

# Acquisition de l'image numérique

## L'appareil photographique

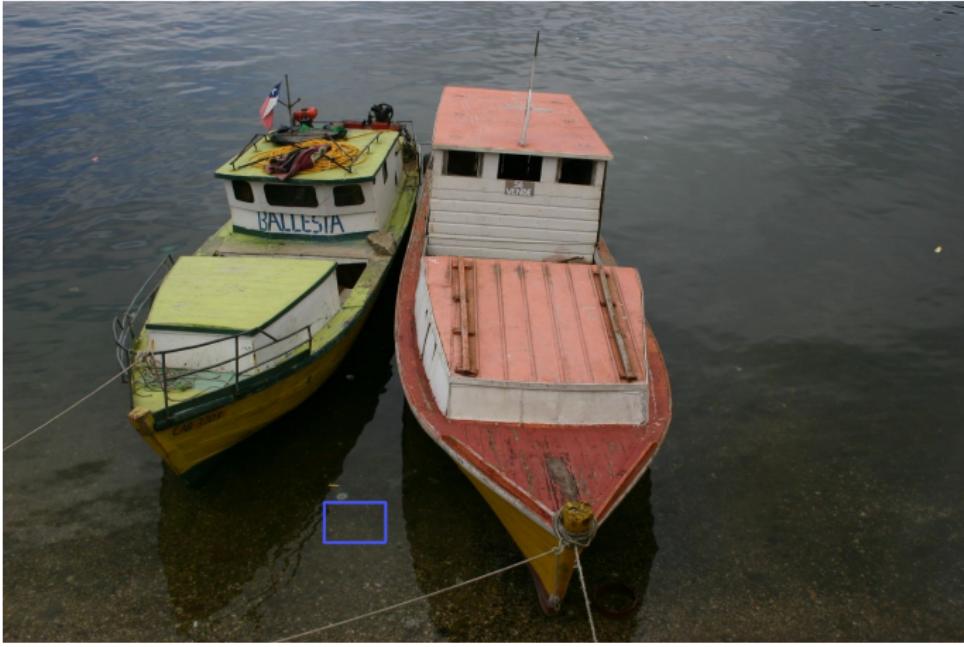
Yann GOUSSEAU  
Télécom ParisTech  
IMA 201  
2021-2022



# Acquisition des images le cas de la photographie numérique

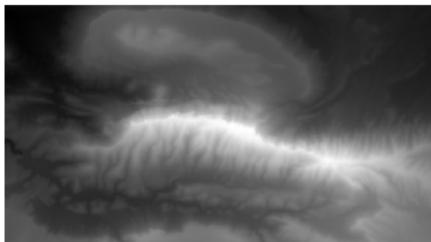
- ① Généralités
- ② Le système optique
- ③ Transformée de Fourier et échantillonnage bidimensionnel
- ④ Sources et modélisation du bruit





# Des dispositifs couverts par ce cours

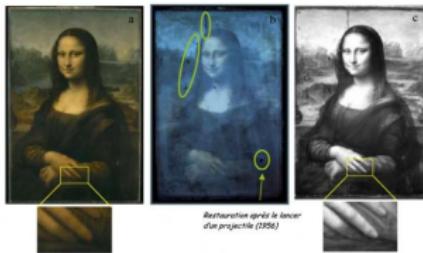




Modèle numérique de terrain



image radar



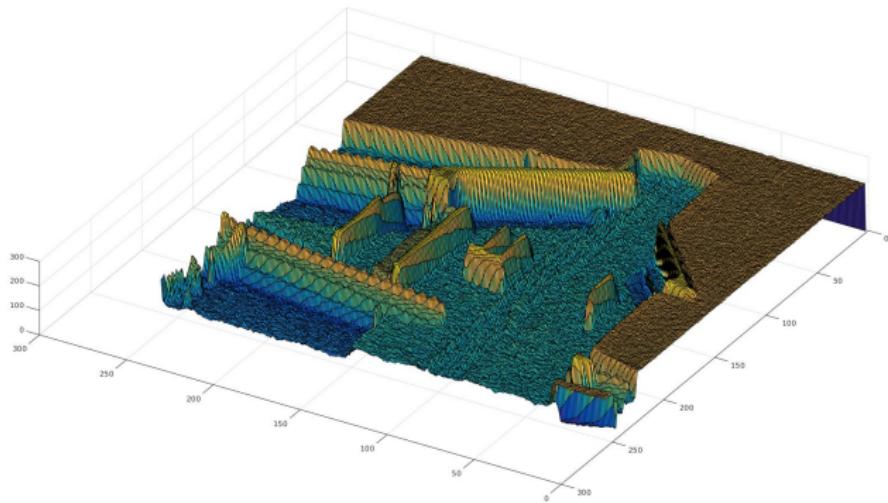
Joconde : visible, UV, IR

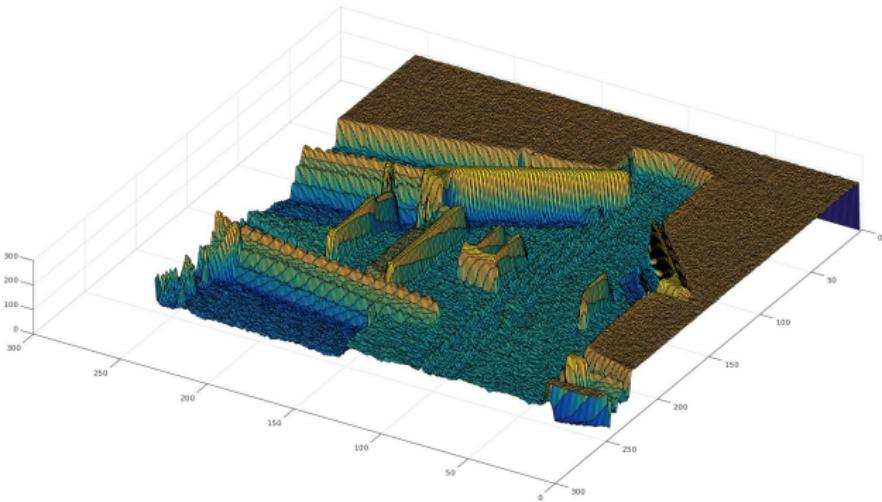


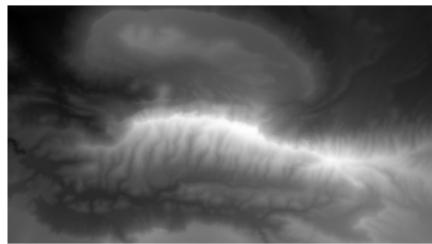
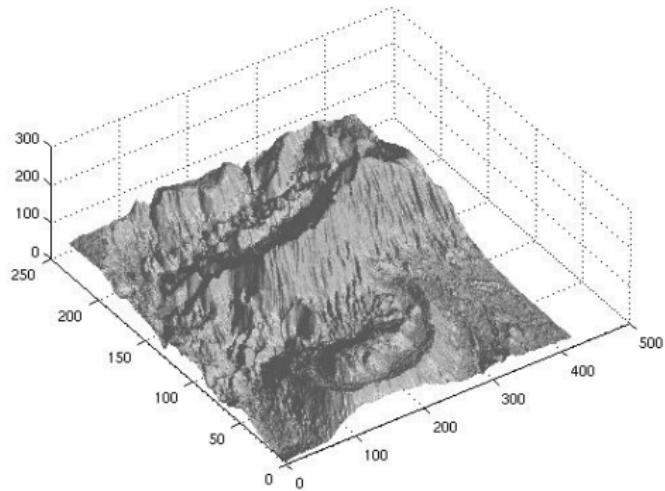
radiographie (rayons X)

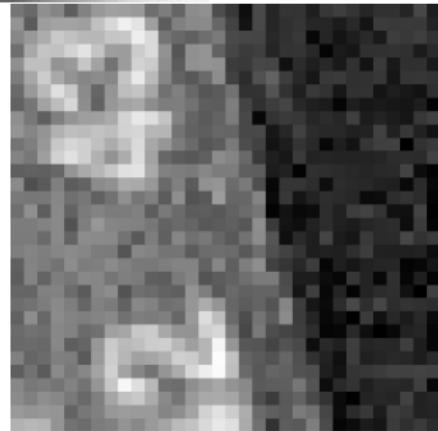
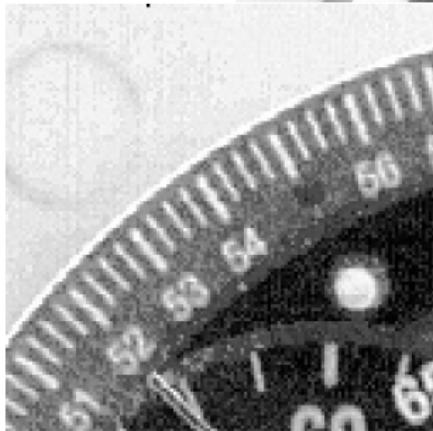


Image de profondeur (laser)









## Le modèle sténopé (pin-hole)

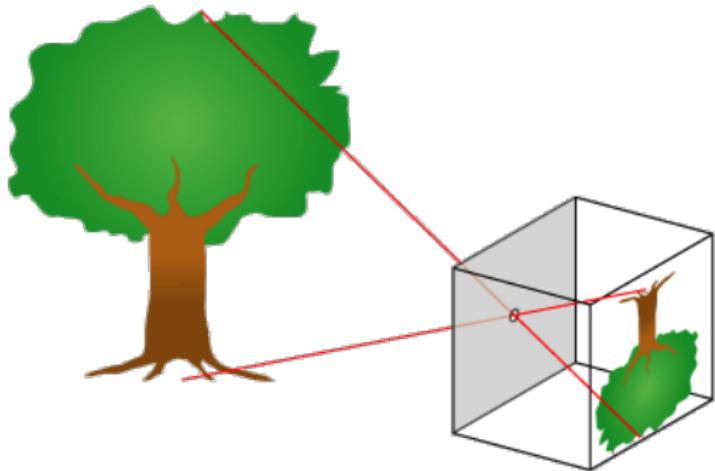


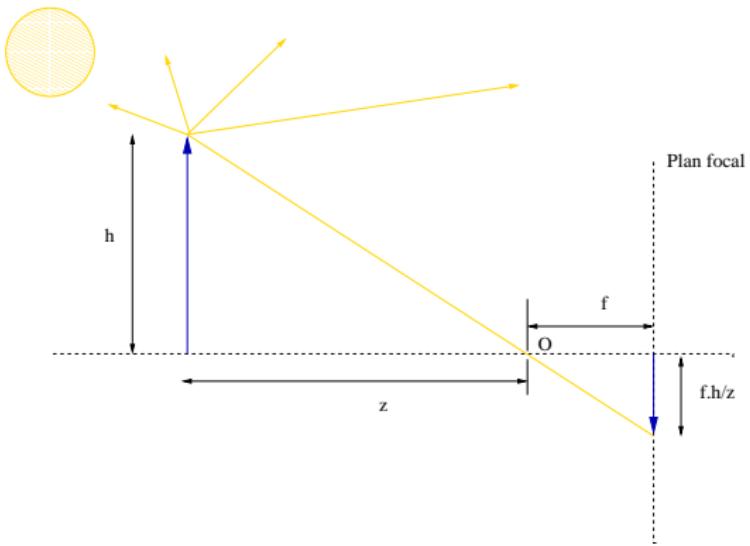
Illustration from wikipedia.org

(Aristote, Alhazen 10<sup>eme</sup> s., Brunelleschi début 15<sup>eme</sup> s., Da Vinci 1500, Kepler 1604 *camera obscura*)



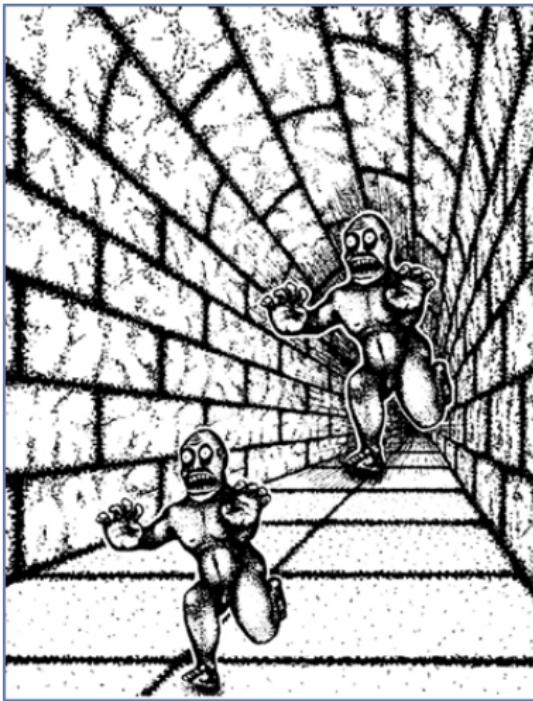
Victoria Cooper - Doug Spowart - creative commons

# Le modèle sténopé (pin-hole)



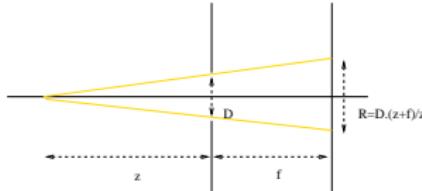
Une partie de la lumière issue de l'objet passe par l'ouverture  $O$  et se projette sur le plan focal. La distance  $f$  est appelée distance focale.

Première conséquence : pas de perception absolue de la taille des objets.



Deuxième conséquence : le phénomène d'occultation → discontinuités (bords)





Ouverture  $O$  finie : rayons → cônes → flou

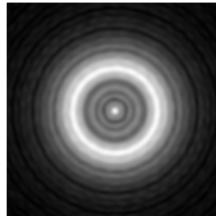
Pour une ouverture de diamètre  $D$  et  $S$  la scène (à distance  $z$ ), on observe (en première approximation)

$$g_z * S,$$

où  $g_z$  est la fonction indicatrice d'un disque de rayon  $(z + f)D/z$ .

Si  $z \gg f$  chaque point donne à peu près la même tache de flou

Modélisation plus fine : prise en compte de la diffraction

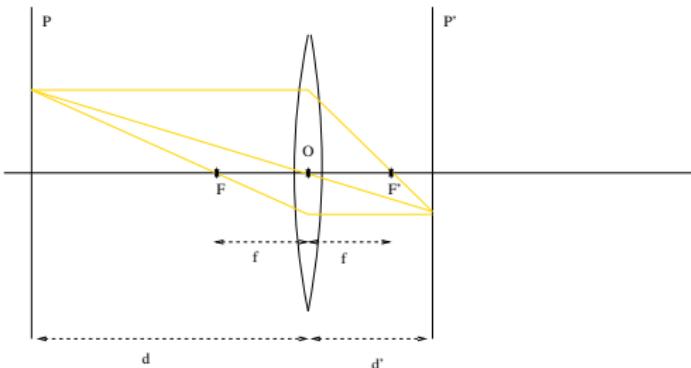




from Sheila Bocchine - creative commons

- Pour obtenir une image plus nette tout en ayant une ouverture non nulle (besoin de lumière) : ajout d'une lentille.
- On observe toujours  $g * S$ , où  $g$  est la réponse impulsionnelle du système optique (en négligeant diverses aberrations et interférences).

# Le modèle "lentille mince"



Hypothèses :

- Rayons lumineux proches de la normale à la surface de la lentille
- On néglige le trajet de la lumière à l'intérieur de la lentille.

Rayons issus de  $P$  se concentrent sur  $P'$ . On note  $f$  la *distance focale*  
→ Relation de Descartes

$$\frac{1}{d} + \frac{1}{d'} = \frac{1}{f}.$$

Mise au point : choix du plan  $P$ .

Si l'objet photographié n'est pas dans le plan  $P$  : flou de mise au point → convolution.

## Modèle d'acquisition I : la partie "optique"

On observe une scène

$$\tilde{s} = g_o * s,$$

où

$$g_o = g_{ouv} * g_{flou} * g_{fil}.$$

- $g_{ouv}$  ouverture finie
- $g_{flou}$  défaut de mise au point
- $g_{fil}$  flou de bougé (ou de filé)

$g_o$  s'appelle la *réponse impulsionnelle* (PSF en anglais pour Point Spread Function) du système optique.

Un phénomène non pris en compte : le flou est variable selon la distance des objets à l'appareil photo



# La profondeur de champ

Distance séparant l'objet net le plus proche de l'appareil de l'objet net le plus lointain.

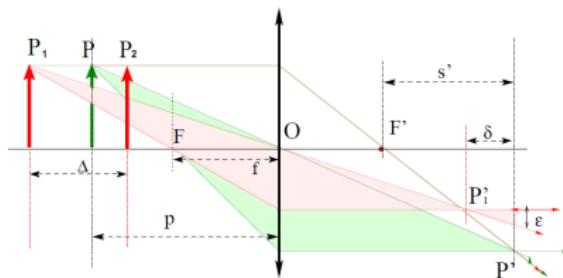


Figure par H. Maître

En utilisant la relation de Descartes et en notant  $D$  l'ouverture du diaphragme, on trouve

$$\Delta = \frac{2\epsilon Dpf(p-f)}{D^2f^2 - \epsilon^2(p-f)^2}$$

En supposant que  $\delta \ll s$  et  $f \ll p$  on a

$$\Delta \approx \frac{2\epsilon p^2}{Df}$$

# Ouverture

$D$  diamètre du diaphragme ;  $f$  distance focale ;  
nombre d'ouverture  $N = f/D$  (généralement progression géométrique)



wikipedia.org

Quand  $N$  augmente (à  $f$  constant) :

- La profondeur de champs augmente (formule précédente)
- Les défauts de diffraction augmentent
- Les défauts de vignettage diminuent

# Profondeur de champs

Nombre d'ouverture  $N = 1,4$



# Profondeur de champs

Nombre d'ouverture  $N = 13$



## Flou artistique



f/2.8 - by bahramr - creative commons

# Focale

Diverses focales en format 24 × 36 depuis un même point de vue : influence sur la taille de l'image



28 mm :  
courte focale,  
image de petite taille.



50 mm :  
focale normale.



70 mm



210 mm :  
longue focale,  
image de grande taille.

**Changement de focale et de point de vue : influence sur la perspective**



24 mm :  
point de vue rapproché,  
perspective exagérée.



50 mm :  
perspective « normale ».



100 mm



200 mm :  
point de vue éloigné,  
perspective écrasée.

(source : Wikipedia)

## Acquisition alternative : le light field

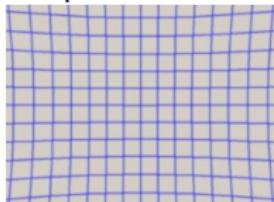


Caméra Lytro

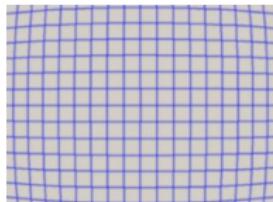
## Autres phénomènes

Parmi les phénomènes qui ne rentrent pas dans le cadre d'un système linéaire invariant par translation :

- Distortion géométriques

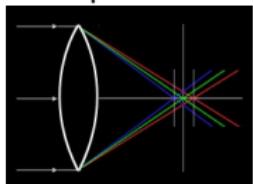


"en coussinet" :

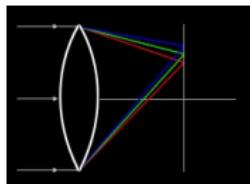


"en barillet" :

- Aberrations chromatiques



"longitudinale" :

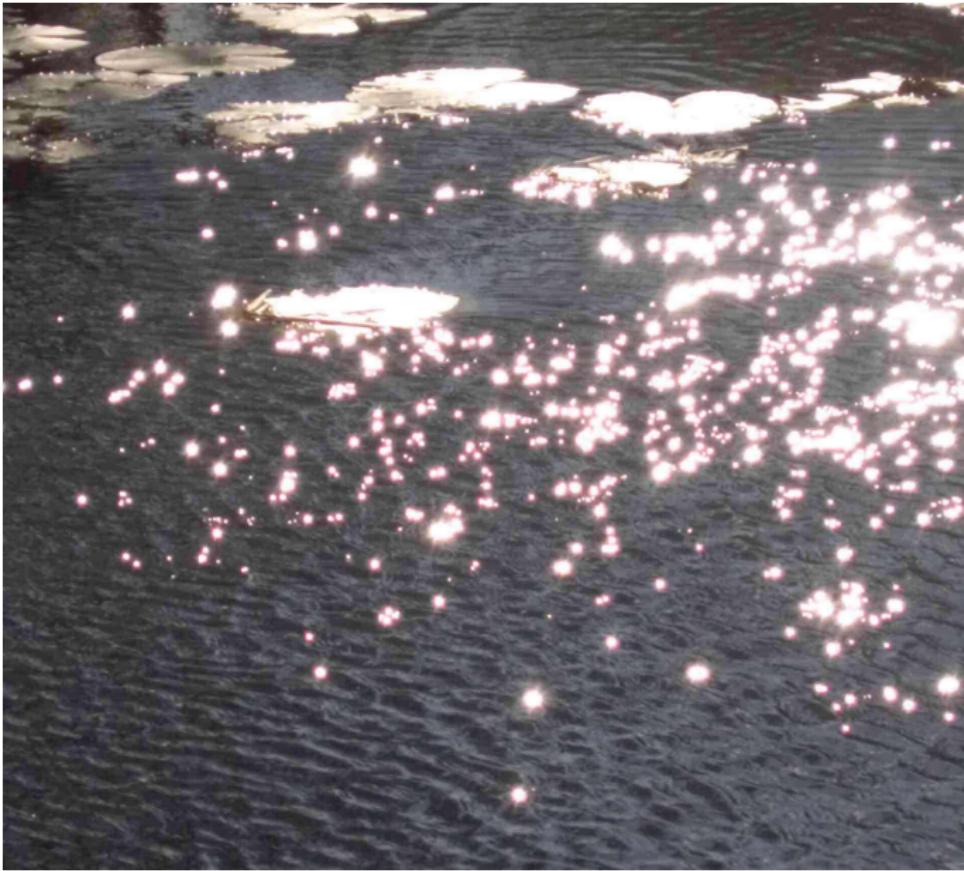


"transversale" :

- Vignetage



Aberration transversale





Avant et après correction du vignettage  
dxo.com

## Autre paramètre d'acquisition : le temps d'exposition

- Peut compenser un manque de lumière ou une ouverture faible
- Doit être court pour les objets en mouvement
- Sinon, flou de mouvement



- Après passage dans le système optique → acquisition de l'information lumineuse du plan focal
- Utilisation d'un capteur numérique : grille de capteurs photosensibles qui convertissent les photons en courant électrique
- *Echantillonnage* du signal lumineux

$$\left( f : \mathbb{R}^2 \mapsto \mathbb{R} \right) \rightarrow (\{f(k)\}_{k \in \Omega}).$$

- Intégration des photons par le capteur (comptage des photons) → convolution avec  $g_{capt}$ , fonction indicatrice du capteur  
Modification de la PSF

# Intégration par le capteur

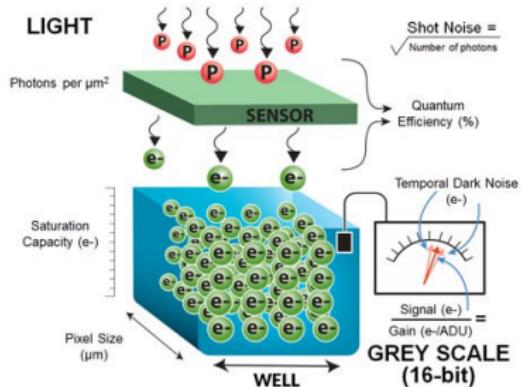
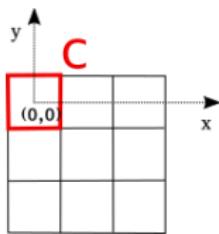


image credit : [www.flir.com](http://www.flir.com)

En chaque cellule photoréceptrice : addition des contributions des photons  
≠ valeurs ponctuelles

# Intégration par le capteur

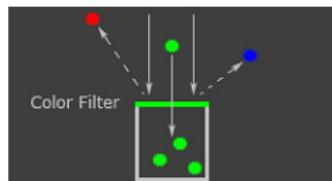
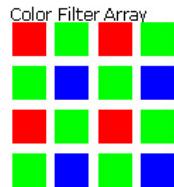
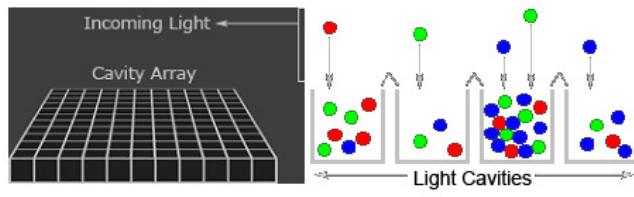


$I_{i,j}$  : image capturée par le capteur à la position  $(i,j)$

$$\begin{aligned} I_{i,j} &= \int_{C+(i,j)} I(x, y) dx dy \\ &= \int_{\mathbb{R}^2} I(x, y) \mathbb{1}_C(x - i, y - j) dx dy \\ &= \int_{\mathbb{R}^2} I(x, y) \mathbb{1}_C(i - x, j - y) dx dy \\ &= (I * \mathbb{1}_C)(i, j) \end{aligned}$$

→ échantillonnage de  $I * \mathbb{1}_C$

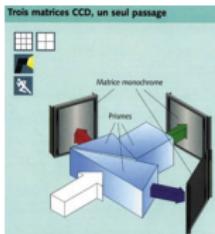
# Trame de Bayer



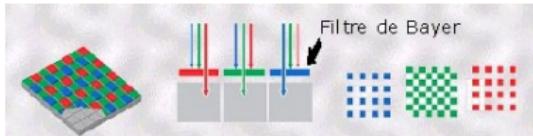
Ordre de grandeur de la taille des capteurs :  
1 (smartphone) à 10 (reflex plein format) microns.

# Alternatives pour capturer la couleur

- Tri-CCD



- Capteur Fovéon



→ cours sur la couleur

- Pour  $f \in L^1(\mathbb{R}^2)$ ,

$$\mathcal{F}(f)(\omega) = \hat{f}(\omega) = \int_{\mathbb{R}^2} f(\mathbf{x}) \exp(-i\mathbf{x} \cdot \omega) d\mathbf{x}.$$

- Définition alternative (traitement du signal) avec  $\exp(-2i\pi \mathbf{x} \cdot \omega)$ .
- $\hat{f}$  est continue et  $\rightarrow 0$  à l'infini.
- On supposera désormais que  $f$  est régulière et décroît rapidement à l'infini :

$$\forall p, q \in \mathbb{N}^2, |\mathbf{x}|^p \partial_q f \leq C.$$

$$\mathcal{F}^{-1}(f) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} f(\mathbf{x}) \exp(i\mathbf{x} \cdot \omega) d\mathbf{x}.$$

# Rappels d'analyse de Fourier

- $\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$  et  $\mathcal{F}(f \cdot g) = (2\pi)^{-2} \mathcal{F}(f) * \mathcal{F}(g)$ .
- $\mathcal{F}(f(\mathbf{x} - \mathbf{a})) = \exp(-i\omega \cdot \mathbf{a}) \mathcal{F}(f)$  et  $\mathcal{F}(\exp(i\mathbf{a} \cdot \mathbf{x}) f)(\omega) = \mathcal{F}(f)(\omega - \mathbf{a})$
- Dans la suite, on appliquera certaines de ces formules à l'impulsion de Dirac ( $\delta$ ) sans plus de justifications  
Cadre mathématique : théorie des distributions (e.g. "Analyse de Fourier et applications", C. Gasquet et P. Witomski, Dunod, 2000)

- $\mathcal{F}(\delta_x)(\omega) = \exp(-ix\omega)$
- Sous certaines conditions, les correspondances convolution-multiplication restent vraies avec les impulsions
- On montre que le *peigne de Dirac*,  $\sum_{j \in \mathbb{Z}} \delta_{aj}$ , vérifie

$$\mathcal{F}\left(\sum_{j \in \mathbb{Z}} \delta_{aj}\right)(\omega) = \sum_{j \in \mathbb{Z}} \exp(iaj\omega) = \frac{2\pi}{a} \sum_{j \in \mathbb{Z}} \delta_{\frac{2\pi j}{a}}.$$

Retour à l'acquisition des images :

$$u = (h * s).\Pi_\Gamma.F$$

où

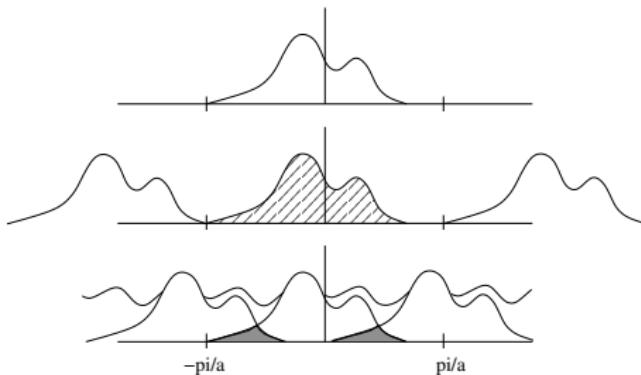
- $s$  scène (fonction de  $L^1$ ),
- $h$  : réponse impulsionnelle du système optique et de l'intégration des capteurs ( $g_{ouv} * g_{flou} * g_{fil} * g_{capt}$ ),
- $\Pi_\Gamma = \sum_{\gamma \in \Gamma} \delta_\gamma$ ,
- $F = \mathbb{1}_\Omega$ ,  $\Omega \in \mathbb{R}^2$ , support de l'acquisition.

# Le repliement de spectre en dimension 1

On considère  $\Gamma = a\mathbb{Z}$ , donc  $\Pi_\Gamma = \sum_{j \in \mathbb{Z}} \delta_{aj}$ ,  
on a alors  $u = f \cdot \Pi_\Gamma$  et

$$\mathcal{F}(u) = \frac{1}{a} \mathcal{F}(f) * \sum_{j \in \mathbb{Z}} \delta_{\frac{2\pi j}{a}},$$

$$\mathcal{F}(u)(\omega) = \frac{1}{a} \sum_{j \in \mathbb{Z}} \mathcal{F}(f) \left( \omega - \frac{2\pi j}{a} \right).$$



# Le repliement de spectre en dimension 1

Conséquence :

- si  $\text{Supp}(\mathcal{F}(f)) \subset [-\pi/a, \pi/a]$ , on a

$$\mathcal{F}(f.\Pi_\Gamma).\mathbb{1}_{[-\pi/a, \pi/a]} = \mathcal{F}(f),$$

donc on retrouve  $f$  car

$$f = \mathcal{F}^{-1}(\mathcal{F}(f.\Pi_\Gamma).\mathbb{1}_{[-\pi/a, \pi/a]}) = f.\Pi_\Gamma * \mathcal{F}^{-1}(\mathbb{1}_{[-\pi/a, \pi/a]})$$

or  $\mathcal{F}^{-1}(\mathbb{1}_{[-\pi/a, \pi/a]}) = \frac{\sin(\pi x/a)}{\pi x/a}$ , donc

$$f(x) = \sum_{j \in \mathbb{Z}} f(ja) \frac{\sin((x - ja)\pi/a)}{(x - ja)\pi/a}.$$

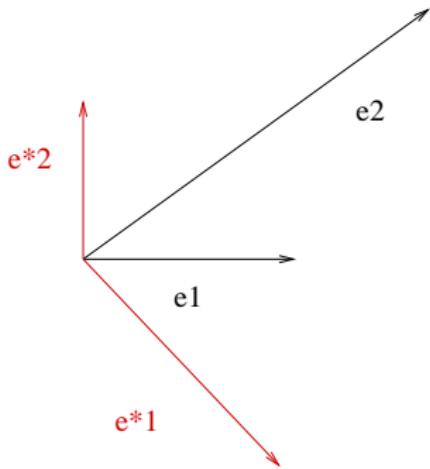
- sinon, le spectre de  $f.\Pi_\Gamma$  est détérioré de manière irréversible.

## La dimension 2 : le réseau dual

Soit  $\Gamma = \mathbb{Z}.\mathbf{e}_1 + \mathbb{Z}.\mathbf{e}_2$ , où  $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^2$  non colinéaires.

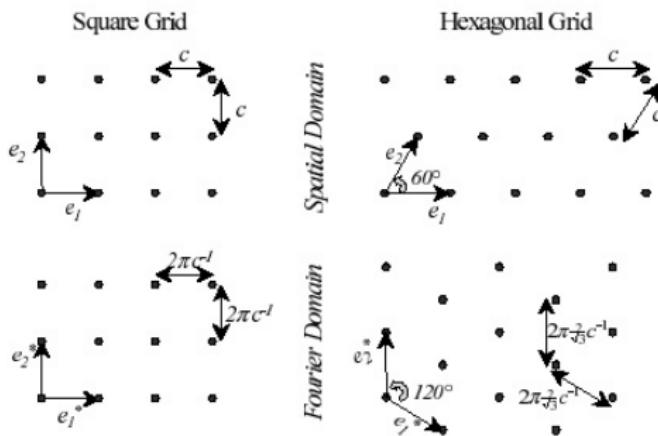
Le réseau dual est  $\Gamma^* = \mathbb{Z}.\mathbf{e}_1^* + \mathbb{Z}.\mathbf{e}_2^*$ ,

où  $\mathbf{e}_i \cdot \mathbf{e}_j^* = 2\pi\delta_{i,j}$ .



# Exemples de réseaux

- Réseau carré  $\leftrightarrow$  réseau carré
- Réseau hexagonal  $\leftrightarrow$  réseau hexagonal



## Le repliement de spectre en dimension 2

On a  $\Gamma = \mathbf{e}_1\mathbb{Z} + \mathbf{e}_2\mathbb{Z}$  et  $\Gamma^* = \mathbf{e}_1^*\mathbb{Z} + \mathbf{e}_2^*\mathbb{Z}$ ,  $\Pi_\Gamma = \sum_{j,k \in \mathbb{Z}} \delta_{j\mathbf{e}_1, k\mathbf{e}_2}$ .

On montre que  $\mathcal{F}(\Pi_\Gamma) = |\mathbf{e}_1^* \wedge \mathbf{e}_2^*| \Pi_{\Gamma^*}$ .

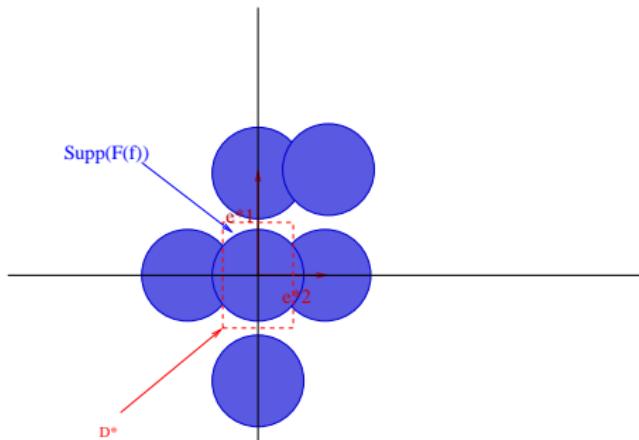
L'image échantillonnée est  $u = f \cdot \Pi_\Gamma$  et

$$\mathcal{F}(u) = |\mathbf{e}_1^* \wedge \mathbf{e}_2^*| \mathcal{F}(f) * \Pi_{\Gamma^*}$$

donc

$$\mathcal{F}(u)(\omega) = |\mathbf{e}_1^* \wedge \mathbf{e}_2^*| \sum_{\gamma \in \Gamma^*} \mathcal{F}(f)(\omega - \gamma)$$

On a repliement de spectre si les  $\gamma + \text{Supp}(\mathcal{F}(f))$  ne sont pas disjoints, pour  $\gamma \in \Gamma^*$ .



Soit  $K \subset \mathbb{R}^2$  tel que  $\forall \gamma, \gamma' \in \Gamma^*$  :

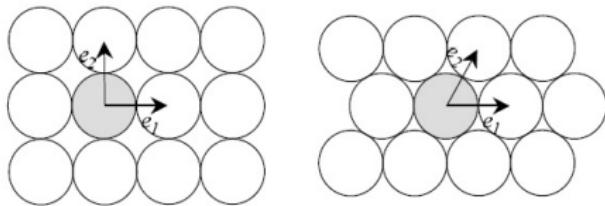
$$(\gamma + K) \cap (\gamma' + K) = \emptyset$$

on veut que  $\text{Supp}(\mathcal{F}(f)) \subset K$ .

- Les hautes fréquences comprises dans la scène  $s$  (par exemple les bords) sont atténuerées par  $g_o$ .
- Si cela n'est pas suffisant (i.e.  $\text{Supp}(\mathcal{F}(f)) \notin K$ ), on peut procéder à un filtrage passe bas. Il suffit de considérer  $h * f$  où  $\text{Supp}(F(h)) \in K$ .
- si on augmente la fréquence d'échantillonnage, il faut diminuer la taille des capteurs → augmentation de la taille du support de  $\mathcal{F}(h_{capt})$  ...

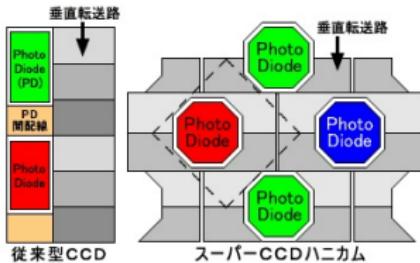
## Exemple 1 : Réseau hexagonal et spectre isotrope

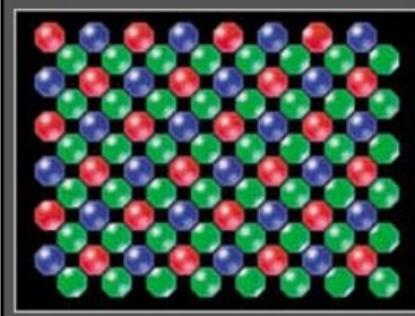
Si l'on suppose que  $\text{Supp}(\mathcal{F}(f))$  est un disque, on montre alors que le réseau hexagonal est le plus efficace (environ 15% de plus que le réseau carré)



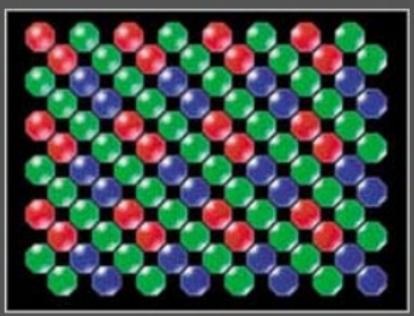
## Exemple 2 : le *super CCD* (Fujifilm)

Agencement en quinquence → possibilité de capteurs de plus grandes dimensions pour une même résolution horizontale et verticale.





Conventional Super CCD color filter array layout



Super CCD EXR color filter array layout

# La transformée de Fourier Discrète

On ne dispose que d'images discrètes :  $\{u_{k,l}\}_{k=1,\dots,M; l=1,\dots,N}$ .

On considère la Transformée de Fourier Discrète (TFD) de  $u_{k,l}$  :

$$\hat{u}_{m,n} = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} u_{k,l} \omega_N^{-mk} \omega_N^{-nl},$$

où  $\omega_N = \exp\left(\frac{2i\pi}{N}\right)$ .

# Une image et sa TFD



# Digression



(a) Image A



(b) Image A

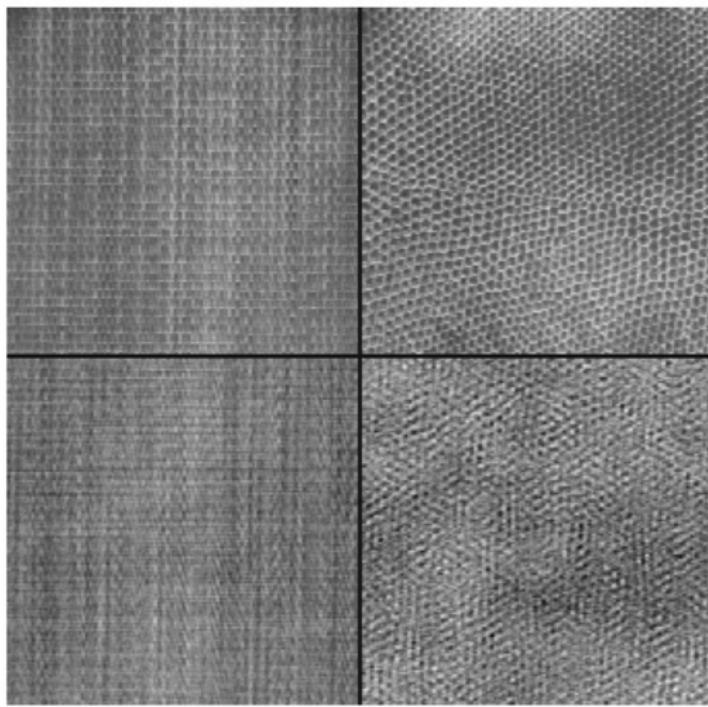


(c) Module de la TF de A et phase de

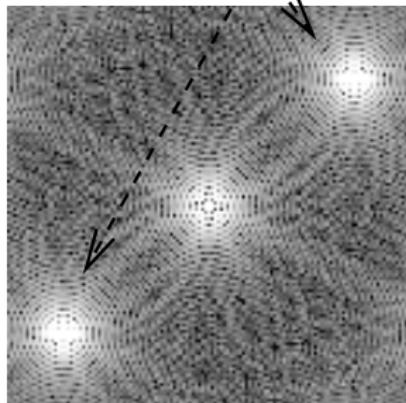
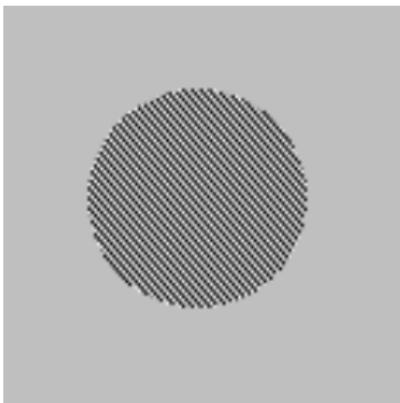
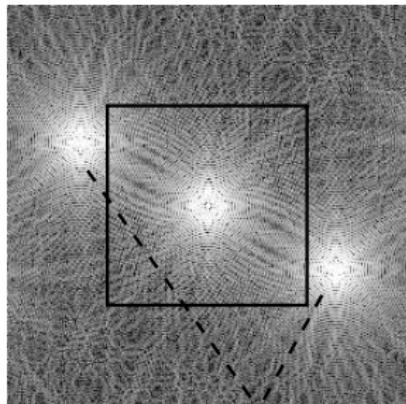
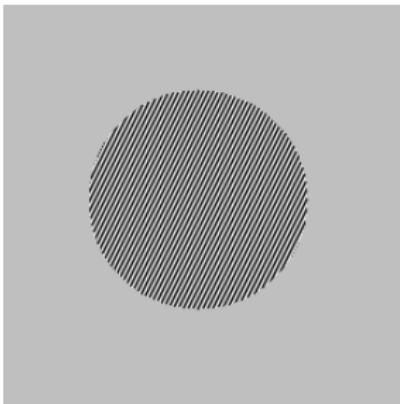


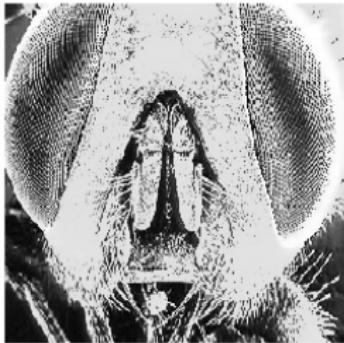
(d) Module de la TF de B et phase de

## Digression (2)

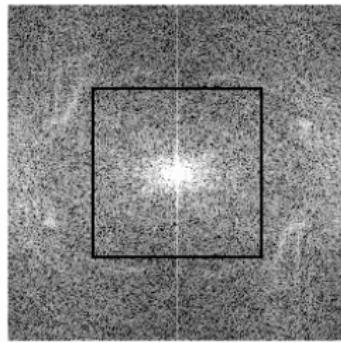


## Un exemple synthétique de repliement





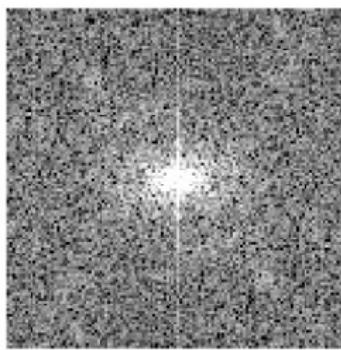
(a) Image originale



(b) Sa TFD, non nulle en dehors du carré visible en surimpression



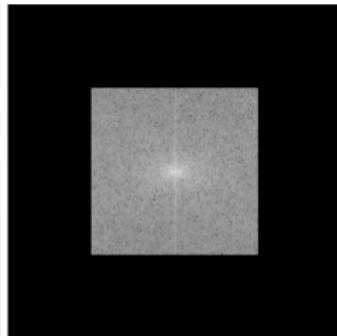
(c) Image sous-échantillonnée d'un facteur 2



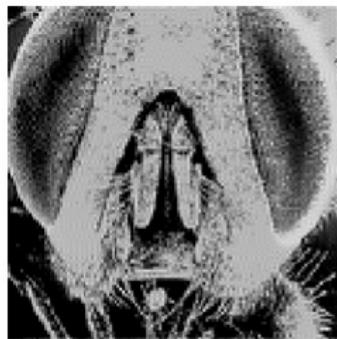
(d) La TFD correspondante, sur laquelle il y a repliement



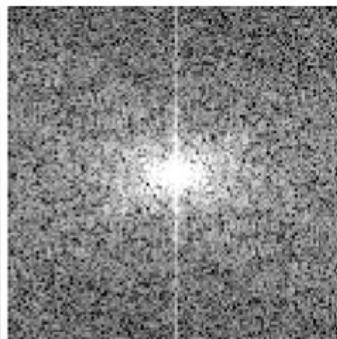
(a) Image obtenue par TFD inverse  
de b



(b) Image obtenue en mettant à zéro  
les hautes fréquences de 11.1.3-a



(c) Sous-échantillonnage : le replie-  
ment a disparu



(d) TFD de c

## Autre exemple



Image issue d'un appareil numérique aux filtres passe-bas (sur les capteurs déficients)



Image issue d'un appareil numérique aux filtres passe-bas (sur les capteurs déficients)



Image issue d'un appareil numérique aux filtres passe-bas (sur les capteurs déficients)

## Autre exemple de repliement



A gauche, image originale ; à droite sous-échantillonnage d'un facteur 2

## Apparition du ringing



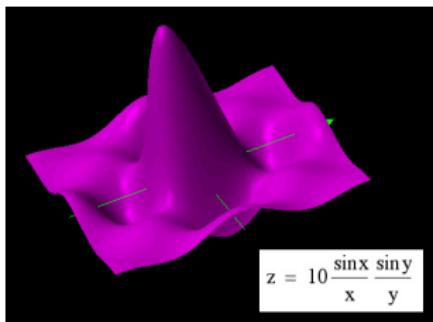
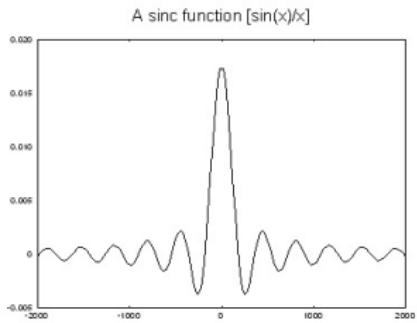
A gauche, image originale ; à droite : sous-échantillonnage d'un facteur 2  
après multiplication par l'indicatrice d'un carré en Fourier

# Le phénomène du ringing

Pour être dans les conditions de Shannon, on peut multiplier le spectre de l'image par  $\mathbb{1}_K$ . L'image résultante est alors

$$\mathcal{F}^{-1}(\mathcal{F}(f)\mathbb{1}_K) = f * \mathcal{F}^{-1}(\mathbb{1}_K).$$

Si  $K$  est la fonction indicatrice d'un carré de côté  $a$ ,  $\mathcal{F}^{-1}(\mathbb{1}_K)$  est le produit de deux fonctions sinc :



# Ringing



Retour à l'acquisition des images :

$$u = (h * s).\Pi_\Gamma.F$$

où

- $s$  scène (fonction de  $L^1$ ),
- $h$  : réponse impulsionnelle du système optique et de l'intégration des capteurs ( $g_{ouv} * g_{flou} * g_{fil} * g_{capt}$ ),
- $\Pi_\Gamma = \sum \gamma \in \Gamma \delta_\gamma$ ,
- $F = \mathbb{1}_\Omega$ ,  $\Omega \in \mathbb{R}^2$ , support de l'acquisition.

$$u = Q[h((g_o * s).\Pi_\Gamma.F + b)]$$

où

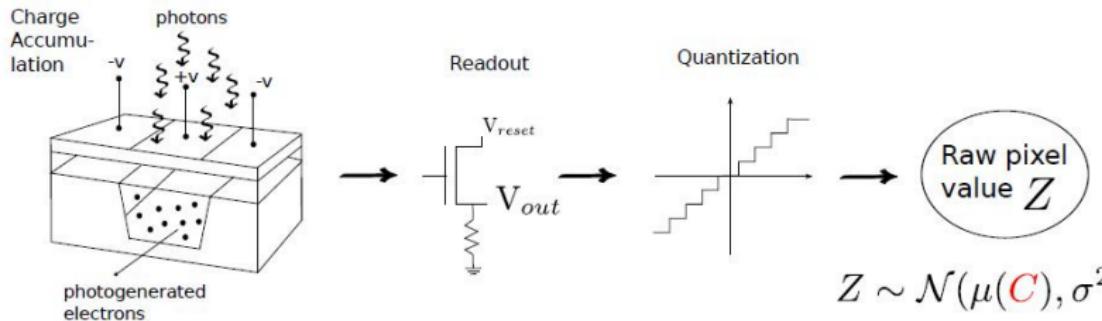
- $s$  scène (fonction de  $L^1$ ),
- $g_o$  : réponse impulsionnelle du système optique et de l'intégration des capteurs ( $g_{ouvr} * g_{flou} * g_{fil} * g_{capt}$ ),
- $\Pi_\Gamma = \sum \gamma \in \Gamma \delta_\gamma$ ,
- $F = \mathbb{1}_\Omega$ ,  $\Omega \in \mathbb{R}^2$ , support de l'acquisition.
- $b$  bruit additif : pour  $(i,j) \in \Gamma$ ,  $\{b(i,j)\}$  est une famille de variables aléatoires i.i.d.
- $h$  est une fonction croissante : un "changement de contraste"
- $Q$  est un opérateur de *quantification*

Pour l'étude détaillée des rôles du changement de contraste et de la quantification :  
→ prochain cours.

On note  $v$  l'image discrète après optique + échantillonnage ;  $u$  l'image bruitée

- Bruit additif :  $u(i,j) = v(i,j) + b(i,j)$   
où les  $b(i,j)$  sont i.i.d.
- Bruit impulsional :  $u(i,j) = v(i,j).A(i,j) + (1 - A(i,j)).B(i,j)$ ,  
où  $A(i,j)$  sont des variables de Bernoulli i.i.d. et  $B(i,j)$  des variables i.i.d.  
("sel et poivre" si  $B$  à valeurs dans  $\{0, M\}$ )
- Bruit multiplicatif (de *chatoiement* ou speckle en anglais) :  
 $u(i,j) = v(i,j).b(i,j)$   
avec les  $b(i,j)$  i.i.d.

# Les causes du bruit



$$Z \sim \mathcal{N}(\mu(C), \sigma^2(C))$$

$C$  irradiance

$\tau$  exposure time

$g$  camera gain

$a$  photo-response non-uniformity factor

$\mu_r, \sigma_r^2$  readout noise mean and variance

$$\mu(C) = ga\tau C + \mu_R$$

$$\sigma^2(C) = g^2 a \tau C + \sigma_r^2$$

Illustration Cecilia Aguerrebere

- **Bruit *shot noise*** Nb de photons émis par la source : loi de Poisson de moyenne  $C\tau$  avec  $C$  radiance (nombre de photons par unité de temps) et  $\tau$  temps d'exposition
- **Courant d'obscurité (Dark current)**  
Emission résiduelle d'électrons d'origine thermique : Poisson de moyenne  $d\tau$  dépendant de  $\tau$
- **Bruit de lecture (readout noise).**  
Erreurs lors de la lecture des électrons (par rapport à un voltage de référence)
- **Non uniformités spatiales**  
de la réponse des photo-senseurs (PRNU)  
du bruit thermique (DCNU)

$$I_{noise} = f([g(Poiss(C\tau) + Poiss(d_\tau))] + N_{out}) + Q$$

avec

- $Poiss(\lambda)$  : variable de Poisson de moyenne  $\lambda$
- $C$  radiance (photons / temps)
- $g$  : gain
- $\tau$  : temps d'acquisition
- $d_\tau$  : valeur moyenne du bruit *dark current*
- $N_{out}$  : bruit de lecture (gaussien de moyenne  $\mu_R$  et variance  $\sigma_R^2$ ).
- $Q$  : bruit de quantification (uniforme)
- $f$  : fonction de réponse de l'appareil

Remarque : on néglige les variations spatiales (DCNU et PRNU)

## Première simplification

Le modèle complet est

$$I_{noise} = f([g(Poiss(C\tau) + Poiss(d_\tau))] + N_{out}) + Q$$

Hypothèses supplémentaires :

- On considère des images RAW, pour lesquelles  $f = Id$  (réponse linéaire au nombre de photons)
- Le bruit de quantification  $Q$  est négligeable devant le bruit de lecture
- Le dark current est négligeable pour temps d'exposition courts ( $< 1s$ )

Alors

$$I_{noise} = gPoiss(C\tau) + N_{out}$$

$$I_{noise} = gPoiss(C\tau) + N_{out}$$

- On a  $P(\lambda) \approx N(\lambda, \lambda)$  pour  $\lambda$  grand (en pratique quelques dizaines)
- En photographie standard : l'approximation gaussienne est suffisante (faux pour e.g. l'astrophotographie)

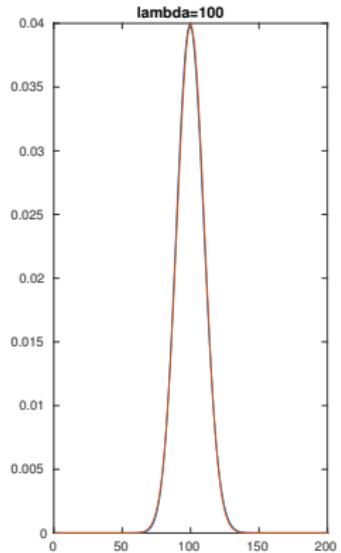
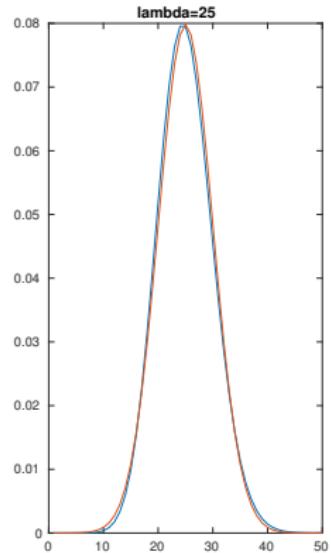
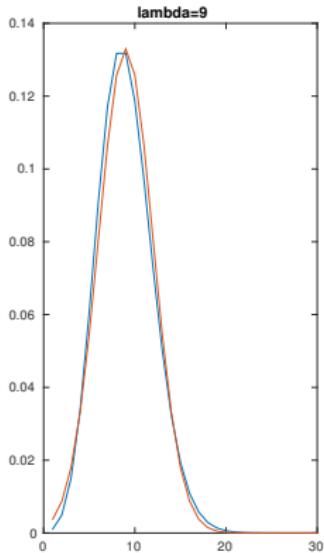
Approximation gaussienne :

$$N(gC\tau + \mu_R, g^2C\tau + \sigma_R^2)$$

avec  $\mu_R$  et  $\sigma_R$  les moyennes et variances du bruit de lecture

Application : création d'images HDR / IMA206

# Approximation gaussienne



Bleu : Poisson ; rouge : normale

Conclusion : la variance du bruit **dépend du signal**.

Pour prendre en compte ce phénomène, deux approches possibles :

- utilisation directe du modèle de bruit (e.g. pour dériver des estimateurs statistiques) : cf cours HDR à venir
- modification du signal pour se ramener à un bruit gaussien i.i.d.  
opération de *stabilisation de la variance* (e.g. transformée de Anscomb)

Compléments :

- *Study of the digital camera acquisition process and statistical modeling of the sensor raw data*, C. Aguerrebere et al., preprint HAL, 2013
- *Optimal inversion of the generalized Anscomb transformation for Poisson-Gaussian noise*, A. Mäkitalo et A. Foi, IEEE Image Processing, 22, 1, pp.91-103, 2013