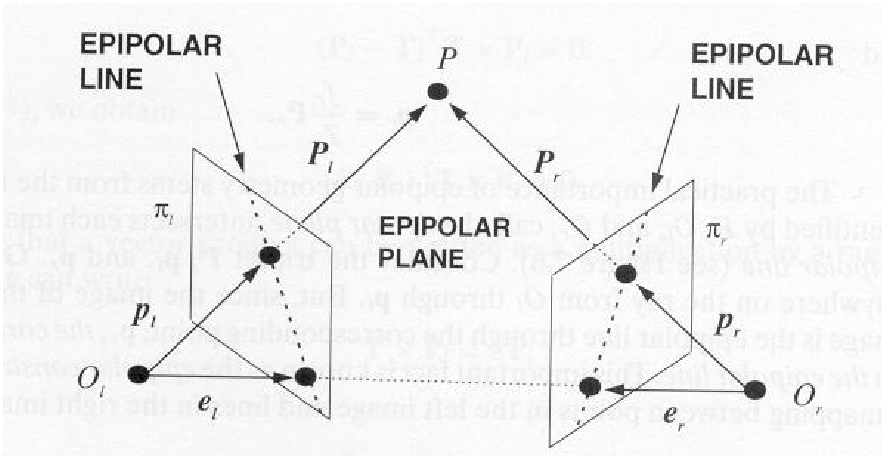
Geometría epipolar

Ahora estudiaremos la geometría de la visión estérea en su total generalidad. Esto nos permitirá clarificar que información es necesaria para llevar a cabo la búsqueda de características correspondientes a lo largo de líneas solamente.

La geometría del estéreo denominada geometría epipolar, se muestra en la siguiente figura



La figura muestra las dos cámaras de tipo "pinhole", sus centros de proyección, O_l y O_r, y sus planos imagen o retinas en coordenadas normalizadas. Las longitudes focales se notan como f_i y f_d . Como es normal, cada cámara identifica un sistema de referencia 3D, el origen del cual se sitúa en el centro de proyección de la cámara y el eje Z con el eje de óptico. Los vectores $\mathbf{P}_i=[X_i,Y_i,Z_i]^T$ y $\mathbf{P}_d=[X_d,Y_d,Z_d]^T$ se refieren al mismo punto 3D, **P**, como vectores en los sistemas de referencia de la cámara izquierda y derecha respectivamente. Los vectores $\mathbf{p}_i=[x_i,y_i,z_i]^T$ y $\mathbf{p}_d=[x_d,y_d,z_d]^T$ definen las proyecciones del punto *P* en la imagen izquierda y derecha respectivamente y están expresados en su correspondiente sistema de referencia. Evidentemente, para todos los puntos de las imágenes tenemos $z_i=f_i$ o $z_d=f_d$ de acuerdo a la imagen que sea.

Los sistemas de referencia de las cámaras izquierda y derecha están relacionados a través de los parámetros extrínsecos. Estos definen una transformación rígida en el espacio 3D definida por un vector de traslación $T=O_d-O_i$ y una matriz de rotación R. Dado un punto P en el espacio la relación entre P_i y P_d es por tanto

$$P_d=R(P_i - T)$$

El nombre de geometría epipolar es debido a que los puntos en los cuales la recta que une los centros de proyección de las cámaras corta a los planos de proyección se llaman epipolos.

Notaremos por $\mathbf{e_i}$ y $\mathbf{e_d}$ el epipolo izquierdo y derecho respectivamente. Por construcción ambos epipolos representan la proyección en su correspondiente plano la imagen del centro de proyección de la otra cámara. En el caso de que uno de los planos imagen sea paralelo a la recta que une los centros de proyección, el epipolo de ese plano estará situado en el infinito.

La relación entre un punto del espacio 3D y su proyección se describe por las ecuaciones usuales de proyección de perspectiva, que en forma vectorial se escriben

$$\mathbf{p}_i = \frac{f_i}{Z_i}\mathbf{P}_i \, , \; \mathbf{p}_d = \frac{f_d}{Z_d}\mathbf{P}_d$$

La importancia práctica de la geometría epipolar arranca del hecho que el plano identificado por $\mathbf{P}$, $\mathbf{O_i}$, $\mathbf{O_d}$ llamado plano epipolar, interseca cada imagen en una línea llamada línea epipolar. Consideremos el triple $\mathbf{P}$, $\mathbf{p_i}$ y $\mathbf{p_d}$. Dado $\mathbf{p_i}$ $\mathbf{P}$ puede caer en cualquier punto del rayo definido por $\mathbf{O_i}$ y $\mathbf{p_i}$. Pero dado que la imagen de este rayo en la imagen derecha es la línea epipolar a través del punto correspondiente $\mathbf{p_d}$ dicho punto debe estar sobre la línea epipolar. Esta correspondencia establece una aplicación entre puntos de la imagen izquierda y rectas de la imagen derecha y viceversa. Una consecuencia de esta correspondencia es, que dado que todos los rayos pasan por construcción por el centro de proyección, todas las rectas epipolares deben pasar por el epipolo.

Por tanto si determinamos la aplicación entre puntos de la imagen izquierda(derecha) y las rectas epipolares de la imagen derecha(izquierda), podemos restringir la búsqueda para el emparejamiento de $\mathbf{p_i}$ a lo largo de la línea epipolar correspondiente.

Así pues la búsqueda de las correspondencias se reduce a un problema 1D. Alternativamente este conocimiento también se puede usar para verificar sí una potencial pareja de puntos correspondientes, lo son de verdad o no. Esta técnica es normalmente una de las más efectivas para detectar las posibles falsas correspondencias debidas a oclusión.

Restricción epipolar: Los puntos correspondientes deben estar sobre líneas epipolares conjugadas.

Ahora queda una cuestión obvia e importante, ¿cómo podemos estimar la geometría epipolar? O equivalentemente ¿cómo podemos establecer una aplicación entre puntos de una imagen y líneas epipolares de la otra?. A continuación estudiaremos técnicas para llevar a cabo dicha estimación.

Matriz esencial y Matriz fundamental

La ecuación del plano epipolar a través de $\mathbf{P}$ puede escribirse como la condición co-planar de los vectores $\mathbf{P}_i$,$\mathbf{T}$ y $\mathbf{P_i - T}$, o $(\mathbf{P_i - T})^T (\mathbf{T} \times \mathbf{P_i}) = 0$. Usando la relación que liga a los vectores $\mathbf{P_i}$ y $\mathbf{P_d}$ obtenemos $(\mathbf{R}^T \mathbf{P_d})^T (\mathbf{T} \times \mathbf{P_i}) = 0$. Teniendo en cuenta que el producto vectorial de dos vectores se puede escribir como la multiplicación de una matriz antisimetrica por un vector, tenemos $\mathbf{T} \times \mathbf{P_i} = [\mathbf{T}]_x \mathbf{P_i} = S\mathbf{P_i}$ donde

$$[T]_\times = S = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$$

obtenemos $\mathbf{P}_d^T \mathbf{E} \mathbf{P}_i = 0$ con $\mathbf{E}=\mathbf{R}S$.

Notemos que por construcción $\mathbf{E}$ siempre tiene rango igual a 2 y se puede probar que su dos autovalores distintos de cero son iguales. La matriz $\mathbf{E}$ se denomina la matriz esencial y establece una unión natural entre la restricción epipolar y los parámetros extrínsecos del sistema estéreo. Más adelante estudiaremos como recuperar los parámetros extrínsecos de la matriz esencial. Ahora si consideramos las ecuaciones vectoriales de la perspectiva y sustituimos en la ecuación anterior obtenemos la ecuación que liga las proyecciones del punto $\mathbf{P}$ en ambas planos imagen y dividimos por Z_i Z_d obtenemos

$$\mathbf{p}_d^T \mathbf{E} \mathbf{p}_i = 0$$

El vector que representa $\mathbf{u}_d=\mathbf{E}\mathbf{p}_i$ puede ser interpretado como el vector director de la recta en que se proyecta, sobre la imagen derecha, el rayo definido por $\mathbf{O}_i\mathbf{P}$. Así pues, y a través de la matriz esencial se establece una aplicación entre los puntos de una imagen y las rectas epipolares de la otra (recordar como se han definido las rectas epipolares asociada a un punto $\mathbf{P}$). Usando la notación vectorial introducida, la ecuación anterior se escribe como $\mathbf{p}_d^T\mathbf{u}_d=0$, que establece la condición de incidencia de que uno de los puntos siempre se encuentra sobre la línea epipolar definida por el otro.

Hasta ahora hemos trabajado en coordenadas de los sistemas de referencia asociados a las cámaras, sin embargo, cuando medimos los puntos proyección estos están medidos en términos de píxeles, por tanto para poder sacar el máximo partido de la matriz esencial será necesario conocer la transformación desde coordenadas de la cámara a coordenadas de píxeles, es decir, conocer los parámetros intrínsecos. Esta restricción puede ser eliminada pero a costa de pagar un precio sobre la información recuperada.

La matriz fundamental

Ahora mostramos que la aplicación entre puntos y líneas puede obtenerse a partir de puntos correspondientes solamente sin necesidad de información a priori sobre el sistema estéreo.

Sean $\mathbf{K}_i$ y $\mathbf{K}_d$ las matrices de los parámetros intrínsecos de las cámaras izquierda y derecha respectivamente. Si notamos por $\overline{\mathbf{p}}_i$ y $\overline{\mathbf{p}}_d$ los puntos en coordenadas píxel correspondientes a $\mathbf{p}_i$ y $\mathbf{p}_d$ respectivamente, tenemos

$$\mathbf{p}_i = K_i^{-1}\overline{\mathbf{p}}_i \quad \text{y} \quad \mathbf{p}_d = K_d^{-1}\overline{\mathbf{p}}_d$$

y sustituyendo en la ecuación de la matriz esencial obtenemos $\overline{\mathbf{p}}_i^T\mathbf{F}\overline{\mathbf{p}}_d = 0$ siendo $\mathbf{F} = K_d^{-T}\mathbf{E}K_i^{-1} = K_d^{-T}\mathbf{R}SK_i^{-1}$, $\mathbf{F}$ se denomina la matriz fundamental. Al igual que con $\mathbf{E}\mathbf{p}_i$, $\mathbf{F}\overline{\mathbf{p}}_i$ puede ser interpretado como el vector de la recta epipolar proyectiva correspondiente al punto $\overline{\mathbf{p}}_i$, $\overline{\mathbf{u}}_d = \mathbf{F}\overline{\mathbf{p}}_i$.

La mayor diferencia entre las ecuaciones en términos de la matriz esencial y la matriz fundamental es que la matriz esencial esta definida en términos de vectores definidos en los sistemas de referencia de las cámaras, mientras que la matriz fundamental esta definida en términos de vectores definidos en términos de coordenadas píxeles de los planos de proyección. Consecuentemente, si se estima la matriz fundamental a partir de puntos en correspondencia en coordenadas píxel se puede reconstruir la geometría epipolar sin absolutamente ninguna información sobre los parámetros intrínsecos o extrínsecos.

Todo lo anterior indica que es posible establecer la correspondencia entre los puntos de una imagen y sus correspondientes líneas epipolares sin ningún conocimiento a priori de los parámetros del sistema estéreo.

La matriz fundamental, al igual que la matriz esencial, tampoco es una matriz de rango completo, siendo de rango igual a 2 ya que la matrices de los parámetros intrínsecos son de rango completo. La matriz fundamental codifica información sobre los parámetros tanto intrínsecos como extrínsecos.

Una consecuencia muy importante de las anteriores propiedades es que si conocemos la calibración de un sistema estéreo o de una cámara, (parámetros intrínsecos y parámetros extrínsecos) conocemos toda la información necesaria para obtener la geometría epipolar y la aplicación que liga puntos de una imagen con sus correspondientes rectas epipolares.

Derivación geométrica

La aplicación de un punto en una imagen a su correspondiente línea epipolar en la otra imagen puede descomponerse en dos pasos. En el primero el punto x se aplica en otro punto x' que cae en la línea epipolar l_i'. Este punto x' es una pareja potencial para el punto x. En el segundo paso la línea epipolar se obtiene a partir del punto x' y el epipolo e'.

La transferencia del punto x se realiza con la ayuda de un plano del espacio π. El rayo definido por el centro de la cámara y el punto x incide en el plano π en el punto $\mathbf{X}$. Este punto $\mathbf{X}$ se proyecta en la otra imagen en el punto $x'=\mathbf{P}'\mathbf{X}$. Es obvio que el punto x' pertenecerá a la línea epipolar de x ya que $\mathbf{X}$ pertenece al rayo definido por x. En este proceso hemos definido una pareja de punto correspondientes $x \leftrightarrow x'$ a través de una homografía inducida por la elección del plano π, $x'=\mathbf{H}_\pi x$.

Usando la construcción de la línea epipolar se tiene

$$l = e' \times x' = e' \times \mathbf{H}_\pi x = [\, e']_\times \mathbf{H}_\pi x = \mathbf{F}\mathbf{x}$$

de lo que se puede enunciar

La matriz fundamental $\mathbf{F}$ puede ser escrita como $\mathbf{F}=[\, e']_\times \mathbf{H}_\pi$ donde $\mathbf{H}_\pi$ es la aplicación de transferencia de una imagen a la otra via un plano π. Además, y dado que $[\, e']_\times$ tiene rango 2 y $\mathbf{H}_\pi$ es no singular $\mathbf{F}$ tiene rango 2.

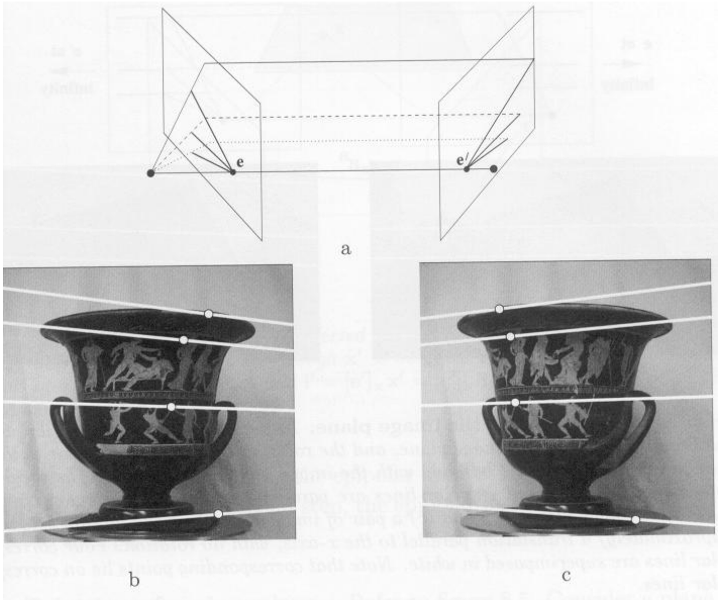
Localización de los epipolos a partir de F

Vamos a establecer la relación que liga a la matriz $\mathbf{F}$ con los epipolos de los planos retinales. Sea $\overline{\mathbf{e}}_i$, el epipólo del plano izquierdo. Dado que es un punto por el que pasan todas las rectas epipolares asociadas a los puntos de la imagen derecha, se verifica que $\overline{\mathbf{p}}_d^T \mathbf{F} \overline{\mathbf{e}}_i = 0$ para todo $\overline{\mathbf{p}}_d$. Pero dado que $\mathbf{F}$ no es idénticamente nula, esto tan solo será posible si se verifica que $\mathbf{F}\overline{\mathbf{e}}_i = \mathbf{0}$. Dado que $\mathbf{F}$ es de rango igual a 2, el epipolo $\overline{\mathbf{e}}_i$ se puede por tanto calcular como el autovector asociado al autovalor nulo de la matriz **F.** Haciendo un razonamiento similar se puede mostrar que el epipolo de la imagen derecha es el autovector asociado al autovalor nulo de la matriz $\mathbf{F}^T$, $\mathbf{F}^T \overline{\mathbf{e}}_d = 0$.

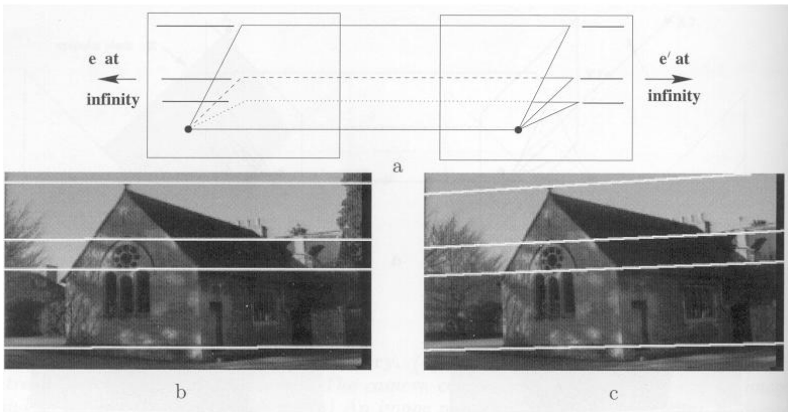
Las consideraciones para la matriz $\mathbf{E}$ son similares a las realizadas para la matriz $\mathbf{F}$.

Las siguientes figuras muestran las líneas epipolares y el epipolo de las imágenes en distintas posiciones de la cámara respecto del objeto:

a) la parte superior de la figura (a) muestra la situación relativa de las cámaras. La parte inferior muestra las imágenes de cada cámara y líneas epipolares conjugada (b) y (c).



La siguiente imagen muestra dos planos prácticamente paralelos de la misma escena. Obsérvese que las líneas epipolares son paralelas, lo que indica que los epipolos están en el infinito,



Relación entre F y las matrices de las cámaras

La relación entre la matriz fundamental F definiendo la geometría epipolar de dos cámaras y las matrices de proyección de dichas cámaras es muy estrecha. Es posible definir la matriz fundamental en función de las matrice de las cámaras y viceversa.

En un razonamiento geométrico, el rayo **X** que se proyecta en un punto de una retina **x**, puede considerarse la solución de la ecuación **PX**= **x** y la notamos $\mathbf{X}(\lambda)= \mathbf{P}^{+}\mathbf{x}+\lambda\mathbf{C}$. La notación $\mathbf{P}^{+}$ indica la inversa generalizada de la matriz **P**, es decir $\mathbf{P}^{+}\mathbf{P}=\mathbf{I}$ debemos recordar que la matriz **P** es rectangular de dimensiones 3x4. Por tanto la recta proyección de dicho rayo sobre la otra retina estará definida por la proyección de dos puntos de dicho rayo en la otra retina (λ=0 y $\lambda=\infty$) . De la expresión $\mathbf{x'}= \mathbf{P'P^{+}x}$ tenemos un punto y de $\mathbf{P'C=e'}$ tenemos el otro. Así pues, la recta estará definida por el vector **l**= $[\mathbf{e'}]_x\mathbf{P'P^+x}$, luego la matriz fundamental será $\mathbf{F}=[\mathbf{e'}]_x\mathbf{P'P^+}$.

La relación inversa que define las matrices de proyección a partir del conocimiento de la matriz fundamental es más ambigua. Es decir, no existe un único conjunto de matrices de proyección asociada a una matriz fundamental. Puede demostrase que cualquier transformación proyectiva del mundo modifica las matrices de proyección pero no modifica la matriz fundamental asociada a dos imágenes. Por tanto y para eliminar la ambigüedad se elige la matriz de proyección de la primera imagen como la matriz de proyección canónica pin-hole y la otra matriz de proyección se calcula en función de esta hipótesis. La solución que se obtiene es correcta salvo una transformación proyectiva global en el mundo.

Resultado. Las matrices **P** y **P'** de las cámaras correspondientes a un matriz fundamental **F** pueden ser escogidas como $\mathbf{P}=[\mathbf{I}|\mathbf{0}]$ y $\mathbf{P'}=[[\mathbf{e'}]_x\mathbf{F} \mid \mathbf{e'}]$.

Puede observarse como los únicos elementos que intervienen en la definición de las matrices son los invariantes geométricos estudiados: epipolos y matriz fundamental.

En esta elección de las matrices **P** y **P'** hemos de notar que el menor 3x3 de **P'** $[\mathbf{e'}]_x\mathbf{F}$ tiene determinante cero, lo cual nos da una cámara con centro en el plano del infinito π_∞ .

La familia más general de posibles cámaras viene dad por el siguiente resultado

Resultado. La expresión general para un par de cámaras canónicas asocadas a una matriz fundamental F esta dada por

$$\mathbf{P}=[\ \mathbf{I}\ |\mathbf{0}] \text{ y } \mathbf{P'}=[[\mathbf{e'}]_x\mathbf{F} + \mathbf{e'v}^T \mid \lambda\mathbf{e'}]$$

Donde **v** es un 3-vector cualquiera y λ un escalar distinto de cero.

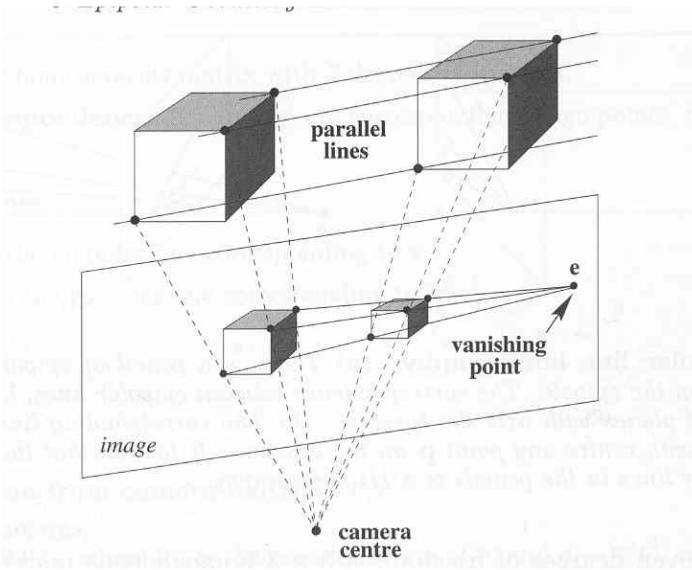
Matrices fundamentales asociadas a movimientos especiales

Estudiaremos dos casos de especial interés en las aplicaciones:

a) traslación pura y
b) movimiento plano puro

Traslación pura

Al considerar este caso podemos suponer la cámara quieta y el mundo moviéndose en una traslación –**t**. En esta situación, los puntos del espacio se mueven siguiendo líneas rectas paralelas a **t** cuyo punto de intersección será el punto de anulación **v** en la dirección **t.**



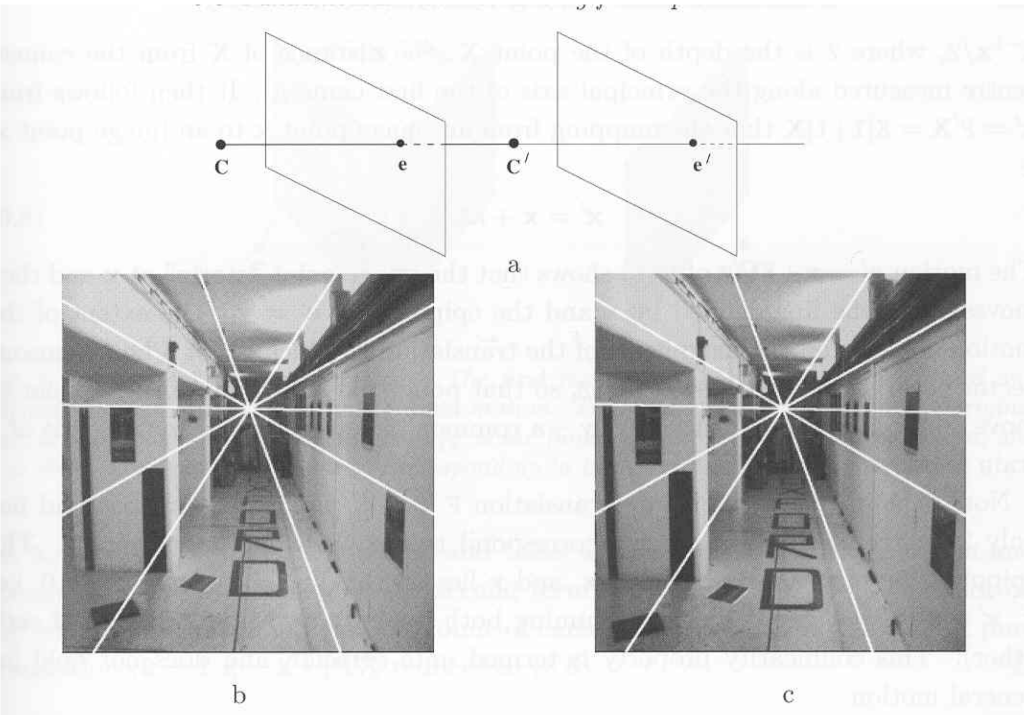
Supongamos que el movimiento de la cámara es de pura traslación sin rotación y sin cambio de parámetros internos. Podemos suponer que las cámaras son **P=K[I |0]** y **P'=K[I | t]**. Entonces la matriz fundamental será

$$\mathbf{F} = [\mathbf{e}']_x \mathbf{K}\mathbf{K}^{-1} = [\mathbf{e}']_x$$

Si la cámara se traslada paralela el eje *x*, entonces $\mathbf{e}'=(1,0,0)^T$, por tanto

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

y la relación $\mathbf{x}'^T\mathbf{F}\mathbf{x}=0$ se reduce a y=y', es decir las líneas epipolares se corresponden con las filas de la imagen.



Si el punto imagen $\mathbf{x}$ se normaliza a $\tilde{\mathbf{x}} = (x,y,1)^T$ entonces de la ecuación de la proyección $\mathbf{x}=\mathbf{P}\mathbf{X}=\mathbf{K}[\mathbf{I} \mid \mathbf{0}]\mathbf{X}$ se puede calcular las coordenadas no-homogéneas (X,Y,Z) de los puntos del espacio como $(X,Y,Z)^T= \mathbf{K}^{-1}\tilde{\mathbf{x}} = \mathbf{K}^{-1}\mathbf{x}/Z$ siendo Z la profundidad del punto $\mathbf{X}$ (distancia medida desde el punto al centro de la cámara a lo largo del eje principal). Se sigue entonces de la proyección del mismo punto $\mathbf{X}$ en la otra cámara, $\mathbf{x}'=\mathbf{P}'\mathbf{X}=\mathbf{K}[\mathbf{I} \mid \mathbf{t}]\,\mathbf{X}$ que la aplicación de un punto $\mathbf{x}$ de una imagen a una punto $\mathbf{x}'$ de la otra es

$$\mathbf{x}' = \mathbf{x} + \mathbf{K}\mathbf{t}/Z$$

Esta ecuación muestra que el punto comienza en $\mathbf{x}$ y se mueve a lo largo de la línea definida por $\mathbf{x}$ y el epipolo $\mathbf{e}=\mathbf{e}'=\mathbf{v}$. La extensión del movimiento depende del vector de traslación $\mathbf{t}$ y de la profundidad del punto Z, por tanto puntos más cercanos a la cámara aparecen moverse más rápidos que puntos más alejados.

Notemos que en este caso de traslación pura $\mathbf{F} = [\mathbf{e}']_x$ es antisimétrica y tiene por tanto solamente dos grados de libertad que corresponden a la posición del epipolo. La línea epipolar de $\mathbf{x}$ es $\mathbf{F}\mathbf{x}=[\mathbf{e}]_x\mathbf{x}= \mathbf{e}\times\mathbf{x}$, es decir $\mathbf{x}$ pertenece a su línea epipolar, luego se da un caso especial de auto-epipolaridad que no se verifica en un movimiento general

Movimiento general

El caso anterior de traslación pura nos da luz adicional sobre el movimiento general. Ya que dadas dos cámaras arbitrarias es posible mostrar que es posible definir una transformación proyectiva $\mathbf{H}$ tal que aplicada a la primera deje ambas cámaras alineadas (rotación + ajuste de las calibraciones de ambas cámaras). En ese caso la matriz fundamental de las cámaras alineadas seria $\mathbf{F} = [\mathbf{e}']_x$. Consecuentemente y deshaciendo la transformación $\mathbf{H}$ la matriz fundamental de las cámaras iniciales es $\mathbf{F} = [\mathbf{e}']_x \mathbf{H}$.

Si suponemos que las cámaras tienen matrices $\mathbf{P}=\mathbf{K}[\mathbf{I} \mid \mathbf{0}]$ y $\mathbf{P}'=\mathbf{K}'[\mathbf{R} \mid \mathbf{t}]$ entonces la transformación seria $\mathbf{H}= \mathbf{K}'\mathbf{R}\,\mathbf{K}^{-1} =\mathbf{H}_\infty$ y $\mathbf{F} = [\mathbf{e}']_x \mathbf{H}_\infty$.

Al igual que antes podemos calcular las coordenadas 3D de los puntos $(X,Y,Z)^T= \mathbf{K}^{-1}\tilde{\mathbf{x}} =\mathbf{K}^{-1}\mathbf{x}/Z$. Y usando la ecuación de proyección en la otra cámara obtenemos

$$\mathbf{x}'= \mathbf{K}'\mathbf{R}\,\mathbf{K}^{-1}\,\mathbf{x}+ \mathbf{K}'\mathbf{t}/Z = \mathbf{H}_\infty\,\mathbf{x}+ \mathbf{K}'\mathbf{t}/Z = \mathbf{H}_\infty\,\mathbf{x} + \mathbf{e}'/Z$$

Esta expresión muestra que el movimiento de un punto en una imagen por un movimiento general de una cámara se compone de dos partes: una que solo depende del giro y de los parámetros de calibración y otro que depende de la traslación efectuada y de la profundidad del punto en la escena.

Movimiento plano puro

En este caso es de destacar que la dirección del movimiento es siempre perpendicular a la normal al plano, lo cual induce una restricción sobre la matriz fundamental ya que es posible demostrar que la parte simétrica $\mathbf{F}_s=(\mathbf{F}+\mathbf{F}^T)/2$ de $\mathbf{F}$ tiene rango 2 lo cual le resta un grado de libertad dejándola solo con 6 grados de libertad.