

SKYNET

SAV  
Humain VS Artificielle

CYBERDYNE  
SYSTEMS

Romuald Mosqueron

Septembre 2017



# Humain VS artificiel



Dans plusieurs domaines nous pouvons faire ces comparaisons:  
ici

La vision i.e. voir et analyser => œil et cerveau

L'audition i.e. entendre et analyser => oreille et cerveau

L'Humain a ses propres caractéristiques variant d'un sujet à l'autre mais très proche

L'artificiel est défini par l'homme pour avoir ses propres possibilités:

- Faire ce que l'homme ne peut pas faire
- Le remplacer (AI)

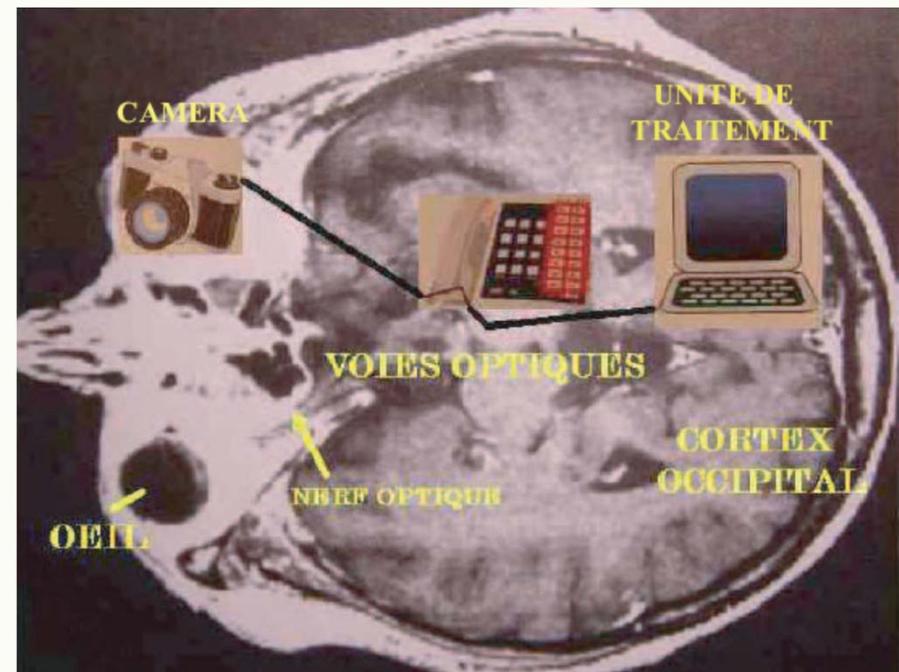
La constitution d'un œil humain a inspiré celle des appareils photographiques.

- Paupière => Obturateur
- Cornée et cristallin => Objectif
- Iris et pupille => Diaphragme
- Rétine => Surface photosensible

# La vision

En fonction des informations extraites des images par le cerveau, une prise de décision est réalisée. Ainsi, on peut comparer la vision humaine à un système composé d'une caméra connectée à une unité de traitement.

On considérera une vision idéale pour rester dans le cas le plus simple, la caméra correspondant à l'oeil et l'unité de traitement au cerveau.

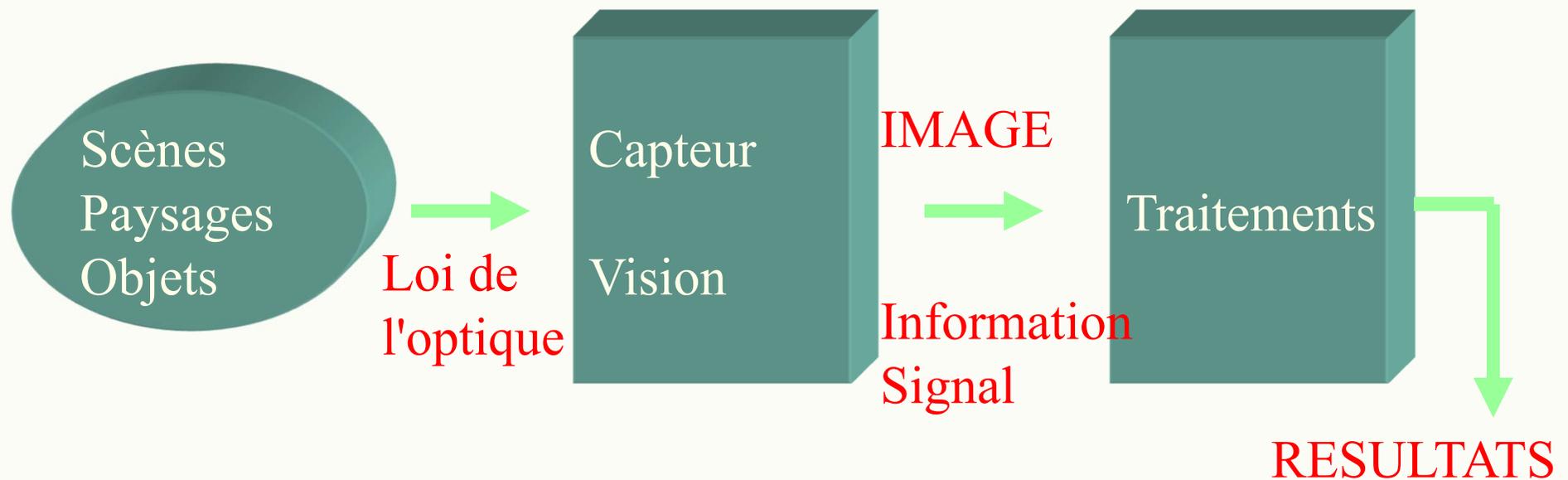


Il n'y a donc pas de couleurs. Si elles existent, c'est grâce à la vue. Notre perception des différentes nuances est rendue possible par la collaboration entre les yeux et le cerveau.

**Notre cerveau reçoit et traite les impulsions lumineuses selon un système de classement qui surpasse de loin n'importe quel ordinateur.** Chaque instant de vision est traité et interprété de manière ininterrompue. C'est l'unique moyen que nous avons pour percevoir la couleur.

# La vision

Dans les deux cas, le fonctionnement est similaire.



La perception d'une image est un processus actif. Le cerveau et l'oeil ajustent en permanence 3 paramètres :

- l'accommodation c'est à dire la mise au point en fonction de la distance de vision (comme l'objectif d'une caméra),
- l'adaptation c'est à dire la quantité de lumière admise à pénétrer l'oeil (comme le diaphragme de la caméra),
- la signification de l'information visuelle en fonction du contexte (comme un logiciel de traitement de l'image dans un ordinateur).

Cette analyse de l'image par la rétine et l'interprétation qu'en fait le cerveau, influencent à leur tour le comportement visuel (orientation du regard, adaptation des gestes, en particulier de la main, adaptation et accommodation sur la cible etc..).

On ne voit donc ce qui à un sens et une importance pour nous à un temps donné (phénomène d'attention).

NB: Le cerveau sert aussi pour le stockage des images et des références comme la mémoire d'un ordinateur.

La vision humaine a plusieurs cordes à son arc:

- Capteur œil, vision achromatique, vision 3D, perception du mouvement.

Et artificiel



## Appareil photo :

Au plus simple, la caméra est un appareil qui permet de faire des images en enregistrant la lumière réfléchiée par le sujet.

- L'objectif capte cette lumière et la dirige vers le film, sur lequel se trouve un matériau sensible à la lumière (l'émulsion).
- Ce film, qui se trouve dans un boîtier étanche, doit recevoir une quantité de lumière précise afin que l'image enregistrée ne soit ni trop pâle, ni trop foncée (ce qu'on appelle la densité de l'image).

## OEil :

- Avant de rentrer dans l'oeil, la lumière doit traverser la cornée et le cristallin, que l'on pourrait assimiler à deux lentilles convergentes comme celles que l'on trouve dans l'objectif d'une caméra.
- La quantité de lumière pénétrant dans l'oeil est régulée par l'iris qui fonctionne de la même manière que le diaphragme d'une caméra.
- Ensuite l'image vient se former sur la rétine de l'oeil ou sur le capteur de la caméra (ou sur la pellicule pour les anciens modèles).
- L'information est ensuite acheminée jusqu'au cerveau par le nerf optique. Avec une caméra l'information circule dans un fil électrique jusqu'au téléviseur ou l'ordinateur.

La vision des couleurs nous vient de l'interprétation qu'en fait notre système visuel (rétine, puis cerveau).

## Caméra couleur :

Dans une caméra couleur, on décompose la lumière par l'intermédiaire de filtres en trois composantes Rouge, Vert et Bleu. Ces composantes sont ensuite traitées et analysées séparément.

Il existe cependant, des différences significatives entre une caméra et notre oeil, certaines sont à l'avantage de l'oeil et d'autres à l'avantage de la caméra.

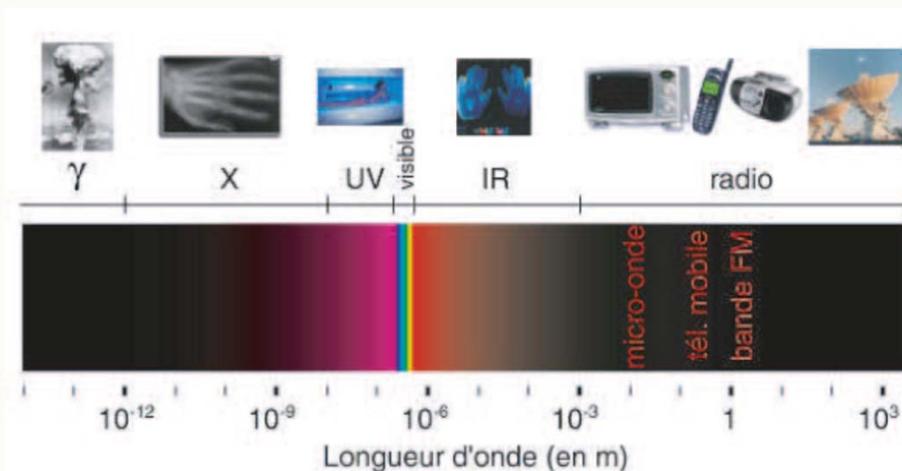
- La définition? Résolution
- Orientation?
- Traitement d'image?
- Spectre?
- Dynamique?
- Frame rate?

La rétine contient environ 125 millions de capteurs photosensibles à l'intensité lumineuse et à l'information couleur, chacun d'entre eux correspondant à un point de l'image. Dans les films photographiques, le nombre de points élémentaires est de 25 millions et de 25 millions au maximum pour la photographie numérique (pixels) mais l'image est **fixe**. Les capteurs des caméscopes grand public ont un nombre plus petit de l'ordre de 8 millions. On peut donc comparer l'oeil à une caméra très haute définition.

La vision humaine inclut aussi un traitement d'image très rapide effectué par le cerveau, cela parait naturel de détecter un contour ou une forme, le traitement est immédiat, alors que dans un système de vision, le traitement doit être fait par une unité dédiée tel qu'un ordinateur. Pour effectuer le traitement en temps réel, le temps de calcul devra être inférieur à celui de l'acquisition d'une image.

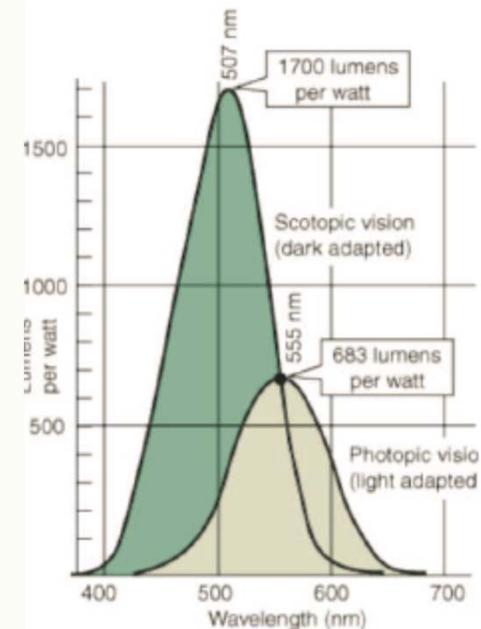
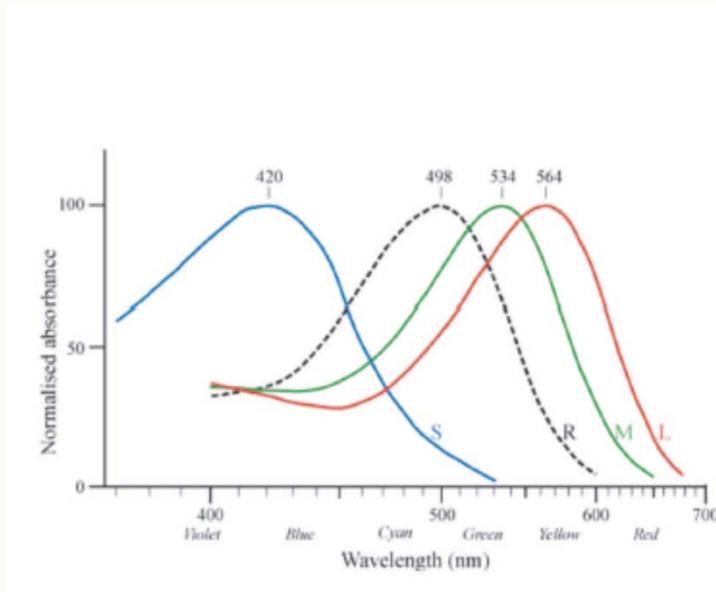
# Spectre

Le spectre électromagnétique est l'ensemble des rayonnements de nature électromagnétique allant des plus grandes ondes radio, jusqu'aux rayons X et gamma. La lumière visible n'est qu'une toute petite partie de ce spectre. Le spectre visible est la gamme de longueurs d'ondes visible par l'oeil humain. Elle s'étend classiquement de 380 à 780 nm. Ces extrêmes correspondent respectivement aux couleurs bleu et rouge. En dessous se trouvent les ultraviolets (UV) et au-dessus, les infrarouges (IR).



# Spectre

Les trois différentes sortes de cônes ainsi que les bâtonnets couvrent chacun une certaine gamme de spectre. En distinguant la vision nocturne (scotopique, principalement lié aux bâtonnets) et celle du jour (photopique, lié aux cônes), on peut établir un diagramme de spectre correspondant avec deux maxima vers 507 et 555 nm.



Les capteurs utilisés par les caméras => le spectre peut être plus large ou plus sélectif suivant le type de matériau utilisé. Les détecteurs peuvent être classés en trois grands groupes dépendant du spectre qu'ils couvrent:

- Ultraviolet : GaP (Phosphore de Gallium), GaN (Nitrure de Gallium), SiC (Carbure de Silicium) ou TiO<sub>2</sub> (Dioxyde de Titane).
- Infrarouge et au-delà: InGaAs (Arséniure de Gallium et Indium).
- Visible, proche-infrarouge: Si (silicium).

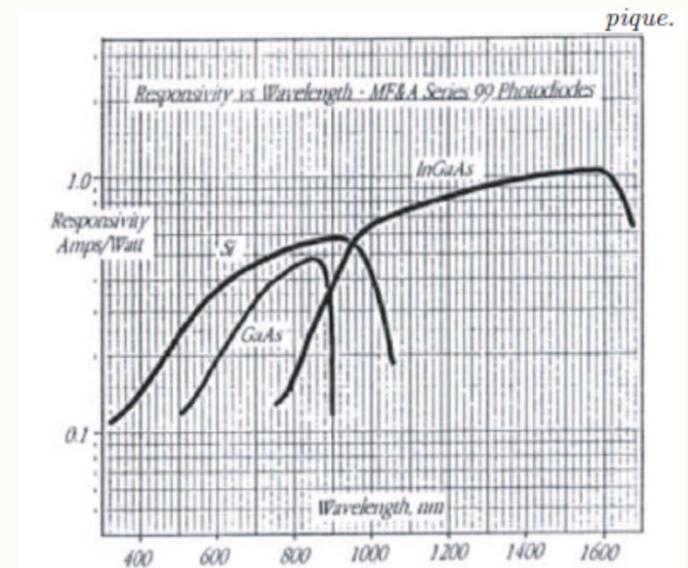
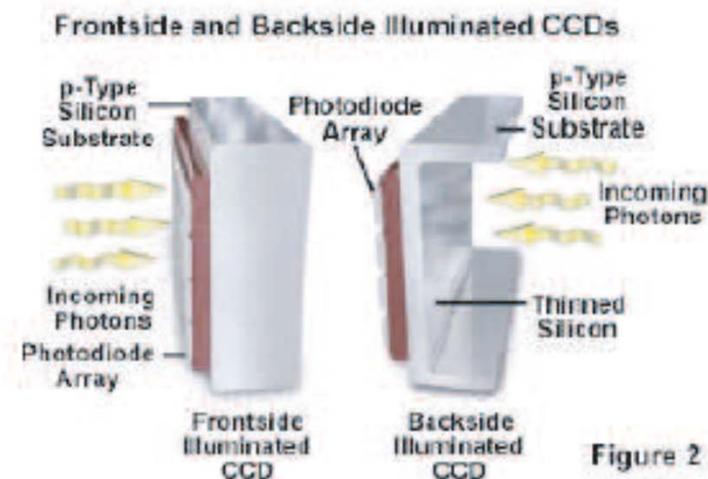
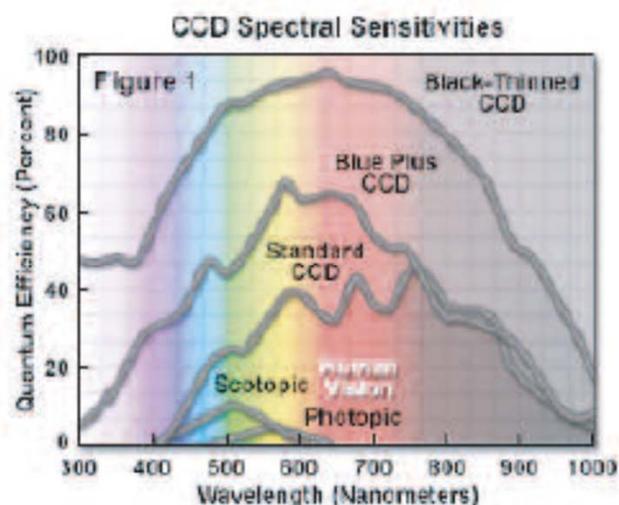


FIG. 1.5 – Spectre des capteurs.

# Efficacité Quantique

Permet de comparer les performances de conversion lumière-signal des différents détecteurs. Il s'agit de déterminer le nombre de photons qui entrent dans le dispositif par rapport au nombre réellement détectés.



# L'étendue dynamique



Pour les capteurs CCDs, l'étendue dynamique se calcule en considérant le rapport entre la capacité de charges d'un pixel en régime linéaire (avant saturation) et la valeur du bruit.

Exemple avec 4096 niveaux

L'oeil, quant à lui peut différencier au plus 128 niveaux de gris dans une chambre bien illuminée ce qui correspond à un codage de 7 bits en numérique.

%\subsection{Limite d'adaptation d'intensité} Le nombre d'électrons que peut stocker

%un pixel est défini par sa surface et sa profondeur.

%

%

%Si cette structure est établie, on a seulement deux possibilités  
%d'éviter les effets de saturation (blooming, smear). Soit en  
%ajoutant un "overflow gate" dans la partie horizontale comme  
%décrit dans la partie sur les détecteurs ou de prévoir un  
%mécanisme d'anti-saturation verticale pour que les électrons  
en

%surplus puissent se diriger vers le substrat mise à terre. La

%vision humaine contient plusieurs méthodes pour éviter la

Un bâtonnet = cylindre, 1  $\mu\text{m}$  de diamètre et 50  $\mu\text{m}$  de longueur. Les disques membraneux sont espacés de 20 nm.

La surface des pixels CCDs entre 2 et 14  $\mu\text{m}$  de côté.

L'épaisseur des couches stockant les charges se situe aux alentours de 200 à 300 nm. La répartition des zones photosensibles est parfaitement homogène.

La rétine étant courbée l'image obtenue sera donc non uniforme et sera corrigée par le cerveau.

Le cristallin n'étant pas extensible à volonté, on ne peut distinguer des détails inférieurs au millimètre. Un principe similaire est appliqué aux objets lointains.

Lors de l'observation d'une scène animée, il n'est pas possible à l'être humain de distinguer des changements trop brefs.

En effet, chez l'homme, l'acquisition d'images est limitée à 24 images par seconde pour des raisons physiologique. Par ailleurs, les capacités humaines en termes d'analyse temps réel de l'image, bien que très importante, ne permettent pas d'effectuer l'analyse d'une image en moins de cent millisecondes.

Dans le cas de caméras standards, l'acquisition d'images s'effectue à 25 images par seconde. En ce qui concerne les traitements, leurs temps de calcul dépendent entièrement des unités de calcul associées et bien entendu de leurs complexités.

	Œil	camera
résolution (pixels)	125 10 <sup>6</sup>	10 10 <sup>6</sup>
Spectre (nm)	400 à 700	300 à 1000
Efficacité quantique	3% photopic 10% scotopic	>95%
Etendue dynamique (bits)	6 à 7	12
Fréquence (images)	24	plusieurs millions

# Audition?



De la même façon que la vue, nous pouvons dire que des éléments non perceptibles peuvent être traités par des systèmes. De plus, certains systèmes vont améliorer notre sensation.

Les limites:

- Limite dans les fréquences.
- Limite de la pression.

=> Appareil pouvant aider à les entendre avec des traitements.

Pour la vue et l'audition, des systèmes audio-vidép sont utilisés pour rendre perceptibles ce qui ne l'est pas pour l'Homme et peut améliorer nos sens.

Par exemple:

- Vision Industrielle
- Militaire
- Enhanced Media (Musique, cinema, TV)
- Augmenter l'immersion. (4K, VR, HDR, HFR...)