

Capítulo 1

Segmentación de imágenes con base en color

De acuerdo con [?] la visión es la detección de la luz del mundo. Esa luz empieza cuando un rayo emana desde una fuente hacia un objeto. Cuando la luz choca con el objeto, mucha de la luz es absorbida y la que no es absorbida, nosotros la podemos percibir como el color de la luz, de esta manera, la luz reflejada hace su camino hacia nuestra cámara. Esto es de particular importancia para la visión computacional.

El modelo más simple de cómo sucede el anterior fenómeno es el de la cámara estenopeica o "pinhole". Una cámara estenopeica se puede imaginar como una habitación sin ventanas en donde la luz únicamente entra por una pequeña apertura en el centro de la pared y que proyecta una imagen dentro de la habitación.

Creo que estos dos párrafos deberían ir dentro de la sección sobre el modelo Pinhole

1.1. El modelo de cámara Pinhole

Como se había dicho anteriormente el modelo de cámara estenopeica (o pinhole) es el modelo más simple de todos, en donde se considera que la luz entra desde la escena o desde un punto un punto distante, pero sólo un rayo de luz entra desde un punto en particular, este punto es luego “proyectado” sobre una superficie. Como resultado, la imagen en este plano (también llamado plano proyectivo), está siempre en el foco y el tamaño de la imagen se relaciona a la distancia del objeto dado por un sólo parámetro de la cámara: "la distancia focal". La principal diferencia entre el objeto que aparece en una cámara Pinhole, es que la imagen aparece invertida.

El punto dentro del Pinhole es reinterpretado como el centro de proyección. Para este tipo de cámara, la distancia desde la abertura del Pinhole hacia la

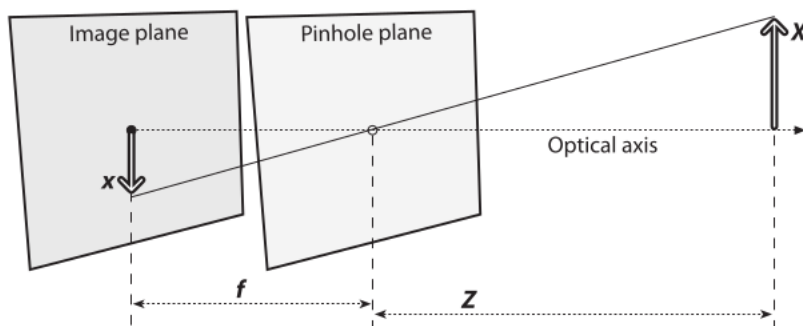


Figura 1.1: Esquema del modelo de cámara Pinhole Si copiaste la imagen de algún lado, menciona de dónde la tomaste.

pantalla, es precisamente la distancia focal. Como puede verse en la Figura 1.

En la Figura 1, f es representada como la distancia focal de la cámara, Z es la distancia de la cámara al objeto, X es la longitud del objeto, y x es la longitud de la imagen proyectada en el plano. Dentro de la figura, se pueden ver dos triángulos semejantes los cuales cumplen que $-x/f = X/Z$, o expresado de otra manera:

$$x = -f \frac{X}{Z}$$

A fin de obtener ecuaciones equivalentes, pero sin signos negativos (correspondientes a la inversión de la imagen) se propone hacer un rearrreglo en el cual se coloca al frente de un centro de proyección al plano proyectivo. Tal y como se ve en la Figura 2.

Revisa la forma de correcta de poner comillas, aunque en textos académicos se estila más usar cursivas en lugar de comillas Con este nuevo cambio, el punto en el Pinhole es reinterpretado como “el centro de proyección”, así cada rayo de luz deja un punto en el objeto distante y se dirige hacia el centro de proyección. El punto de intersección del plano proyectivo y el eje óptico es conocido como *el punto principal*. El nuevo plano de imagen frontal es equivalente al anterior plano proyectivo, el el cual la imagen del objeto distante tiene exactamente el mismo tamaño que en el esquema mostrado en la Figura 1, la diferencia es que en este caso la imagen no queda invertida, por lo tanto la relación de los triángulos quedaría de la siguiente manera:

$$x/f = X/Z$$

Para poner letras en cursiva usa el comando `textit` Uno podría pensar Se podría pensar que “el punto principal” es equivalente al centro de la imagen, sin embargo, este centro usualmente no está en el eje óptico, es por eso que se introducen dos nuevos parámetros, c_x y c_y para modelar un posible desplazamiento

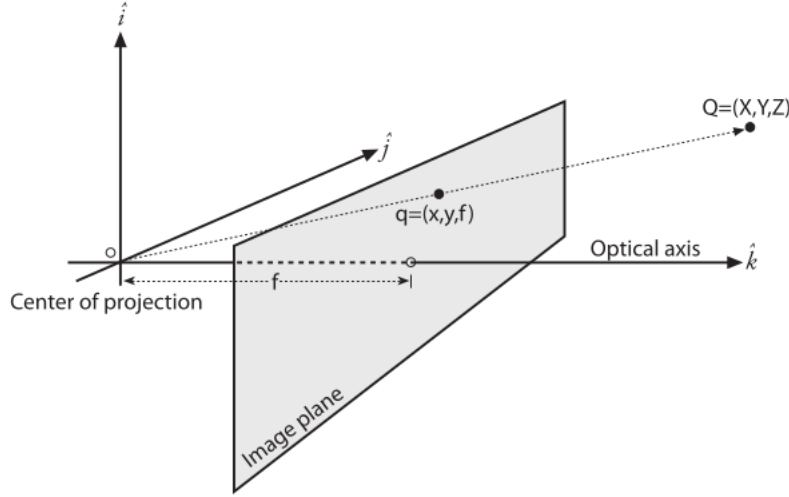


Figura 1.2: Rearreglo de una cámara Pinhole **Todo debe estar en español**

(perpendicular al eje óptico) del centro de coordenadas en el plano de proyección. El resultado es que un punto Q en el mundo físico cuyas coordenadas son (X,Y,Z) es proyectado dentro de la pantalla en una localización de píxeles dada por (x_{screen}, y_{screen}) de acuerdo con las siguientes ecuaciones:

$$x_{screen} = f_x \left(\frac{X}{Z} \right) \quad , \quad y_{screen} = f_y \left(\frac{Y}{Z} \right)$$

Y aquí dónde entraron c_x y c_y

Nótese que se han introducido dos diferentes distancias focales, la razón es porque generalmente los píxeles son rectangulares.

1.2. Corrección de la distorsión

Desafortunadamente una verdadera cámara pinhole no es una muy buena forma de obtener imágenes porque no reúne suficiente luz para crear una imagen aproximada a la real. Es por eso que los ojos y las cámaras usan lentes para reunir más luz que la que estaría disponible en un solo punto.

1.2.1. Geometría básica proyectiva

La relación que mapea los puntos Q_i en el mundo físico con coordenadas (X_i, Y_i, Z_i) a los puntos en el plano de proyección con coordenadas (x_i, y_i) es

4CAPÍTULO 1. SEGMENTACIÓN DE IMÁGENES CON BASE EN COLOR

llamado una "transformación proyectiva". Cuando se trabaja con estas transformaciones es conveniente usar "coordenadas homogéneas". El plano de imagen es el espacio proyectado y tiene dos dimensiones, con lo que se puede representar puntos en el plano como un vector tridimensional $q = (q_1, q_2, q_3)$ Si tengo un plano, por qué usaría tres dimensiones?. Una forma de hacer un arreglo con los parámetros que definen a la cámara f_x, f_y, c_x y c_y dentro de una matriz de 3x3 el cual es llamado matriz de parámetros intrínsecos. La proyección q de los puntos del mundo físico dentro de la cámara es resumida dentro de la siguiente forma ...del mundo físico en el plano de la imagen se puede calcular como:.

$$q = MQ$$

donde

$$q = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

1.2.2. Distorsión de las lentes

La estimación de los parámetros internos es importante en la calibración de las cámaras [?].

Dada la manufactura y diversos factores, las lentes no son perfectas, es por eso que se introducen dos nuevos conceptos a la distorsión de las lentes, las "Distorsiones radiales" y las "Distorsiones tangenciales". Las primeras surgen como resultado de la forma de las lentes y las segundas como resultado del proceso de ensamblado de la cámara.

Te falta agregar todas tus referencias