

Perception Avancée et Robotique Mobile

UPJV, Département EEA

Master 2 3EA, EC31

Parcours RoVA

Fabio MORBIDI

Laboratoire MIS

Équipe Perception Robotique

E-mail: fabio.morbidi@u-picardie.fr



Electronique

Energie Electrique

Automatique

CM, TD: Mardi et Mercredi 13h30-16h00, salle CURI 305

TP: Jeudi 9h00-12h00, salle TP204

AU 2024-2025



Ch. 1: Perception pour la robotique

- Introduction

Partie 1

- Classification des capteurs

Partie 2

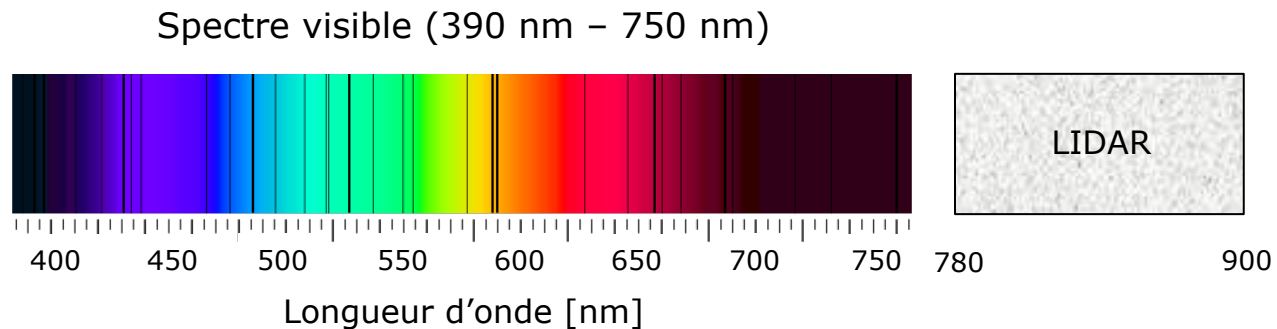
- Typologies de capteur

Partie 3

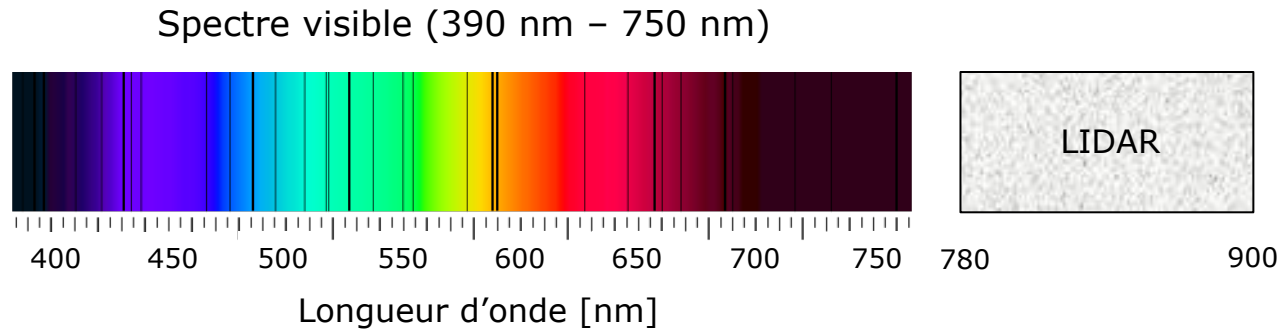
Capteurs Extéroceptifs

Télémètres laser

- LIDAR (« Light Detection And Ranging »)
En réalité, portemanteau de *light* et *radar* !
 - Technologie pour créer des cartes à haute résolution
 - Mesure de distance en éclairant une cible avec un faisceau de lumière laser
- Rayons émis/reçus coaxiaux
- Longueur d'onde: typiquement 780-900 nm



Télémètres laser



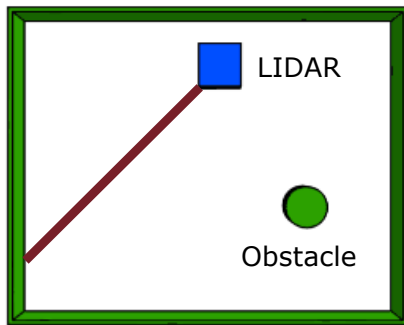
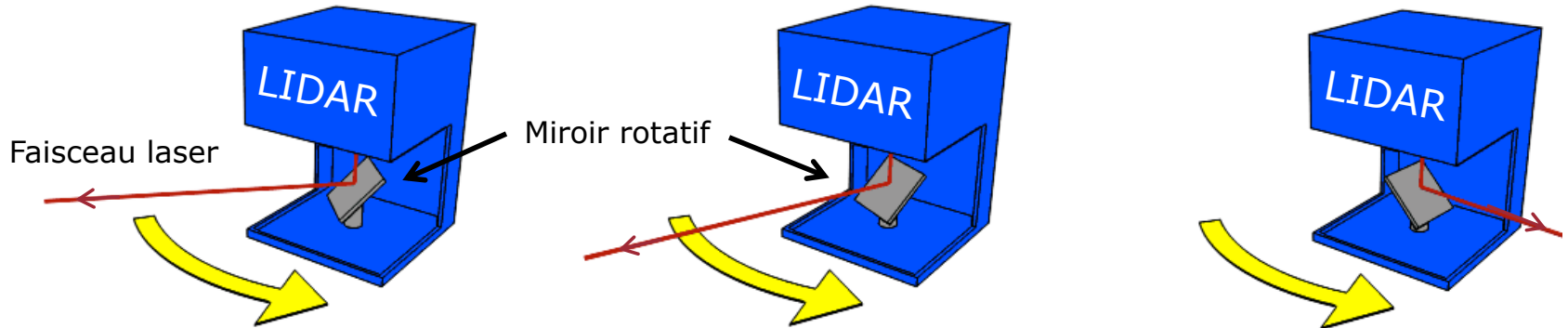
- Bandes spectrales des radiations infrarouges, IR (découpage ISO 20473:2007 pour l'optique et la photonique)



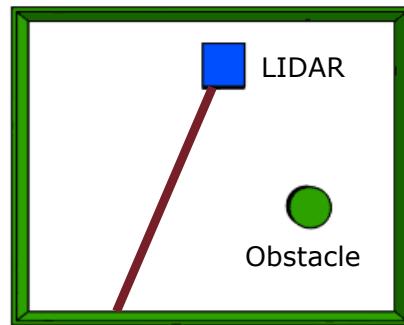
Désignation	Abréviation	Longueur d'onde
Infrarouge proche	PIR (IR-A, NIR)	780 nm – 3 μ m
Infrarouge moyen	MIR (IR-B)	3 μ m – 50 μ m
Infrarouge lointain	LIR (IR-C, FIR)	50 μ m – 5 mm

www.iso.org/standard/39482.html

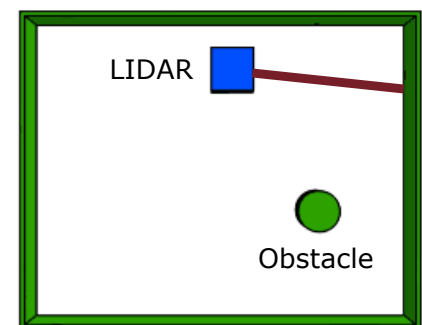
Télémètre laser : principe de fonctionnement



● LIDAR

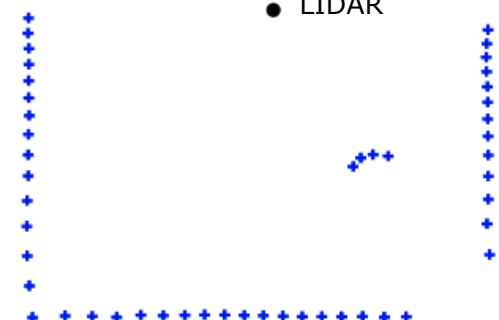


● LIDAR



● LIDAR

Mesures de distance
à des intervalles
angulaires réguliers
(points bleus)



Télémètres laser : méthodes de mesure

1. *Laser pulsé* (ex. SICK)

- Mesure directe du temps de vol
- Besoin d'une mesure à la picoseconde (10^{-12} s)
(En fait, 3 m correspondent à 10 ns)

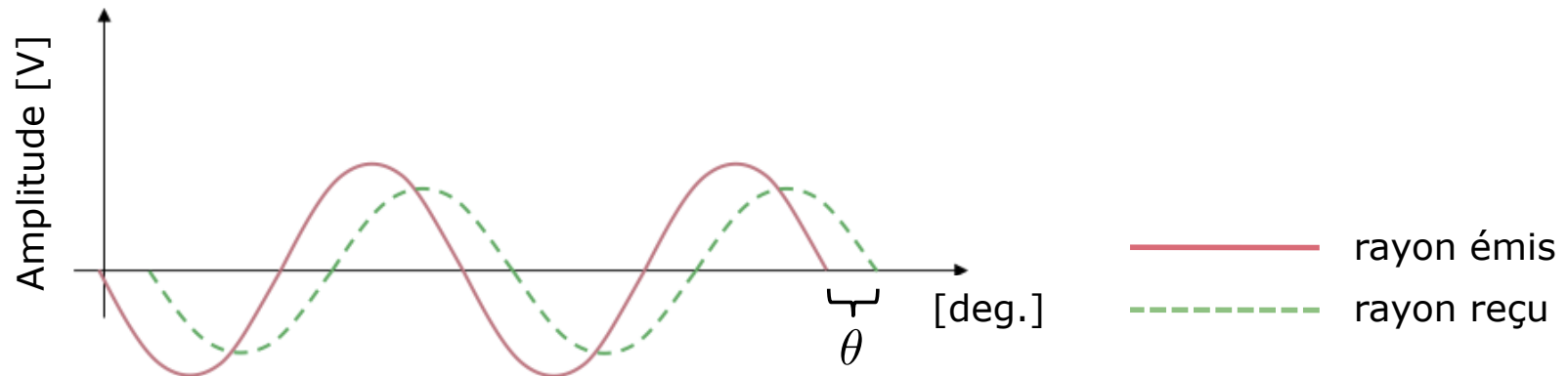
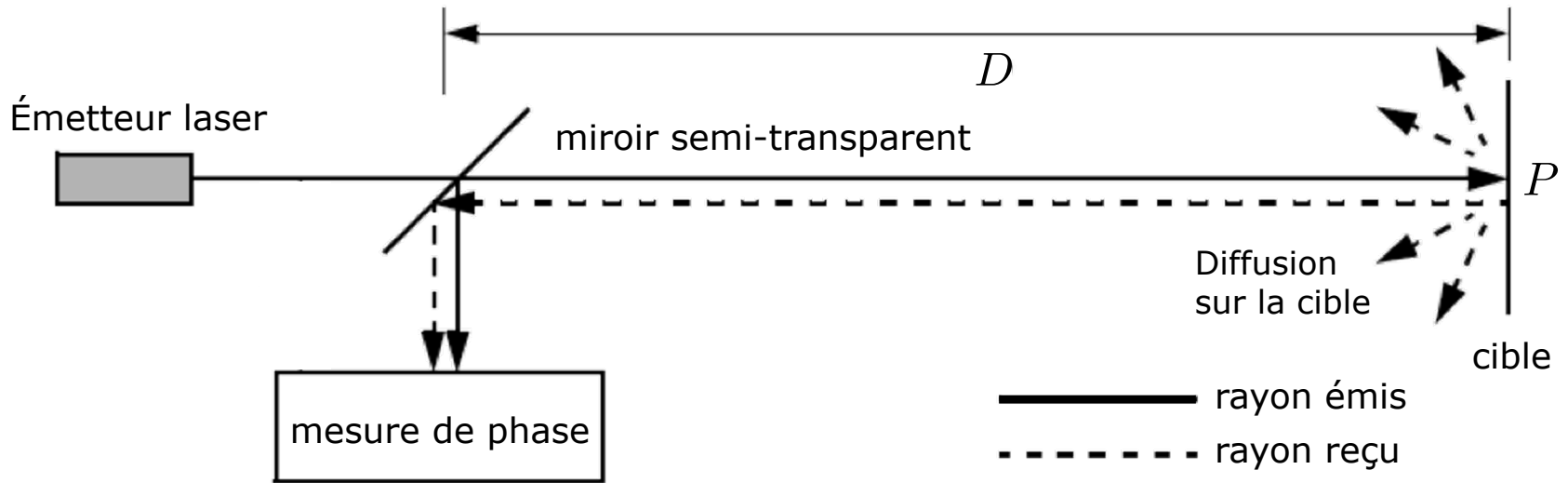


2. *Mesure de déphasage* (ex. Hokuyo)

- Émission d'onde modulée à fréquence connue
- Mesure de déphasage entre les signaux émis et reçus
- Techniquement « plus simple »

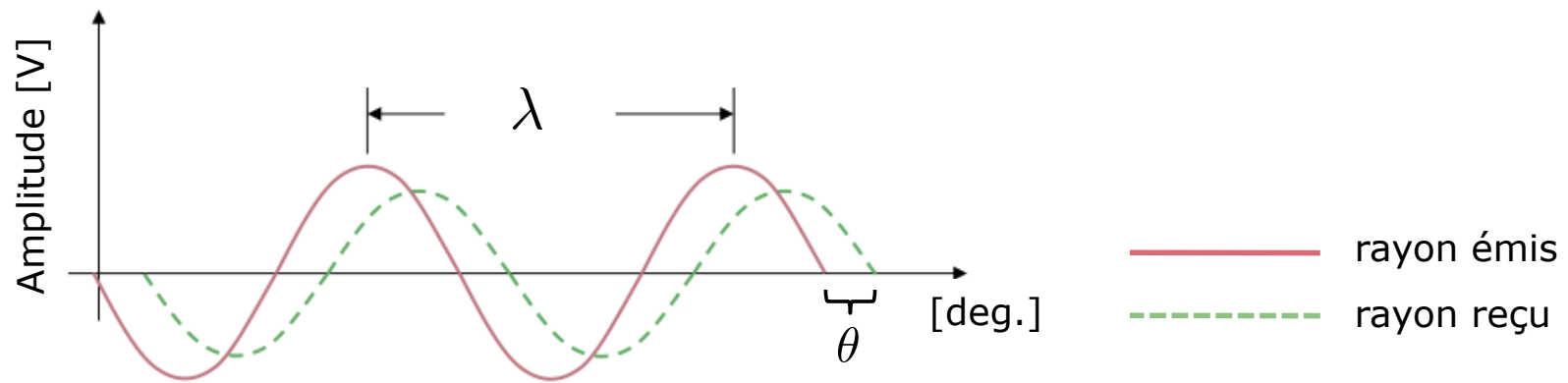


Mesure de déphasage



Problème: jamais pure réflexion. On a toujours de la diffusion sur la cible

Mesure de déphasage



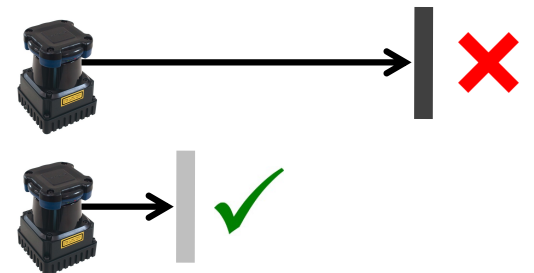
Distance totale D' parcourue par la lumière émise (dans un milieu uniforme):

$$D' = 2D = \lambda \frac{\theta}{2\pi}, \quad \lambda = \frac{c}{f}$$

- θ : différence de phase
- c : vitesse de la lumière (3×10^8 m/s)
- f : fréquence de modulation
- λ : longueur d'onde

La confiance dans la distance mesurée est proportionnelle à l'amplitude du signal reçu, donc:

- Objets *sombres* et/ou *lointains*: estimation « imprécise »
- Objets *proches* et/ou *clairs*: estimation « précise »



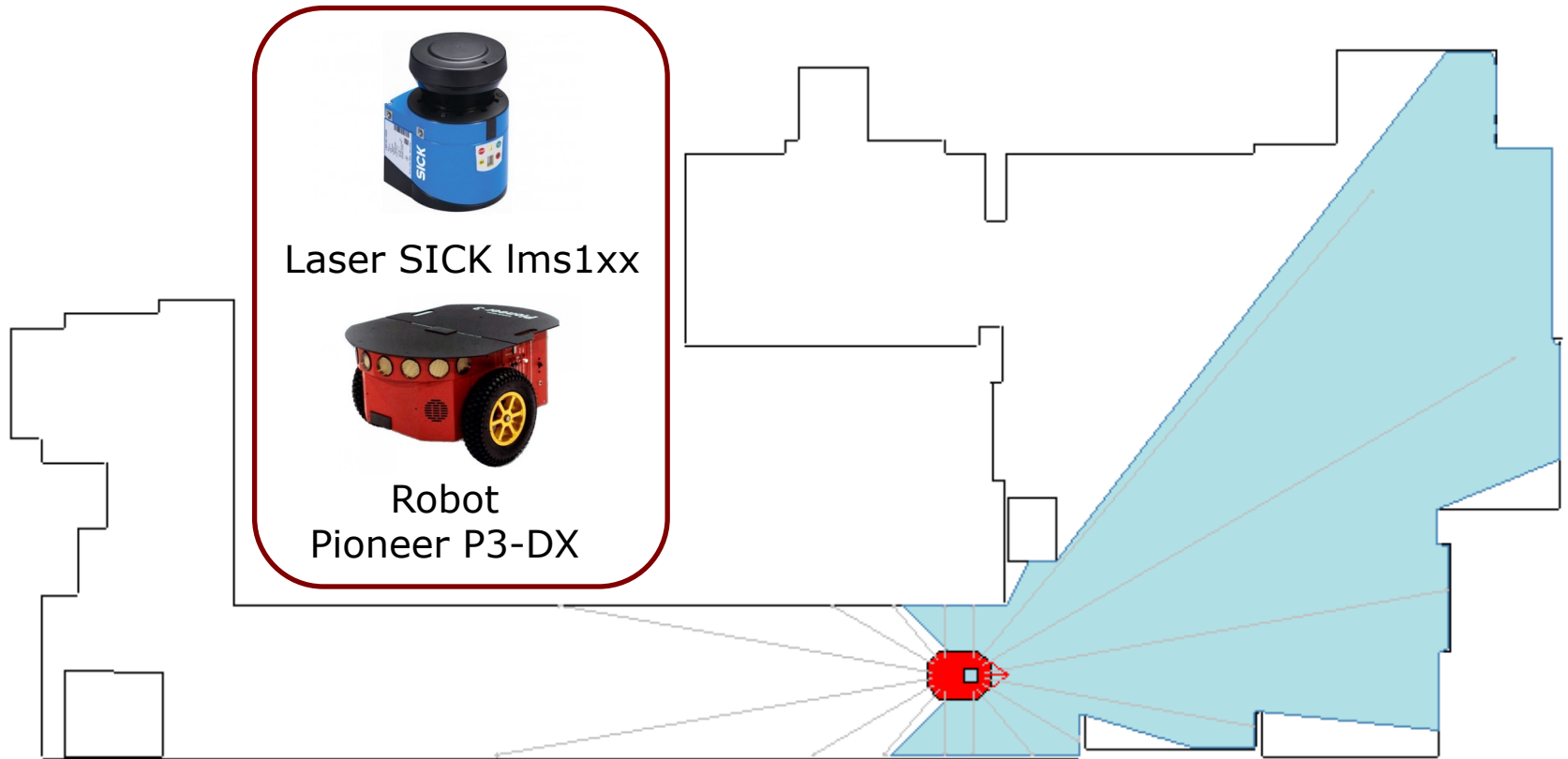
Télémètres laser 2D : résultats de mesure



Laser SICK Ims1xx

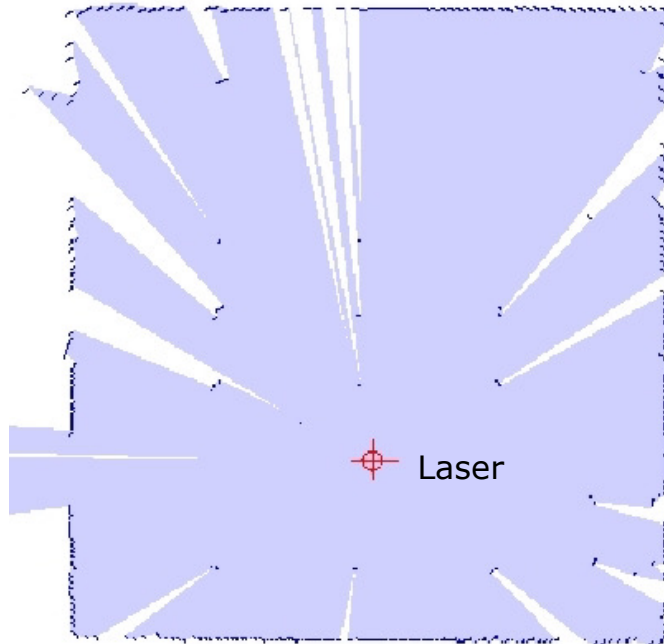


Robot
Pioneer P3-DX

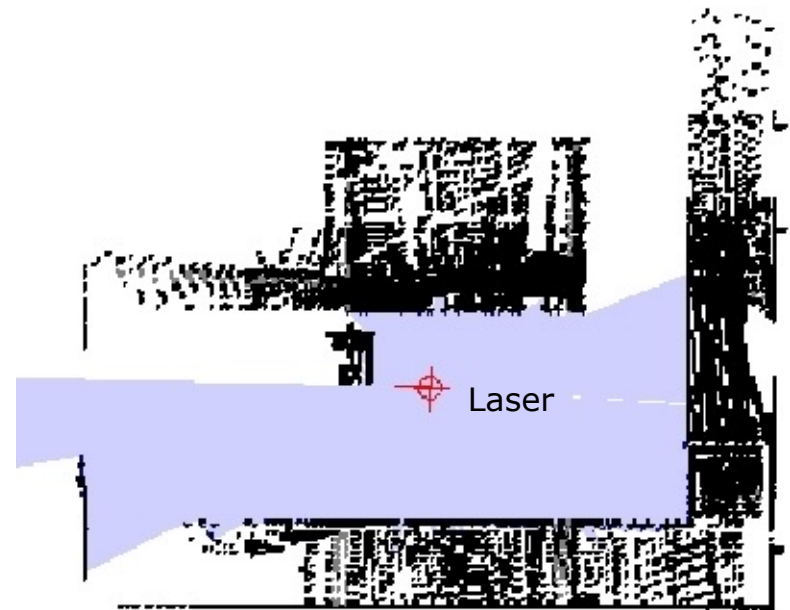


Simulateur MobileSim d'Omron Adept Technology

Télémètres laser 2D : résultats de mesure

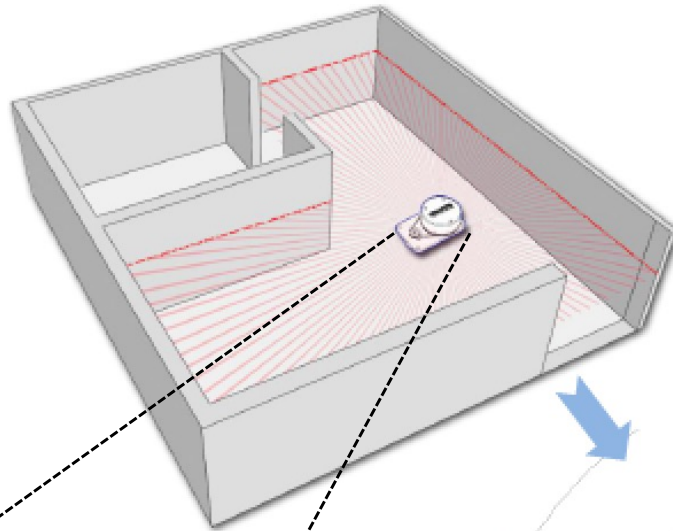


2 lasers SICK combinés:
champ de vue $180^\circ + 180^\circ$
(une salle)



4 lasers Hokuyo combinés
(un couloir)

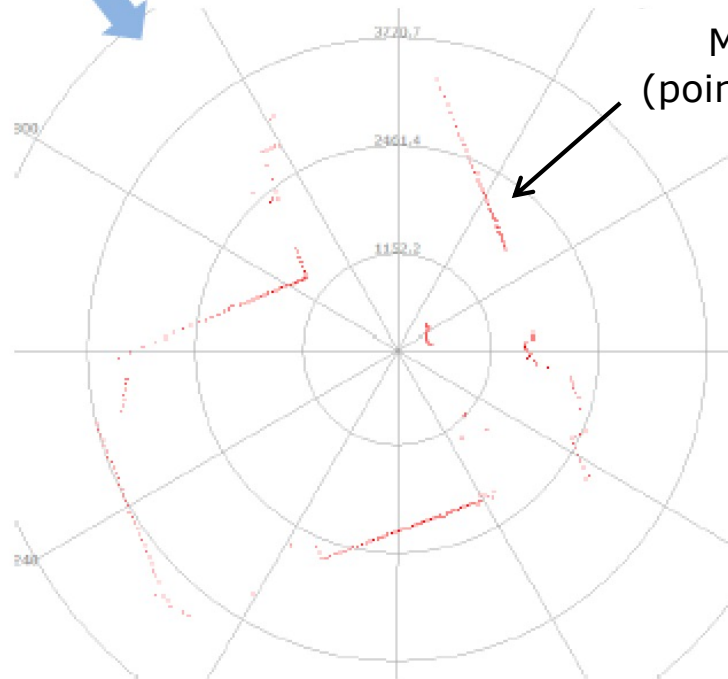
Télémètres laser 2D : résultats de mesure



Emetteur IR Récepteur



RPLIDAR de RoboPeak

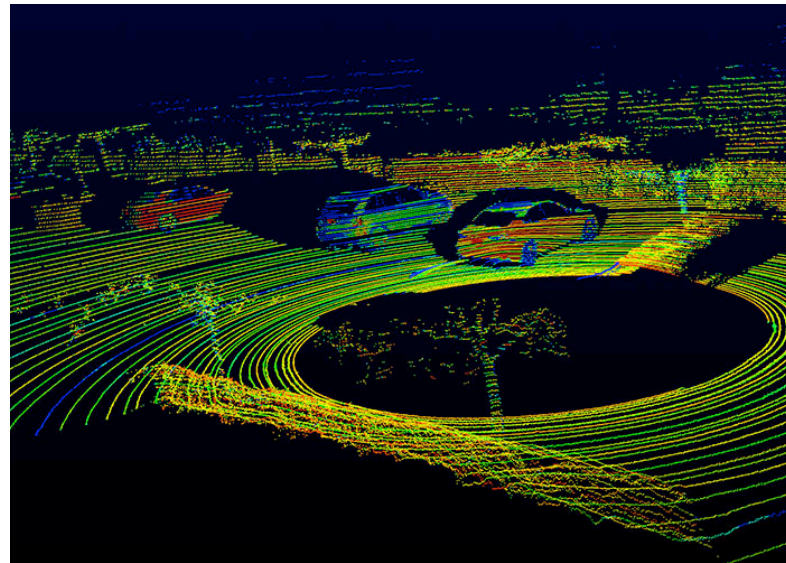


Mesures
(points rouges)

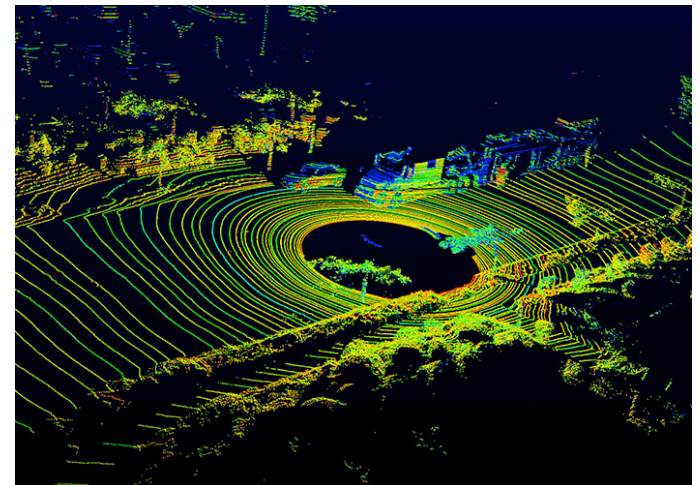
Télémètres laser 3D : résultats de mesure



360°, fréquence de rotation jusqu'à 15 Hz



Données brutes
du Velodyne
HDL-64E



Vidéo Velodyne

Télémètres laser commerciaux

Capteurs 2D :

- RPLIDAR de *RoboPeak* (Chine)
 - 1 laser \times 360°, portée max 6 m
 - 1.15 W, 170 g, 5.5 Hz
 - Résolution angulaire: $\leq 1^\circ$
 - Prix: 400 €
- URG-04LX de *Hokuyo* (Japon)
 - 1 laser \times 240°, portée max 5.6 m
 - 2.5 W, 160 g, 10 Hz
 - Résolution angulaire: 0.352°
 - Prix: 3500 €
- LMS291 de *SICK* (Allemagne)
 - 1 laser \times 90°, portée max 80 m
 - 30 W, 4.5 kg, 75 Hz
 - Résolution angulaire: 0.5°
 - Prix: environ 5000 €

RPLIDAR



RPLIDAR A2



URG-04LX



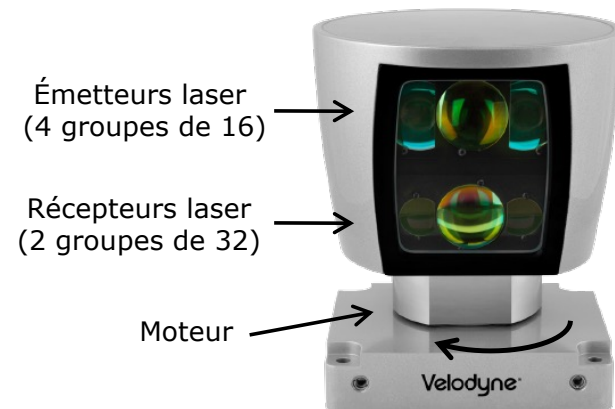
LMS291



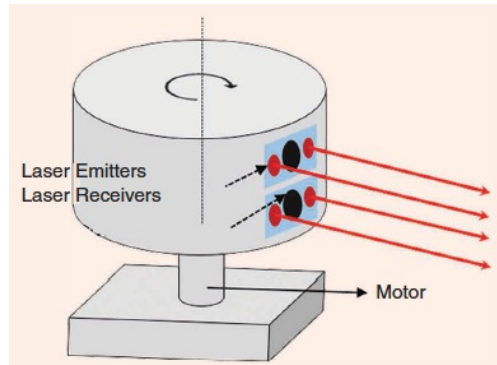
Télémètres laser commerciaux

Capteurs 3D :

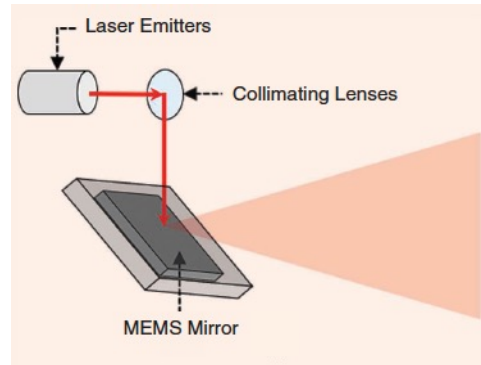
- LUX HD de *IBEO*, app. automobiles (Allemagne)
 - 4 lasers \times 110° (3.2° vertical), portée max 30 m
 - 10 W max, 1 kg, 25 Hz
 - Résolution angulaire H/V: $0.25^\circ/0.8^\circ$
 - Prix: environ 22 k€
- HDL-64E de *Velodyne* (USA)
 - 64 canaux \times 360° (26.9° vertical), portée max 120 m
 - 60 W, 12.7 kg, 5-20 Hz, max 2.2 millions de points/s
 - Résolution angulaire H/V: 0.08° - $0.35^\circ/0.4^\circ$
 - Prix: environ 75 k\$
- OS2-128 de *Ouster* (USA) [« solid-state »]
 - 128 canaux \times 360° (22.5° vertical), portée max 240 m
 - 14-20 W, 1.1 kg, 10 ou 20 Hz, max 2.62 millions de points/s
 - Résolution angulaire verticale: 0.18°
 - Prix: environ 24 k\$



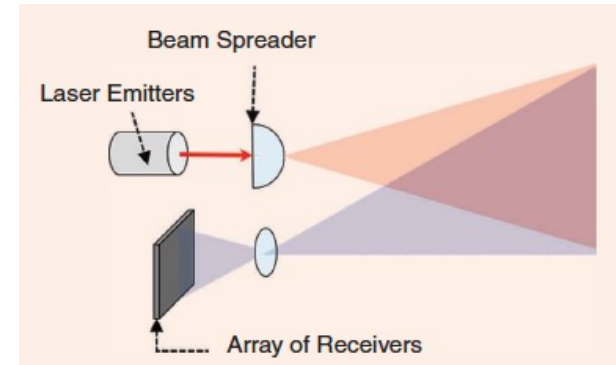
Télémètres laser : mécaniques vs « solid-state »



LIDAR mécanique



LIDAR « quasi-solid-state »
Utilisation de miroirs MEMS



LIDAR « solid-state »
Deux technologies: *optical phased array (OPA)* et *flash LIDAR*. Utilisation de modulateurs de phase et de diffuseurs optiques, respectivement

“Suitability of Various Lidar and Radar Sensors for Application in Robotics: A Measurable Capability Comparison”, H. Gim, S. Baek, J. Park, H. Lee, C. Sung, K.-T. Kim, S. Han, IEEE Robotics & Automation Magazine, vol. 30, no. 3, pp. 28-43, 2023

Télémètres laser : mécaniques vs « solid-state »

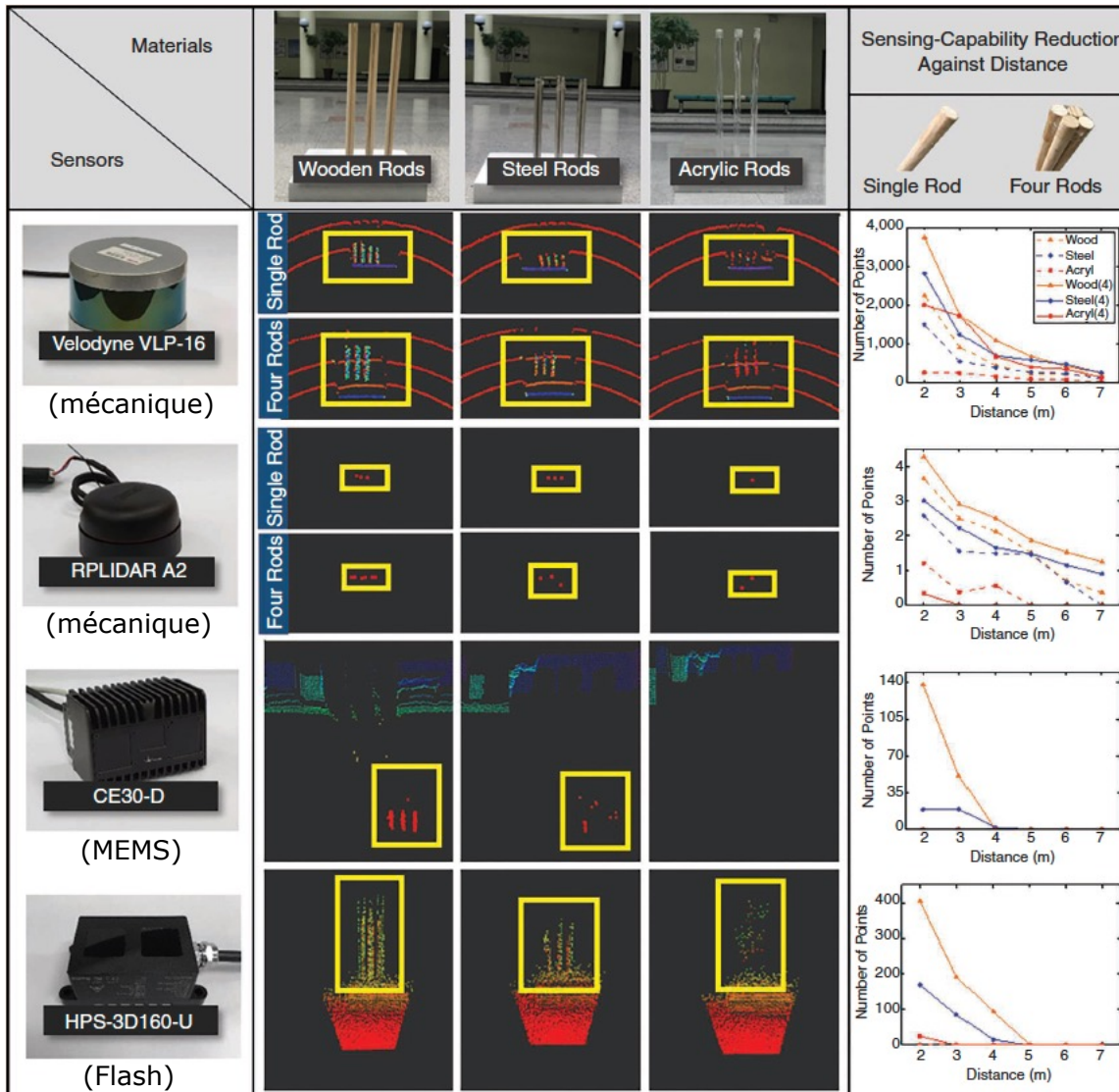
- Capteurs testés dans l'article de Gim *et al.*, RAM'23

Capteur	Type	Champ de vue (deg)	Portée (m)	Résol. angul. (deg)	Résol. dist. (mm)	Résol. pixels	Points/s	Coût (\$)	Long. d'onde (nm)	Freq. (Hz)	Poids (g)
VLP-16	Mécan.	360 × 30	Jusqu'à 100	0.1-0.4 (H), 2 (V)	30	–	600k	5k-6k	905	5-20	830
Ouster OS1 (16 canaux)	Mécan.	360 × 33.2	0.8-120	0.01 (H), 0.01 (V)	30	–	328k	5k-6k	850	10/20	380
RPLIDAR-A2	Mécan.	360	Jusqu'à 18	0.9	0.5	–	8k	319	785	10	190
CE30-D	MEMS	60 × 4	Jusqu'à 30	0.2	10	320 × 20	192k	1.4k	850	–	334
HPS-3D160-U	Flash	76 × 32	Jusqu'à 12	N/D	N/D	160 × 60	336k	269	850	–	140

N/D: Non disponible

“Suitability of Various Lidar and Radar Sensors for Application in Robotics: A Measurable Capability Comparison”, H. Gim, S. Baek, J. Park, H. Lee, C. Sung, K.-T. Kim, S. Han, IEEE Robotics & Automation Magazine, vol. 30, no. 3, pp. 28-43, 2023

Télémètres laser : mécaniques vs « solid-state »



- Étude comparative: capacité à détecter des objets (barres verticales de matériaux différents: bois, acier, acrylique)

Télémètres laser : mécaniques vs « solid-state »

- Résumé des avantages et inconvénients des capteurs testés

Capteurs	Avantages	Inconvénients	Applications adaptées
LIDAR mécanique 3D	<ul style="list-style-type: none"> • Large amount of data (point clouds) • Large FoV (Field of View) • Intensity information • High resolution • Relatively high laser power 	<ul style="list-style-type: none"> • Low vertical resolution • Heavy and large size • Expensive price 	<ul style="list-style-type: none"> • Wide-range 3D scanning • Mapping and localization • Object detection and semantic segmentation
LIDAR mécanique 2D	<ul style="list-style-type: none"> • High distance resolution (2D) • 360° coverage • Low price • Lightweight structure • Relatively high laser power 	<ul style="list-style-type: none"> • Weak to external contact • Semantics are hard to recognize 	<ul style="list-style-type: none"> • 2D mapping and localization • Simple geometry indoor/outdoor driving
LIDAR MEMS (quasi-solid-state)	<ul style="list-style-type: none"> • High durability • Evenly distributed resolution 	<ul style="list-style-type: none"> • Vulnerability to fog • Noise • Narrow FoV 	<ul style="list-style-type: none"> • Mapping and localization in small or narrow indoor environments
Flash LIDAR (solid-state)	<ul style="list-style-type: none"> • High durability • Evenly distributed resolution • Image-like data • Low price • Lightweight 	<ul style="list-style-type: none"> • Vulnerability to fog • Narrow FoV 	<ul style="list-style-type: none"> • Precision recognition • Pose recognition

“Suitability of Various Lidar and Radar Sensors for Application in Robotics: A Measurable Capability Comparison”, H. Gim, S. Baek, J. Park, H. Lee, C. Sung, K.-T. Kim, S. Han, IEEE Robotics & Automation Magazine, vol. 30, no. 3, pp. 28-43, 2023

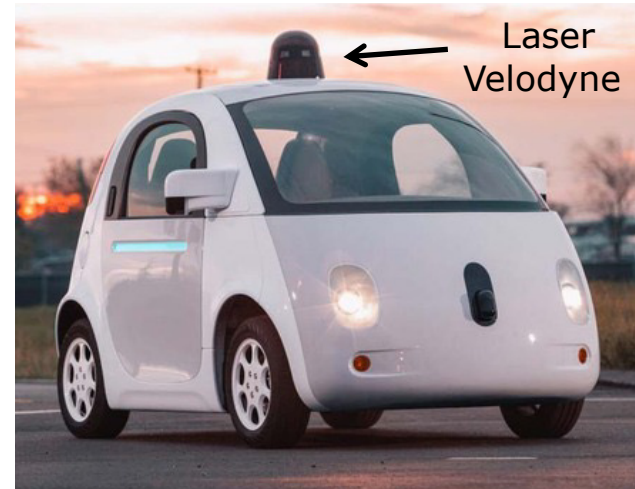
Télémètres laser : applications

- **Voitures sans conducteur**

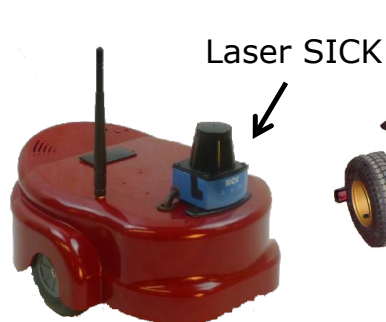
Stanley (Stanford University) gagnant de la DARPA Grand Challenge en 2005



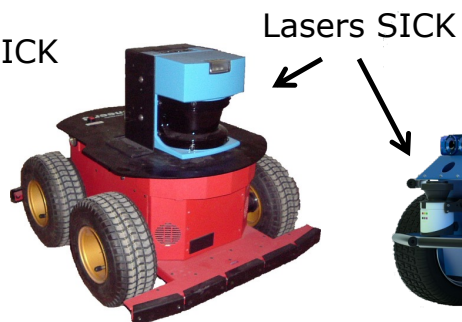
Google car : 2015-aujourd'hui
Waymo, Cruise, Zoox, etc.



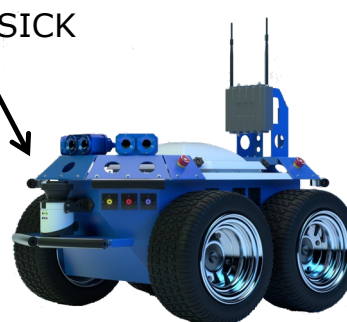
- **Robotique mobile**



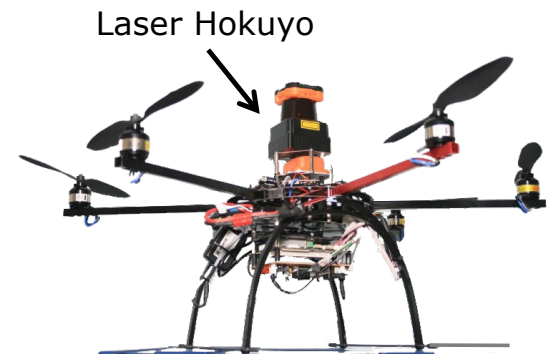
AmigoBot



Pioneer 3-AT



Seekur Jr



HexaKopter

Télémètres laser : applications

- **Modélisation 3D**: numérisation de bâtiments, vestiges archéologiques, etc.



Focus3D X 330 de *Faro*



ScanStation C10 de *Leica*



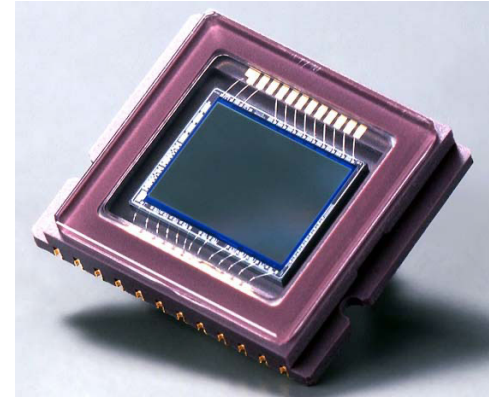
Numerisation 3D du portail sud
de la cathédrale d'Amiens
(Lab. MIS, programme E-cathédrale)

<https://mis.u-picardie.fr/e-cathedrale>

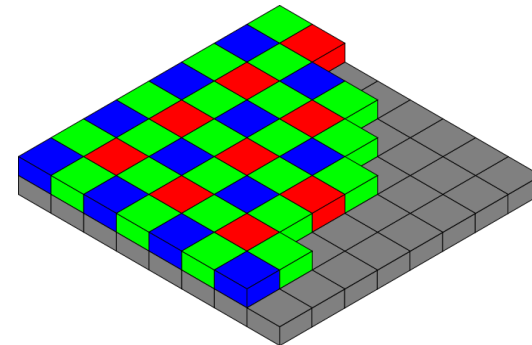
Capteurs de vision



Caméra couleur numérique



Matrice de photorécepteurs (« plane image »)



Matrice de Bayer

Matrice de Bayer: matrice de filtres de couleur placée entre l'*objectif* et la *matrice de photorécepteurs* (technologie CCD ou CMOS) afin de pouvoir enregistrer des images en couleur

Éléments fondamentaux d'une caméra

Caméra = objectif + capteur photosensible



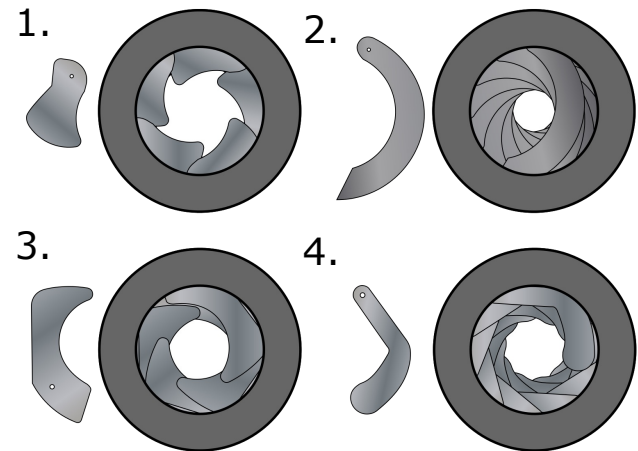
- Objectif

- *Optique* (lentilles)

- Pour changer la mise au point, l'amplitude du champ de vue, etc.

- *Diaphragme* (ouverture à largeur variable)

- Il laisse passer/bloque la lumière vers le capteur photosensible

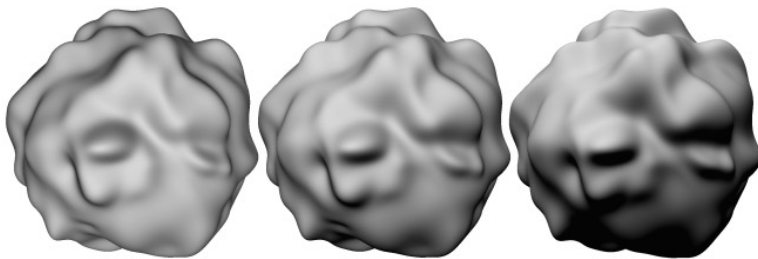


Diaphragme : 4 types d'iris

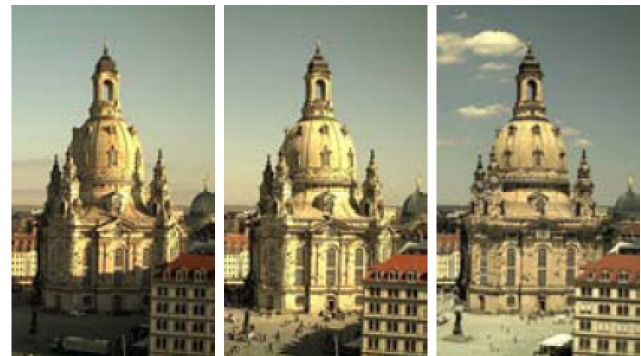
Vision monoculaire

Une image : uniquement information 2D sur la scène

- Manque d'information sur la structure 3D de l'environnement
- Besoin d'information supplémentaire. Plusieurs stratégies possibles:
 - *Taille réelle* des objets observés dans la scène est connue: mise en correspondance 2D-3D
 - *Mouvement* des objets est connu: SfM (« Structure-from-Motion »)
 - *Position de la source lumineuse* est connue:
 - On utilise la position variable des ombres pour percevoir la profondeur (« Shape-from-Shading »)
 - On mesure l'intensité lumineuse pour estimer la profondeur (stéréophotométrie). Hypothèse de base: source lumineuse lambertienne (orthotrope)



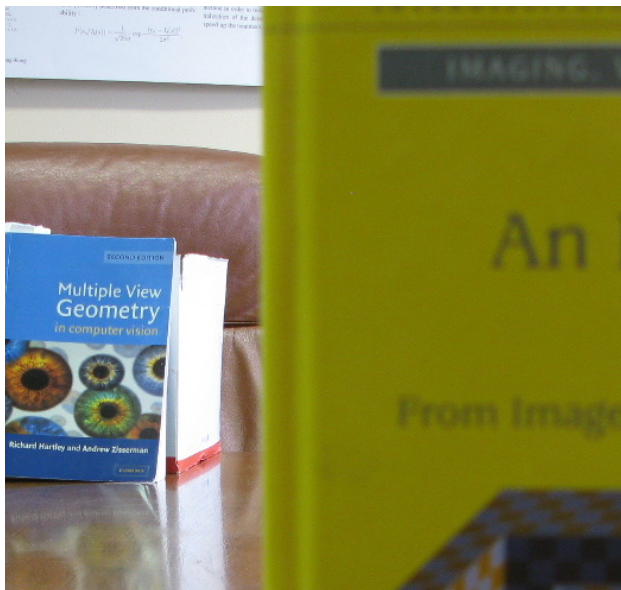
Shape-from-Shading



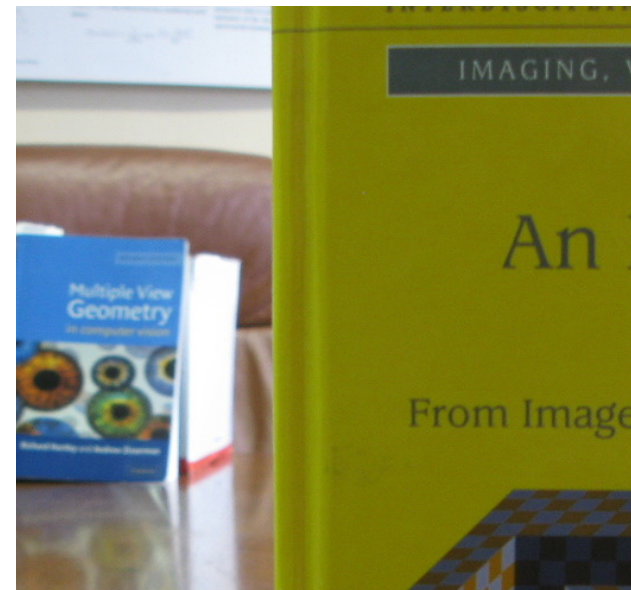
Stéréophotométrie (Ackermann *et al.*, CVPR'12)

Vision monoculaire

- Estimation de la profondeur à partir de la *mise au point* (« Shape-from-Focus »)
 - Il faut quantifier la netteté locale d'une image
 - Lien direct entre la longueur focale f et la profondeur 3D



Mise au point A (arrière-plan)



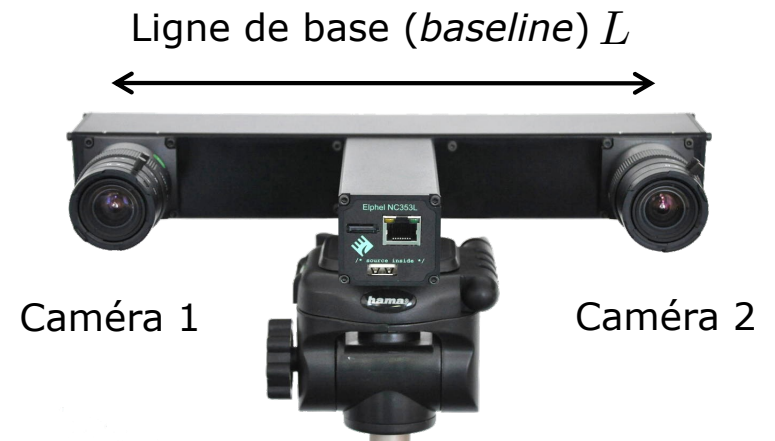
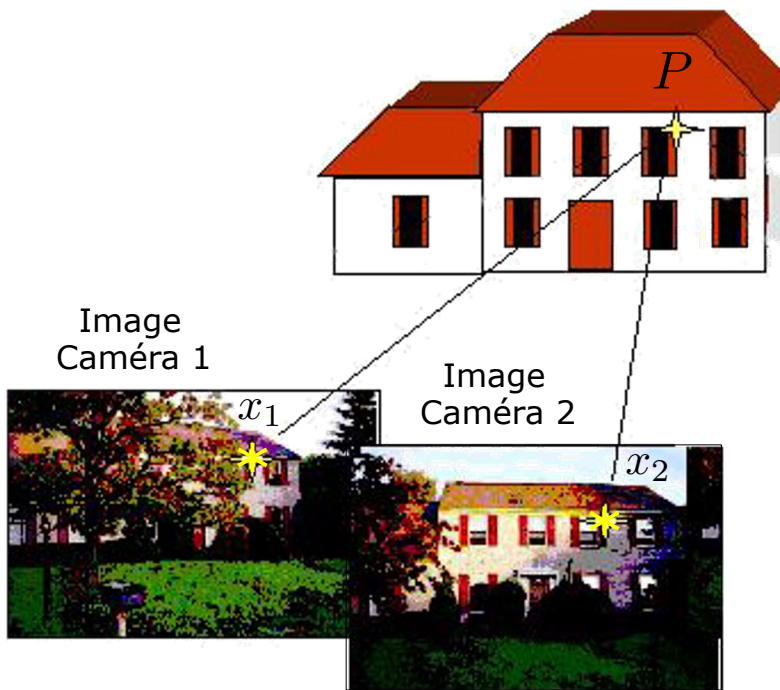
Mise au point B (premier plan)

Vision monoculaire ... et IA (iPhone) !



Vision stéréoscopique

- Deux caméras: **un couple d'images**



- Mise en correspondance
- Triangulation (géométrie)



Information 3D sur la scène

Vision omnidirectionnelle

1. Systèmes multi-caméras

- Image $360^\circ \times 360^\circ$, moyenne/haute résolution
- *Problème*: assemblage des images (aberrations)



Dodeca
2360

(12 caméras, 100 Mpixels,
 $360^\circ \times 360^\circ$, 30 fps*)



Ladybug de *Point Grey*
(0.8 Mpixels \times 6 caméras)



Panono

(36 cameras,
100 Mpixels,
 $360^\circ \times 360^\circ$)

Vidéo



(modèle
2022: pas
à l'échelle)

Google Street View

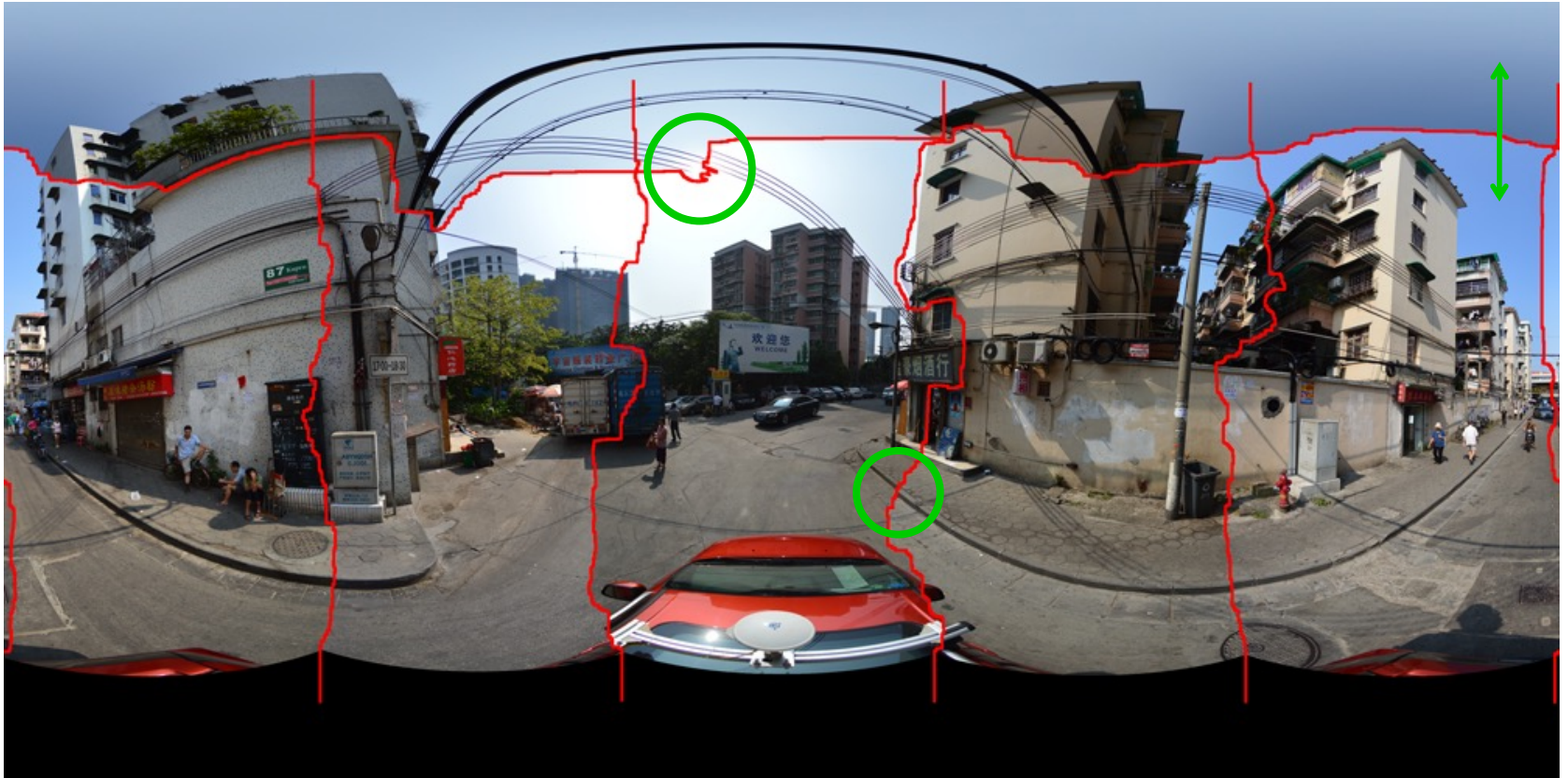
20 Mpixels \times 8 cameras (en 2017)
20 Mpixels \times 7 cameras (en 2022)

*fps = frames per second (ou images/s)

Vision omnidirectionnelle

1. Systèmes multi-caméras

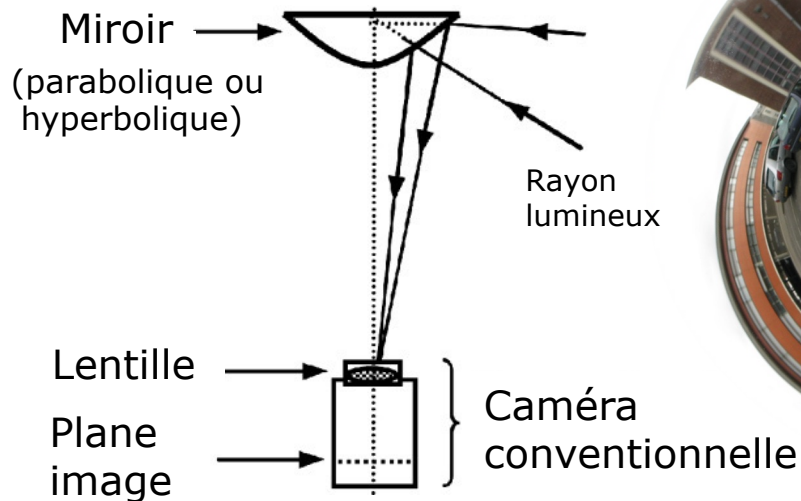
- *Problème*: assemblage (« stitching ») des images



Système multi-caméras: 6 caméras Nikon D7100 (24 Mpixels)
installées sur une voiture (Li *et al.*, Sensors'16)

Vision omnidirectionnelle

2. *Systèmes catadioptriques* (caméra + miroir + lentille)



- Une même primitive (par ex. une droite) est perçue longtemps
- Appariement d'une partie importante de la scène avec l'image



Dépt. EEA (image panoramique ou équirectangulaire)

Vision omnidirectionnelle

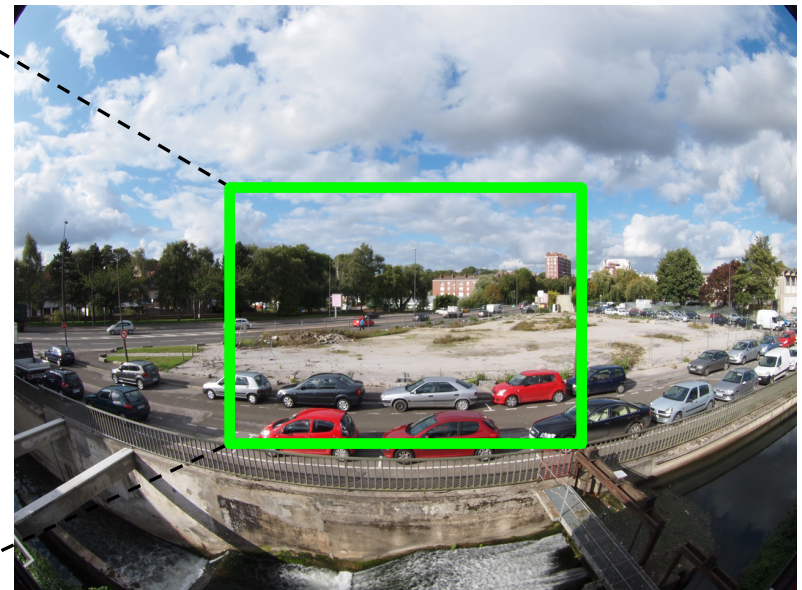
3. *Systèmes à optique « fisheye »*
- Très grand angle ($\geq 180^\circ$)
 - *Problème*: distorsions provoquées par l'optique (lentille)



Nikon 1 V1 avec optique fisheye



Caméra classique (espace devant le département EEA)



Caméra avec optique fisheye

Vision omnidirectionnelle

4. *Systèmes hybrides*: caméras multiples à optique « fisheye »

Salle TP204 du département EEA, prise avec la caméra Theta S 360° montée sur l'effecteur du robot Stäubli TX60

Dual
fisheye



Theta S 360° de Ricoh
(400 \$, 14 Mpixels,
360° × 360°,
enregistrement vidéo)

Vidéo

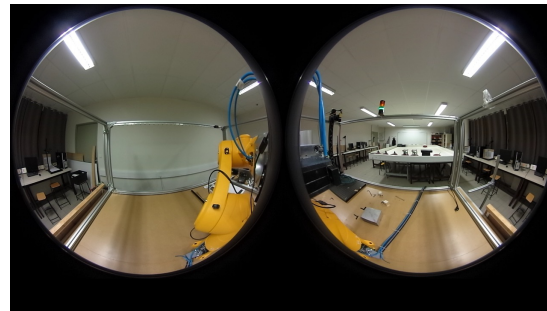


Image dual fisheye



Image équirectangulaire

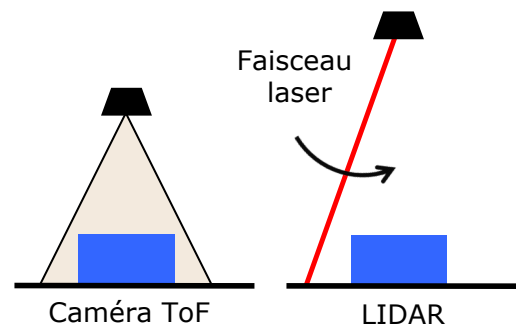
Caméras temps-de-vol

Caméra temps-de-vol (ToF ou Time-of-Flight)

- Projection de *lumière IR modulée*
- Image de profondeur de la scène observée
- Système compact
- Possible perturbation due à d'autres sources de lumière

Exemples:

- Kinect v2 pour XBOX de *Microsoft* (2014-2017)
 - Fréquence: 30 Hz
 - 512 pixels × 424 pixels (caméra de profondeur)
 - 1920 pixels × 1080 pixels (caméra RGB)
 - Champ de vue: 84.1° (H) × 53.8° (V)
 - Prix: 150 €
- RealSense L515 de *Intel* (2021-2022)
 - 30 fps (caméra RGB et de profondeur)
 - Prix: 369 \$
- Helios2 de *Lucid Vision Labs* (2023-)
 - 640 pixels × 480 pixels (caméra de profondeur)
 - Longueur d'onde: 850 nm (à l'intérieur)
 - 30 fps
 - Portée: 0.3-8.3 m
 - Champ de vue: 69° (H) × 51° (V)
 - Prix: 1495 \$

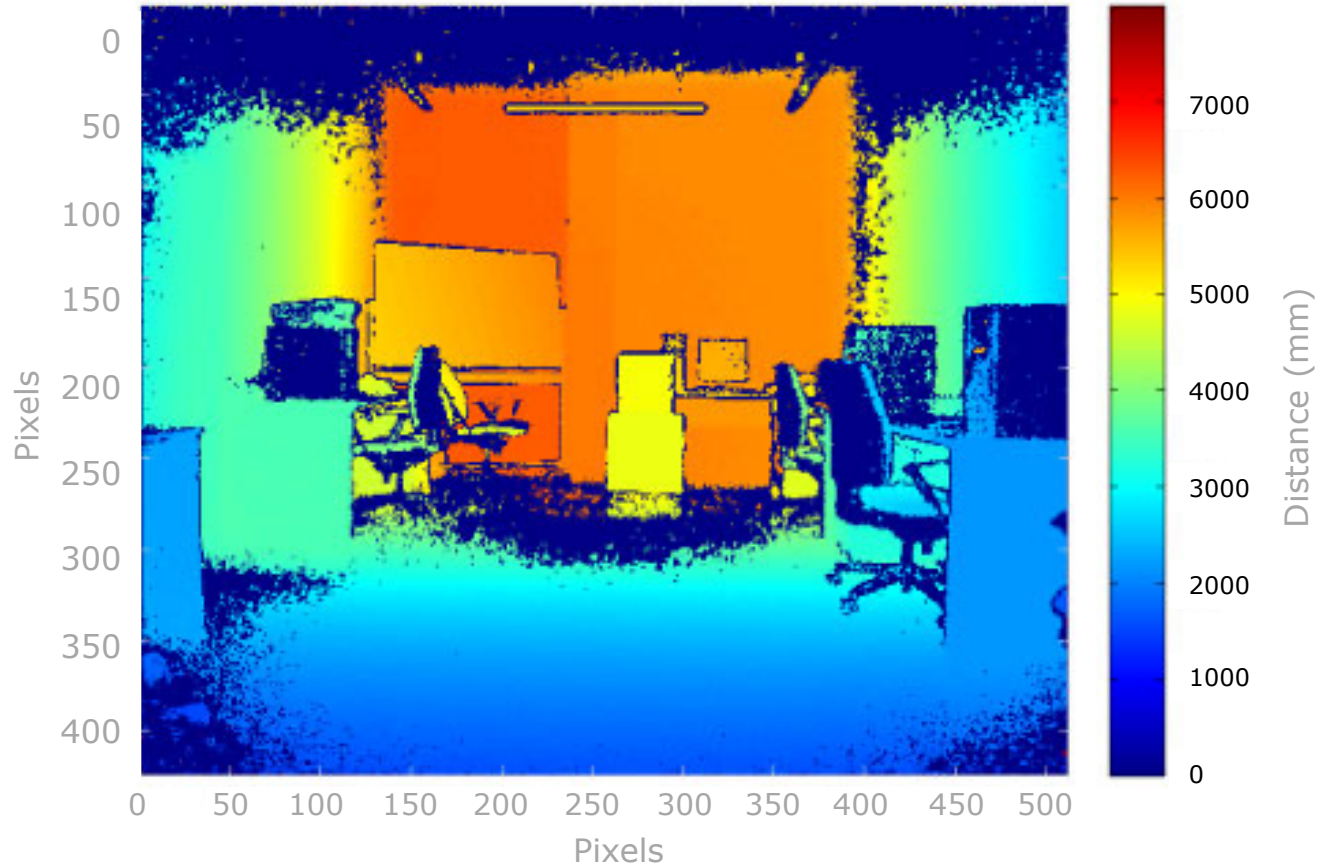


Caméras temps-de-vol

- Kinect v2 de *Microsoft*



Image ou carte de profondeur



"An Empirical Evaluation of Ten Depth Cameras", G. Halmetschlager, M. Suchi, M. Kampel, M. Vincze, IEEE Robotics & Automation Magazine, vol. 26, n. 1, pp. 67-77, 2018

Capteurs de triangulation

- *Triangulation optique* (capteur 1D)
 - Lumière laser (IR)
 - Émission
 - Réflexion
 - Mesure de distance z d'un objet

$$z = \frac{fd}{p + f \tan \theta}, \quad x = z \tan \theta$$

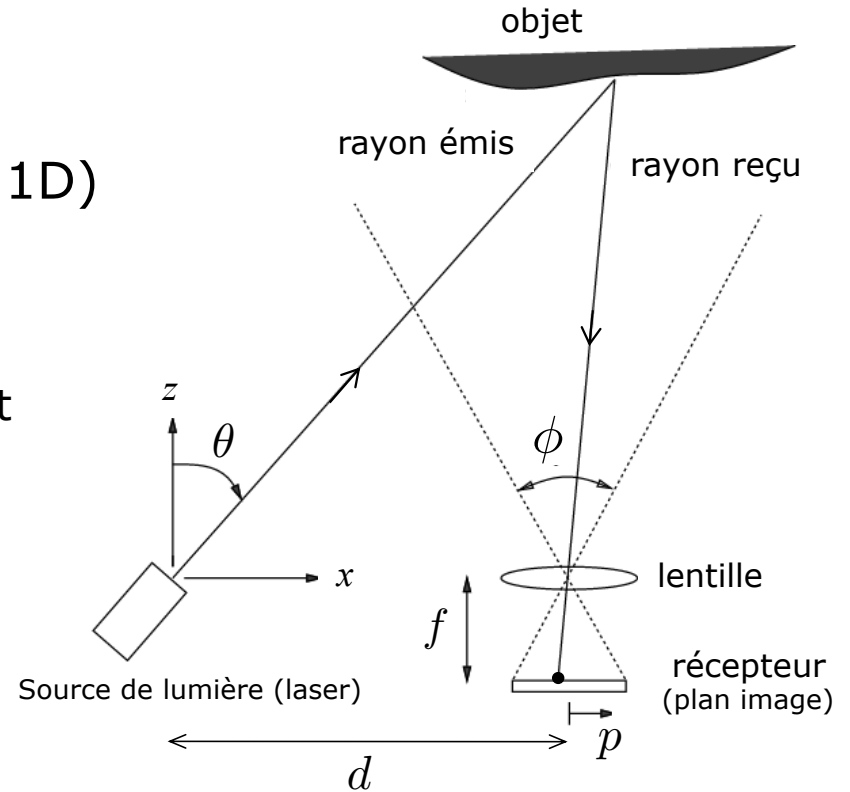
avec

f : distance focale du récepteur (estimée)

d : distance émetteur-récepteur (connue)

Exemple: 16 *proximètres IR*
sur le robot Koala II de *K-Team*

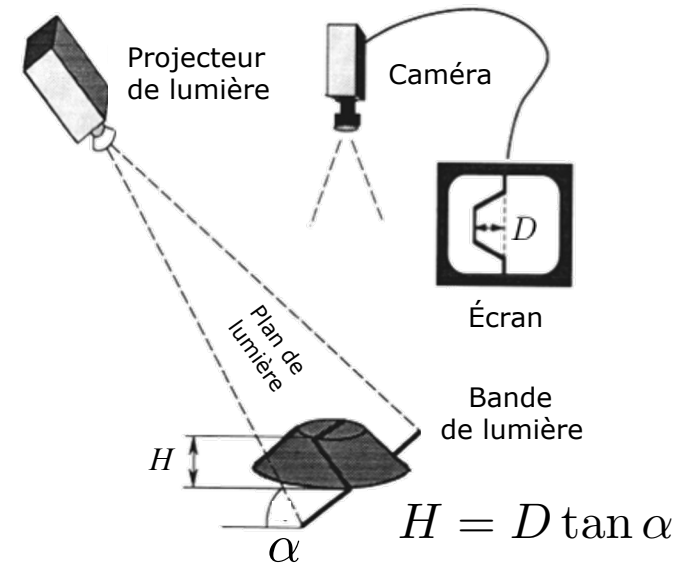
- Portée: 5-20 cm
- Prix: ~ 5 €/capteur
- « Bumper immatériel »



Capteurs de triangulation

- *Lumière structurée* (capteur 2D)

- Une caméra
- Un projecteur de motif connu
 - *Vision active*
 - Source lumineuse modulée dans l'espace, le temps, l'intensité et/ou la couleur
- Calcul de profondeur simplifié par rapport à la stéréovision classique
 - *(Stéréo)vision passive*: corrélation entre images/environnement
 - *Vision active*: mesure directe, fonctionne dans le noir, pas besoin de texture



Exemples: Kinect v1 de Microsoft (2011-2017), Xtion Pro Live de ASUS, ASTRA PRO de Orbbec, RealSense F200 de Intel

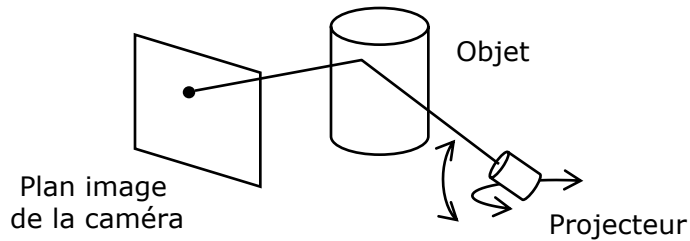


Kinect v1



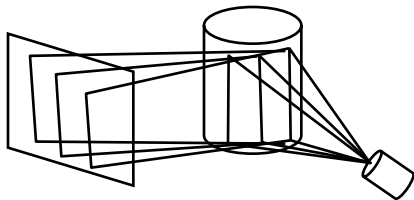
Motif IR projeté par la Kinect v1

Lumière structurée : définitions (D. Fofi)



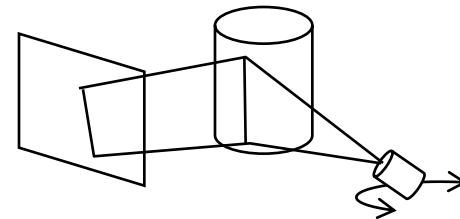
1 - Point de surbrillance :

- Pas de problème de mise en correspondance
- Balayage suivant les deux axes



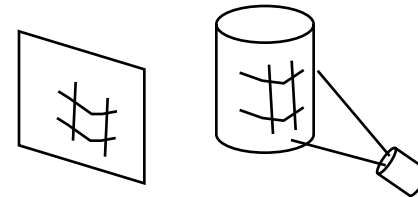
3 - Multi-plans :

- Correspondance entre les plans ?
- Pas de balayage



2 - Plan de lumière :

- Correspondance entre points de la ligne projetée ?
- Balayage suivant un axe



4 - Motif structurant :

- Problème de mise en correspondance
- Pas de balayage

Lumière structurée : classification

3 critères selon les propriétés du motif projeté :

- Dépendance temporelle (**statique** ou **dynamique**)
- Nature de la lumière émise (**binaire**, **niveaux de gris** ou **couleurs**)
- Dépendance aux discontinuités (**périodique** ou **absolu**)

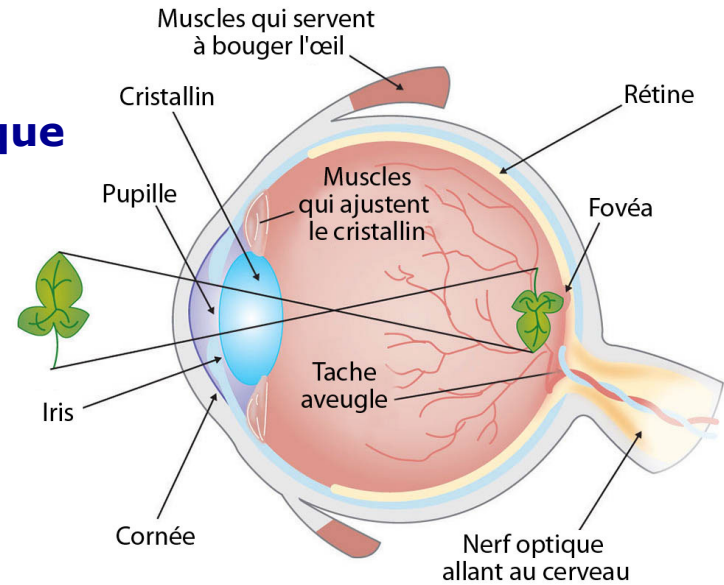
Classification proposée dans l'article :

"Recent progress in coded structured light as a technique to solve the correspondence problem: a survey", J. Batlle, E. Mouaddib, J. Salvi, Pattern Recognition, vol. 31, n. 7, pp. 963-982, 1998

Caméras événementielles

L'œil humain comme une caméra numérique

- 10^8 de photorécepteurs
- 10^6 de fibres en sortie.
Taux d'activation maximum: 100 Hz
- Plage dynamique: 90 dB (10^9)
- Plus de 20 yeux différents
- Consommation énergétique: 3 mW



Sortie: *Flot épars et asynchrone d'événements ponctuels*

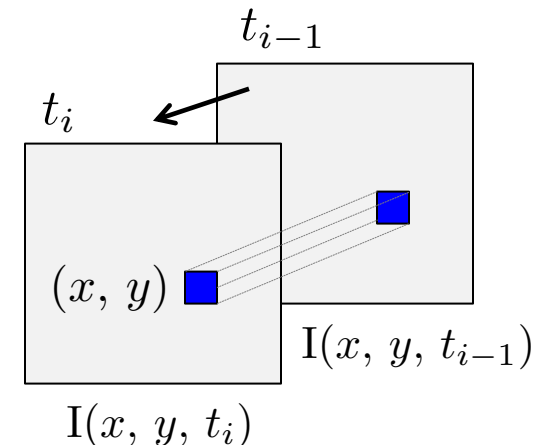
➡ Adoptons une approche neuromorphique !

Caméra événementielle (ou "event-based")

La caméra produit en sortie des *événements asynchrones* (x, y, t, p) où $p \in \{-1, 1\}$ est la polarité, avec une résolution temporelle de l'ordre de la microseconde. On a un événement lorsque:

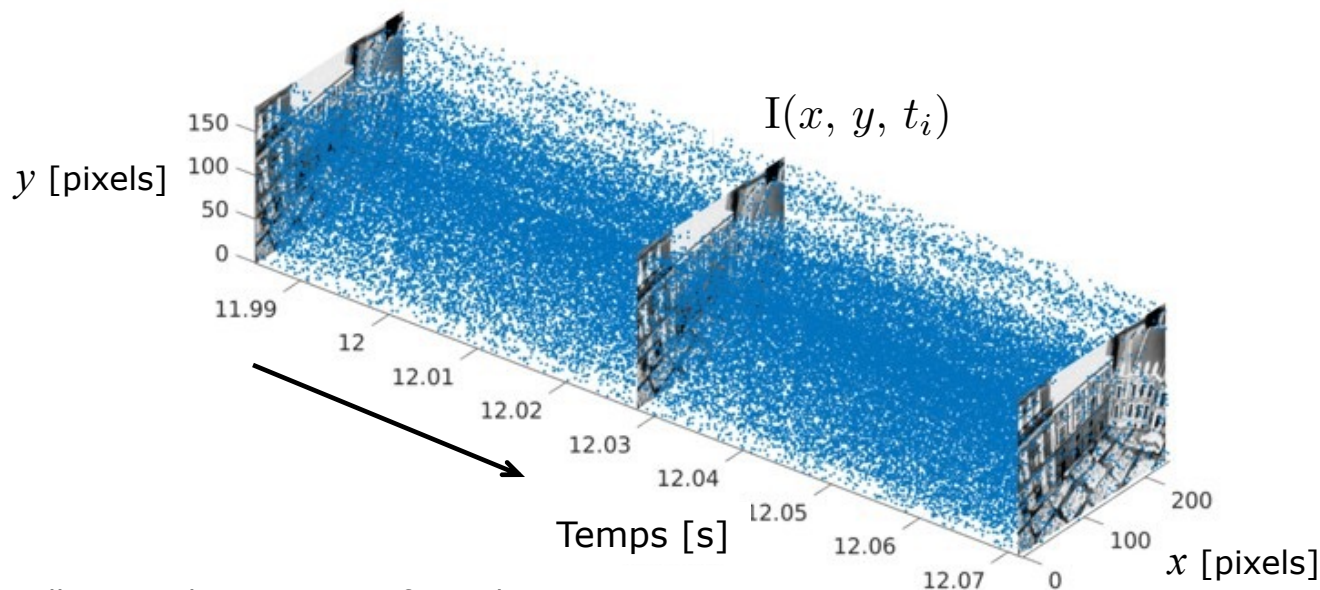
$$|\log(I(x, y, t_i)) - \log(I(x, y, t_{i-1}))| \geq \gamma$$

où γ est un seuil donné



Caméras événementielles

- Équivalente (ou supérieure) à une *caméra standard haute vitesse* à plus de 1000 fps (pas d'éclairage spécial, latence très faible)
- *Réduction drastique* de la consommation énergétique, de la quantité de données à stocker (pas de redondance) et des ressources de calcul
- *Plage dynamique*: augmentation de plusieurs ordres de grandeur grâce au traitement local (supérieure à 120 dB)



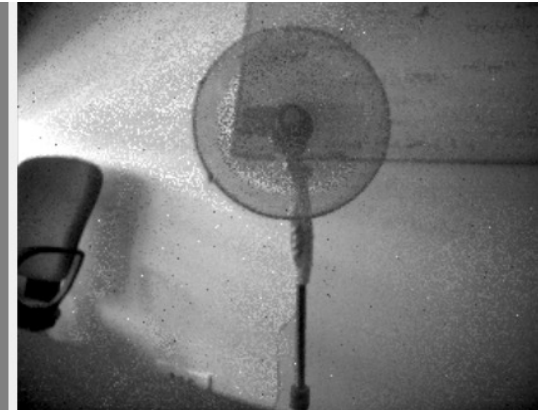
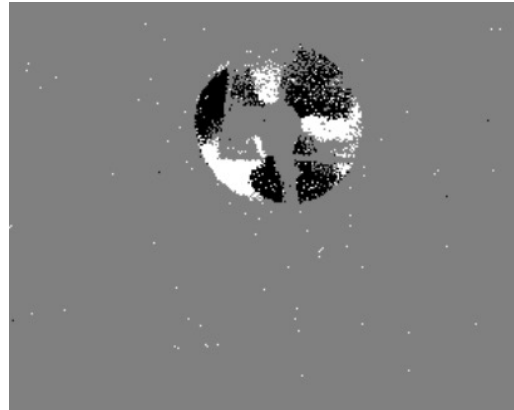
Source: Gallego et al., University of Zurich

Caméras événementielles commerciales

- **Evaluation Kit 1 - VGA** de *Prophesee* (2017), www.prophesee.ai



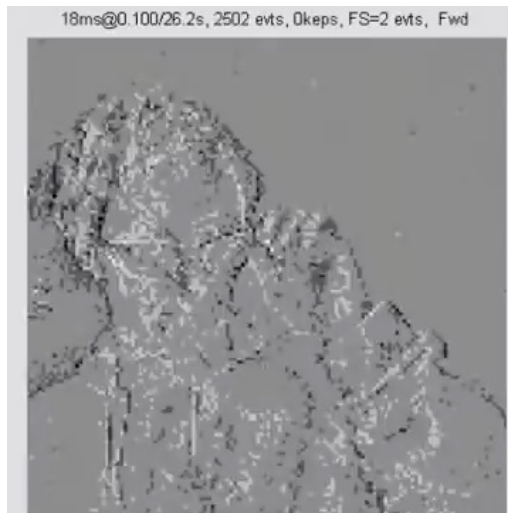
Ventilateur en mouvement



Événements

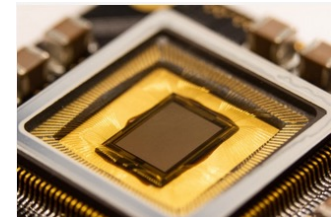
Image en niveau de gris

Résolution [pixels]	640 × 480
Latence [μ s]	40-200
Plage dyn. [dB]	> 120



- **DVS** (Dynamic Vision Sensor) de *IniVation*
<https://inivation.com/dvp>

Latence [μ s]	< 100
Plage dynamique [dB]	> 120

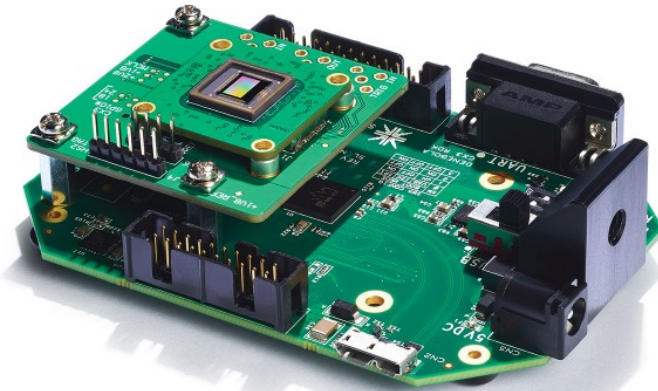


Pour plus d'infos voir :

- "Event-based Vision: A Survey", G. Gallego *et al.*, IEEE Trans. Pattern. Anal. Mach. Intell., vol. 44, n. 1, pp. 154-180, 2022
- "Recent Event Camera Innovations: A Survey", B. Chakravarthi, A.A. Verma, K. Daniilidis, C. Fermuller, Y. Yang, in Proc. Europ. Conf. Comp. Vis., 2024

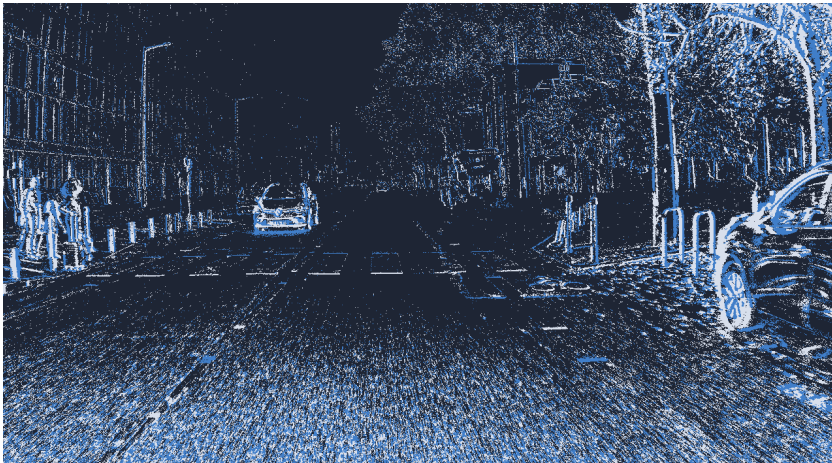
Caméras événementielles commerciales

- **Evaluation Kit 3 – HD (EVK3)** de *Prophesee* (2021)



Spécifications de la caméra EVK3

Résolution [pixels]	1280 × 720
Latence [μ s]	220
Plage dynamique [dB]	> 110
Taille des pixels [μ m]	4.86 × 4.86
Max. larg. de bande [Gb/s]	1.6
Puissance absorbée [W]	4.5
Poids (optique exclue) [g]	112
Alimentation	USB 3.0



“Image événementielle” acquise par la caméra EVK3 embarquée sur un véhicule en mouvement

Polarité: blanc $p = 1$, bleu $p = -1$

Typologies de capteur : Tableau récapitulatif

A : Actif
P : Passif
PC : Proprioceptif
EC : Extéroceptif

Classification (utilisation typique)	Capteur/système de perception	PC ou EC	A ou P
Capteurs tactiles (détection de contact physique ou de proximité, interrupteurs de sécurité)	Interrupteurs de contacts, bumpers	EC	P
	Barrière optique	EC	A
Capteurs de roue/moteur (vitesse et position de roue/moteur)	Encodeurs à balais	PC	P
	Potentiomètres	PC	P
	Encodeurs optiques	PC	A
	Encodeurs magnétiques	PC	A
	Encodeurs inductifs	PC	A
	Encodeurs capacitifs	PC	A

Classification (utilisation typique)	Capteur / système de perception	PC ou EC	A ou P
Capteurs d'orientation (orientation du robot en relation à un référentiel fixe)	Compas	EC	P
	Gyroscopes	PC	P
	Inclinomètres	PC	A/P
Basé balise (localisation dans un référentiel fixe)	GPS	EC	A
	Balises actives optiques ou radio	EC	A
	Balises actives à ultrasons	EC	A
	Balises réfléchives	EC	A

Classification (utilisation typique)	Capteur / système de perception	PC ou EC	A ou P
Télémétrie active (réflectivité, temps de vol et triangulation géométrique)	Capteurs de réflectivité	EC	A
	Capteurs à ultrasons	EC	A
	Télémètres laser	EC	A
	Triangulation optique (1D)	EC	A
	Lumière structurée (2D)	EC	A
Capteurs de mouvement/vitesse (vitesse relative à des objets statiques ou fixes)	Doppler radar	EC	A
	Doppler sonore	EC	A
Capteurs de vision (télémétrie visuelle, analyse de l'image complète, segmentation, reconnaissance d'objet)	Caméras à capteur CCD/CMOS	EC	P