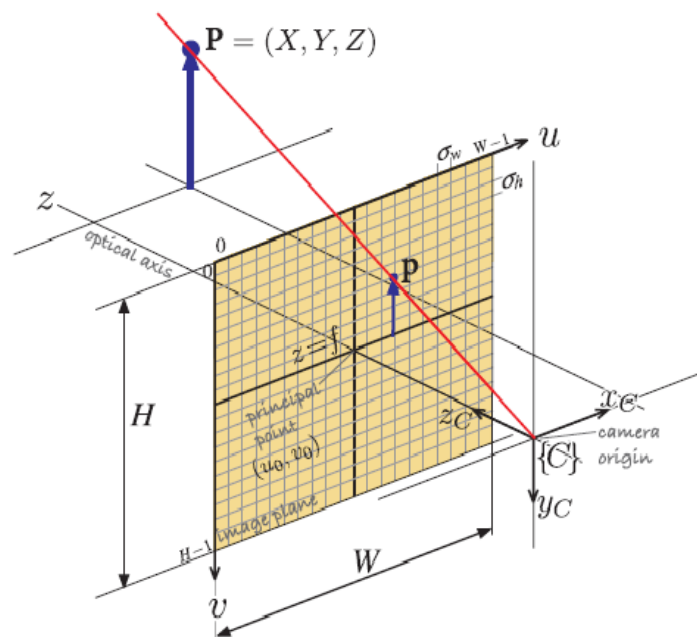


Figura 3: Esquema del modelo pin-hole.

La relación que establece el modelo entre un punto en coordenadas del mundo (X, Y, Z) y las coordenadas (u, v) de su proyección en la imagen es la siguiente:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t \quad (1)$$

Donde R es la matriz de rotación del marco de referencia del mundo al de la cámara y t el vector de traslación también rotado que trasladan el punto (X, Y, Z) al sistema de referencia de la cámara (x, y, z) .

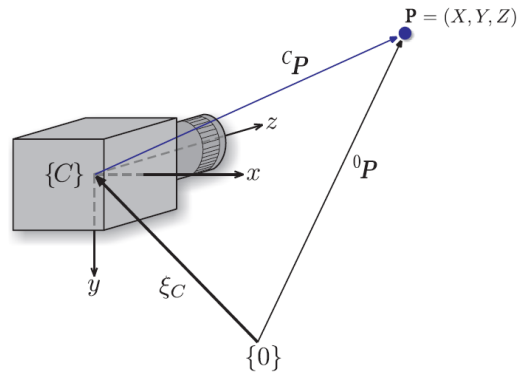


Figura 4: Marcos de referencia de la cámara.

Luego, las ecuaciones que describen la proyección de un punto con coordenadas en el mundo $P = (X, Y, Z)$ al punto de imagen $p = (u, v)$ son:

$$\begin{aligned} u &= \frac{x}{\rho_w} + u_0 \\ v &= \frac{y}{\rho_h} + v_0 \end{aligned}$$

Donde :

(u, v) son las coordenadas sobre la imagen medidas en píxeles.

(ρ_w, ρ_h) son el ancho y la altura de cada píxel.

(u_0, v_0) son las coordenadas del punto principal, el punto en el que el eje óptico interseca el plano imagen.

El campo de visión puede determinarse a partir de la geometría de la Figura (6) de la siguiente manera:

$$\theta_x = 2 \tan^{-1} \left(\frac{W/2\rho_w}{f} \right) \quad (2)$$

$$\theta_y = 2 \tan^{-1} \left(\frac{H/2\rho_h}{f} \right) \quad (3)$$

Donde:

W y H son el número de píxeles en la posición horizontal y vertical de la imagen.

3.2. Calibración de la cámara PTZ.

3.2.1. Calibración intrínseca de la cámara PTZ.

Los parámetros intrínsecos de la cámara se obtienen a partir de la relación que existe entre las coordenadas 3D de puntos en el mundo y los píxeles asociados a ellos en una imagen. El repositorio de la librería *OpenCV* de *Python*, provee un set de funciones capaces de llevar a cabo la calibración utilizando un patrón similar a un tablero de ajedrez [3].

La calibración intrínseca de la cámara PTZ HECKER se realizó en un trabajo final de carrera anterior [[1]] , por lo que para las simulaciones realizadas en este trabajo se utilizan los datos obtenidos en el citado trabajo.

El procedimiento de calibración intrínseca utilizado fue el siguiente:

1. Capturar 10 imágenes o más del tablero desde diferentes ángulos.
2. Encontrar automáticamente las esquinas interiores del tablero en cada imagen mediante la función `findChessboardCorners`.
3. Asignarle a cada punto una posición en el mundo, con $z = 0$ por encontrarse todos en el mismo plano.
4. Establecer la relación entre los píxeles de las esquinas detectadas y su posición en el mundo mediante la función `calibrateCamera`, obteniendo como resultado la matriz con los parámetros intrínsecos de la cámara.

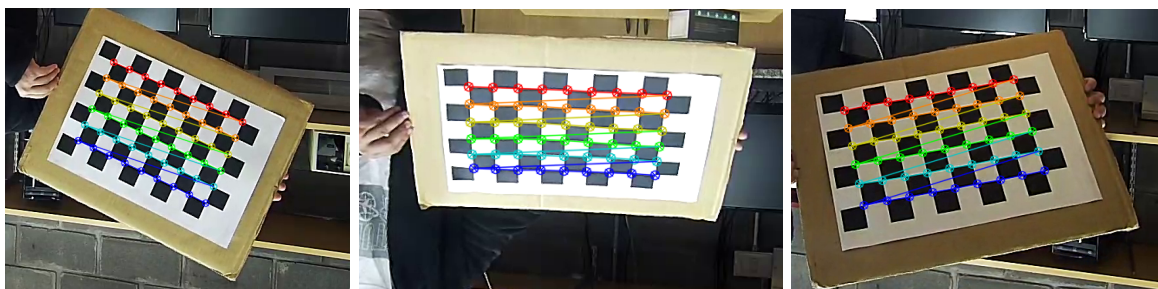


Figura 5: Capturas del tablero desde diferentes ángulos y los puntos de interés del paso 2 del procedimiento de calibración detectados.

Como la cámara posee la variable del zoom, este algoritmo se repite para cada posible nivel de zoom.

Los parámetros intrínsecos que se obtienen con el algoritmo de calibración son:

- f_x : Distancia focal en x expresada en píxeles.
- f_y : Distancia focal en y expresada en píxeles.
- u_0 : Coordenada del centro de la imagen en x expresada en píxeles.
- v_0 : Coordenada del centro de la imagen en y expresada en píxeles.

<i>zoom</i> [%]	<i>f_x</i> [píxel]	<i>f_y</i> [píxel]	<i>fov_x</i> [grados]	<i>fov_y</i> [grados]
0	1246.33	1246.33	54.36	42.12
10	33244.96	3244.96	22.31	16.82
20	7119.92	7119.92	10.27	7.71
30	10004.30	10004.30	7.32	5.49
40	13151.33	13151.33	5.57	4.18
50	16478.57	16478.57	4.45	3.34
60	18598.74	18598.74	3.94	2.95
70	21702.63	21702.63	3.38	2.53
80	27223.82	27223.82	2.69	2.02
90	27707.66	27707.66	2.64	1.98
100	26728.01	26728.01	2.74	2.05

Figura 6: Valores de distancia focal y campo de visión de la cámara PTZ para diferentes valores de zoom.

A partir de los datos de la calibración intrínseca, se define la función que mapea para cada posible nivel de zoom de la cámara PTZ, su correspondiente distancia focal (medida en píxeles).

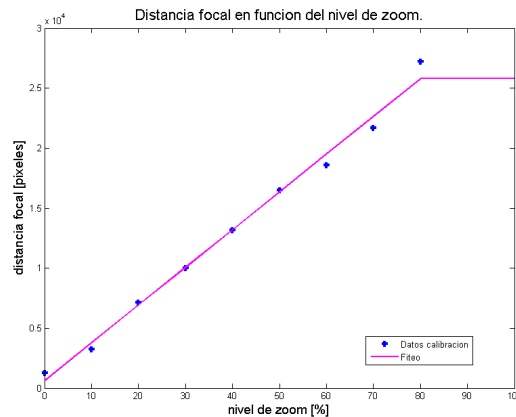


Figura 7: Función de la distancia focal para cada nivel de zoom, fiteada a partir de los datos obtenidos en la calibración intrínseca de la cámara PTZ..(*fiteo_zoom.m*)

3.2.2. Simulación de los parámetros intrínsecos.

Una vez obtenidos los parámetros intrínsecos, se procede a crear el simulador para la cámara mediante el *RoboticsToolbox* de Peter Corke, [2] para *MATLAB*, se crea un objeto cámara de la clase *CentralCamera*, el cual al setearle dichos parámetros simula el comportamiento de la cámara PTZ.

```
% CAMARA PTZ
global camptz
camptz = CentralCamera('name', 'PTZ', 'focal', fptz, ...
    'resolution', [Npixptzw Npixptzh], ...
    'centre', [Npixptzw/2 Npixptzh/2], ...
    'noise', 0.00, ...
    'focal', fptz);
```

Figura 8: Definición del objeto *CentralCamera* mediante el *RoboticsToolbox* de *MATLAB*, seteada con los parámetros intrínsecos calculados para la cámara PTZ.

3.2.3. Incerteza de los parámetros intrínsecos.

3.2.4. Calibración extrínseca de la cámara PTZ.

La calibración extrínseca consiste en hallar los parámetros que determinan la posición de la cámara con respecto al mundo, están descriptos por la matriz de rotación R y el vector de traslación t , vistos en la ecuación 1.

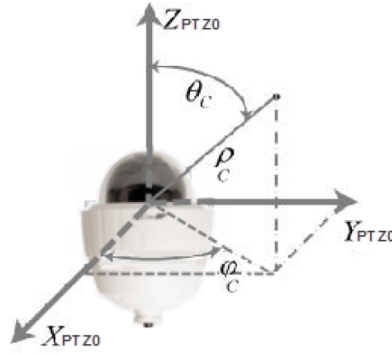


Figura 9: Coordenadas esféricas de la cámara PTZ.

Los parámetros extrínsecos de la cámara PTZ son:

- $(x_{PTZ0}, y_{PTZ0}, z_{PTZ0})$: Posición de la base de la cámara en el mundo.
- $(\alpha_{PTZ0}, \beta_{PTZ0}, \gamma_{PTZ0})$:: Ángulos de rotación de la base de la cámara con respecto a los ejes Z_m , Y_m y X_m del mundo.
- $(\phi_{PTZ}, \theta_{PTZ})$: Valores de tilt y pan respecto a la base de la cámara PTZ.

3.2.5. Simulación de los parámetros extrínsecos.

Para simular estos parámetros en el simulador ,se calcula la proyección que lleva los puntos de coordenadas del mundo a coordenadas respecto a la base de la cámara,teniendo en cuenta la particularidad de que la cámara PTZ posee una horquilla sobre la cual el sensor CCD puede girar sobre la base de la cámara mediante dos mecanismos,los cuales permiten dos tipos de movimientos que reciben los nombres de $PAN(\phi)$ (movimiento horizontal respecto al eje Z) y $TILT(\theta)$ (movimiento vertical respecto al eje Z).

Por lo que la traza de la base de la PTZ respecto al mundo es:

$$T_{PTZ} = T_{XYZ} R_{XYZ} R(\phi_{PTZ}, \theta_{PTZ})$$

Donde:

- R_{XYZ} : Rotación de la base de la cámara respecto al mundo.
- $R(\phi_{PTZ}, \theta_{PTZ})$:Rotación del sensor CCD debido al mecanismo pan-tilt Z-X.
- T_{XYZ} :Traslación de la base de la cámara al mundo.

Primero se determinan los parámetros extrínsecos en el simulador, es decir, se determina la pose de la cámara que se va a simular:

```
%PARAMETROS EXTRINSECOS - CAMARA -PTZ-
xptz=40; % coord. X de PTZ en trama {m=0} medida en [m]
yptz=23; % coord. Y de PTZ en trama {m=0} medida en [m]
zptz=-30; % coord. Z de PTZ en trama {m=0} medida en [m]

%Angulos de Euler ZYX+ZX para la orientacion de la camara PTZ en trama {0}
%Notar que la orientacion descripta por ZYX corresponde a la base de la
%camara PTZ que denomino trama {PTZ0} y luego se le agrega la rotacion ZX
%que corresponde al mecanismo PAN y TILT.
alfa_ptz=0;
beta_ptz=-0;
gama_ptz=+0;
phiptz=-0*pi/180; %PAN
titaptz=0*pi/180; %TILT

%Finalmente Tptz es la Tptz_r_0 es la trama de {PTZ} respecto de {0}
Tptz = transl(xptz, yptz, zptz) * eulZYX2tr_dami_ptz(alfa_ptz,beta_ptz,gama_ptz,phiptz,titaptz);
```

Figura 10: Simulación de la cámara PTZ :Determinación de los parámetros extrínsecos de la cámara.

3.2.6. Incerteza de los parámetros extrínsecos.

3.3. Simulación del nivel de zoom.

Para poder estimar el nivel de zoom necesario para poder captar al objetivo (por ej. la patente de un auto), es necesario conocer el campo de visión con el que dispone la cámara para cada nivel de zoom, de ésta manera, y conociendo las dimensiones del objeto y la distancia de éste hasta la cámara en forma aproximada, se busca obtener imágenes del objeto con la mayor resolución posible y que capten al objetivo en su totalidad.

Mediante la función fiteada en la figura 7 se simula el nivel de zoom de la cámara, simplemente ingresando un nivel de zoom de entrada, la función la mapea al valor de distancia focal que la cámara tendría si hubiera ajustado realmente su lente para aumentar/reducir el nivel de zoom.

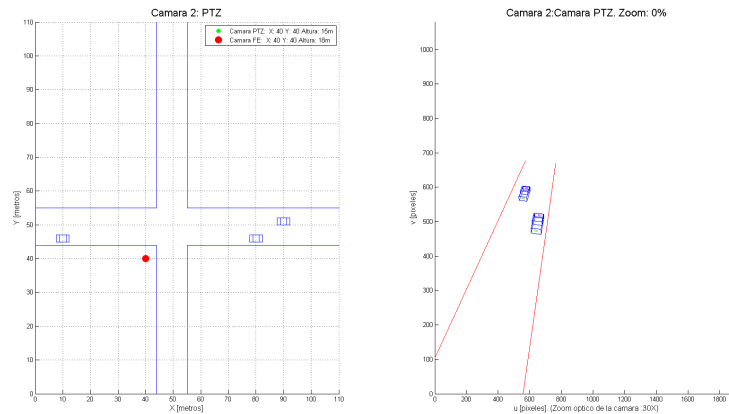


Figura 11: Simulación de la cámara PTZ :(1) Se simula escena de tránsito con tres automóviles, en coordenadas en la trama mundo. (2) Simulación del plano imagen proyectado por la cámara PTZ con un nivel de zoom de 0 % (se visualiza parte de la intersección y dos automóviles apenas visibles.) . (SIMULADOR_PTZ_FE.m)

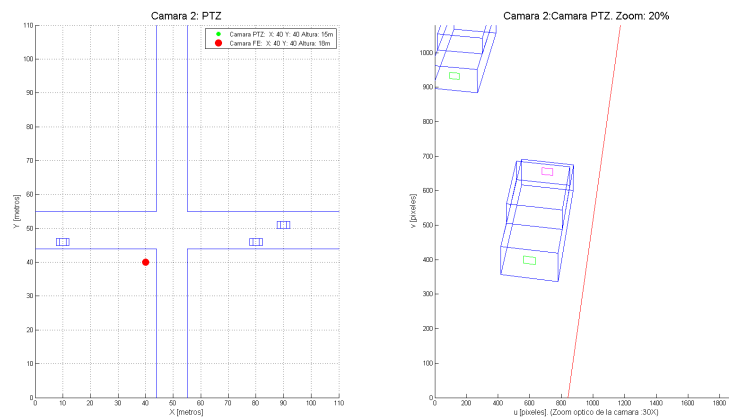


Figura 12: Simulación de la cámara PTZ :(1) Se simula escena de tránsito con tres automóviles, en coordenadas en la trama mundo. (2) Simulación del plano imagen proyectado por la cámara PTZ con un nivel de zoom de 20 %.(se visualiza parte de la intersección y los automóviles son mas visibles.).(SIMULADOR_PTZ_FE.m)

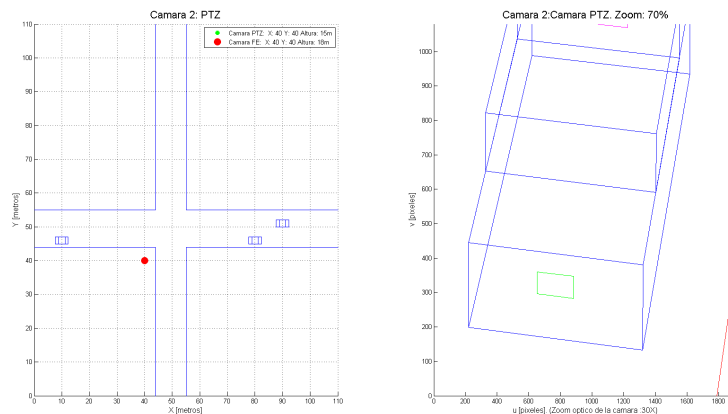


Figura 13: Simulación de la cámara PTZ :(1) Se simula escena de tránsito con tres automóviles, en coordenadas en la trama mundo. (2) Simulación del plano imagen proyectado por la cámara PTZ con un nivel de zoom de 70 %. (se visualiza la patente del automóvil.).(SIMULADOR_HERNAN.m)

De las ecuaciones (2) y (3) , se observa que el campo de visión es inversamente proporcional a la distancia focal. Por lo tanto, conociendo la relación entre el zoom y la distancia focal f , se puede encontrar la correspondencia que existe entre ellos y los ángulos del campo de visión.

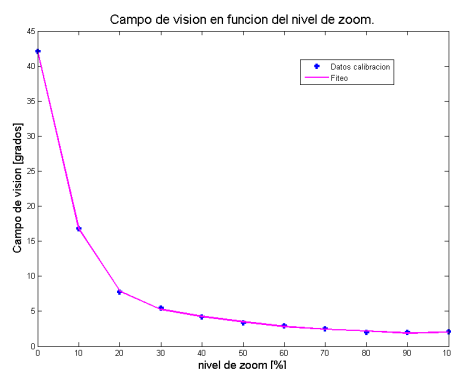


Figura 14: Función del campo de visión para cada nivel de zoom ,fiteada a partir de los datos obtenidos en la calibración intrínseca de la cámara PTZ.(fiteo_zoom.m)

3.4. Simulación del proceso de calibración. Minimización del error de ajuste sobre el mundo.

En esta sección se realizan simulaciones para analizar el proceso de calibración extrínseca. Se asume que un operador instalará la cámara PTZ según especificaciones previas tanto de posición como de orientación de la base de la cámara. Estas son:

$$\alpha = 0, \beta = 0, \gamma = 0$$

$x_{ptz} = 43,5$ (coord. X de PTZ en trama m=0 medida en [m])
 $y_{ptz} = 43,5$ (coord. Y de PTZ en trama m=0 medida en [m])
 $z_{ptz} = -15,0$ (coord. Z de PTZ en trama m=0 medida en [m])

Sin embargo, siempre habrá una diferencia entre la pose especificada y la pose real de instalación. Por lo tanto, en esta sección se simulará que la base de la cámara PTZ tiene una pose real desconocida (pero relativamente cercana a la pose especificada). Se asume que el proceso de instalación introduce un error alrededor de la pose especificada. Esto se realiza sumando un valor aleatorio que representa el error de instalación.

Para la calibración extrínseca, se utiliza un patrón de 6 puntos como se muestra en la figura 15A. Se posicionan los ángulos de pan, tilt y el nivel de zoom tal que el patrón de calibración pueda visualizarse completamente (ver figura 15B). Se asume que el error de posicionamiento del punto de calibración en la imagen es de 1 pixel.

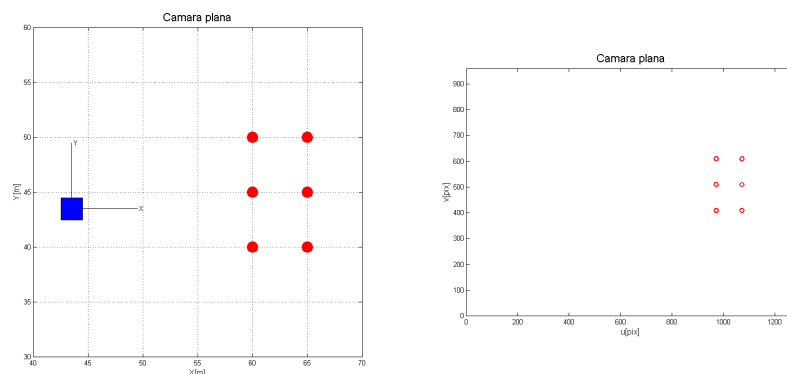


Figura 15: Simulación de calibración extrínseca para cámara PTZ. (Izquierda) Patrón de calibración de 6 puntos. (Derecha) Ejemplo de imagen observada por la cámara PTZ.

Se simularon un total de 20 situaciones con distintos ángulos de orientación de la base y con un desconocimiento de alrededor de 10° en los ángulos de pose y de 0.5 metros en las coordenadas (x, y, z) de la cámara.

Inicialmente, debido al error de instalación, existe cierta diferencia entre la pose real y la deseada, luego se aplica un proceso de optimización que minimiza la distancia cuadrática entre los puntos de calibración en el mundo y la predicción según el modelo de proyección. La optimización de la pose de la cámara, se realiza con la función *nlinfit* de *Matlab*. En el siguiente cuadro, se observan los errores de estimación de la pose de la cámara antes y después del proceso de optimización:

Parámetro	Error estimación inicial	Error estimación final
$\alpha[^\circ]$	-2.50	-0.25
$\beta[^\circ]$	-6.75	-0.24
$\gamma[^\circ]$	-9.17	0.16
$X_m[m]$	0.11	0.11
$Y_m[m]$	-0.71	0.12
$Z_m[m]$	1.01	-0.05

Cuadro 1: Errores obtenidos en la estimación de la pose de la cámara PTZ (antes y después de la optimización).

Se observa una mejora drástica en los ángulos de orientación y en la coordenada z de la cámara. Finalmente, es importante destacar que se simularon numerosas poses de instalación diferentes a la descrita previamente. En todos los casos los resultados obtenidos fueron similares.

3.4.1. Incerteza de los parametros extrinsecos usando la tecnica de bootstrap.

3.5. Simulación de cámara PTZ.

Una vez descrita y seteada la pose de la cámara, se originan puntos en coordenadas del mundo tal que describan una escena de tránsito (una intersección de calles y varios automóviles), para luego simular la proyección de dichos puntos al plano imagen mediante el toolbox de *MATLAB*, verificando los resultados obtenidos mediante un script propio.

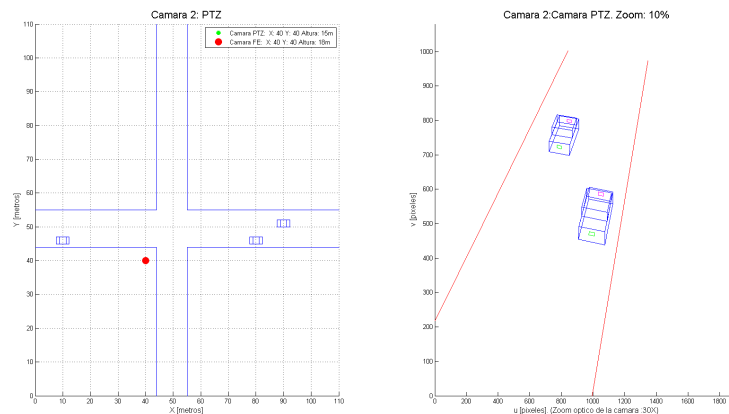


Figura 16: Simulación de la cámara PTZ :(1) Se simula escena de tránsito con coordenadas en la trama mundo. (2) Simulación del plano imagen proyectado por la cámara.(*SIMULADOR_PTZ_FE.m*)

4. Simulación de un sistema de adquisición con una cámara fisheye (FE). Calibración y estimación de posición.

4.1. Modelado de la cámara fisheye.

Los parámetros que describen el funcionamiento de una cámara se denominan intrínsecos, éstos parámetros convierten las coordenadas en el marco de referencia del dispositivo a las coordenadas en píxeles de la imagen, dichos parámetros son; la distancia focal y el punto principal.

A su vez, son los parámetros extrínsecos los que definen la posición y la orientación del marco de referencia de la cámara con respecto al mundo se denominan y están determinados por una rotación y una traslación en dicho sistema de coordenadas.

Para modelar el comportamiento de la cámara fisheye, se opta por el modelo estereográfico de proyección de lente fisheye [2], en el cual un punto $P = (X, Y, Z)$ en el mundo se representa en coordenadas esféricas $R(\theta, \phi)$ con respecto al origen de la cámara.

4.1.1. Modelo de proyección.

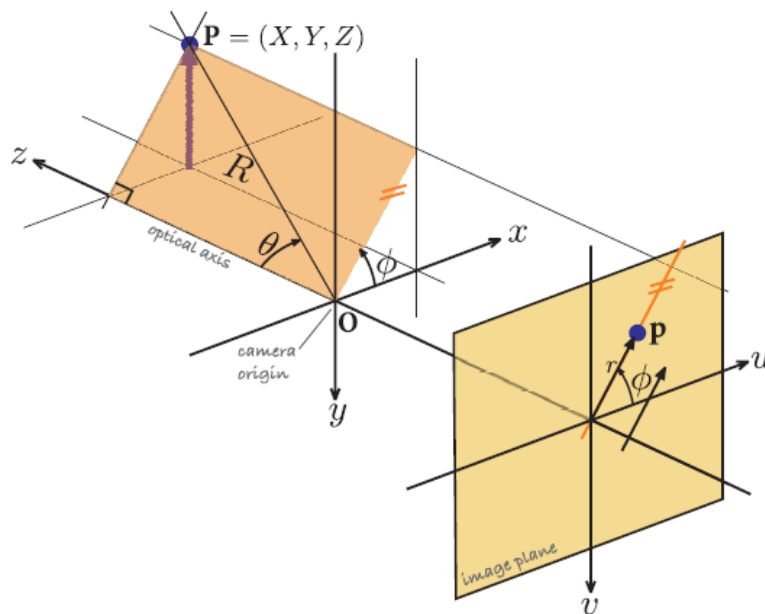


Figura 17: Modelo de la cámara fisheye.

Nombre del modelo	Ecuación	Error absoluto medio
Estereográfico	$r(\theta) = k \tan(\theta/2)$	4.91 píxeles
Equiangular	$r(\theta) = k \theta$	36.77 píxeles
Equisólido	$r(\theta) = k \sin(\theta/2)$	39.40 píxeles

Figura 18: Tabla I. Modelos de proyección.

Por lo que el mapeo desde la trama mundo a la base de la cámara FE para el modelo estereográfico es el siguiente:

$$R = \sqrt{X^2 + Y^2 + Z^2}, \phi_{VCA} = \arctan(Y/X), \theta_{VCA} = \arccos(R/Z) \quad (4)$$

$$r_{VCA} = k \tan\left(\frac{\theta_{VCA}}{2}\right)$$

$$x_{VCA} = r_{VCA} \cos(\phi_{VCA})$$

$$y_{VCA} = r_{VCA} \sin(\phi_{VCA})$$

$$u_{VCA} = x_{VCA}/\rho_w + u_0 \quad (5)$$

$$v_{VCA} = y_{VCA}/\rho_h + v_0 \quad (6)$$

Donde:

θ_{VCA} es el ángulo hacia afuera del eje óptico.

ϕ_{VCA} es el ángulo de rotación alrededor del eje óptico.

(u_{VCA}, v_{VCA}) son las coordenadas sobre la imagen medidas en píxeles.

(ρ_w, ρ_h) son el ancho y la altura de cada píxel.

(u_0, v_0) son las coordenadas del punto principal, el punto en el que el eje óptico interseca el plano imagen.

4.2. Calibración de la cámara fisheye.

4.2.1. Calibración intrínseca de la cámara fisheye.

A partir de las ecuaciones (5) y (6), queda definida la proyección del mundo al plano imagen de la cámara FE, la cual depende de los parámetros intrínsecos, k (relación entre la distancia de un punto al centro de la imagen y su coordenada esférica) y las coordenadas del punto principal.

La calibración intrínseca de la cámara fisheye VIVOTEK fue realizada en un trabajo anterior [4], por lo que para las simulaciones de este trabajo, se utilizan los datos obtenidos en el citado trabajo.

El proceso realizado en [4] para la calibración intrínseca de la cámara fisheye fue el siguiente :

Para calibrar la cámara fisheye se utiliza un patron uniforme de puntos en coordenadas cartesianas. Se miden las coordenadas esféricas (θ, ϕ) (ver figura 17) y sus coordenadas polares en la proyección a través de la lente fisheye (r, θ) . En la figura 19) se muestran las mediciones de r vs θ y los ajustes por cuadrados mínimos segun los modelos de la tabla I. Se comparan los ajustes de los modelos por cuadrados mínimos y se concluye que el modelo que mejor ajusta la cámara fisheye es el modelo estereográfico.

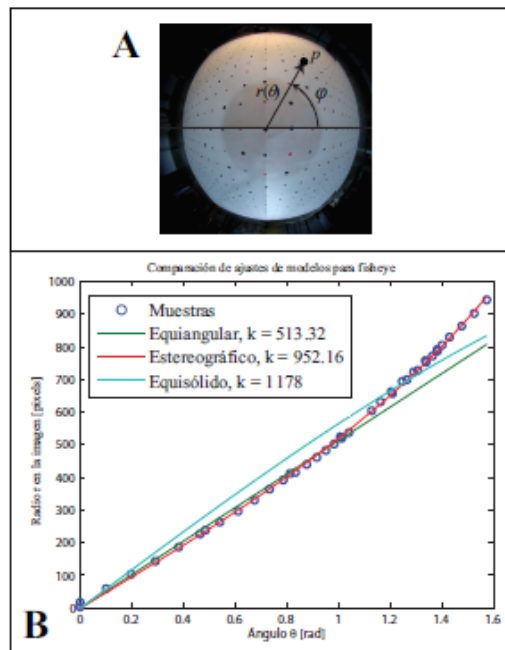


Figura 19: A) Patron uniforme de puntos en coordenadas cartesianas vistos a través de la camara fisheye. B) Mediciones de r vs θ y los ajustes por cuadrados mínimos segun los modelos de la tabla I.

4.2.2. Simulación de los parámetros intrínsecos.

```
% CAMARA VCA (FISHEYE)
global camvca
camvca = FishEyeCamera('name', 'Fisheye',...
    'projection', 'stereographic',...
    'resolution', [Npixvcaw Npixvcvh],...
    'centre', [Npixvcaw/2 Npixvcvh/2], ...
    'noise', 0.00, ...
    'k', kvca);
```

Figura 20: Definición del objeto *FishEyeCamera* mediante el *RoboticsToolbox* de *MATLAB*, seteada con los parámetros intrínsecos calculados para la cámara FE.(*SIMULADOR_PTZ_FE.m*)

4.2.3. Incerteza de los parámetros intrínsecos.

4.2.4. Calibración extrínseca de la cámara fisheye.

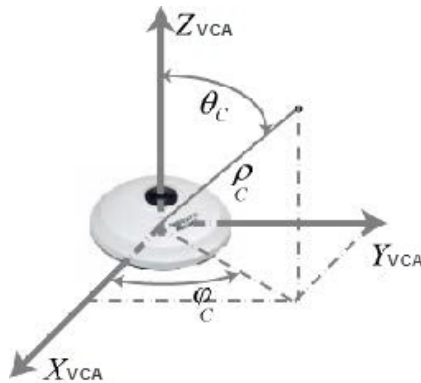


Figura 21: Coordenadas esféricas de la cámara fisheye.

Los parámetros extrínsecos de la cámara FE son:

- $(x_{VCA}, y_{VCA}, z_{VCA})$: Posición de la base de la cámara en el mundo.
- $(\alpha_{VCA}, \beta_{VCA}, \gamma_{VCA})$: Ángulos de rotación de la base de la cámara con respecto a los ejes Z_m , Y_m y X_m del mundo.

4.2.5. Simulación de los parámetros extrínsecos.

Para llevar a cabo este proceso en el simulador, se calcula la proyección que lleva los puntos de coordenadas del mundo a coordenadas respecto a la base de la cámara.

Por lo que la traza de la base de la cámara FE respecto al mundo es:

$$T_{FE} = T_{XYZ} R_{XYZ}$$

Donde :

- R_{XYZ} : Rotación de la base de la cámara respecto al mundo.
- T_{XYZ} : Traslación de la base de la cámara al mundo.

Siendo los coeficientes de estos, los parámetros extrínsecos que determinan la posición de la cámara fisheye con respecto al mundo.

Primero se determinan los parámetros extrínsecos en el simulador, es decir, se determina la pose de la cámara que se va a simular seteando los valores de $x_{VCA}, y_{VCA}, z_{VCA}, \alpha_{VCA}, \beta_{VCA}, \gamma_{VCA}$:

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%POSE CAMARA FISHEYE

%desplazamientos de las camaras en trama mundo {0}
xvca=40; % coord. X de FE en trama {m=0} medida en [m]
yvca=40; % coord. Y de FE en trama {m=0} medida en [m]
zvca=-18; % coord. Z de FE en trama {m=0} medida en [m]

%Angulos de Euler ZYX para la orientacion de la camara FE en trama mundo {0}
alfa_vca=0.15*pi/180;
beta_vca=0.2;
gama_vca=-0.9;

%Finalmente Tvca es la trama de {FE} respecto del mundo {0}

Tvca = transl(xvca, yvca, zvca) * eulZYX2tr_dami(alfa_vca,beta_vca,gama_vca);

```

Figura 22: Simulación de la cámara FE :Simulación de los parámetros extrínsecos.(*SIMULADOR_PTZ_FE.m*)

4.2.6. Incerteza de los parámetros extrínsecos.

4.3. Simulación del proceso de calibración. Minimización del error de ajuste sobre el mundo.

En esta sección se utilizan simulaciones para analizar el proceso de calibración extrínseca de la cámara FE. Se asume que un operador instalará la cámara FE según especificaciones previas tanto de posición como de orientación de la base de la cámara. Estas son:

$$\alpha = 0, \beta = 0, \gamma = 0$$

$$\begin{aligned} x_{fe} &= 43,5 \text{ (coord. X de FE en trama m=0 medida en [m])} \\ y_{fe} &= 43,5 \text{ (coord. Y de FE en trama m=0 medida en [m])} \\ z_{fe} &= -15,3 \text{ (coord. Z de FE en trama m=0 medida en [m])} \end{aligned}$$

Sin embargo, como en el caso de la cámara PTZ, se asume que siempre habrá una diferencia entre la pose especificada y la pose real de instalación. Por lo tanto, en esta sección se simulará que la base de la cámara FE tiene una pose real desconocida pero relativamente cercana a la pose especificada. Se asume que el proceso de instalación introduce un error alrededor de la pose especificada. Esto se realiza sumando un valor aleatorio que representa el error de instalación:

Para la calibración extrínseca, se utiliza un patrón de 6 puntos como se muestra en la figura ??A. Se posicionan los ángulos de pan, tilt y el nivel de zoom tal que el patrón de calibración pueda visualizarse completamente (ver figura 15B). Se asume que el error de posicionamiento del punto de calibración en la imagen es de 1 pixel.

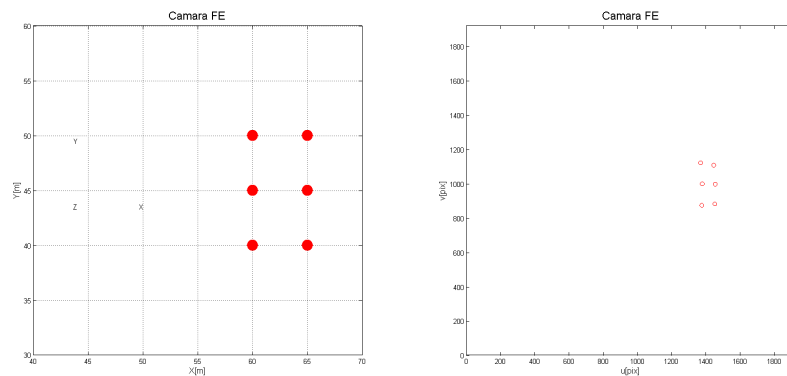


Figura 23: Simulación de calibración extrínseca para cámara FE. (Izquierda) Patrón de calibración de 6 puntos. (Derecha) Ejemplo de imagen observada por la cámara FE.

Se simularon un total de 20 situaciones con distintos ángulos de orientación de la base y con un desconocimiento de alrededor de 10° en los ángulos de pose y de 0.5 metros en las coordenadas (x, y, z) de la cámara.

Inicialmente, debido al error de instalación, existe cierta diferencia entre la pose real y la deseada, luego se aplica un proceso de optimización que minimiza la distancia cuadrática entre los puntos de calibración en el mundo y la predicción según el modelo de proyección. La optimización de la pose de la cámara, se realiza con la función *nlinfit* de *Matlab*. En el siguiente cuadro, se observan los errores de estimación de la pose de la cámara antes y después del proceso de optimización:

Parámetro	Error estimación inicial	Error estimación final
$\alpha[^\circ]$	-9.63	0.20
$\beta[^\circ]$	-13.85	0.33
$\gamma[^\circ]$	10.88	-0.17
$X_m[m]$	-0.67	-0.06
$Y_m[m]$	-0.08	-0.11
$Z_m[m]$	-0.26	0.13

Cuadro 2: Errores obtenidos en la estimación de la pose de la cámara FE (antes y después de la optimización).

Se observa una mejora drástica en los ángulos de orientación y en la coordenada z de la cámara. Finalmente, es importante destacar que se simularon numerosas poses de instalación diferentes a la descrita previamente. En todos los casos los resultados obtenidos fueron similares.

4.3.1. Incerteza de los parametros extrinsecos usando la tecnica de bootstrap.

4.4. Simulación de cámara FE.

Una vez descripta y seteada la pose de la cámara, se originan puntos en coordenadas del mundo tal que describan una escena de tránsito (una intersección de calles y varios automóviles), para luego simular la proyección de dichos puntos al plano imagen mediante el toolbox de *MATLAB*, verificando los resultados obtenidos realizando el cálculo de las proyecciones mediante un script propio.

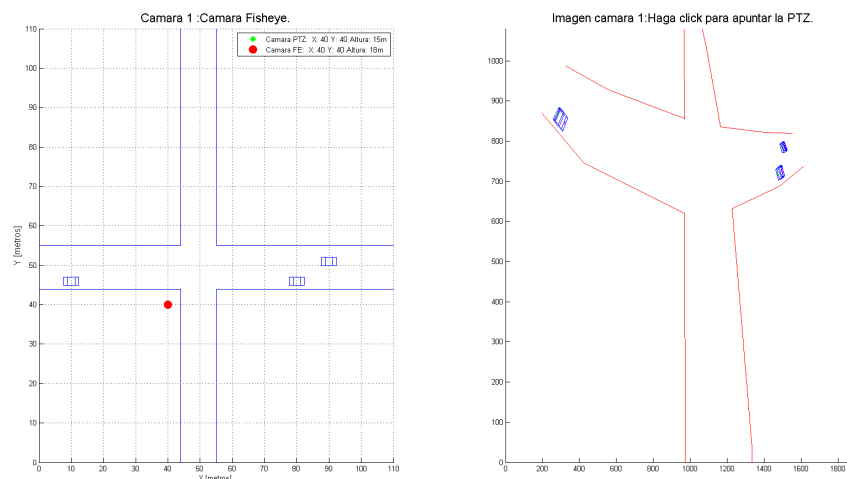


Figura 24: Simulación de la cámara FE :(1) Se simula escena de tránsito con coordenadas en la trama mundo. (2) Simulación del plano imagen proyectado por la cámara.(*SIMULADOR_PTZ_FE.m*)

5. Simulación de los delays de movimiento y de comunicación para las cámaras PTZ y FE.

5.1. Retardos en la captura.

5.1.1. Retardo en la captura de la cámara PTZ.

5.1.2. Retardo en la captura de la cámara fisheye.

5.1.3. Diferencia de retardo en la captura de ambas cámaras.

5.1.4. Frecuencia de capturas de las cámaras.

5.2. Modelado del retardo de movimiento de la cámara PTZ.

Desde las primeras pruebas en las que se apuntó con la cámara PTZ, a través de movimientos de pan, tilt y de zoom, se observó que existía un retraso notable entre el inicio y el fin de cada movimiento realizado por la cámara, por lo tanto, para proporcionar información acerca de la eficiencia del dispositivo para realizar aplicaciones

dinámicas en tiempo real , como la intercepción de vehículos en movimiento, se decidió realizar pruebas con el fin de estimar un modelo matemático que permita predecir el tiempo de retardo que le lleva a la cámara realizar cada uno de estos movimientos.

Para llevar esto a cabo se realizaron tres experimentos, uno para cada movimiento posible de la cámara:

1. Usando la librería *OpenCV* y un script de Python basado en el protocolo ONVIF, se manipula la cámara para moverla un cierto ángulo (en el caso de PAN o TILT), o mover el lente de ésta en el caso del ZOOM.
2. Al instante la cámara comienza a grabar un video de unos pocos segundos, mientras que el script guarda en memoria una imagen frame por frame junto con un valor de tiempo asociado para cada frame.
3. Mediante inspección visual, se verifica frame por frame hasta encontrar el frame de inicio y el del final del movimiento, y teniendo los valores de tiempo asociados a cada frame, se calcula cuanto tiempo le tomó a la cámara realizar el movimiento.
4. Este proceso se repite varias veces para diferentes ángulos y niveles de zoom para generar un conjunto de datos.
5. A partir de los datos del paso anterior, se ajusta una curva y se obtiene el modelo de retardo de movimiento de la cámara para cada movimiento.

Los modelos obtenidos para el retraso de la cámara PTZ HECKER fueron:

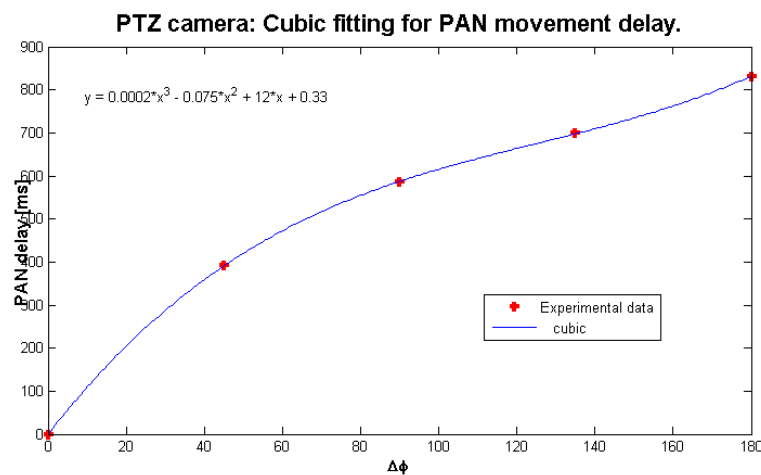


Figura 25: Modelo obtenido experimentalmente para el retardo del movimiento de PAN de la cámara HECKER PTZ.

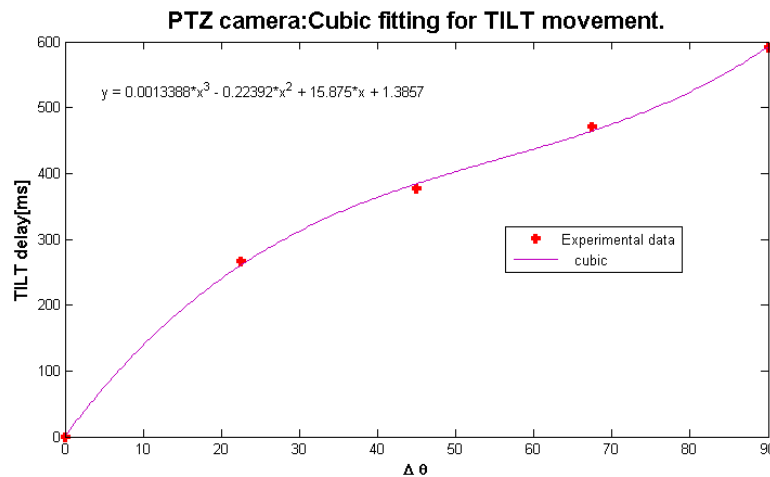


Figura 26: Modelo obtenido experimentalmente para el retardo del movimiento de TILT de la cámara HECKER PTZ.

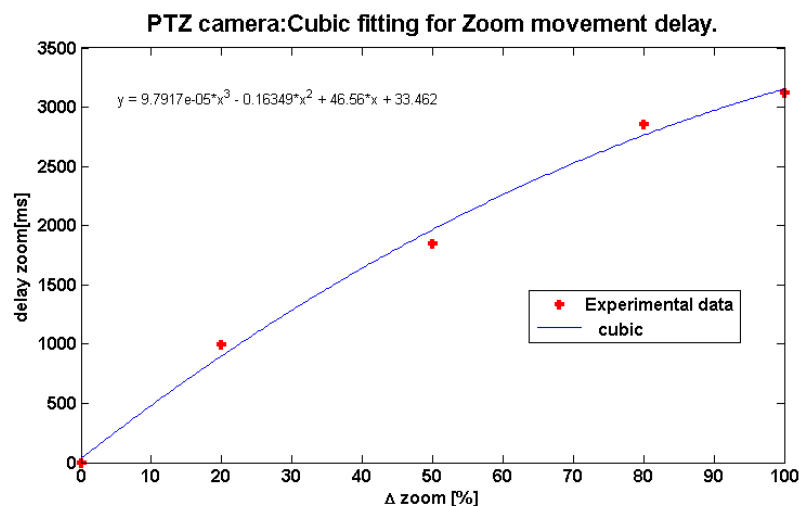


Figura 27: Modelo obtenido experimentalmente para el retardo del movimiento de ZOOM de la cámara HECKER PTZ.

6. Simulación del sistema PTZ+FE.

Lo que se pretende para diseñar el simulador de este sistema es que un operario pueda marcar un punto en la imagen de la cámara fisheye, para luego calcular los ángulos PAN-TILT a donde hay que apuntar la cámara PTZ para que el objeto quede en el centro de la imagen capturada por ella y variar el campo de visión a través del ZOOM para que éste se vea en mayor detalle.

6.1. Cinemática de apunte de la cámara PTZ.

Hasta ahora, se cuenta con el modelado intrínseco de ambas cámaras y se demostró como pueden simularse los parámetros extrínsecos que dependen de la pose en la que se encuentran. Con estos datos, se calcula la retroproyección de un punto en la imagen fisheye hacia el mundo y la proyección de una coordenada del mundo a la imagen PTZ.

El problema a resolver es que la cámara PTZ se mueva y apunte hacia un objeto que se visualiza en la imagen fisheye y haga zoom sobre él.

Este problema fue resuelto en un trabajo anterior [1], cuya solución propuesta fue la siguiente:

La solución puede ser dividida en tres partes, una primera de calibración de la base de la cámara PTZ, otra de cálculo de ángulos de apunte y una última de determinación del nivel de zoom.

6.1.1. Calibración de la base de la cámara PTZ.

La rototraslación que convierte coordenadas en el mundo a coordenadas en el marco de referencia de la cámara PTZ se puede descomponer en dos:

$${}^C_0T = {}^C_BT {}^B_0T, \quad (7)$$

donde:

- C_BT : matriz de transformación de las coordenadas de la base al lente de la cámara.
- B_0T : matriz de transformación de las coordenadas del mundo a las de la base de la cámara.

La simulación de la pose de la cámara establece una B_0T , mientras que C_BT se puede calcular a partir de los ángulos pan y tilt ya que representa justamente la rotación relativa entre la base de la cámara y la lente. Para definir esta rotación en símbolos se considera primero a los dos marcos de referencia (lente y base) coincidentes, es decir, sin todavía aplicar ninguna rotación. En la figura 28 se indican las direcciones de rotación pan y tilt. Como puede verse, el pan es una rotación alrededor del eje vertical z fijo a la base de la cámara y el tilt es una rotación alrededor del eje x móvil.

La composición de estas rotaciones resulta en la matriz de transformación sin componente de traslación:

$$R(\alpha, \beta) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \cos(\beta) & -\sin(\alpha) \sin(\beta) & 0 \\ \sin(\alpha) & \cos(\alpha) \cos(\beta) & \cos(\alpha) \sin(\beta) & 0 \\ 0 & -\sin(\beta) & \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (8)$$

Esta matriz es la que transforma coordenadas expresadas en la lente a coordenadas en el marco de referencia de la base de la cámara.

En consecuencia las coordenadas de un punto en el marco de referencia de la lente se transforma al marco de referencia de la base haciendo

$$X_B = R(\alpha, \beta) X_C = {}^C_BT^{-1} X_C. \quad (9)$$

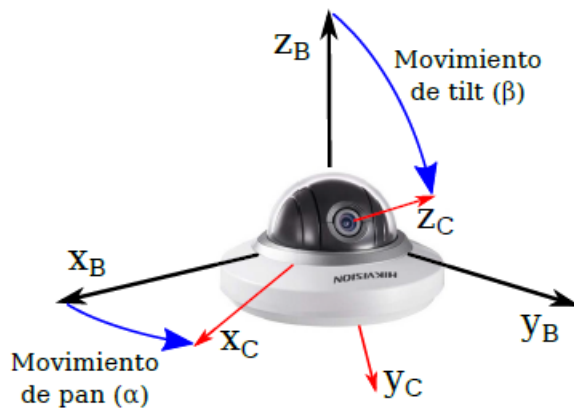


Figura 28: Movimiento en los planos horizontal y vertical de cámara PTZ.

Es decir que la transformación de un punto en el marco referencia del mundo a la lente es

$$X_C = {}^C T_0^B T X_0 = R(\alpha, \beta)^{-1} {}^B T_0 X_0 = {}^C T X_0 \quad (10)$$

donde queda claro que $R(\alpha, \beta)^{-1} {}^B T_0 = {}^C T$ y por lo tanto la trama de la base puede despejarse como

$${}_0^B T = R(\alpha, \beta) {}^C T. \quad (11)$$

Una vez conocida la matriz de transformación ${}_0^B T$, es posible transformar un punto X_0 descrito en las coordenadas del mundo a las coordenadas de la base de la cámara X_B .

$$X_B = {}_0^B T X_0 \quad (12)$$

6.1.2. Cálculo de los ángulos de apunte.

Para saber a donde debería apuntar la cámara PTZ de manera que el punto de interés elegido se vea en el centro de la imagen se realiza la cinemática inversa de apunte. Esta operación consiste en resolver α y β en función de una posición en el mundo. X_C se denota al punto en coordenadas de la cámara que se encuentra en el origen del plano (x, y) y se representa por un vector en la dirección del eje z ,

$$X_C = \begin{bmatrix} 0 \\ 0 \\ ||X_B|| \end{bmatrix} = {}^C T X_B. \quad (13)$$

Como se mencionó antes, la operación inversa de ${}^C_B T$ es la rotación en los ángulos pan y tilt y siendo estos los datos que se desea conocer, se realiza la sustitución:

$$\begin{aligned} {}^C_B T^{-1} \begin{bmatrix} 0 \\ 0 \\ ||X_B|| \end{bmatrix} &= X_B \\ R(\alpha, \beta) \begin{bmatrix} 0 \\ 0 \\ ||X_B|| \end{bmatrix} &= X_B \\ ||X_B|| \begin{bmatrix} -\sin(\alpha) \sin(\beta) \\ \cos(\alpha) \sin(\beta) \\ \cos(\beta) \end{bmatrix} &= \begin{bmatrix} X_{Bx} \\ X_{By} \\ X_{Bz} \end{bmatrix} \end{aligned} \quad (14)$$

Quedando un sistema de 3 ecuaciones con dos incógnitas:

$$\begin{aligned} ||X_B|| (-\sin(\alpha) \sin(\beta)) &= X_{Bx} \\ ||X_B|| (\cos(\alpha) \sin(\beta)) &= X_{By} \\ ||X_B|| (\cos(\beta)) &= X_{Bz} \end{aligned} \quad (15)$$

En las dos primeras ecuaciones resultantes del sistema 15 se obtiene el seno y el coseno de α directamente, multiplicados por un factor que es irrelevante en la solución. Despejando se llega a:

$$\alpha = \arctan(-X_{Bx}, X_{By}) \quad (16)$$

Para hallar β se cuenta con el coseno en la tercera ecuación y para llegar al seno se realiza el módulo de las dos primeras ecuaciones y se utiliza la identidad $\sin^2(\alpha) + \cos^2(\alpha) = 1$ derivando en:

$$\beta = \arctan(\sqrt{X_{Bx}^2 + X_{By}^2}, X_{Bz}) \quad (17)$$

6.1.3. Determicación del nivel de zoom.

La última fase del sistema de apunte consiste en conocer qué valor de zoom tiene que tener la cámara PTZ para ver el objetivo de interés con más detalle. Una vez que apunta correctamente se define con qué ángulo de campo de visión se obtiene una imagen en la que el objeto se vea lo mejor posible.

La calibración intrínseca de la cámara PTZ se realizó para cada valor de zoom en el que se podía configurar la lente. Junto con los parámetros inherentes a la cámara se obtuvieron los ángulos del campo de visión en grados que se tenían en cada valor de zoom probado.

La cuestión ahora es definir cómo se obtiene el ángulo del campo de visión que se necesita. Se plantea la situación como la que se muestra en la figura 29, donde el objeto de interés se encuentra a una distancia $|x|$ de la cámara. Se decide qué *radio* alrededor del objetivo se quiere incluir en la imagen y se arma de esta manera el ángulo de visión fov .

La relación geométrica que se establece es:

$$\tan\left(\frac{fov}{2}\right) = \frac{radio}{|x|}. \quad (18)$$

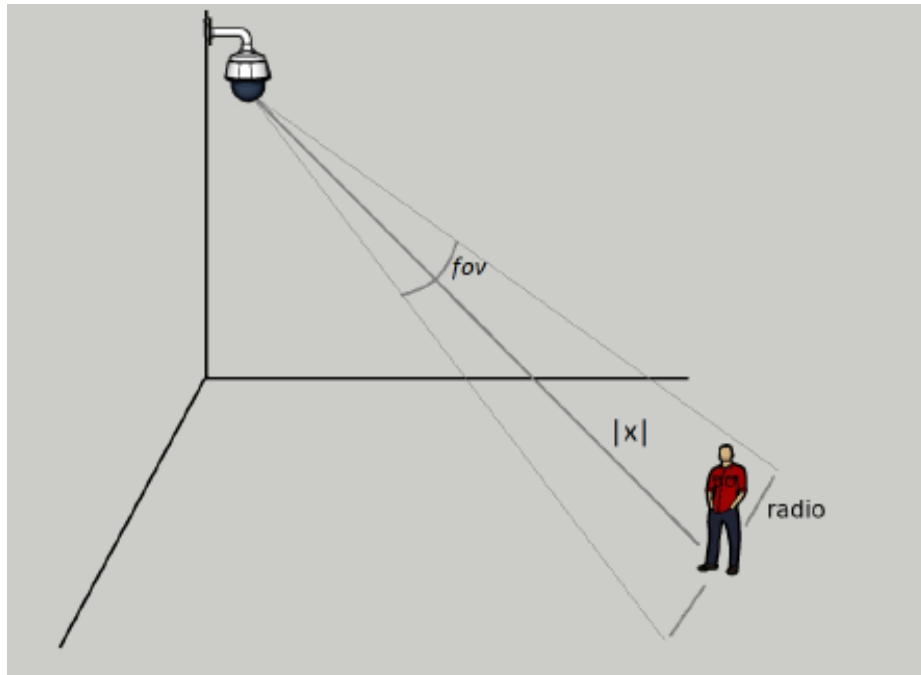


Figura 29: Relación geométrica entre el campo de visión y la distancia a la que se encuentra el objetivo.

El *radio* se considera conocido ya que puede ser elegido por el operador.

Usando la ecuación 18 y los datos de calibración se llega a una relación directa entre la distancia a la que se encuentra el objeto $|x|$ y el zoom a configurar en la lente.

En el sistema se obtiene $|x|$ del módulo de la distancia entre la cámara y la posición en la que se encuentra el objeto en el mundo. (Esta posición es conocida cuando se convierte el punto en la imagen de la fisheye en píxeles a las coordenadas del mundo en metros.)

De esta manera, partiendo de X_0 y evaluando la función anterior se llega a un valor de zoom. En este punto reaparece la limitación del hardware de la cámara que solo permite modificar los valores con un paso de 10 %. Es por este motivo que al resultado se le aplica una función de redondeo para que lo acerque al valor admisible más cercano.

7. Métodos para la estimación de velocidad y trayectorias de los vehículos con la cámara PTZ y FE.

7.1. Modelos del movimiento de los vehículos en el mundo.

7.2. Estimación de velocidad con Filtro Kalman, Cuadrados mínimos recursivos y Filtro de partículas.

8. Análisis de estrategias de posicionamiento de la cámara PTZ en función del posicionamiento inferido con la cámara FE.

8.1. Diseño de una estrategia de apunte de la PTZ intersectar el vehículo y para capturar una imagen de la patente.

8.2. Diseño de una estrategia para optimizar el tiempo en el que la cámara graba el video.

9. Bibliografía.

Referencias

- [1] Rojas,L.(2018).Sistema de Videovigilancia integrado por una Cámara Fisheye y una Cámara Pan-Tilt-Zoom ,UNQ.
- [2] Peter Corke.(2017),Robotics,Vision and Control Fundamental Algorithms in MATLAB.Springer International Publishing.
- [3] https://docs.opencv.org/master/d4/d94/tutorial_camera_calibration.html
- [4] Stanganelli,,Oliva,Noblia,Safar.(2014) ,Calibración de una Cámara Fisheye Comercial con el Modelo Unificado para la Observación de Objetos Múltiples,UNQ.