

Vision par ordinateur - Reconnaissance Visuelle

Cours 1 : Introduction.

Céline Hudelot - Mention IA - CentraleSupélec

2022-2023

Avant propos

Petit sondage

La vision par ordinateur pour vous ? Pouvez-vous citer des applications connues de vision par ordinateur ?



Plan et sujets abordés

- Vision humaine / Vision Artificielle - Sciences cognitives
- Formation de l'image - Géométrie de l'image.
- Introduction au traitement de l'image :
 - ▶ Filtrage.
 - ▶ Contours.
 - ▶ Extraction de primitives.
- Extraction et mise en correspondance de caractéristiques.
- Segmentation.
- Reconnaissance.
- Mouvement.
- Calibration - Reconstruction 3D.
- (Cas des données multimédias.)

Organisation

Déroulement des séances

- ~ 1h30 de cours
- ~ 1h30 de mise en pratique (sous forme de notebooks, majoritairement avec OpenCV)
- Sauf pour les deux premières séances : 3H de cours puis 3H de TDs.

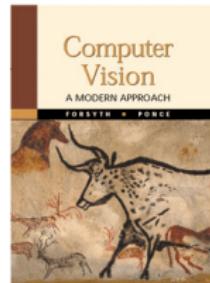
Evaluation

- Un TP noté lors de la dernière séance (OU)
- Un projet libre pouvant être mutualisé avec d'autres cours (notamment cours de Deep Learning).
 - ▶ rapport sous la forme d'un article ainsi que le code.
 - ▶ individuel, ou groupe de deux à trois étudiants maximum.

Ouvrages conseillés pour les fondements

- **Un ouvrage de référence**

Computer vision : a modern approach. Forsyth and Ponce



<http://luthuli.cs.uiuc.edu/~daf/book/book.html>

Quelques sites web :

- The computer vision homepage

<http://www.cs.cmu.edu/~cil/vision.html>

- Machine Vision Online

<http://www.machinevisiononline.org/> (pour le monde de l'industrie)

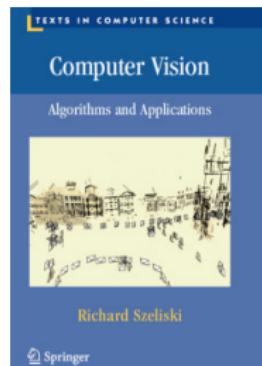
- Certains articles WP https://fr.wikipedia.org/wiki/Portail:Imagerie_num%C3%A9rique

Ouvrages conseillés

- Un ouvrage un peu plus récent, très axé sur les applications

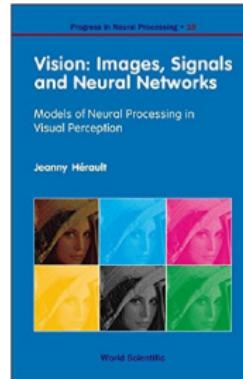
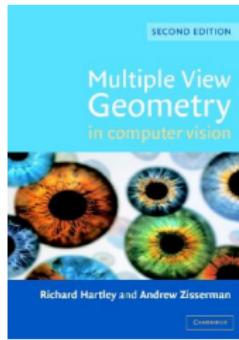
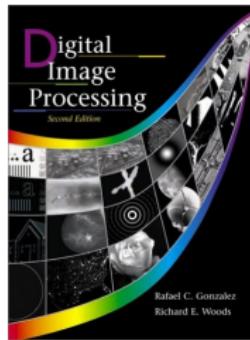
Computer vision : Algorithm and Applications. R.Szeliski (Microsoft Research - Interactive Visual Media Group) Séconde édition en date de février 2021

<http://szeliski.org/Book/>



Ouvrages conseillés

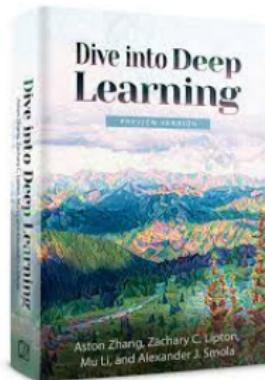
- **Multiple View Geometry in Computer Vision**, R. Hartley, A. Zisserman
 - ▶ <http://www.robots.ox.ac.uk/~vgg/hzbook/>
- **Digital Image Processing**, R. Gonzales and R. Woods
 - ▶ <http://www.imageprocessingplace.com/>
- **Vision : Images, Signals and Neural Networks : Models of Neural Processing in Visual Perception**, J. Hérault



Ouvrages conseillés

Pour le deep learning :

- Dive into Deep Learning, Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola
- <https://d2l.ai/index.html>



Plan

1 Organisation du cours

2 Introduction

- Définition

3 Vision humaine - Vision par ordinateur

- Vision humaine
- Vision par ordinateur

4 Qu'est ce qu'une image ?

- Vision par ordinateur : les difficultés

5 Applications de la vision par ordinateur

Qu'est ce que la vision par ordinateur ?

Définitions

- Traitement automatisé par un ordinateur des informations visuelles.
- Inférer des propriétés du monde réel (3D, 3D+t), l'**interpréter**, le **comprendre** à partir d'une ou plusieurs images.

Qu'est ce que la vision par ordinateur ?

Définitions

- **Trucco & Verri** : *computing properties of the 3D world from one or more digital images.*
- **Stockman & Shapiro** : *to make useful decisions about physical objects and scenes based on sensed images.*
- **Ballard & Brown** : *the construction of explicit, meaningful descriptions of physical objects from images.*
- **Forsyth & Ponce** : *extracting descriptions of the world from pictures or sequences of pictures.*
- **Szeliski** : *write computer programs that can interpret images.*

Qu'est ce que la vision par ordinateur ?

Définitions

- Computer Vision (CV) is a field of Artificial Intelligence (AI) that deals with computational methods to help computers understand and interpret the content of digital images and videos.
- Computer vision aims to make computers see and understand visual data input from cameras or sensors.

Qu'est ce que la vision par ordinateur ?

Vision par ordinateur vs Traitement de l'image

- Traitement d'images

- ▶ Production d'une nouvelle image
- ▶ Travail au niveau *signal*
- ▶ Peu de relations avec la réalité extérieure, pas d'interprétation

- Vision par ordinateur

- ▶ Emulation de la vision humaine ou robotique
- ▶ Prise en compte de l'environnement, la réalité du monde extérieur
- ▶ L'objectif est d'extraire de l'information pour une prise de décision
- ▶ L'image est une source d'information (parmi d'autres)

Qu'est ce que la vision par ordinateur ?

Lien avec d'autres disciplines.

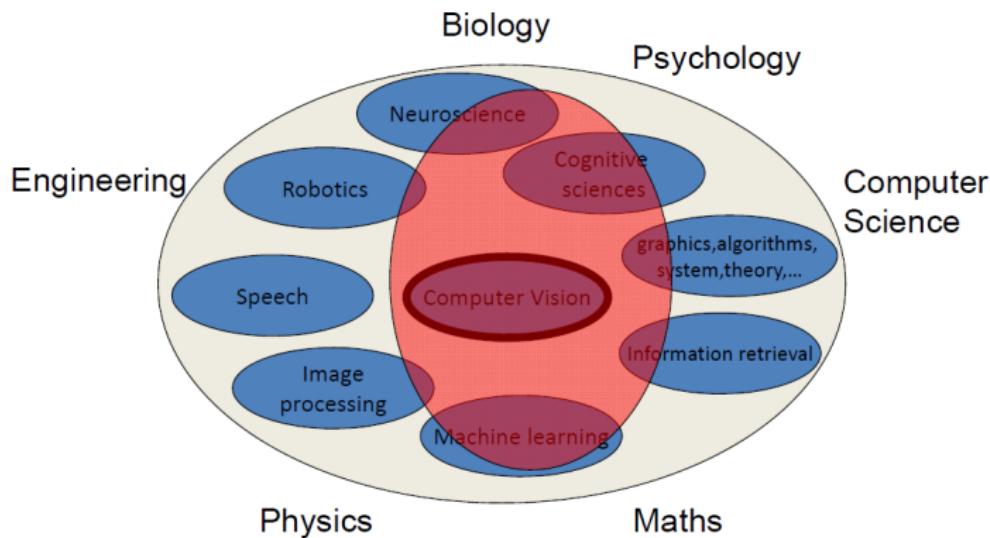


FIGURE – Source : Fei-Fei Li, Computer Vision Course, Stanford

But de la vision par ordinateur

Reduire le fossé entre les pixels et leur *signification*.



La Gare Montparnasse, 1895

What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
8	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

FIGURE – Source : Fei-Fei Li, Computer Vision Course, Stanford

La caméra est un dispositif de mesure (un capteur), pas un système de vision !

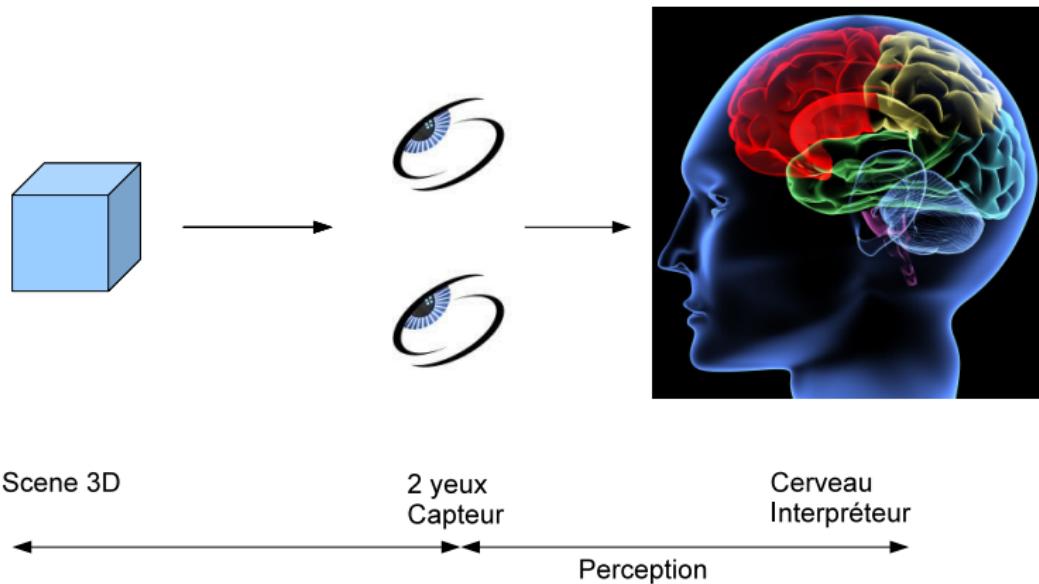
Plan

- 1 Organisation du cours
- 2 Introduction
 - Définition
- 3 Vision humaine - Vision par ordinateur
 - Vision humaine
 - Vision par ordinateur
- 4 Qu'est ce qu'une image?
 - Vision par ordinateur : les difficultés
- 5 Applications de la vision par ordinateur

Vision humaine

Définition

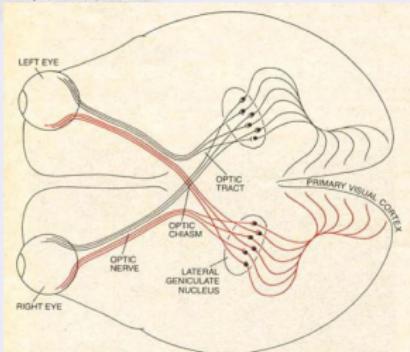
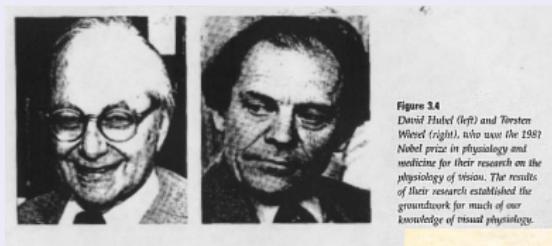
Perception par l'organe de la vue



Un pas de la science

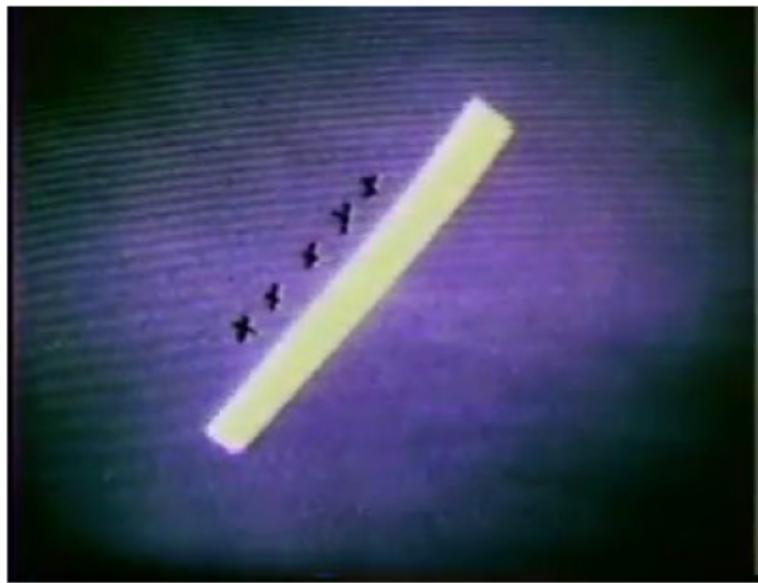
1981 : prix Nobel de Médecine

- Torsten Wiesel et David Hunter Hubel reçoivent le prix Nobel de médecine pour leurs découvertes concernant le traitement de l'information dans le système visuel (1959).



Un pas de la science

Hubel et Wiesel

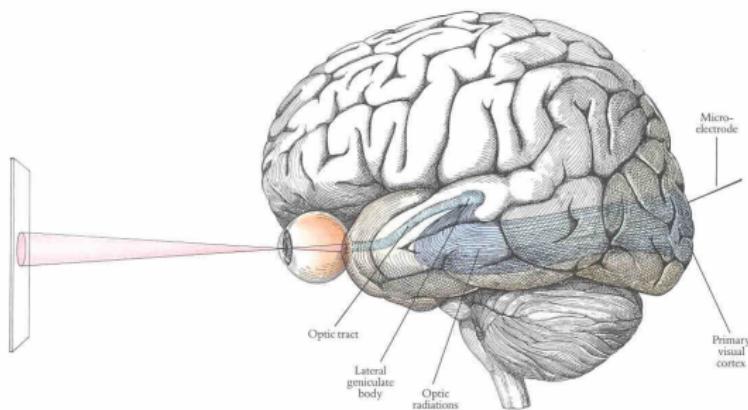


Hubel and Wiesel. JP 1959. Receptive Fields of Single Neurones in the Cat's Striate Cortex.

<https://www.youtube.com/watch?v=jw6nBWo21Zk>

Un pas de la science

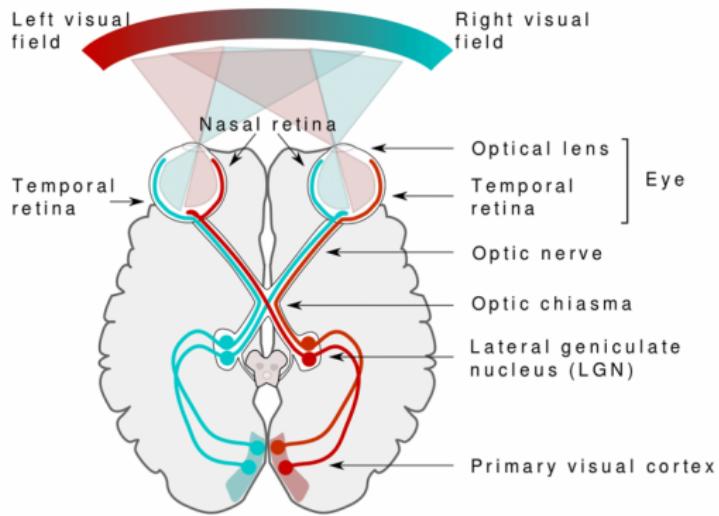
Hubel et Wiesel : système visuel des mammifères



Hubel 1995. Eye, Brain, and Vision.

Un pas de la science

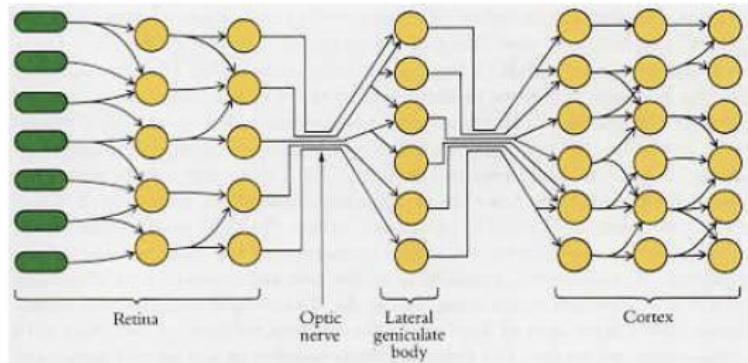
Hubel et Wiesel : visual pathway



Hubel 1995. Eye, Brain, and Vision.

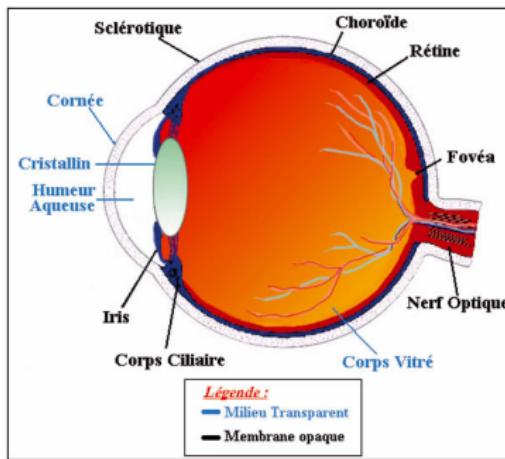
Un pas de la science

Hubel et Wiesel : représentation topographique



- En se déplaçant le long de la rétine, les points correspondants dans le cortex tracent un chemin continu.
- Chaque colonne représente un tableau bidimensionnel de cellules.

L'oeil humain : un capteur puissant et complexe



© Jacques Livage

- Fonction optique : focalisation d'un stimulus de couleur sur sa partie photosensible, la **rétine**.
- **Iris** et son ouverture centrale la **pupille** : fonctionne comme un diaphragme, i.e. contrôle la quantité de lumière qui pénètre dans l'œil.
- **Cristallin** : optique : lentille biconvexe molle.
- **Nerf optique** : transport de l'information.

L'oeil humain

Caractéristiques

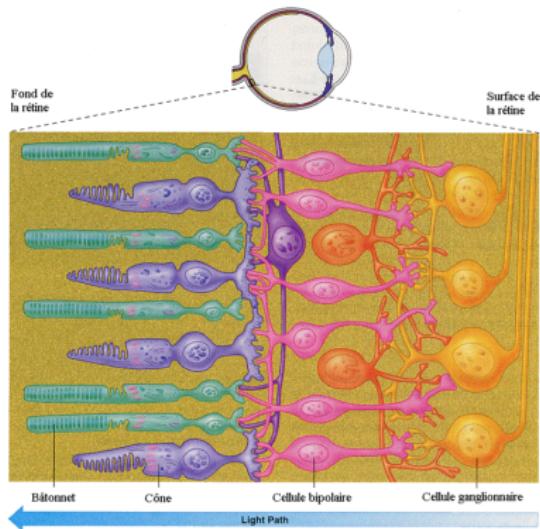
- Résolution angulaire : $0,017^\circ$ (un détail d'environ 1 mm pour un objet ou une image situé à 3 m de distance.)
- Résolution temporelle : de l'ordre de 20 images par seconde (non réglable).
- Résolution spectrale : de 400 nm à 700 nm.

Des caractéristiques qui peuvent varier avec l'âge et la personne.

Ordres de grandeur de la finesse des détails observables en fonction de la distance (diamètre d'un cheveu : 40 à 100 nm) :

Distance	10 cm	1 m	10 m	100 m
Taille de l'object en m	1.744E-05	1.744E-04	1.744E-03	1.744E-02
Distance	1 km			terre-lune
Taille de l'object en m	1.744E-01			6.71E+04

L'oeil humain : la rétine



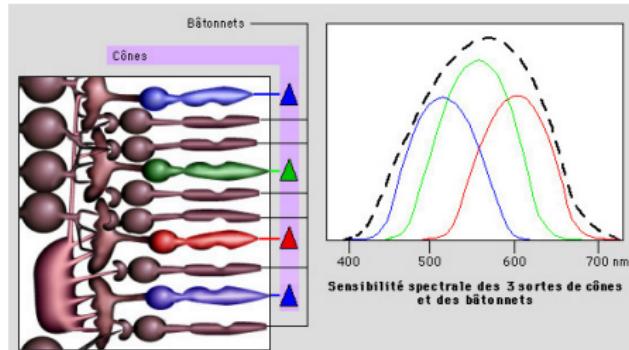
Composée de 3 couches

- Couche des photorécepteurs : cônes et bâtonnets.
- Couche granuleuse interne.
- Couche des ganglionnaires.

L'oeil humain : la rétine

Couche des photorécepteurs

- Cônes (6 à 7 millions)
 - ▶ Trois familles : courtes, moyennes et grandes longueurs d'onde
 - ▶ Vision photopique (diurne)
 - ▶ Couleur
 - ▶ Zone fovéale
- Bâtonnets (120 millions)
 - ▶ Tous identiques
 - ▶ Vision scotopique (nocturne)
 - ▶ Faible intensité, achromatique
 - ▶ Zone extra-foviale



La rétine

La rétine est composée de deux zones :

- Partie centrale : le fovea
 - ▶ Ne comporte que peu de bâtonnets.
 - ▶ Saisit les couleurs et les mouvements.
 - ▶ Zone d'acuité maximale de l'oeil.
- Reste de la rétine
 - ▶ Principalement formée de bâtonnets.
 - ▶ Permet la vision en faible luminosité.
 - ▶ Permet de saisir les nuances de gris.

La rétine

Couche granuleuse interne

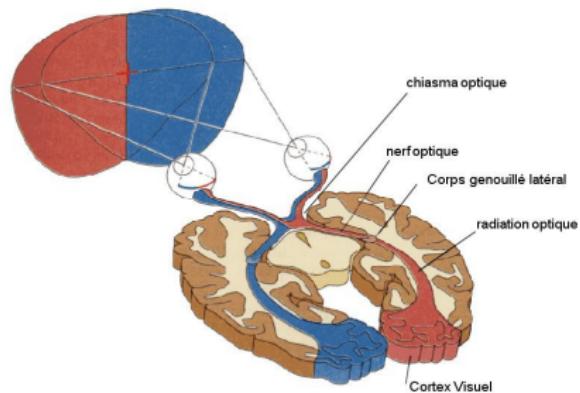
- Formée de cellules bipolaires et amacrines.
- Elles servent à grouper les messages (signal nerveux électrique émis lorsque la lumière les frappe) provenant des photorécepteurs.

Couche ganglionnaire

- Formée de cellules ganglionnaires (environ 1 million), prolongées par des fibres nerveuses.
- Ces fibres optiques naturelles se réunissent pour former un gros câble, le nerf optique qui transmet l'influx nerveux au cerveau.

Notre matière grise peut alors analyser les messages et composer les images ... et les **interpréter**.

TraITEMENT PAR LE CERVEAU



Le cortex visuel

- Occupe 15 % de la surface cérébrale.
- Centaines de millions de neurones avec des milliards de connexions.
- Composé d'une trentaine d'aires corticales qui travaillent de manière complémentaire pour former une image nette.
- Mécanismes pas entièrement compris à l'heure actuelle.

Traitement par le cerveau

Lobe occipital

V1 : aire primaire

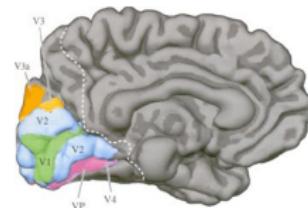
- 1ère analyse des informations : forme, couleur, mouvement.
- Distribue aux autres aires.

V2

- Trie les informations de V1.
- Traite les contours, textures, orientations et couleurs.

V3

- Analyse des formes en mouvement.
- Appréciation des distances.



V4

- Traitement des couleurs.
- Formes immobiles.

V5

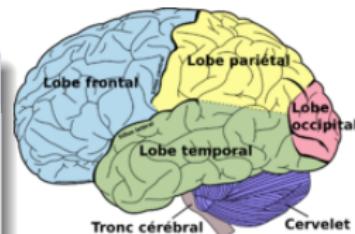
- Perception des mouvements.

TraITEMENT par le cerveau

L'information visuelle se sépare en deux :

Une partie se dirige vers le lobe temporal

- Identification des formes (contours et couleurs).
- En pratique, permet de reconnaître les objets et les personnes.



L'autre partie se dirige vers le lobe pariétal

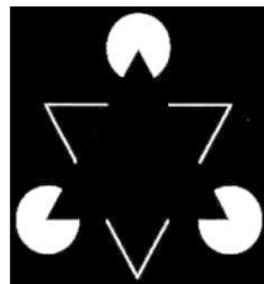
- Localisation spatiale des objets.
- Coordination entre la vision et la motricité.

Le reste du traitement est très complexe, implique la mémoire et permet de reconnaître les objets (les interpréter). Pour plus de détails, voir la page Wikipedia¹

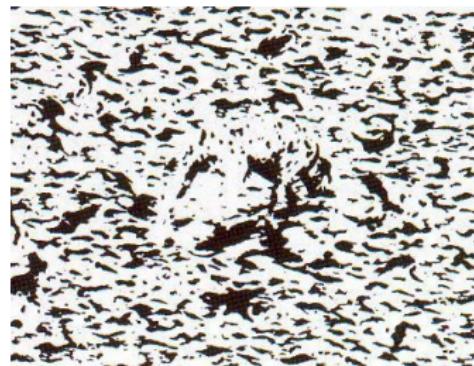
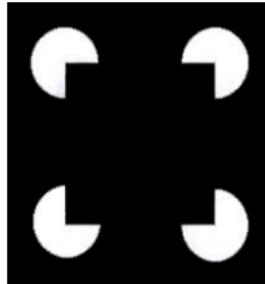
1. https://fr.wikipedia.org/wiki/Cortex_visuel

Une vision humaine non parfaite

Illusions d'optiques



Contours illusoires



Un dalmatien ? Le cerveau préfère une information structurée

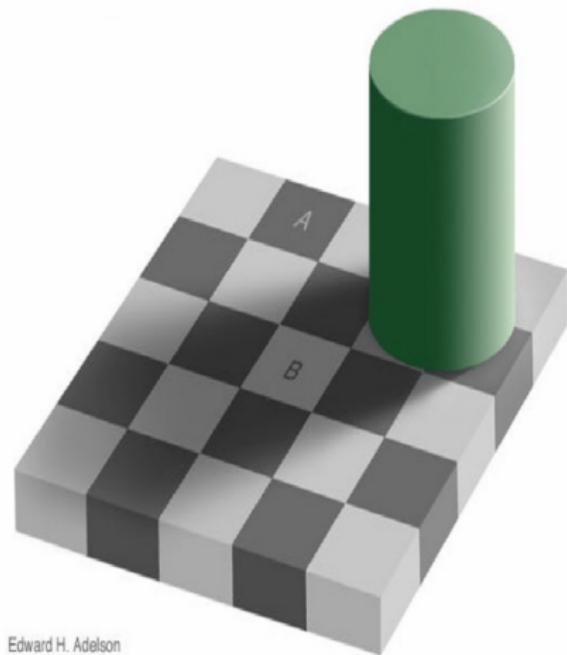
<http://www.michaelbach.de/ot/index.html>

<https://www.illusionsindex.org/illusions>

<http://www.ritsumei.ac.jp/~akitaoka/index-e.html>

Une vision humaine non parfaite

Ombrage : l'échiquier d'Adelson.

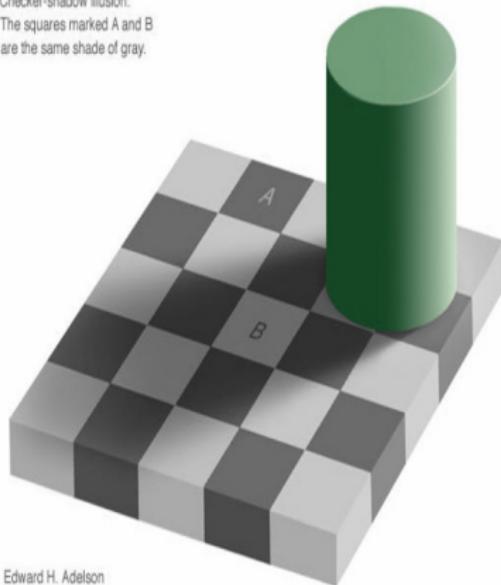


Edward H. Adelson

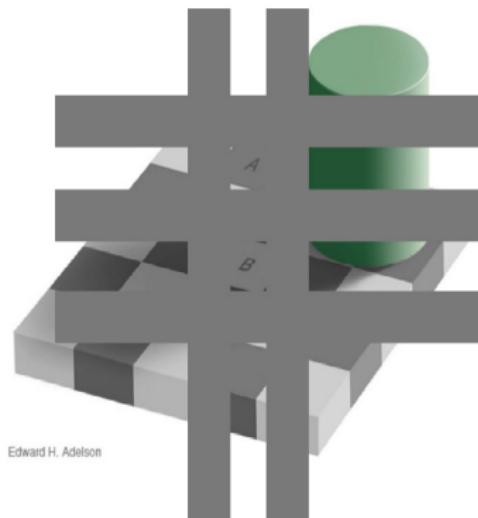
Une vision humaine non parfaite

Ombrage : l'échiquier d'Adelson.

Checker-shadow illusion:
The squares marked A and B
are the same shade of gray.



Edward H. Adelson



Edward H. Adelson

Une vision humaine non parfaite

Ombrage : l'échiquier d'Adelson.

- Les carrés ont la même couleur.
- Notre cerveau se force à imaginer les carrés tels qu'ils doivent être : un sombre et l'autre clair.
- Importance de la perception dans la manière dont nous interprétons le monde.

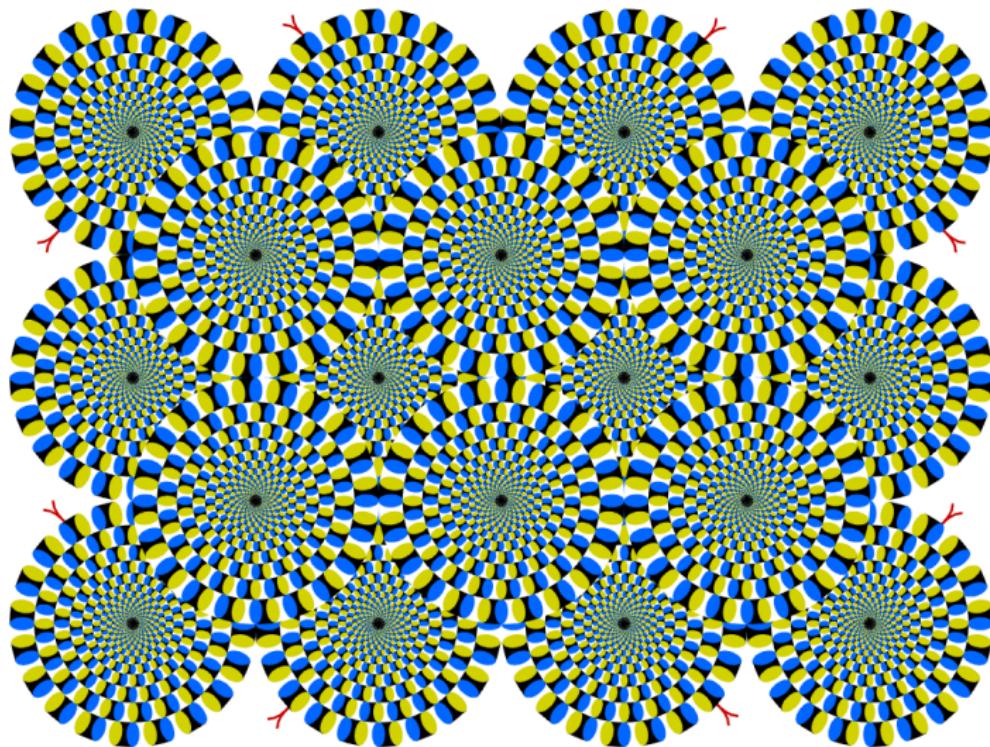
Une vision humaine non parfaite

Quel âge a-t-elle ?



Une vision humaine non parfaite

Impression de mouvement (Professeur Kitoaka)



Une vision humaine non parfaite : Change Blindness



FIGURE – Source : http://www.medien.ifi.lmu.de/fileadmin/mimuc/mmi_ws0405/uebung/essays/sebastian.kraemer/Change%20Blindness.html

- <http://www.cogsci.uci.edu/~ddhoff/cb.html>
- <http://www.cogsci.uci.edu/~ddhoff/cbvenice.html> [Rensik et al, 99]²

2. http:

[//nivea.psych.univ-paris5.fr/RensinkOReganClarkVisCog/RensORClark.pdf](http://nivea.psych.univ-paris5.fr/RensinkOReganClarkVisCog/RensORClark.pdf) ↗ ↘ ↙ ↘

Une vision humaine non parfaite mais l'homme est très performant en interprétation

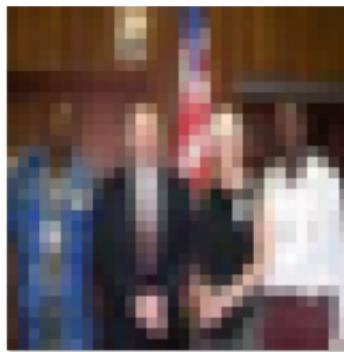
Interprétation d'images : très facile et très rapide pour un humain [Thorpe et al., Nature 1996]³ : reco animal < 150 ms



3. <https://www.ncbi.nlm.nih.gov/pubmed/8632824>

Une vision humaine non parfaite mais l'homme est très performant en interprétation

Les humains peuvent dire beaucoup sur une scène même à partir d'une information très pauvre.



Source: "80 million tiny images" by Torralba, et al.

A propos de ce dataset : 80 million tiny images - Supprimer en 2020 car contenant des images litigieuses⁴

4. <https://groups.csail.mit.edu/vision/TinyImages/>

Petite parenthèse éthique

Jeu de données d'images à grande échelle : à quel prix ? - Une victoire à la Pyrrhus⁵.

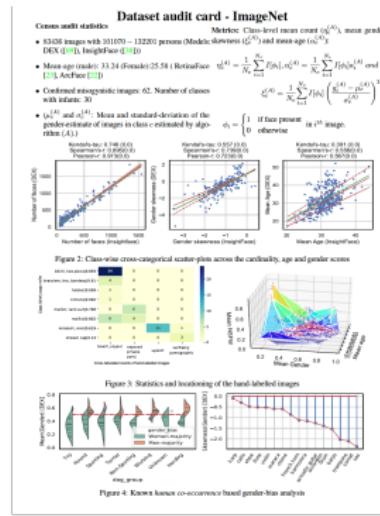


FIGURE – Source : [Birhane et al, WACV 2021] Large Image datasets : A pyrrhic win for computer vision ?⁷

5. victoire obtenue au prix de pertes si lourdes pour le vainqueur qu'elle équivaut quasiment à une défaite
6. <https://arxiv.org/abs/2006.16923>

Une vision humaine non parfaite mais l'homme est très performant en interprétation

Importance du contexte.

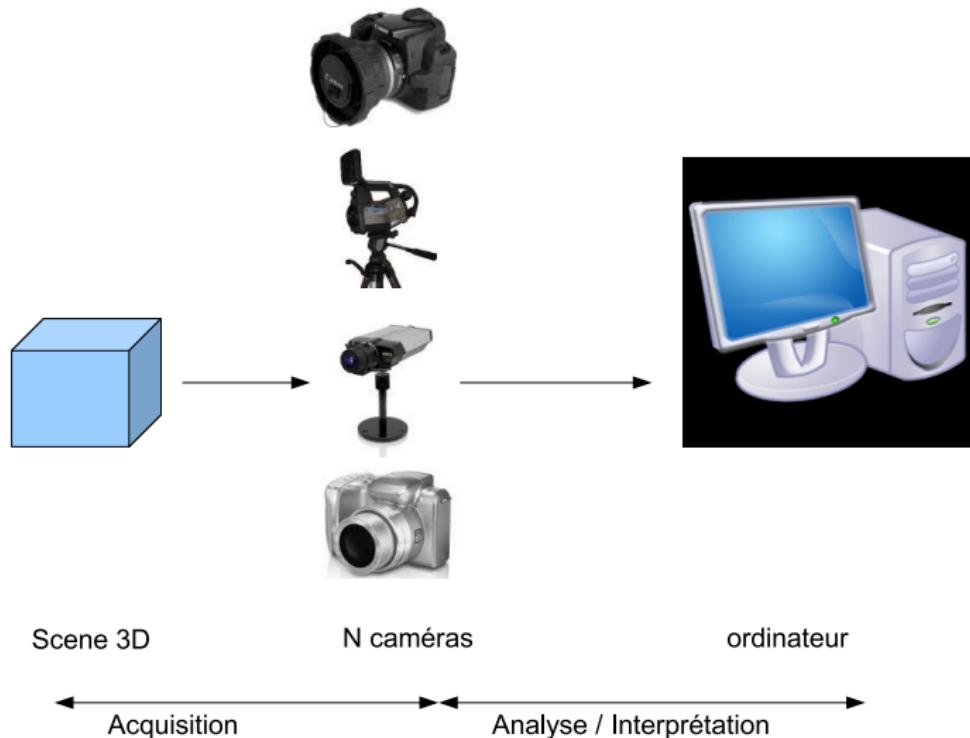


Une vision humaine non parfaite mais l'homme est très performant en interprétation

Importance du contexte.



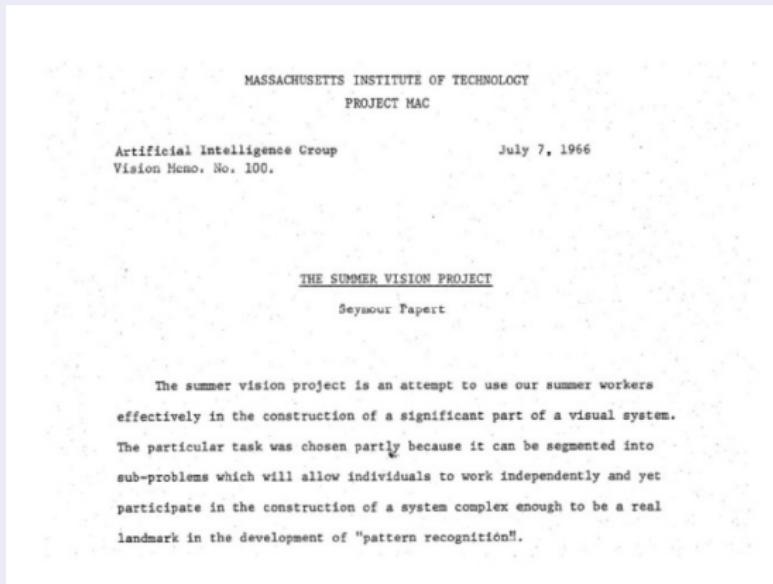
Vision par ordinateur



Vision par ordinateur

Les origines

Un projet d'été du MIT

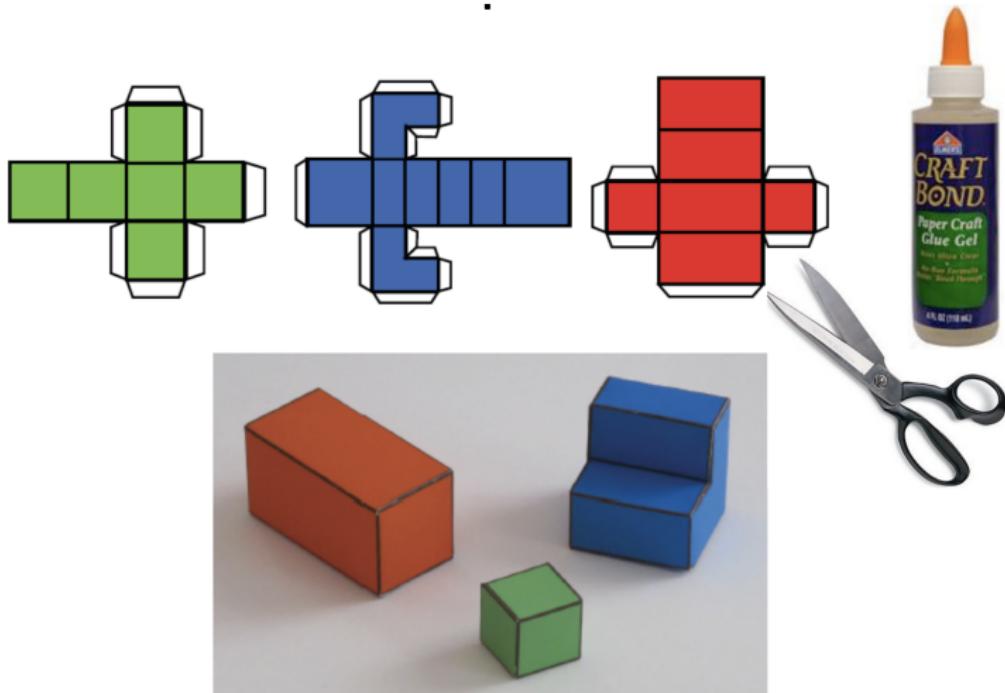


Voir <https://dspace.mit.edu/handle/1721.1/6125>



Vision par ordinateur

Pour aller plus loin : Simple vision system⁸.

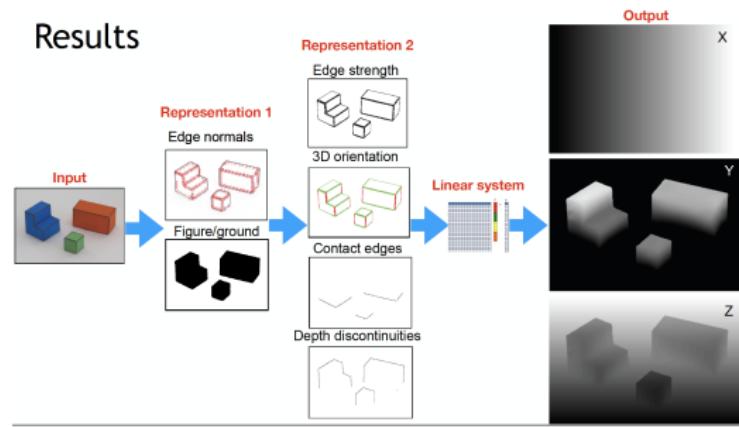


8. http://6.869.csail.mit.edu/fa13/lectures/chapter_01_simplesystem.pdf

Vision par ordinateur

Pour aller plus loin : Simple vision system⁹.

- Un monde simple
- Un but simple : reconstruction de l'information 3D à partir d'images 2D
- Un modèle simple de formation d'images.



9. http://6.869.csail.mit.edu/fa13/lectures/chapter_01_simplesystem.pdf

Vision par ordinateur

La vision humaine est

- Une grande source de motivation pour les recherches en vision par ordinateur : différentes philosophies.
- Un modèle mais pas exclusif pour la vision par ordinateur :
 - ▶ L'oeil : une caméra performante.
 - ▶ Le cerveau : un ordinateur performant.

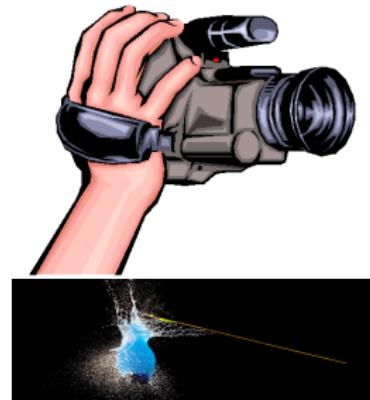
Une longue chaîne de traitements complexes

- Acquisition des images (assez facile).
- Interprétation automatique des images par l'ordinateur (très difficile car l'image n'est qu'un ensemble de pixels).

Caméra

Caractéristiques

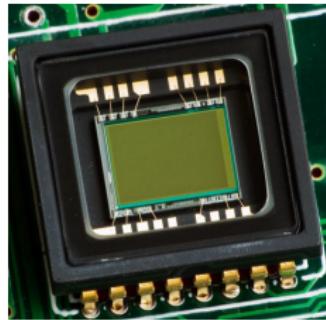
- Résolution spatiale : 60K pixels à 2800 M Pixels (source Wikipedia).
- Résolution angulaire : dépend de l'objectif.
- Résolution temporelle : de 0,5 images par seconde à 4000 (?) images par seconde.
- Résolution spectrale : infrarouge à ultraviolet selon les capteurs.
- Extrême sensibilité aux conditions d'éclairage



Caméra

Composée de deux blocs

- Le capteur : CMOS ou CCD (transfert de charge)
- L'objectif : assemblage de lentilles fixes ou mobiles



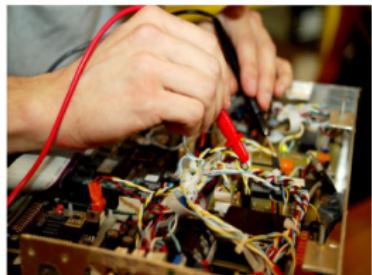
Caméra

Résolution

- Généralement exprimée sous forme de PixLar × PixLon ou nombre de pixels.
- La résolution (angulaire...) dépend du capteur et de l'objectif

Caméra

Influence de la résolution du capteur avec une focale fixe



800 × 600



100 × 75



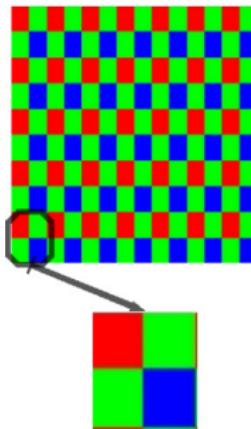
50 × 38



25 × 19

Caméra : capteur

- Une matrice de points qui mesurent la quantité de lumière reçue pendant un temps donné.
- Capteur noir et blanc : tous identiques.
- Capteur couleur : arrangement des pixels comme dans l'image de gauche.
- Arrangement lié à la sensibilité de l'oeil au vert.



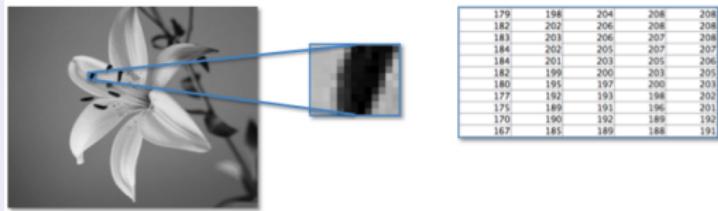
Plan

- 1 Organisation du cours
- 2 Introduction
 - Définition
- 3 Vision humaine - Vision par ordinateur
 - Vision humaine
 - Vision par ordinateur
- 4 Qu'est ce qu'une image ?
 - Vision par ordinateur : les difficultés
- 5 Applications de la vision par ordinateur

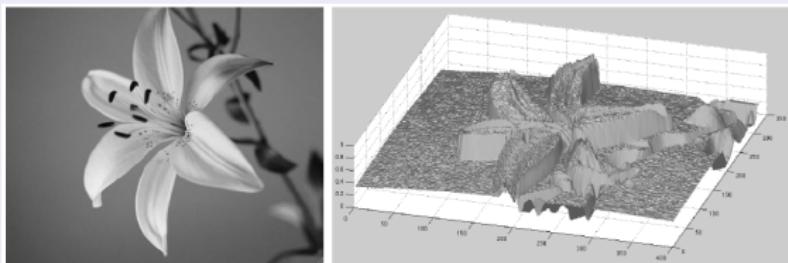
Qu'est-ce qu'une image ?

Deux visions

- Vision discrète



- Vision surfacique



Image

Une image NB est :

Un tableau de pixels (picture elements) de dimension NxM qui contient la luminosité de chaque pixel.

0	130	255
130	86	10
255	18	10



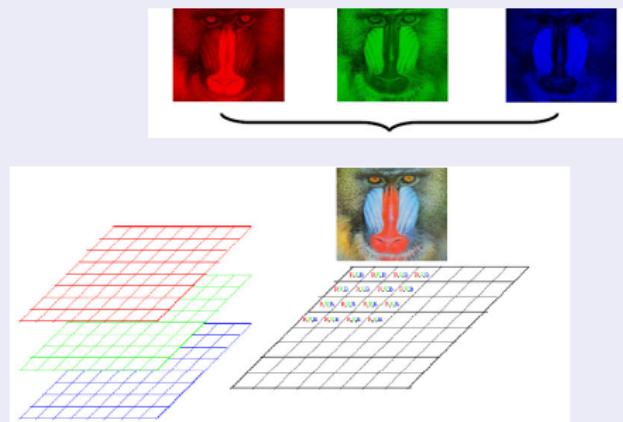
Une image numérique est un tableau de pixels. Un pixel s est décrit par :

- ses coordonnées dans l'image (i, j) .
- sa valeur $I(i, j)$, représentant son niveau de gris.

Image

Une image couleur est :

- Un tableau de dimension NxMx3.
- Dans chacun des trois plans la valeur exprime la luminosité d'une des couleurs de référence.
- Référence de base : RVB
- Plusieurs bases : HSV, Luv,...



Source : Alain Boucher

Image : codage

La valeur $I(i,j)$ d'un pixel $s = (i,j)$ représente son intensité lumineuse (ou sa couleur).

- En niveau de gris
 - ▶ binaire : $I(i,j) = 0$ noir ou $I(i,j) = 1$ blanc.
 - ▶ codage 8 bits (le plus classique) : $I(i,j) = 0, \dots, 255$ du plus foncé au plus clair.
- En couleur
 - ▶ codage dans l'espace RGB : trois intensités lumineuse rouge, vert, bleu.
 - ▶ codage 24 bits (le plus classique) : $I_R(i,j) = 0, \dots, 255$;
 $I_G(i,j) = 0, \dots, 255$; $I_B(i,j) = 0, \dots, 255$

Qu'est ce qu'une image ?

On peut voir l'image comme une fonction :

$$\begin{aligned} I : \quad S &\rightarrow \Omega \\ (i, j) &\rightarrow x = I(i, j) \end{aligned}$$

- En discret :

$$S = \{0, 1, \dots, n_l - 1\} \times \{0, 1, \dots, n_c - 1\}$$

$$\Omega = \{0, 1, \dots, 255\}$$

- En continu :

$$S = [0, h] \times [0, l]$$

$$\Omega = [0, P]$$

Expression de la dérivée... Varie.

Qu'est ce qu'une image ?

Une image est avant tout un signal 2D (x, y) qui représente une réalité 3D (x, y, z)

Du point de vue mathématique

- Une matrice de nombres représentant un signal.
- De nombreux outils permettent de manipuler ce signal.

Du point de vue humain

- Une image contient plusieurs informations sémantiques.
- Il faut interpréter le contenu au delà de la valeur des nombres.

Image numérique

- Les valeurs de $I(x, y)$ sont la réponse du capteur au phénomène observé (énergie lumineuse).
- Les valeurs de $I(x, y)$ sont des valeurs de voltage continu.
- Les valeurs de $I(x, y)$ doivent être converties vers le domaine du numérique.
- **Numérisation** = Echantillonnage + Quantification
 - ▶ Echantillonnage spatial : discréétisation des coordonnées de l'image réelle.
 - ▶ Quantification des luminances : discréétisation de l'intensité de l'image réelle.

Echantillonnage et quantification

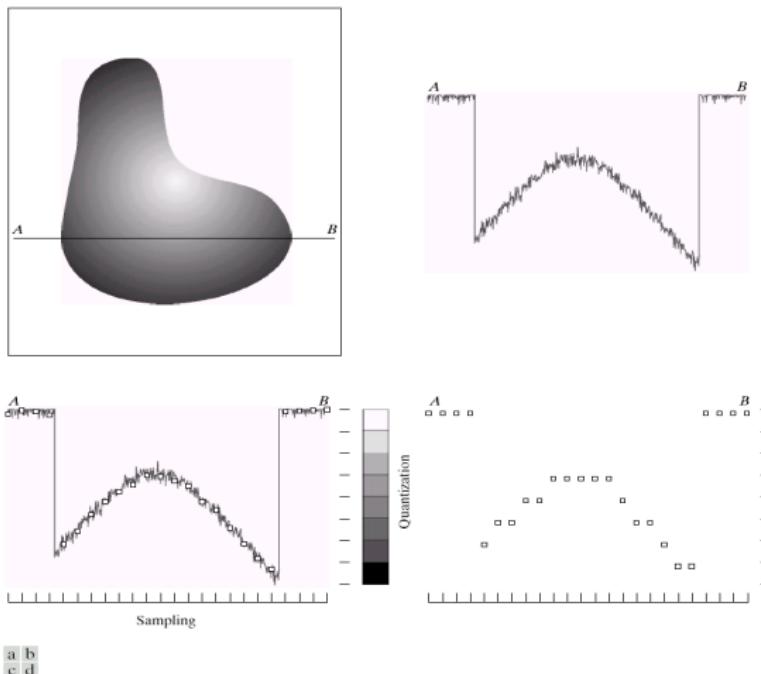
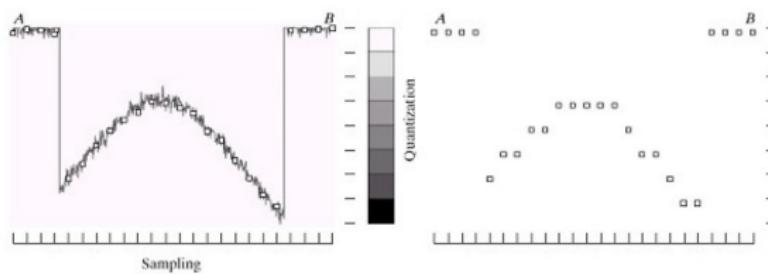


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

Source : Gonzalez and Woods. Digital Image Processing. Prentice-Hall, 2002.

Echantillonnage et quantification

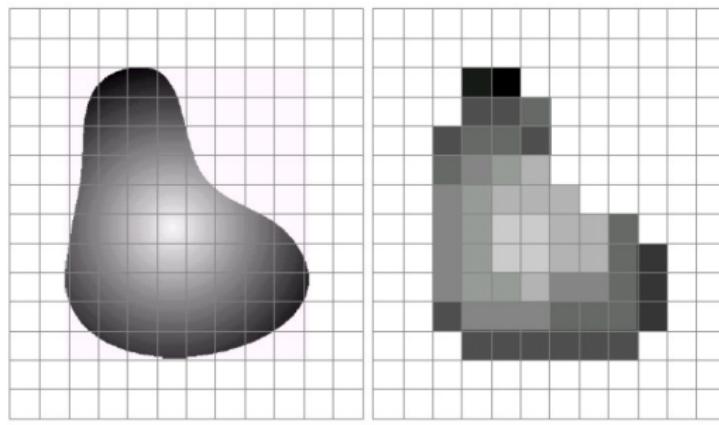
- L'échantillonnage est limité par la capacité du capteur, i.e. le nombre de pixels disponibles.
- La quantification est limitée par la quantité de tons (de gris) définie dans l'intervalle.



Source : Gonzalez and Woods. Digital Image Processing. Prentice-Hall, 2002.

Echantillonnage et quantification

Avec un capteur à matrice.



a b

FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

Echantillonnage et quantification

Echantillonnage

Résolution spatiale : le plus petit détail discernable



Quantification

Résolution tonale : le plus petit changement discernable



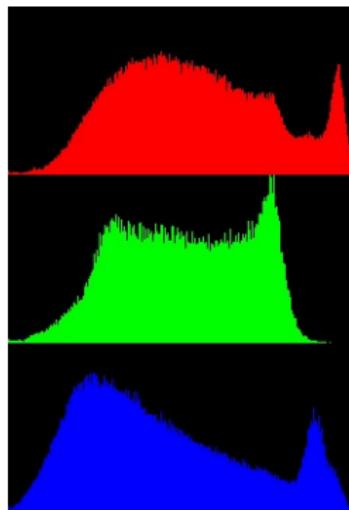
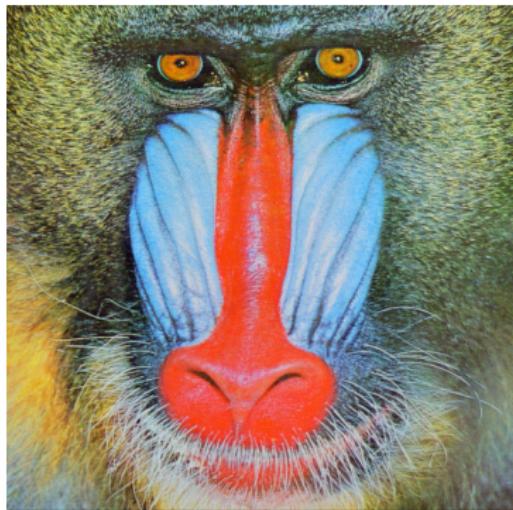
Représentation des images

- Matrice de dimension $M \times N$.
- Chaque élément à une valeur entière dans l'intervalle $[L_{min}, L_{max}]$.
- Le nombre de bits requis pour représenter les niveaux de gris dans l'intervalle L est K , avec $L = 2^K$.
- Le nombre de bits pour stocker une image est donc : $b = M \times N \times K$.

Une image a donc une résolution spatiale de $M \times N$ pixels et une résolution tonale de K bits ou de L niveaux.

Histogramme d'une image

- Distribution des niveaux de gris (ou de couleurs) dans une image.
- $H(k)$ = nombre de pixels ayant la valeur k .

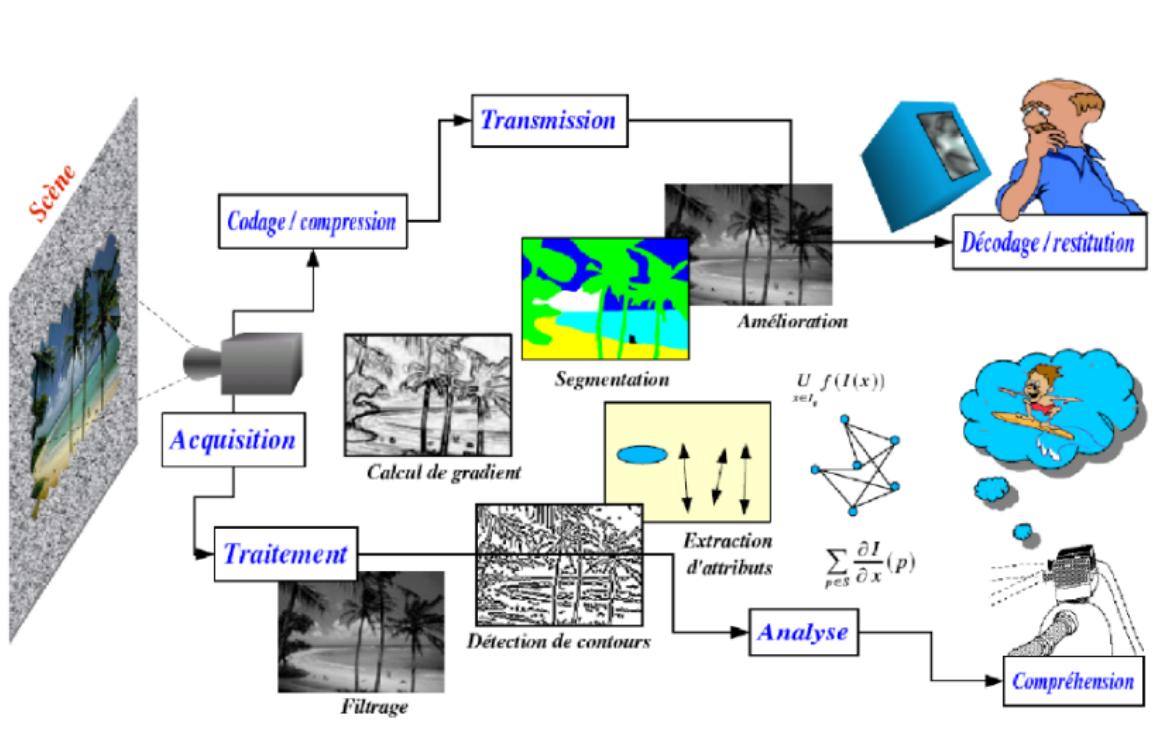


Histogramme d'une image

Histogramme normalisé

- Proportion de pixel en fonction du niveau de gris (densité de probabilité)
- $x \rightarrow H_n(x) = \frac{H(x)}{nbpixels}$

Vision par ordinateur : chaîne de traitements



D'après A. Manzanera

Vision par ordinateur : plusieurs phases

Les différentes étapes - différents niveaux

① Acquisition de l'image

- ▶ Quels capteurs ?
- ▶ Quel espace colorimétrique ?

② Traitement de bas niveau (traitement d'images).

- ▶ Filtrage.
- ▶ Extraction de primitives : points d'intérêts, contours, régions, ...

③ Traitement de haut niveau.

- ▶ Reconnaissance de forme (2D,3D).
- ▶ Reconstruction 3D.
- ▶ Analyse de mouvements.
- ▶ Interprétation d'images.

Vision par ordinateur : plusieurs phases

Les différentes étapes - différents niveaux

- Traitement bas niveau
 - ▶ Entrée : image
 - ▶ Sortie : image
- Traitement moyen niveau
 - ▶ Entrée : image ou représentation préalable
 - ▶ Sortie : une nouvelle représentation
- Traitement haut niveau
 - ▶ Entrée : image ou représentation préalable
 - ▶ Sortie : sémantique de l'image

Chaine de traitements : différents niveaux d'analyse

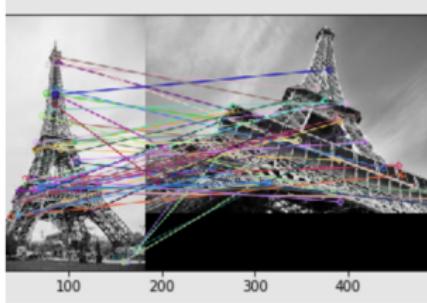
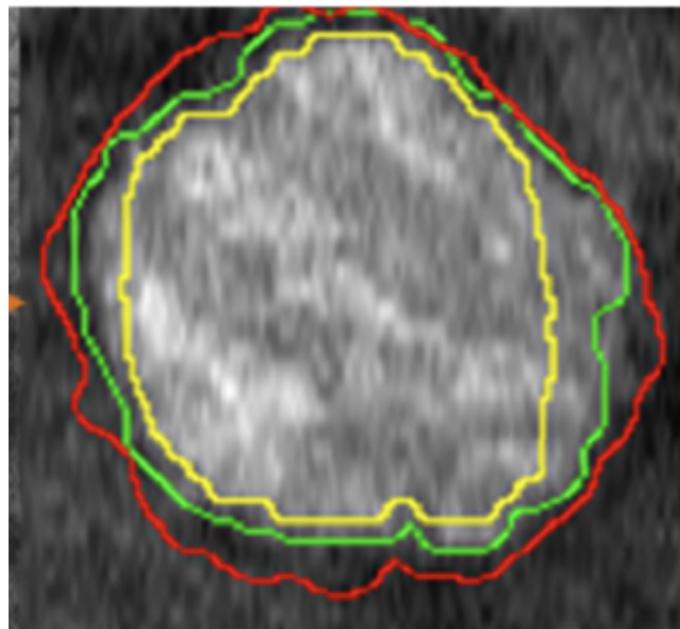
Exemple bas niveau : améliorer les caractéristiques d'une image.



D'après E. Arnaud.

Chaine de traitements : différents niveaux d'analyse

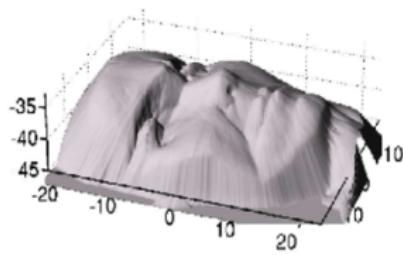
Exemple moyen niveau : detection de contours, détection de points d'intérêt.



D'après E. Arnaud.

Chaine de traitements : différents niveaux d'analyse

Exemples haut niveau : suivi, shape from X, reconstruction, ..



Plan

- 1 Organisation du cours
- 2 Introduction
 - Définition
- 3 Vision humaine - Vision par ordinateur
 - Vision humaine
 - Vision par ordinateur
- 4 Qu'est ce qu'une image ?
 - Vision par ordinateur : les difficultés
- 5 Applications de la vision par ordinateur

Perte de la profondeur

Passage d'une scène 3D à une scène 2D



Distorsion d'images liée à la perte de profondeur et au point de vue

Une multitude de transformations possibles

Transformations géométriques

- Translations, rotations de l'image.
- Changements d'échelle.
- Changements de points de vue.
- Occultations, fonds différents.

Transformations photométriques

- Changements d'illuminations (défaut de caméras, déplacement des sources).

Une multitude de transformations possibles

Changement d'échelle



<http://www.cs.cmu.edu/~chuck/lennapg/>

Une multitude de transformations possibles

Changement d'illumination

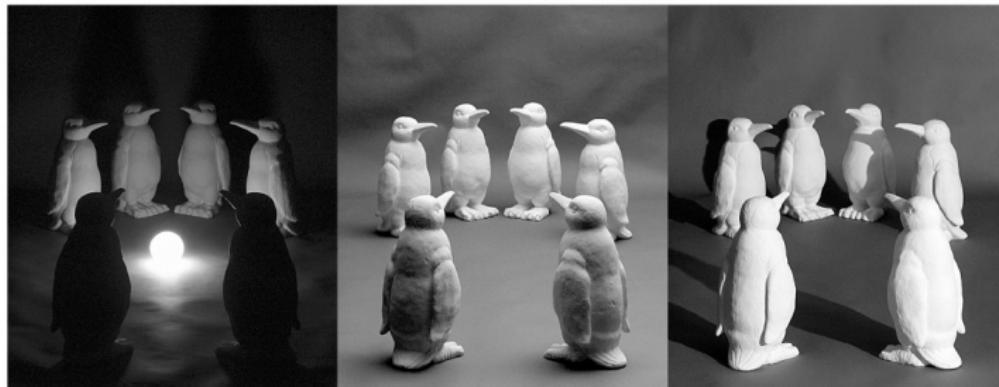


image credit: J. Koenderink

Une multitude de transformations possibles

Variation des points de vue



Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Une multitude de transformations possibles

Occlusions



Bruit dans l'image



Une réalité complexe

Fond texturé



slide credit Svetlana Lazebnik

Une réalité complexe

Mouvement lors de la prise de vue



slide credit Svetlana Lazebnik

Une réalité complexe

Diversité visuelle d'un même objet



slide credit Fei-Fei, Fergus & Torralba

Pas de sémantique

Sémantique de cette image ?

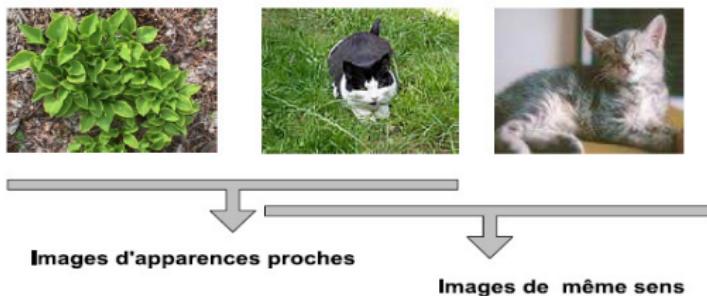
- *Un objet blanc sur un fond vert*
- *Un insecte*
- *Une aleurode sur une feuille de rosier*



Conséquences

- La sémantique n'est pas dans l'image elle-même.
- L'interprétation dépend de connaissances a priori.
- => Knowledge-based computer vision.

Fossé sémantique



Fossé sémantique

Manque de concordance entre les informations que l'on peut extraire des données visuelles et l'interprétation qu'ont ces mêmes données pour un utilisateur dans une situation déterminée (Smeulders et al., 2000).

Pourquoi étudier la vision par ordinateur ?

Les images et les vidéos sont partout.

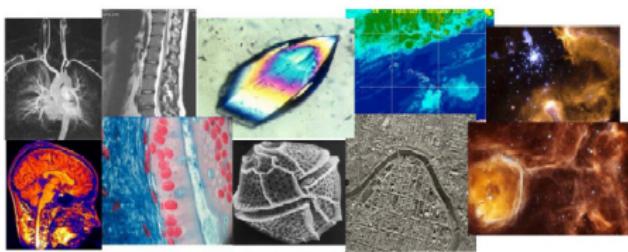


Google™
Image Search Picasa™

flickr™ webshots™ picsearch™ YouTube™
Broadcast Yourself™



Surveillance and security



Medical and scientific images

FIGURE – Source : Fei-Fei Li, Stanford

Systèmes de vision par ordinateur industriels

Exigences et caractéristiques

- S'adressent en général à des **tâches répétitives**, simples et bien définies.
- Nécessitent que les caractéristiques des scènes à observer et l'information à extraire soient connues.
- Doivent fonctionner de manière autonome, non supervisée ou semi-supervisée.
- Les méthodes de mesure quantitatives ou qualitatives doivent être fiables et reproductibles.
- Les traitements doivent suivre la cadence imposée par la chaîne de fabrication.
- Se contentent souvent de l'information en deux dimensions sans passer par une reconstruction.

Principaux avantages

- Permettent une grande variabilité d'utilisations (adaptation à des pièces différentes).
- Plusieurs mesures peuvent être réalisées simultanément.
- Mesures sans contact et non destructives.

Inspection

De nombreux procédés industriels nécessitent un tri automatisé en différentes catégories.

Conforme - Non Conforme : contrôle qualité

- Embouteillage de boissons.
- Pièces plastiques moulées.



Calibre et couleur

- Fruits et légumes.
- Médicaments.



Pixel©

Gestion de production

Objectif

Surveillance d'un procédé de fabrication.

Avantages

- Contrôle qualité global et non plus seulement sur un prélèvement de la production.
- Permet d'atteindre des cadences que l'oeil humain ne peut plus gérer.
- Permet d'établir des statistiques multicritères très précises sur la production (% de rebus, temps d'arrêt).

Gestion de production : Identification automatique par la vision

Arcelor Mittal à Fos sur Mer



Gestion de production : Reconnaissance de caractères

OCR : Optical character recognition

- Tri postal.
- Encaissement des chèques.
- Analyse de formulaires.
- Lecture et reconnaissance de codes barre.
- Reconnaissance de plaque d'immatriculation
(http://en.wikipedia.org/wiki/Automatic_number_plate_recognition)

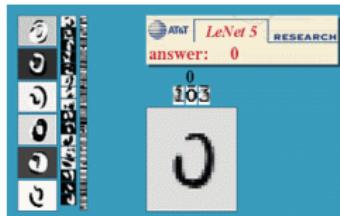


FIGURE – Reconnaissance de nombres, AT&T labs

<http://yann.lecun.com/exdb/lenet/index.html>

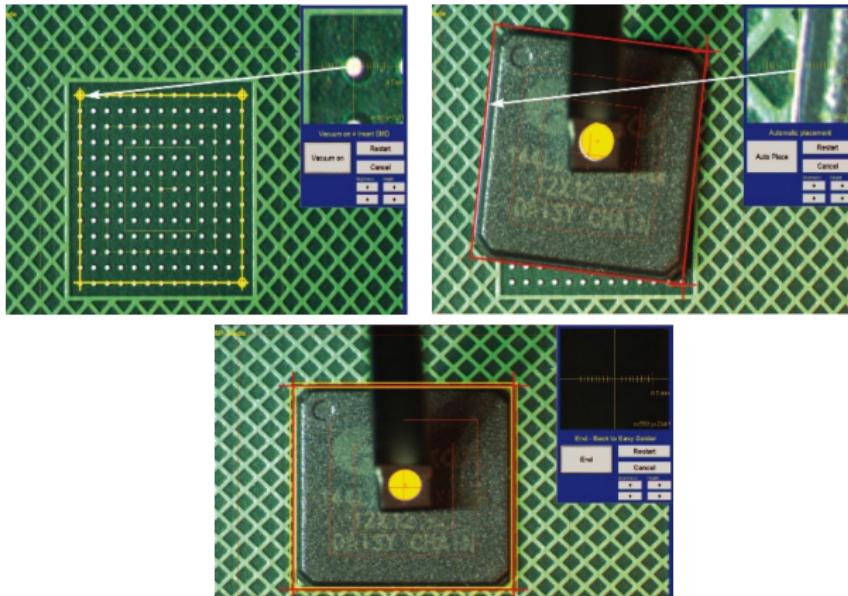
Gestion de production : Positionnement

Lorsque l'oeil humain n'est plus capable (pour des raisons de cadence et de précision), d'effectuer une tâche, il peut être intéressant de lui adjoindre une assistance visuelle ou d'automatiser complètement la tâche.

Fabrication de composants électroniques

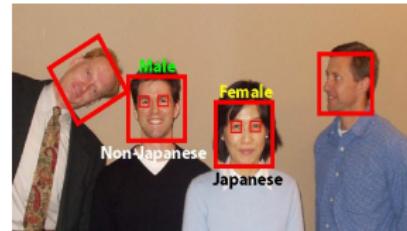
- Placement des composants de surface

Gestion de production



Sécurité et surveillance : biométrie

- Détection de visages.
- Reconnaissance de visages.
- Reconnaissance d'empreintes.
- Reconnaissance de l'iris.
- ...

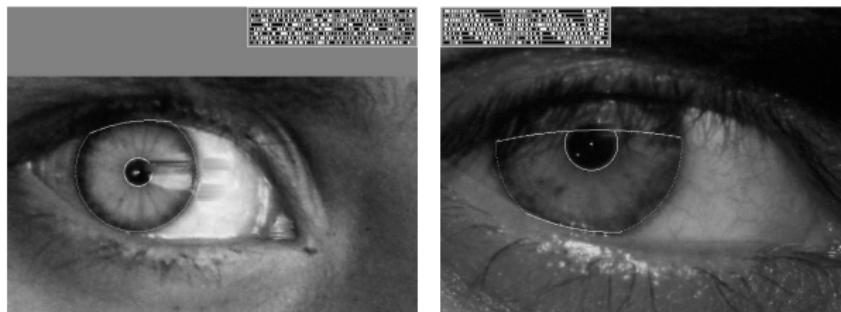


Fingerprint Analysis Software



La biométrie et en particulier l'identification de personnes trouve aussi de plus en plus d'applications dans le domaine de la vente ciblée

Sécurité et surveillance : biométrie

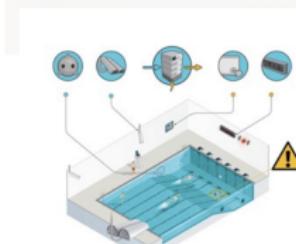


Source : Steve McCurry : <http://www.cl.cam.ac.uk/~jgd1000/afghan.html>

Sécurité et surveillance : vidéo-surveillance



Axis : Détection d'intrusion
Analyse de mouvement



Poseidon: détecteur
noyade

Poseidon, système d'IA pour la détection des
noyades à l'aide de caméras aériennes
et subaquatiques



Axis : préservation de la vie privée

Industrie automobile

How Does Mobileye SuperVision™ Work?

Mobileye SuperVision's end-to-end system is enabled by:

- Full surround high-definition computer vision perception (1 cameras)
- Road Experience Management™ (REM)-based AV maps
- Responsibility-Sensitive-Safety™ (RSS) safety model & driving policy
- 2 EyeQ™ S6/High SoCs

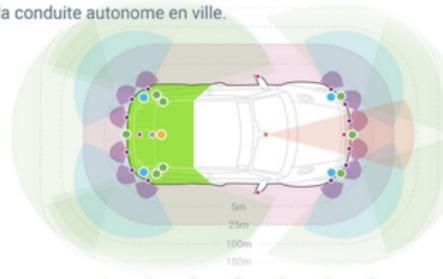
Camera Set

1 X Main	120°	8MP
1 X Narrow	20°	8MP
2 X WideFront	100°	8MP
2 X WideRear	100°	8MP
1 X Rear	40°	8MP
4 X Side	100°	8MP
1 X LR/Radar		

Legend: Camera, Parking Camera, Front LiDAR/NIR, Corner Radar (Optional)

Les capteurs Valeo DRIVE4U®

Configuration technique pour la conduite autonome en ville.



- MobilEye : <https://www.mobileye.com/>
- Autopilot Tesla : https://www.tesla.com/fr_FR/autopilot
- Valeo : conduite à Paris : <https://www.youtube.com/watch?v=ZxqBKqEQpZ8>

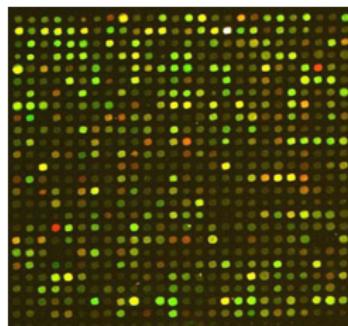
Applications médicales



Aide au diagnostic



Chirurgie guidée



Génétique

Applications robotiques



Cycab : INRIA
<http://www.lara.prd.fr/en/>

Aide à la conduite
Conduite automatique



Nao (Aldebaran)

Robots de service



BrainTech
Random bin picking

Reconstruction 3D : patrimoine

Building Rome in a day :



<http://grail.cs.washington.edu/rome/>

<http://www.cs.cornell.edu/~snavely/bundler/>

Dans le domaine du sport

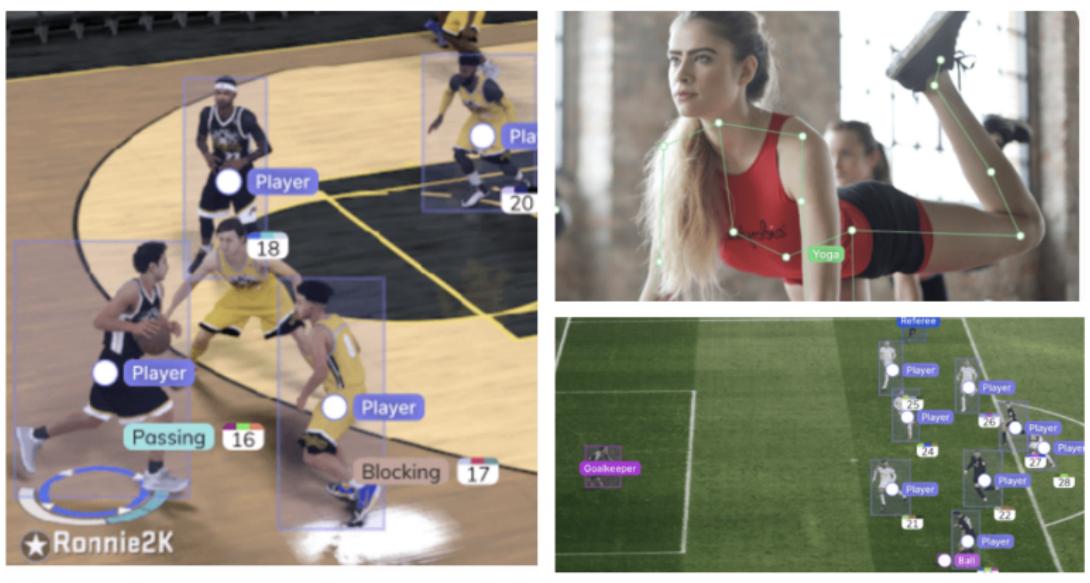


FIGURE – Source : V7 : <https://www.v7labs.com/>

Multimédia



Industrie du cinéma



Microsoft's Kinect

Industrie du jeu

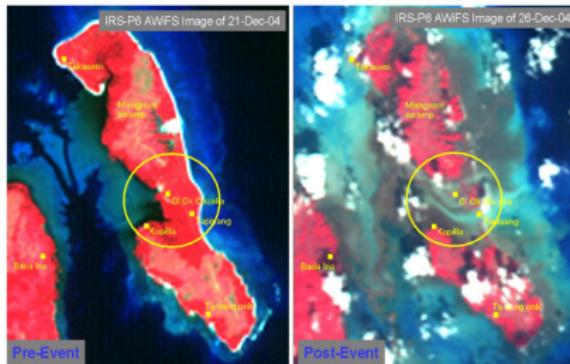


Multimédia

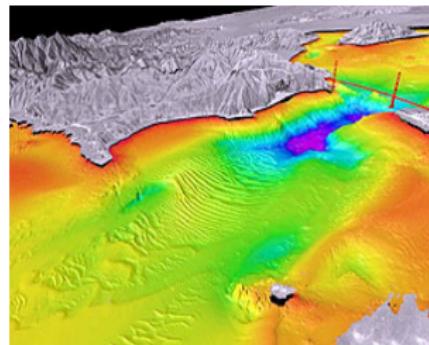
Observations de la terre

Tsunami - 2004

A Close View of Trinkat Island



Gestion des séismes



Exploration des fonds marins

Réalité augmentée



Présentation de produit interactif - Maintenance (credit : Diota)

Nouvelles applications

- Pinterest : Visual discovery :Pinterest LensBETA
<https://newsroom.pinterest.com/en/post/search-outside-the-box-with-new-pinterest-visual-discovery-lens/>
- Surveillance d'une marque : relomeric
<https://relometrics.com/>
- Reconnaissance d'émotions : <https://www.realeyesit.com/>
- D'une image à une recette : Pic2Recipe :
<http://im2recipe.csail.mit.edu/>
- Pour les régimes : <http://www.caloriemama.ai/>
- The visual microphone :
<http://people.csail.mit.edu/mrub/VisualMic/>

Applications

- <https://viso.ai/blog/applications/>
- <http://www.cs.ubc.ca/~lowe/vision.html>

En résumé

- Différences vision humaine / vision par ordinateur.
- Une image = une matrice de pixels qu'il faut **analyser** et **interpréter**.
- Un domaine avec de nombreuses applications présentes et futures.