

(continuación)

immerger todo en foco (extended depth-of-field)
EDoF

Caso $N=2$

$$s(x,y) = f_1(x,y) + f_2(x,y)$$

teníamos:

$$i_1(x,y) = f_1(x,y) + h_{12}(x,y) * f_2(x,y)$$

T.F. \downarrow

$$I_1(u,v) = F_1(u,v) + H_{12}(u,v) F_2(u,v) \leftarrow$$

$$i_2(x,y) = f_2(x,y) + h_{21}(x,y) * f_1(x,y)$$

T.F. \downarrow

$$I_2(u,v) = F_2(u,v) + \underbrace{H_{21}(u,v)}_{\substack{\text{simétrica} \\ H_{12}(u,v)}} F_1(u,v) \leftarrow$$

$$S(u,v) = \mathcal{F}\{s(x,y)\}$$

$$S(u,v) = F_1(u,v) + F_2(u,v)$$

$$= \vec{e} \cdot \vec{F}(u,v) \quad \text{con } \vec{e} = (1,1)$$

\uparrow producto interno

a forma matricial:

$$\underbrace{\begin{bmatrix} I_1(u,v) \\ I_2(u,v) \end{bmatrix}}_{\vec{I}(u,v)} = \underbrace{\begin{bmatrix} 1 & H_{12}(u,v) \\ H_{12}(u,v) & 1 \end{bmatrix}}_{H(u,v)} \underbrace{\begin{bmatrix} F_1(u,v) \\ F_2(u,v) \end{bmatrix}}_{\vec{F}(u,v)}$$

Para N imágenes en el stack:

$$\underbrace{\begin{bmatrix} I_1(u,v) \\ \vdots \\ I_N(u,v) \end{bmatrix}}_{\vec{I}(u,v)} = \underbrace{\begin{bmatrix} \Delta & H_{12}(u,v) & \dots & H_{1N}(u,v) \\ & \ddots & & \\ & & \ddots & H_{N-1,N}(u,v) \\ H_{2N}(u,v) & \dots & \dots & \Delta \end{bmatrix}}_{H(u,v) \text{ (simétrica)}} \underbrace{\begin{bmatrix} F_1(u,v) \\ \vdots \\ F_N(u,v) \end{bmatrix}}_{\vec{F}(u,v)}$$

luego esto para cada par de frecuencias (u,v)

Para hallar $\vec{F}(u,v)$ tenemos que invertir $H(u,v)$

→ No es invertible para $(u,v) = (0,0)$

$H_{(u,v)}(0,0) = \Delta$ (por ser una OTF)

→ Problema "ill-posed"

⇒ Posible solución → $H^T(u,v)$
pseudoinversa de Moore-Penrose

(se reduce a $H^{-1}(u,v)$ cuando $H(u,v)$ es invertible)

$H^T(u,v)$ provee el conjunto de vectores que minimizan la norma Euclídea

$$\|H(u,v)\vec{F}(u,v) - \vec{I}(u,v)\| \text{ siendo}$$

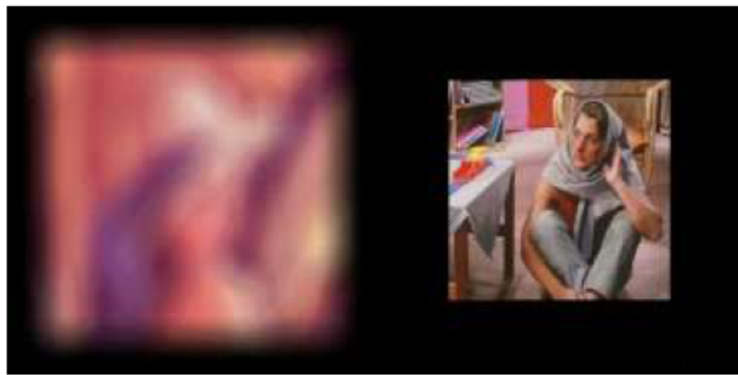
$H^T(u, v) \vec{I}(u, v)$ el vector de norma mínima dentro de ese conjunto.

Aproximamos la solución por:

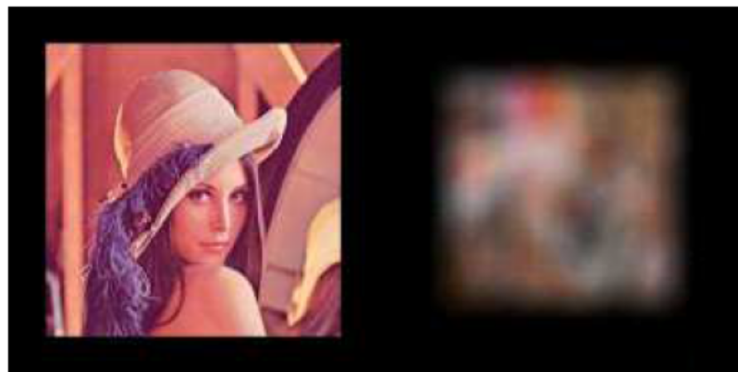
$$\Rightarrow S(u, v) = \vec{e} \cdot [H^T(u, v) \vec{I}(u, v)]$$

y la imagen todo en foco:

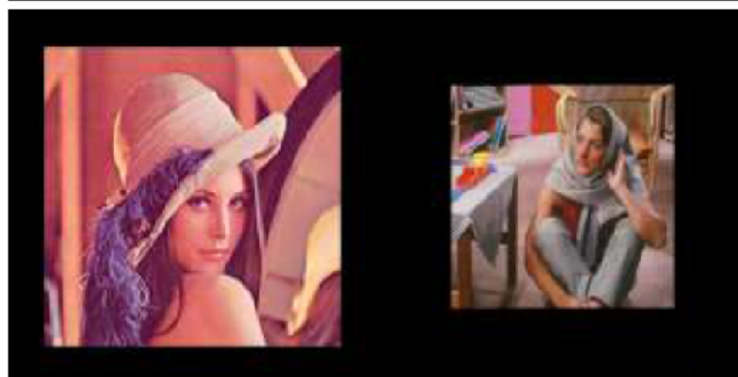
$$s(x, y) = \mathcal{F}^{-2} \{ S(u, v) \}$$



$\leftarrow i_2(x, y)$



$\leftarrow i_1(x, y)$



$\leftarrow s(x, y)$

fusión

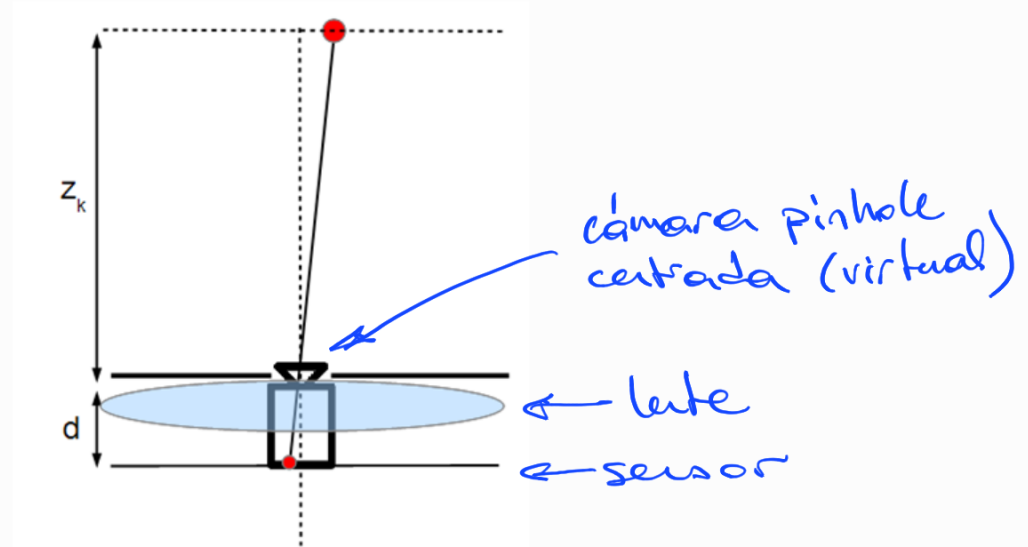
o

todo en foco

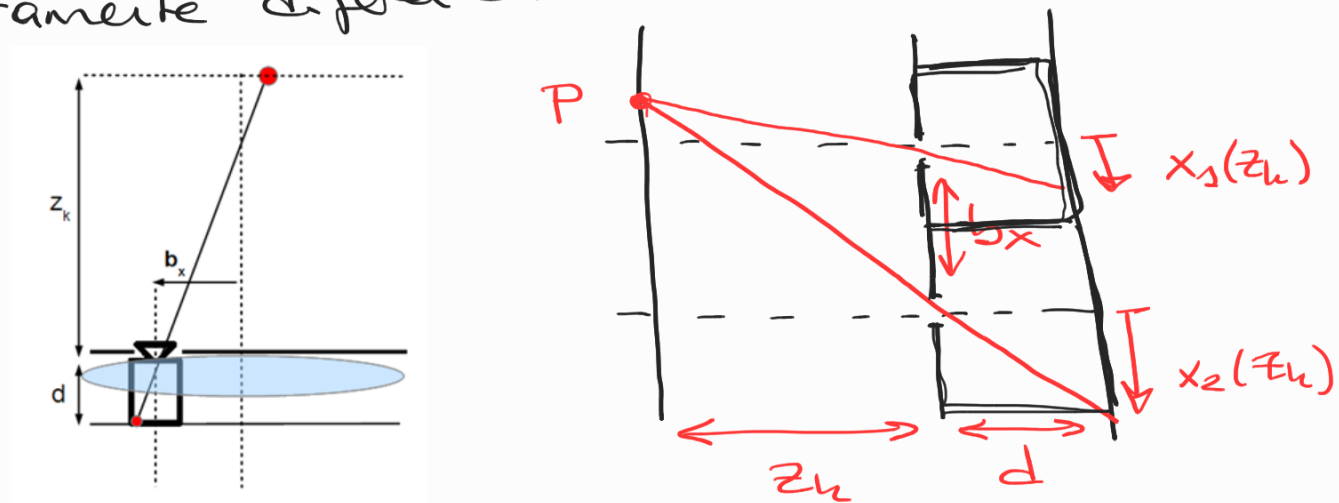
2) síntesis de puntos de vista

Se puede generar pequeños cambios de perspectiva simulando desplazamientos horizontales y verticales de una cámara pinhole virtual.

La imagen todo en foco es equivalente a la imagen que hubiéramos capturado con una cámara pinhole virtual centrada.



Si consideramos una cámara pinhole virtual desplazada, vemos una imagen todo en foco de la escena 3D pero desde una perspectiva ligeramente diferente.



obs: región en foco (f_k) debe desplazarse una cantidad correspondiente a la disparidad asociada con el baseline (b_x) y la distancia z_k

$$x_2 - x_1 = \frac{d b_x}{z_k}$$

← baseline

$$\Rightarrow S_{b_x, b_y}(x, y) = \sum_{k=1}^N f_k \left(x - \frac{d b_x}{z_k}, y - \frac{d b_y}{z_k} \right)$$

↑
 (b_x, b_y) ← baseline (posición de la nueva cámara pinhole virtual con respecto a la cámara pinhole centrada)

Fourier-Transform Shift Theorem

$$\mathcal{F}\{g(x, y)\} = G(u, v)$$

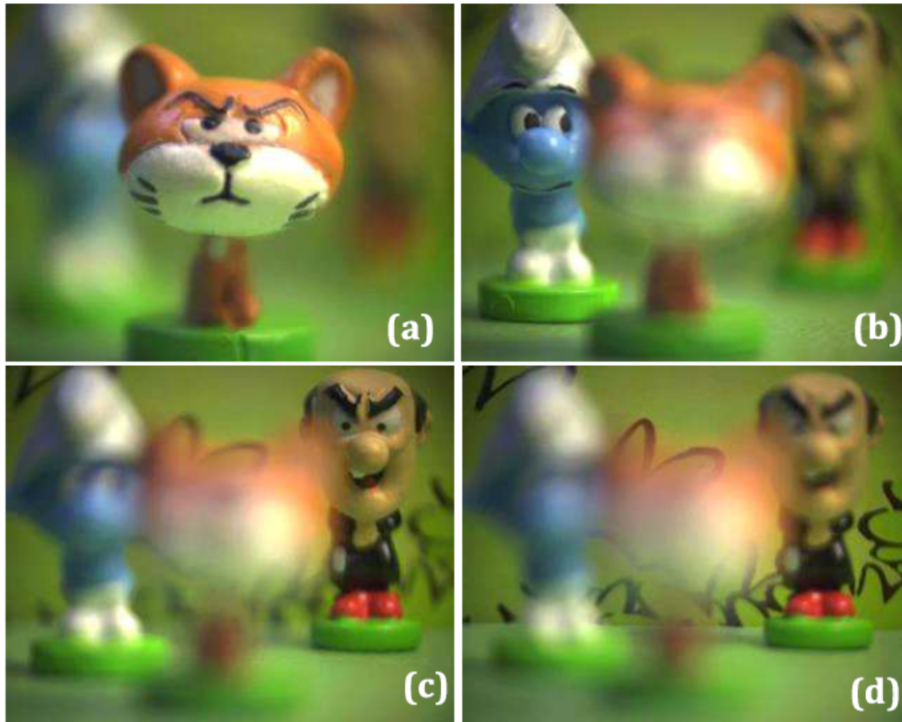
$$\Rightarrow \mathcal{F}\{g(x-a, y-b)\} = G(u, v) e^{-j2\pi(ua + vb)}$$

Entonces,

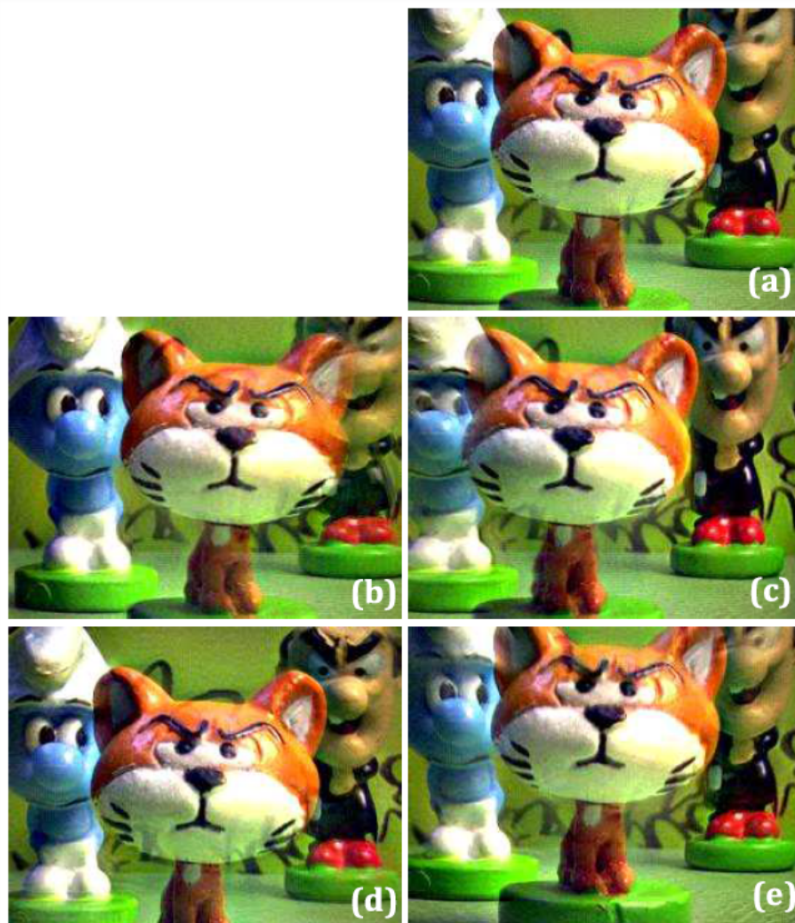
$$S_{b_x, b_y}(u, v) = \sum_{k=1}^N e^{-j \frac{2\pi d}{z_k} (b_x u + b_y v)} \left(H^+(u, v) \vec{I}(u, v) \right)_k$$

$$\Rightarrow S(x, y)_{b_x, b_y} = \mathcal{F}^{-1} \left\{ S(u, v)_{b_x, b_y} \right\}$$

ej. stack multifoco



perspectivas sintetizadas:



En particular, se pueden generar pares estereoc (una perspectiva para cada ojo)



Figure 6.14: Stereoscopic pair from *cat* set. Right and left images for cross-eyed view.



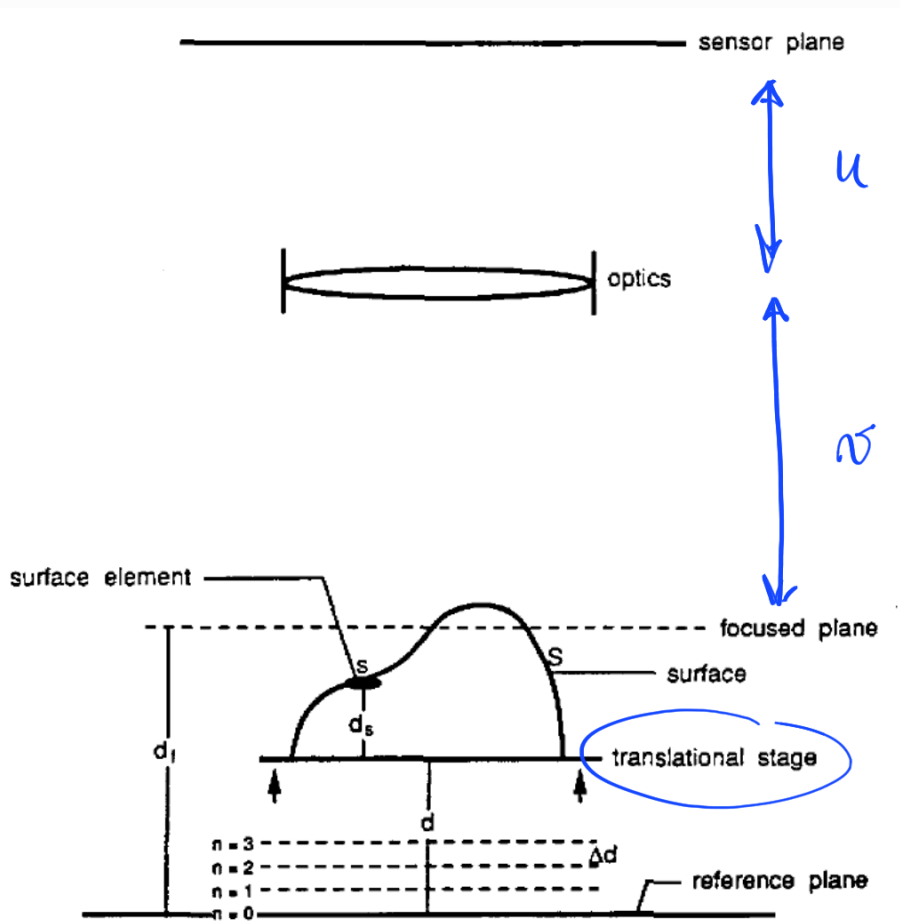
Figure 6.15: Red cyan anaglyph from *cat* set. (Red/cyan glasses are needed.)

Mapa de profundidad por Foco

Depth From Focus - DFF / Shape From Focus

No es trivial cómo estimar el nivel de nitidez, en particular en zonas de poca textura.

La mayoría de estas aproximaciones, intentan asociar a $\text{pixel } (x, y)$ un valor de profundidad basado en maximizar una función como medida de foco.



u = dist. lente-sensor

v = dist. lente a plano de enfoque

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

Una medida de foco utilizada es el Laplaciano de suma modificada (sum-modified Laplacian / Modified Laplacian).

Se aplica la función medida de foco a la secuencia multifoco

Laplaciano:

$$\nabla^2 I(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

$I(x,y)$ = intensidad de la imagen en el punto (x,y)

Laplaciano modificado (SML):

$$\nabla_{\pi}^2 I = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right|$$

La medida de foco en el punto (x, y) se calcula como el laplaciano modificado en una pequeña ventana de tamaño $K \times K$ ($K=3, K=5, K=10, \text{etc.}$) alrededor de (x, y)

$$M(x, y, k) = \sum_{i=x-k}^{x+k} \sum_{j=y-k}^{y+k} \text{SML}(i, j, k)$$

↖ medida de
foco en (x, y)

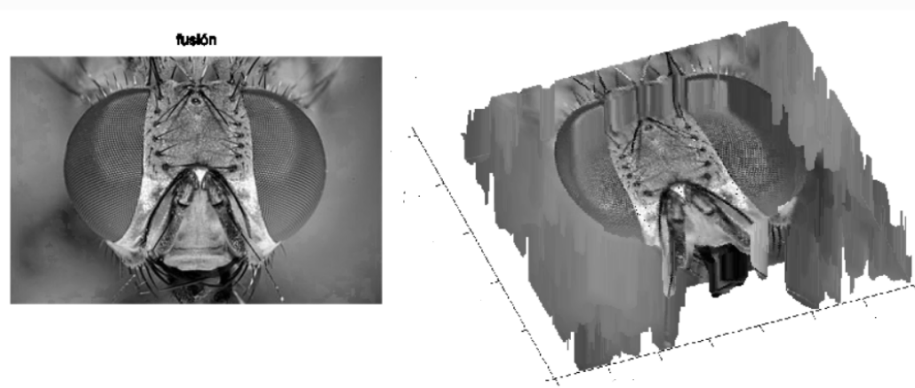
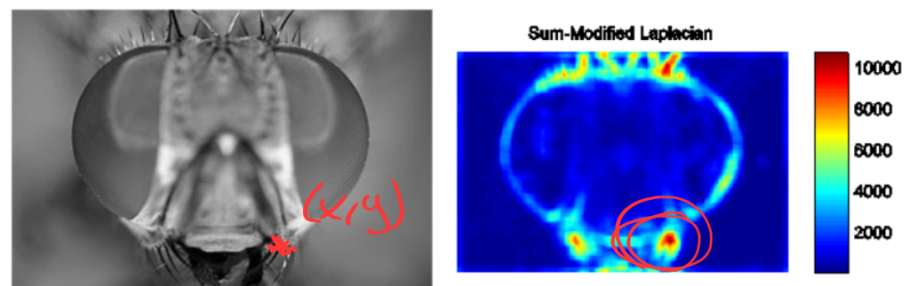
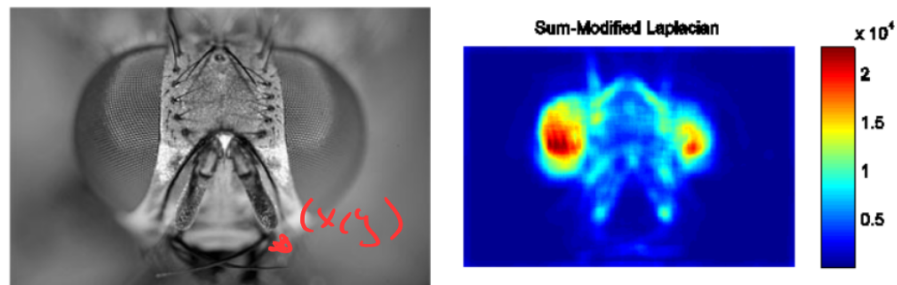
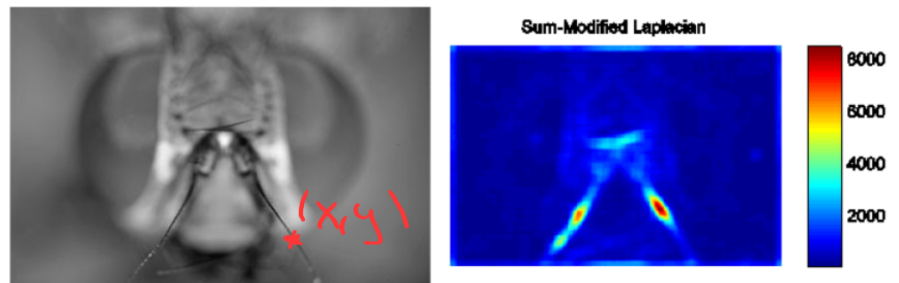
⇒ la profundidad de la escena para γ pixel (x, y) , corresponde a calcular el índice k dentro del stack multifoco donde la medida de foco se hace máxima:

$$D(x, y) = \underset{k}{\text{argmax}} M(x, y, k)$$

↖ asigno el pixel (x, y) a aquella distancia (o imagen) del stack en que la medida de foco se hace máxima.

dos: existen distintas medidas de foco
(ej $(\nabla^2 I)^2$, etc.). A veces se considera
un umbral T para asignar el pixel.

Una vez determinado el mapa, es inmediato
sintetizar la imagen todo en foco.



Métodos multi-foco vs. estéreo

¿qué tan diferentes son en realidad?

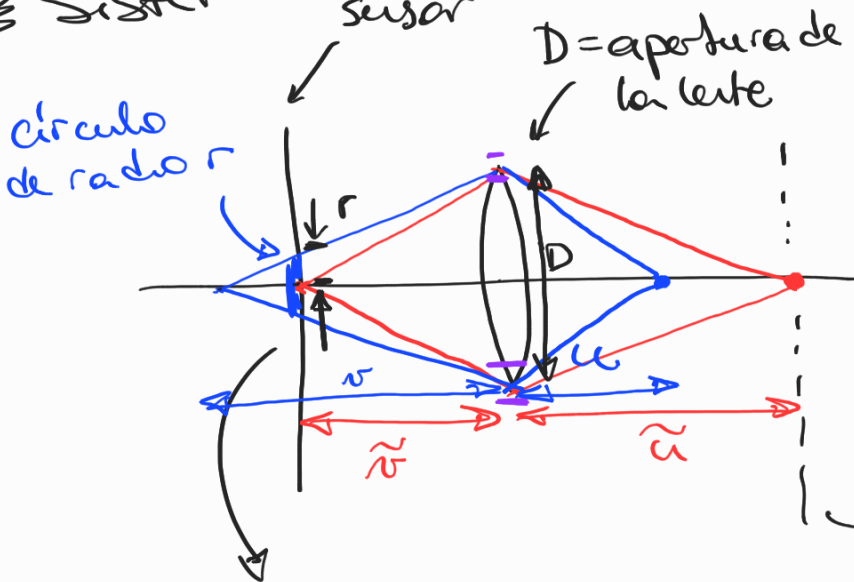
Profundidad por foco (DFF)

Profundidad por desenfoque (DFD)

Los métodos DFF reconstruyen el mapa de profundidad de una escena 3D a partir de secuencias de imágenes adquiridas cambiando las distancias de enfoque y usar medidas de foco.

Los métodos DFD buscan obtener el mapa de profundidad a partir de dos imágenes (frente y fondo) y usar el radio de "blureo".

≡ Sistema multi-foco sensor



$$\frac{1}{\tilde{u}} + \frac{1}{\tilde{v}} = \frac{1}{F} \quad \text{distancia focal de la lente}$$

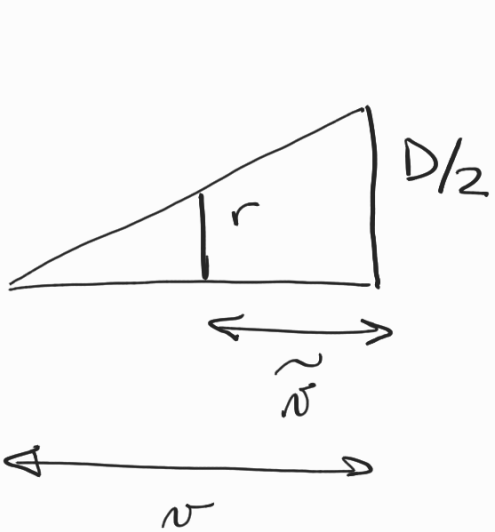
$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

→ el objeto a dist. \tilde{u} de la lente está en foco en el sensor

el objeto a dist. u de la lente, se ve fuera de foco, con una PSF dada por el círculo de radio r en el plano del sensor

$$\frac{1}{2r} + \frac{1}{2r} = \frac{1}{F}$$

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{F} \rightarrow \frac{1}{v} = \frac{1}{F} - \frac{1}{u}$$



$$\frac{r}{D/2} = \frac{|v - \tilde{v}|}{v} = |1 - \tilde{v}/v|$$

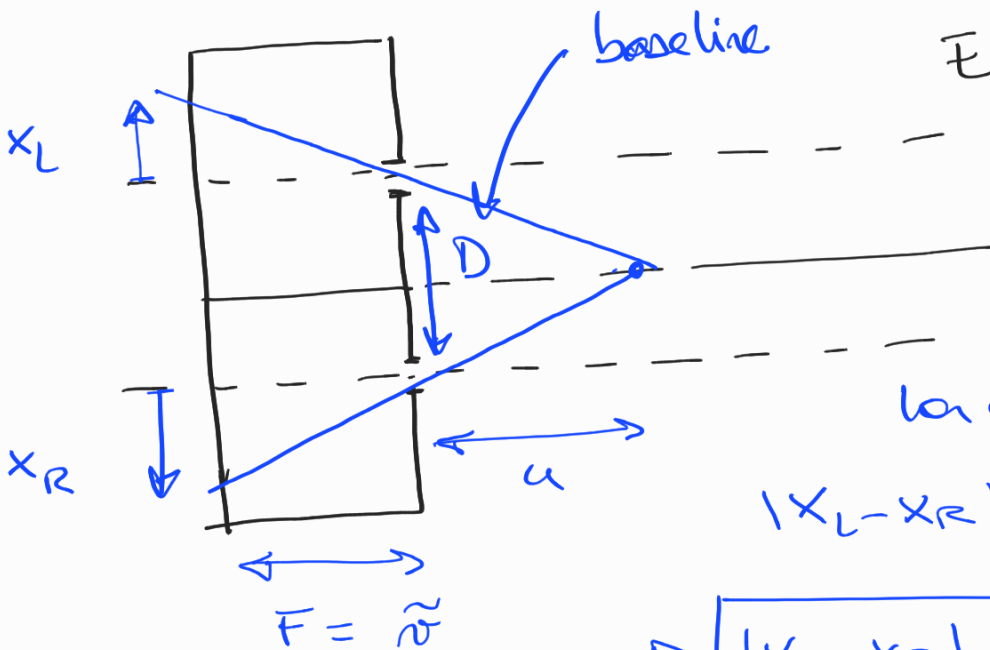
$$= \left| 1 - \left(\frac{2r}{F} - \frac{2r}{u} \right) \right|$$

$$= \frac{|Fu - u\tilde{v} + F\tilde{v}|}{Fu}$$

$$r = \frac{D}{2} \frac{|Fu - u\tilde{v} + F\tilde{v}|}{Fu} \quad (*)$$

≡ sistema estéreo (2 cámaras pinhole con baseline D)

Cámara pinhole
 $\tilde{v} = F$



En (1):

$$r \approx \frac{D}{2} \frac{2r}{u}$$

Comparando con la disparidad:

$$|x_L - x_R| = \frac{D\tilde{v}}{u}$$

$$\Rightarrow |x_L - x_R| \approx 2r$$

Obs: dados un sistema multi-foco y un sistema estéreo con las mismas dimensiones físicas (tamaño de la apertura de la lente D_{FD}/D_{FE} y baseline en estéreo), la disparidad es igual al diámetro de la PSF geométrica (círculo de "blureo")