Usando geometría proyectiva para corregir una cámara.

Parte I (En el Volumen 28.2 continúa el desarrollo de este artículo)

No hay nada particularmente profundo en este problema o en su solución, pero espero que muestre el placer que se puede encontrar cuando usamos nuestra intuición matemática.

David Austin[1]

Introducción

No hace mucho tiempo, recibí la llamada de un ingeniero que trabaja para la industria automotriz que necesitaba corregir la imagen producida por una cámara que fue mal alineada. Coincidentemente, la semana anterior había leído una de las columnas que escribe Jim Blinn sobre gráficos de computadora, donde planteaba una explicación simple y elegante para este problema.

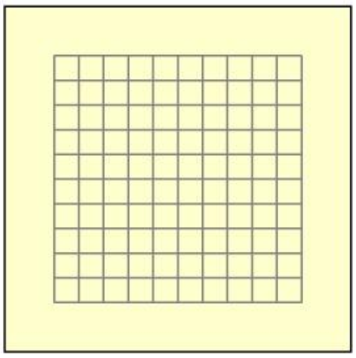
En este artículo, describiré el problema y mostraré como las ideas de la columna de Blinn nos dan la solución. No hay nada particularmente profundo en este problema o en su solución, pero espero que muestre el placer que se puede encontrar cuando usamos nuestra intuición matemática.

Primero, el problema. La compañía estaba construyendo un nuevo sistema de suspensión para automóviles y usaba una cámara para medir el movimiento de la cubierta asociada con el sistema de suspensión en sus pruebas de laboratorio. La cámara sigue el movimiento de la cubierta registrando periódicamente la posición de ciertos puntos de la misma.

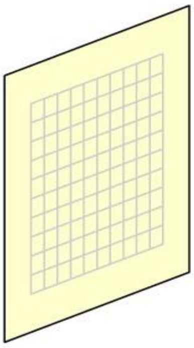
[1]Impreso con autorización del autor.

Para obtener la precisión requerida por estas pruebas es necesario que la cámara este perfectamente alineada en un plano paralelo al de la cubierta, pero en la práctica esto es muy difícil de lograr.

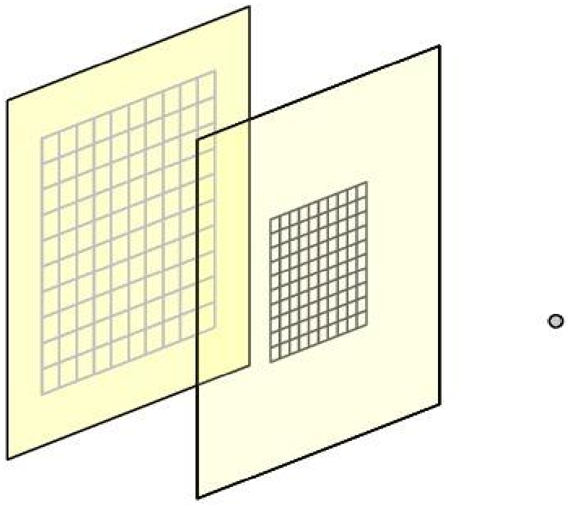
Aquí vemos una versión simplificada del problema. Supongamos que queremos tomarle una fotografía a este plano,



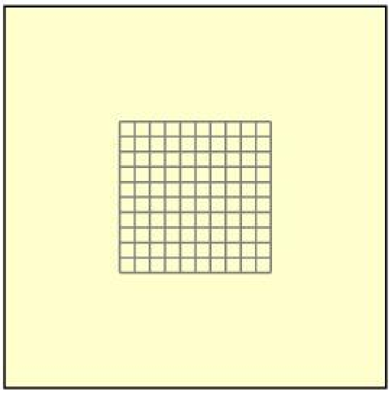
que está ubicado en el espacio 3-dimensional.



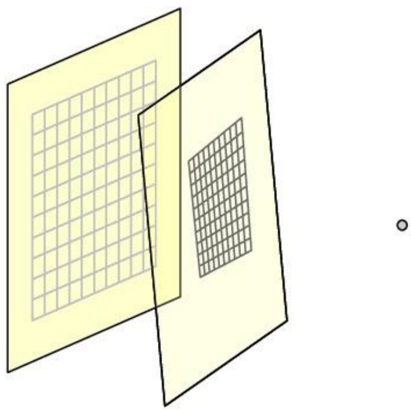
Si la cámara está perfectamente alineada con el plano



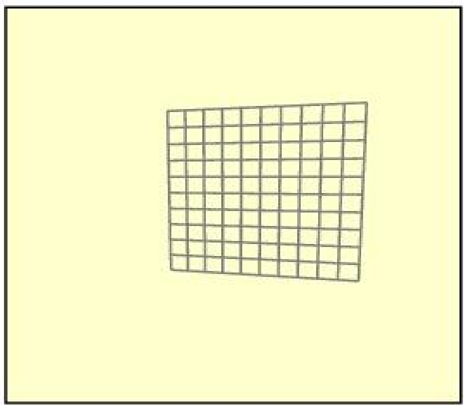
entonces la imagen va a ser registrada fielmente pero en escala.



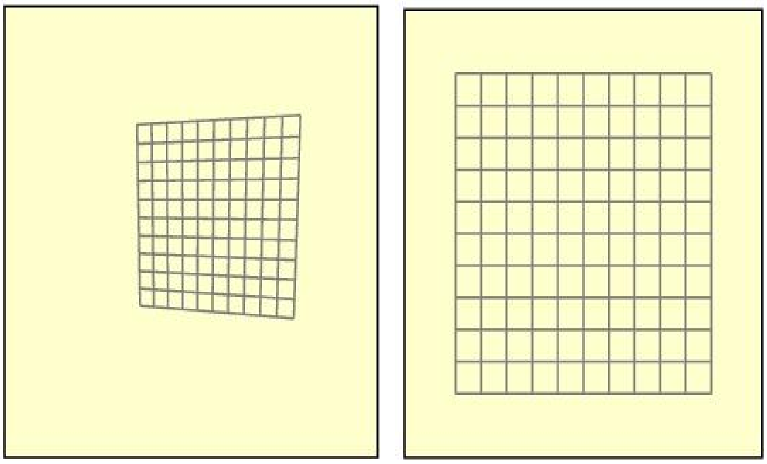
Sin embargo, si el plano de la cámara no es paralelo al plano,



entonces la imagen registrada por la cámara será distorsionada.

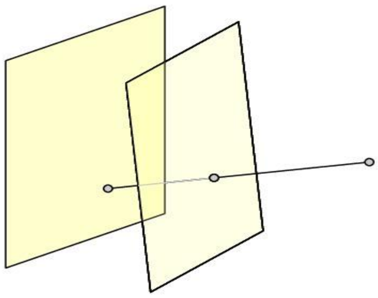


Nuestro objetivo es tomar la imagen distorsionada registrada por la cámara, como se muestra abajo a la izquierda, y reconstruir la imagen original como si la cámara hubiese estado correctamente alineada, como en la representación ubicada abajo a la derecha.

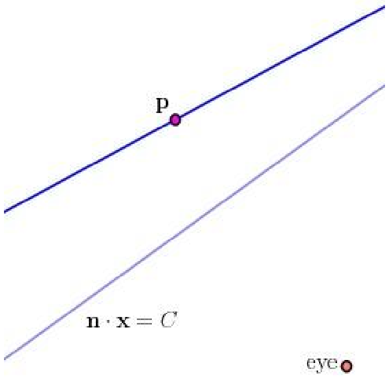


Transformaciones proyectivas.

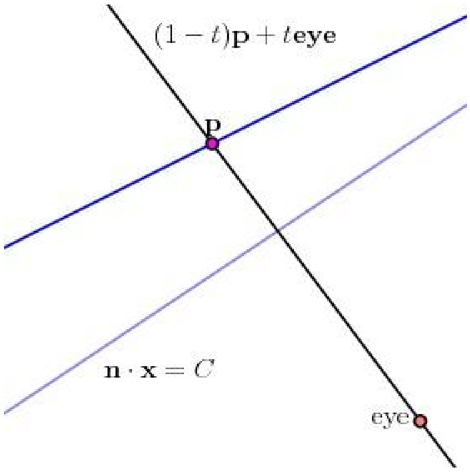
Para empezar, describiremos la función que proyecta un plano sobre el otro. No importa si consideramos la proyección del plano sobre el plano de la cámara o la proyección inversa.



Supongamos que proyectamos el punto **p** sobre el plano dado por la ecuación $\mathbf{n}\cdot\mathbf{x}=C$. La siguiente figura muestra una versión 2-dimensional de este problema



Parametrizamos los puntos de la recta que pasa por **p** y el punto *eye* (ojo) con la ecuación **r**=(1−*t*)**p**+*t* **eye** y nos preguntamos para que valores de *t* estos puntos yacen en el plano **n·x**=*C*.

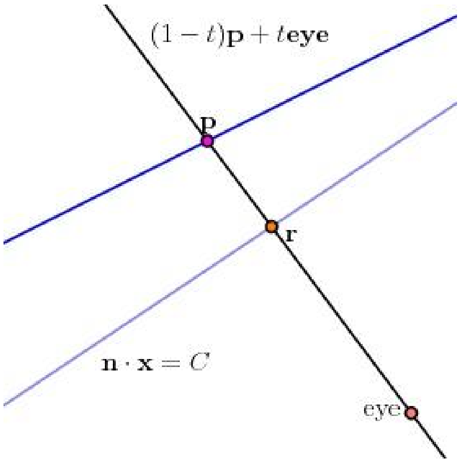


La condición que encontramos sobre *t* es la siguiente

$$t = \frac{C - \boldsymbol{n}\cdot\boldsymbol{p}}{\boldsymbol{n}\cdot\boldsymbol{eye} - \boldsymbol{n}\cdot\boldsymbol{p}},$$

y así

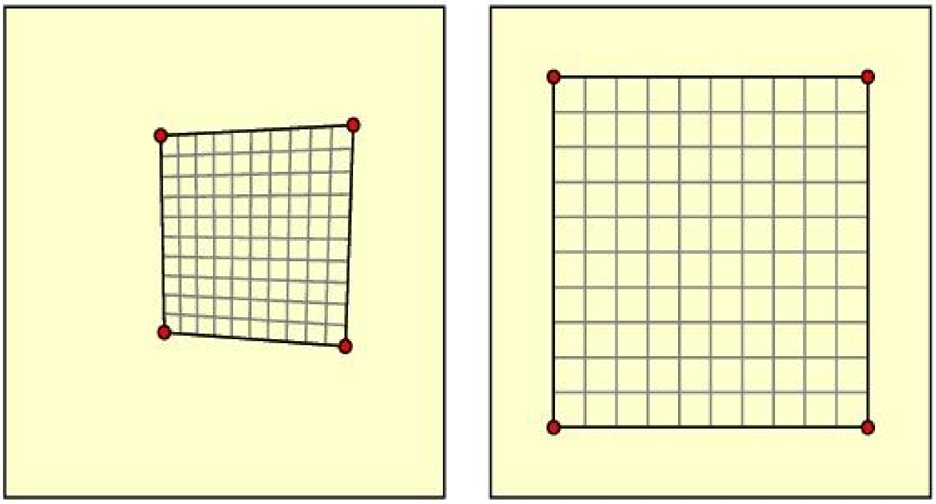
$$\mathbf{r} = \frac{\mathbf{n} \cdot \mathbf{eye} - C}{\mathbf{n} \cdot \mathbf{eye} - \mathbf{n} \cdot \mathbf{p}}\,\mathbf{p} + \frac{C - \mathbf{n} \cdot \mathbf{p}}{\mathbf{n} \cdot \mathbf{eye} - \mathbf{n} \cdot \mathbf{p}}\,\mathbf{eye}$$



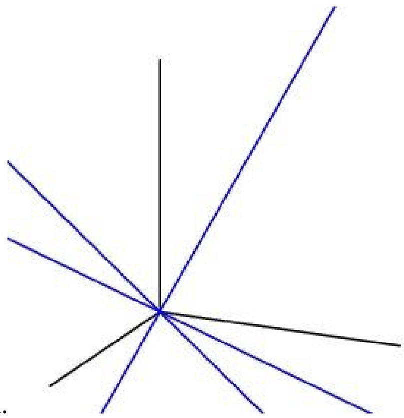
Si suponemos que tenemos sistemas coordenados en los planos, la función tiene la forma

$$(x, y) \mapsto (u, v) = \left(\frac{ax + by + c}{gx + hy + j}, \frac{dx + ey + f}{gx + hy + j}\right).$$

Dadas nuestras dos imagenes, elegiremos cuatro puntos de la imagen distorsionada y los correspondientes cuatro puntos en la imagen corregida y a partir de estos datos, reconstruiremos la función.

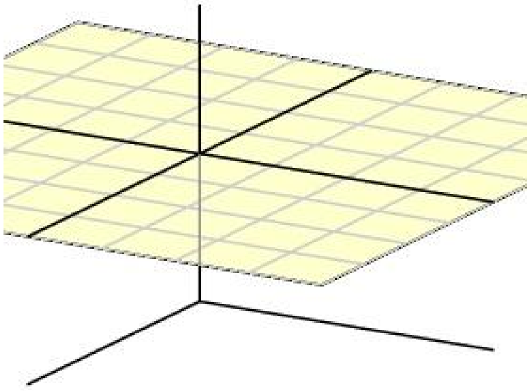


La forma de esta función, conocida como *transformación proyectiva*, es particularmente llamativa, ya que aunque sus componentes son funciones racionales, los numeradores y los denominadores son lineales. Por esta razón, es natural pensar en esta función como una trasformación del plano proyectivo en el conjunto de rectas

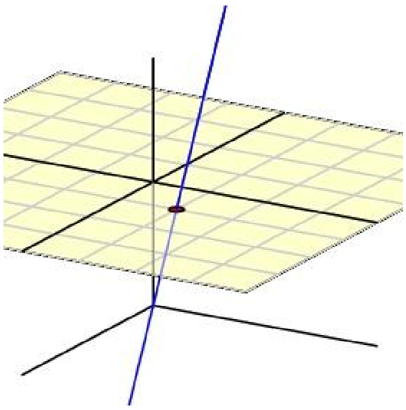


de $\boldsymbol{R}^3$ que pasan por el origen.

Si pensamos al plano Euclido dentro de $\boldsymbol{R}^3$ como el plano z=1,



vemos que el plano Euclideo está dentro del plano proyectivo ya que una recta interseca el plano z=1 en a lo sumo un punto.



Los puntos del plano proyectivo se describen usando coordenadas homogéneas: denotamos una recta en el plano proyectivo por [x,y,z] donde (x,y,z) es un punto de la recta. Observemos que cualquier múltiplo escalar de (x,y,z) es también un punto de esta misma recta, lo que significa que las coordenadas homogéneas se definen salvo un múltiplo escalar, es decir que [x,y,z]=[wx,wy,wz].

Notemos también que los puntos del plano Euclideo son de la forma $[x,y,1]$. Los otros puntos del plano proyectivo tienen la forma $[x,y,0]$ y corresponden a las líneas que están en el plano *xy*. Estos puntos se llaman *puntos en el infinito*, ya que son límite de una sucesión no acotada de puntos en el plano Euclideo.

Si usamos coordenadas homogéneas podemos escribir nuestra transformación de la siguiente forma,

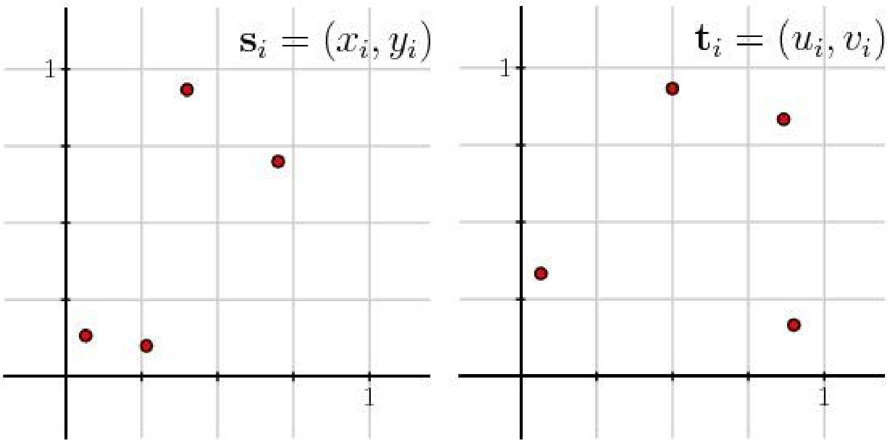
$$[x,y,1] \mapsto [u,v,1] = [ax+by+c,\ dx+ey+f,\ gx+hy+j],$$

y observemos que es convenientemente lineal, lo que será de gran utilidad. Para determinar la transformación necesitamos determinar la matriz formada por las constantes $a,b,\ldots,j$. Esta matriz, sin embargo está definida salvo un múltiplo escalar ya que usamos coordenadas homogéneas.

Una solución ingénua.

Nuestro problema no es tan difícil, ya que admite una solución directa que describiremos a continuación.

La imagen de la izquierda se llama *espacio de partida* y el de la derecha *espacio de llegada.* Supongamos que tenemos cuatro puntos $\mathbf{s}_i=(x_i,\ y_i)$ en el espacio de partica y sus correspondientes imágenes a través de la transformación $\mathbf{t}_i=(u_i,\ v_i)$ en el espacio de llegada. Nuestro objetivo es determinar la transformación.



Tenemos que

$$[x_i\ y_i\ 1]\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & j \end{bmatrix} = w_i\ [u_i\ v_i\ 1],$$

donde $w_i = gx_i + h\,y_i + j$ es un factor de reescalamiento. En otras palabras,

$$\begin{aligned} ax_i + by_i + c &= w_i u_i \\ dx_i + ey_i + f &= w_i v_i \\ gx_i + hy_i + j &= w_i \end{aligned}$$

o

$$\begin{aligned} ax_i + by_i + c &= (gx_i + hy_i + j)u_i \\ dx_i + ey_i + f &= (gx_i + hy_i + j)v_i \end{aligned}$$

o

$$x_i + y_i b + c - x_i u_i g - y_i u_i h - u_i j = 0$$
$$x_i\, d + y_i e + f - x_i v_i g - y_i v_i h - v_i j = 0$$

o

$$\begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i u_i & -y_i u_i & -u_i \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i v_i & -y_i v_i & -v_i \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ j \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Con los cuatro pares de puntos $\{\mathbf{s}_i\}$ y $\{\mathbf{t}_i\}$, obtenemos un sistema homogéneo de 8×9 ecuaciones que nos determinaran las constantes *a*, *b*,…,*j* salvo un múltiplo escalar. Esto explica porqué necesitamos cuatro puntos para determinar la transformación.

La transformación resultante se denota **Mst** y está definida por la matriz

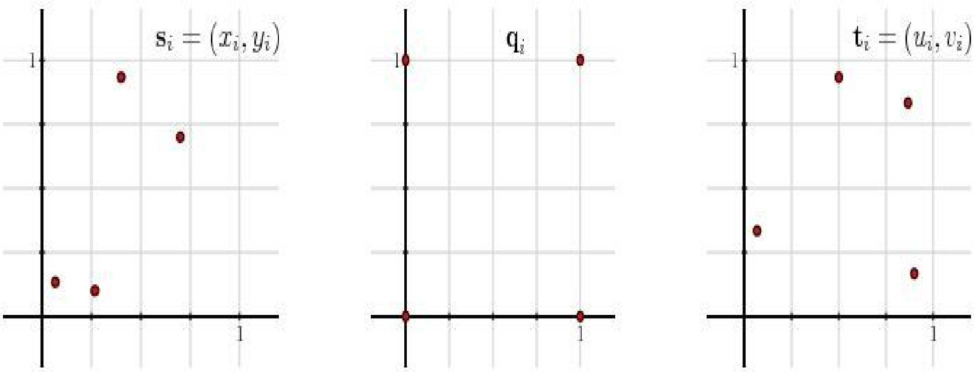
$$\mathbf{M}_{\mathbf{st}} = \begin{bmatrix} a & d & g \\ b & e & g \\ c & f & j \end{bmatrix}.$$

Esta es la solución obvia del problema, aunque computacionalmente es desagradable resolver un sistema de 8x9 ecuaciones no es demasiado difícil. La crítica más seria que se puede hacer a esta solución es esencialmente que hemos perdido de vista la intuición geométrica que tiene el problema.

Haciéndolo en dos pasos

Una solución mas interesante fue propuesta por Paul Heckbert. En lugar de encontrar la transformación que convierte los puntos **s***i* directamente en los $\mathbf{t}_i$, construyamos un proceso de dos etapas con un conjunto intermedio de puntos $\mathbf{q}_i$ que son los vértices del cuadrado unidad. Es decir, mapearemos

$$\mathbf{s}_i \quad \mapsto \quad \mathbf{q}_i \quad \mapsto \quad \mathbf{t}_i$$



Definimos transformaciones **Msq** y **Mqt** tales que **s***i* **Msq** $= \mathbf{q}_i$ y $\mathbf{q}_i$ **Mqt** $= \mathbf{t}_i$. Nuestro objetivo final es la composición

$$\mathbf{Mst}=\mathbf{MsqMqt}.$$

Observemos sin embargo que $\mathbf{M}_{\mathbf{sq}} = \mathbf{M}_{\mathbf{qs}}^{-1}$ y encontrar $\mathbf{M}_{\mathbf{qs}}$ es esencialmente lo mismo que encontrar $\mathbf{M}_{\mathbf{qt}}$.

La idea de Heckbert es buena porque los vértices del cuadrado unidad tienen muchos ceros: $(0,0)$, $(1,0)$, $(1,1)$ y $(0,1)$, lo que nos lleva a tener que resolver un sistema más sencillo. Por ejemplo, para $\mathbf{q}_0 = (0,0)$, dos de las ecuaciones son

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -u_0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -v_0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ j \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

De hecho Blinn muestra cómo resolver estas ocho ecuaciones explícitamente en términos de las coordenadas, $\mathbf{t}_i = (u_i, v_i)$ para encontrar la transformación **Mqt**. La solución no es muy linda, pero no requiere ningún esfuerzo computacional. Así,

$$\mathbf{Mst} = \mathbf{Msq}\ \mathbf{Mqt} - \mathbf{M}_{\mathbf{qs}}^{-1}\mathbf{Mqt}.$$

Como podemos multiplicar la matriz **Mst** por un escalar y obtener la misma transformación proyectiva, reemplazamos $\mathbf{M}_{\mathbf{qs}}^{-1}$ por la adjunta $\mathbf{M}_{\mathbf{qs}}^{*}$, la cual difiere de la inversa por la multiplicación del determinante de **Mqs**. Esto nos da

$$\mathbf{Mst} = \mathbf{M}_{\mathbf{qs}}^{*}, \mathbf{Mqt}.$$

Colaboración de la Dra. Carina Boyallián. Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba.