

# Algoritmos de navegación para robots móviles usando visión estéreo

Agustín Olmedo, *Estudiante, UBA*  
 Amit Stein, *Estudiante, UBA*,  
 y Fernando Bugni, *Estudiante, UBA*,

**Abstract**—El presente trabajo desarrolla una implementación para la navegación de un robot utilizando visión estéreo. Al mismo utiliza conceptos sobre vision computacional: calibración estéreo, mapa de disparidad, rectificación. La implementación fue desarrollada íntegramente en C++ utilizando la librería OpenCV, además de otras librerías para la movilidad del robot.

**Index Terms**—Computer Society, L<sup>A</sup>T<sub>E</sub>X, Visión Computacional, OpenCV, Mapas de Disparidad



## 1 INTRODUCCIÓN

LA navegación de robots es el área encargada de desarrollar robots capaces de desplazarse de forma autónoma. La utilización de esta tecnología se utiliza en distintos ámbitos como por ejemplo: misiones espaciales, misiones militares, asistencia para personas con discapacidades e inclusive como guías en una exposición.

El objetivo del presente trabajo es desarrollar un algoritmo de navegación de robots utilizando visión estéreo. En la actualidad, existen diversas técnicas de visión por computadora que nos brindan la información necesaria sobre el entorno en el que se encuentra el robot para que este pueda tomar la mejor decisión posible. Algunos ejemplos de estas técnicas son: utilización de cámaras tipo eyebird enfocando todo el terreno, generación de mapas a través de las imágenes tomadas o también reconocimiento de objetos a través de imágenes. El enfoque de este trabajo consiste, principalmente, en utilizar mapas de disparidad para obtener información sobre la distancia entre el robot y los objetos que luego será utilizada por el robot para decidir el próximo movimiento a realizar. Un paper que ha influido fuertemente el presente trabajo fue Region of Interest in Disparity Mapping for Navigation of Stereo Vision Autonomous Guided Vehicle. Este paper posee un enfoque similar a este, en cuanto a las técnicas de visión utilizadas. Las diferencias radican en la interpretación de la información que nos brindan los mapas de disparidad.

El presente trabajo describe los fundamentos teóricos y prácticos utilizados para lograr la navegación de un robot, en particular del exabot (poner link, foto o alguna referencia). Cabe destacar que el entorno utilizado juega un papel fundamental en el buen desempeño del algoritmo, utilizamos un entorno llano iluminado para no perder ningún dato al tomar cada fotografía. La cámara se encuentra en la parte superior del frente del robot y su movilidad se reduce a avanzar, retroceder y rotar sobre su propio eje.

Por último, dejaremos ideas para alguna posible continuación.

palabras clave: explicacion del problema, estado del arte (antes hablar del arte), entorno de prueba, approach del paper, arquitectura de hardware

## 2 CALIBRACIÓN DE LA CÁMARA ESTÉREO

Cuando capturamos una escena del mundo (tres dimensiones) con una cámara, ésta se proyecta sobre una imagen en dos dimensiones.

Para conocer la posición de cada punto es necesario definir un sistema de referencias, es decir, los tres ejes principales con su origen, sus orientaciones y sus unidades.

Existen tres sistemas de referencias posibles:

- El sistema del mundo: independiente de la cámara. Unidades físicas, por ejemplo, metros.
- El sistema de la cámara: origen en el centro óptico. Unidades físicas.
- El sistema de la imagen: origen en la esquina de la imagen. Unidades en píxeles.

Los parámetros intrínsecos son los que relacionan el sistema de referencia de la cámara con el sistema de referencia de la imagen y los parámetros extrínsecos son los que relacionan el sistema de referencia del mundo con el sistema de referencia de la cámara.

El sistema de referencia de la cámara y el sistema de referencia del mundo tienen las mismas unidades. Lo que cambia es el origen y la orientación de los ejes. Esto significa que la conversión de un sistema al otro implica una traslación (cambio de origen) y una rotación (cambio de orientación). Ver figura 1.

Por otro lado, el sistema de referencia de la cámara y el sistema de referencia de la imagen no tienen las mismas unidades. Con lo cual hay que realizar un cambio de unidad, dividiendo por el tamaño del pixel. Además hay que cambiar el lugar del origen sumando la posición del centro de la imagen. Ver figura 2.

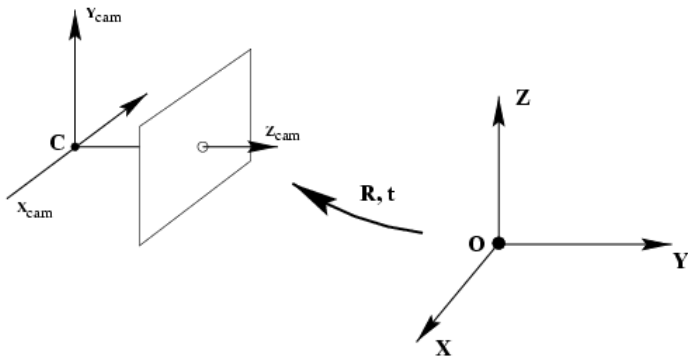


Fig. 1. Parámetros extrínsecos

Los parámetros intrínsecos de una cámara son los atributos que posee la misma, como la distancia focal, punto principal, skew, distorsión de los lentes, etc.

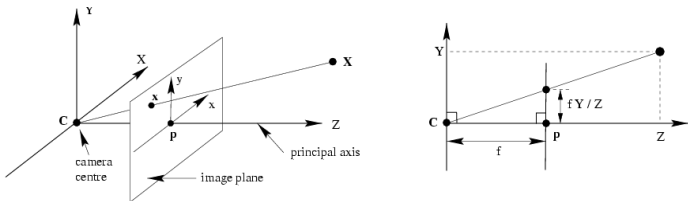


Fig. 2. Parámetros intrínsecos

La calibración de una cámara es el procedimiento por el cual obtenemos los parámetros intrínsecos y extrínsecos de la cámara a partir de imágenes tomadas por la misma.

Una vez que obtenemos los parámetros intrínsecos y extrínsecos, podemos decir que la cámara está calibrada, pudiendo representar cualquier punto del mundo real sobre el plano 2D de la cámara.

Para el desarrollo del presente trabajo utilizamos una cámara estéreo, con lo cual no nos alcanza solo con calibrar cada una de las cámaras, sino que además debemos calibrar las dos cámaras de forma conjunta. A esto se lo denomina calibración estéreo.

La calibración estéreo nos provee los parámetros intrínsecos de cada cámara (que son los obtenidos en la calibración de cada una) y los parámetros extrínsecos, que nos definen la posición relativa de las cámaras.

Para realizar dicha calibración utilizamos un toolbox de calibración de cámaras para matlab <sup>1</sup>. Dicha herramienta necesita una serie de fotos para realizar la calibración.

### 3 IMAGEN ESTÉREO

Al proyectarse los objetos, de un espacio tridimensional a una imagen bidimensional se pierde la información de la distancia a la cámara o profundidad (eje Z) de cada punto. Una forma de tratar de recuperar esta información es mediante el uso de dos cámaras, en lo que se conoce como visión estéreo.

Libelas<sup>2</sup> (basado en Learning Open cv pag 415 Stereo Imaging)

### 4 MAPAS DE DISPARIDAD

Un mapa de disparidad representa la distancia que existe entre el centro de las cámaras y después de leer lo explico bien¿

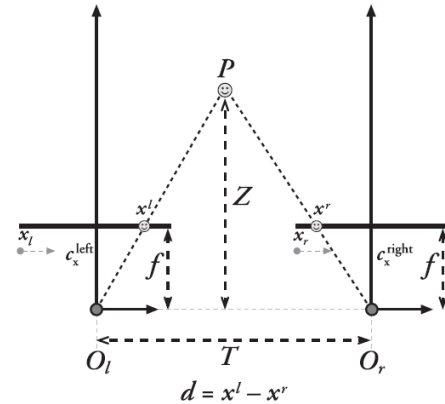
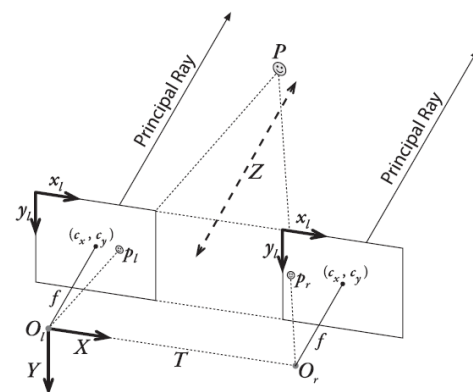
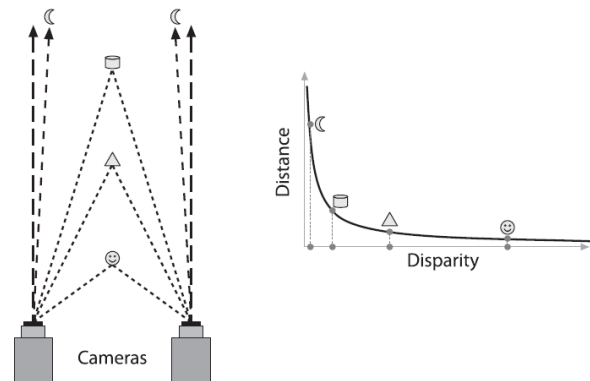


Fig. 3. figura 1:



1. [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/)

2. <http://www.rainsoft.de/software/libelas.html>

## 4.1 Calculo del mapa de disparidad

En esta seccion explicaremos los pasos ha realizar para obtener un mapa de disparidad a partir de dos imágenes capturadas por la cámara estéreo.

Básicamente, se buscan correspondencias entre los puntos que se encuentran en una imagen y los mismos puntos en la otra imagen. Con estas correspondencias y conociendo la distancia entre las camaras(linea base) podemos calcular la ubicacion en 3 dimensiones de los puntos.

Para conseguir el mejor resultado posible eliminaremos la distorsion radial y tangencial de las imagenes obtenidas con la camara minoru. Ademas se ajusta los angulos y distancias entre las camaras para que los planos de las imagenes queden coplanares. A este ultimo paso se lo conoce con el nombre de rectificacion. (mejorar)

una vez que tenemos las dos imágenes utilizamos la información de las dos imágenes capturadas para encontrar las correspondencias entre ellas, como lo realiza nuestra vista.

Los pasos son:

1 - Remover distorsiones radiales y tangenciales del lente; este proceso se llama (en ingles) undistortion

2 - Ajustar los ángulos y distancias entre imágenes; este proceso se llama rectificación. En este paso ya obtenemos imágenes rectificadas y alineadas

3 - Encontrar iguales puntos en las dos imágenes, esto se llama correspondencia. La salida de este proceso es una imagen llamada mapa de disparidad, donde la disparidad se refiere a la diferencia entre las dos imágenes en un mismo punto.

4 - Por ultimo, podemos pasar el mapa de disparidad a distancias utilizando triangulación; este proceso se llama reproyección y la salida es una imagen con profundidad.

Realizamos estos pasos utilizando fuertemente la librería Open CV. Esta librería nos facilita mucho la realización de cada uno de ellos. Vamos a dedicarnos fuertemente en explicar como utilizamos el mapa de disparidad.

## 4.2 Mapas de disparidad

(agregar imagenes 12-5 pag 417 libro open cv y 12-6 pag 418 del mismo libro, tambien puede ir la imagen 12-4)

Realizando los pasos anteriores sucesivamente obtenemos el mapa de disparidad. Dicha imagen nos permite utilizar la información obtenida de las dos imágenes para saber la profundidad de cada objeto (o punto?). Veamos el concepto que se utiliza para poder armar esta imagen.

Asumamos que la imagen no posee distorsión y esta rectificada. En la figura (12-6) podemos ver la representación de la visión estéreo. Los puntos Ol y Or representan los centros de la cámara izquierda y derecha respectivamente. Los puntos pl y pr son los proyectados en cada imagen respectivamente y corresponden a un unico punto P. Este punto debe ser alineado por las dos

imágenes. Si definimos como  $x_l$  y  $x_r$  como las coordenadas de  $x$  en cada punto respectivamente, podemos definir disparidad como  $d = x_l - x_r$ .

Utilizando geometría podemos calcular la distancia  $Z$ , que es inversamente proporcional al valor de disparidad. Podemos apreciar esto en la figura (12-5). Si la distancia del objeto es mayor, el valor de disparidad va a ser mas chico. Por ejemplo, si vemos un objeto a mucha distancia, la visión de cada imagen va a ver ese objeto sin poca diferencia, es por eso que su disparidad tiende a cero. Ocurre lo contrario si el objeto esta mas cercano. La relación de estas dos variables claramente no es lineal.

Cabe aclarar que para obtener buenos resultados se debe sacar las imágenes al mismo tiempo así como también rectificarlas.

## 4.3 Geometría epipolar

(agregamos figura 12-8 del mismo libro)

Para entender un poco mas como encuentra la relación entre cada punto que representan el mismo objeto vamos a ver la geometría se sucede en las cámaras. La figura (12-8) nos muestra los planos que observa cada cámara y su relación con el punto en común P. Lo mas interesante es ver que para todo punto que se ve en las dos imágenes existe una linea epipolar que atraviesa por un punto epipolo y que tiene su correlación en la otra imagen. Esta es la base de la correspondencia entre cada imagen.

Cada punto en 3D visto desde las cámaras esta contenido en un plano epipolar que intersecta cada imagen en la linea epipolar. Cuando un punto se corresponde con otro en la otra imagen debe apoyarse sobre su linea epipolar correspondiente. Gracias a esta regla, la búsqueda de cada punto con respecto a su correspondencia en la otra imagen se reduce de una búsqueda en dos dimensiones a una busqueda en una dimension , o sea sobre una linea. Esto reduce el computo significativamente.

(redactar mejor , la idea es esa)

introduccion: para que nos sirve el mapa de disparidad

geometria epipolar: grafico geometria epipolar con su explicación

como calculamos (como lo hace libelas) el mapa de disparidad (la teoria)

mencionar libelas (referencia)

teoria 417 open cv implementacion 438 ... 442

## 5 NAVEGACIÓN PARA ROBOT MÓVILES UTILIZANDO LOS MAPAS DE DISPARIDAD

Una vez que obtuvimos el mapa de disparidad debemos correr nuestro algoritmo de navegación. Dado un mapa de disparidad, vamos a explicar paso a paso que haría con ella.

Sobre toda la imagen de disparidad nos quedamos con una región. Esta región esta determinada entre las filas 300 y 340 de la imagen. Este corte horizontal lo utilizamos para poder decidir para que lugar moverse.

Realizamos varios experimentos y llegamos a la conclusión de que es la mejor region para definir el posible camino dada la arquitectura y las condiciones utilizadas. Esta región la llamaremos región de interes, con esta vamos a continuar con los pasos de navegacion.

Cuando obtuvimos esta porción de imagen, vamos a procesarla para poder tomar una decisión con respecto a la navegación. Recorremos toda la imagen por columnas realizando la mediana de los datos por cada columna de información. (tenemos que hacer un grafico de esto.) Si la columna posee algún valor muy lejano se le asigna un valor muy grande parametrizado en el programa. Realizamos distintos experimentos y llegamos a la conclusion que la mediana representa bastante bien la distancia de cada columna. En un principio utilizamos el promedio de los valores pero fue descartado por lo siguiente: Supongamos que poseemos  $1/3$  de la columna con pixeles negros. El promedio de esta columna va a estar cerca del valor de negro, descartando esa columna de informacion. Si utilizamos la mediana, los valores mayores van a estar concentrados en la parte final. Al utilizar los dos elementos del medio vamos a obtener un valor de disparidad real y no va a estar afectado a los pixeles negros. Es por eso que nos inclinamos a utilizar la mediana. Muchas veces experimentamos esto por tener imagenes con poca luminosidad por ejemplo y pudimos evitarlo de esta forma.

Luego de realizado este procesamiento sobre cada columna de la imagen, obtenemos un array de todas las distancias sugeridas por cada columna.

Si no se obtiene una cadena de datos con información suficiente como para elegir un camino a seguir, el algoritmo define una política de escape.

region de interes funcion de mediana justificacion puntos 3D arquitectura de software y de hardware(figura 3 y 4 paper base)

## 6 EXPERIMENTOS

Experimentos.

- arquitectura del robot(paq udp, netbook)
- lugares donde se probo
- limitaciones encontradas (obstaculos grandes, luminosidad, objetos fijos, indoor, distancia de acercamiento)
- experimentos

## 7 TRABAJOS FUTUROS

chamuyo mapas vistos de arriba mejorar algunas cosas utilizar la información obtenidas de las imágenes anteriores para tomar mejores decisiones.(con objetos fijos) thread(futuro cercano y lejano a la vez)

## 8 CONCLUSION

The conclusion goes here.