



IMAGES NUMÉRIQUES : ACQUISITION

Julie Delon

MAP5, Université Paris Descartes

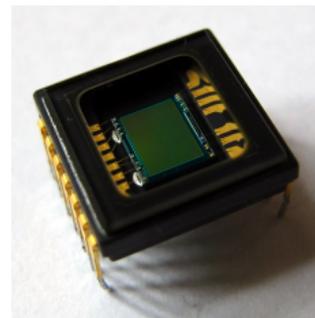
M2 Mathématiques et modélisation

1 Introduction

2 Modèles d'acquisition d'images

QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?

QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?

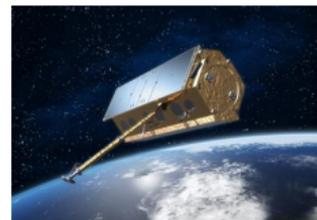
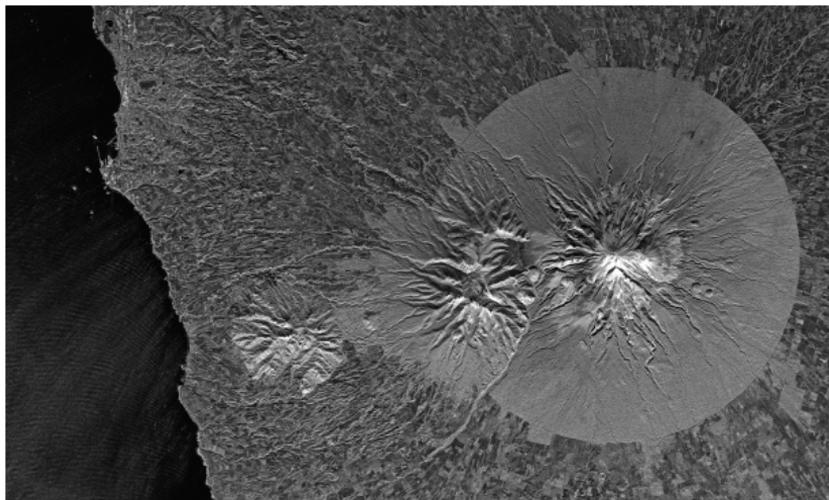


©Disney Studios, RED

QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



©DLR, EADS Astrium

QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



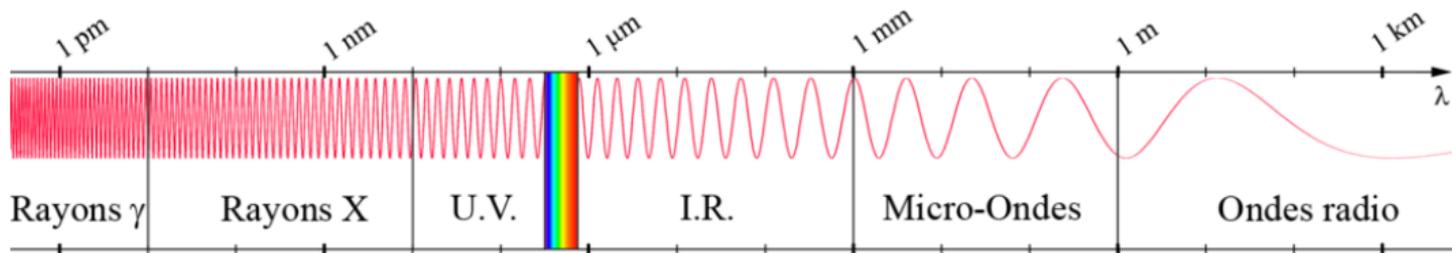
QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



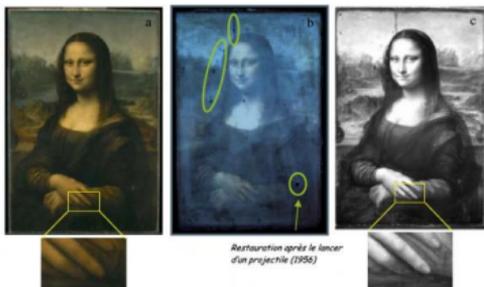
QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



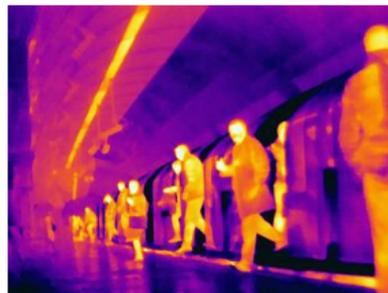
Spectre électromagnétique



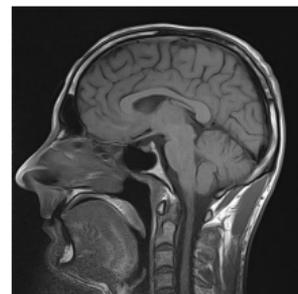
rayons X



visible, UV, IR

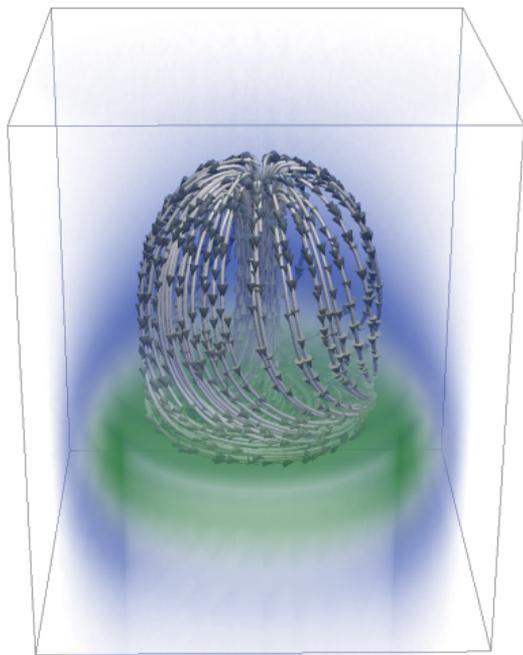


infrarouge



ondes radio

QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?

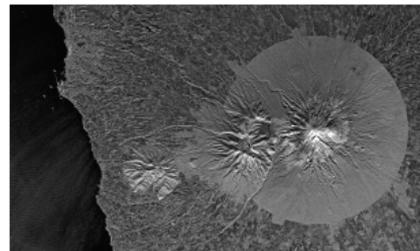
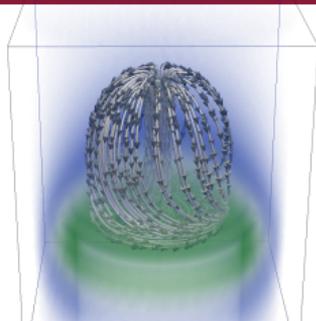


QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



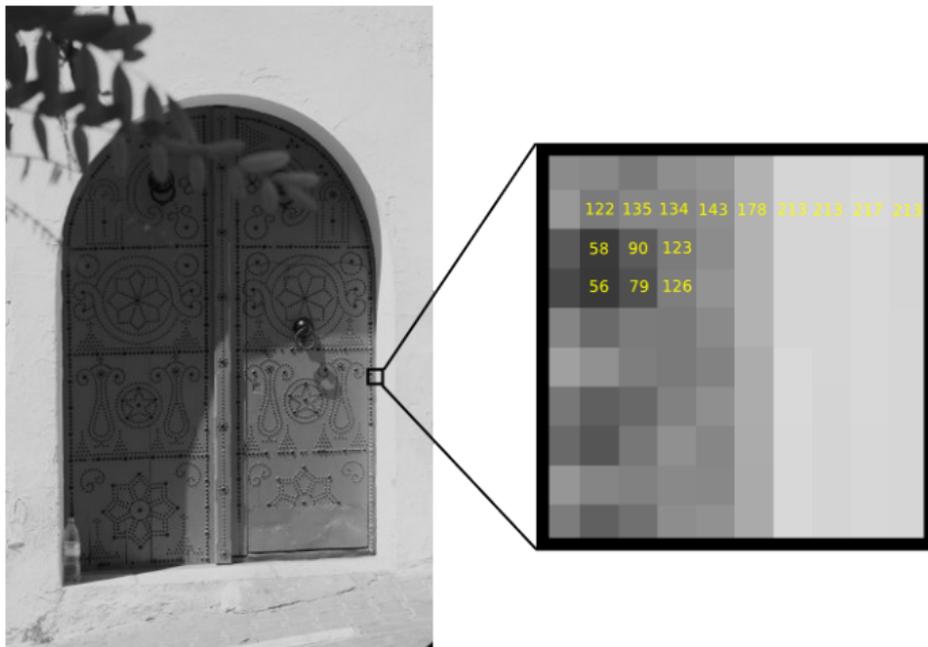
©Disney studio, gettyimages

QU'EST-CE-QUE L'IMAGERIE NUMÉRIQUE ?



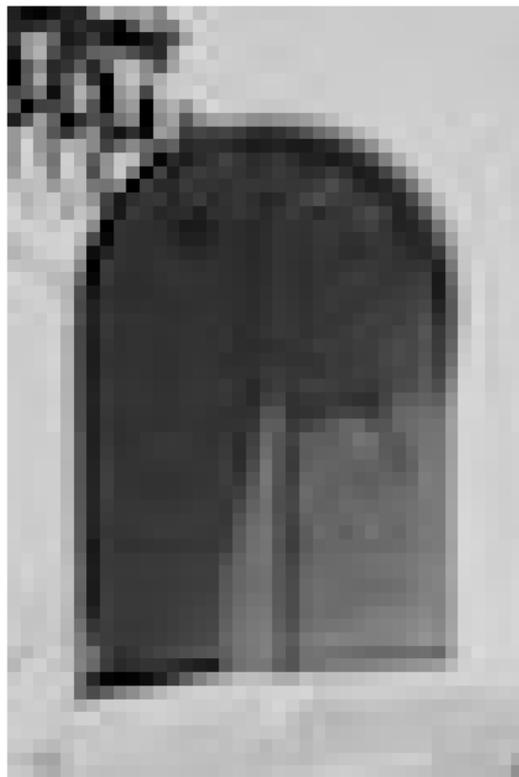


$u \in \mathbb{R}^{|\Omega|}$ avec $\Omega \subset \mathbb{Z}^2$, rectangle de N lignes et M colonnes.



$u \in \mathbb{R}^{|\Omega|}$ avec $\Omega \subset \mathbb{Z}^2$, rectangle de N lignes et M colonnes.

- **Résolution** : nombre de pixels de l'image ⚠
- **Quantification** : nombre de valeurs possibles que peuvent prendre les pixels de l'image
cas typique : 8 bits = $\{0, \dots, 255\}$



A gauche, image de taille 324×486 , à droite taille divisée par 8 dans chaque dimension.

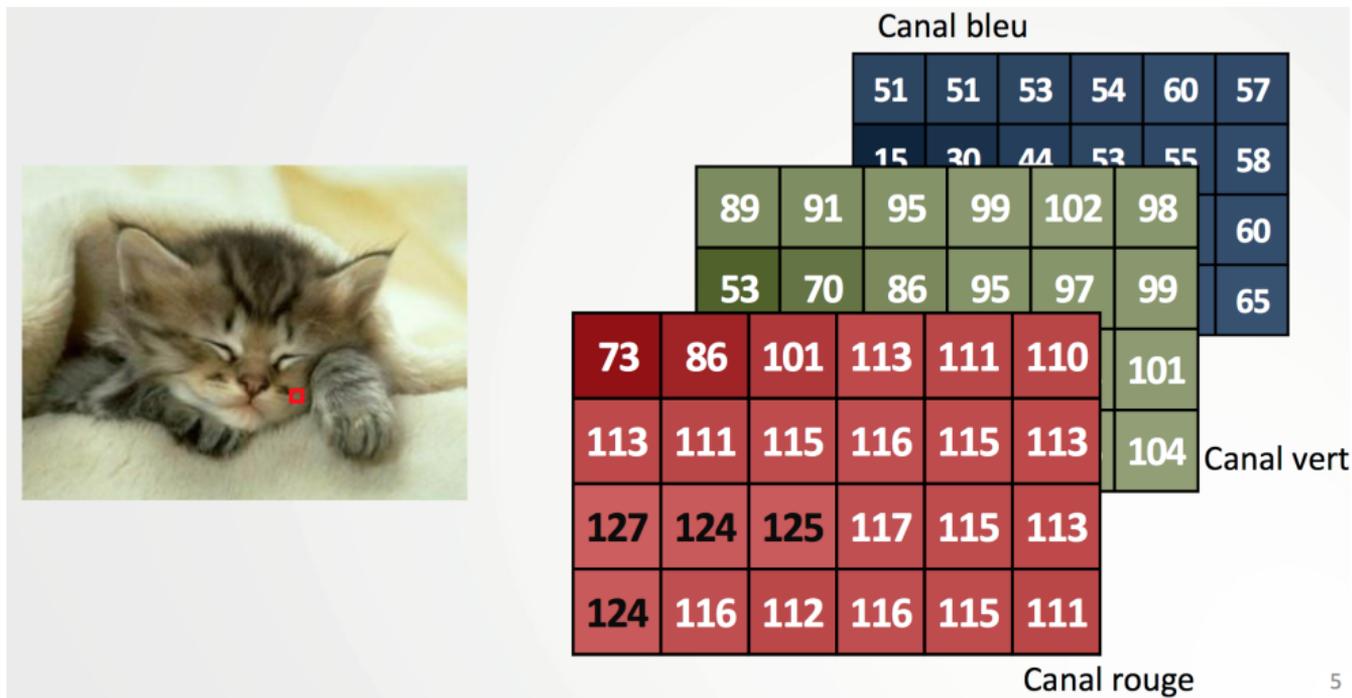


A gauche, image avec 256 niveaux de gris, à droite image avec 3 niveaux de gris.



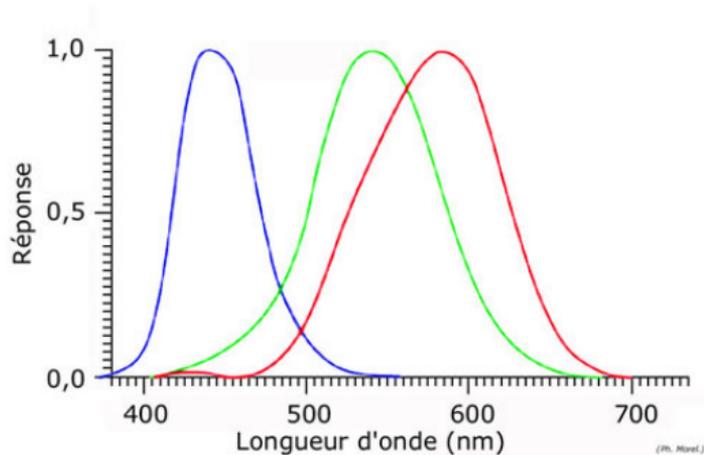
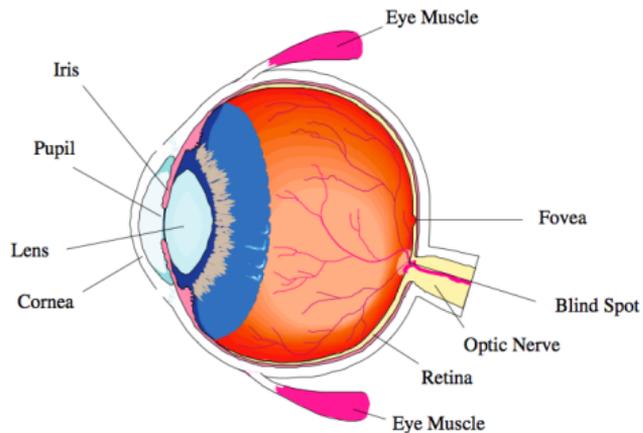
A gauche, image avec 256 niveaux de gris, à droite image avec 3 niveaux de gris.

→ Pour aller plus loin : <http://images.math.cnrs.fr/Le-traitement-numerique-des-images.html>



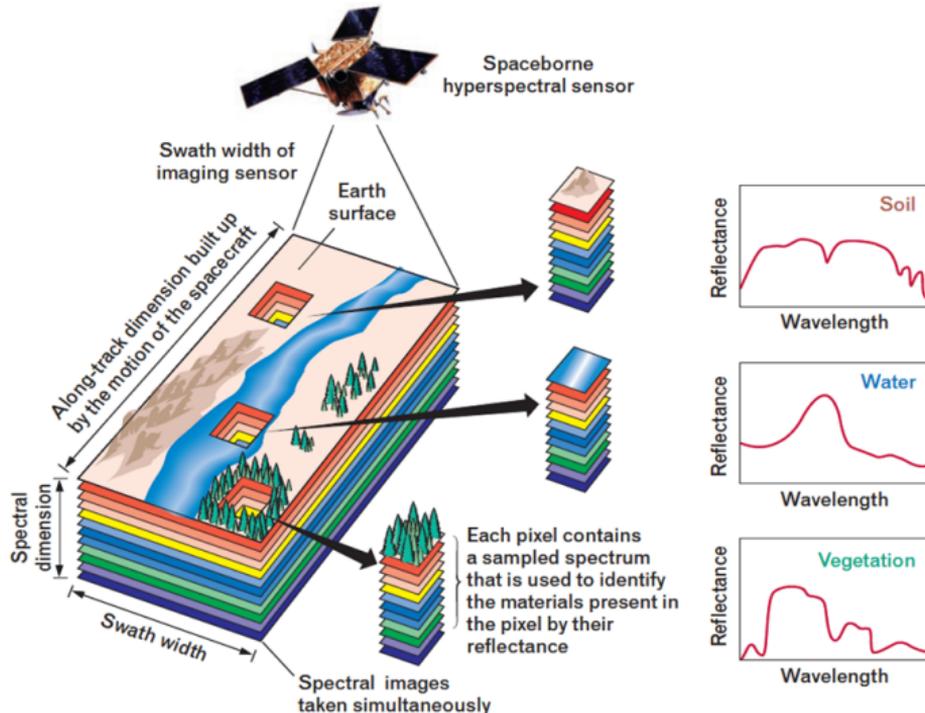
5

$u \in \mathbb{R}^{|\Omega| \times 3}$ avec $\Omega \subset \mathbb{Z}^2$ domaine rectangulaire discret de l'image



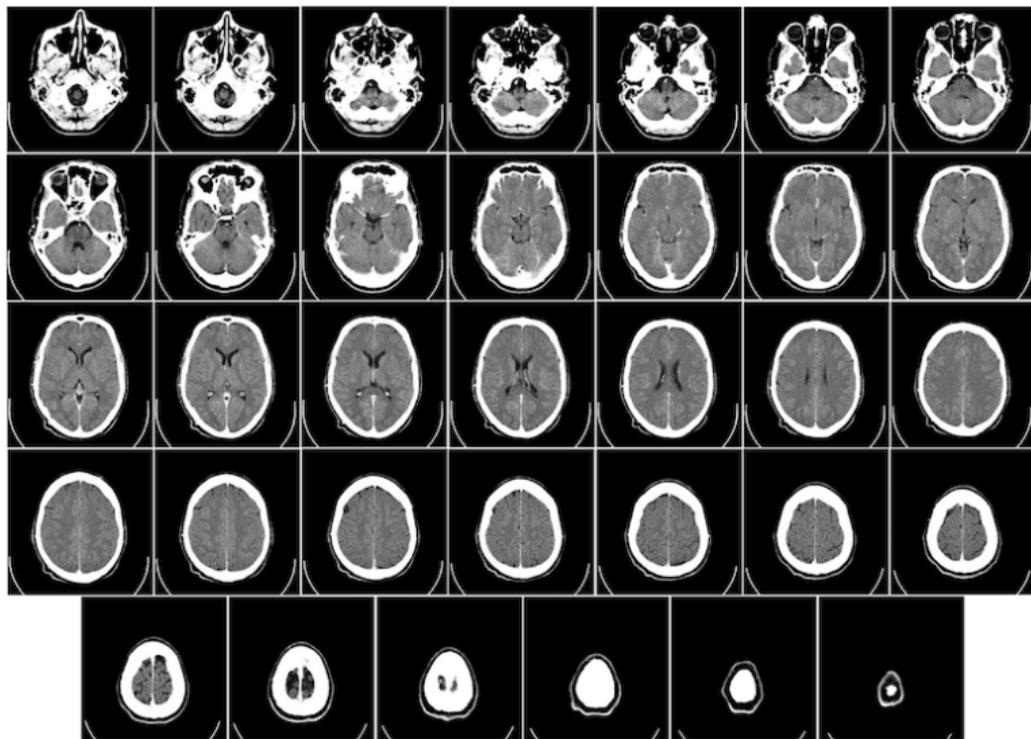
Spectre visible : longueurs d'ondes entre $\lambda_{min} = 380\text{nm}$ et $\lambda_{max} = 700\text{nm}$.





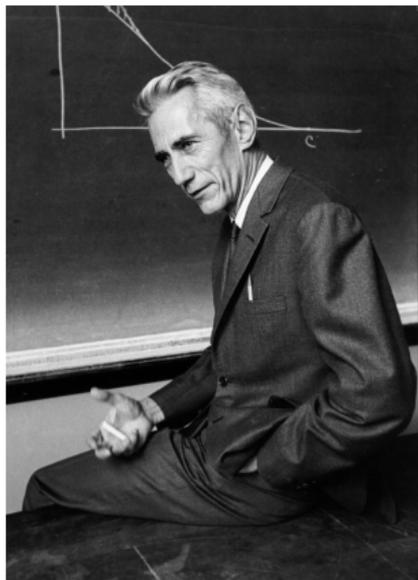
©<http://www.markelowitz.com/Hyperspectral.html>

Image **hyperspectrale** : acquisition de plusieurs longueurs d'onde, dans ou hors du spectre visible, $u \in \mathbb{R}^{|\Omega| \times H}$ avec H la dimension spectrale



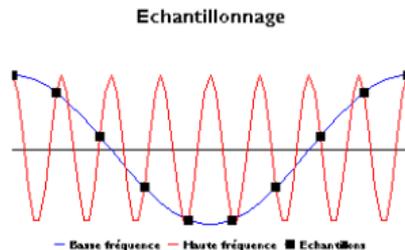
©Wikimedia Commons

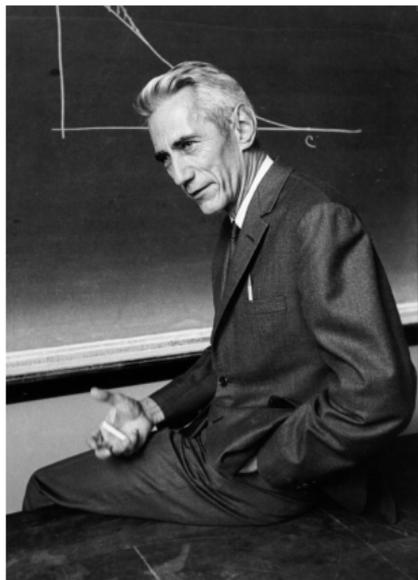
$u \in \mathbb{R}^{|\Omega|}$ avec $\Omega \subset \mathbb{Z}^3$



Claude Shannon (1916-2001)

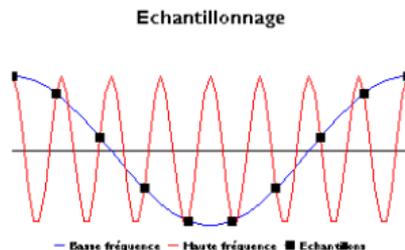
Combien faut-il d'échantillons pour représenter un signal continu ?



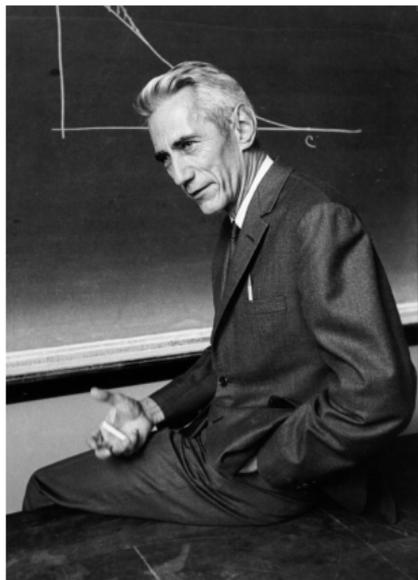


Claude Shannon (1916-2001)

Combien faut-il d'échantillons pour représenter un signal continu ?

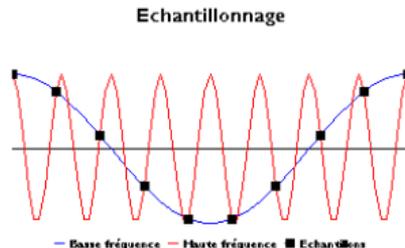


Théorème d'échantillonnage (1949) : La fréquence d'échantillonnage doit être plus grande que le double de la fréquence maximale présente dans le signal.



Claude Shannon (1916-2001)

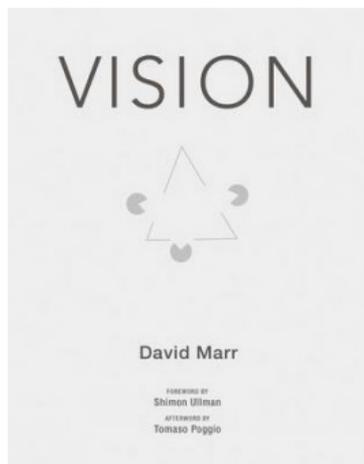
Combien faut-il d'échantillons pour représenter un signal continu ?



Théorème d'échantillonnage (1949) : La fréquence d'échantillonnage doit être plus grande que le double de la fréquence maximale présente dans le signal.

Traitement de l'image : discipline dont le but est de diriger et contrôler la qualité de l'acquisition de l'image et de la restaurer.

Quelles maths ? : analyse spectrale, distributions, optimisation, probabilités, estimation statistique, etc.



David Marr (1945-1980)

Vision par ordinateur (branche de l'IA, début des 80's)

- Discipline scientifique issue de la psychophysique, de la neurophysiologie, des mathématiques et de l'informatique
- But : algorithmes partant d'images ou de vidéos brutes et arrivant à une interprétation sémantique de la scène, permettant des actions automatiques

Applications : analyse de scène et alertes en vidéo surveillance, recherche d'une image par son contenu, planification du mouvement d'un robot, conduite automatique, etc.

Quelles maths ? : géométrie projective, statistique, optimisation, apprentissage, etc.



A word cloud of various imaging disciplines. The words are arranged in a roughly horizontal shape, with 'Vision par ordinateur' and 'Graphisme' being the largest. Other prominent words include 'Traitement d'image', 'Imagerie médicale', and 'Imagerie sismique'. The colors range from dark red to light orange.

Interférométrie Imagerie infrarouge Imagerie sismique
Traitement d'image Imagerie médicale
Visualisation scientifique Graphisme Vision par ordinateur
Imagerie satellitaire Imagerie biologique Imagerie aérienne Photographie Computationnelle
Imagerie radar



Vision par ordinateur + Informatique graphique + Traitement d'images



Géométrie projective, Algèbre linéaire, Synthèse de texture (Probas / Stat), etc...

1 Introduction

2 Modèles d'acquisition d'images

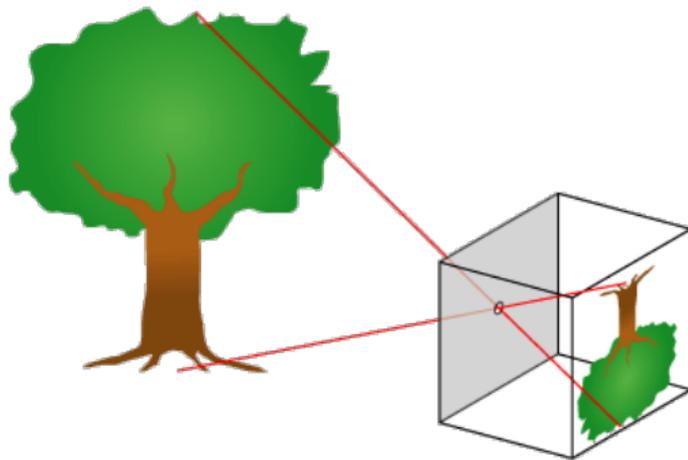
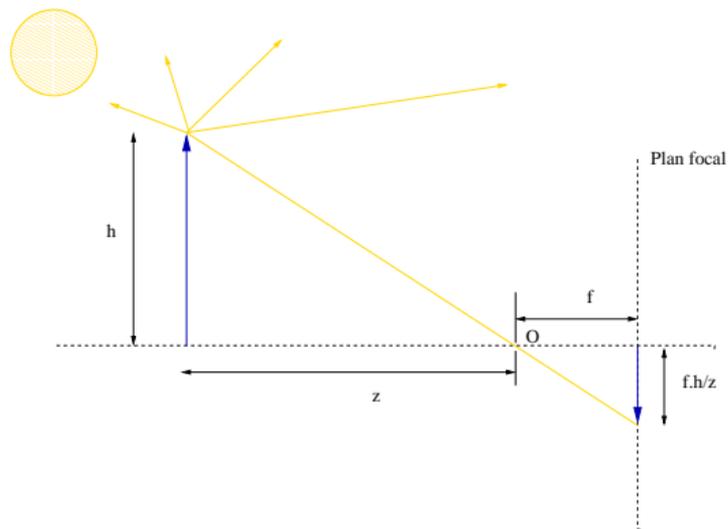


Illustration from wikipedia.org

(Aristote, Alhazen 10^{eme} s., Brunelleschi début 15^{eme} s., Da Vinci 1500, Kepler 1604 *camera obscura*)

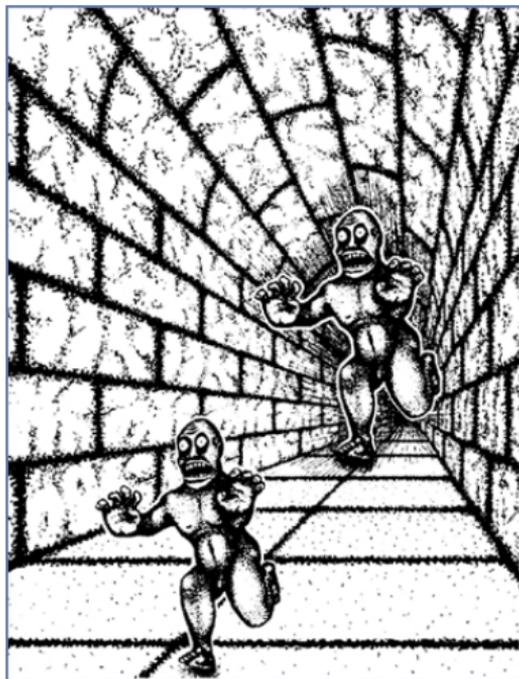


Victoria Cooper - Doug Spowart - creative commons

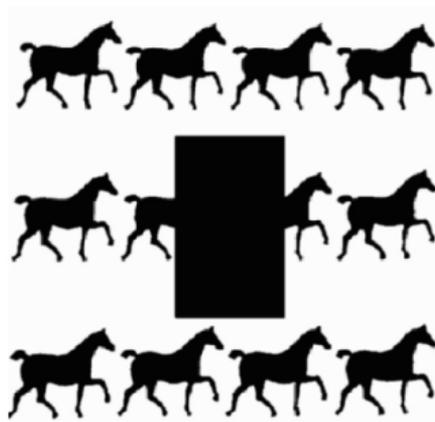


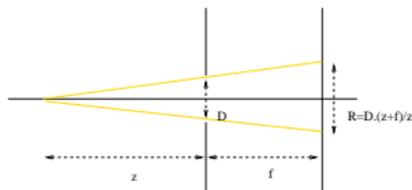
Une partie de la lumière issue de l'objet passe par l'ouverture O et se projette sur le plan focal. La distance f est appelée distance focale.

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



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





Ouverture O finie : rayons \rightarrow cônes \rightarrow flou

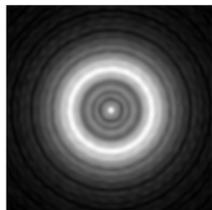
Pour une ouverture de diamètre D et S la scène (à distance z), on observe (en première approximation)

$$g_z * S,$$

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

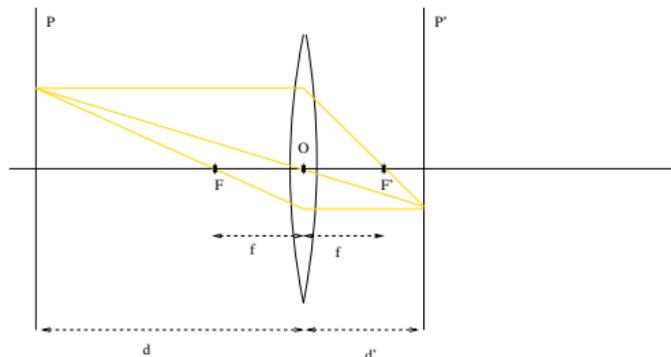
Modélisation plus fine : prise en compte de la diffraction





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).



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.

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



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

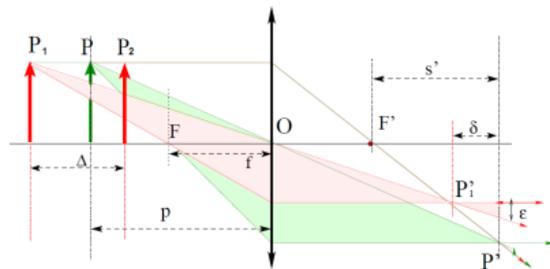


Figure par H. Maître

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

$$\Delta = \frac{4\epsilon D p f (p - f)}{4D^2 f^2 - \epsilon^2 p^2}$$

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

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

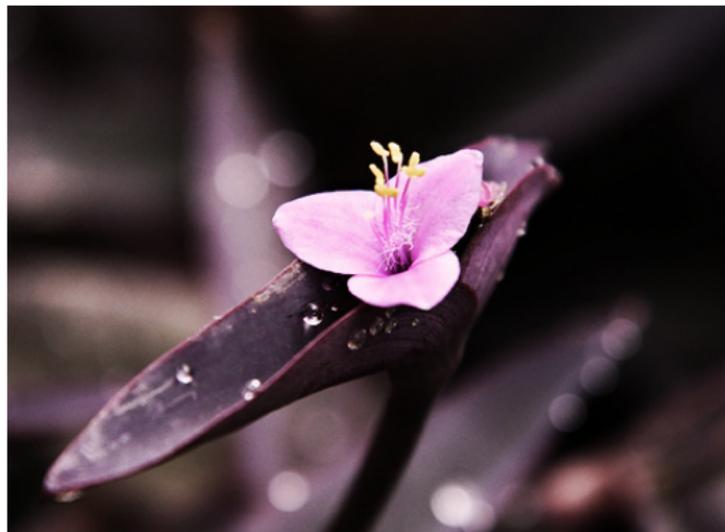
ouverture D diamètre du diaphragme ; f distance focale ; nombre d'ouverture $N = f/D$
(généralement progression géométrique)



images 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



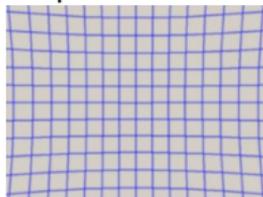
f/2.8 - by bahramr - creative commons



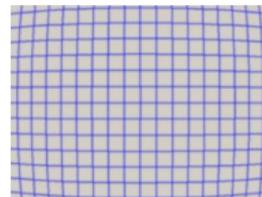
Caméra Lytro

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

- Distorsion géométriques

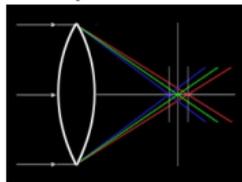


"en coussinet" :

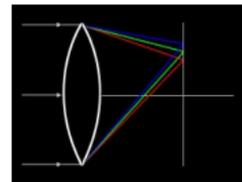


"en barillet" :

- Aberrations chromatiques



"longitudinale" :



"transversale" :

- Vignetage



Aberration latérale



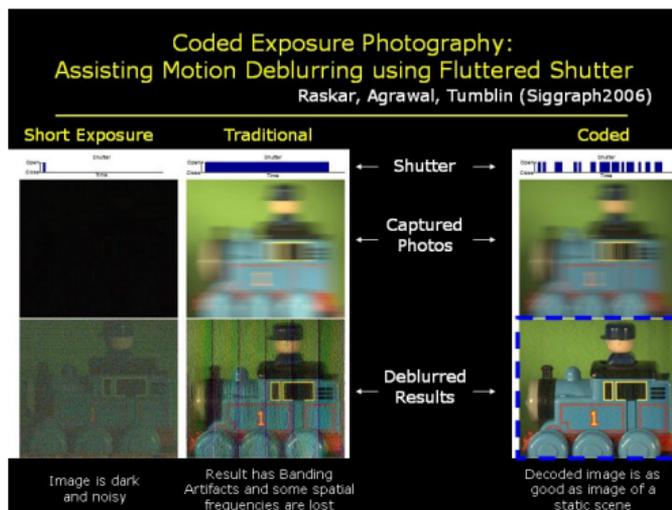


Avant et après correction du vignettage
dxo.com

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

Alternative à l'obturation unique : l'ouverture papillonante (flutter shutter) → démo ipol



- Après passage dans le système optique \rightarrow 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

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

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

Retour à l'acquisition des images :

$$u = (g_o * s). \Pi_{\Gamma}. F$$

où

- s scène (fonction de L^1),
- g_o : 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 ($g_{ouv} * g_{flou} * g_{fil} * g_{capt}$),
- $\Pi_\Gamma = \sum \gamma \in \Gamma \delta_\gamma$,
- $F = \mathbb{1}_\Omega$, $\Omega \in \mathbb{R}^2$, support de l'acquisition.
- 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 non linéaire (un "changement de contraste")
- Q est un opérateur de *quantification*

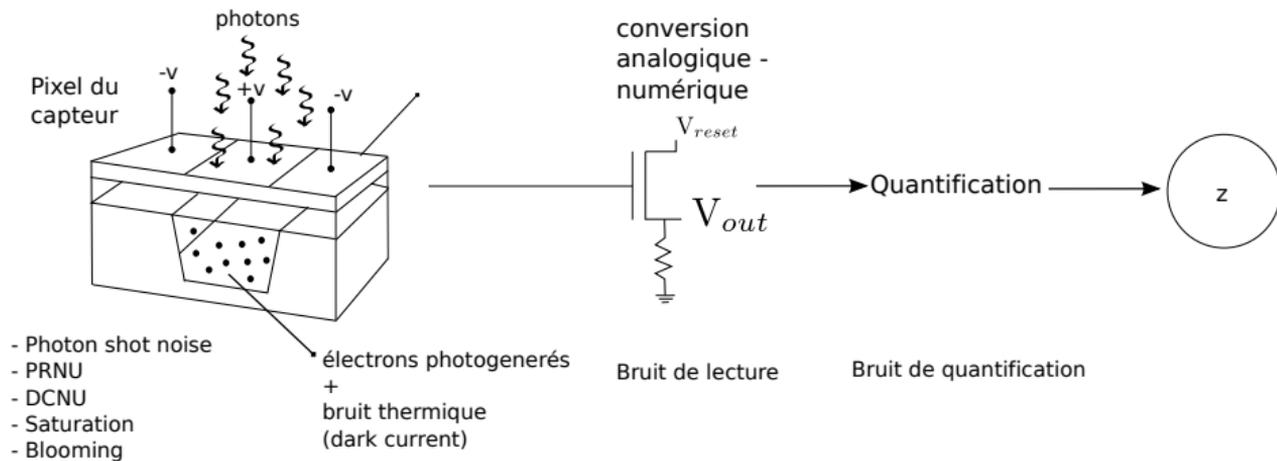


Illustration par Cecilia Aguerrebere

- **Bruit *shot noise*** Nb de photons emis par la source : loi de Poisson de moyenne $C\tau$ avec C radiance (nombre de photons par unité de temps) et τ temps d'exposition
- **Courant d'obscurité (Dark current)**
Emission résiduelle d'électrons d'origine thermique : Poisson de moyenne $d\tau$ dépendant de τ
- **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 λ
- C radiance (photons / temps)
- g : gain
- τ : temps d'acquisition
- d_τ : valeur moyenne du bruit *dark current*
- N_{out} : bruit de lecture (gaussien de moyenne μ_R et variance σ_R^2).
- Q : bruit de quantification (uniforme)
- f : fonction de réponse de l'appareil

Remarque : on néglige les variations spatiales

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

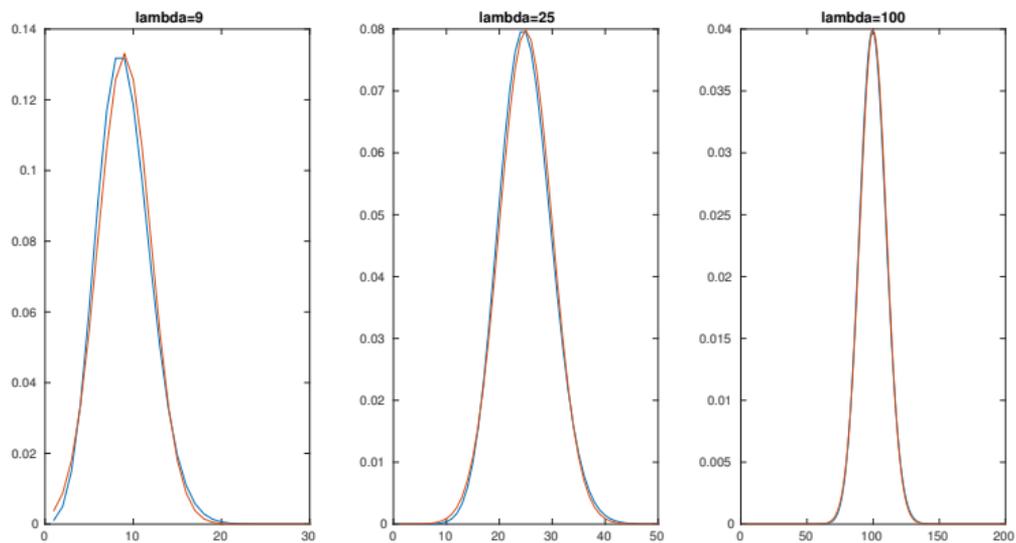
- On a $P(\lambda) \approx N(\lambda, \lambda)$ pour λ grand (en pratique quelques dizaines)
- En photographie standard : l'approximation gaussienne est suffisante (faux pour e.g. l'astrophotographie)
- Bruit de quantification négligeable devant le bruit de lecture
- Dark current négligeable pour temps d'exposition courts ($< 1s$)

Approximation gaussienne :

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

avec μ_R et σ_R les moyennes et variances du bruit de lecture

Application : création d'images HDR (cours à venir)



Bleu : Poisson ; rouge : normale

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

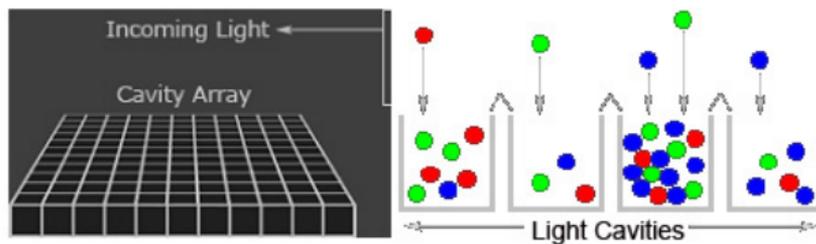
Pour prendre en compte ce phénomène, deux approches possibles :

- 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 Anscombe)

Compléments :

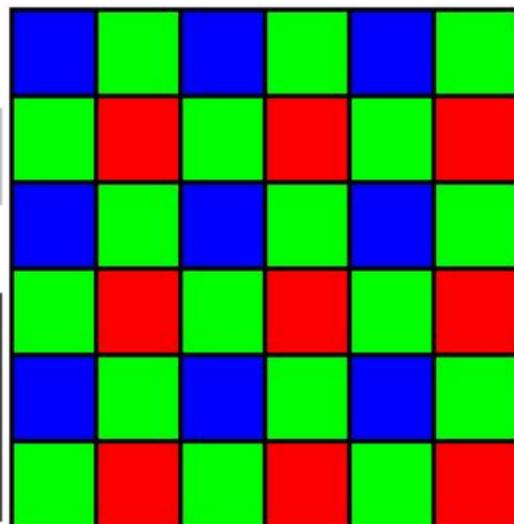
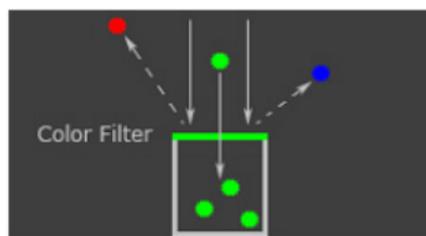
- *Study of the digital camera acquisition process and statistical modeling of the sensor raw data*, C. Aguerrebere et al., preprint HAL, 2013
- *Optimal inversion of the generalized Anscomb transformation for Poisson-Gaussian noise*, A. Mäkitalo et A. Foi, IEEE Image Processing, 22, 1, pp.91-103, 2013

Et la couleur ?



courtesy of cambridgeincolour.com

Color Filter Array



Exemple d'image obtenue par la trame de Bayer (par ex. fichier RAW d'un reflex numérique).



Fichier RAW (12 bits) compressé sans perte



Fichier TIFF ou JPEG 24 bits.

Interpolation possible



Fichier RAW (12 bits) compressé sans perte



Fichier TIFF ou JPEG 24 bits.

① Modèle d'acquisition : de la scène à l'image RAW

$$s \text{ scène} \xrightarrow[\text{+capt+flou}]{\text{PSF+mouvement}} s * g_o \xrightarrow[\text{quantization}]{\text{sampling}} u = Q((s * g). \Pi_{\Gamma}. F + b)$$

① Modèle d'acquisition : de la scène à l'image RAW

$$s \text{ scène} \xrightarrow[\text{+capt+flou}]{\text{PSF+mouvement}} s * g_o \xrightarrow[\text{quantization}]{\text{sampling}} u = Q((s * g). \Pi_{\Gamma}. F + b)$$

② Demosaicking de u

Acquisition d'images par un appareil photo numérique (suite)

1 Modèle d'acquisition : de la scène à l'image RAW

$$s \text{ scène} \xrightarrow[\text{+capt+flou}]{\text{PSF+mouvement}} s * g_o \xrightarrow[\text{quantization}]{\text{sampling}} u = Q((s * g) \cdot \Pi_{\Gamma} \cdot F + b)$$

2 Demosaicking de u

3 De l'image RAW à l'image RGB

$$u \xrightarrow{\text{white balance}} T_w \cdot u \xrightarrow[\text{sRGB}]{\text{linear}} T_s T_w \cdot u \xrightarrow{\text{tone mapping}} h[T_s T_w \cdot u] \xrightarrow[\text{denoising}]{\text{+compression}} u_{jpg}$$

