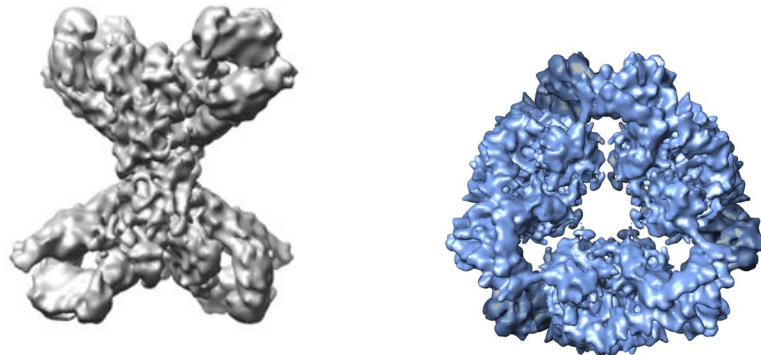


# Microscopie électronique et traitement d'images

## Cours N°2 Traitement d'images Structure d'objets biologiques dans les 3D

Master informatique spécialité IMA



Prof. **Catherine Vénien-Bryan**

Institut de minéralogie et de physique des milieux condensés (IMPMC)

CNRS UMR 7590, UPMC

E-mail: [catherine.venien@impmc.upmc.fr](mailto:catherine.venien@impmc.upmc.fr)

## Traitement d'images

### Particules isolées, protéines macromolécules, petits virus (haute résolution)

- Numérisation de l'image, cameras CCD
- Sélection des particules et normalisation du contraste
- Analyse dans les 2 dimensions:
  - Alignement
  - Classification des images
- Création du modèle dans les 3 dimensions et raffinement
  - Méthode des séries coniques aléatoires
  - Lignes communes (transformée de Radon)
  - Raffinement

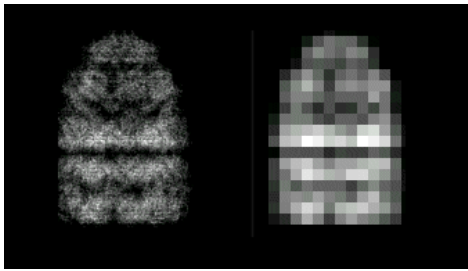
### Cellules entières, bactéries virus (moyenne résolution)

- Tomographie électronique cellulaire

# Digital images

Sampling and grey level  
Discrétisation spatiale et quantification

## Image Digitization



Original image      Image sampled at low resolution

### Pixel size

The image must be divided up into pixels (sampled) at a spacing at least twice as fine as the finest detail (highest frequency) to be analysed.  
(in practice, 3-4x as fine).

Each picture element stored in the computer, with its own grey level, is called a **pixel**.



1	1	3	5	4	3
1	1	2	5	5	3
1	1	5	5	5	4
1	1	3	2	1	3
1	1	2	5	3	1
1	1	3	5	4	3
2	1	5	4	5	5
...					

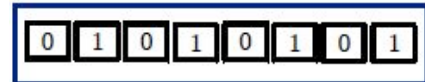
Example Data Matrix

### Grey level resolution:

one bit can code for 2 states:  
"0" or "1"

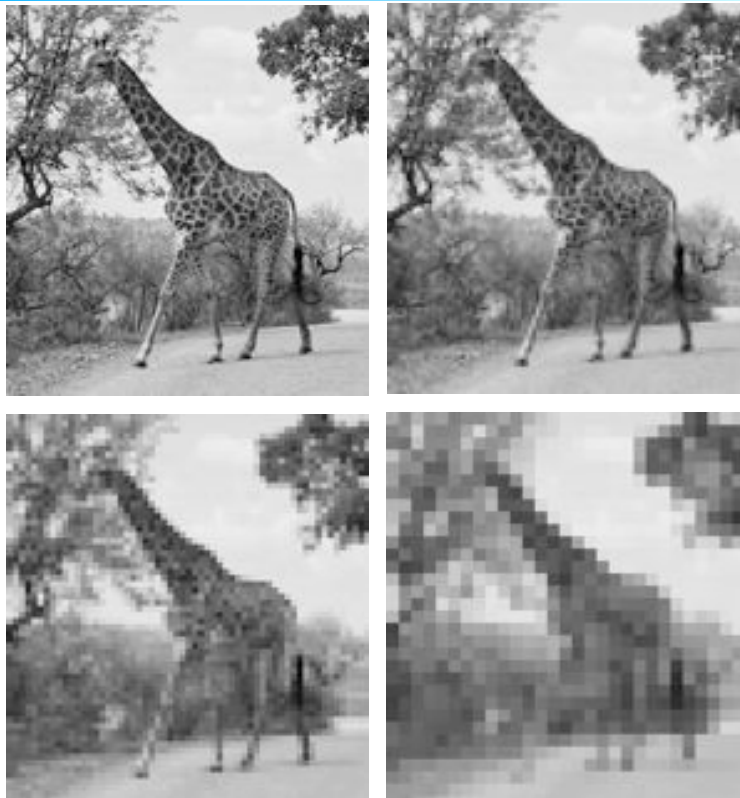


Byte (octet): a string of 8 bits can code for  $2^8 = 256$  different states



## Digital images Sampling-1

More is better!!!



## Digital images Sampling -3 - spatial resolution Théorème de Nyquist-Shannon

Remember that the image can be represented as the sum of a series of sinusoidal waves. The highest frequency wave term present defines the resolution limit. When an image is digitised, the wave components are sampled at an interval defined by the scanner step size.



Nyquist-Shannon frequency: the sampling frequency is twice the highest frequency term being represented.

If the same sampling frequency is used for a higher frequency wave, the sample points do not follow the oscillations. The features will not be correctly represented in the image.

- In practice, it is necessary to sample at 3-4x the resolution, to avoid numerical rounding errors.

## Digital images Scanning – number of grey levels-1



144



72



16



8



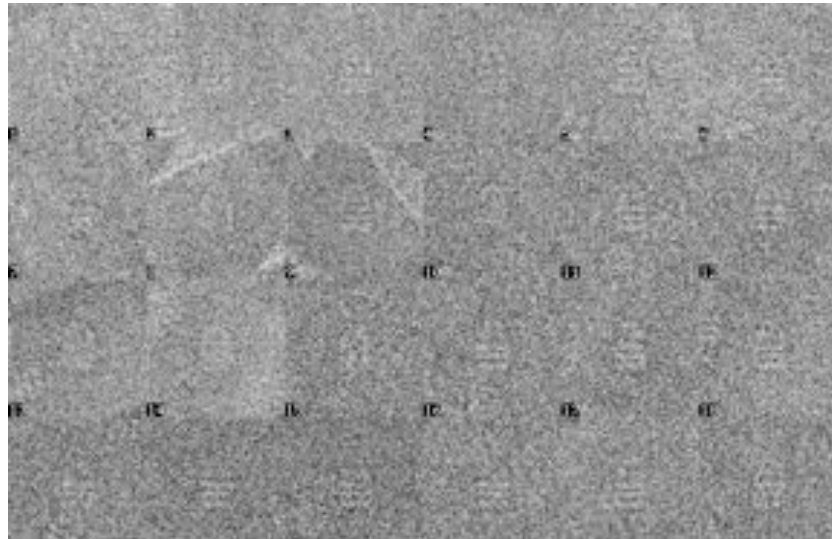
4



2

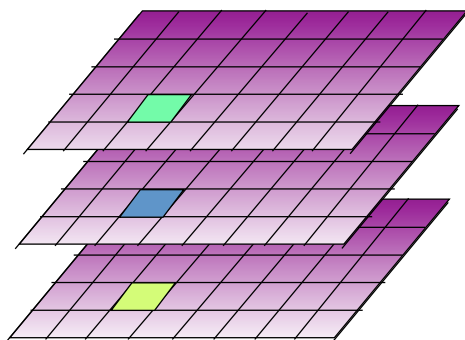
## Noise reduction by averaging

Raw images



Averages of 2 5 10 25 200 images

## Variance



For a stack of aligned images, the variance can be calculated for each pixel, to give a map of variations between the images in the data set. This can help to assess the reliability of features seen on the average image, and can reveal if images of different structures are mixed up in the same data set.

The variance is determined for each pixel as the difference between the pixel value in a given image and the average value of that pixel in all the images. This difference is squared and the sum of these squares is calculated for all the images in the stack.

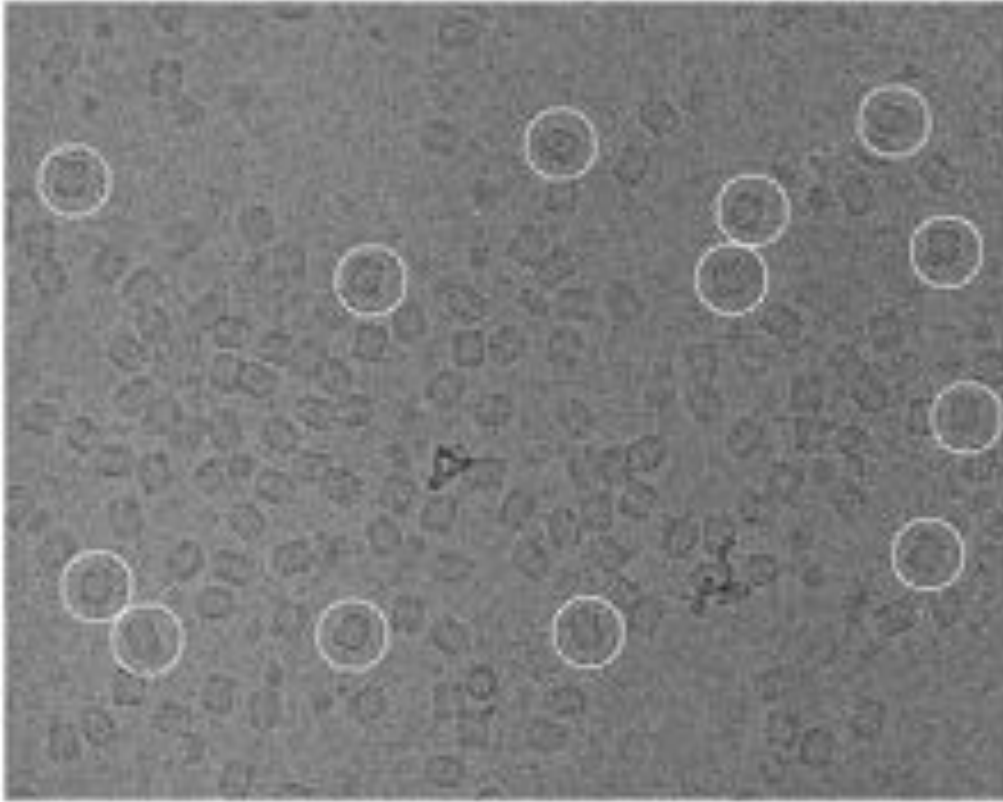
$$\text{Variance} = [1/(N-1)] \sum_{i,j} [P_i(r_j) - P_{av}(r_j)]^2$$

where  $P_i(r_j)$  is the value of pixel  $j$  in image  $i$  and  $P_{av}(r_j)$  is the average value of pixel  $j$  in all the images, for a set of  $N$  images.



## Analyse d'images dans les 2D

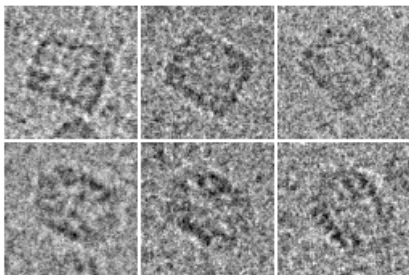
### Sélection des particules (hemocyanin *Helix pomatia*)



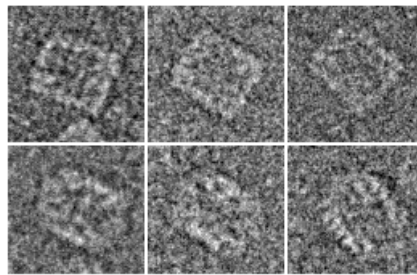
## Normalisation du contraste

**Soustraction de la moyenne et division par l'écart type**

**ORIGINAL CONTRAST**

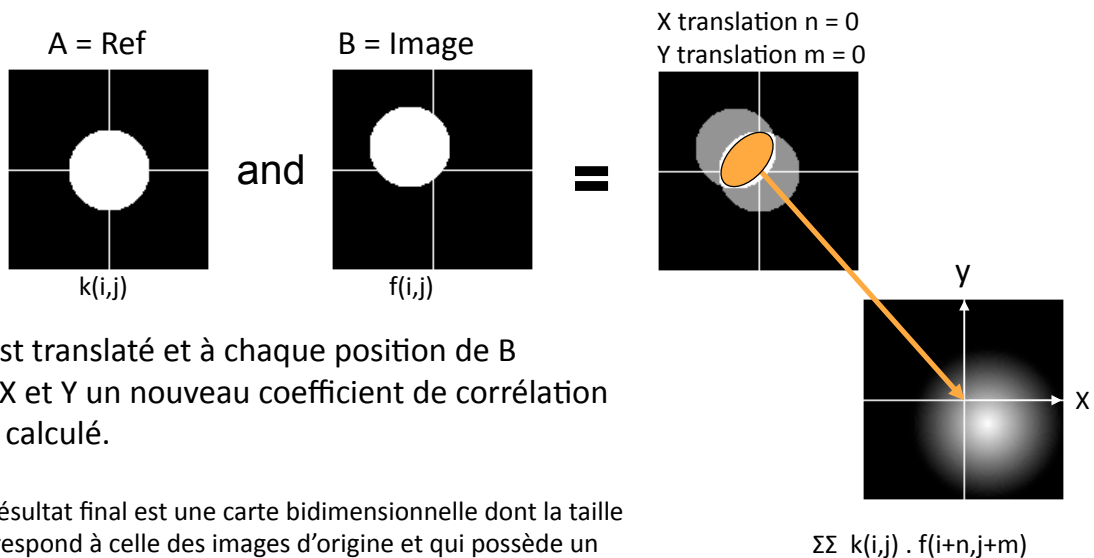


**NORMALIZED**



## Alignement et fonctions de corrélation

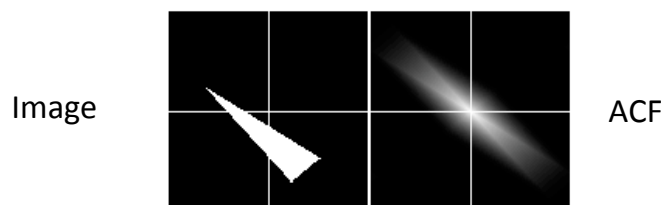
### Corrélation croisée 2D



## Alignement et fonctions de corrélation

### Fonction d'Auto-Corrélation = FAC

C'est la corrélation croisée d'une image avec elle-même



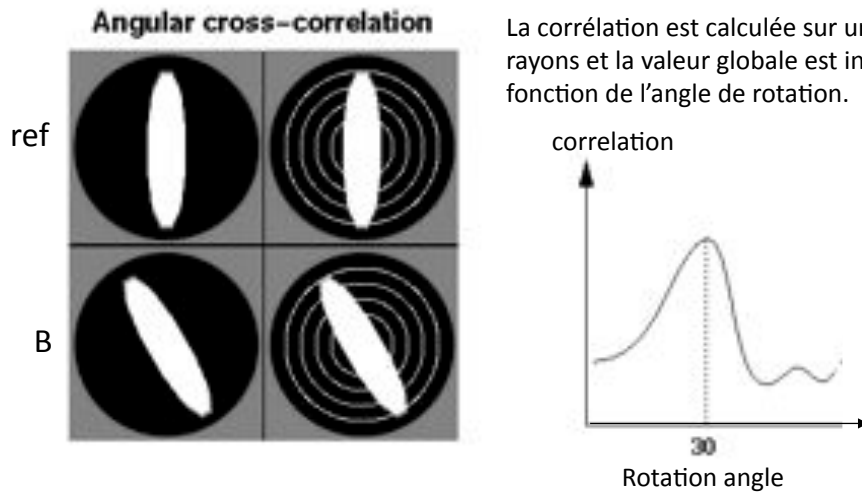
La FAC est une fonction "Centrée"  
 $F_{max}$  toujours à la position  $X = 0$  et  $Y = 0$

La FAC est une fonction "Paire"  
 $F(x,y) = F(-x,-y)$

## Alignement et fonctions de corrélation

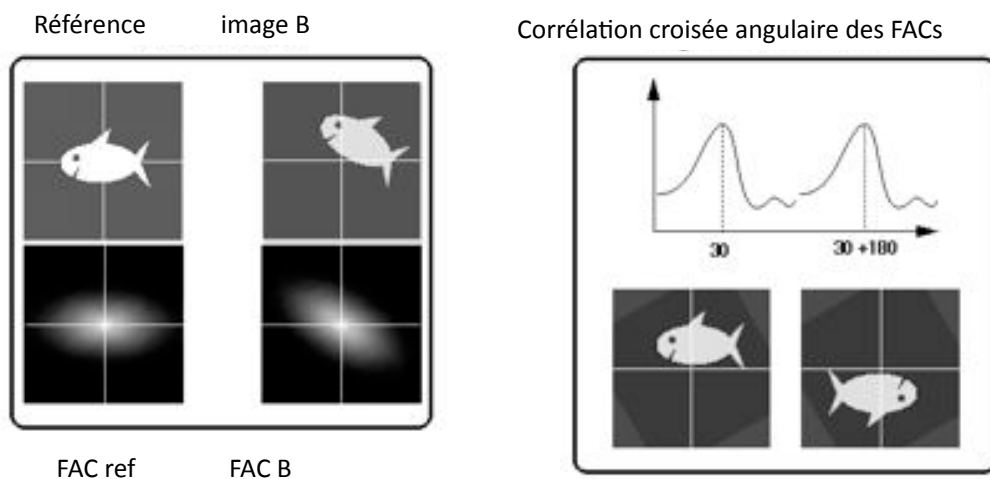
### Corrélation croisée angulaire

Une image B (supposée centrée) est placée sur une image de référence, et pour chaque rotation de B, on mesure leur coefficient de corrélation.



## Alignement et fonctions de corrélation

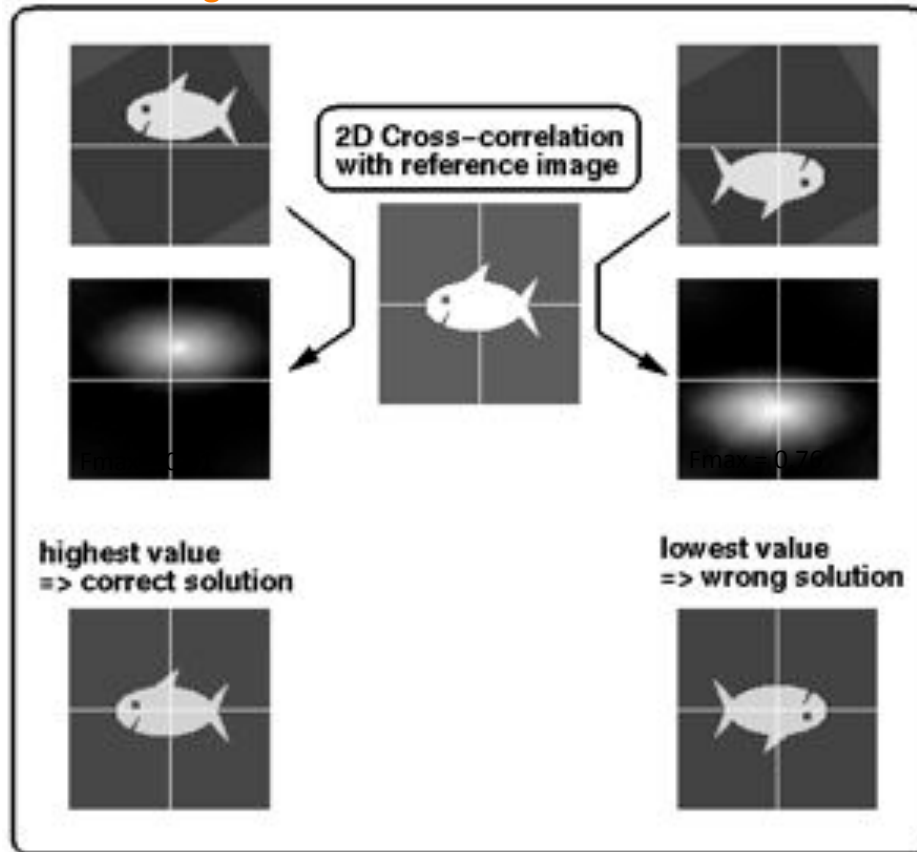
### Exemple d'alignement sur une référence



On calcule les FACs des images pour obtenir une représentation centrée et reflétant partiellement la structure des particules.

Mais comme les FACs sont des fonctions paires, il y a une ambiguïté de 180° sur l'angle de rotation qu'on trouve en calculant leur corrélation croisée angulaire.

## Alignement et fonction de corrélations Alignement sur une référence-2



## D'autres méthodes d'alignement ont été proposées pour minimiser l'influence de l'image de référence.

**"Reference free" iterative alignment** (Penczek *et al.*, 1992) :

1) Ici deux images sont prises au hasard, alignées et leur moyenne est alors utilisée comme nouvelle référence pour aligner une troisième image. Le processus se reproduit itérativement jusqu'à ce que toutes les images soient alignées.

2) Pour minimiser l'influence de l'ordre dans lequel les images ont été choisies pendant l'alignement, on repart ensuite à l'envers, en réalignant la première image et en la réalignant sur (Moyenne totale - l'image 1). Puis la seconde image est réalignée sur (Moyenne totale - l'image 2), etc ...

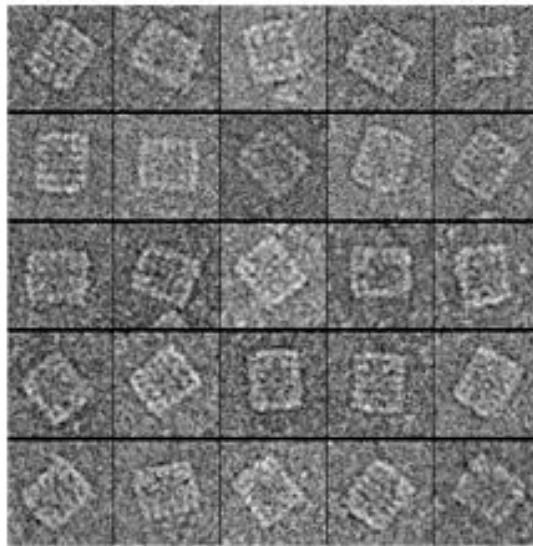
3) Le processus entier est recommencé à nouveau sur les images issues de ce premier cycle d'alignement (étapes 1 et 2), jusqu'à ce qu'aucune amélioration ne soit constatée d'un cycle au suivant.

**"multi-reference alignment"** et d'autres méthodes d'alignement utilisent la classification en parallèle du processus d'alignement.



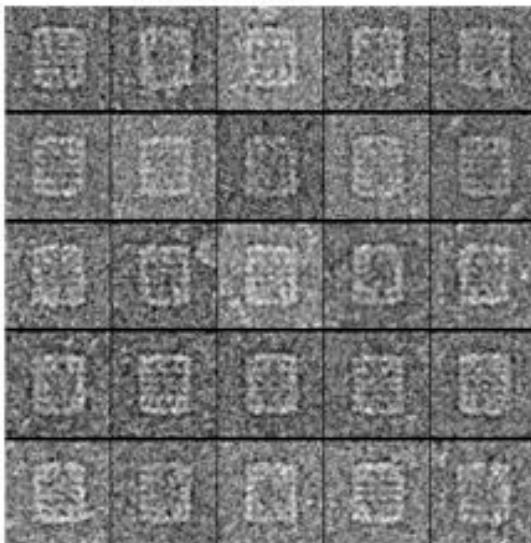
## Exemple de centrage et d'alignement , hemocyanin-1

*Helix pomatia* side views  
before alignment



## Exemple de centrage et d'alignement , hemocyanin-2

*Helix pomatia* side views  
after alignment



Averages	Variances	
		Raw images
		Cycle No.1
		Cycle No.2
		Cycle No.3 (rotation)

## Classification 2D-1

Projections 2D

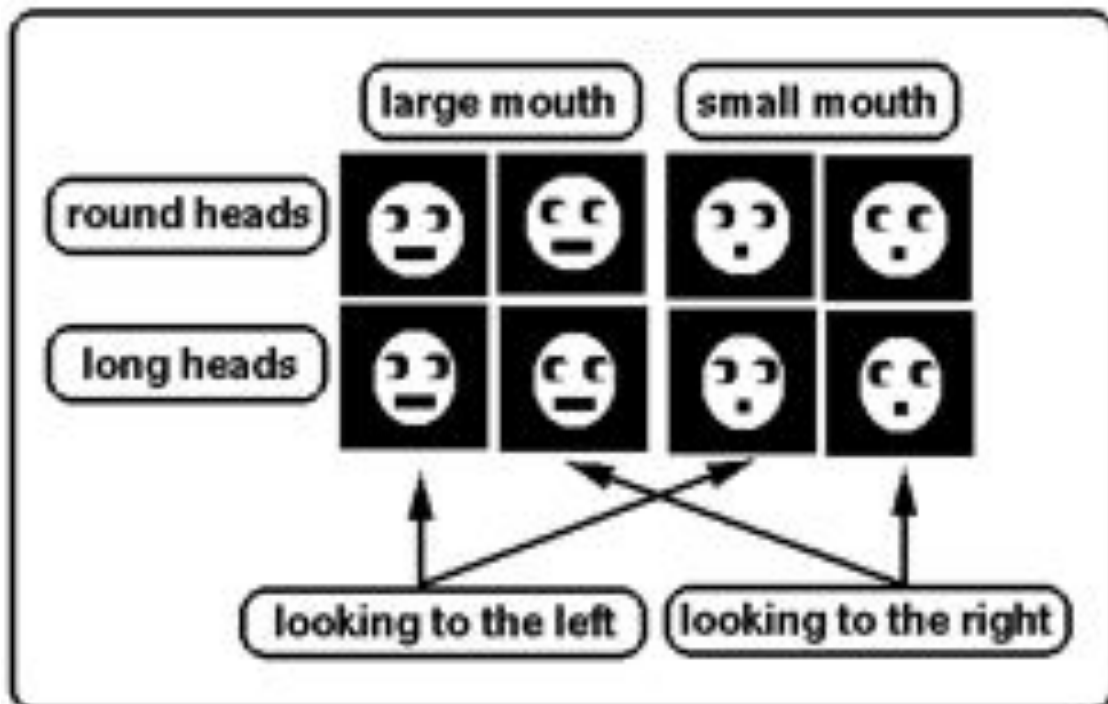
Un objet : différentes proj. 2D : OK

Différents objets : différentes proj. 2D : Problème !

EPFL\_Mech07 / JD 41

## Classification 2D-2

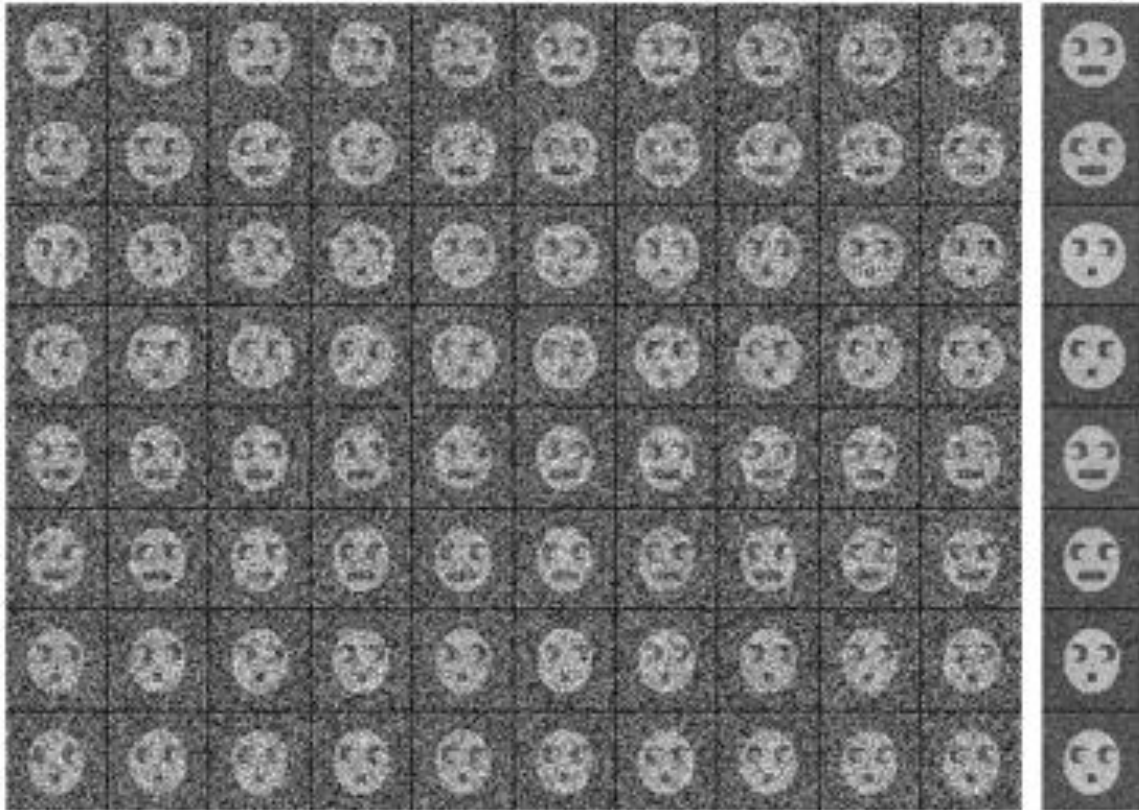
Brétaudière JP and Frank J (1986) Reconstitution of molecule images analyzed by correspondence analysis: A tool for structural interpretation. *J. Microsc.* 144, 1-14.



# Classification 2D-3

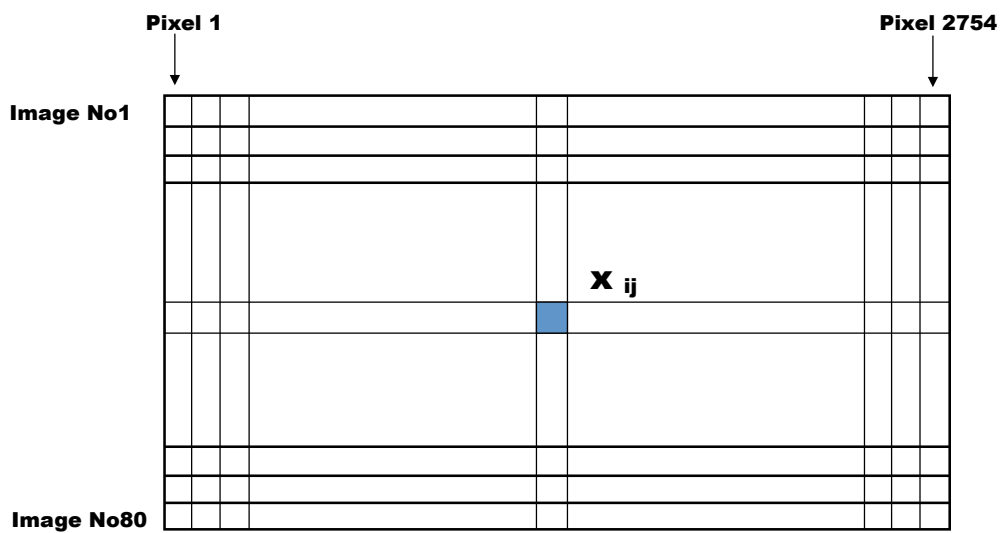
10 copies of the 8 types of heads + random noise

Averages



# Classification 2D-4

Matrice

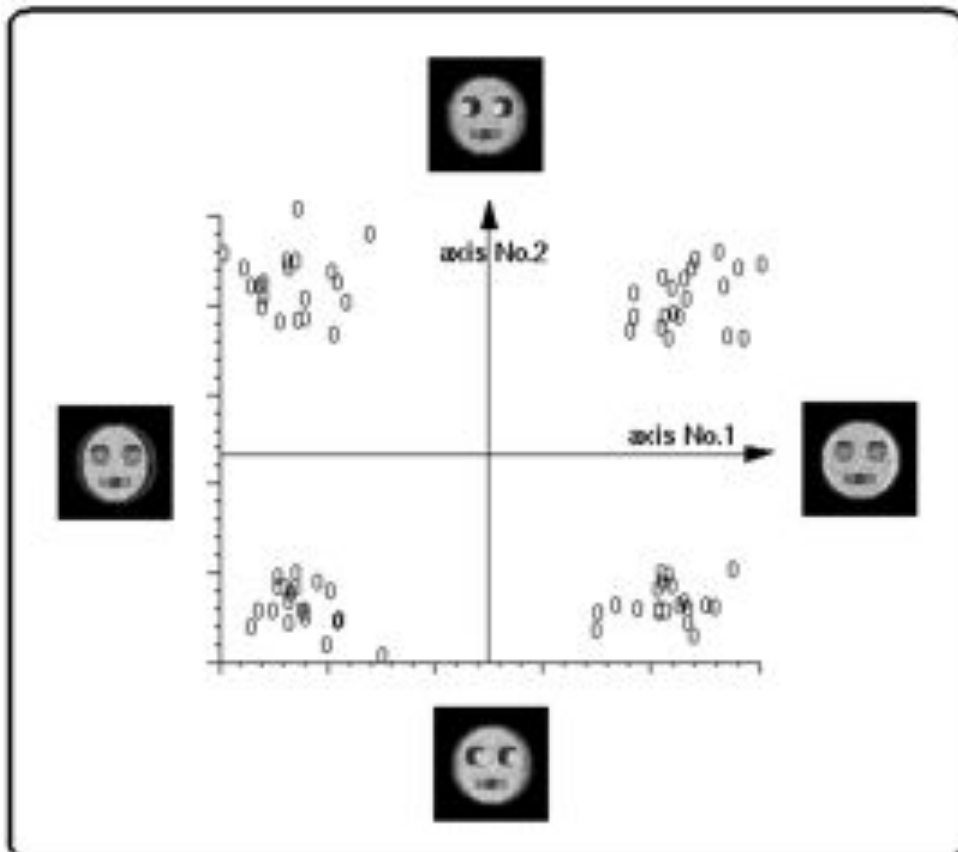
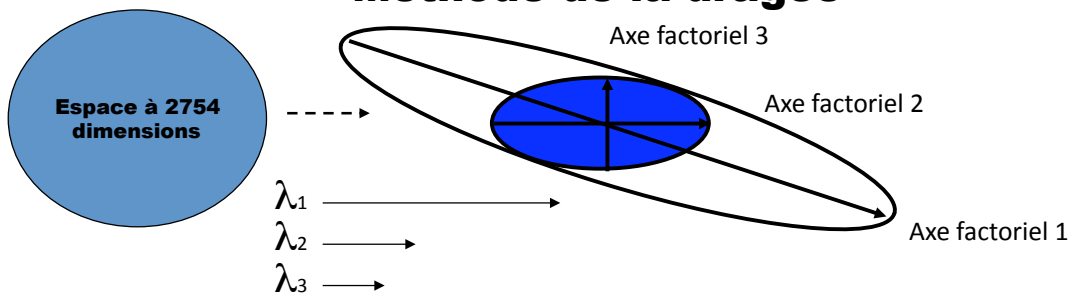


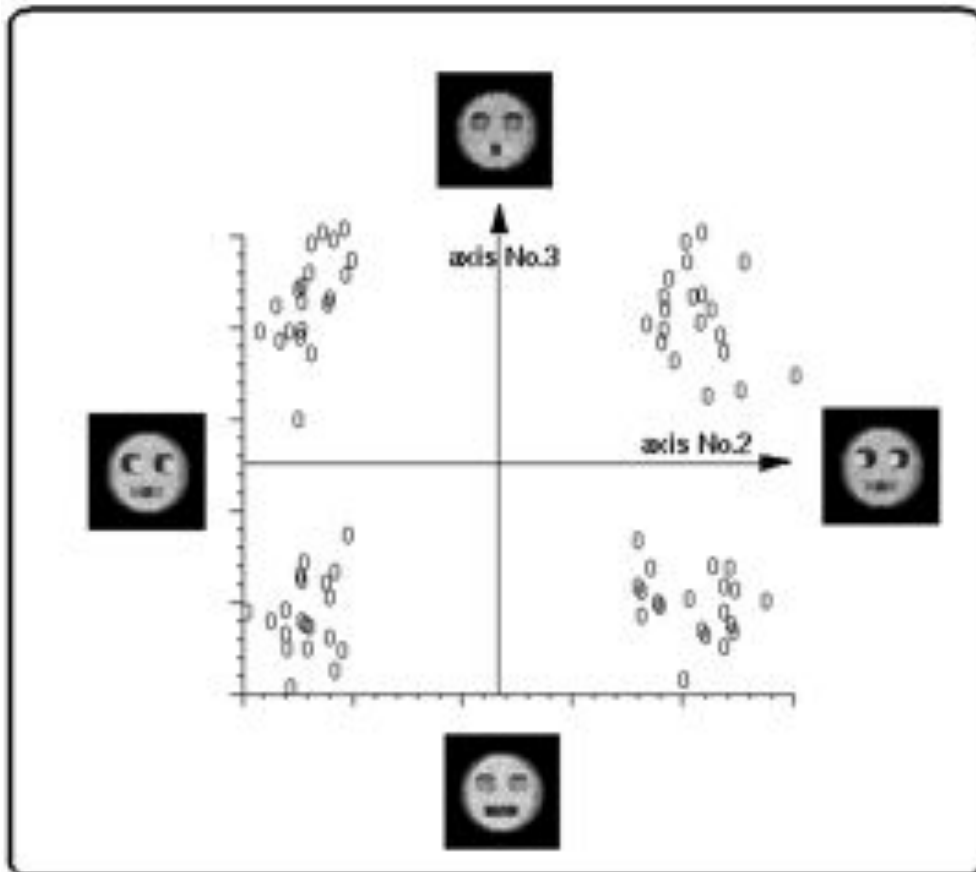
# Classification 2D-5

## Création d'une matrice a 2758 dimensions

La diagonalisation de la matrice carrée  $\mathbf{T} = \mathbf{X}'\mathbf{X}$  permet de déterminer la plus grande direction d'extension de nos données (nuage de points) dans l'espace multi-dimensionnel. Cette direction d'extension correspond à la plus grande « variation » ou « tendance » au sein de nos données. C'est l'axe factoriel N°1 dont l'amplitude est caractérisé par la valeur propre  $\lambda_1$ . On effectue alors un changement de repère pour déterminer la position de chacune de nos données (images) par rapport à cet axe factoriel. Puis, on recherche la seconde plus grande direction d'extension « orthogonale » à la première pour définir l'axe factoriel N°2 caractérisé par une amplitude de  $\lambda_2$ . Le fait que les axes factoriels N°1 et 2 soient orthogonaux, indique qu'ils caractérisent des variations indépendantes (non-corrélées). On exprime ainsi par ordre décroissant toutes les variances indépendantes de nos données sur les axes factoriels N°1, 2, 3, 4 ,etc ...

### Méthode de la dragée

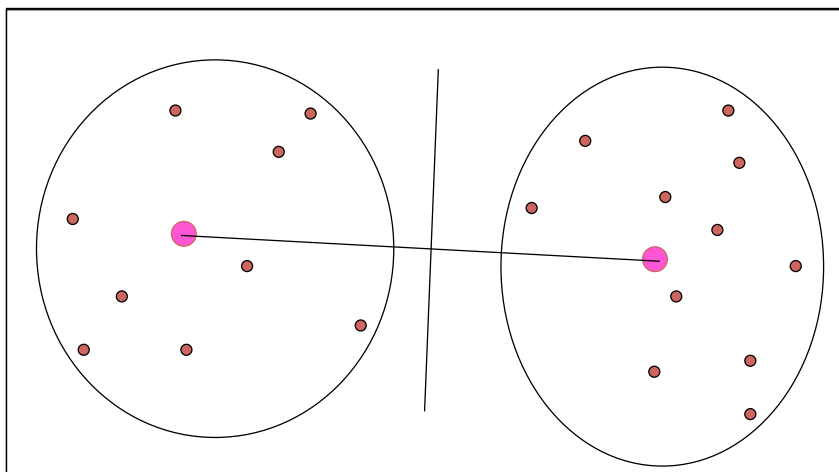




## Classification 2D

### a) Méthodes par partition: e.g. "Moving seeds" method

Diday E (1971) La méthode des nuées dynamiques. *Rev. Stat. Appl.* 19, 19-34.

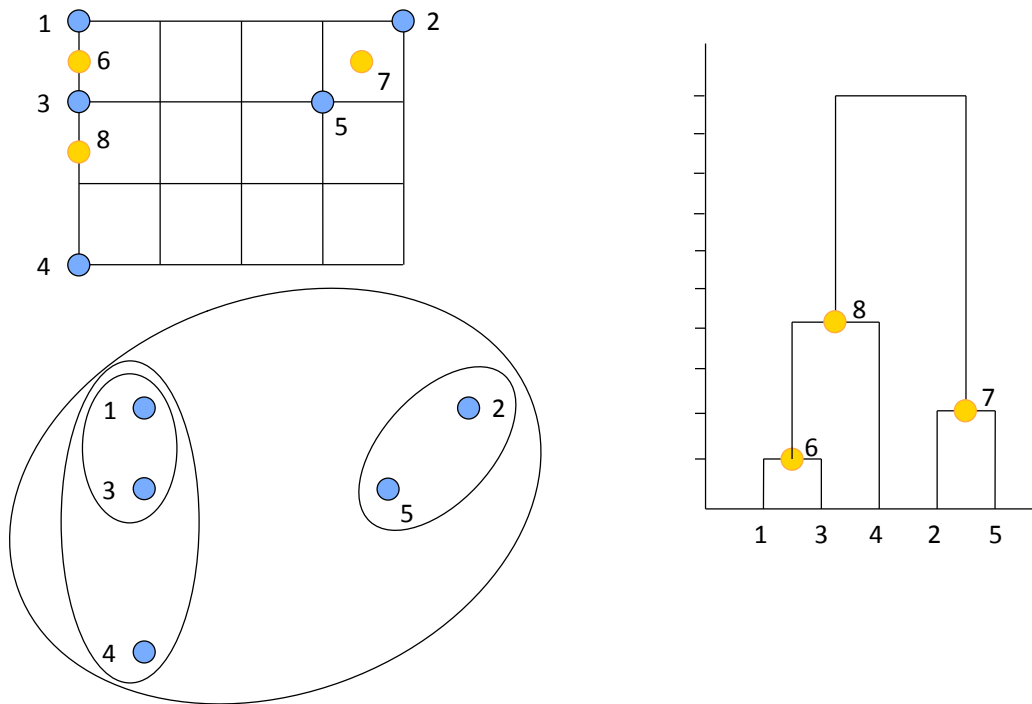


Deux images prises au hasard servent de centres d'agrégation pour la partition. Les centres de gravité de chaque classe servent de nouveaux centres d'agrégation pour un nouveau cycle de partition. Arrêt lorsque les centres d'agrégation ne bougent plus d'un cycle à l'autre, ou après un nombre déterminé d'itérations.

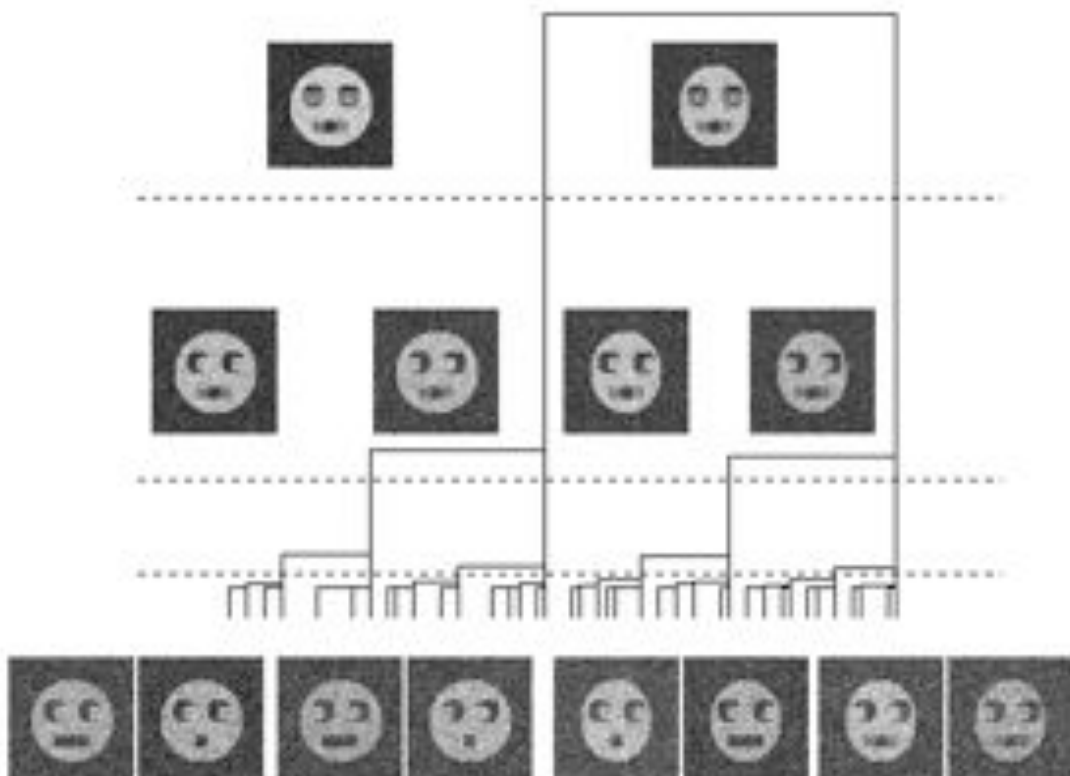


## Classification 2D

### b) Classification Ascendante Hiérarchique



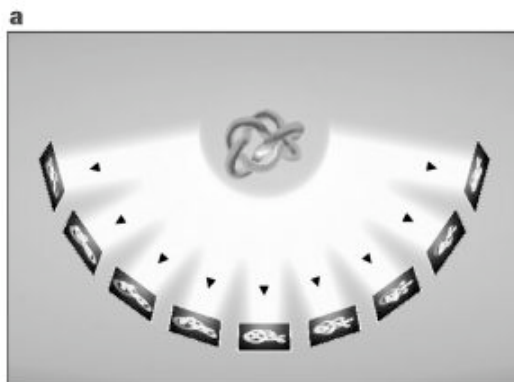
### HIERARCHICAL ASCENDENT CLASSIFICATION



## References

- Shaikh T., Gao, H., Baxter, W.T., Asturias, F.J., Boisset, N., Leith, A., and Frank, J. (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs *Nature Protocols* 3 (12) 1941-1974.
- Carragher B, Kisseberth N, Kriegman D, Milligan RA, Potter CS, Pulokas J, Reilein A (2000) Legion: an automated system for acquisition of images from vitreous ice specimens. *J. Struct. Biology* 132, 33-45.
- Potter CS, Chu H, Frey B, Green C, Kisseberth N, Madden TJ, Miller KL, Nahrstedt K, Pulokas J, Reilein A, Tchong D, Weber D, Carragher B. (1999) Frank, J. (1975). Averaging of low exposure electron micrographs of nonperiodic objects. *Ultramicroscopy* 1, 159-162.
- Frank, J., and Al-Ali, L. (1975). Signal-to-noise ratio of electron micrographs obtained by cross-correlation. *Nature* 256, 376-378.
- Frank J. (1996) Three-dimensional microscopy of macromolecular assemblies. *Academic Press*, San Diego.
- Frank J. (1990) Classification of macromolecular assemblies studied as 'single particles'. *Quart. Rev. Biophys.* 23, 281-329.
- Henderson, R., and Glaeser, R.M. (1985). Quantitative analysis of image contrast in electron micrographs of beam-sensitive crystals. *Ultramicroscopy* 16, 139-150.
- Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quart. Rev. Biophys.* 28, 171-193.
- Saxton, W.O. (1994) Accurate alignment of sets of images for super-resolving applications, *J. Microsc.* 174, 61-68.
- van Heel, M. (1989). Classification of very large electron microscopical image data sets. *Optik* 82, 114-126

## Principe de la tomographie moléculaire (1)

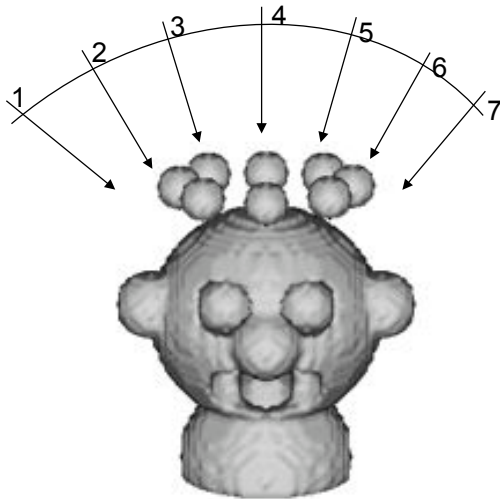


Série de projections inclinées

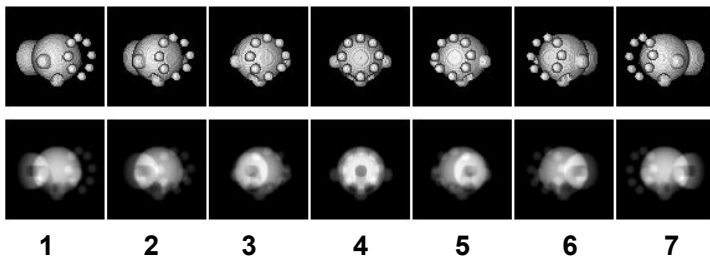


Calcul d'un volume de reconstruction 3D par rétroprojection

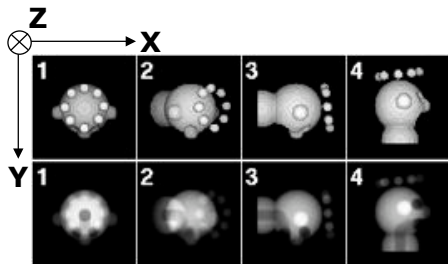
# Principe de la tomographie (séries inclinées) (2)



Espace de Fourier

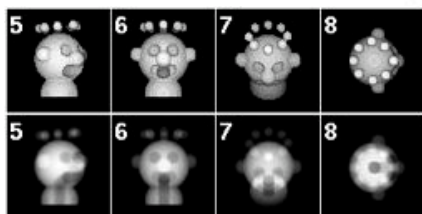
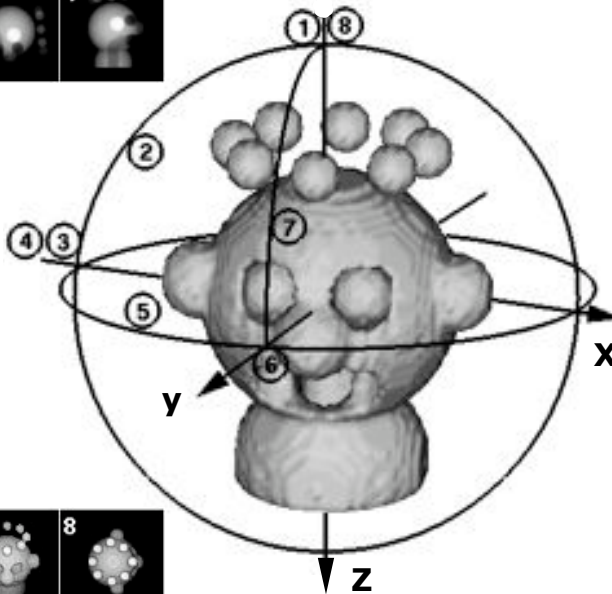


**Théorème de la section centrale:**  
 En espace réciproque, toute projection bi-dimensionnelle d'un objet correspond à une section centrale dans la transformée de Fourier 3D de l'objet. Chaque section centrale est orientée perpendiculairement à la direction de projection (direction du faisceau d'électrons).



Les conventions SPIDER concernant les angles eulériens.

	$\varphi$	$\theta$	$\psi$
1:	0	0	0
2:	0	45	0
3:	0	90	0
4:	0	90	90
5:	-45	90	90
6:	-90	90	90
7:	-90	45	90
8:	0	0	90



$\varphi$  = phi = rotation autour de Z  
 $\theta$  = theta = rotation autour de Y  
 $\psi$  = psi = rotation autour de Z

### Contrainte de la cryoMET sur objets biologiques :

Travailler à faible dose d'électrons ( $10e^-/\text{\AA}^2$ )  
Seulement 1 (ou 2) prise(s) de vue(s) par zone  
Collecter N images à faible rapport signal/bruit  
Avec une forte sous-focalisation (contraste)  
Calculer des moyennes numériques 2D ou 3D  
Corriger la fonction de transfert de contraste (CTF)

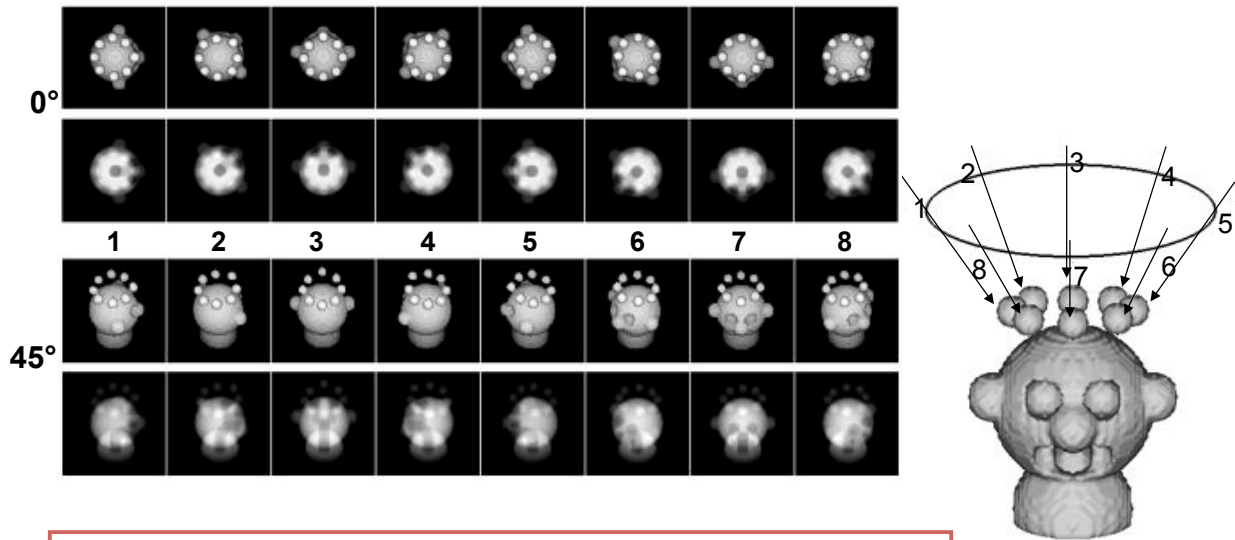
Utiliser les symétries et/ou disposition des particules:  
Symétrie hélicoïdale, icosaédrale, réseaux 2D,  
ou pas de symétrie.

### La reconstruction d'un modèle préliminaire 3D

**L'acquisition des séries coniques aléatoires**

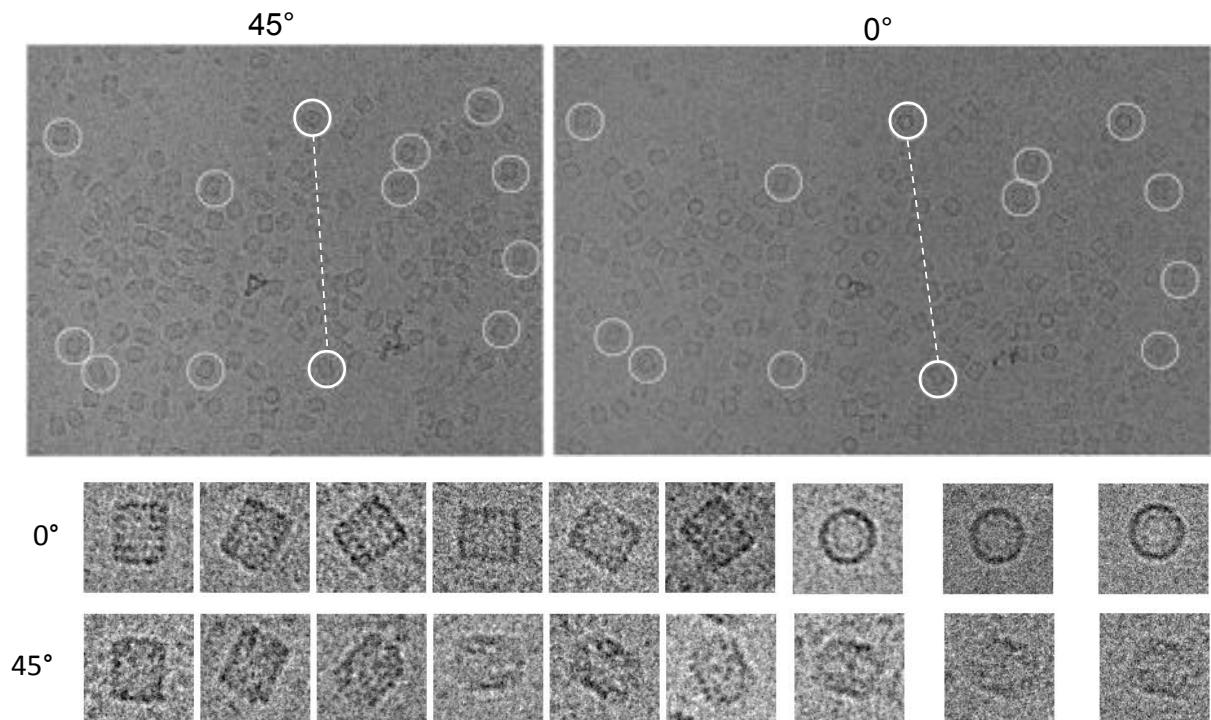
**Les méthodes de reconstructions**

## Séries coniques aléatoires



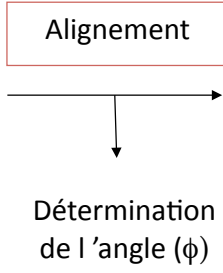
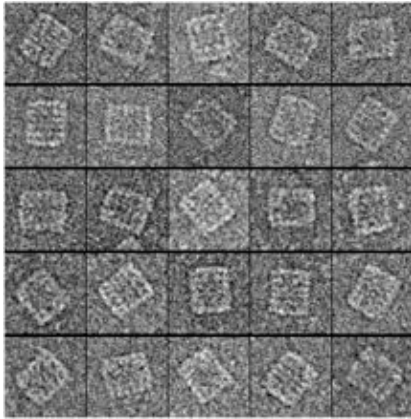
Particules isolées: Avec seulement deux prises de vues on peut calculer un volume de reconstruction

## Exemple de reconstruction à partir des images expérimentales (Hémocyanine d'*Helix pomatia*, symétrie $D_5$ )

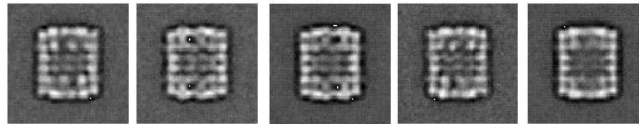
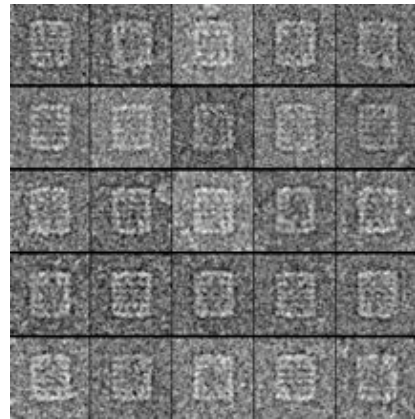




Vues de côté brutes

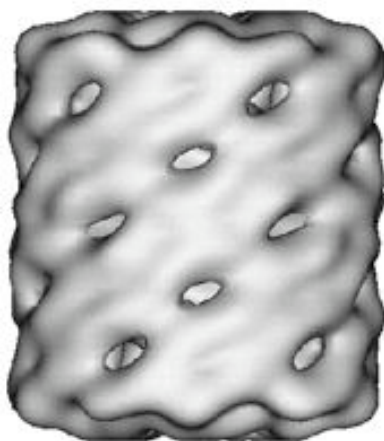


Mêmes images alignées

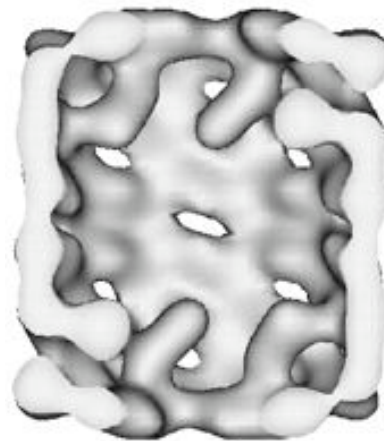


Classification automatique → 5 vues  
Calcul d'un volume par classe à partir des images inclinées  
Assemblage des volumes alignés

Hémocyanine d'*Helix pomatia*



Surface externe



Cavité centrale

## **A retenir sur l'approche des série coniques aléatoires**

- 1 Cette méthode nécessite de collecter des images avec et sans inclinaison dans le microscope. La microscopie a un faible rendement car il y a des effets de charge et de superposition partielle des particules qui ne sont pas maîtrisables.
- 2 La technique fonctionne même sur les images à faible rapport signal / bruit de fond  
→ on travaille donc sur les images brutes.
- 3 Il y a toujours un artéfact du cône manquant qui peut être supprimé si on regroupe des volumes correspondant à des vues différentes de la particules, ou si la particules a des symétries internes favorables. (exemple: particule cylindrique à symétrie  $D_5$ : Les vues de dessus produisent un volumes anisotrope avec un artéfact du cône manquant alors que les vues de côté produisent un volume isotrope).

**Les méthodes d'affinement de l'alignement et de la reconstruction**

**ART-SIRT**

## Methods for refinement of alignment and reconstruction

Require a reference 3D model.

### Real space :

Shaikh T., Gao, H., Baxter, W.T., Asturias, F.J., Boisset, N., Leith, A., and Frank, J. (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs *Nature Protocols* 3 (12) 1941-1974. SPIDER software

Penczek, P.A., Grassucci, R.A., Frank, J. (1994) The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy* 53: 251–270. Spider software

### Fourier transform:

Grigorieff, N. (1998). Three-dimensional structure of bovine NADH: ubiquinone oxidoreductase (complex I) at 22 Å in ice. *J. Mol. Biol.* 277: 1033–1046.

Jonic, S., Sorzano, C.O., Thevenaz, P., El-Bez, C., De Carlo, S. & Unser, M. (2005). Spline-based image-to-volume registration for three-dimensional electron microscopy. *Ultramicroscopy* 103(4), 303-17.

Radon transform: Radermacher, M. (1994). Three-dimensional reconstruction from random projections: orientational alignment via Radon transforms. *Ultramicroscopy* 53: 121–136.

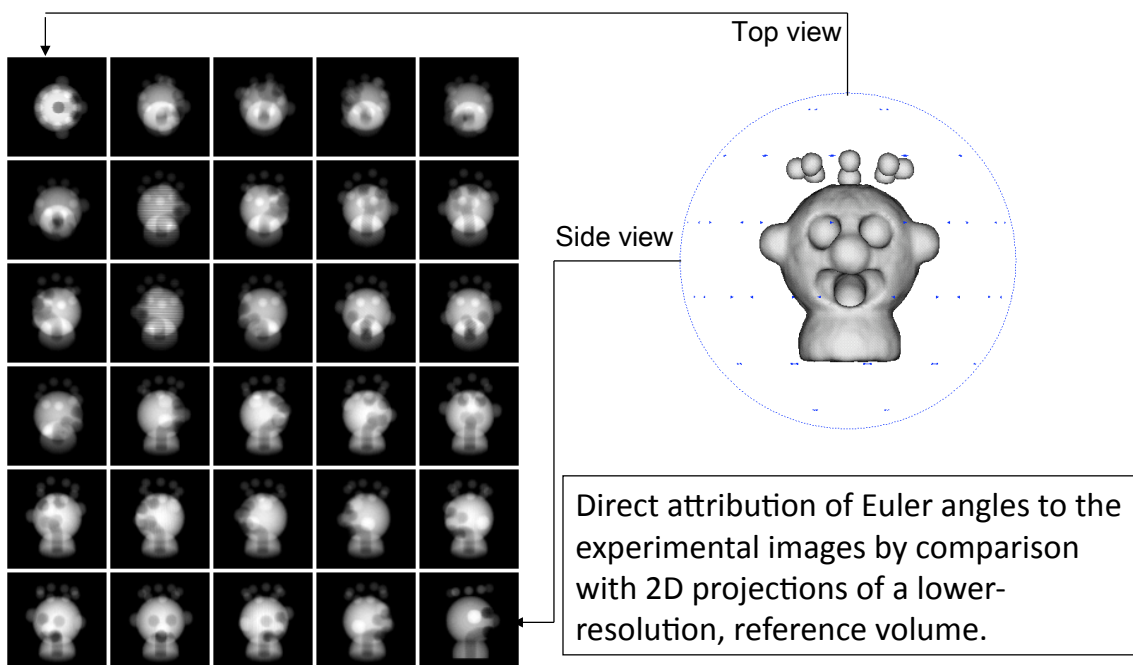
Wavelet transform: Sorzano, C.O., Jonic, S., El-Bez, C., Carazo, J.M., De Carlo, S., Thevenaz, P. & Unser, M. (2004). A multiresolution approach to orientation assignment in 3D electron microscopy of single particles. *J Struct Biol* 146(3), 381-92.

Polar Fourier transform: Baker, T.S. & Cheng, R.H. (1996). A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J Struct Biol* 116(1), 120-30.

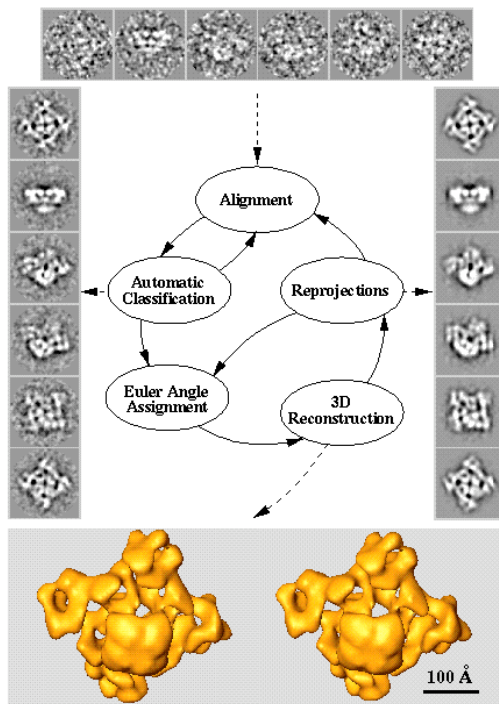
Software: SPIDER, IMAGIC, XMIPP, EMAN, etc.

## Methods that quantize the alignment parameters (1)

**Reference projections** (projections of a reference volume, discrete number of projection directions)



## Methods that quantize the alignment parameters (2)



### Global refinement

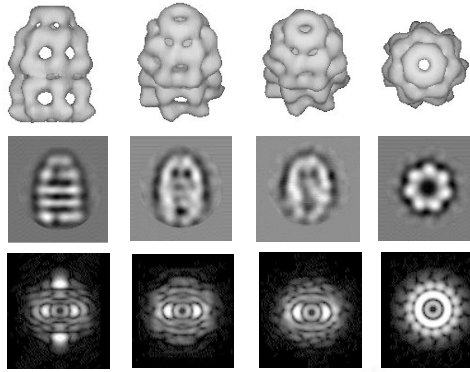
Angular sampling step for computation of the library of reference projections is reduced in each iteration of the refinement.

One of the global refinement strategies.  
Egelman 2000, Ultramicroscopy, 85: 225-234.

### Autres méthodes de reconstruction d'un modèle préliminaire 3D

L'approche des lignes communes, sinogrammes, transformée de Radon

La méthode des lignes communes (Elena Orlova EMBO 2003)



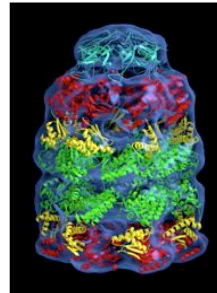
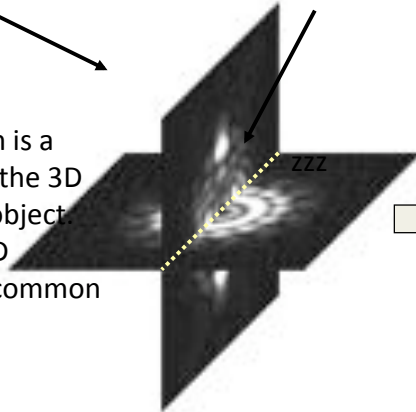
4 orientations du GroEL-GroES

Les projections 2D correspondantes

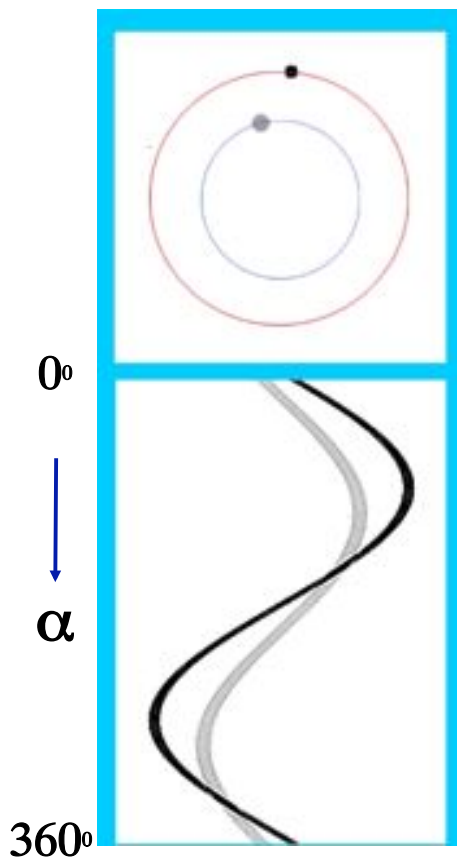
Les transformées de Fourier 2D

Transformée de Fourier 3D.

Each 2D transform is a central section of the 3D transform of the object. So each pair of 2D transforms has a common line



Lorsque toutes directions de l'espace sont représentées, la transformée inverse produit un volume de reconstruction 3D isotrope.



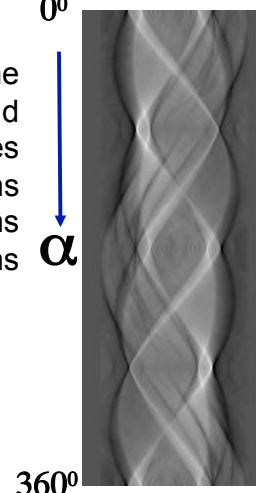
Projection 2D



2D → 1D y

Le « sinogram » d'une image 2D correspond à l'empilement de ses différentes projections 1D (lignes) dans toutes les directions ( $0^\circ$  à  $360^\circ$ ).

$$0^\circ \leq \alpha \leq 360^\circ$$

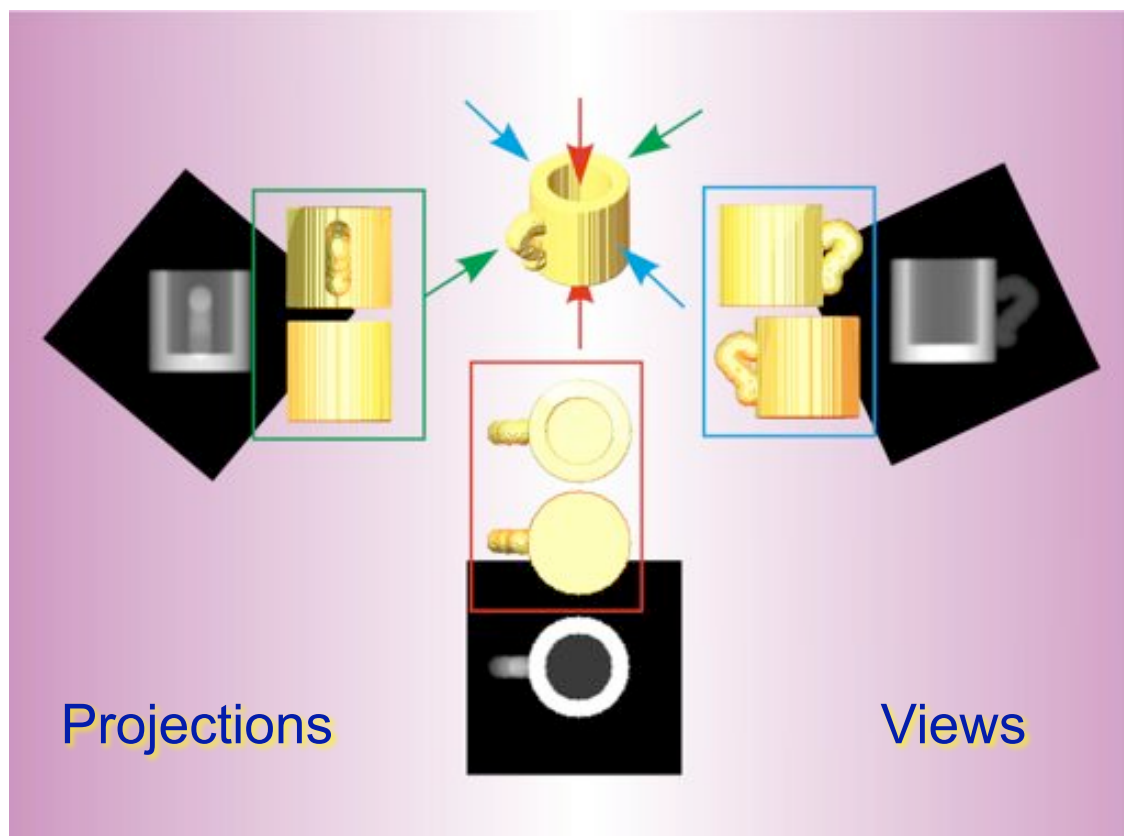


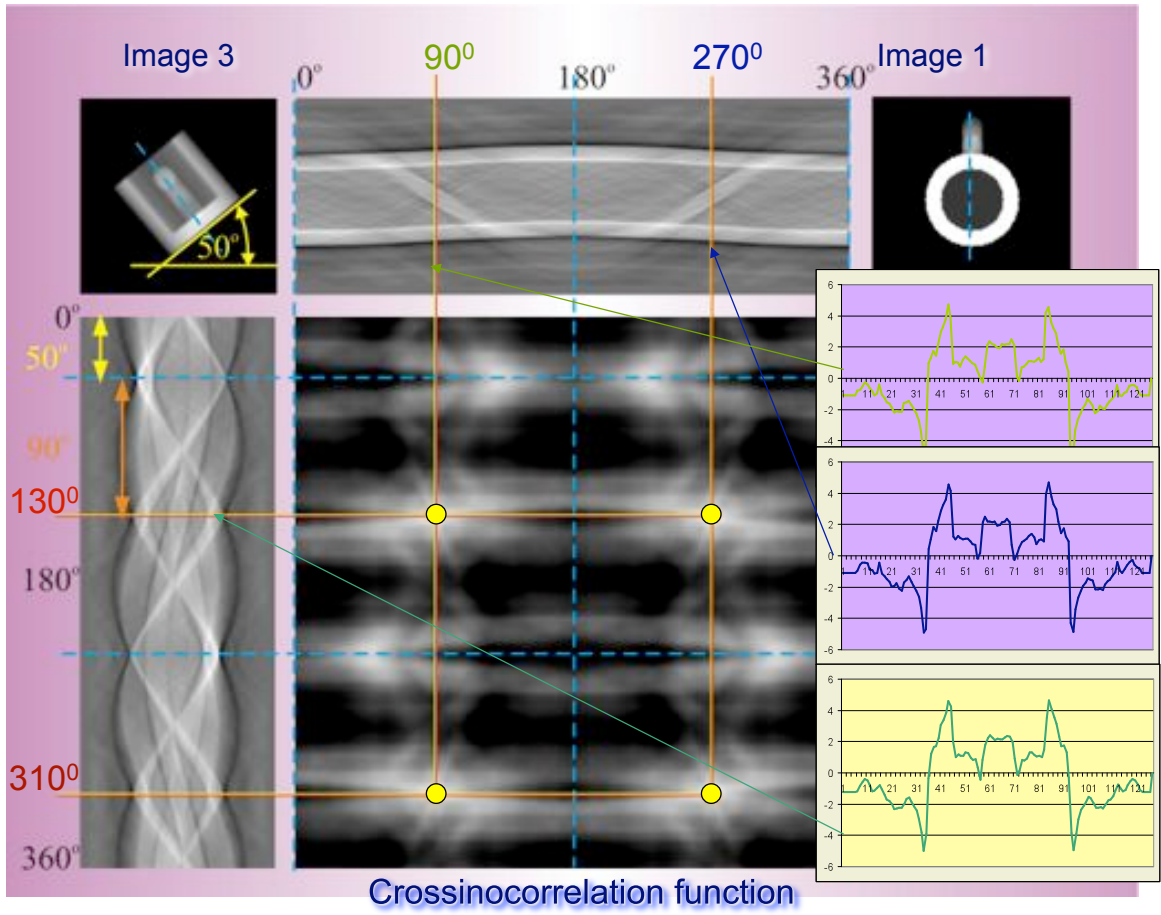


## Les lignes communes en espace réel:

La méthode des lignes communes est basée sur le fait que :

*N'importe quelle paire de projections 2D d'un objet 3D a au moins une ligne commune dans leurs projections 1D respectives.*

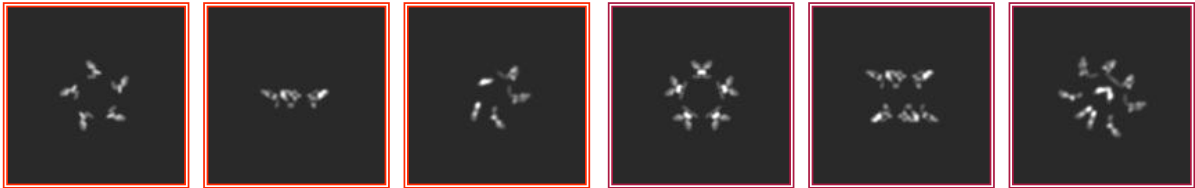




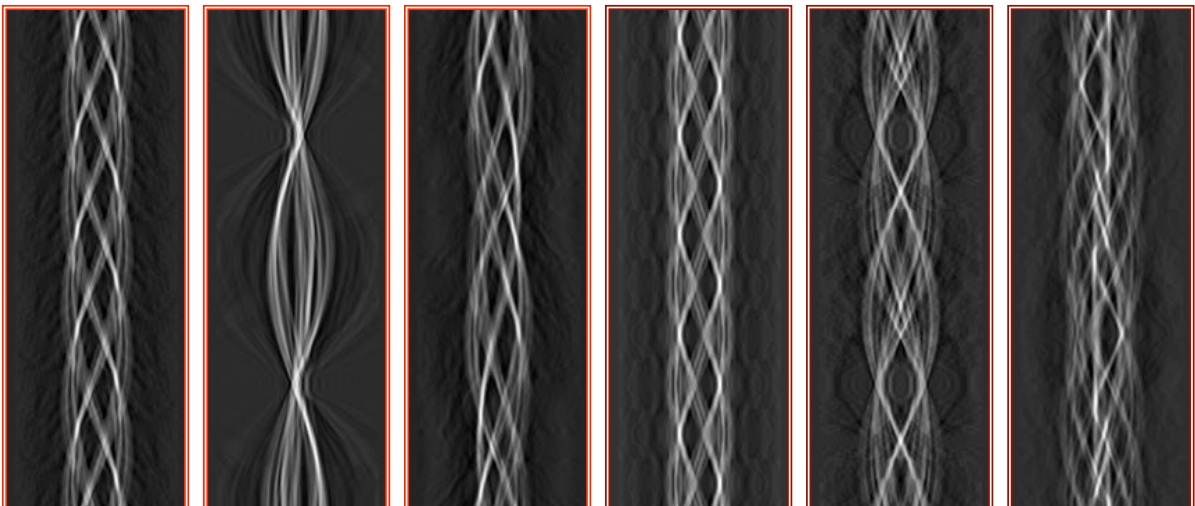
**C5**

**Projections**

**D5**



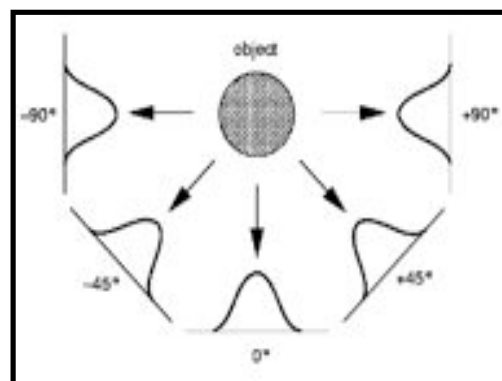
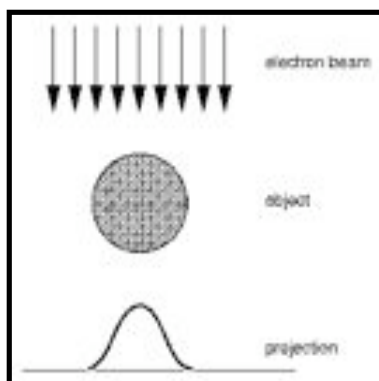
**Sinograms**



## A retenir sur l'approche des lignes communes et des sinogrammes

- 1 Cette méthode permet de collecter des images sans inclinaison dans le microscope.
- 2 Les sinogrammes ne marchent que sur des images à fort rapport signal / bruit de fond  
→ on travaille donc sur des moyennes 2D
- 3 Le centre de masse des moyennes 2D doit coïncider avec le pixel central des images.  
Le plus petit décalage introduit un biais dans la détermination des lignes communes.
- 4 La structure 3D d'une particule produite par cette approche a une chance sur deux d'être le mauvais énantiomorphe. Il faut donc réaliser au moins une paire de prises de vues avec le porte-objet incliné à  $0^\circ$  et à  $15\sim 30^\circ$  pour vérifier qu'on a bien produit le bon énantiomère.
- 5 La technique marche d'autant mieux que la particule possède des symétries internes.

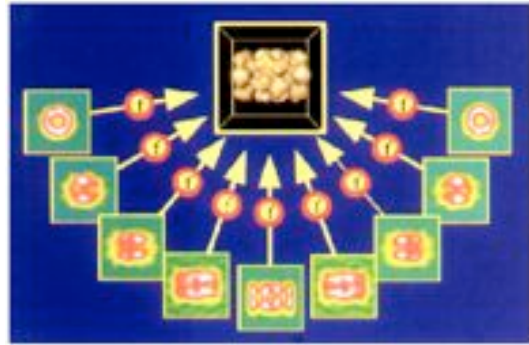
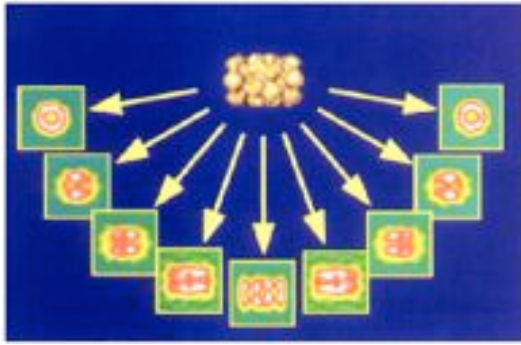
### Tomographie électronique cellulaire



An object viewed from many different angles will generate slightly different images.

## Tomographie cellulaire

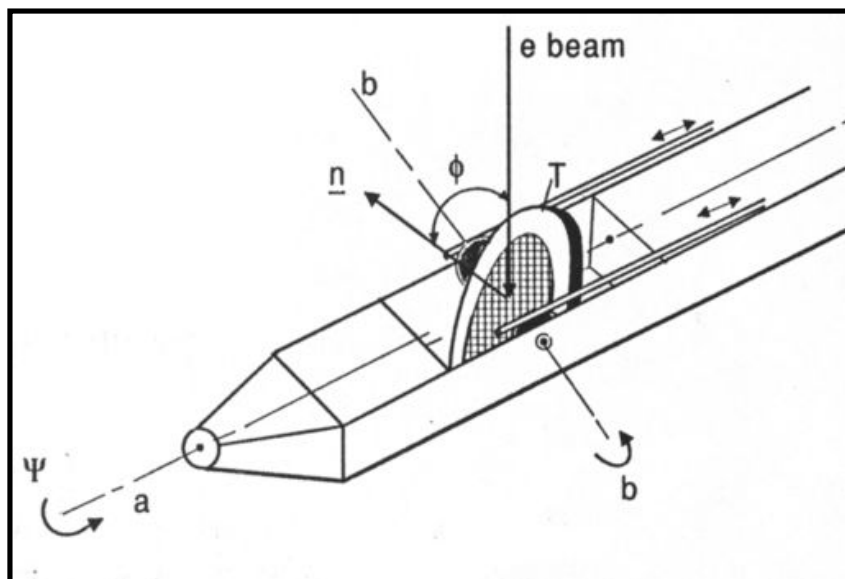
T



*Object creates multiple images so  
the inverse is also possible*

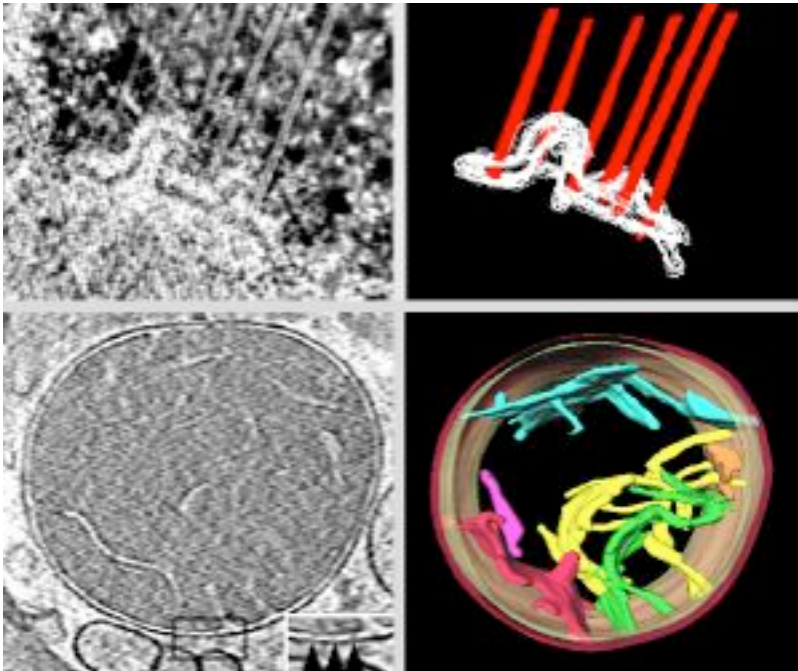
These images can be recorded and analyzed to create a tomographic rendering of the specimen.

## *Electron Tomography*

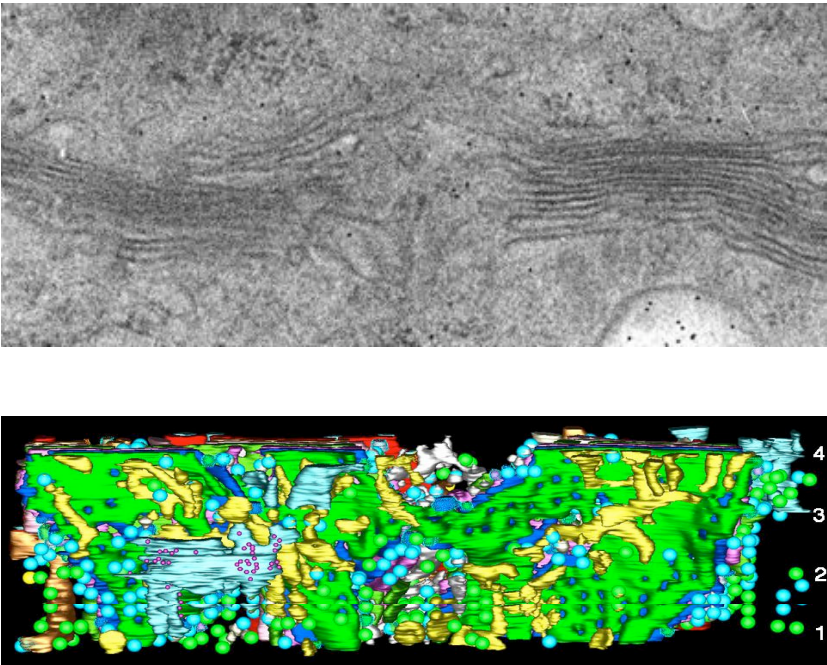


Double tilt holder

# Electron Tomography of kinetochores



# Reconstruction of Golgi



# Microscopie électronique

## Ce qu'il faut retenir

-

### -Pourquoi utiliser cet outil?

La cryo-microscopie

-Avantage: environnement naturel Inconvénient: peu de contraste, et très fragile (faible dose nécessaire)

### -Résolution

Quelles sont les limitations pour obtenir une résolution élevée?

(microscopes peu performant, aberration du ME, échantillon pas très homogène)

### Les méthodes pour aligner dans les 2D et reconstruire dans les 3D un volume

-digitaliser l'image échantillonnage et quantification normalisation

Centrage référence free multiréférence

Classification K-mean

Reconstruction d'un modèle 3D.

Méthode des séries coniques aléatoire

-sinogramme

-Raffinement de la structure 3D, Back projection, ART SIRT