

Commmande par vision

EXPLOITATION DE L'INFORMATION VISUELLE

JACQUES.GANGLOFF@UNISTRA.FR

Plan du chapitre

Exploitation de l'information visuelle

- Technologie des caméras
 - Les technologies de capteurs
 - Transfert de l'information visuelle
- Les dispositifs d'acquisition d'images
 - Technologies d'interfaces
 - Accès à l'image numérisée
 - Modèle dynamique du processus d'acquisition
- Critères de qualité de l'image

Technologie des caméras

Mesure de la luminance

- Les capteurs à semi-conducteurs sont constitués d'une matrice de photo-sites ou pixels.
- Cette matrice peut être rectangulaire (caméra) ou linéaire (capteur à balayage, ex: scanner).
- La lumière est convertie en charges électriques au niveau de chaque pixel : chaque photon produit un certain nombre de paires électron-trou.
- Les charges sont intégrées durant un certain temps puis sont converties en tension.
- Le temps d'intégration est réglable au moyen d' un obturateur électronique (*electronic shutter*).

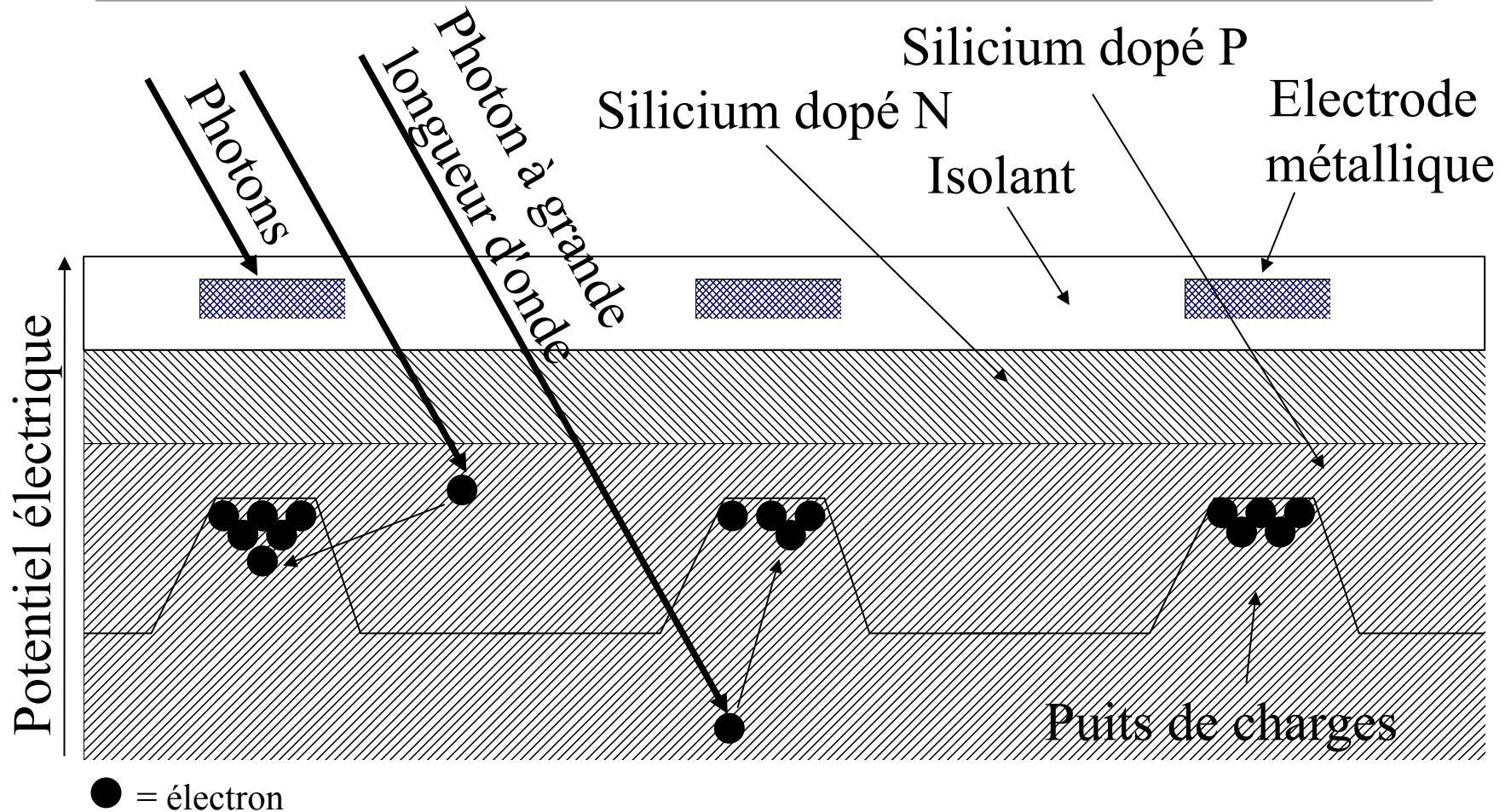
Technologie des caméras

Mesure de la luminance

- Les capteurs à semi-conducteurs mesurent la luminance. Pour reconstituer l'information de chrominance, il faut former l'image sur 3 capteurs munis de filtres laissant passer chacun une longueur d'onde différente (en général rouge, vert et bleu).
- La façon dont les charges sont stockées puis converties en tension diffère suivant les technologies

Technologie des caméras

Mesure de la luminance par photodiodes



Technologie des caméras

Mesure de la luminance par photodiodes

- Lorsque l'intensité lumineuse est ponctuellement très importante, le trop-plein de charge peut déborder d'un puits vers les puits adjacents (*blooming*). Il existe des dispositifs de drainage de ces charges excédentaires (*anti-blooming*).
- Les rayons infra-rouge (grande longueur d'onde) ont tendance à pénétrer plus profondément dans le substrat du capteur. Ils génèrent ainsi des charges qui peuvent, suivant l'angle incident, être capturées par des puits voisins produisant ainsi une image floue (*blurr*).

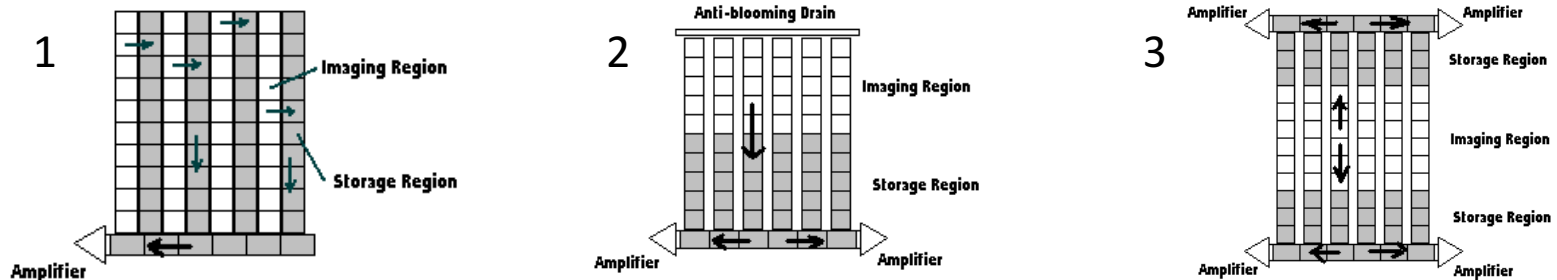
Technologie des caméras

Capteur CCD

- CCD = **C**harge **C**oupled **D**evice
- Chaque photon incident crée une paire électron-trou
- L'électron est capturé par un champ électrique dans un puits de charge
- Les électrons sont accumulés dans les puits durant le temps d'exposition puis sont transférés, **tous en même temps** et très **rapidement**, dans des registres à décalage avant d'être **convertis séquentiellement** en tension.
- Il existe 3 architectures de transfert de charges : par interligne, par image et par portion d'image.
- La technologie CCD est ancienne et tend à disparaître.

Technologie des caméras

Capteur CCD



1. Transfert par inter-ligne (*Interline transfer*) : les zones de stockage sont entrelacées avec les zones recevant la lumière : moins de la moitié de la surface du capteur est sensible à la lumière → le facteur de remplissage (*fill factor*) est $< 50\%$.
2. Transfert par image (*Frame transfer*) : l'image est transférée très rapidement vers une zone de stockage adjacente à la zone sensible → le facteur de remplissage est proche de 100%.
3. Transfert par portion d'image (*Split-frame transfer*) : afin d'augmenter la rapidité de transfert et diminuer ainsi l'ajout intempestif de charges dans des mauvais pixels pendant le transfert (*smear*) on transfère simultanément plusieurs portions de l'image.

Technologie des caméras

Capteur CCD

Caractéristiques des capteurs CCD

- Tous les pixels sont échantillonnés en même temps.
- Les charges sont converties séquentiellement en tension.
- On ne peut accéder à un pixel donné qu'après avoir décalé tous les pixels qui le précède.
- L'électronique d'amplification est située en-dehors du capteur.
- Il y a un nombre limité d'amplificateurs.
- Le transfert vers les registres à décalage génère du « *smear* ».
- Les charges sont transférée analogiquement par des registres à décalage : c'est un transfert avec pertes.
- Les capteurs CCD consomment beaucoup d'énergie.

Technologie des caméras

Capteur CMOS

- CMOS = **C**omplementary **M**etal **O**xide **S**emi-conductor
- Le principe physique utilisé est le même que pour les capteurs CCD : les photons incidents sont convertis en une paire électron-trou.
- La grande différence réside dans la façon de convertir les charges en tension :
 - Avec la technologie CCD, les charges sont transférées puis converties séquentiellement en tension
 - Avec la technologie CMOS :
 - Les charges sont converties en tension directement au niveau du pixel grâce à un réseau de transistors CMOS.
 - Il est possible de réaliser des prétraitements au niveau du pixel.
 - Il est possible d'accéder directement à la tension de chaque pixel grâce à un système d'adressage.

Technologie des caméras

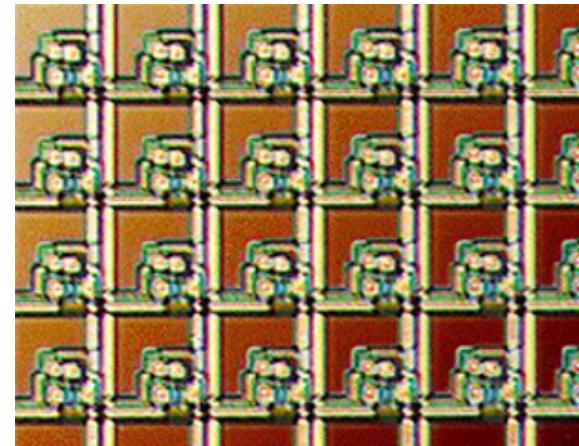
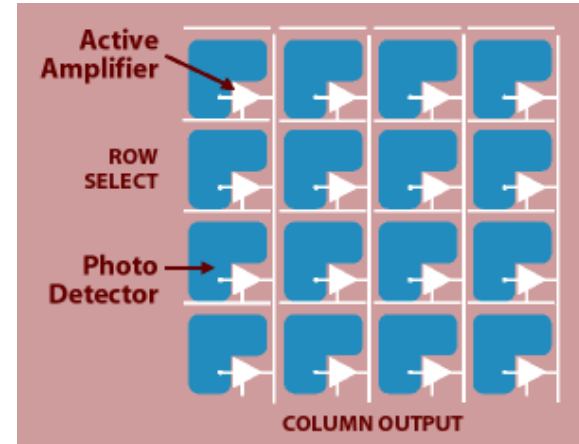
Capteur CMOS

Avantages

- Emploi la technologie CMOS de fabrication de circuit : réduction des coûts
- Faible consommation (100 fois moins que la technologie CCD)
- Accès aléatoire aux pixels comme une mémoire RAM
- Par d'effet *smear*
- Possibilité de réaliser des fonctions évoluées sur le même substrat : filtrage, générateur d'adresse, conversion analogique-numérique, ...

Inconvénients

- Bruit sur les pixels (tend à diminuer)
- Facteur de remplissage moyen



Technologie des caméras

Capteurs particuliers

Vision chip

- Chaque pixel embarque un CPU.
- Le capteur est également un processeur de signal 2D parallèle.
- Tend à disparaître car difficile à programmer.

Capteur plénoptique

- Capteur recouvert de microlentilles
 - Grande profondeur de champ.
 - Permet des acquisition à grande ouverture donc faible lumière.
 - La mise au point peut être logicielle.
 - Correction logicielle d'aberrations.
 - Possibilité d'image 3D
- Exemples de photos : <https://pictures.lytro.com/>

Technologie des caméras

Capteurs particuliers

Capteur Time of flight

- Une source lumineuse pulsée permet en plus de la luminance, de mesurer la distance du point de la scène par mesure du temps de vol de la lumière.



Technologie des caméras

Echantillonnage

Global shutter

- Entrelacé : les lignes paires puis impaires sont échantillonnées alternativement. Chaque $\frac{1}{2}$ image est appelée une trame.
- Intégration par trame : l'accumulation des charges est active uniquement pendant la période d'une trame.
- Intégration par image : l'intégration se fait pendant la période d'une image.
- Progressif : tous les pixels de l'image sont échantillonnés simultanément.

Rolling shutter

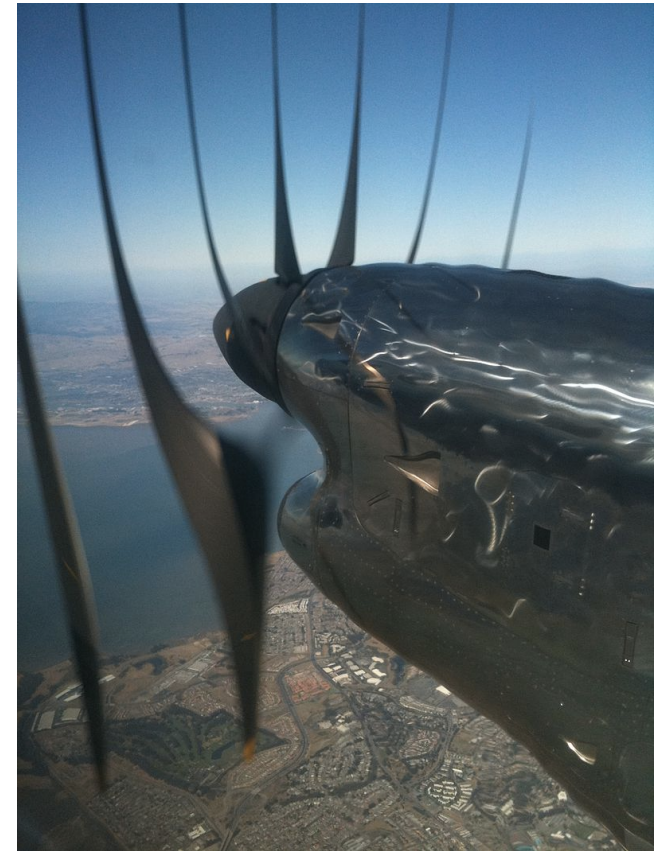
- Les pixels sont échantillonnés en balayant l'image de gauche à droite et de haut en bas.

Technologie des caméras

Echantillonnage

Effets du *rolling shutter*

- Permet d'augmenter la sensibilité.
- Propre à la technologie CMOS.



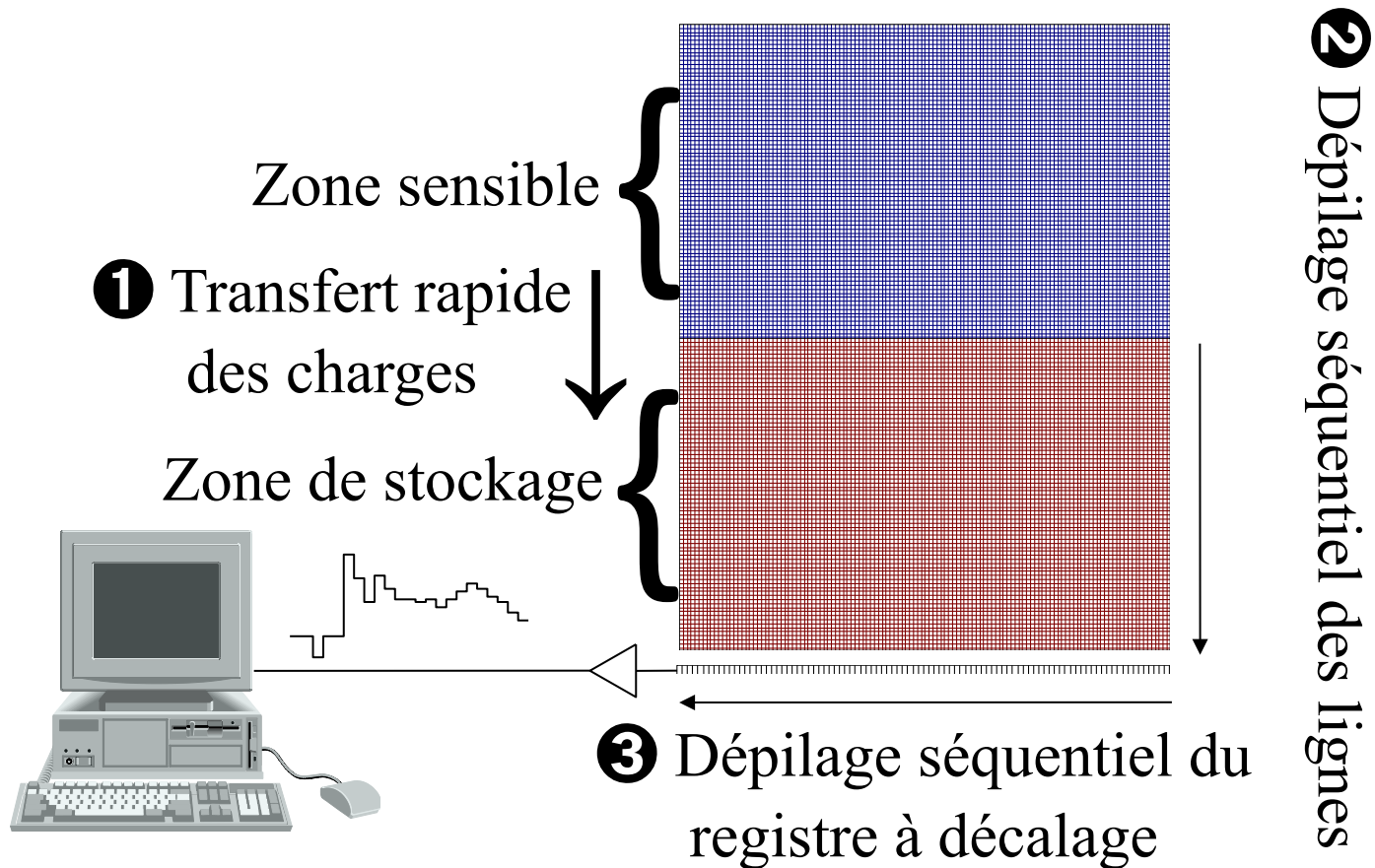
Technologies des caméras

Transfert de l'information

- L'image doit être transférée vers un ordinateur en vue de son traitement.
- L'information analogique doit être échantillonnée en temps et en amplitude.
- L'échantillonnage peut être fait dans l'ordinateur grâce à une carte d'acquisition. Dans ce cas, l'information est transmise depuis la caméra par voie analogique (habituellement un câble coaxial). Cette approche tend à disparaître.
- L'échantillonnage peut intervenir dans la caméra puis l'information numérisé est transmis à l'ordinateur via un bus de données. Cette approche tend à se généraliser.

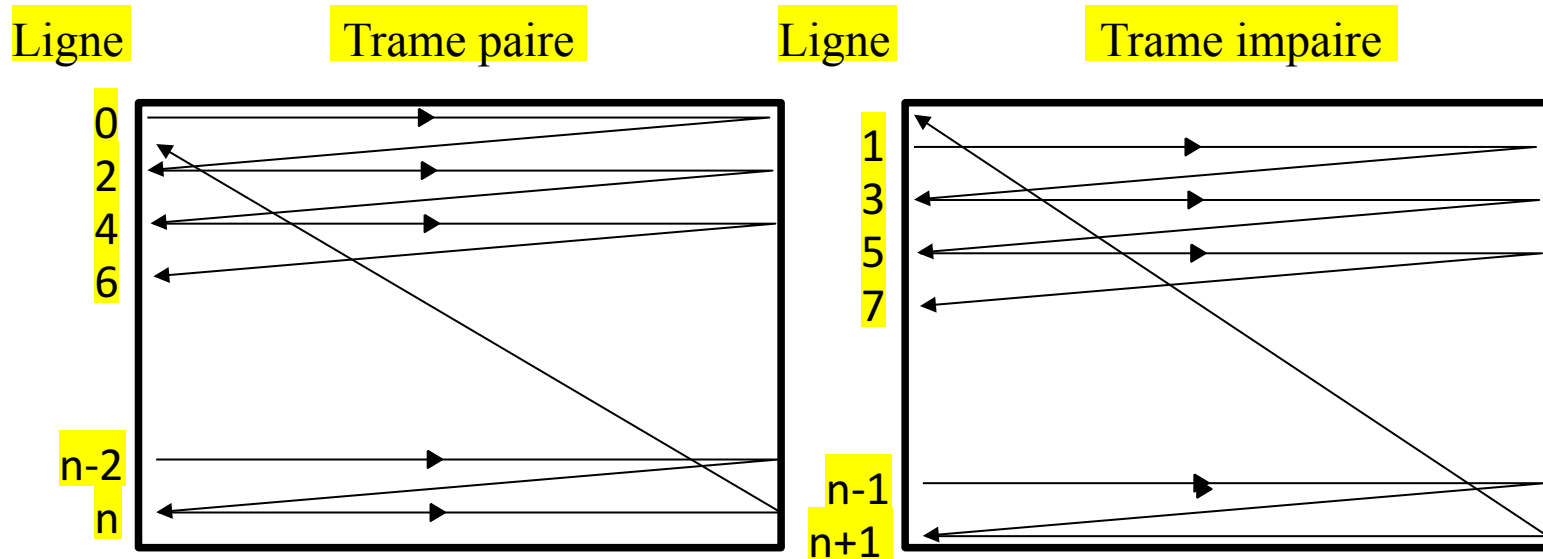
Technologies des caméras

Transfert analogique



Technologies des caméras

Transfert analogique

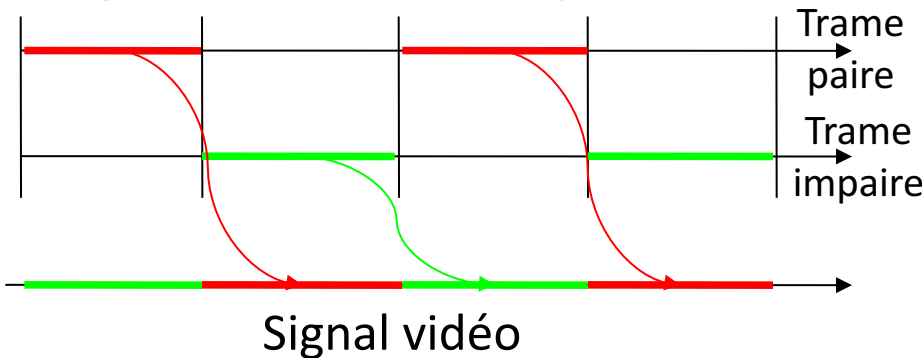


- Balayage de l'image de gauche à droite de haut en bas.
- Caméras progressives : toute l'image est balayée.
- Caméras entrelacées : les lignes paires puis les lignes impaires sont balayées alternativement.

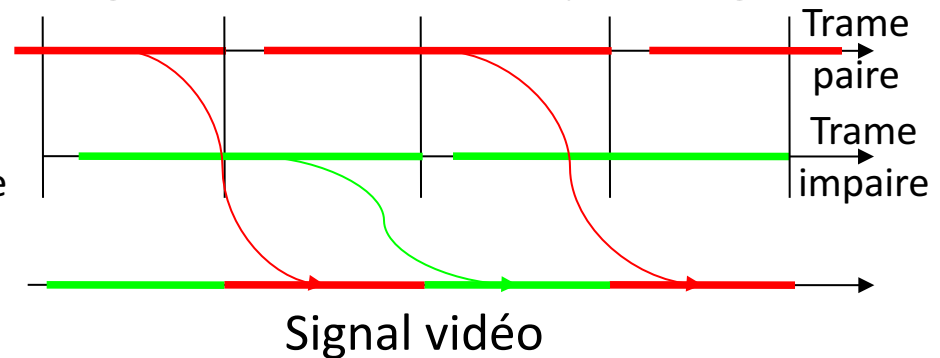
Technologies des caméras

Transfert analogique

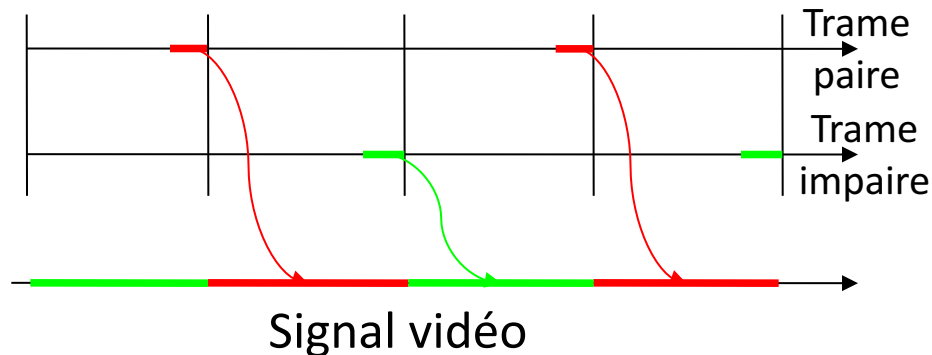
Intégration de la lumière par trame



Intégration de la lumière par image



Intégration de la lumière en utilisant un obturateur électronique



Technologies des caméras

Transfert analogique

Standards N&B :

- RS170 (USA, Japon) : 30 images entrelacées de 525 lignes par seconde.
- CCIR (Europe, Australie) : 25 images entrelacées de 625 lignes par seconde.

Standards couleur :

- PAL, SECAM, NTSC

Limites du transfert analogique :

- Sensibilité aux perturbations électromagnétiques.
- Débit limité à 50 Mega-pixels par seconde.

- Fréquence maximum en VGA N&B : $N_{max} = \frac{50 \cdot 10^6}{640 \times 480} \approx 162 \text{ fps}$

Technologies des caméras

Transfert numérique

Les niveaux de gris des pixels sont numérisés au niveau de la caméra puis sont transférés par une liaison numérique vers l'ordinateur.

Types de liaisons :

- Camera Link :
 - Liaison série
 - Jusqu'à 850 Mo/s
- GigE Vision :
 - Basé sur Gigabit Ethernet (100 Mo/s)
 - Protocol de configuration, de découverte automatique et de transfert de flux vidéo par UDP.
- Autres : USB3 (625 Mo/s), Firewire, Thunderbolt (5 Go/s), ...



Technologies des caméras

Transfert numérique

Exemple : Dalsa-Teledyne HC640

Active Resolution	640 x 480
Frame Rate	300 fps
Pixel Size	7.4 μm
Data Format	8-bits Raw Bayer
Exposure Control	Programmable, or via External Trigger
Dynamic Range	48 dB
Nominal Gain Range	0 dB to +12 dB
Output	Gigabit Ethernet, also supports 100 Mbps
I/O Ports	2 opto-isolated input, 2 opto-isolated output
Lens Mount	C or CS-mount / Right-Angle option is also available
Size	44 mm x 29 mm x 67 mm (including lens adapter)
Mass	~115 g (without lens)
Operating Temp	0°–45°C
Power Supply	12 to 24V
Power Dissipation	4W
Data Connector	Standard or Screw mount RJ-45
Power and I/O	Hirose 12-Pin
Software Platform	Teledyne DALSA Sopera LT
GigE Vision Compliant	Yes
Regulatory Compliance	FCC Class A, CE, RoHs 2002/95/EC
Part Number	CR-GC00-H640x



Calcul du débit maximum D_{max}

Dispositifs d'acquisition

Technologies d'interfaces

Interface dédiée : *frame grabber*

- Analogique : reconstruction de l'horloge pixel, conversion A/N, stockage local à double accès, générateur de zone d'intérêt.
- Numérique : interface de liaison dédiée (Camera Link), génération d'horloge image, stockage local à double accès, générateur de zone d'intérêt.
- Carte au format PCIe (500Mo/s par ligne).



Feature	Specification
Part Number	OR-Y4C0-XPX00
Camera Interface	1 Full, 80-bit (Deca), Medium or Base CameraLink
Camera Format	Base, Med., or Full Camera Link, PoCL
Pixel Clock	up to 85 MHz
Bits/Pixel	up to 16
Number of Camera Taps	up to 10 taps/8-bit
Host Bus	PCIe Gen2 x4
Frame Buffer	512 MB
Advanced Features	ILUTS, Flat Field, Flat Line, Dead Pixel, Bayer, PoCL
OS Support	Windows 7 and 8 (32/64-bit) compatible
Software	Sapera Vision Software
GPIO	On-board 4-in/4-out

Dispositifs d'acquisition

Technologies d'interfaces

Interfaces génériques :

- USB3, GigE, Thunderbolt, Firewire, ...
- Pas de fonctions matérielles dédiées à la vision.
- Fonctions logicielles :
 - Libres : par ex. "OpenCV".
 - Propriétaires : par ex. "Sapera Nitrous" de Dalsa Teledyne.

Key Features

- Seamless integration with Sapera Essential with over 400 image processing primitives and application tools
- GPU: based on Nvidia's CUDA Rev 2.3 or higher
- MCO: based on Intel's latest instruction set technology and supports for multi-core CPUs
- Includes demos and source code examples to delineate optimum coding practices
- Supports for line and area scan, color and monochrome cameras with – GigE Vision®, Camera Link®, LVDS and Analog interfaces
- Support 32/64-bit platforms for Windows® 7, Windows XP and Windows Vista

Dispositifs d'acquisition

Accès à l'image numérisée

Interface dédiées :

- Transfert DMA de toute l'image ou d'une portion programmable de l'image.
- Accès en lecture à la mémoire pendant l'acquisition.
- Fonctions avancées de synchronisation (interruption en fin d'acquisition d'image par exemple).

Interfaces génériques :

- La programmation de la zone d'intérêt se fait directement dans la caméra.
- Certaines caméras embarquent des fonctions de prétraitement.
- L'accès à l'image se fait par pointeur sur un double ou un triple *buffer*.

Dispositifs d'acquisition

Modèle dynamique

Le modèle dynamique du processus de transfert et de traitement de l'image est un retard D .

Ce retard dépend de l'architecture choisie :

- Série
- Parallèle
- Pipeline
- À la volée

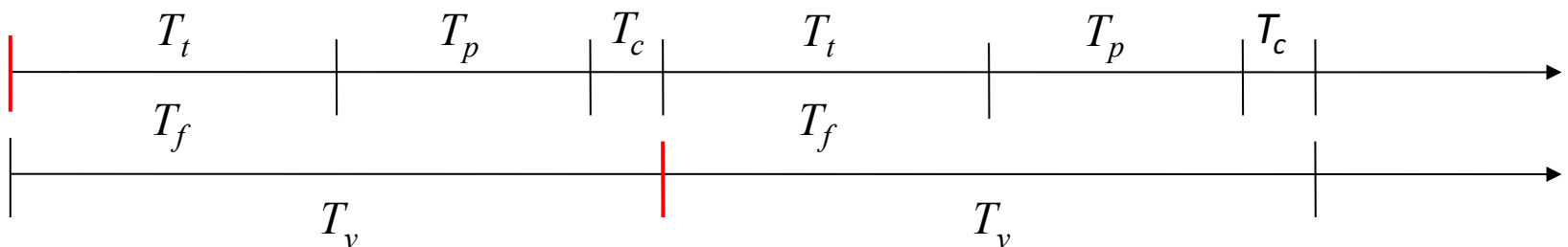
On fait l'hypothèse que le temps de transfert est similaire à la période d'une image.

Dispositifs d'acquisition

Modèle dynamique : architecture série

Soient :

- T_t : temps de transferts de l'image de la caméra vers l'ordinateur.
- T_p : durée du traitement d'image.
- T_c : durée du calcul de la commande appliquée au robot.
- T_f : période de l'acquisition d'image.
- T_v : période de l'asservissement visuel.

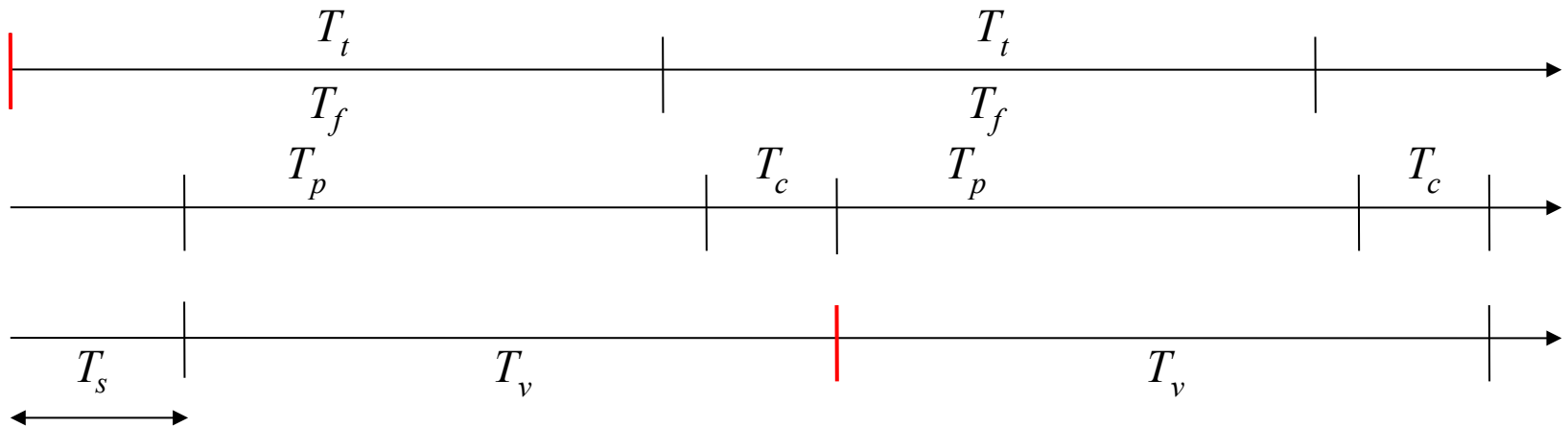


On applique la commande à la période suivante d'où $D = T_t + T_p + T_c$. A chaque période T_v on « saute » une ou plusieurs images. La fréquence de l'asservissement visuel = fréquence de la caméra / (nb images sautées + 1)

Dispositifs d'acquisition

Modèle dynamique : architecture parallèle

Le traitement commence pendant le transfert :



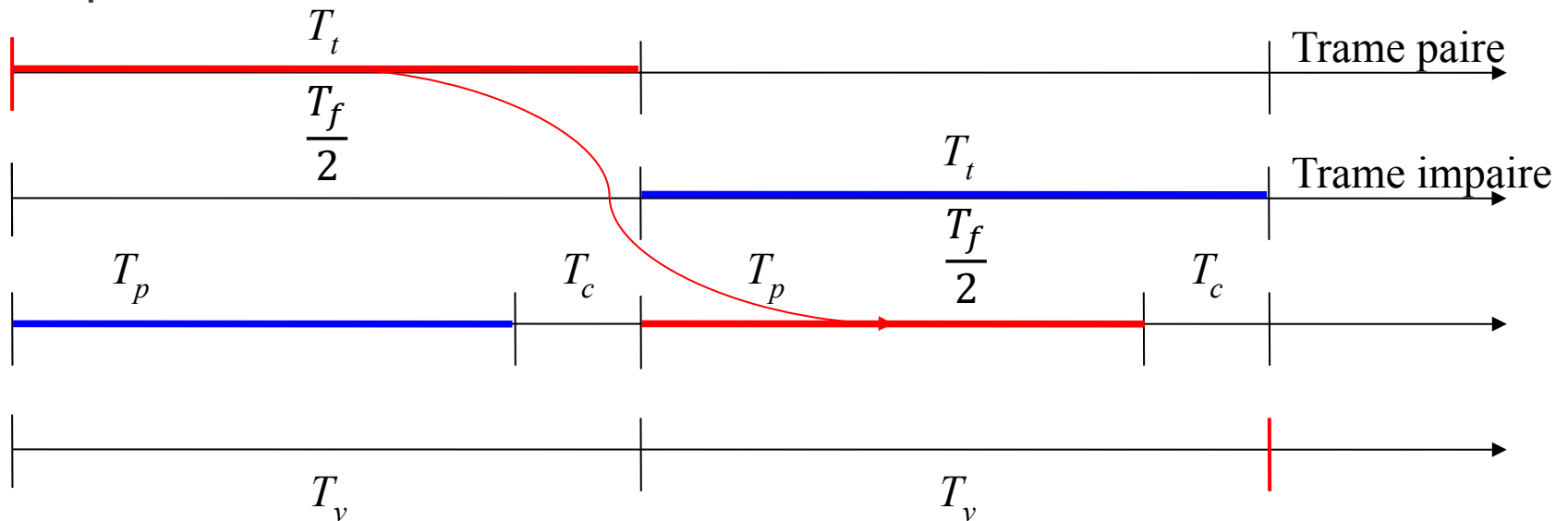
$$D = T_s + T_p + T_c < T_t + T_p + T_c$$

La fréquence de l'asservissement visuel peut être la même que celle de la caméra.

Dispositifs d'acquisition

Modèle dynamique : architecture parallèle

Exploitation de l'entrelacement :



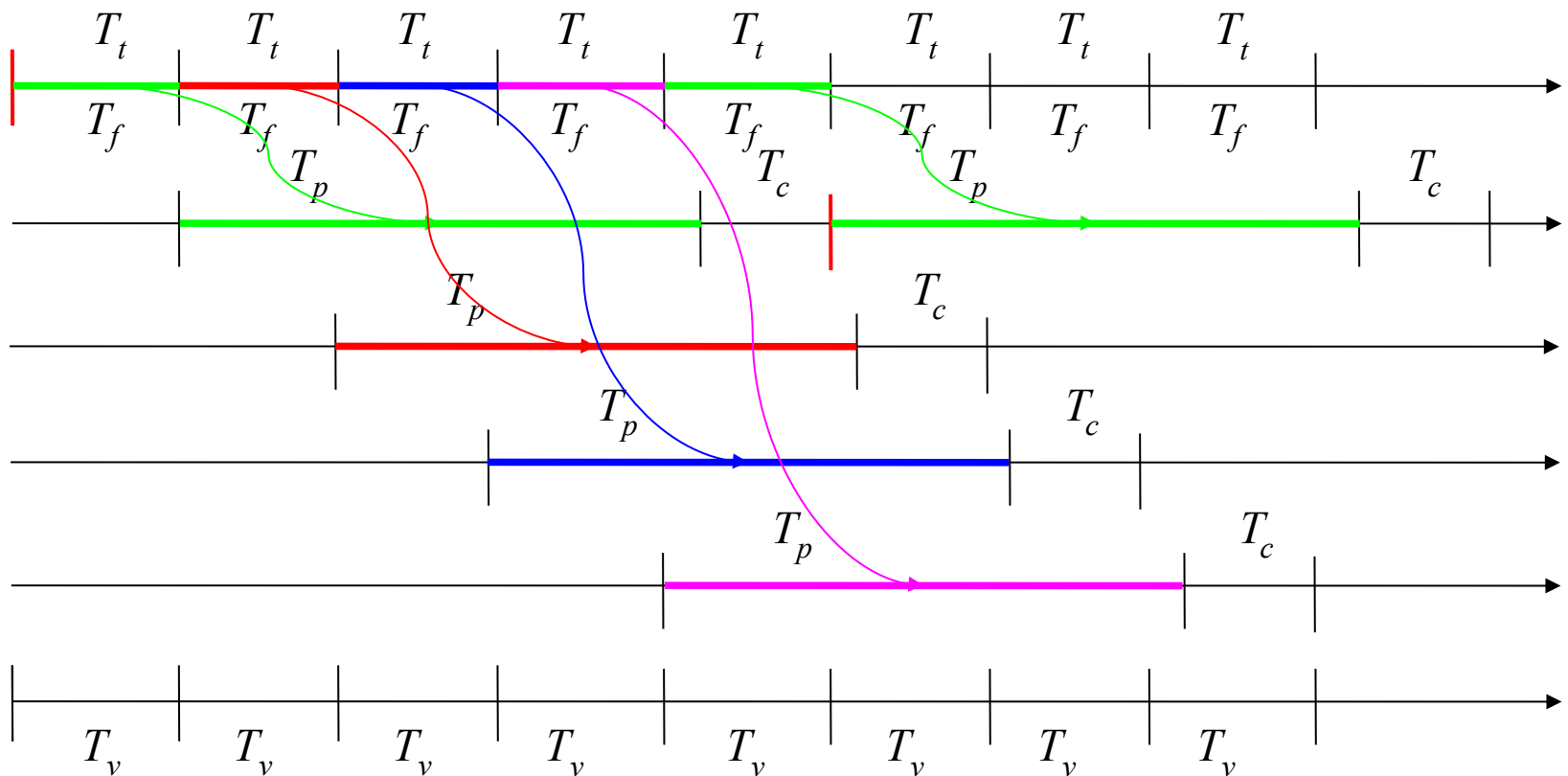
$$D = T_f/2 + T_p + T_c = T_f$$

La fréquence de l'asservissement visuel est le double de celle de la caméra.

Dispositifs d'acquisition

Modèle dynamique : architecture pipeline

Soit N le nombre de processeurs :



Dispositifs d'acquisition

Modèle dynamique : architecture pipeline

$$D = T_t + NT_f$$

La fréquence de l'asservissement visuel est égale à celle de la caméra.

Si l'algorithme de traitement d'image est parallélisable, il vaut mieux paralléliser le traitement sur le N processeurs.

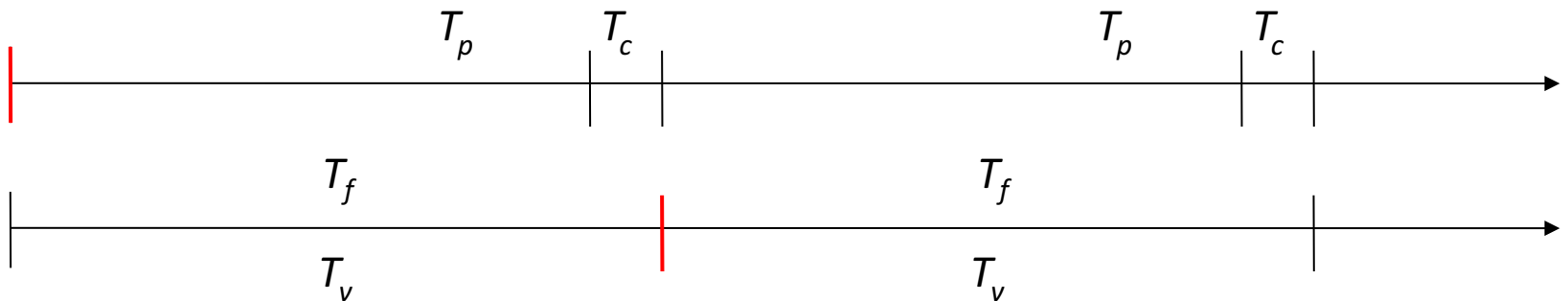
Dans ce cas, $D = T_t + T_f$ sans changement de la fréquence de l'asservissement visuel.

Dispositifs d'acquisition

Modèle dynamique : architecture à la volée

Le temps de transfert est négligeable :

- *Vision chip*.
- Liaison numérique très rapide et fréquence d'acquisition lente.



$$D = T_p + T_c$$

La fréquence de l'asservissement visuel est la même que celle de la caméra.

Critères de qualité de l'image

Qualité spatiale

Les aberrations diminuent la netteté de l'image en introduisant un flou :

- Les aberrations géométriques : variations de netteté dans l'image dues aux défauts géométriques de la lentille.
- Les aberrations chromatiques : sont dues au fait que des rayons de différentes longueur d'onde ne sont pas projetés au même endroit.
- Les aberrations dues à la diffraction : la diffraction au niveau du diaphragme fait qu'une source lumineuse ponctuelle se projette en une tache. Cet effet est d'autant plus prononcé que l'ouverture est faible.

La profondeur de champ : intervalle de distance le long de l'axe optique dans lequel les objets sont nets. La profondeur de champ augmente lorsque l'ouverture diminue.

Critères de qualité de l'image

Qualité spatiale

Les distorsions géométriques déforment l'image tout en conservant sa netteté :

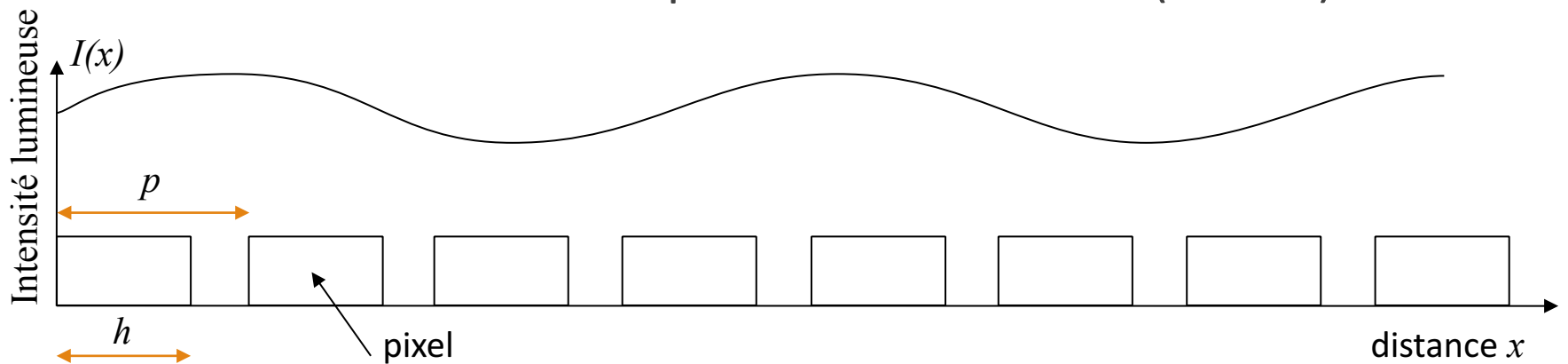
- *La distorsion radiale* contracte ou dilate l'image suivant des demi-droites qui ont pour origine le point principal de l'image.
- *La distorsion tangentielle* déforme l'image suivant une direction perpendiculaire aux demi-droites issues du point principal. Il en résulte une torsion de l'image. Généralement la distorsion tangentielle est nettement plus faible que la distorsion radiale.

La fonction de transfert de modulation (MTF) : donne la réponse fréquentielle du capteur à un motif sinusoïdal d'intensité lumineuse.

Critères de qualité de l'image

Qualité spatiale : MTF

Soit un éclairage de répartition sinusoïdale (cas 1D) :



- On a : $I(x) = A \sin(\omega x) + B \mid B > A > 0$
- Au point d'abscisse x , la quantité de charge dq générée par une lumière d'intensité $I(x)$ sur une longueur dx au bout du temps d'exposition T_s est donnée par : $dq = KT_s I(x)$
- K est un coefficient modélisant la sensibilité du capteur.

Critères de qualité de l'image

Qualité spatiale : MTF

La quantité de charge $Q[k]$ accumulée dans le pixel n° k est donnée par :

$$Q[k] = \int_{kp}^{kp+h} KT_s I(x) dx = KT_s \int_{kp}^{kp+h} A \sin(\omega x) + B dx$$

En normalisant par h , on obtient :

Critères de qualité de l'image

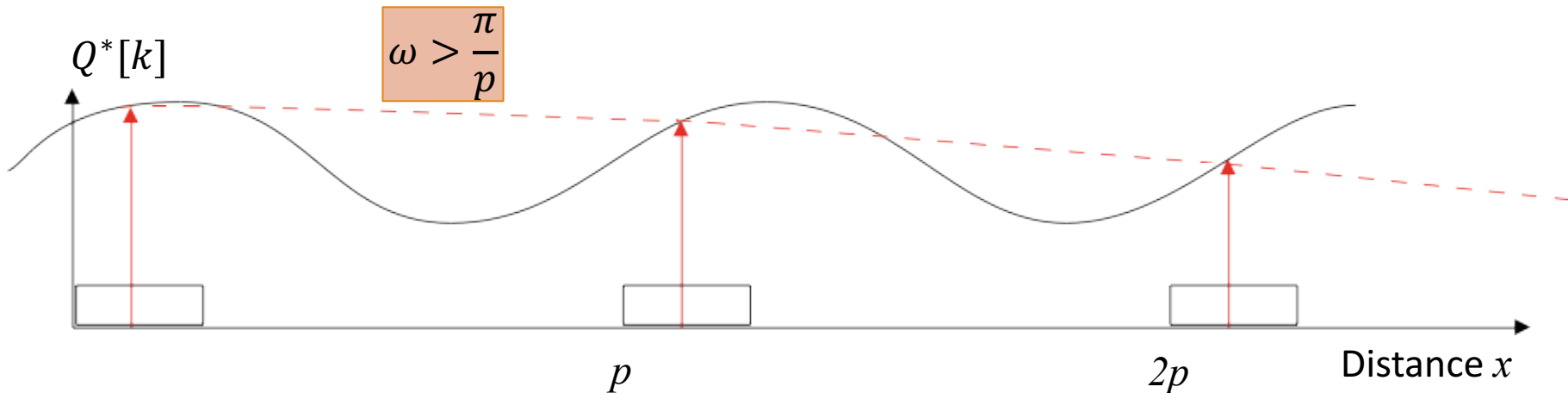
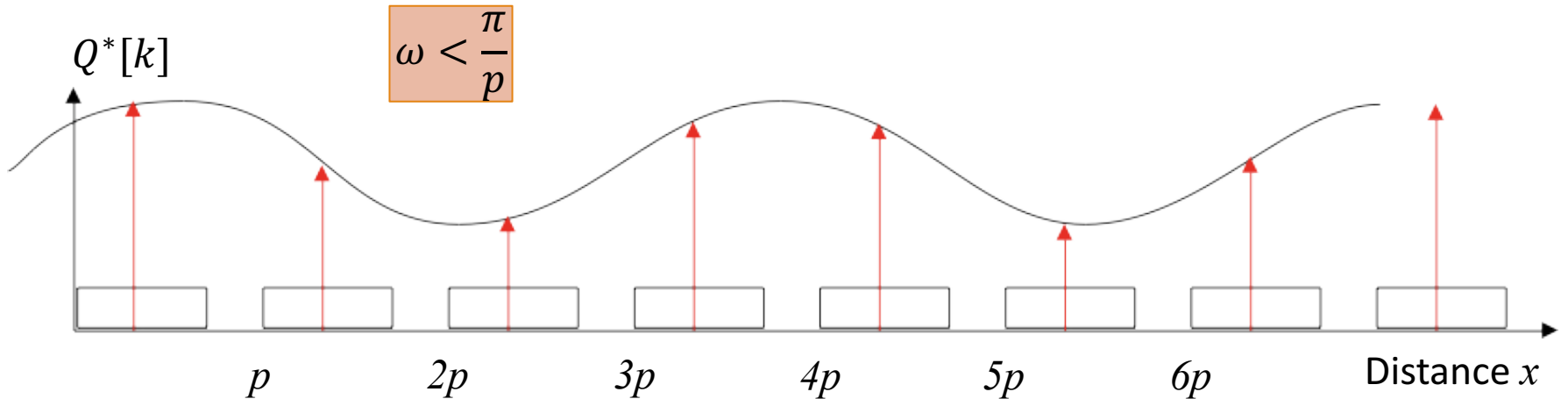
Qualité spatiale : MTF

Effets de l'échantillonnage spatial :

- Le fondamental du motif sinusoïdal échantillonné subit un décalage vers les x décroissants d'une distance $\frac{h}{2}$. Ce décalage ne pose aucun problème car il est le même quelque soit la pulsation.
- L'amplitude du fondamental du motif sinusoïdal échantillonné varie suivant une fonction sinus cardinal de la pulsation :
 - si $h \rightarrow 0$, pas d'atténuation : échantillonnage idéal.
 - Si $\omega > \frac{\pi}{p}$ (fréquence de Nyquist), d'après le théorème de Shannon, il y a repliement spectral dans la bande de base et donc apparition de nouvelles composantes fréquentielles basse fréquence. Ceci provoque l'apparition de franges de Moiré dans l'image.

Critères de qualité de l'image

Qualité spatiale : MTF

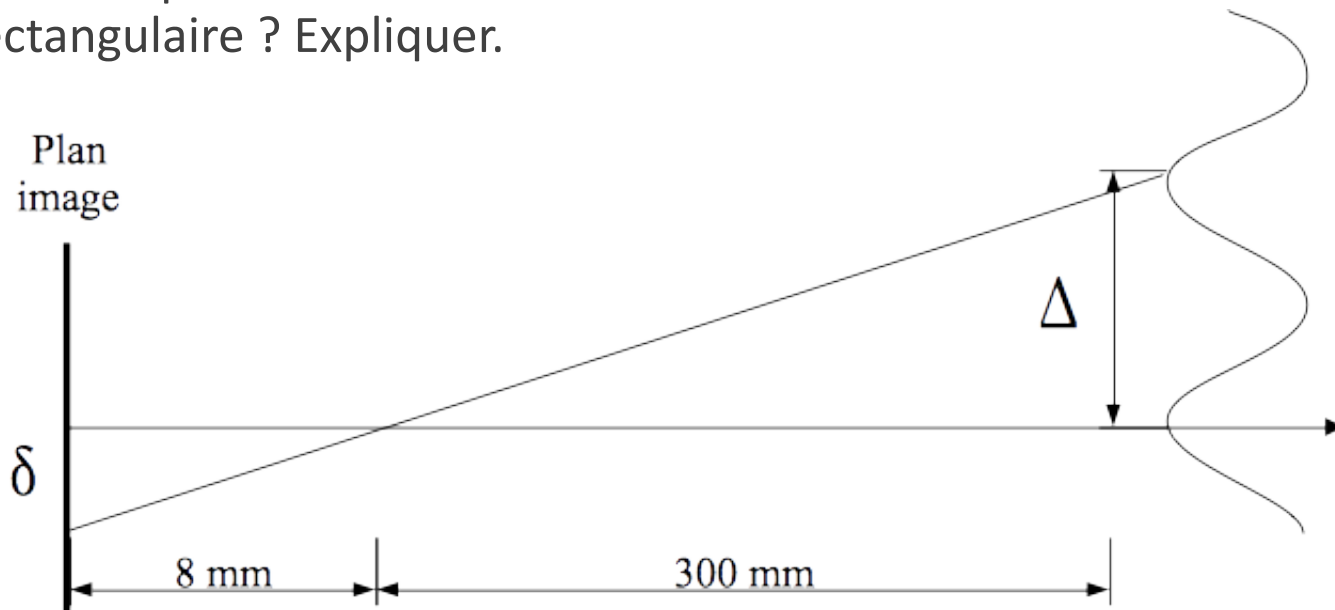


Critères de qualité de l'image

Qualité spatiale : MTF

Exercice :

- Soit une caméra CCD munie d'un objectif de 8 mm . Le pixel est un carré de 8 microns de côté. Quelle est la période spatiale minimale d'un motif sinusoidal placé à 30 cm de la caméra qui peut être mesurée par la caméra ? En est-il de même avec un motif rectangulaire ? Expliquer.



Critères de qualité de l'image

Qualité spatiale : MTF

- Il n'en est pas de même avec un motif rectangulaire. En effet, un signal rectangulaire peut être décomposé en série de Fourier. Il en résulte un spectre qui s'étend jusqu'à l'infini. Donc, il existe des composantes fréquentielles harmoniques qui sont au-delà de la fréquence de Nyquist et qui se replient dans le bande de base.
- Un moyen de pallier à ce problème est de défocaliser la lentille. Il en résulte un filtrage analogique passe-bas anti-repliement de l'image qui élimine les hautes fréquences (en rendant l'image floue).

Critères de qualité de l'image

Qualité temporelle : bruit

Différentes sources de bruit s'additionnent au cours de la chaîne d'acquisition :

- Bruit dû à l'arrivée aléatoire des photons
- Bruit de conversion des photons en charges
- Bruit de génération thermique de charges
- Bruit de l'amplificateur qui converti les charges en tension
- Bruit de quantification

Il en résulte, pour une scène fixe, un bruit blanc gaussien sur la valeur au cours du temps du niveaux de gris des pixels.

Critères de qualité de l'image

Qualité temporelle : bruit

Le bruit sur les pixels se répercute sur la mesure de l'asservissement visuel, il est donc important de le réduire au maximum.

Pour maximiser le rapport signal/bruit il faut maximiser la quantité de charges produites en :

- Augmentant la l'éclairage de la scène (compromis avec la saturation du capteur et l'effet de blooming).
- Augmentant l'ouverture (compromis avec la profondeur de champ).
- Augmentant le temps d'exposition (compromis avec le bruit de génération thermique de charges).

Critères de qualité de l'image

Qualité temporelle : échantillonnage

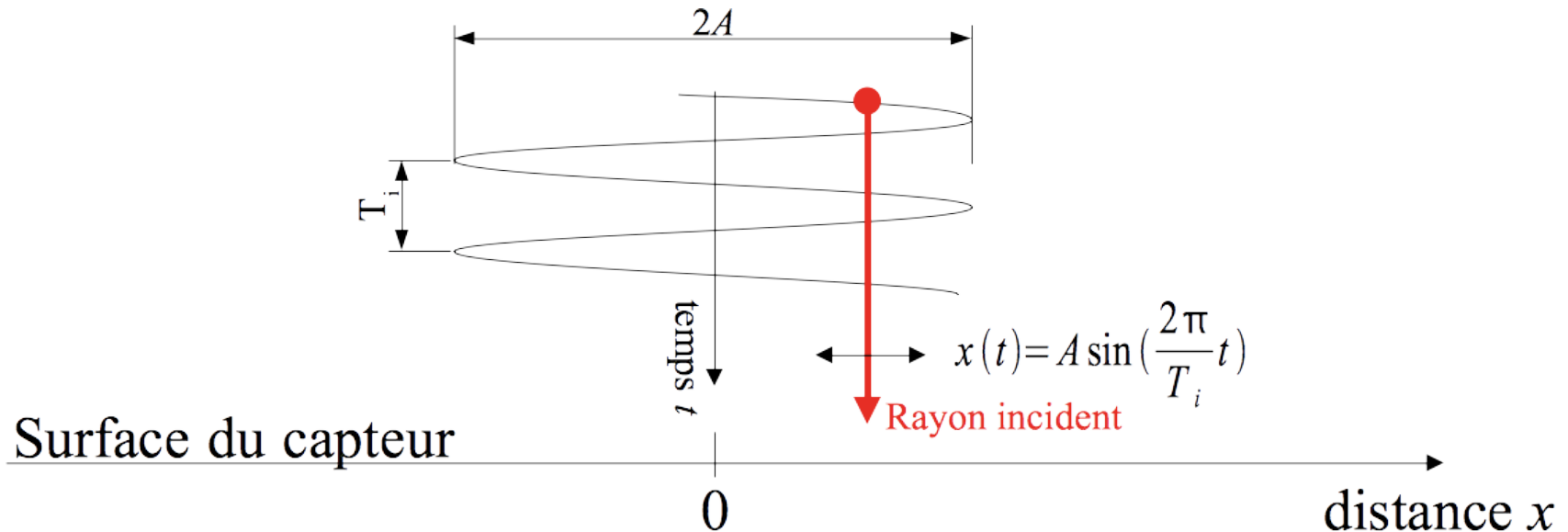
Effet de l'échantillonnage temporel :

- Tout comme l'échantillonnage spatial, l'échantillonnage temporel introduit une déformation du signal initial.
- Intuitivement, on peut aisément comprendre que si la scène évolue durant la période d'accumulation des charges, il en résultera un flou dans l'image.

Dans cette partie, nous allons établir la fonction de transfert approchée du processus d'échantillonnage de l'image. Pour cela nous considérons une source lumineuse ponctuelle se déplaçant le long d'une droite parallèle au plan optique. La position de cette source sur cette droite est une fonction sinusoïdale du temps. Pour simplifier, nous négligeons l'effet de l'échantillonnage spatial.

Critères de qualité de l'image

Qualité temporelle : échantillonnage



Soient :

- T_s le temps d'intégration de la lumière
- $\omega_i = 2\pi/T_i$ la pulsation du signal incident
- T la période l'une image

Critères de qualité de l'image

Qualité temporelle : échantillonnage

- La position $X[k]$ de la primitive extraite de l'image k est donnée par :

$$\sin(a - b) + \sin(a + b) = 2 \sin a \cos b$$

- Lorsque $T_s \rightarrow 0$ l'amplitude tend vers A .
- Le signal échantillonné est retardé de $T_s/2$
- Si $T_s = T$, le signal échantillonné est retardé de la moitié d'une période image. De plus, dans ce cas, le signal échantillonné est nul pour $\omega_i = \pi/T$, la pulsation de Nyquist.

Critères de qualité de l'image

Qualité temporelle : échantillonnage

Fonction de transfert de l'échantillonnage lorsque $T_s = T$:

- Soit le filtre moyennneur :
- Equation temporelle :
- En remplaçant l'expression de x :