

# Analyse et Reconnaissance d'Images

James L. Crowley

M2R IVR

Premier Semestre 2007/2008

Séance 2

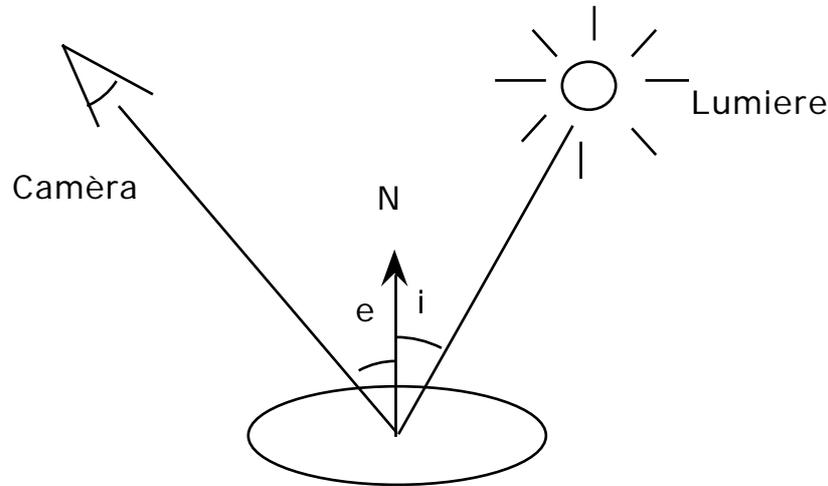
15 Octobre 2007

## Reflection et Couleur

### Plan de la Séance :

Couleur et Lumière.....	2
La Spectre .....	2
La Direction d'un rayon.....	3
Réflexion.....	4
Les fonctions de reflection.....	5
Le Modèle de réflexion di-chromatique.....	6
Les Espace de la Couleur .....	7
La Perception de la couleur.....	7
Les Espaces de la Couleur .....	12
l'espace RVB.....	12
l'espace CMY.....	12
L'espace YIQ.....	13
L'espace TLS.....	14
La Distribution de la Couleur d'une Surface .....	15
Analyse Probabiliste au niveau Pixel.....	16
Séparation de Luminance et Chrominance.....	17
Histogrammes de Couleur.....	18
Détection par ratio d'histogramme .....	19
Caractérisation par moments.....	20
Composantes principales.....	21

## Couleur et Lumière



Angles entre :

- e : Caméra et Normale
- i : Rayon incident (entre source et normale)
- g : Caméra et Source
- λ : Le longuer d'onde du photon

## La Spectre

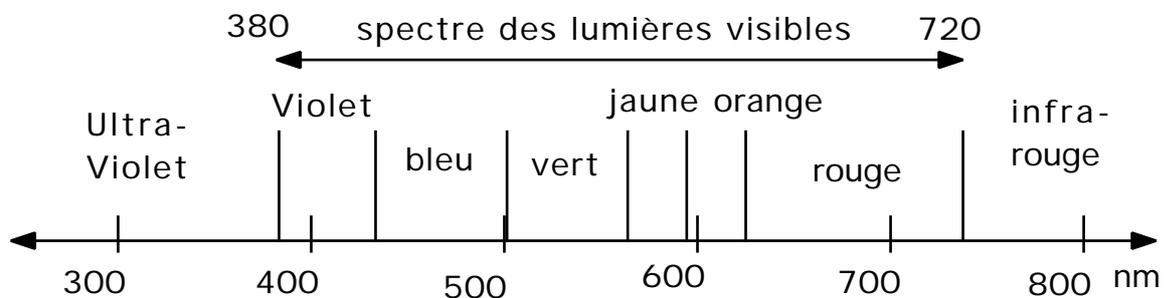
Une source lumière est caractérisé par une spectre :

Il s'agit de l'amplitude (quantité de photons en moyenne) pour chaque

longeur d'onde  $= \frac{1}{f}$ .

Les hommes peuvent percevoir les photons entre 380 nm et 720 nm.

Les caméras disposant d'une rétine en silicium peuvent percevoir d'une spectre de photons plus large.

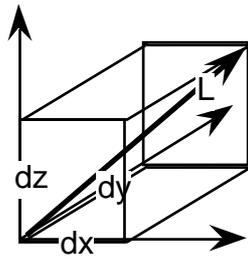


## La Direction d'un rayon

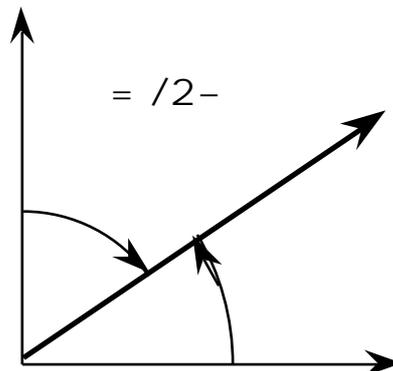
La direction d'une rayone de lumière est définit par une vecteur directeur composé de cosines d'angles.

$$D = \begin{pmatrix} \cos(\alpha) \\ \cos(\beta) \\ \cos(\gamma) \end{pmatrix} = \begin{pmatrix} x/L \\ y/L \\ z/L \end{pmatrix}$$

Les cosines de la direction sont la projection de la rayon sur 3 axes, normalisé à une longuer de unité.



Par exemple, en 2D

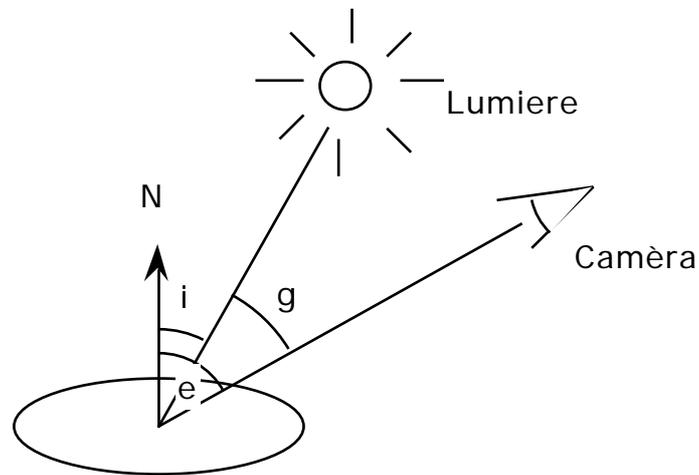


$$D = \begin{pmatrix} x/L \\ y/L \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \cos(\beta) \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \cos(\pi/2 - \alpha) \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$$

## Réflexion

### L'Albédo d'une Surface

Albédo : propriété de réflectance d'une surface



Angles entre :

e : Caméra et Normale

i : Rayon incident (entre source et normale)

g : Caméra et Source

Les angles sont mesuré en degrés ou en rayons (sur ).

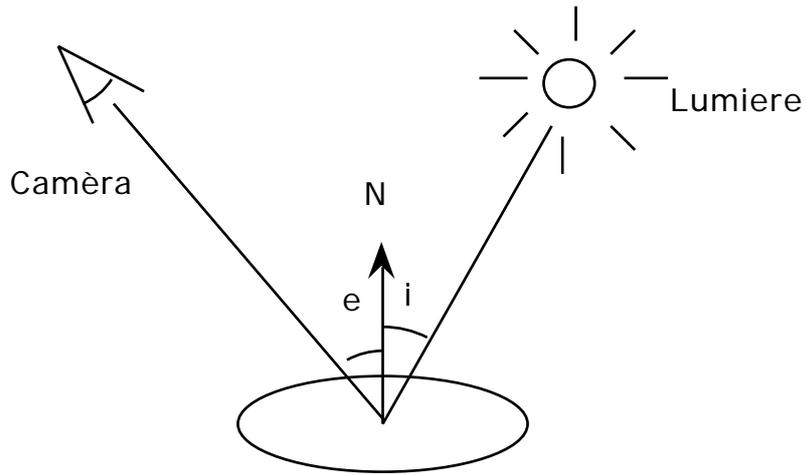
Carte de Réflectance : Ratio

$$R(i, e, g) = \frac{\text{Radiance}}{\text{Irradiance}} = \frac{\text{lumière émise}}{\text{lumière reçue}}$$

## Les fonctions de réflexion

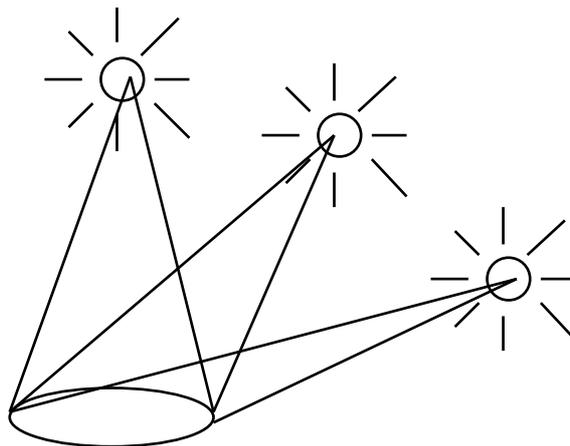
Réflexion spéculaire (la cas d'un miroir)

$$R(i, e, g, \lambda) = \begin{cases} 1 & \text{Si } i=e \text{ et } i+e = g \\ 0 & \text{Sinon} \end{cases}$$



Réflexion Lambertienne (exemples : neige, papier)

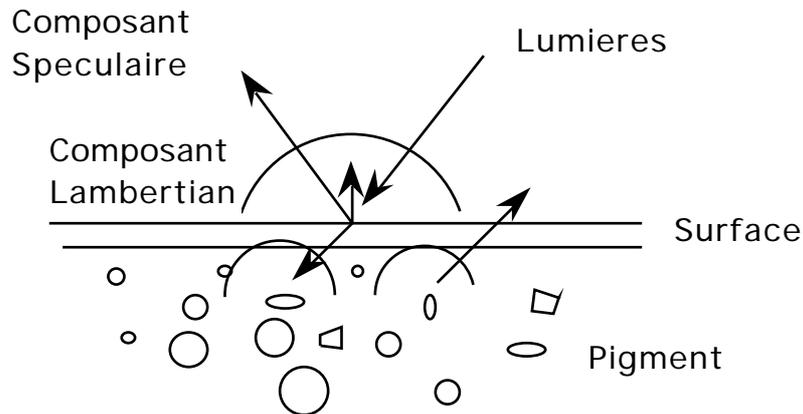
$$R(i, e, g, \lambda) = P(\lambda) \text{Cos}(i) \quad (P(\lambda) - \text{Spectre du pigment})$$



## Le Modèle de réflexion di-chromatique

L'albédo d'un objet non-métallique peut être approximé par la composition d'une réflexion "spéculaire" et d'une réflexion "lambertienne".

$$R(i, e, g, \lambda) = \rho_s R_s(i, e, g, \lambda) + \rho_l R_l(i, e, g, \lambda)$$



La composante "spéculaire" a le même spectre que la source de la lumière.  $S(\lambda)$ .

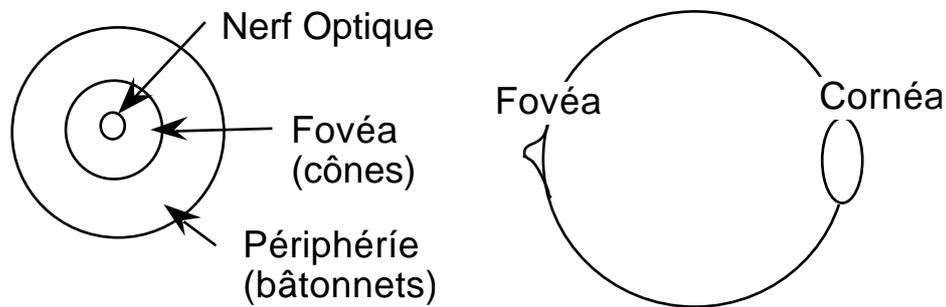
La composante "lambertienne" est "filtrée" par les pigments de la matière  $S(\lambda) P(\lambda)$

La composante "luminant" est déterminé par l'orientation de la surface.

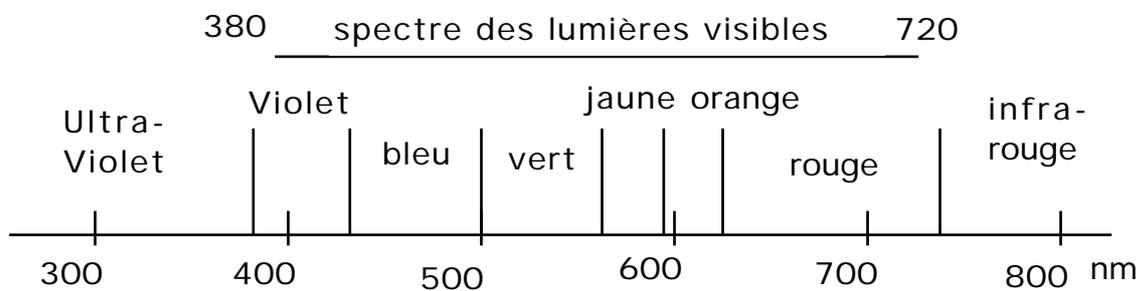
La composante "chrominant" est déterminé par la composition de la spectre de la source et le spectre d'absorption des pigments de la surfaces.

## Les Espace de la Couleur

### La Perception de la couleur



Spectre de la lumière visible pour l'homme :



Le système humain est composé de bâtonnets et de cônes.

Les bâtonnets sont responsables de la perception achromatique et des lumières atténuées.

Ils sont formés d'une pigmentation, la rhodopsine, sensible à tout le spectre, avec un sensibilité maximum de 510 nm (vert).

Ils sont dans la périphérie de la rétine, et donnent une perception de plus faible résolution.

Les cônes sont responsables de notre vision chromatique.

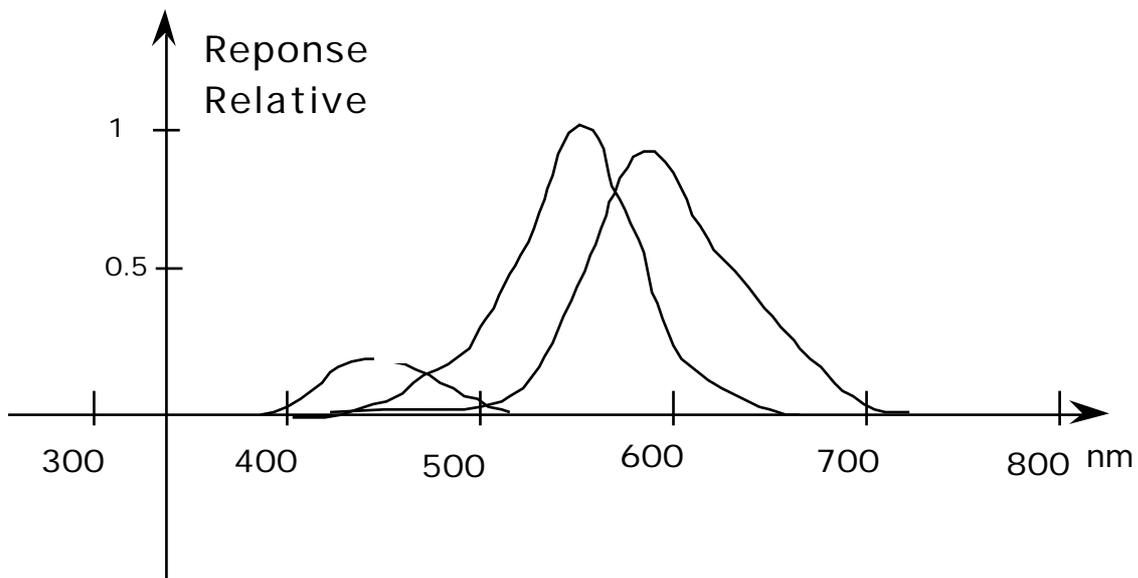
Ils sont concentrés dans une zone appelée la fovéa.

Il y a trois sortes de cônes, distingués par leurs pigments :

cyanolabe (445 nm),

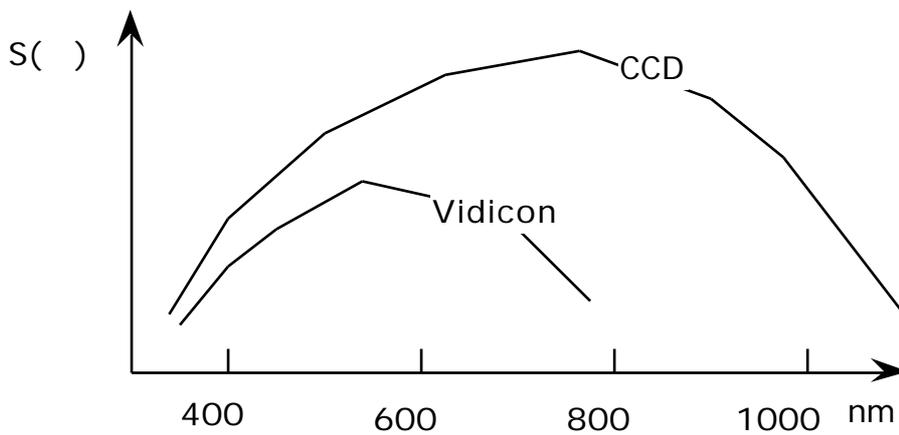
chlorolabe (535 nm), et

erythrolabe (570 nm)



La transformation entre stimuli de cônes et perception de couleur est un phénomène perceptuel qui n'est pas encore bien modélisé.

La sensibilité des caméras



Intensité de lumière perçue :

elle dépend :

du spectre de la source :  $S(\lambda)$

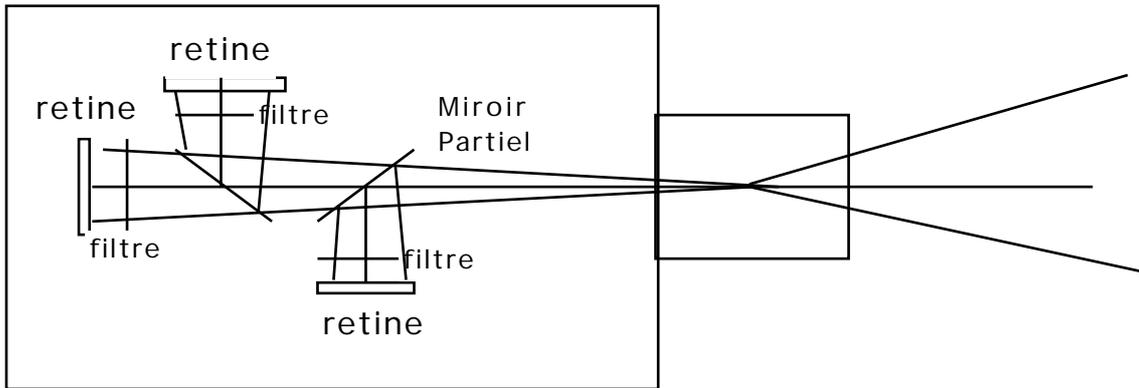
de la réflectance d'un point  $i, j$  de la scène vue :  $R(i, e, g, \lambda)$

et du spectre de réceptivité de la caméra,  $C(\lambda)$

$$p(x,y) = p_0 \int R(i, e, g, \lambda) S(\lambda) C(\lambda) d\lambda$$

$p_0$  est le Gain.  $R(i, e, g, \lambda)$  est le cône "vue" par le pixel  $x, y$ .

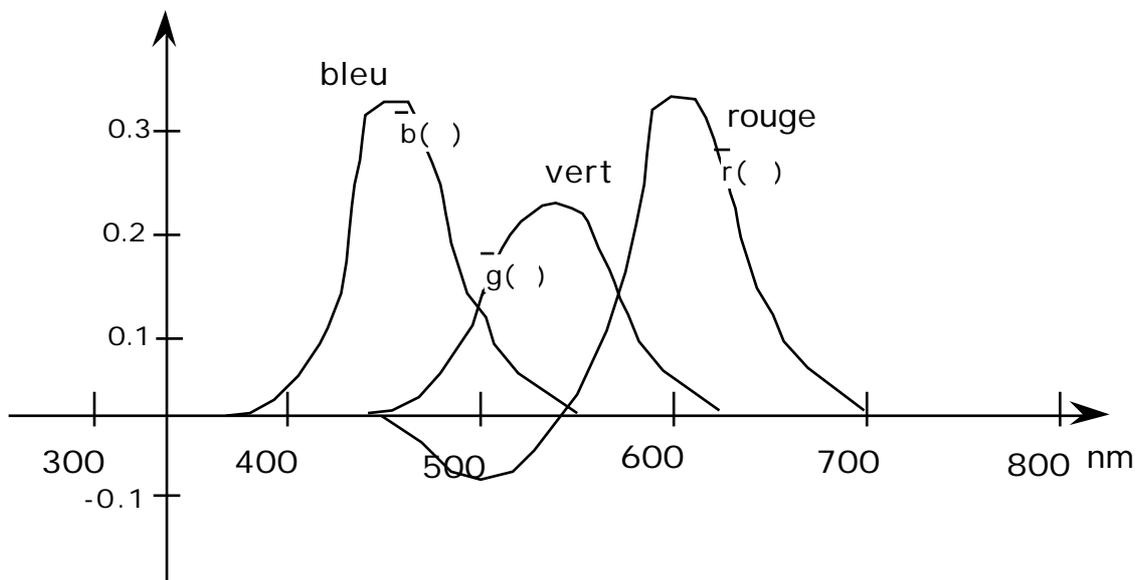
**Caméra couleur "classique" :**



Les filtres suivent un standard défini par la IIC (International Illumination Commission).

Ils sont des fonctions de  $\lambda$  :  $b(\lambda)$ ,  $g(\lambda)$ , et  $r(\lambda)$ .

Note : ils sont composés d'un autre ensemble de fonctions,  $x(\lambda)$ ,  $y(\lambda)$ ,  $z(\lambda)$ . Ceci permet la valeur négative de  $r(\lambda)$ .



Un pixel "couleur" est défini par :

$$P(x,y) = \begin{pmatrix} R(x,y) \\ V(x,y) \\ B(x,y) \end{pmatrix}$$

où :

$$R(x,y) = r_0 \int_0^{\lambda} R(i, e, g, \lambda) S(\lambda) C(\lambda) r(\lambda) d\lambda$$

$$V(x,y) = g_0 \int_0^{\lambda} R(i, e, g, \lambda) S(\lambda) C(\lambda) g(\lambda) d\lambda$$

$$B(x,y) = b_0 \int_0^{\lambda} R(i, e, g, \lambda) S(\lambda) C(\lambda) b(\lambda) d\lambda$$

ou bien :

$$R = r_0 \int_0^{\lambda} R(i, e, g, \lambda) S(\lambda) C(\lambda) r(\lambda) d\lambda$$

$$V = g_0 \int_0^{\lambda} R(i, e, g, \lambda) S(\lambda) C(\lambda) g(\lambda) d\lambda$$

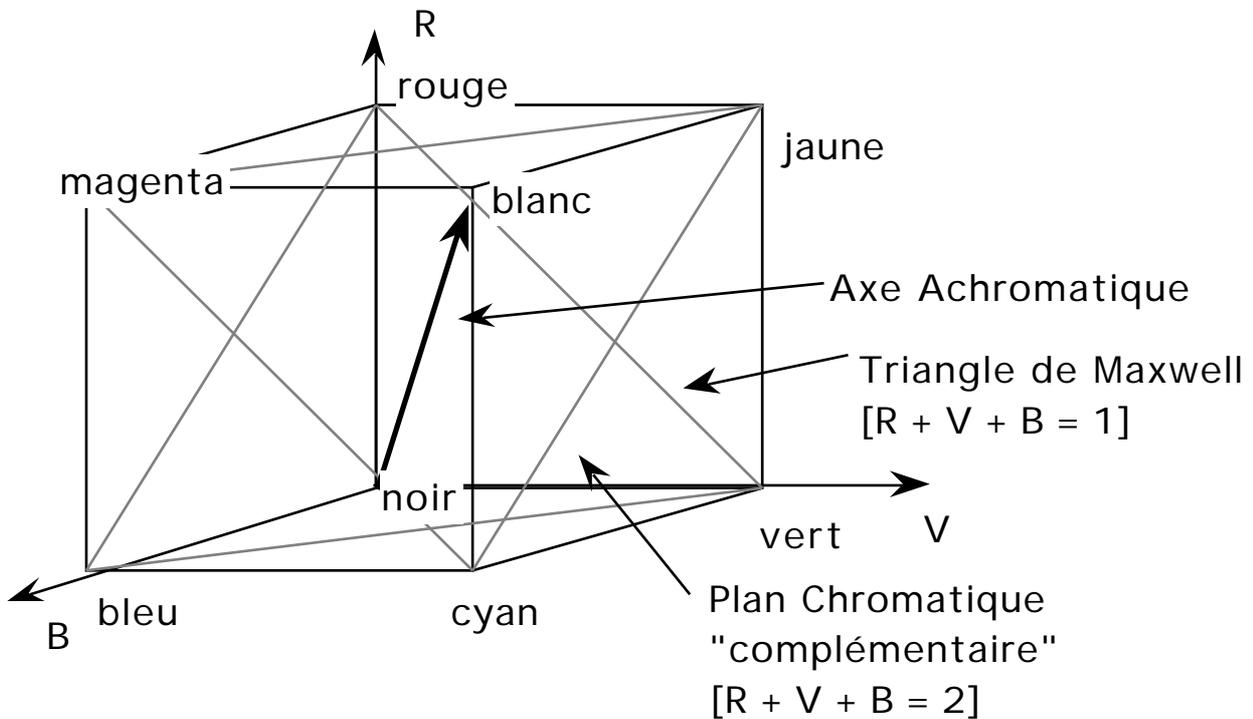
$$B = b_0 \int_0^{\lambda} R(i, e, g, \lambda) S(\lambda) C(\lambda) b(\lambda) d\lambda$$

Donc, la couleur est une projection de la spectre sur les spectres des récepteurs.  
On ne peut pas percevoir la spectre d'une manière directe.

## Les Espaces de la Couleur

### l'espace RVB

Les signaux d'une caméra vidéo nous donnent un espace RVB.  
RVB est un espace ADDITIF.



L'axe achromatique est  $R = V = B$

### l'espace CMY

Une page est imprimé avec les pigments.  
Les pigments sont "soustractifs". Ils absorbent les photons.

On utilise l'espace : CMY : "Cyan-Magenta-Yellow".

$$\begin{array}{l} C \\ M \\ Y \end{array} = \begin{array}{l} R_{\max} \\ V_{\max} \\ B_{\max} \end{array} - \begin{array}{l} R \\ V \\ B \end{array}$$

CMYK est une variation de CMY dans lequel on ajoute du noir (k).

**L'espace YIQ**

Le norme TV NTSC des États-Unis utilise l'encodage Y I Q :

$$\begin{array}{rcccc} Y & 0.30 & 0.59 & 0.11 & R \\ I & = & 0.60 & -0.28 & -0.32 & G \\ Q & & 0.21 & -0.52 & -0.31 & B \end{array}$$

Approximation de :

Y = Luminance

I = rouge - cyan

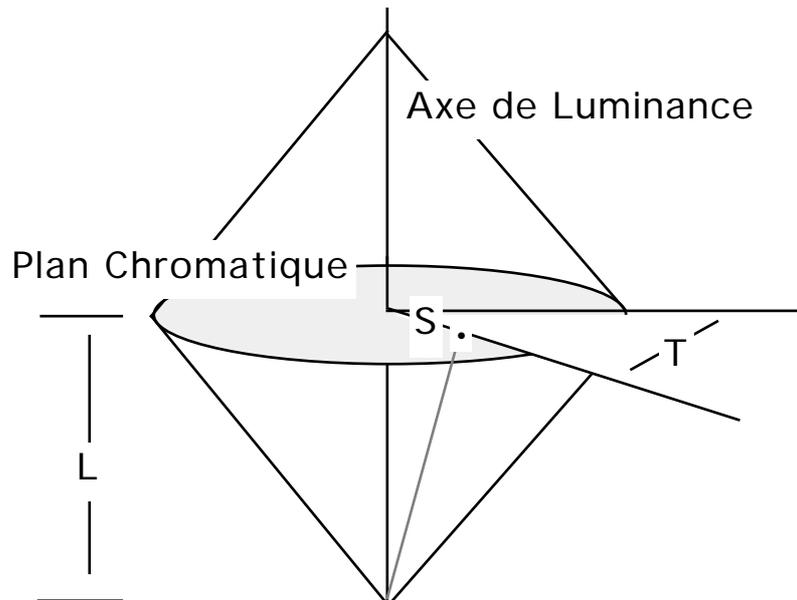
Q = magenta - vert

Les téléviseurs noir et blanc n'affiche que Y.

## L'espace TLS

(en anglais Hue Luminance Saturation - HLS)

Pour mieux modéliser la perception de la couleur, nous pouvons transformer l'espace RVB dans l'espace TLS, composé de teinte, luminance, et saturation :



luminance :  $L = (R + V + B)$

Normalisation des couleurs :

$$r = R / L$$

$$v = V / L$$

$$b = B / L$$

saturation :  $1 - 3 * \min(R, G, B) / L$

teinte : 
$$x = \text{Cos}^{-1} \frac{\frac{1}{2} [(R-V) + (R-B)]}{\sqrt{(R-V)^2 + (R-B)(V-B)}}$$

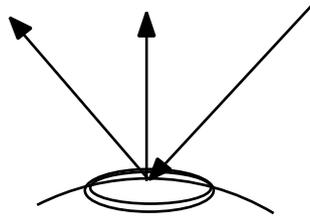
puis si  $b < v$  alors  $T = x$   
 si  $b > v$  alors  $T = 2 - x$

## La Distribution de la Couleur d'une Surface

Qu'est-ce que la couleur d'un reflet ?

Réponse : la couleur de la source.

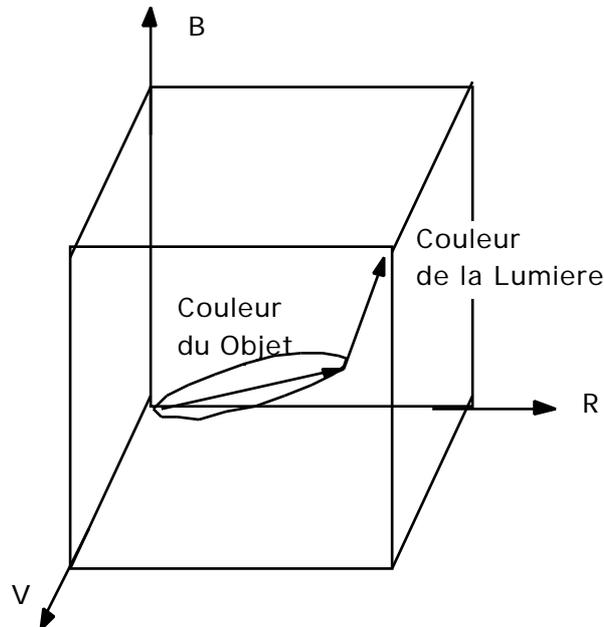
Qu'est-ce qui fait la frontière d'un reflet ?



Le corps d'une surface a la couleur de ses pigments.

Le reflet a la couleur de la source.

La frontière donne une transition :



On peut détecter les deux axes à partir d'une histogramme de couleur.

Une histogramme de couleur est un tableau listant la fréquence d'occurrence pour chaque triple  $(R, G, B)$  dans l'image.

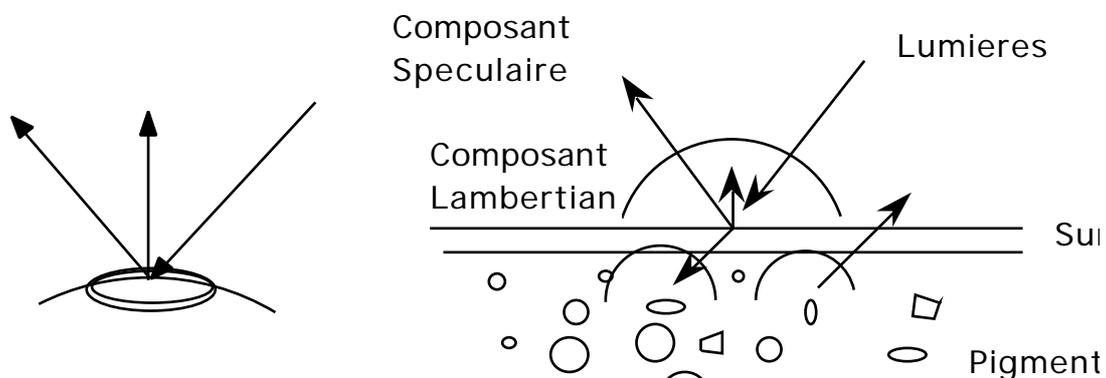
On alloue un tableau 3D de taille  $N_h$  (exemple  $32 \times 32 \times 32 = 32 \text{ K}$  cellules :

$h(R, V, B)$ . Pour chaque pixel  $C = c(i, j) = \begin{matrix} R(i, j) \\ V(i, j) \\ B(i, j) \end{matrix}$ , on incrémente la cellule de

l'historgramme qui correspond à  $(R, V, B)$

$$h(R, V, B) := h(R, V, B) + 1$$

## Analyse Probabiliste au niveau Pixel



Rappelons que l'albédo d'un objet non-métallique peut être approximé par la composition d'une réflexion "spéculaire" et d'une réflexion "lambertienne".

$$R(i, e, g, \lambda) = \rho_s R_s(i, e, g, \lambda) + \rho_l R_l(i, e, g, \lambda)$$

Hors reflets spéculaire, pour une réflexion lambertienne :

$$R_l(i, e, g, \lambda) = p(\lambda) \cos(i)$$

Ceci peut être vu comme un produit de luminance terme  $\cos(i)$  qui module la luminance et  $p(\lambda)$  modulant la chrominance.

Le composant "chrominance"  $p(\lambda)$  est déterminé par la composition du spectre de la source et le spectre d'absorption des pigments de la surface. Si le spectre de la source est constant, la chrominance indique l'identité de l'objet.

Ceci donne un indice pour la détection d'objet au niveau pixel.

## Séparation de Luminance et Chrominance

L'axe luminance, L, peut être défini par

$$L = R + V + B$$

La chrominance  $C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$  est une signature pour l'objet.

La chrominance peut être définie par plusieurs codages.

Par exemple, pour la détection de la peau, il est fréquent de voir Normalisation par la luminance laisse deux axes chromatiques : r, v

$$c_1 = r = \frac{R}{R+V+B} \quad c_2 = v = \frac{V}{R+V+B}$$

Une autre codage fréquente est un codage en "couleur opposée", par exemple :

$$L = \frac{R+G+B}{3}, \quad c_1 = \frac{R-G}{2} \quad \text{et} \quad c_2 = \frac{R+G}{2} - B.$$

$$\begin{array}{rcccl} L & & 0.33 & 0.33 & 0.33 & R \\ C_1 & = & 0.5 & -0.5 & 0 & G \\ C_2 & & 0.5 & 0.5 & -1 & B \end{array}$$

En tout cas, on peut apprendre et détecter la signature de la chrominance avec une manière statistique.

## Histogrammes de Couleur

Un histogramme est une table de fréquence. Il peut fournir une estimation d'une densité de probabilités.

On peut utiliser les histogrammes pour calculer la probabilité qu'un pixel est une projection d'un objet.

Par exemple, construisons un histogramme pour le vecteur de chrominance  $(r, v)$ .

Chaque pixel est un vecteur  $(r, v)$  :  $C(i, j) = \begin{pmatrix} r \\ v \end{pmatrix} (i, j)$

On alloue un tableau 2D de taille  $N_h$  (exemple  $32 \times 32 = 1024$  cellules :  $h(r, v)$ ).

Pour chaque pixel  $C = C(i, j)$  dans l'image, on incrémente la cellule de l'histogramme qui correspond à  $C = (r, v)$

$$h(r, v) := h(r, v) + 1 \quad \text{c'a-dire} \quad h(C) := h(C) + 1$$

Soit  $M$  Pixels dans l'image. Un histogramme des chrominances,  $h(C)$ , des  $M$  pixels dans une image donne leurs fréquences d'occurrence.

$$P(C) = \frac{1}{M} h(C)$$

Pour que l'estimation soit "raisonnable", il faut assurément que  $M \gg N_h$

Considérons une région  $W$  de  $M_o$  pixels de la même image correspondant à l'objet  $O$ .

$$(i, j) \in W : h_o(C(i, j)) := h_o(C(i, j)) + 1$$

Ensuite: pour tout pixel  $C(i, j) = \begin{pmatrix} r \\ v \end{pmatrix} (i, j)$  :  $p(C | \text{objet}) = \frac{1}{M_o} h_o(C)$

Parce que  $W$  est dans l'image, la probabilité de rencontrer un pixel de  $W$ ,

$$P(W) = \frac{M_o}{M}$$

## Détection par ratio d'histogramme

L histogramme permet d'utiliser la règle de Bayes afin de calculer la probabilité qu'un pixel corresponde à un objet.

Pour chaque pixel  $C(i, j)$   $p(\text{objet} | C) = p(C | \text{objet}) \frac{p(\text{objet})}{p(C)}$

Soit  $M$  images de  $I \times J$  pixels. Ceci fait  $N = I \times J \times M$  Pixels.

Soit  $h(r, v)$ , l'histogramme de tous les  $N$  pixels.

Soit  $h_o(r, v)$ , l'histogramme des  $N_o$  pixels de l'objet "o".

$$p(\text{objet}) = \frac{M_o}{M}$$

$$p(C) = \frac{1}{M} h(C)$$

$$p(C | \text{objet}) = \frac{1}{M_o} h_o(C)$$

$$\text{Donc } p(\text{objet} | C) = p(C | \text{objet}) \frac{p(\text{objet})}{p(C)} = \frac{1}{M_o} h_o(C) \frac{\frac{M_o}{M}}{\frac{1}{M} h(C)}$$

$$p(\text{objet} | C) = \frac{h_o(C)}{h(C)}$$

Par exemple, voici une image de la probabilité de peau fait par ratio d'histogramme de  $r, v$



## Caractérisation par moments

Les ensemble connexes de pixels s'appelles les "blobs".

On peut décrire une blob par une vecteur de caractéristiques "invariantes" à l'orientation grâce aux "moments"

Les moments sont invariants aux transformations affines.

Pour une febetre (imagette)  $w(i, j)$  de taille  $N \times M$

$$\text{Somme des Pixels :} \quad S = \sum_{i=1}^M \sum_{j=1}^N w(i, j)$$

Premiers moments :

$$\mu_i = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot i \quad \mu_j = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot j$$

Le premier moment est le centre de gravité de la forme :

Deuxième moment :

$$i^2 = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N (w(i, j)) \cdot (i - \mu_i)^2$$

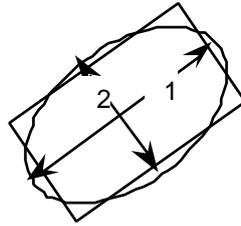
$$j^2 = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot (j - \mu_j)^2$$

$$ji^2 = \frac{1}{S} \sum_{i=1}^M \sum_{j=1}^N w(i, j) \cdot (i - \mu_i)(j - \mu_j)$$

Ceci permet de définir les "axes", majeur,  $\mu_1$  et mineur,  $\mu_2$ , de la forme par analyse des composantes principales de la deuxième moment

$$C_o \cong \begin{pmatrix} i^2 & ij^2 \\ ij^2 & j^2 \end{pmatrix}$$

## Composantes principales



Les deuxièmes moments sont "invariants" à l'orientation

Les axes sont calculés par une analyse en composantes principales de la matrice  $C$ . Il s'agit de trouver une rotation,  $\Phi$ , dans l'espace de caractéristiques  $\Phi C_P \Phi^T = \Lambda$  telles que  $\Lambda$  soit diagonale.

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{tel que } 1 > 2 \quad \Phi = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$\Phi C_P \Phi^T = \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \Phi^T \Phi = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Phi C_P \Phi^T \Phi = \Phi C_P = \Lambda \Phi = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Les lignes du  $\Phi$  sont des vecteurs propres du  $C$ .

La longueur des axes majeurs et mineur est les valeurs propres de la matrice  $C$ .

$\theta$  est l'orientation de l'axe "majeur" et  $1/2$  est le rapport entre la longueur et la largeur.

$1/2$  est une caractéristique invariante de la taille et de l'orientation.