Initiation à la Vision par Ordinateur

Bertrand VANDEPORTAELE

Département GEII TOULOUSE

bertrand.vandeportaele@iut-tlse3.fr
BUT 3 AII

December 4, 2023

Sommaire

1. Introduction
2. Capteurs, plateformes et calculateurs
   - Capteurs
   - Plateformes
   - Calculateurs
   - Systèmes de vision industriels
3. Traitement d'image
4. Géométrie pour la vision
   - Modélisation Géométrique
   - Géométrie projective
   - Étalonnage Géométrique
   - Stéréovision
   - Applications
5. Considérations annexes

Section 1

Introduction

Objectifs

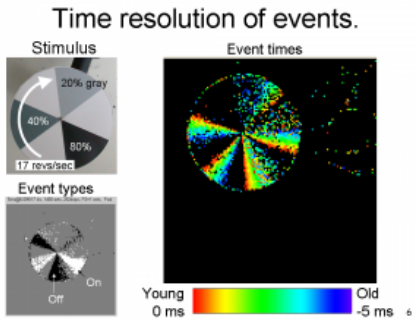
- La perception pour la robotique est un vaste domaine.
- La vision par ordinateur (au sens large) (Computer Vision) recouvre:
  - Traitement du signal
    - 2D: image
    - 3D: séquence d'images
  - Géométrie
  - Modélisation de l'incertitude
  - I.A. et apprentissage
  - Sémantique
  - Couplage Action et Perception
  - ...
- Ce cours est une introduction

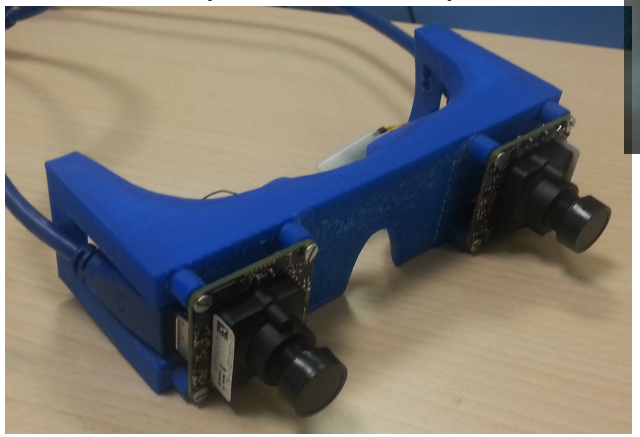
Section 2

Capteurs, plateformes et calculateurs

Perception pour la robotique: Capteurs Passifs

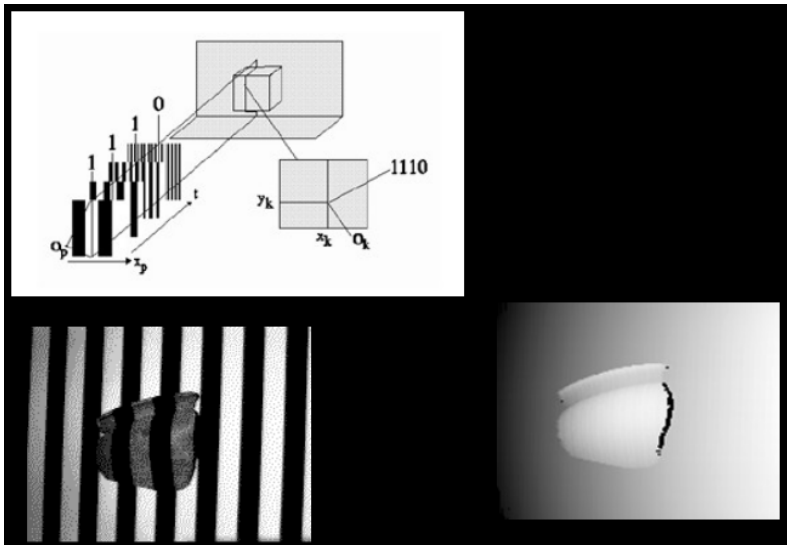
- Mesure la lumière émise ou réfléchie par la scène depuis l'éclairage présent
  - caméra monoculaire
  - banc stéréo (caméra multi-oculaire)
  - caméras DVS
  - Caméra infrarouge proche
  - Caméra infrarouge lointain (thermique)

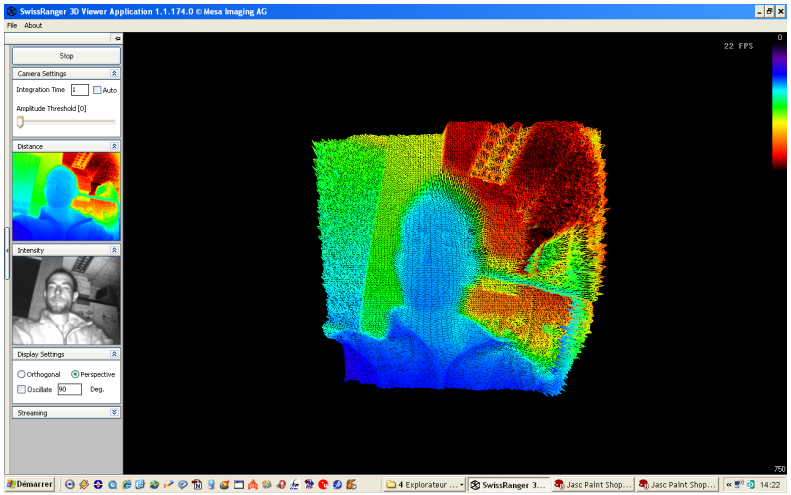




Perception pour la robotique: Capteurs Actifs

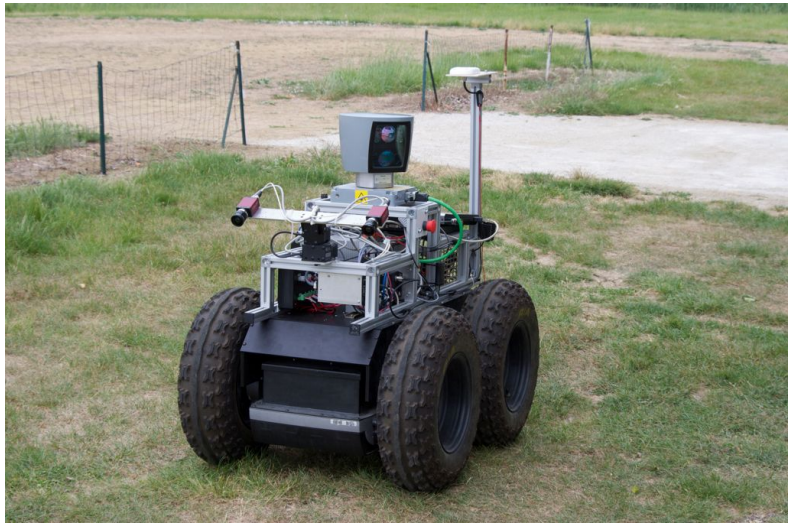
- Emet sa propre source de lumière et mesure celle réflechie par la scène
  - Vision avec projection de lumière structurée ou non
  - Temps de vol (TOF)
    - Ultrasons
    - Télémètres Lidar à balayage
    - Capteurs matriciels RGB-D (Kinect 2), plutôt en intérieur
- Shape From Shading

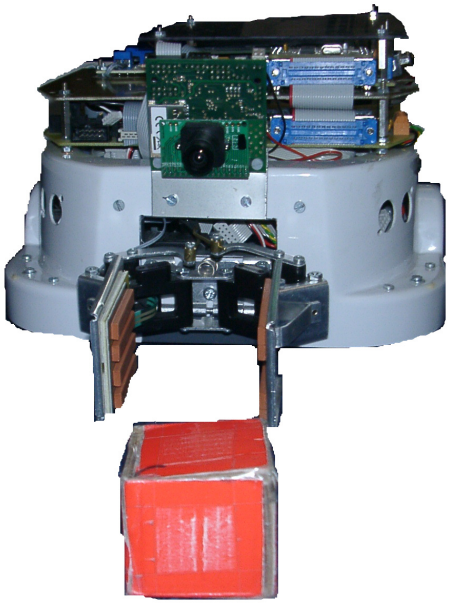




Perception pour la robotique

- Déplacement et orientation (Pan-Tilt)
- Zoom motorisé
- Différentes contraintes
  - Prix, poids, compacité
  - Grande dynamique
  - Cadence, latence
  - Robustesse

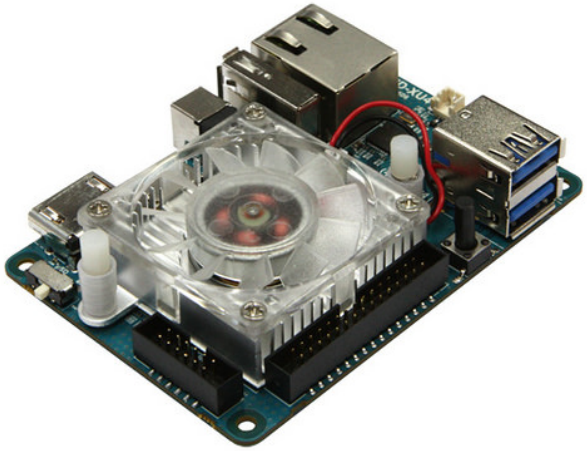


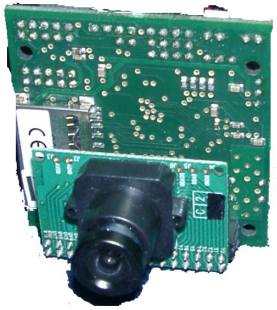


Perception pour la robotique: Calculateur

Architecture matérielle

- Considérer:
  - Calculateur
  - Mémoire (taille et bande passante)
  - Interfaces
- CPU(s)
  - Micro-controleur 8bits 20Mhz
  - Core i7 64bits 4 coeurs 4Ghz
  - Multi-threading pour traitement en //
- DSP
- GP-GPU : Traitement massivement //
- FPGA ou ASIC : Architecture plus ou moins configurable optimisée
- Smart-Camera: Capteur et calculateur imbriqués





Perception pour la robotique: Systèmes de vision industriels

Exemples de cameras industrielles

- https://www.baumer.com/fr/fr/apercu-des-produits/cameras-industrielles-traitement-dimage/cameras-industrielles/c/331
- https://www.cognex.com/fr-fr/products/machine-vision/vision-software/vision-accessories/cameras
- https://www.raspberrypi.com/products/raspberry-pi-global-shutter-camera/
- https://www.zebra.com/fr/fr/products/industrial-machine-vision-fixed-scanners/smart-sensors-and-cameras.html?page=1&tactic_type=SEMG&tactic_detail=AV_BRXFR_KWD_EMEA_FR-FR_Search_7170000011421803 2&ds_eid=700000002161721&ds_cid=7170000011421803 2&ds_agid=5870000853081914 3&ds_kids=p77815490162&gclid=

Perception pour la robotique: Cameras industrielles

- Robustesse aux environnements durs
  - Température
  - Poussière
  - Vibrations
  - Humidité
- Optique intégrée ou monture pour l'optique
  - Diamètre de l'optique (luminosité vs encombrement+coût)
  - Réglages de l'optique
  - Focale (objectifs grand angle ou fort zoom)
  - Ouverture diaphragme (quantité de lumière vs profondeur de champ)
- Filtres
  - Longueurs d'onde (visible/UV/IR)
  - Polarisants (réflexion)
- Type d'obturateur (global ou rolling shutter)
- Nombre de canaux de couleurs (bayer/multi)

Perception pour la robotique: Cameras industrielles

- Surface et résolution du capteur
- Paramétrage d'exposition: durée d'exposition et gain
- Dynamique du capteur
- Cadence (fps) et latence (ms)
- Taille du buffer d'images (mémoire)
- Entrée/sorties de synchronisation (caméras ou illuminateurs)
- Interfaces de communication pour récupérer les images (GigE, ethernet M12, USB 2/3,...)
- Unité de (pré)traitement intégrée ou non : capteurs de vision intégrés (smart...)
- Drivers pour récupérer les images ou les résultats de traitement (OS système et langage API)
- Logiciels propriétaires ou ouverts pour faire les traitements sur PC

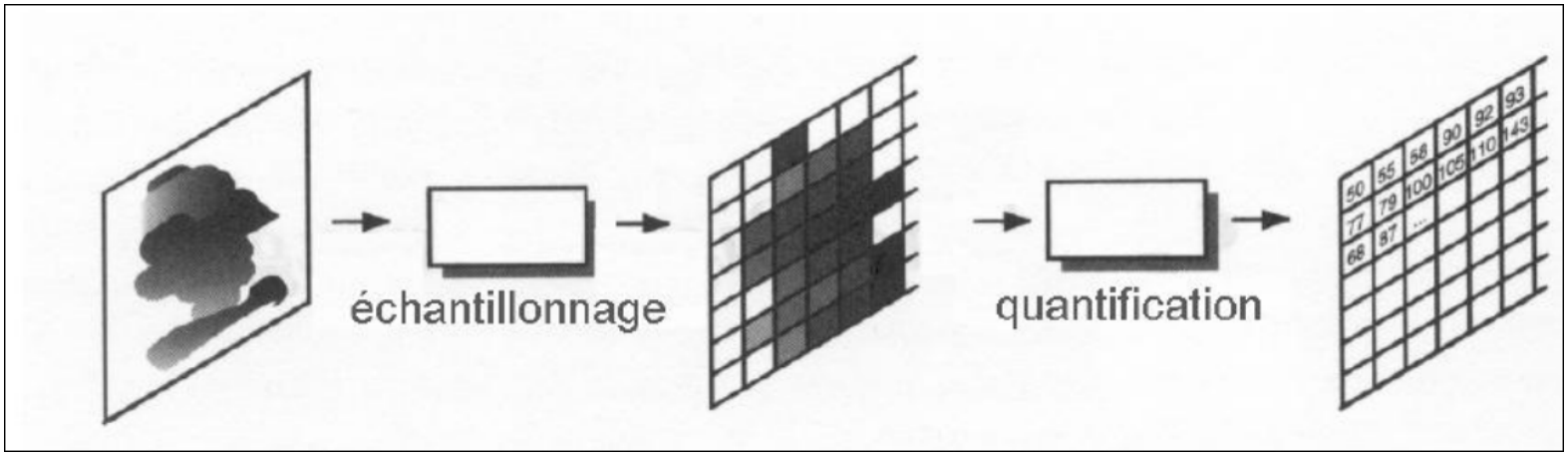
Perception pour la robotique: Systèmes d'éclairage

- Différentes longueurs d'ondes
- Diffus:
  - Panneaux ou anneaux de LEDs
  - Film diffuseur
- Directifs
- Continus ou flash
- Laser avec motifs:
  - Ligne(s)
  - Point(s)
  - Mouchetis
- Position par rapport à la caméra (intégré ou déporté)
  - Ex:caméra endoscopique pour l'inspection

Section 3

Traitement d'image

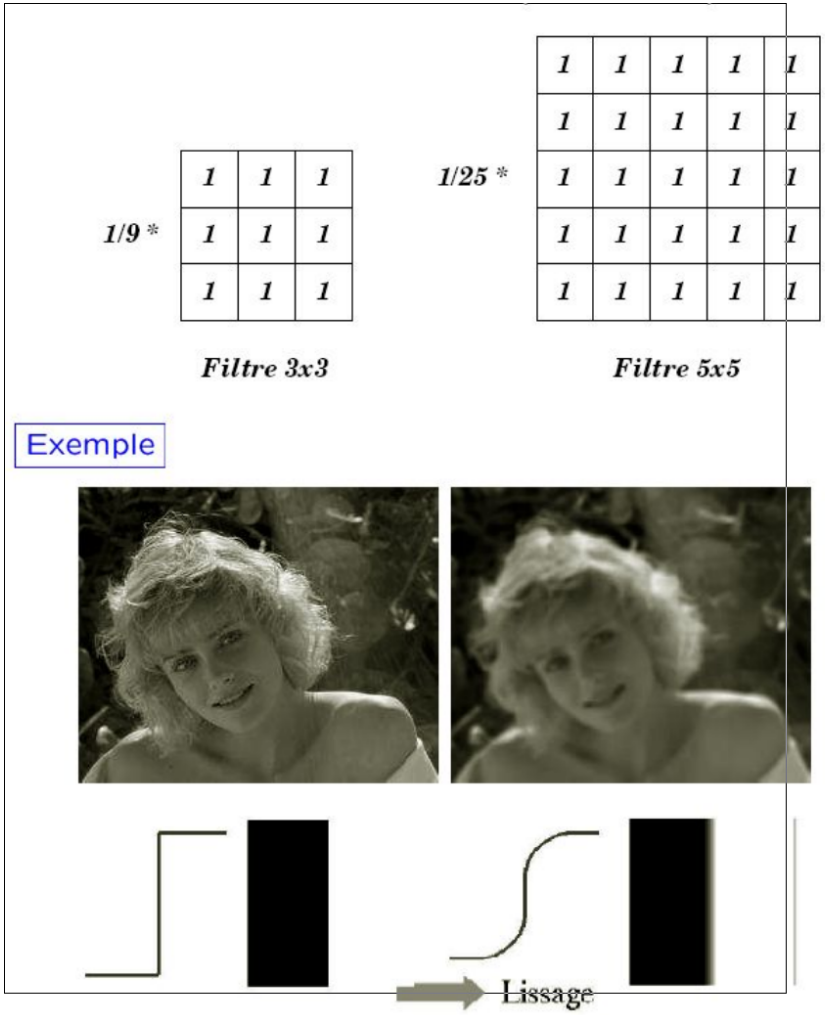
Qu'est ce qu'une image numérique



1

[1] http://www.isima.fr/~vbarra/IMG/pdf/diapos_cours.pdf

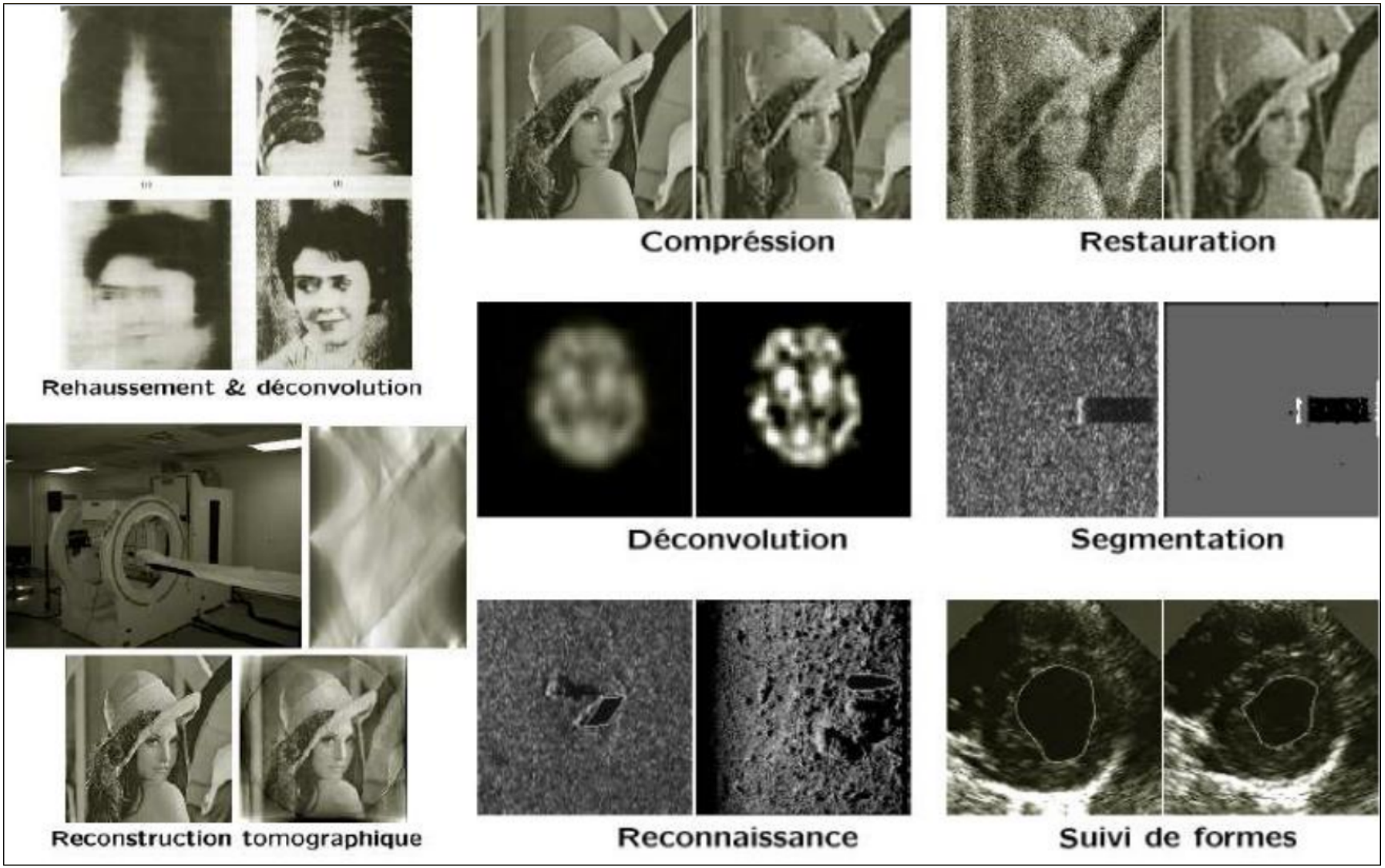
Exemple de traitement: filtrage



2

[2] http://www.isima.fr/~vbarra/IMG/pdf/diapos_cours.pdf

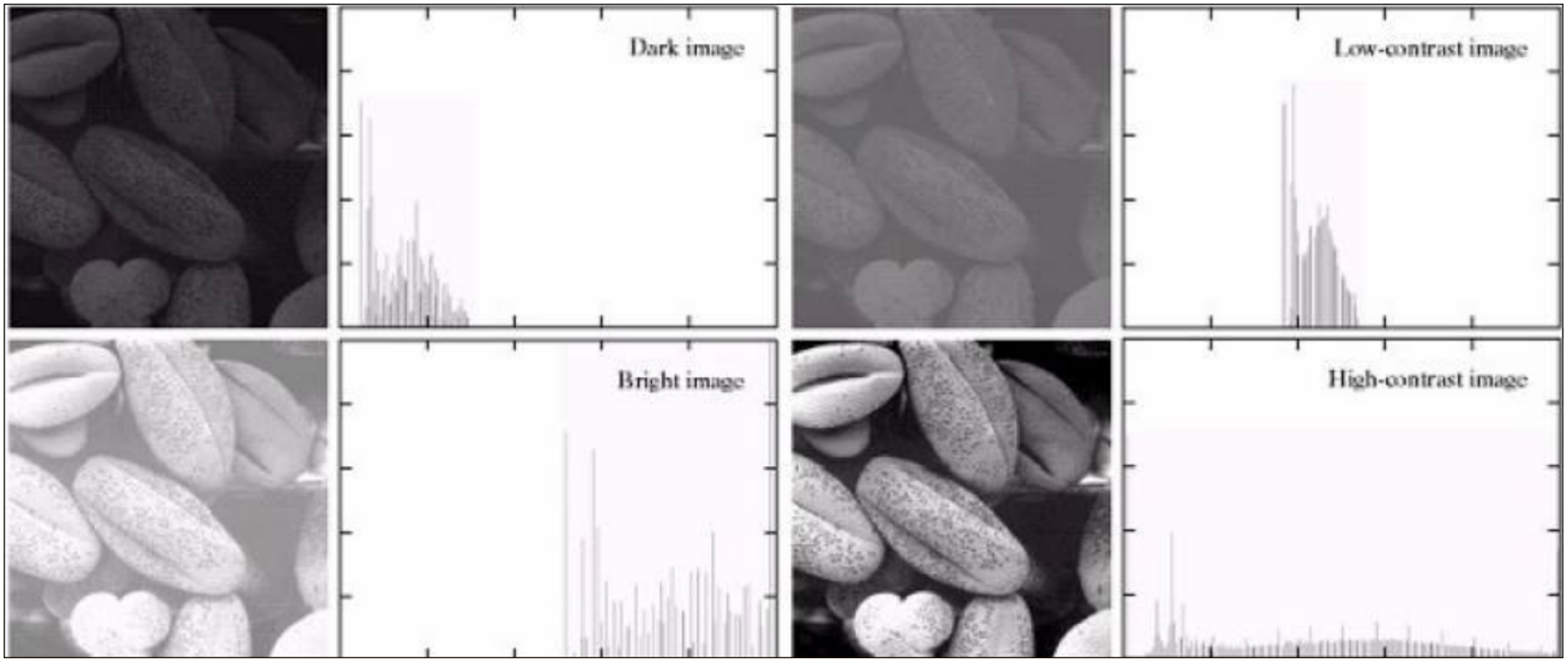
Autres exemples de traitements



[3]

[3] http://www.isima.fr/~vbarra/IMG/pdf/diapos_cours.pdf

Histogramme

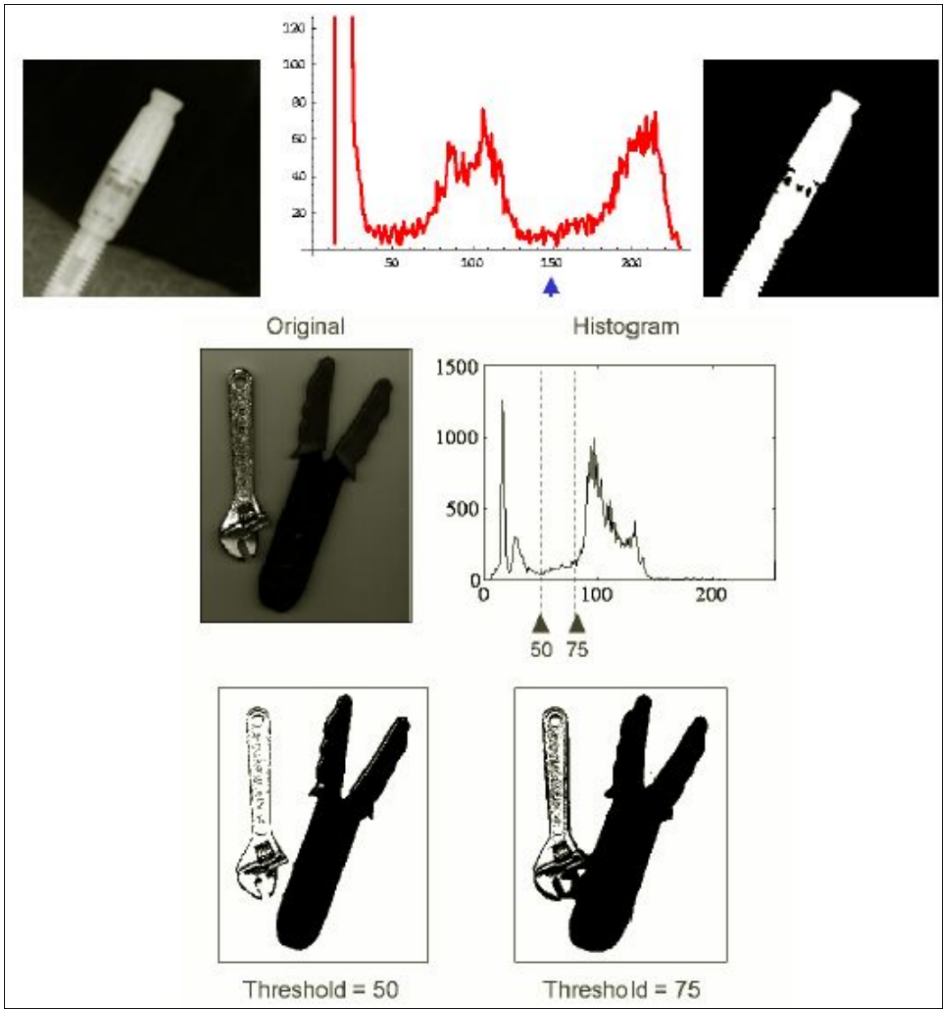


- Correction par calcul et application d'une table de correspondance (LUT)
- N'améliore l'image que pour la visualisation

[4]

[4] http://www.isima.fr/~vbarra/IMG/pdf/diapos_cours.pdf

Seuillage

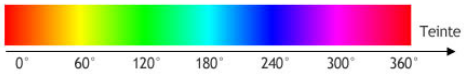


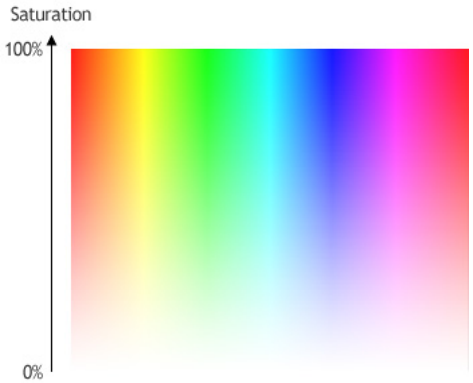
[5]

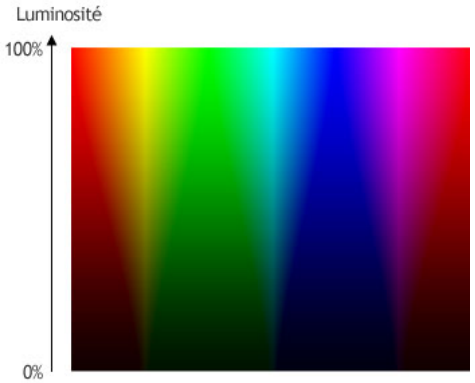
[5] http://www.isima.fr/~vbarra/IMG/pdf/diapos_cours.pdf

Couleur, niveau de gris, image binaire

- GRAY
- RGB, BGR...
- RGBA
- Mais aussi d'autres espaces colorimétriques
  - Teinte Saturation Luminosté
  - YUV
  - CMJN
  - ...







Opérateur de morphologie mathématique

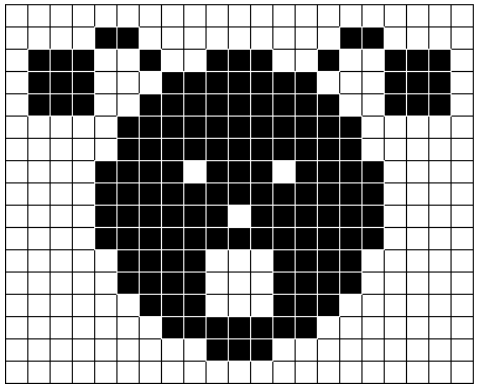
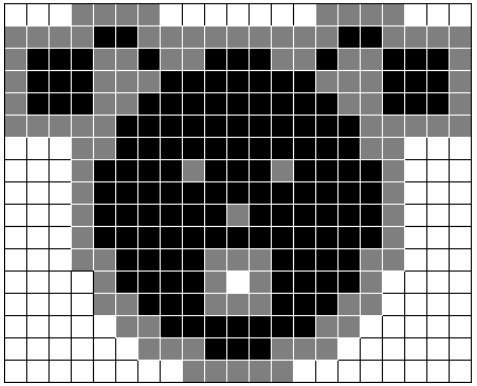
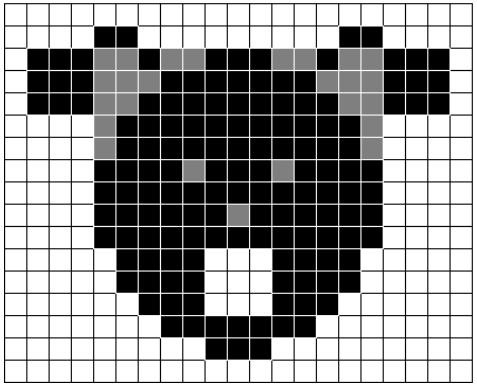


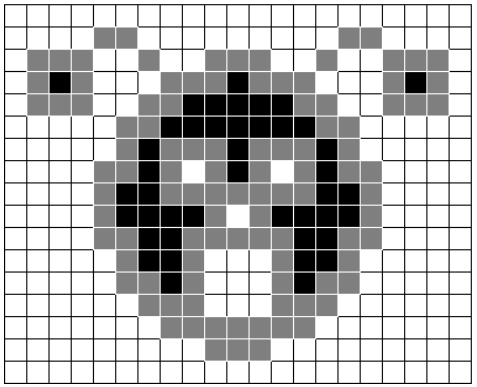
image originale



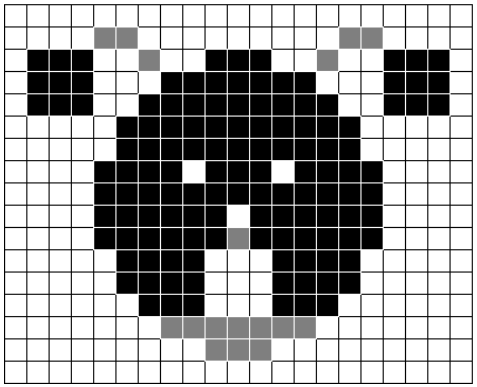
Dilatation



Fermeture



Erosion



Ouverture

[6]

[6] https://fr.wikipedia.org/wiki/Morphologie_math%C3%A9matique

Segmentation et étiquettage d'image

Attribution d'un numéro à chaque pixel = label

Image d'origine

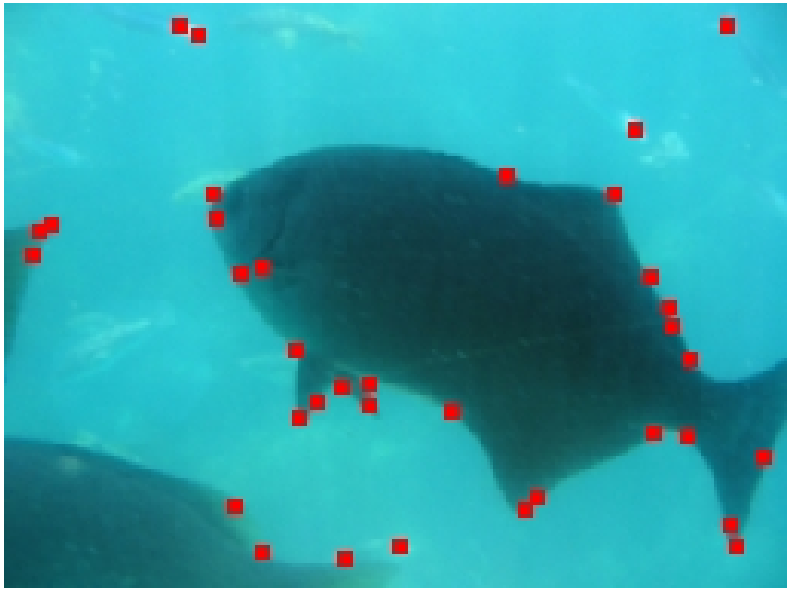
Image segmentée

[7]

[7] http://www.umiacs.umd.edu/~mishraka/activeSeg.html

Extraction et appariement de points

- Détecteur de points invariants à des changements:
  - de point de vue (répétabilité), d'échelle, de rotation, de transformation affine, d'illumination
- Ces points d'intérêt sont généralement des coins: Harris, KLT, SIFT, FAST, ORB, BRIEF...
- On associe parfois un descripteur au point pour pouvoir l'apparier avec une observation dans une autre image



[8]

[8] http://www.roborealm.com/help/Harris_dest.jpg

Extraction de contours et segments

Extraction de structures géométriques

Exemple de détection de coniques

Extraction de plans

- Analyse des points de fuite associés aux directions détectées
- Ce n'est plus du traitement d'image, mais de la Vision, traitement qui nécessite une analyse plus globale, à "plus haut niveau"

Section 4

Géométrie pour la vision

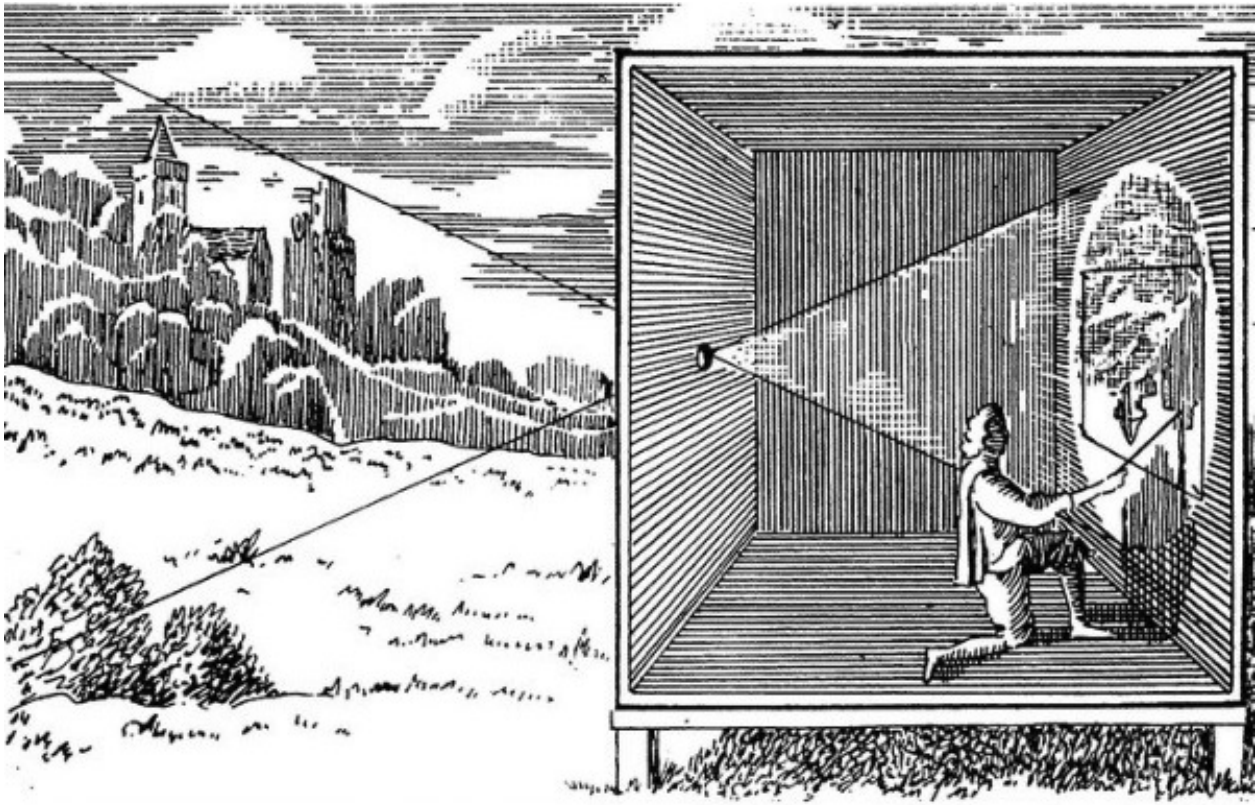
Modèle paramétrique: Exemple de la balance 1/2

- On souhaite mesurer la masse m d'un objet
- La force générée est le poids P, g est l'accélération à un endroit donné ($g \approx 9,8m/s^2$): $P = g.m$
- Longueur L d'une épreuve se déformant sous l'effet du poids: $L = L0 + K_1.P$
- Montage dynamomètre piézoélectrique fournit tension V dont la valeur varie linéairement avec la déformation: $V = V_i + K_2.L$
- Quantification à l'aide d'un Convertisseur Analogique Numérique (version simplifiée): $N = (V_{out} - V_0).K_3.V$
- Au final N est observé comme fonction paramétrique de m et des différents paramètres
  - Modèle direct: $N = f(m)$
  - Modèle inverse (le modèle direct est inversible, il n'y a pas de "perte" d'information): $m = g(N)$

Modèle paramétrique: Exemple de la balance 2/2

- Paramètres du modèle, par exemple 2 scalaires K et N_0 tq: $N = f(m) = K.m - N_0$ et $m = g(N) = (N + N_0)/K$
- Estimation en faisant 2 mesures de masses de réference connues
  - résolution d'un système de 2 équations à 2 inconnues
- Estimation en faisant plus de mesures
  - Moindres carrées permet d'estimer mieux les paramètres si les masses de réference et les valeurs lues sont imprécises:
  - Méthodes directes (inversion de matrice, SVD...)
  - Méthode itérative (Gauss-Newton, Descente de gradient...)
  - Méthodes "Robustes" aux données erronées (M-estimateurs, RANSAC...)
- Impossible de discerner dans K les valeurs de g, K_1, K_2 et K_3 car ils sont observés de manière couplée.
- Modélisation et étalonnage de toute la chaine d'acquisition (environnement, capteur, mise en forme du signal, quantification....)

Les caméras



9

- Taille du trou d'épingle (pinhole): Image floue ou sombre
- Projection perspective
- Centre de projection=centre optique

[9] http://www.albertdemun.fr/college/art2mun/histoire-des-arts-4eme-la-physique-et-lart-la-camera-obscura

Les caméras

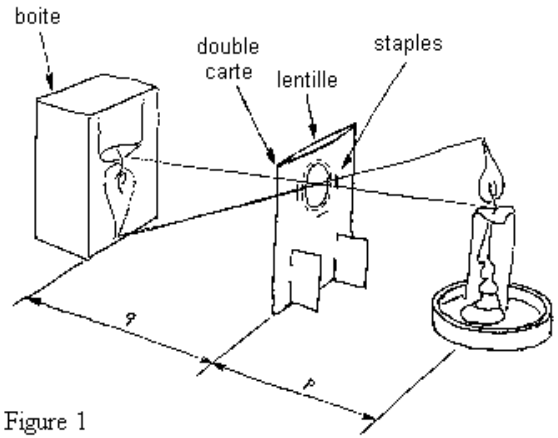


Figure 1
Banc d'optique pour la formation des images étudiées.
p = distance lentille objet; q = distance lentille écran

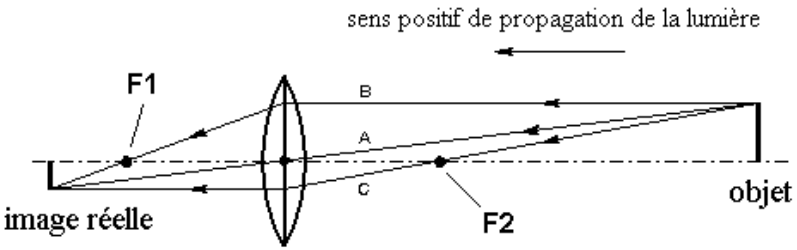


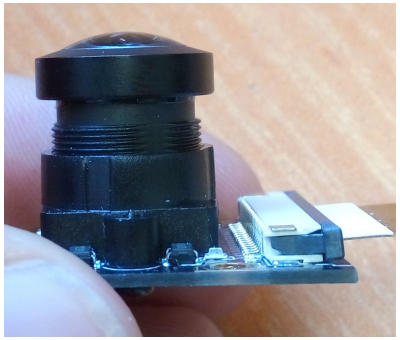
Figure 2 - Formation de l'image par une lentille convergente.

[10]

- En pratique, utilisation de systèmes dioptriques (assemblages de lentilles)
  - Permet de capter plus de lumière
  - Image nette pour un intervalle de distance (profondeur de champ)
  - Distorsions optiques

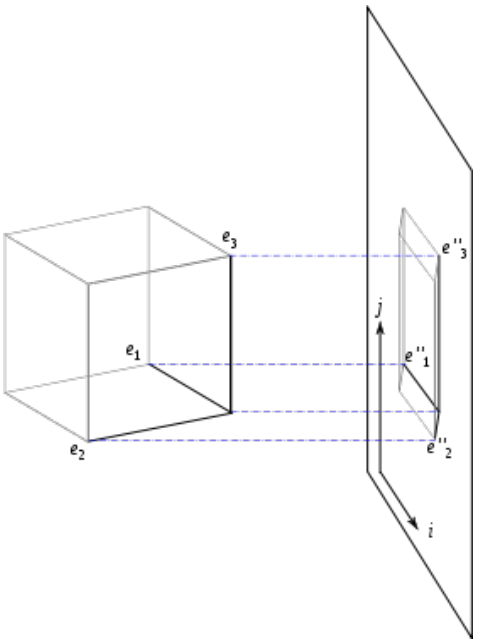
[10] http://www.funsci.com/fun3_fr/lentilles/lentilles.htm

Les caméras



- Optiques fisheye jusqu'à 240 x360 degrés de champ de vision
- Très courte distance focale
- Attention aux abhérations chromatiques
- Projection sur un plan?

Les caméras

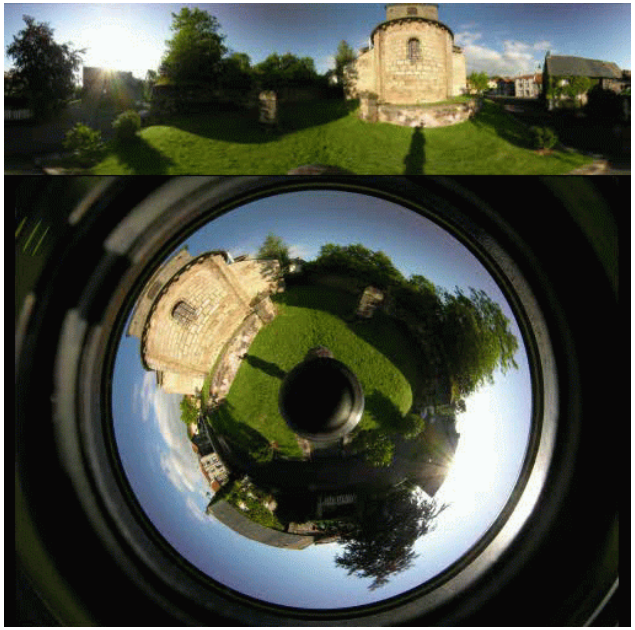


[11]

- Très longue distance focale (Centre optique à l'infini)
- Perçoit des rayons perpendiculaires au plan image
- Taille des objets perçus indépendante de la distance
- Microscope
- Lunette astronomique

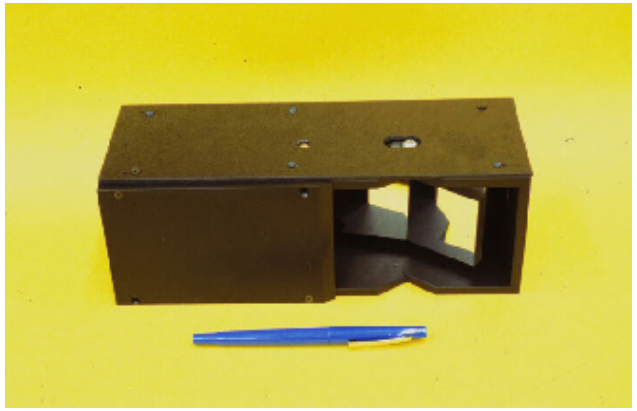
[11] https://fr.wikipedia.org/wiki/Projection_orthogonale

Les caméras



- Systèmes catadioptriques (lentilles + mirroir(s))
  - Augmentation du champ de vision

Les caméras

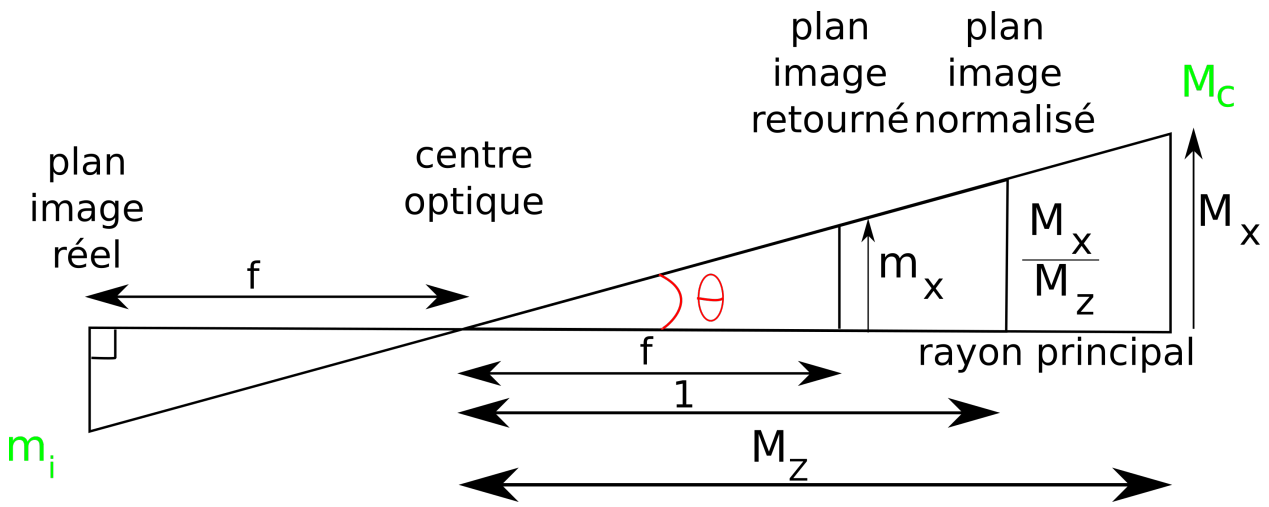


- Systèmes catadioptriques (lentilles + mirroir(s))
  - Stéreovision avec une seule caméra observant différents miroirs
  - Pas de problème de synchronisation

Les caméras numériques

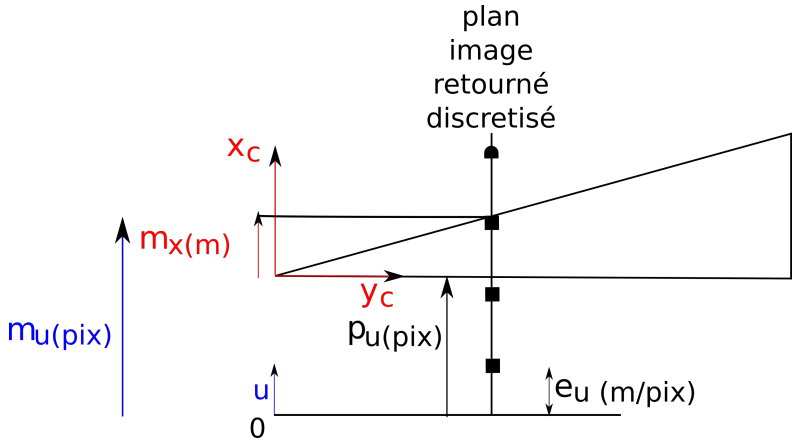
- Discrétisation du plan image en pixel
- Résolution en pixel
- Fréquence d'acquisition (FPS)
- Bus de communication
- Type d'obturateur:
  - Global (global shutter)
  - Tournant (rolling shutter), webcam et obturateur mécanique à double rideaux
- Dynamique: quantification de la quantité de lumière reçue pendant la durée d'intégration sur un nombre de bits donné
  - réponse type linéaire
  - réponse type log
  - High Dynamic Range:
    - capteur dédié
    - caméras DVS (120dB)
    - intégration de plusieurs images avec des expositions différentes (enfuse)

Modèle paramétrique: Thalès ou la règle de 3



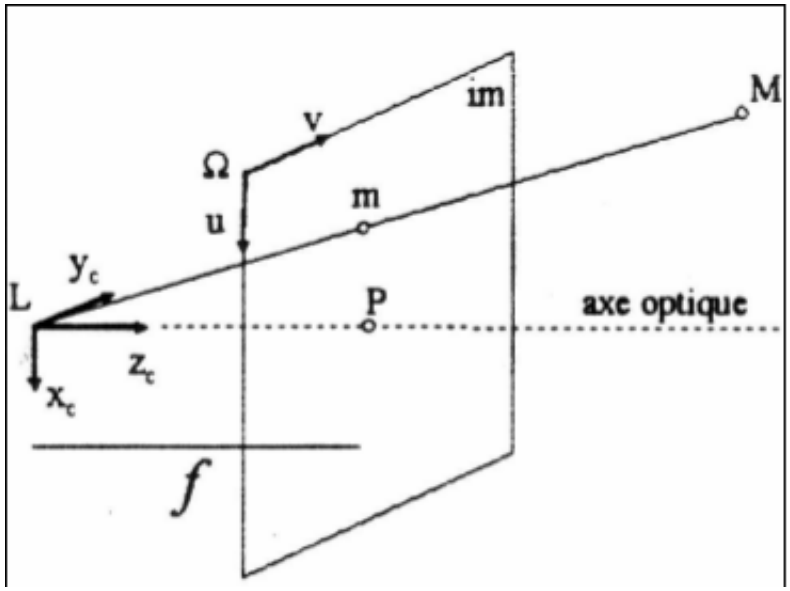
- Le rayon principal (axe optique) est l'orthogonal au plan image passant par le centre optique
- D'après Thales: $m_x = f.M_x/M_z$ (tout en unité métrique)
- La caméra est un appareil de mesure d'angle: θ
- Même raisonnement sur les 2 axes du plan image

Modèle paramétrique: Discrétisation



- Discrétisation du plan image: m_u position en pixel de la projection dans l'image
  $m_u = (m_x/e_u) + p_u = f.e_u^{-1}.M_x/M_z + p_u = \alpha_u.M_x/M_z + p_u$
- $\alpha_u = f.e_u^{-1}$ focale pixellique
- p_u point principal

Modèle paramétrique: La matrice de projection



Généralisation aux axes u et v et écriture sous forme matricielle. L'image $m_i = (m_w.m_u \ m_w.m_v \ m_w)^T$ d'un point $M = (M_{x_c} \ M_{y_c} \ M_{z_c} \ 1)^T$ exprimé dans le repère caméra est obtenue:

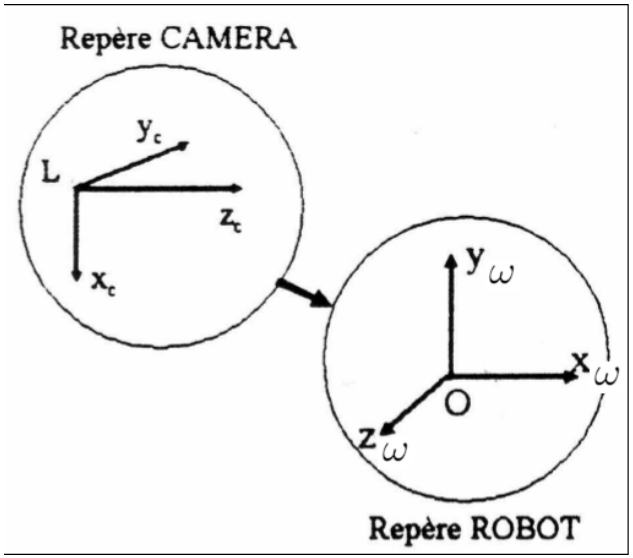
$$\begin{pmatrix} m_w.m_u \\ m_w.m_v \\ m_w \end{pmatrix} = \begin{pmatrix} \alpha_u \ 0 \ p_u \ 0 \\ 0 \ \alpha_v \ p_v \ 0 \\ 0 \ 0 \ 1 \ 0 \end{pmatrix} . \begin{pmatrix} M_{x_c} \\ M_{y_c} \\ M_{z_c} \\ 1 \end{pmatrix}$$

$${}^i\tilde{C}_C = \begin{bmatrix} \alpha_u \ 0 \ p_u \ 0 \\ 0 \ \alpha_v \ p_v \ 0 \\ 0 \ 0 \ 1 \ 0 \end{bmatrix} \ ; \ m_i = {}^i\tilde{C}_c \ M_c$$

Modèle paramétrique: Notations pour les changements de repères

- Choix des notations de: `https://www.gdr-robotique.org/cours_de_robotique/online/Khalil-Dombre_Modelisation/Khalil-Dombre_Modelisation.pdf`
- Matrice de rotation dont les colonnes sont les axes du repère i exprimés dans le repère j: ${}^{i}R_{j}$
- Vecteur translation exprimant l'origine du repère j dans le repère i: ${}^{i}t_{j}$
- Matrice de changement de repère: ${}^{i}M_{j} = \begin{bmatrix} {}^{i}R_{j} & {}^{i}t_{j} \\ 0 & 1 \end{bmatrix}$

Modèle paramétrique: Changements de repère Monde vers camera



Changement de repère entre repère de travail (ROBOT ou Monde) et repère caméra:

$$\begin{pmatrix} M_{x_c} \\ M_{y_c} \\ M_{z_c} \\ 1 \end{pmatrix} = \begin{pmatrix} {}^cR_\omega & {}^ct_\omega \\ 0_3^T & 1 \end{pmatrix} \cdot \begin{pmatrix} M_{x_\omega} \\ M_{y_\omega} \\ M_{z_\omega} \\ 1 \end{pmatrix}$$

avec ${}^ct_\omega = -{}^cR_\omega . \begin{pmatrix} L_{x_\omega} \\ L_{y_\omega} \\ L_{z_\omega} \end{pmatrix}$

Position du centre optique dans le repère ROBOT:

$\begin{pmatrix} L_{x_\omega} & L_{y_\omega} & L_{z_\omega} \end{pmatrix}^T$

Modèle paramétrique: La matrice caméra

Combinaison des matrices de projection et de changement de repère:
L'image $m_i = (m_w.m_u \ m_w.m_v \ m_w)^T$ d'un point $M = (M_{x_\omega} \ M_{y_\omega} \ M_{z_\omega} \ 1)^T$ exprimé dans le repère ROBOT est obtenue à partir de la matrice caméra: ${}^i\tilde{C}_\omega = {}^i\tilde{C}_c \begin{pmatrix} {}^cR_\omega & {}^ct_\omega \\ 0_3^T & 1 \end{pmatrix}$.

$$ {}^i\tilde{C}_\omega = \begin{bmatrix} \alpha_u \ r_1^T + p_u \ r_3^T & \alpha_u \ t_x + p_u \ t_z \\ \alpha_v \ r_2^T + p_v \ r_3^T & \alpha_v \ t_y + p_v \ t_z \\ r_3^T & t_z \end{bmatrix} \ ; \ m_i = {}^i\tilde{C}_\omega \ M_\omega $$

- Paramètres intrinsèques: α_u, α_v, p_u, p_v (4ddl)
- Paramètres extrinsèques: ${}^cR_\omega$ (3 ddl), ${}^ct_\omega$ (3ddl)

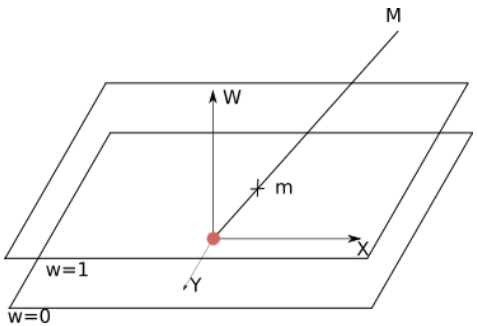
Paramétrisation des rotations

- Matrice 3x3 (avec 3 contraintes de norme et 3 contraintes d'orthogonalité)
- Euler (composition de 3 rotations successives)
- Rodrigues (axe de rotation de norme 1 multiplié par angle en radian)
- Quaternion unitaire (vecteur 4 de norme 1)

- En général non commutatif sauf pour les rotations infinitésimales

Espaces projectifs

Définition: $\mathbb{P}^n = \mathbb{R}^{n+1} - \{0\}$



- L'espace $\mathbb{P}^2$ contient les points $M = (\ M_x\ M_y\ M_w\ )^T$:
  - les points finis ($M_w \neq 0$) dont les coordonnées sont $m = (\ \frac{M_x}{M_w}\ \frac{M_y}{M_w}\ )^T$ dans $\mathbb{R}^2$
  - les points à l'infini ($M_w = 0$) dont la direction est $m = (\ M_x\ M_y\ )^T$ dans $\mathbb{R}^2$.
- Permet d'appliquer des rotations et translations (et échelle et projection...) par produit matriciel
- Une Homographie est une application projective qui
  - conserve les droites
  - conserve les bi-rapports

La matrice caméra est une application projective

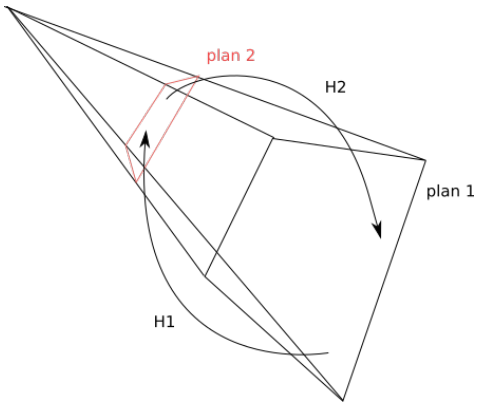
La matrice ${}^{i}\tilde{C}_{\omega}$ est une homographie de $\mathbb{P}^3$ vers $\mathbb{P}^2$

$$m_i = \begin{pmatrix} m_w.m_u \\ m_w.m_v \\ m_w \end{pmatrix} = {}^{i}\tilde{C}_{\omega}.\begin{pmatrix} M_{x_\omega} \\ M_{y_\omega} \\ M_{z_\omega} \\ 1 \end{pmatrix}$$

- Elle permet de projeter des points de $\mathbb{R}^3$
- Elle permet de projeter des directions de $\mathbb{R}^3$, dont l'image s'appelle un point de fuite
- L'image d'un point (fini ou à l'infini) est un point (fini ou à l'infini)

Homographie de $\mathbb{P}^2$ vers $\mathbb{P}^2$

On considère souvent les homographies de $\mathbb{P}^2$ vers $\mathbb{P}^2$



- Ces homographies sont des matrices 3x3
- Si H_1 est inversible alors $H_2 = H_1^{-1}$
- Les homographies peuvent être composées par produit matriciel
- Cas particulier de matrice caméra qui s'applique à un plan unique
- Choix arbitraire du plan $z = 0$ possible en appliquant une rotation au repère de travail (éliminer la 3° colonne de ${}^i\tilde{C}_\omega$)

Distorsions optiques: La caméra réelle n'est pas un trou d'épingle!

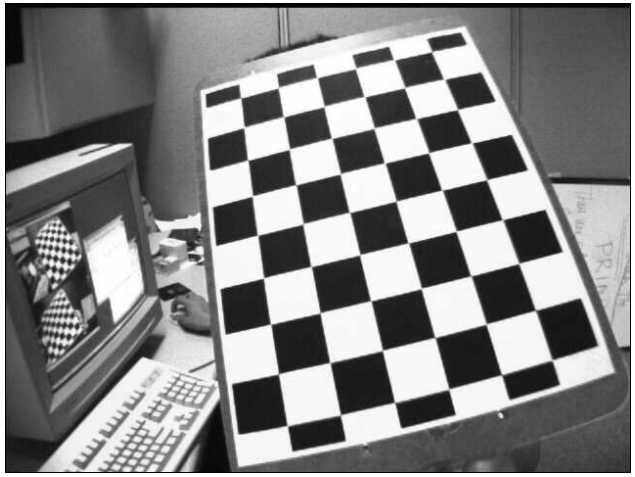
Paramètres intrinsèques supplémentaires pour modéliser les distorsions:

- Radiales (Coussinet ou Barillet)
  - Fonction polynomiale (Taylor)
  - Appliquées dans le plan image normalisé (indépendantes de la taille des pixels)
- Tangentielles

Choisir le nombre de degrés de liberté dans le modèle en fonction de la caméra et de l'application!

Rectification des distorsions optiques

Rectifier = obtenir une image qui aurait été acquise sans distorsion et pour laquelle la matrice caméra est utilisable



Originale

Rectifiée

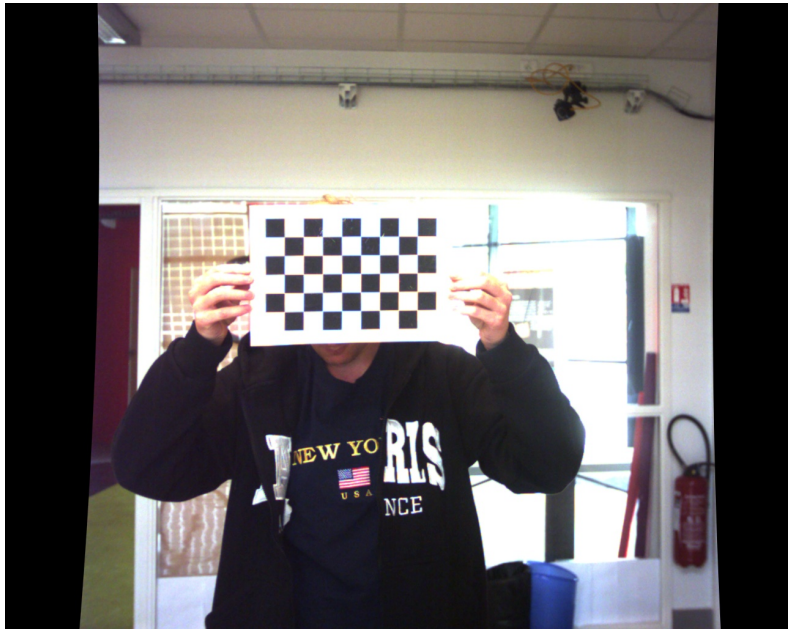
Dans l'image rectifiée, les images de droites sont droites!

Estimation des paramètres du modèle

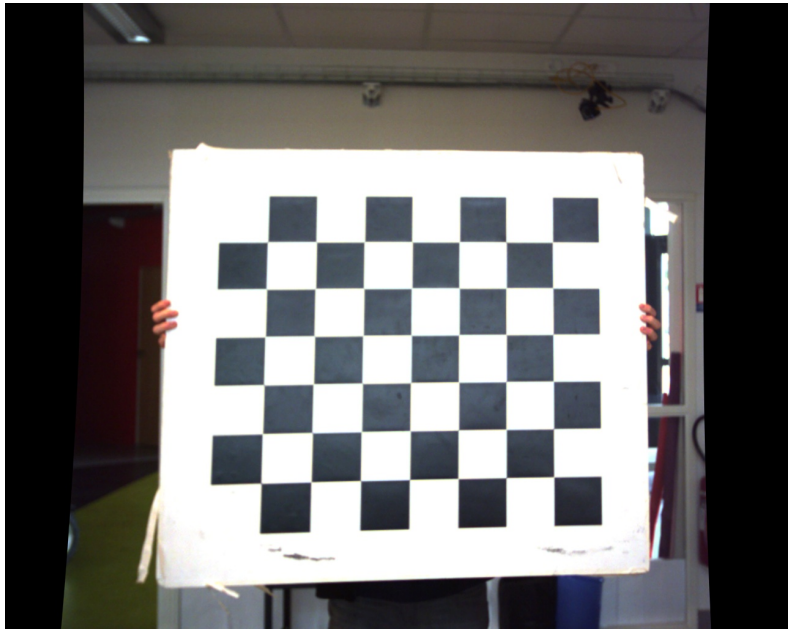
- Observation d'une mire connue:
  - 3D à une pose
  - Plus généralement 2D à différentes poses
- Détection subpixellique des coins
- Minimisation de l'erreur de reprojection entre les points observés et les points prédits par le modèle paramétrique: moindres carrés
- Grand nombre d'inconnues (+6 extrinsèques par pose de mire)

Mires de différentes tailles

Permet d'échantillonner pour différentes distances: le volume de travail



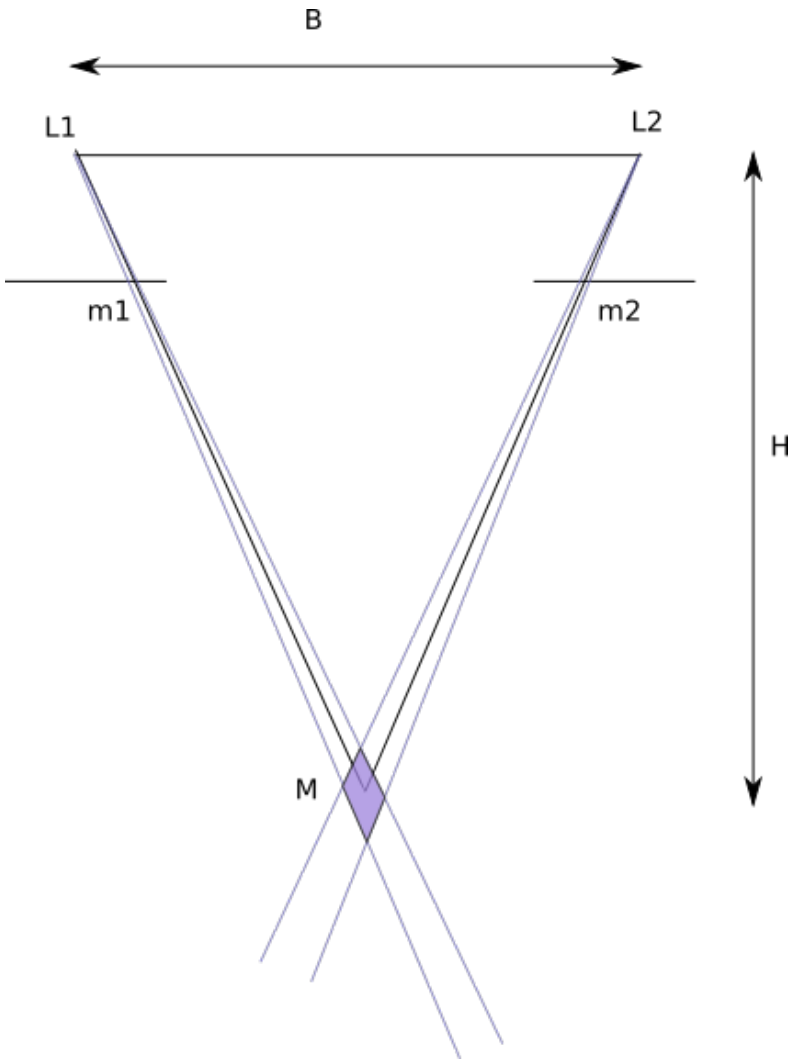
Petite mire pour acquérir des points proches



Grande mire pour acquérir des points lointains

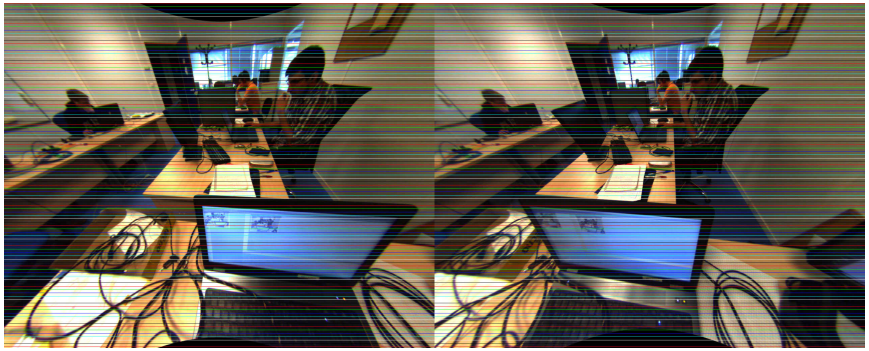
Principe de la Stéréovision

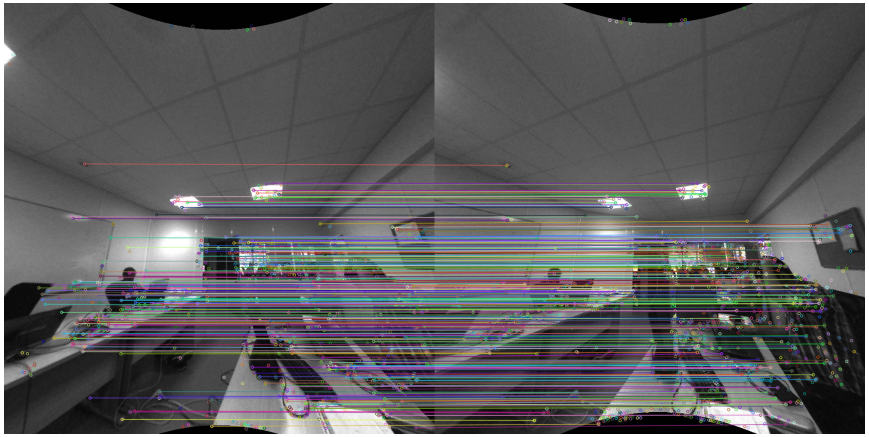
- Lever l'indétermination de distance grâce à plusieurs observations (par triangulation)
- Écart intraoculaire (Baseline)
- Rapport B/h
- Plus de 2 caméras pour:
  - Augmenter le champ de vision stéréo
  - Augmenter la précision
  - Augmenter la robustesse



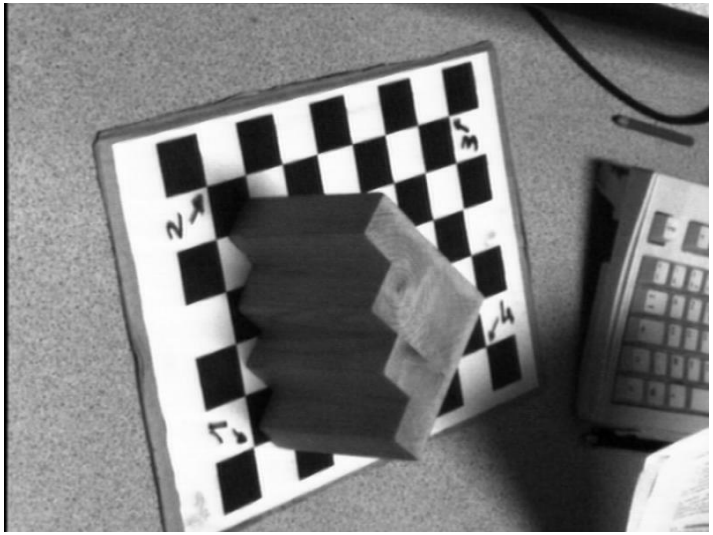
Rectification épipolaire

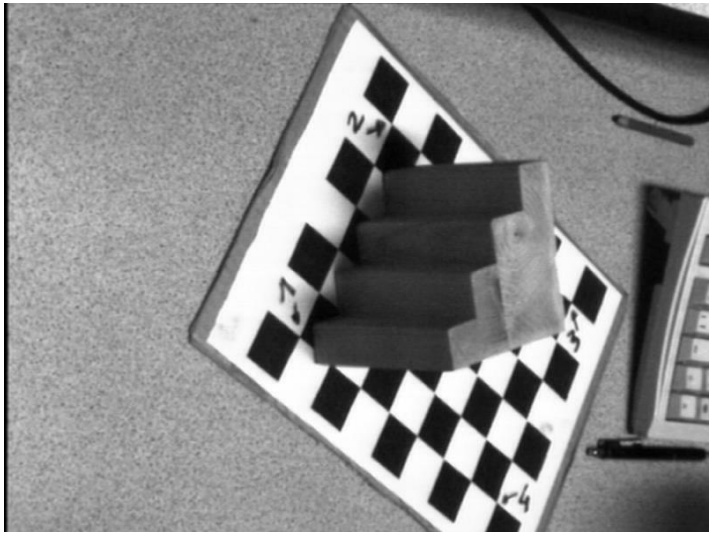
- Reprojection de la paire d'images sur un plan commun (Homographie)
- Epipoles à l'infini
- Rectification = obtenir une paire d'images sur laquelle les pixels homologues sont sur les mêmes lignes
- Appariement automatique (épars ou dense)

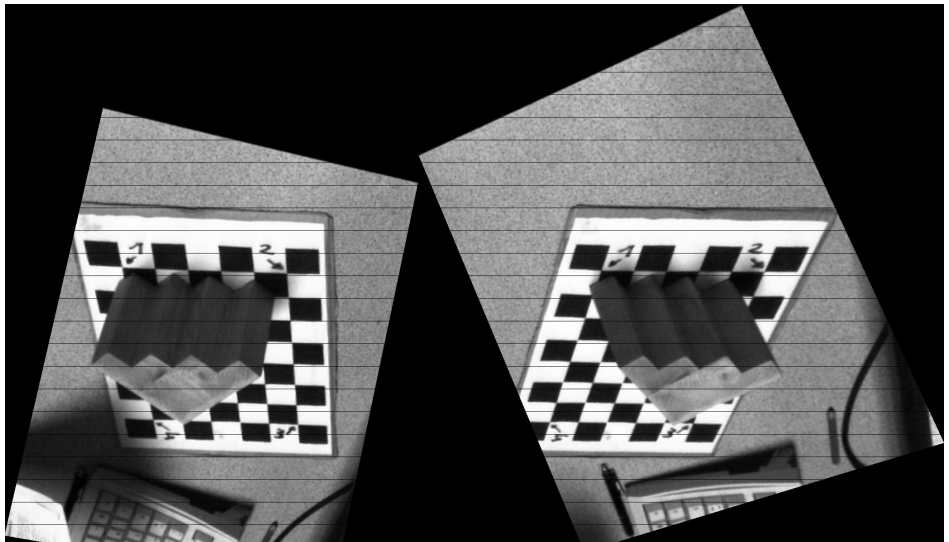




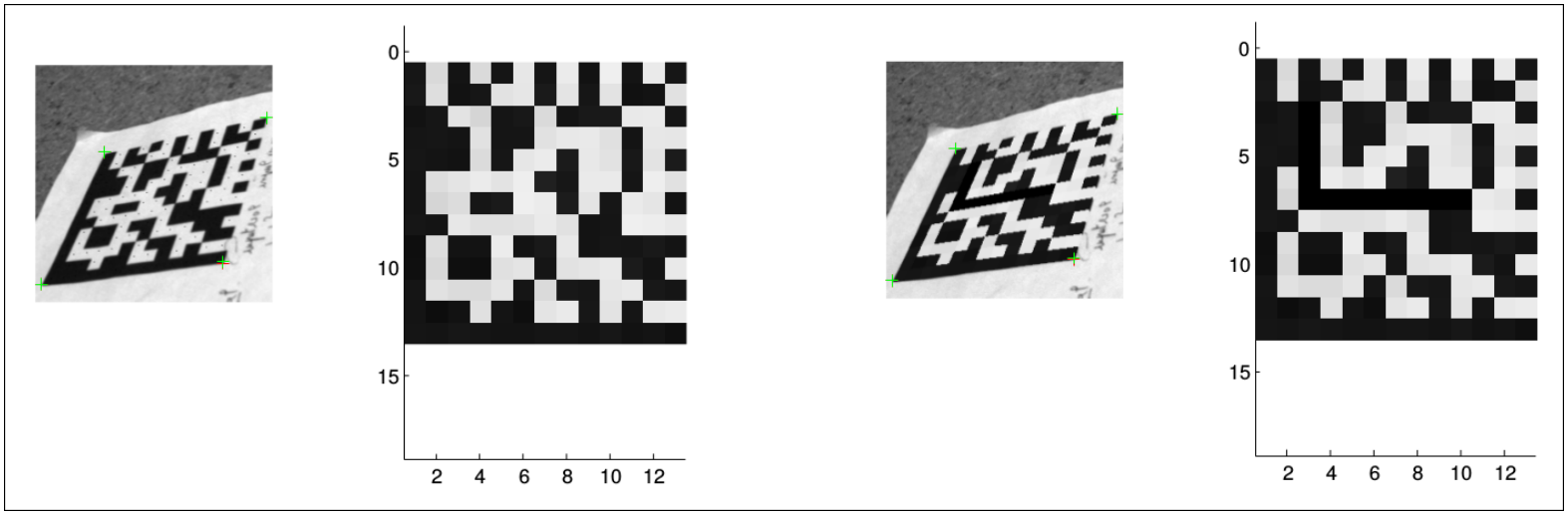
Rectification épipolaire







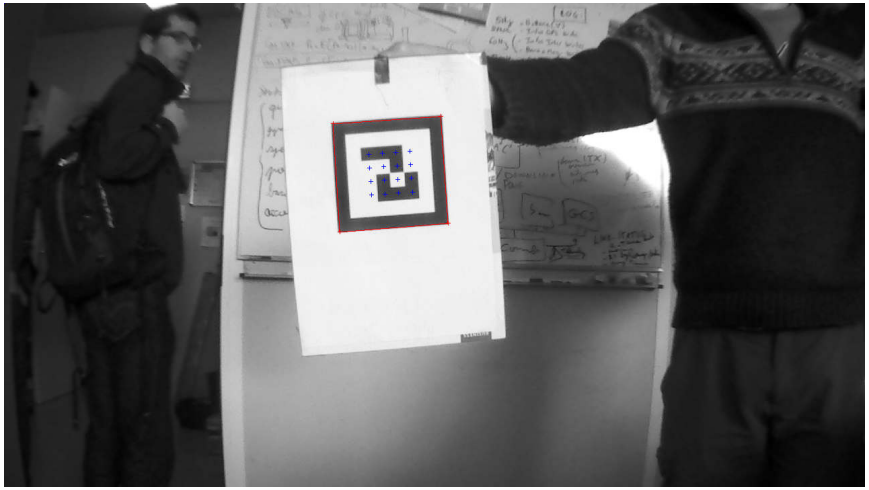
Homographies pour la rectification d'image

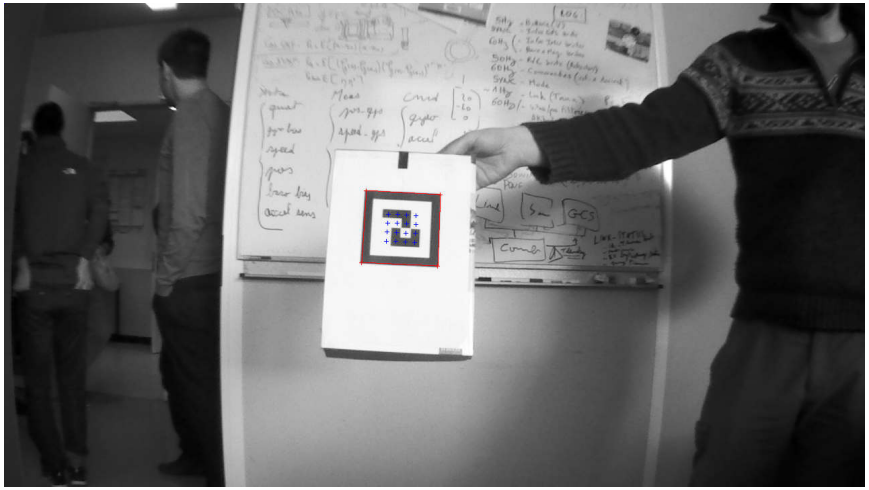


- Reprojection de l'image sur le plan du tag
- Modification de l'image du tag
- Reprojection du tag sur le plan de l'image

Homographies pour la localisation de TAG

Algorithme Perspective-N-Points





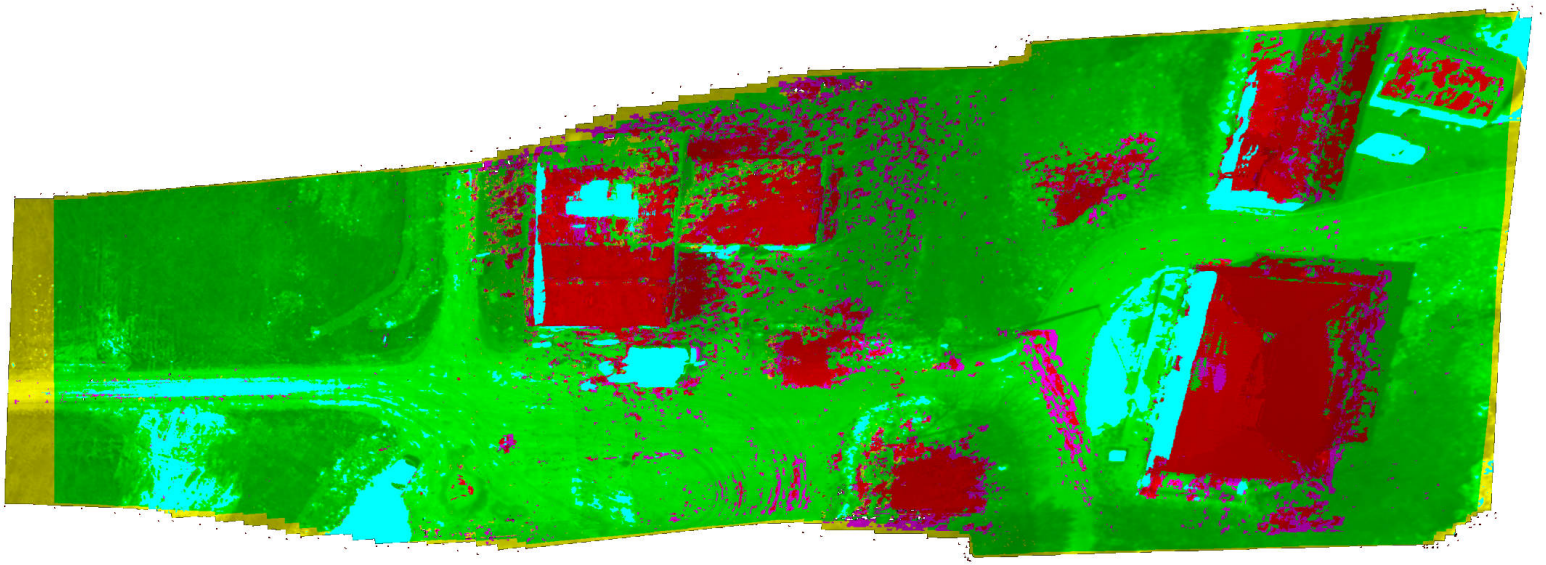
- Détection automatique des coins du tag
- Vérification du code par reprojection
- Calcul des 6ddl de la pose à partir de l'observation des 4 coins

Mosaiquage d'images par Homographies

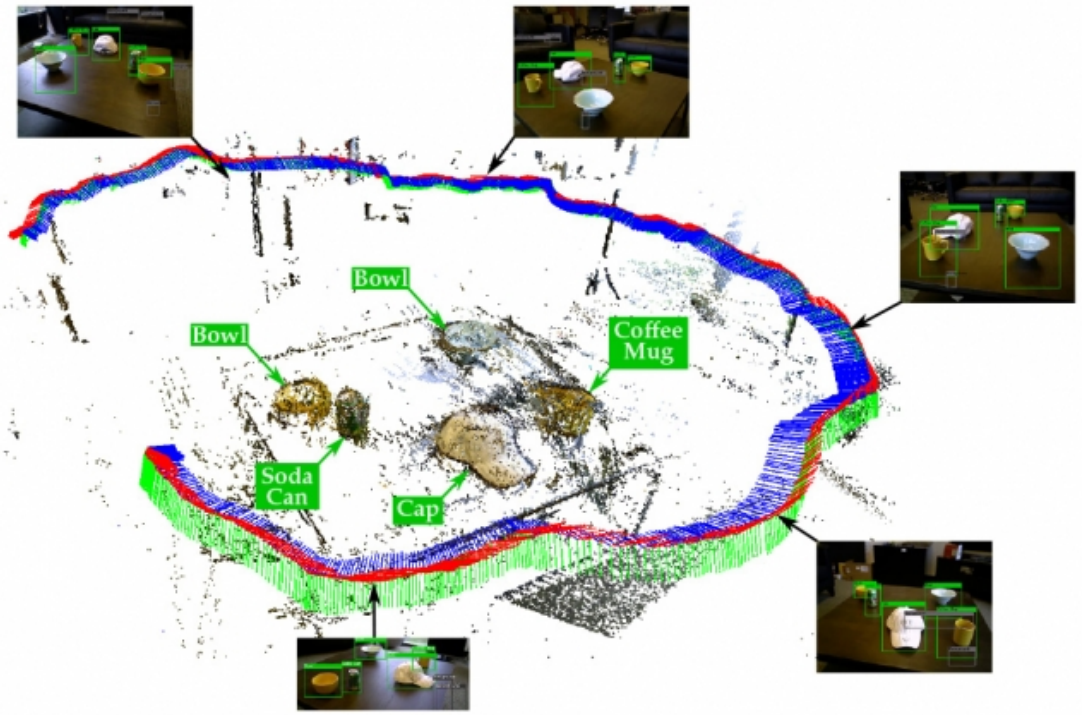
- Intégratrion des images sur un plan lors du déplacement/rotation de la caméra.
- Valide uniquement pour un plan, modèle 2D de l'environnement.

Mosaiquage d'images par Homographies

- Exploitable pour déterminer l'élévation.
- Sensible aux sol non plans.



Localisation et cartographie simultanée (SLAM)



[12]

- Détermination de la géométrie de la scène et de la pose de la caméra

[12] http://news.mit.edu/2015/object-recognition-robots-0724

Section 5

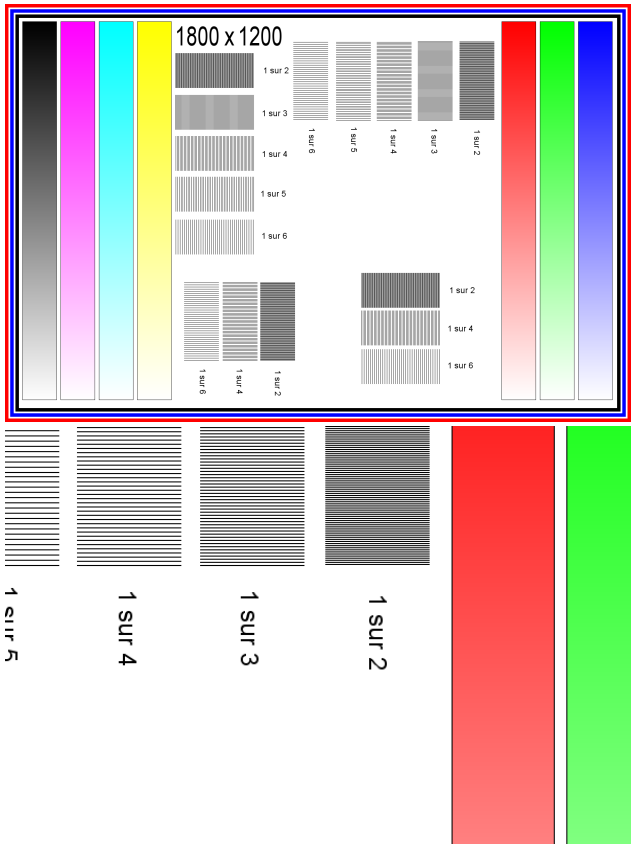
Considérations annexes

Quelques faits généraux sur les traitements

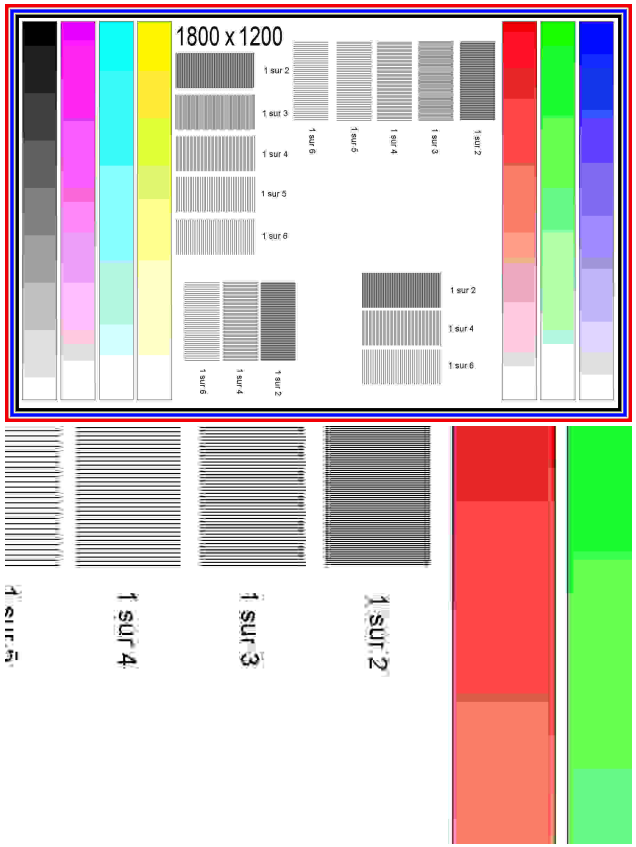
- Traitement d'image "lourd" en nombre d'opérations
- Attention aux fausses impressions:
  - Ce n'est pas parce que votre système visuel sait faire qu'un système automatique saura et inversement...
- Attentions aux paramètres à régler "à la main": préférer des méthodes adaptatives qui calculent elles mêmes, par exemple seuil pour la méthode d'Otsu

Importance de traiter des données brutes

Eviter la compression avec perte (JPEG)



Originale



Comprimée JPEG

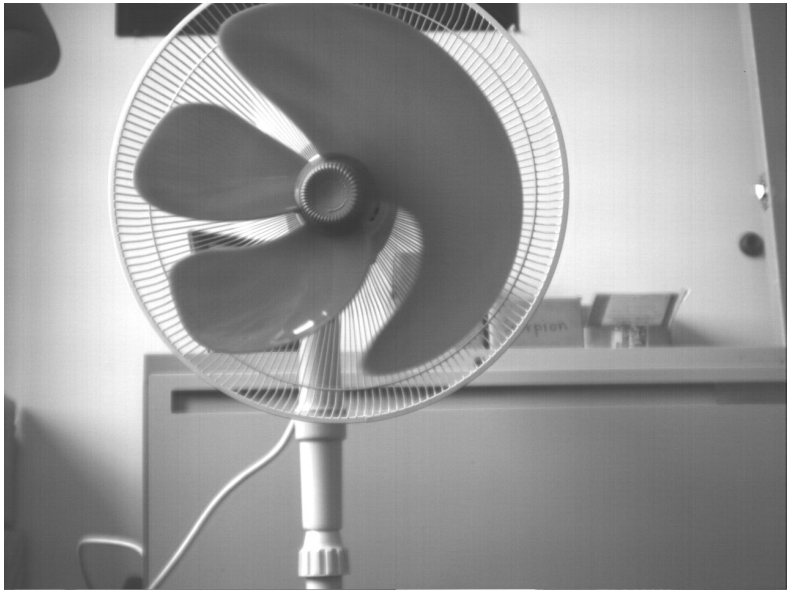
Choix de la caméra

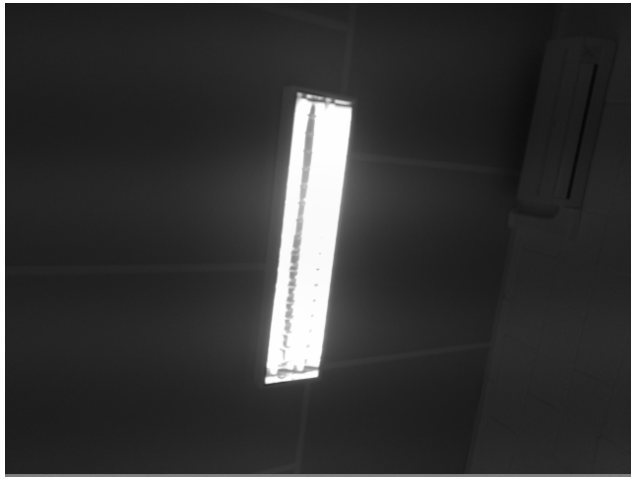
Eviter l'autofocus

Change les paramètres intrinsèques en permanence

Choix de la caméra

Eviter le Rolling shutter





Bibliothèque

- Ne pas réinventer la roue... sauf si il existe des contraintes (matérielles, temps réel...)
- OpenCV
  - Binding pour protoypage en Python
  - Attention aux versions (actuelle 12/2023= 4.8.1)
- Numpy (ou Eigen en C++)

Architecture logicielle

- La vision est une fonctionnalité d'un robot parmi plusieurs : motricité, communication, interaction, tâche principale, gestion de l'énergie, décision...
- Architecture modulaire
- MiddleWare ROS (Robot Operating System)
  - Modules (Noeuds) et Serveur
  - Publisher/Suscriber et Services