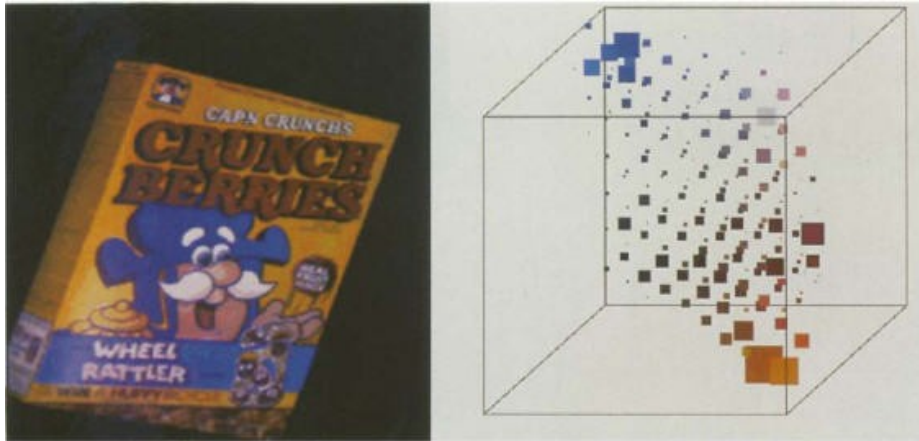


Reconnaissance d'objets couleur et suivi

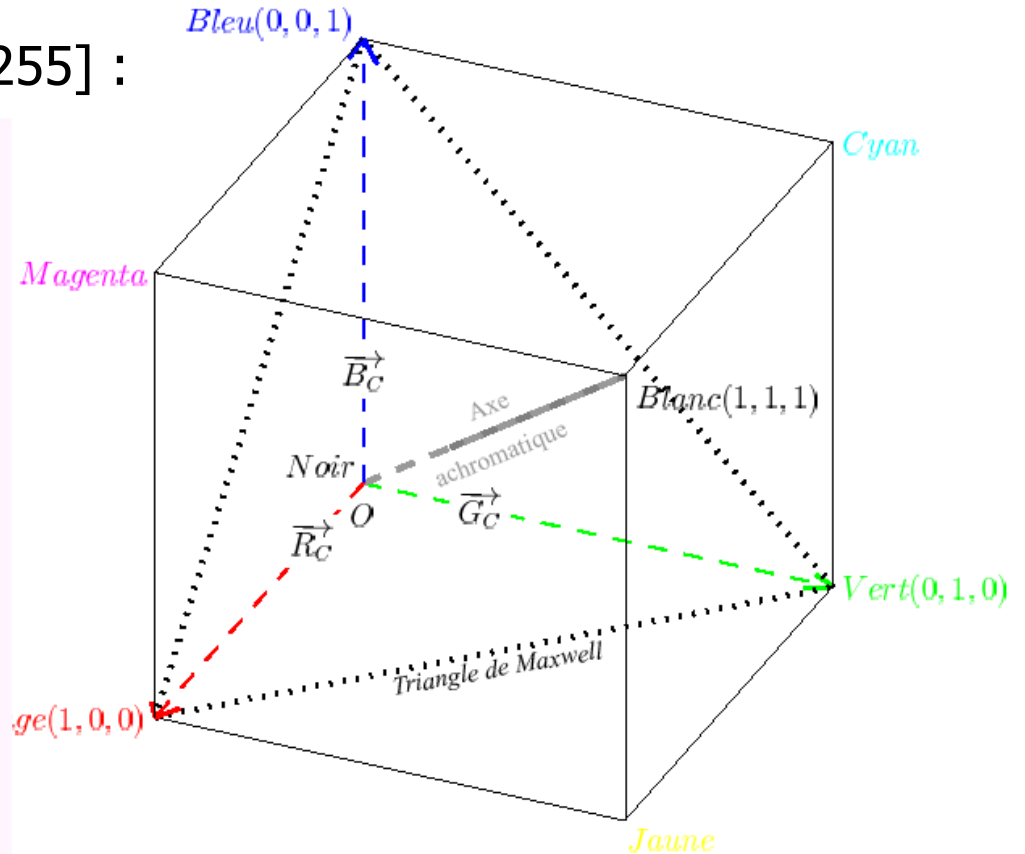
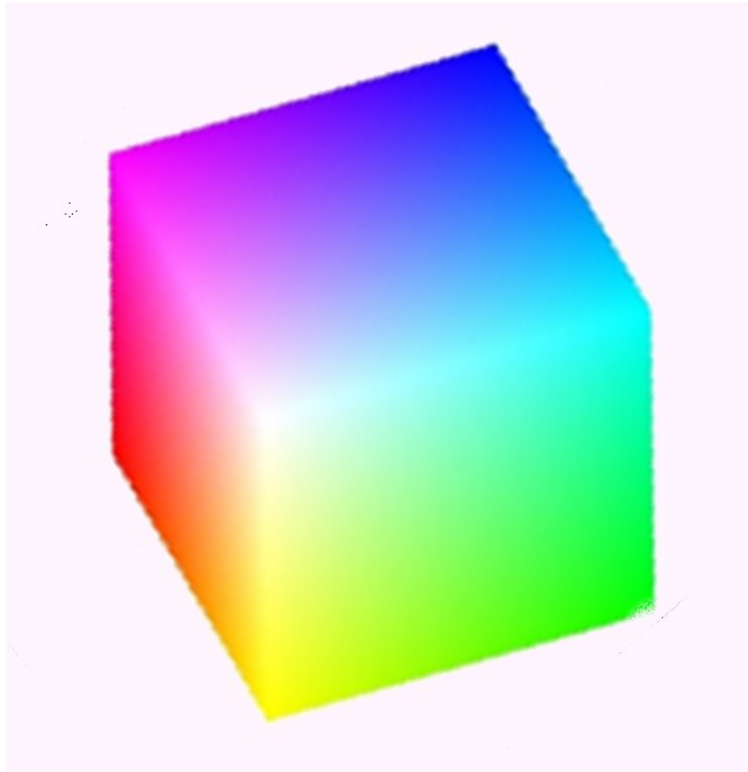
Histogrammes couleur

- Détection, segmentation (et suivi) d'objets par la couleur.
- Méthode classique : Swain and Ballard '91

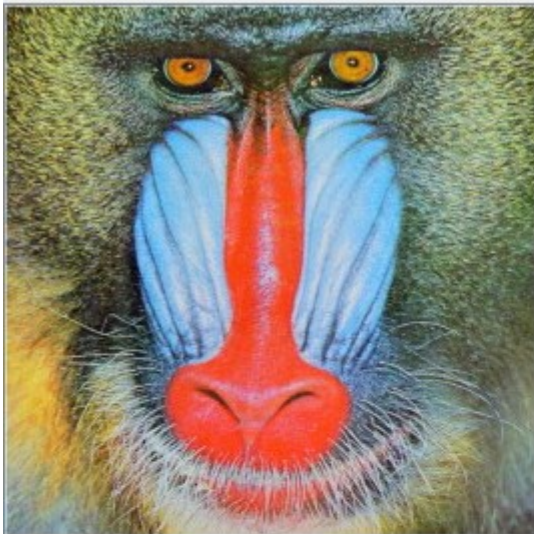


RGB (RVB)

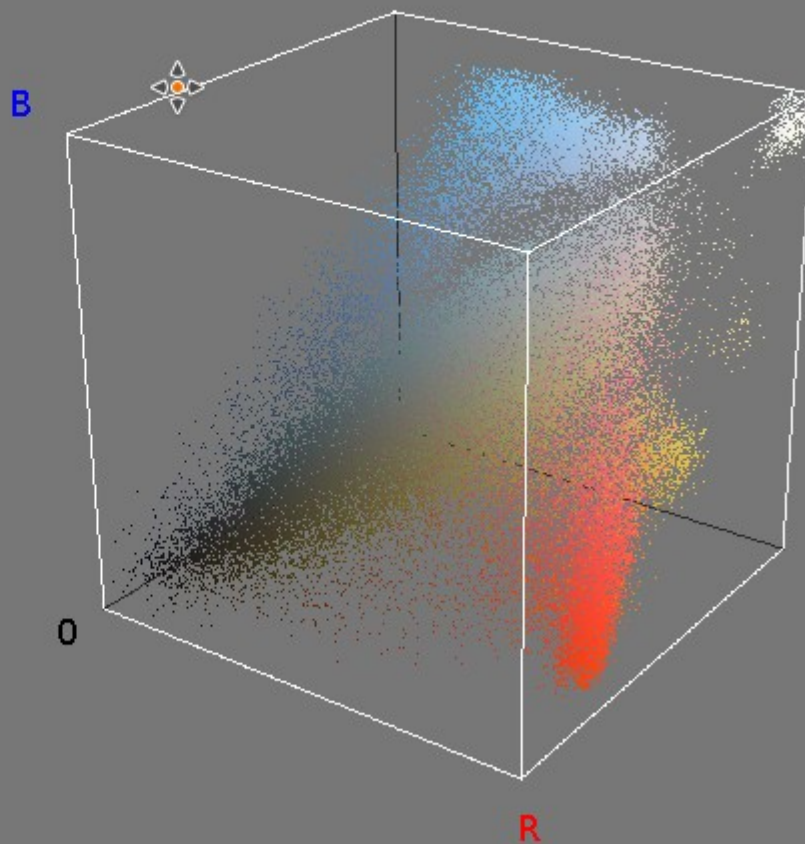
Cube des couleurs ([0,1] ou [0 255] :



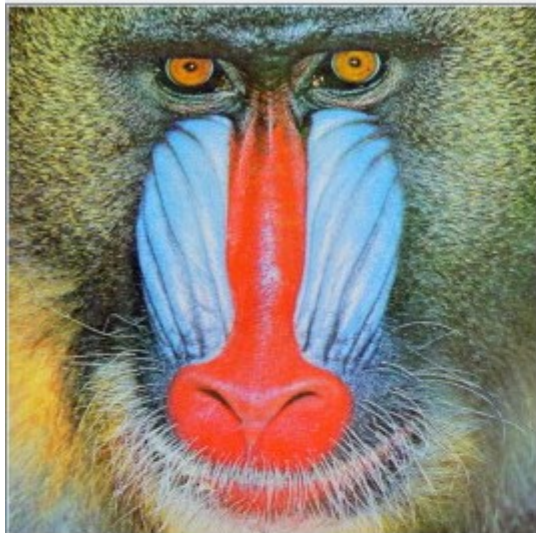
71022 Pixels, 58124 Colors



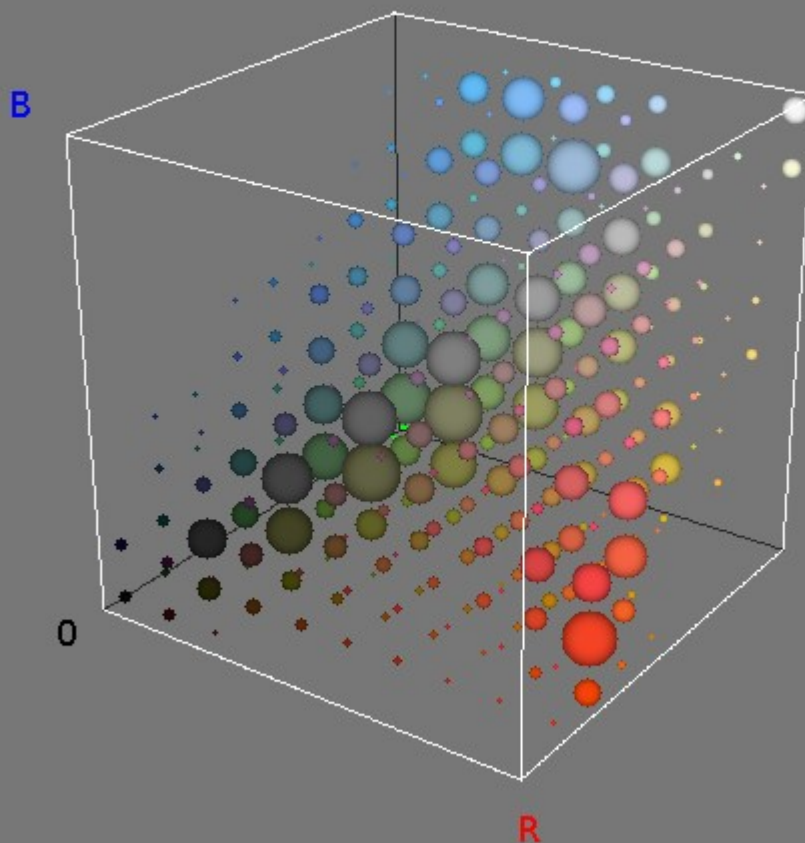
Pixels



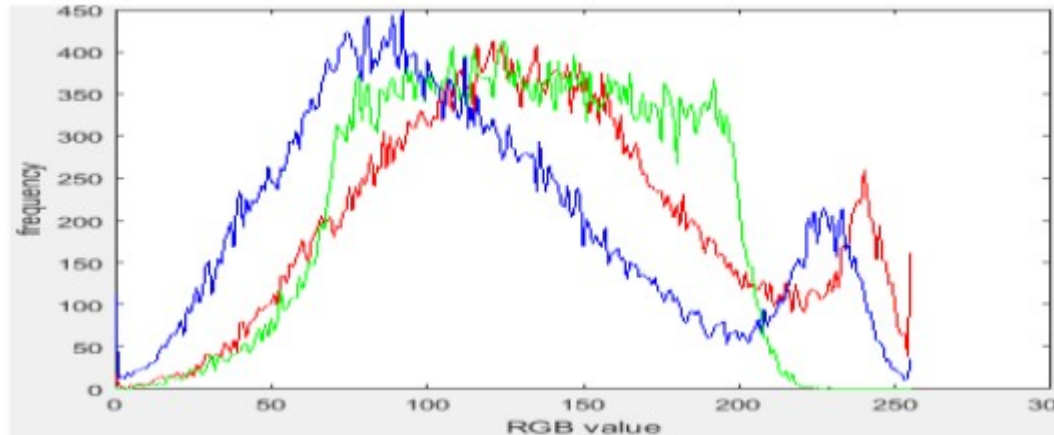
71022 Pixels, 58124 Colors



Histogramme



- Le plus souvent on utilise les histogrammes marginaux : perte de l'information de corrélation



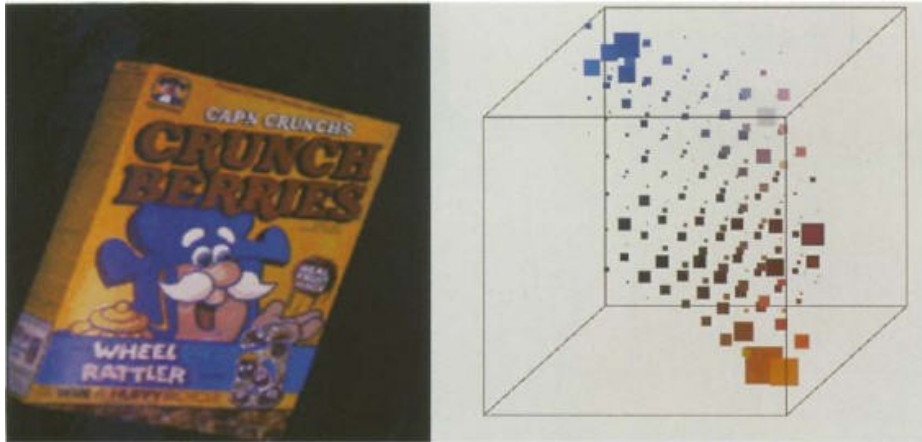
Histogrammes couleur

Color indexing

MJ Swain, DH Ballard

International journal of computer vision 7 (1), 11-32, 1991

- Histogrammes r,g,b :

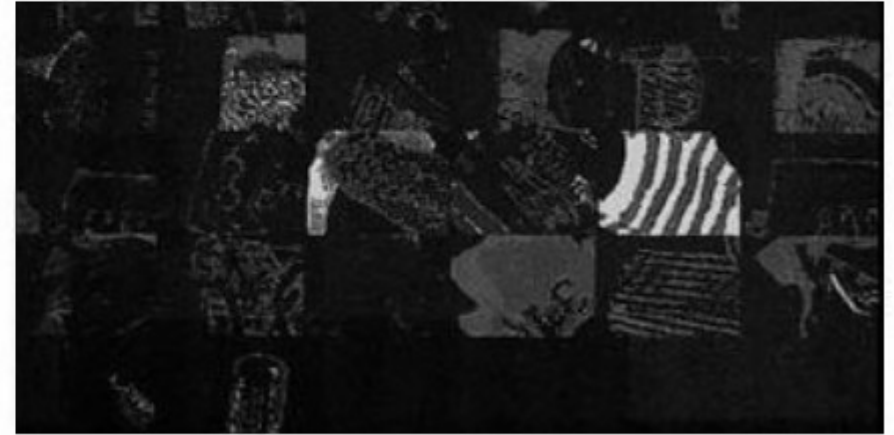


- Idée : calculer l'histogramme d'un objet de référence puis comparer aux histogrammes d'autres objets.

Algorithme de rétroprojection de l'histogramme

- Calculer l'histogramme M de l'objet de référence :
 - M_i = fréquence de la $i^{\text{ème}}$ couleur
- Calculer l'histogramme I de l'image à examiner
- Calculer $R_i = \min(M_i/I_i, 1)$
 - R_i indique l'importance de la couleur i pour l'objet de référence.
- On remplace chaque couleur i par R_i
- L'image résultat est convolée avec un masque circulaire de la taille de l'objet recherché (à voir...) :
- Les pics indiquent la présence (probable) de l'objet.

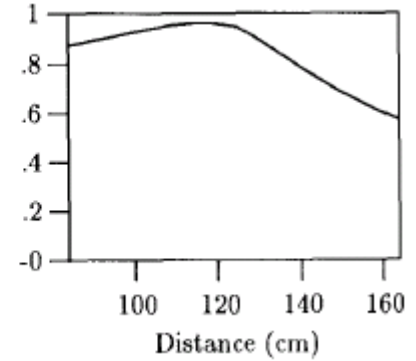
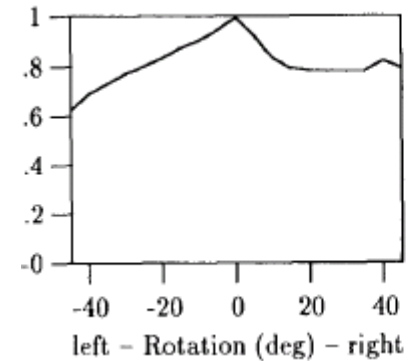
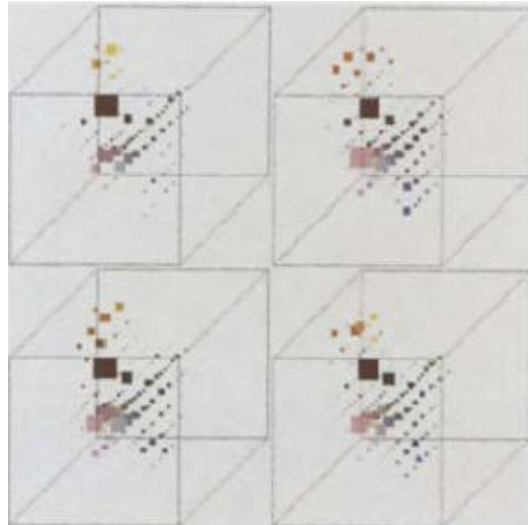
Object Localization



Résultat pour la recherche du pull bleu

Attention : niveaux de gris dans l'image résultat d'autant plus clairs qu'il est probable que le pixel appartienne au pull bleu.

- Relativement tolérant aux changements géométriques (rotation 3D), occultations partielles etc:



Valeur du pic entre les différentes images et l'original

- Excellent resultats pour 30 objets (...) :

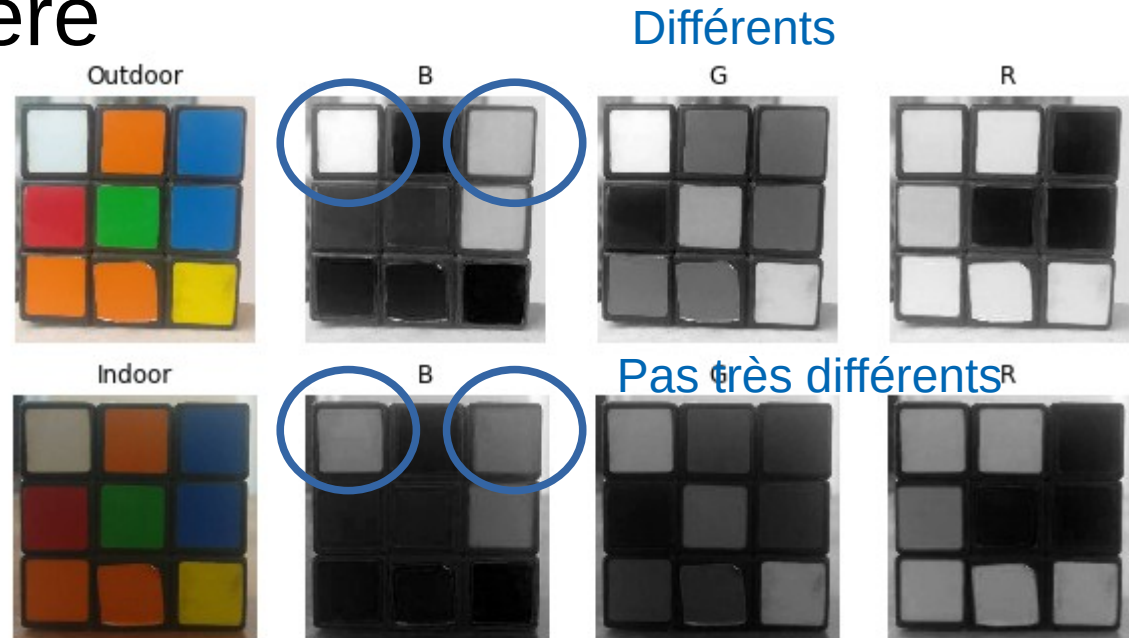


Image Size	Correct Match Placement				Average Match Percentile
	1st	2nd	3rd	> 3rd	
128 x 90	29	3	0	0	99.9
64 x 45	27	5	0	0	99.8
32 x 22	24	7	1	0	99.6
16 x 11	15	6	4	0	97.8
8 x 5	4	4	3	21	78.1

Mais dans des conditions irréalistes. En pratique, RGB est mal adapté aux conditions réelles...

- Corrélation/lumière

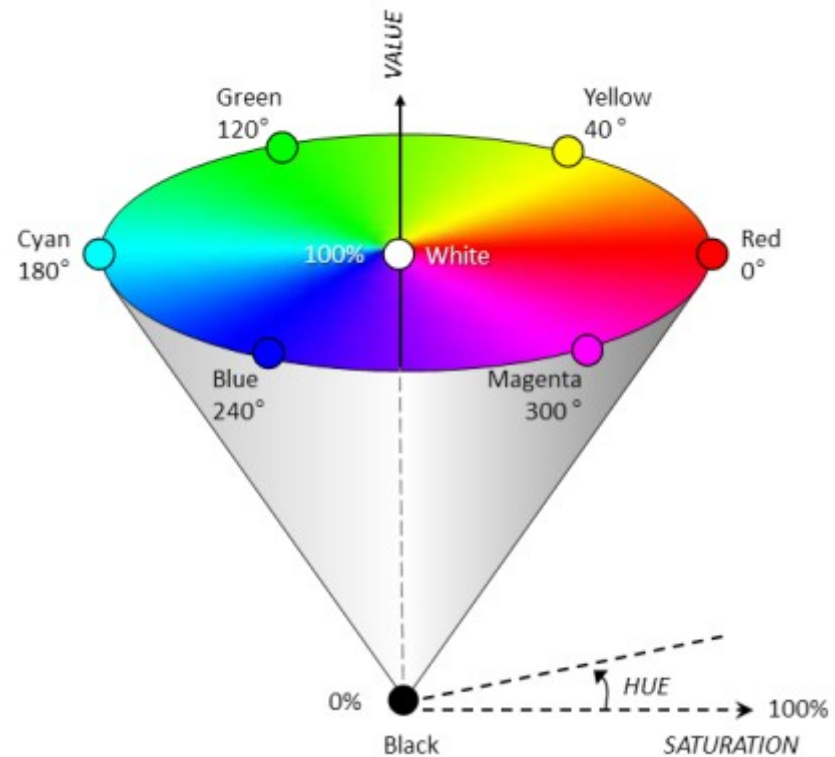
Exemple :



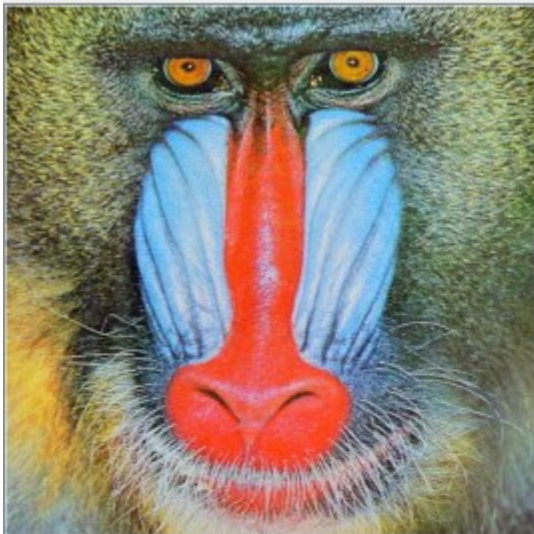
- Non uniformité perceptuelle
- Mélange chrominance et luminance
- Mal adapté à la segmentation

HSV

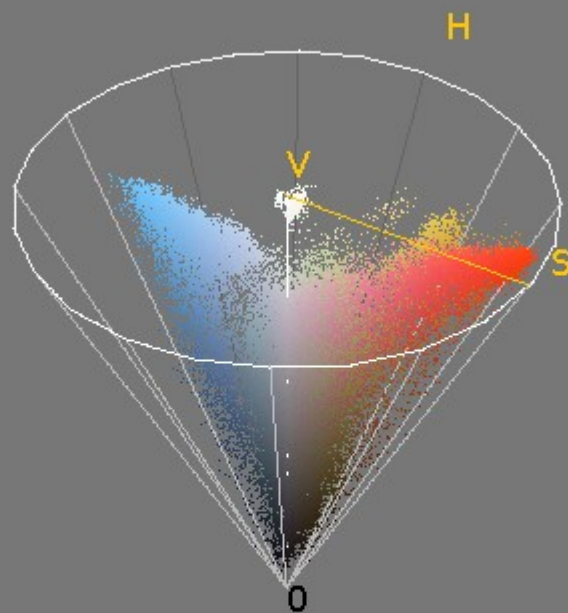
- Hue :
 - teinte
- Saturation :
 - intensité de la couleur
- Value :
 - brillance



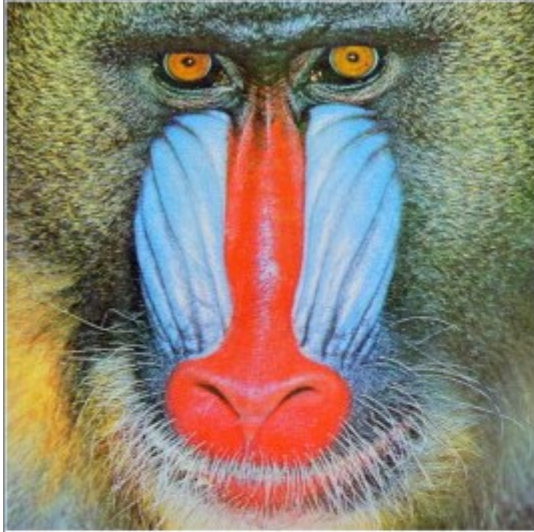
71022 Pixels, 58124 Colors



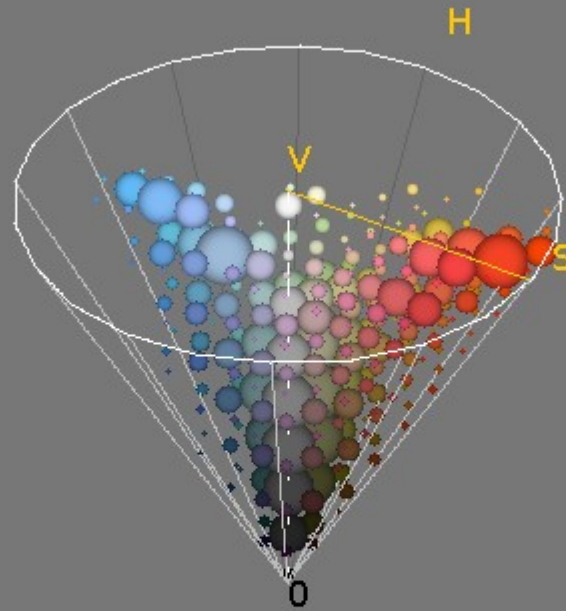
Pixels



71022 Pixels, 58124 Colors



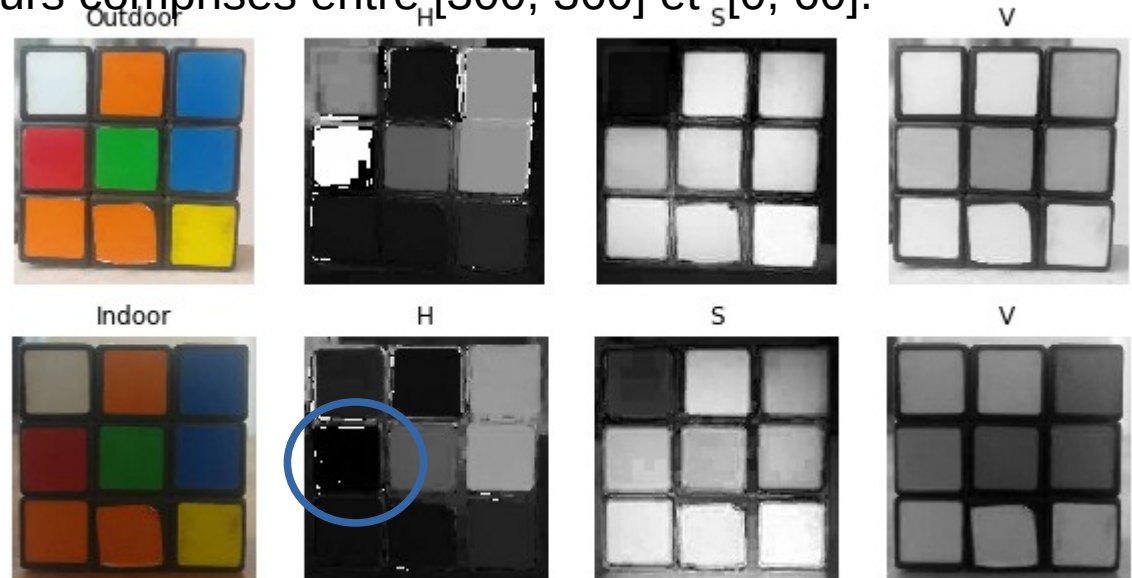
Histogramme



- H est très similaire dans les deux images : informations de couleur (relativement) intactes, même en cas de modification de l'éclairage.
- S est également très similaire dans les deux images.
- V capture la quantité de lumière : change en fonction des changements d'éclairage.

Pour le rouge, H est une unité circulaire avec le rouge = 0.

Donc, il peut prendre des valeurs comprises entre $[300, 360]$ et $[0, 60]$.

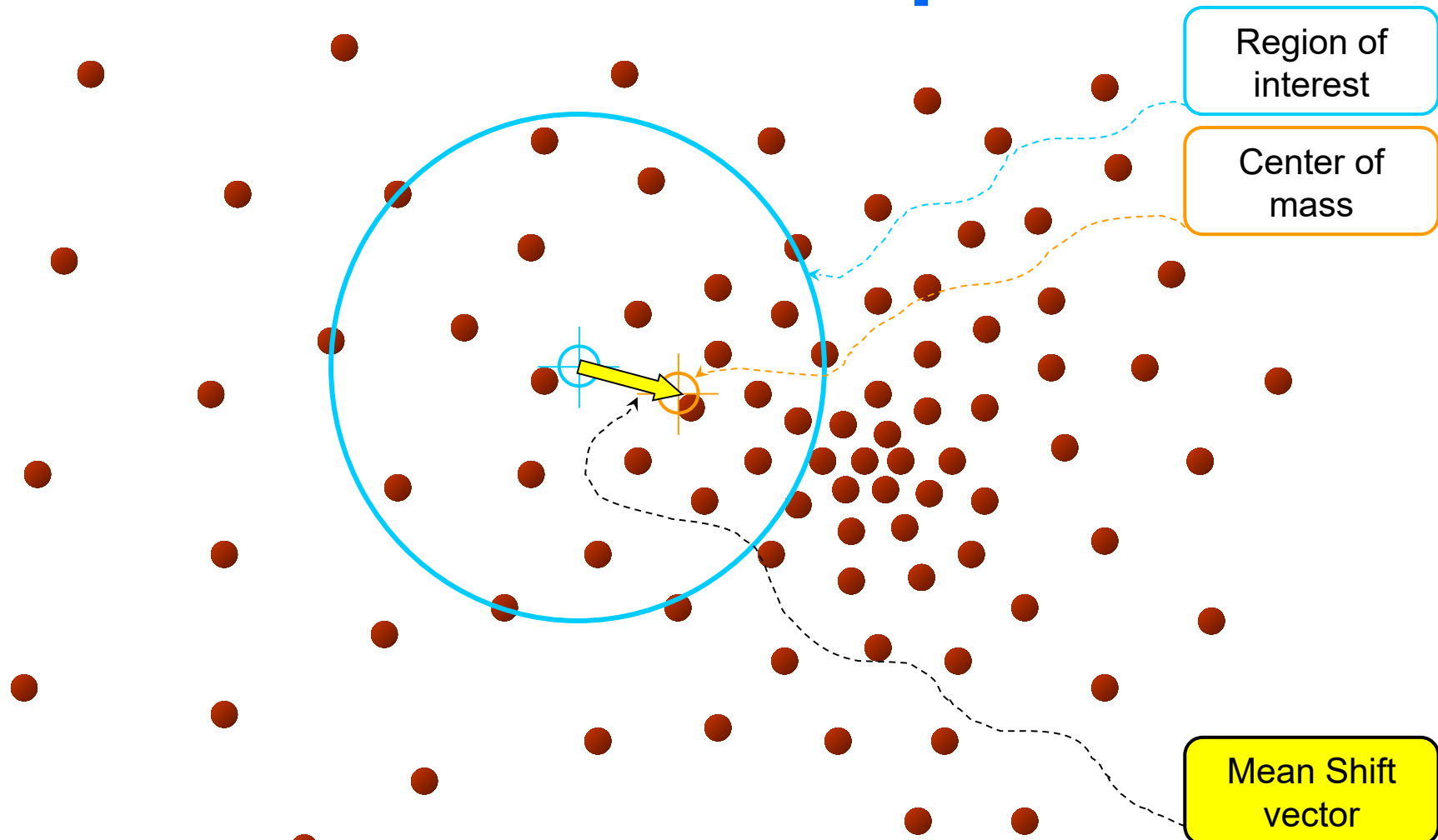


- Pour obtenir des histogrammes (relativement) invariants aux changements d'éclairage on utilise donc souvent :
 - H : Histogramme 1D
 - H et S : Histogramme 2D
- La rétroprojection de l'histogramme dans H ou (mieux) est efficace sur les teintes « chair » : $p(peau|x,y)$
- Algo. classique de détection et suivi de visage utilisé dans les webcams :
 - **Camshift**

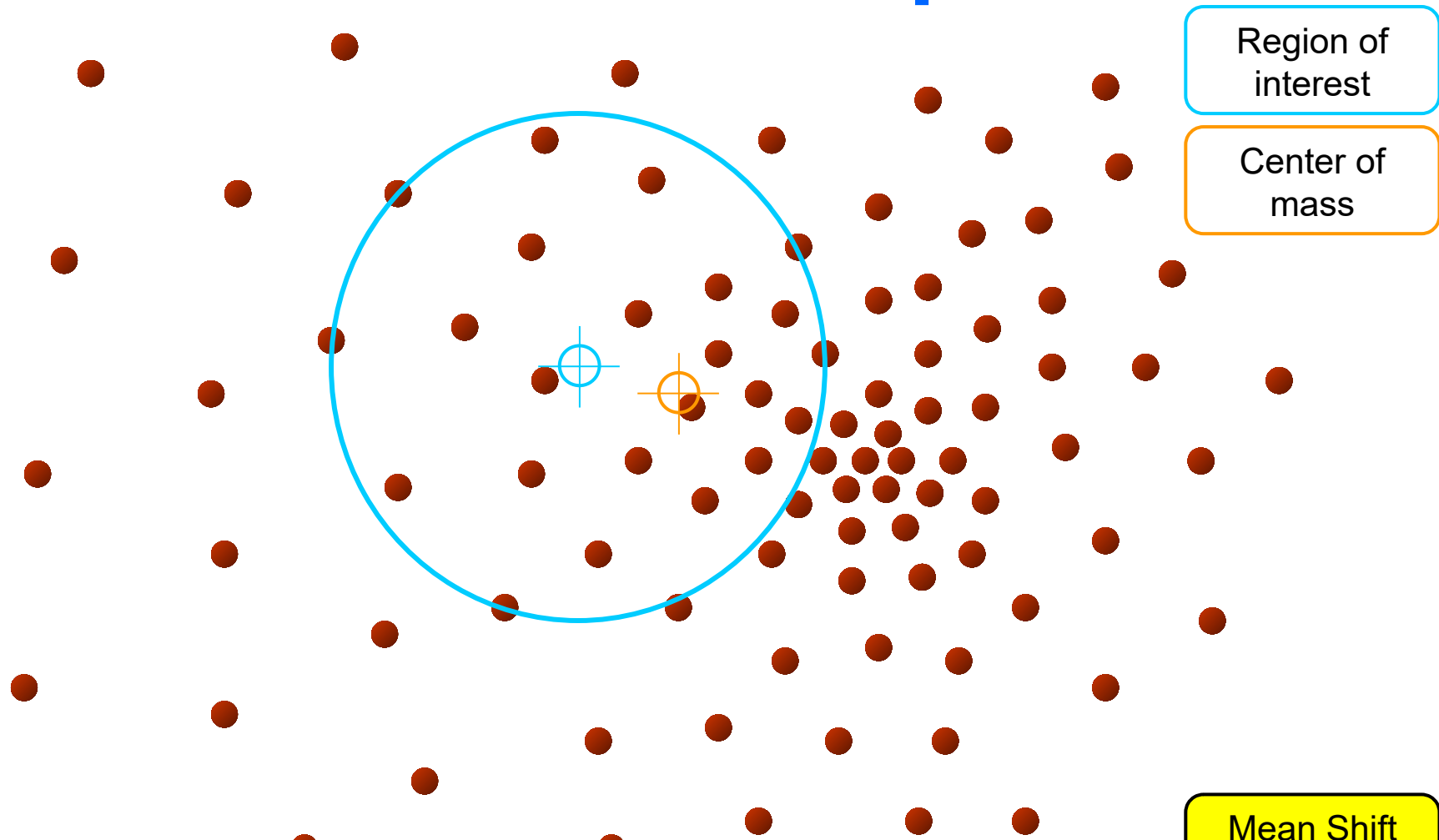
Mean-shift – motivation et intuition

- Recherche de la zone la plus dense de points.
- Exemple 2D case (cf. prochain slide):
 - On part d'un point quelconque
 - On place un cercle centré sur ce point (« région d'intérêt »)
 - A chaque itération on calcule le centre de gravité (moyenne) de la région d'intérêt.
 - On déplace le centre vers le centre de gravité
 - Continuer jusqu'à convergence

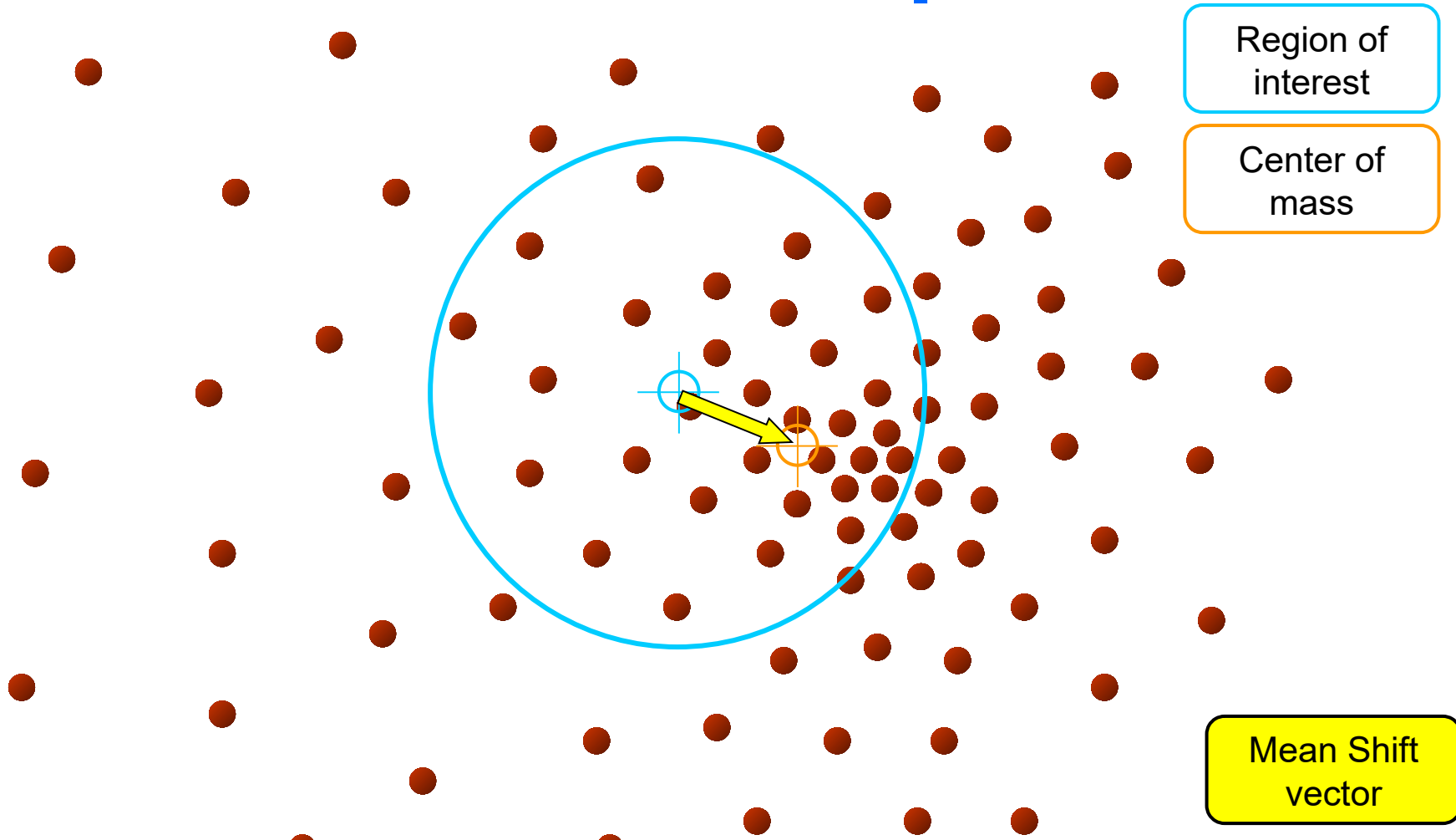
Intuitive Description



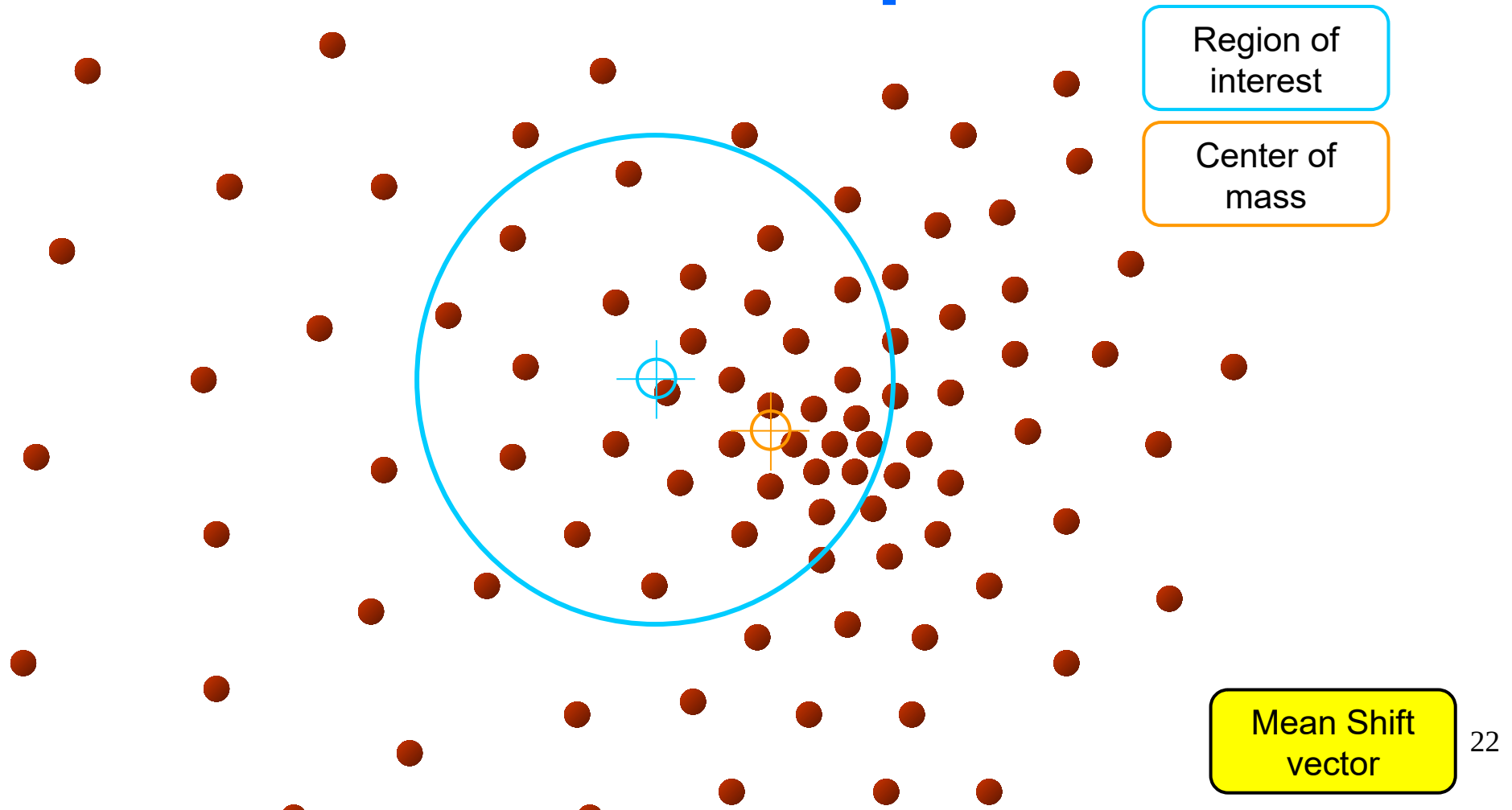
Intuitive Description



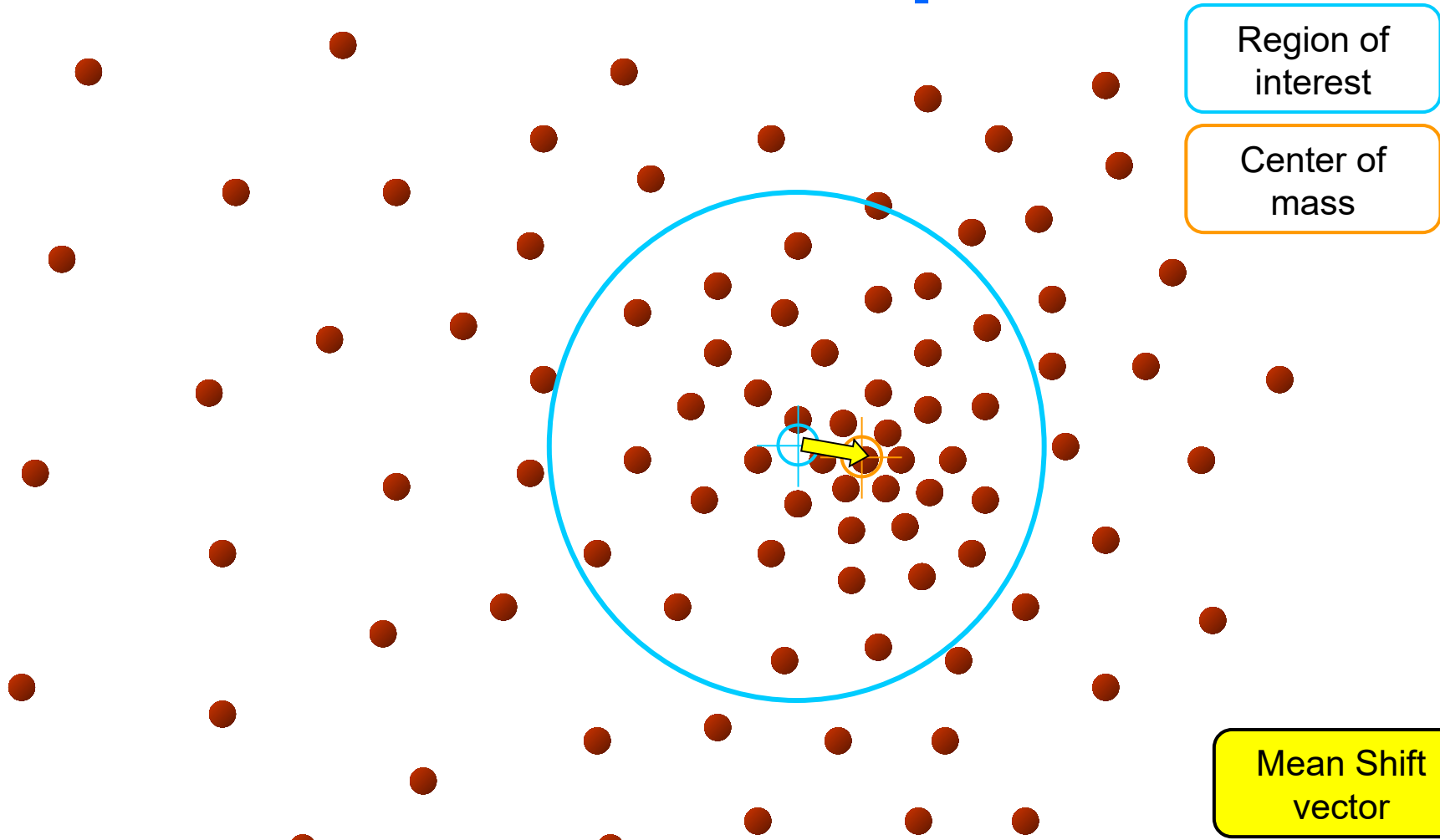
Intuitive Description



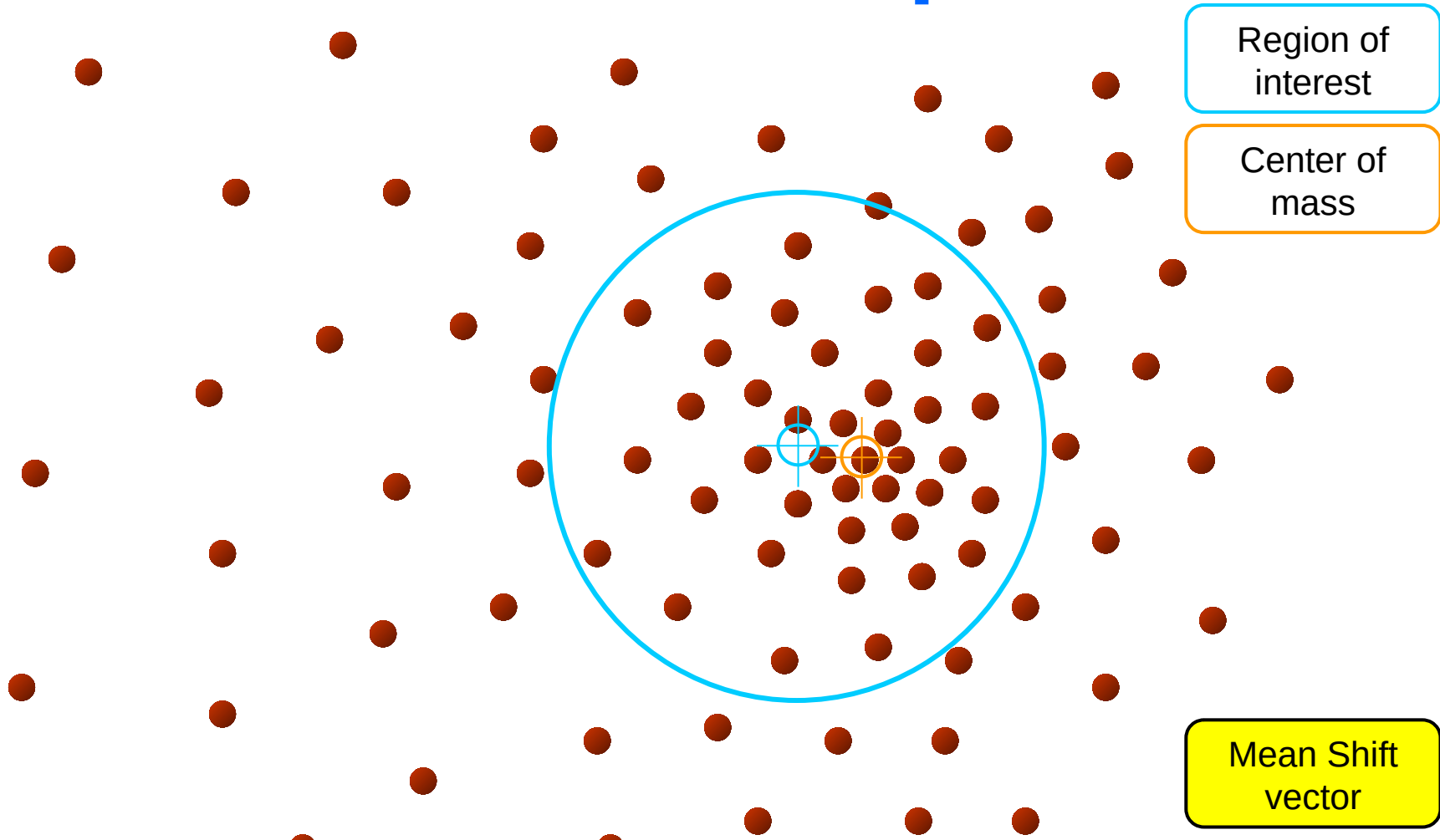
Intuitive Description



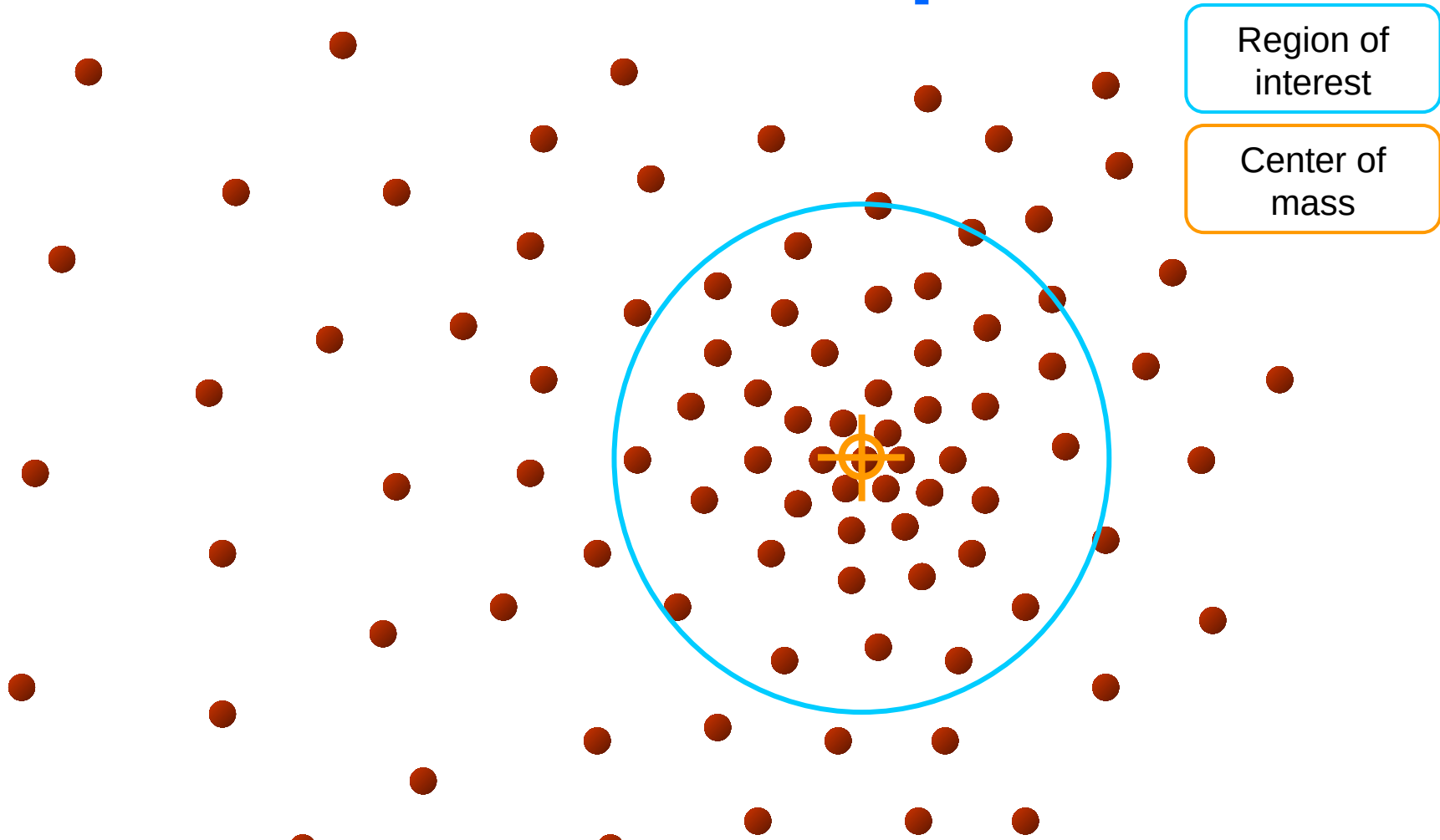
Intuitive Description

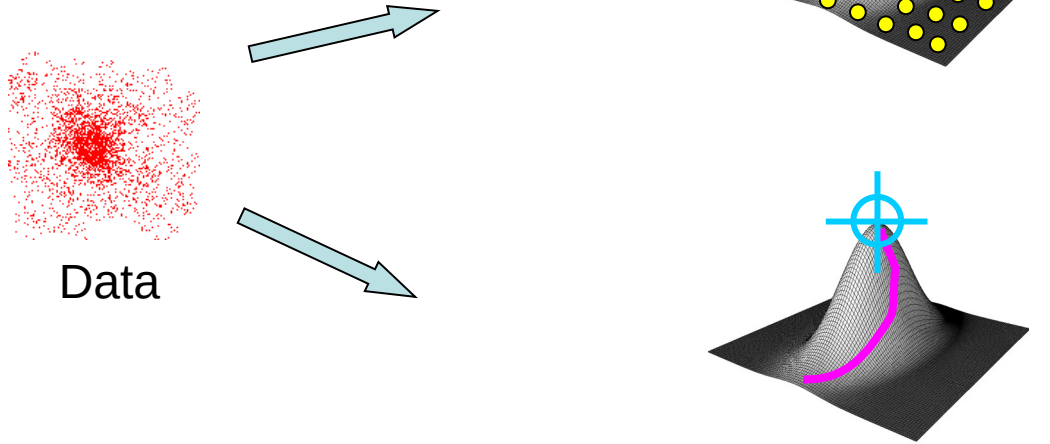


Intuitive Description

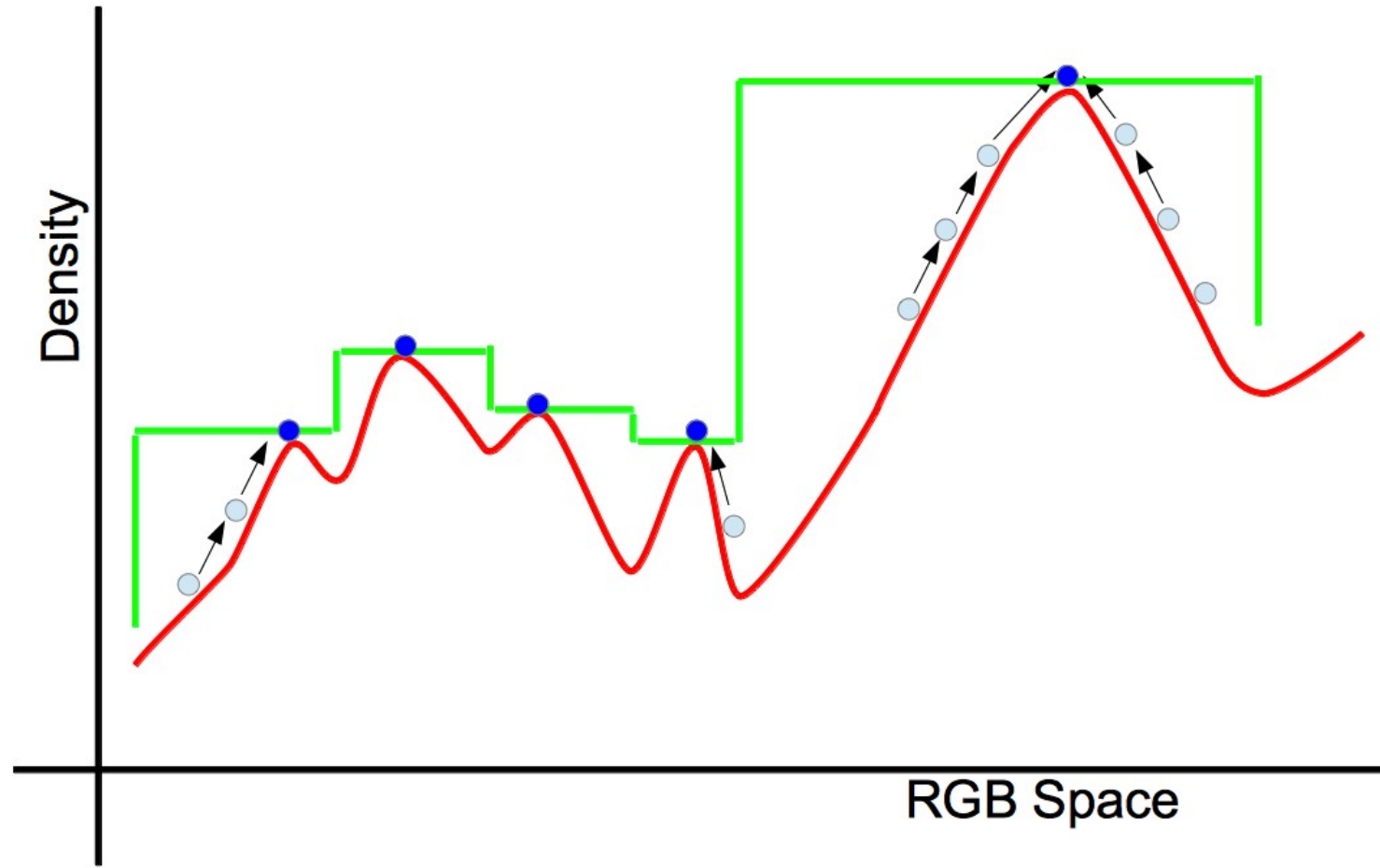


Intuitive Description





- Non-parametrique – pas d'hypothèse sur la densité (e.g. distribution normale)

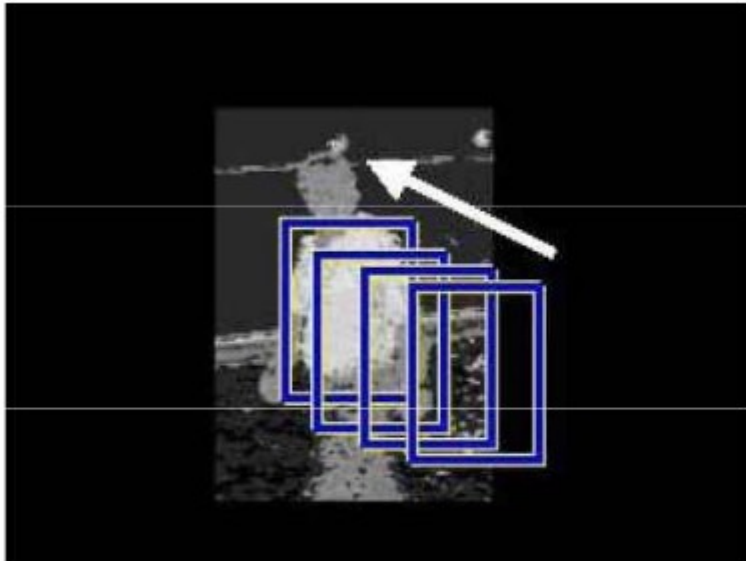


Mean-shift – algorithm

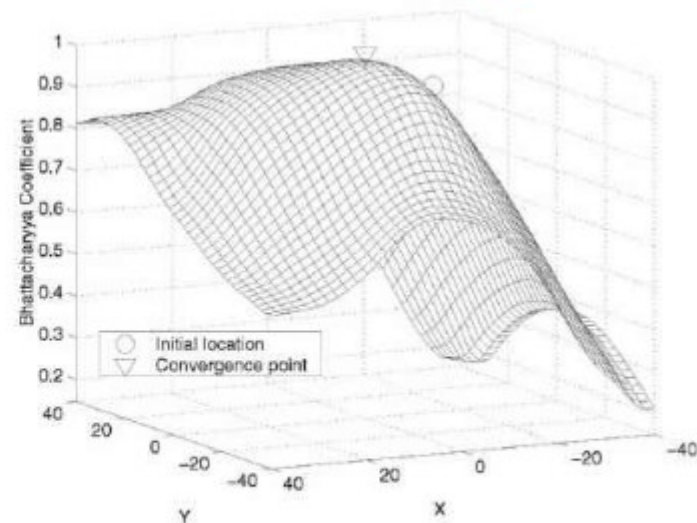
D. Comaniciu, P. Meer: Mean Shift Analysis and Applications,
IEEE Int. Conf. Computer Vision (ICCV'99), Kerkyra , Greece ,
1197-1203, 1999

Continuously Adaptive Mean-SHIFT (CAMSHIFT) Tracking

- Meanshift appliqué sur l'algo. de rétroprojection



@R. Collins CVPR 2003



@Comaniciu PAMI 2003

- On initialise l'objet à suivre (ex : visage)
- On calcule la rétroprojection (« teinte chair ») sur l'image courante.
- On applique le meanshift pour trouver le maximum
- On passe à l'image suivante en initialisant la région d'intérêt à la position (estimée) de l'objet à l'image courante



Algorithme du meanshift

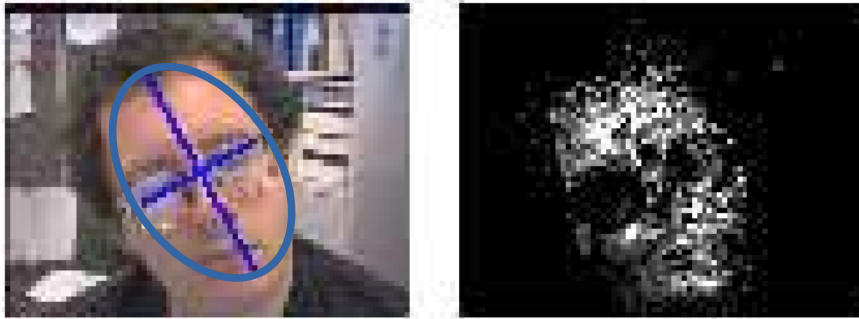
- Soit I l'image de rétroprojection de l'histogramme. $I(x,y)$ est la valeur au pixel (x,y)
- Initialisation de la fenêtre w centrée en (x_c, y_c) , « au hasard » (par exemple au centre de l'image;)
- Tant que w est modifiée :
 - Calculer le barycentre des pixels de la fenêtre w :

$$M_{00} = \sum_{x,y} I(x,y) \quad M_{10} = \sum_{x,y} xI(x,y) \quad M_{01} = \sum_{x,y} yI(x,y) \quad \Rightarrow x_c = \frac{M_{10}}{M_{00}}, \quad y_c = \frac{M_{01}}{M_{00}}$$

- La notation $\sum_{x,y} I(x,y)$ correspond à la somme des pixels de w .
- Déplacer w en (x_c, y_c)

Algorithme *Continuously Adaptive Mean Shift* (CAMSHIFT)

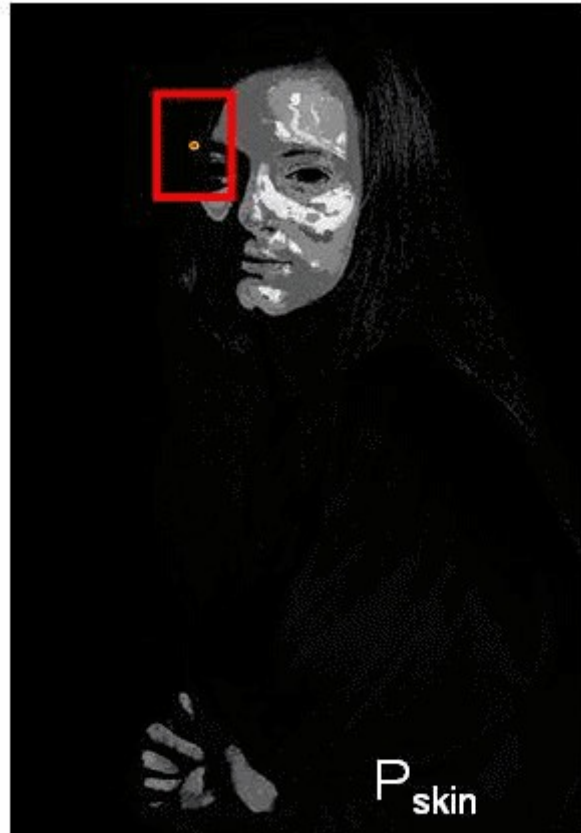
- Adaptation et amélioration à la vidéo –
estimation de l'orientation de la tête (*head roll*)



Bradski, G.R., "Real time face and object tracking as a component of a perceptual user interface," Applications of Computer Vision, 1998. WACV '98.

CAMSHIFT pour le suivi video

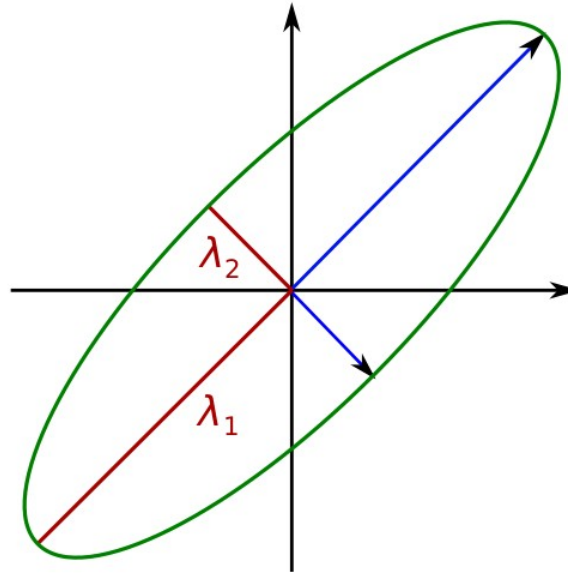
- Principe
 - le meanshift est appliqué à l'image, avec une taille de fenêtre fixe et jusqu'à convergence, puis la taille de w est modifiée pour l'image suivante.
 - Les nouvelles dimensions de w sont définies à partir de **l'ellipse** déduite des **moments centraux du second ordre** de la rétroprojection de l'histogramme (teinte chair), calculés sur la fenêtre w produite par le meanshift agrandie (exemple + 5 pixels)
- Algorithme :
 - Initialiser la fenêtre w
 - Calculer **I (rétroprojection de l'histogramme)** sur une **région de recherche centrée sur w mais plus grande** (ex: +20%).
 - Appliquer le meanshift jusqu'à convergence.
 - Augmenter la taille de w (ex : **+ 5 pixels**) pour calculer l'ellipse puis calculer la nouvelle taille de w
 - Recommencer pour l'image suivante.



Mean shift window
initialization

Calcul de l'ellipse

Le principe est de calculer la matrice de covariance de l'image rétro-projetée (ou de sa version binarisée) et de calculer ses **vecteurs propres** et **valeurs propres** qui donnent l'orientation et l'allongement du nuage de points.



Le centre (x_c, y_c) de l'image est donné par : $x_c = M_{10} / M_{00}$ et $y_c = M_{01} / M_{00}$ où

$$M_{00} = \sum_x \sum_y I(x, y)$$

$$M_{10} = \sum_x \sum_y x I(x, y)$$

$$M_{01} = \sum_x \sum_y y I(x, y)$$

Pour trouver l'angle θ par rapport à l'horizontale, la longueur l et la largeur L de l'ellipse on effectue les calculs suivants :

$$a = M_{20} / M_{00} - x_c^2$$

$$b = 2 * (M_{11} / M_{00} - x_c \cdot y_c)$$

$$c = M_{02} / M_{00} - y_c \cdot y_c$$

$$d = b^2 + (a - c)^2$$

$$l = \sqrt{\frac{(a+b) + \sqrt{d}}{2}}$$

$$L = \sqrt{\frac{(a+b) - \sqrt{d}}{2}}$$

$$\theta = \frac{1}{2} \arctan \left(\frac{2 \left(\frac{M_{11}}{M_{00}} - x_c y_c \right)}{\left(\frac{M_{20}}{M_{00}} - x_c^2 \right) - \left(\frac{M_{02}}{M_{00}} - y_c^2 \right)} \right)$$

Détection de la peau

- Il existe de nombreux algorithmes définissant une teinte « chair » pour isoler facilement les visages, les mains...
- Les plus simples utilisent un espace colorimétrique, parfois plusieurs
- Les plus complexes : deeplearning
- Dans tous les cas ces méthodes sont déterminées à partir de bases d'exemples.

Détection de la peau

$0 \leq H \leq 17$ et $15 \leq S \leq 170$ et $0 \leq V \leq 255$

et

$0 \leq Y \leq 255$ et $135 \leq Cr \leq 180$ et $85 \leq Cb \leq 135$

Utilise deux espaces : HSV et YCbCr

Djamila Dahmani, Mehdi Cheref, Slimane Larabi, Zero-sum game theory model for segmenting skin regions, Image and Vision Computing, Volume 99, 2020, 103925, ISSN 0262-8856,

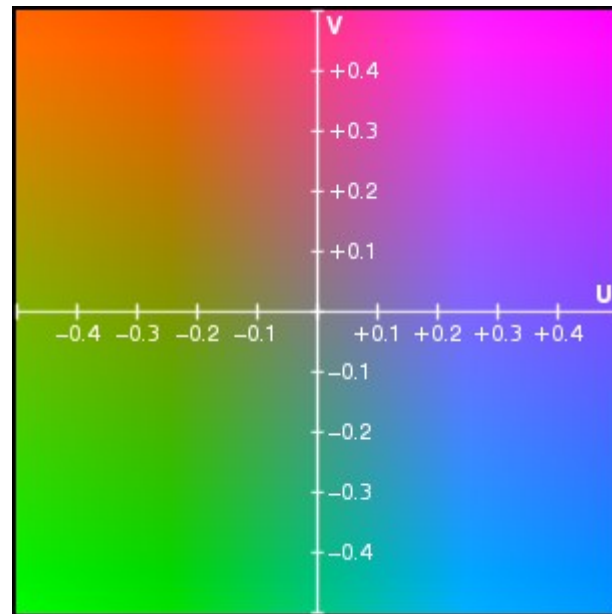
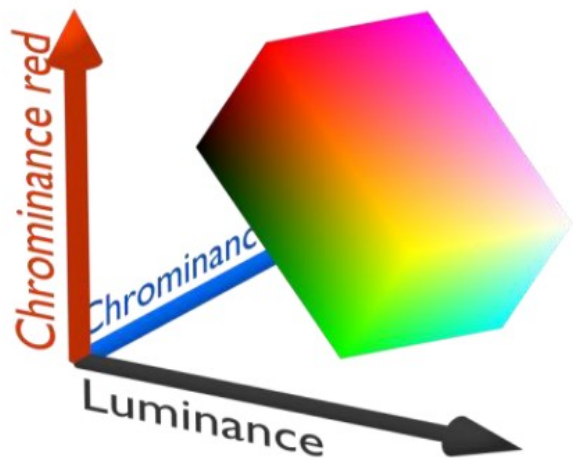
YCbCr (YUV)

- Format issu du domaine analogique (PAL/SECAM)
- Y : Luminance
- Cb+Cr=Chrominance

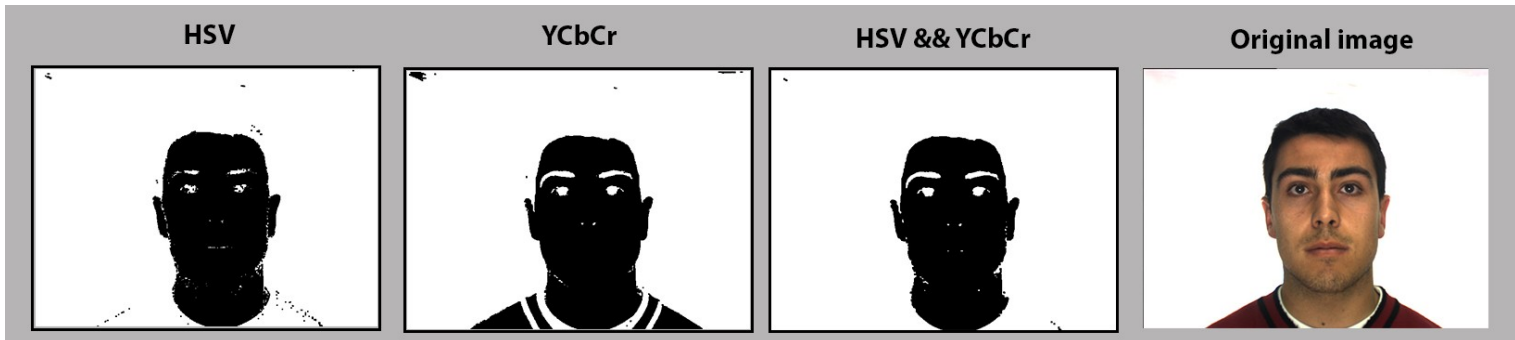
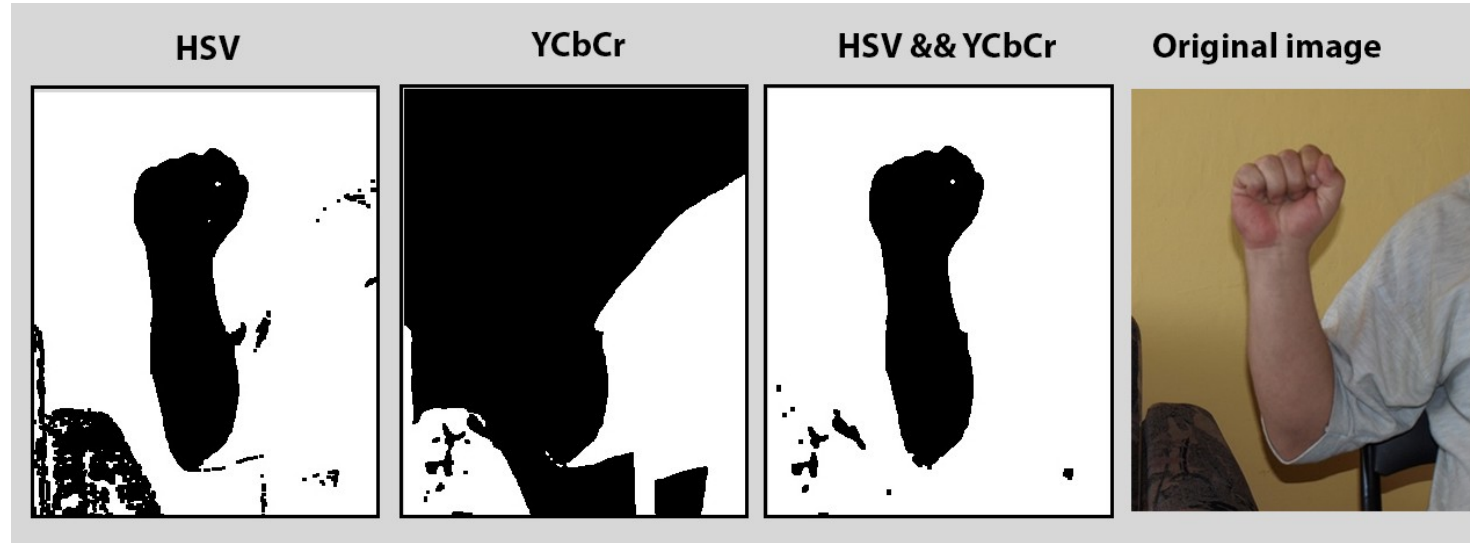
Conversion RGB → YCbCr

- $Y = 0.299R + 0.587 G + 0.114 B$
- $Cb = U = -0.147R - 0.289 G + 0.436B = 0.492(B - Y)$
- $Cr = V = 0.615R - 0.515G - 0.100B = 0.877(R - Y)$

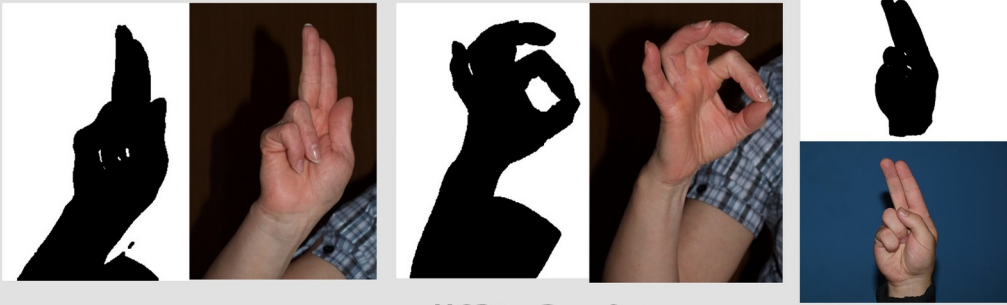
- Plan CbCr pour $Y = 0.5$



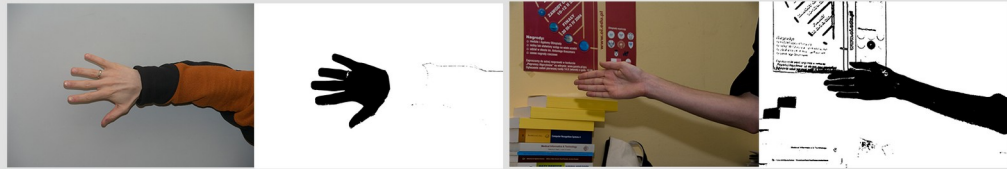
Résultats



HGR1_DataSet



HGR2a_DataSet



SFA_DataSet



