

CAD

**Mémoires reçus : Classification d'objets dans
des micrographies électroniques brouillées, au
moyen de l'analyse des correspondances**

Les cahiers de l'analyse des données, tome 6, n° 1 (1981),
p. 101-107

http://www.numdam.org/item?id=CAD_1981__6_1_101_0

© Les cahiers de l'analyse des données, Dunod, 1981, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

MÉMOIRES REÇUS :
Classification d'objets dans des micrographies électroniques
brouillées, au moyen de l'analyse des correspondances
[CLASS. FIGURES]

par Marin Van Heel ⁽¹⁾ et Joachim Frank ⁽²⁾;
et d'autres mémoires sur le même sujet par J. Franck et coll.

Nous avons reçu de J. Frank, une suite d'articles écrits en collaboration avec divers auteurs. Bien que l'un de ces articles, qui fournit un titre à la présente note, nous paraisse contenir l'essentiel de ce qui regarde l'analyse des données, nous donnons en bibliographie les titres complets, pour la commodité des lecteurs qui disposent plus ou moins facilement des diverses sources.

La microscopie électronique permet aujourd'hui de distinguer des détails dont les dimensions sont de l'ordre de 10 \AA ($10 \text{ \AA} = 1 \text{ nm} = 10^{-9}$ mètre) ; à cette échelle, il semble possible de connaître la forme de macromolécules de la chimie biologique. Toutefois, le flux électronique détériorant rapidement des objets aussi fragiles que ces macromolécules, il est impossible d'obtenir des images d'une clarté satisfaisante : l'exploitation des micrographies électroniques, requiert donc leur élaboration numérique sur ordinateur.

1 L'alignement des images : Selon J. Frank et coll. une voie particulièrement prometteuse est l'alignement des images ; par quoi on entend la superposition d'images représentant des objets identiques vus sous le même angle : ainsi les véritables détails s'affirment, et le bruit aléatoire s'estompe. Il faut concevoir que, dans le champ du microscope (et donc sur la photographie qui en est prise), on a comme un semis d'objets tous égaux entre eux, mais diversement orientés et inégalement espacés. Il s'impose donc de donner d'abord à ces objets une même orientation, puis de les superposer par translation.

Pour cela, Frank & coll. ont recours aux calculs de corrélation d'image. A la base de tels calculs est la notion de produit scalaire d'image : si e.g. on représente comme des fonctions f et g de deux variables x et y , l'intensité de deux images, le produit scalaire de celles-ci n'est autre que :

$$\langle f, g \rangle = \int f(x, y) g(x, y) dx dy.$$

On calcule d'abord pour chaque objet individuel sa fonction d'autocorrélation :

$$A(\xi; \eta) = \int f(x, y) f(x + \xi, y + \eta) dx dy$$

(1) *Biochemisch Laboratorium, Rijksuniversiteit Groningen ; 9747 AG. Groningen, Hollande.*

(2) *D. of Lab. and Research, N.-Y. State Dep. of Health, Albany, N.-Y. 12201 U. S. A.*

i.e. le produit scalaire de l'objet avec tous ses translatés. Fonction dont il est facile de s'assurer qu'elle est paire ; i.e.

$$\forall \xi, \eta : A(\xi; \eta) = A(-\xi; -\eta).$$

Pour orienter un objet f' en accord avec l'objet f , F & coll. cherchent l'angle de rotation sur ces êtres mathématiques déjà centrés (donc ne requérant aucune translation préalable) que sont les fonctions d'autocorrélation A et A' : cela requiert simplement le calcul, pour une suite de valeurs de φ , du produit scalaire :

$$\int A(\xi; \eta) A'(\xi \cos \varphi - \eta \sin \varphi ; \xi \sin \varphi + \eta \cos \varphi) d\xi d\eta,$$

où l'on note dans A' un changement de coordonnées définissant une rotation. Il importe toutefois de noter que de par la symétrie de A et A' , l'angle de rotation optimal, φ , ne peut être défini qu'à π près ; autrement dit à un demi-tour près. Les objets étant orientés de façon cohérente (à π près toutefois) il devient possible d'achever l'alignement par une simple translation rendant maximum le produit scalaire.

Ici inspiré en partie par l'analyse de données dont nous rendons compte plus loin, nous croyons bon de suggérer à Frank et coll. quelques essais.

1 Pour aligner deux objets dont l'orientation n'est constante qu'à π près, il s'impose de tenter d'effectuer sur l'un de ces objets une symétrie centrale : on cherchera donc en fonction de la translation (ξ, η) le maximum des deux produits scalaires :

$$\text{Dir} : \int f(x, y) f'(x + \xi, y + \eta) dx dy ;$$

$$\text{Rot} : \int f(x, y) f'(-x + \xi, -y + \eta) dx dy ;$$

et si le maximum de Rot dépasse celui de Dir, on conclura que l'orientation de l'objet f' est à corriger d'un demi-tour.

2 De même dès l'étape de la comparaison entre les fonctions d'autocorrélation A et A' , il serait sage de tenter une symétrie axiale (miroir, ou changement de signe d'une seule coordonnée) ; certains objets pouvant être retournés. On comparera donc les maxima (pour φ variant de 0 à π) des deux produits scalaires :

$$\text{Sic} : \int A(\xi; \eta) A'(\xi \cos \varphi - \eta \sin \varphi ; \xi \sin \varphi + \eta \cos \varphi) d\xi d\eta ;$$

$$\text{Mir} : \int A(\xi; \eta) A'(-\xi \cos \varphi + \eta \sin \varphi ; \xi \sin \varphi + \eta \cos \varphi) d\xi d\eta ;$$

et si la maximum de Mir dépasse celui de Sic, on conclura que l'objet f' doit être regardé dans un miroir avant d'être ajusté à l'objet f .

3 Les calculs de fonction d'autocorrélation étant coûteux, (Frank et coll. parlent d'heures sur leur micro-ordinateur ; contre des minutes pour une a. de correspondance), on devrait tenter des ajustements d'après les centres de gravité, ou les moments et axes principaux d'inertie : ces axes déterminant l'orientation de l'objet (à π près), à condition toutefois qu'ils soient bien déterminés ; i.e. que l'objet n'ait pas à trop peu près l'inertie d'un disque ($\lambda_1 \# \lambda_2$). Les moments d'ordre supérieur à 2 pouvant également être calculés. D'ailleurs l'analyse de correspondance (cf. *infra* §§ 2 et 3) peut elle-même servir à améliorer un alignement imparfait : si un décalage subsiste entre les images les translations et rotations par rapport à l'image moyenne (qui résulterait d'un alignement optimum) ne manqueront pas de se manifester comme des facteurs. Il est facile de s'en assurer en analysant un ensemble I d'images (cf. § 2) engendré en déplaçant légèrement dans un cadre une image unique.

De ce point de vue, il apparaît qu'un alignement approché devrait suffire, avant de soumettre à l'a. des corr. les images codées.

2 Superposition ou analyse factorielle : Supposons désormais réalisé l'alignement d'un ensemble I d'images : on peut supposer que le champ plan de chacune de ces images est découpé en un ensemble J de cellules : Frank et coll. considèrent e.g. 46 images découpées en $32 \times 32 = 1024$ cellules. Ceci posé la fonction d'intensité de l'image i peut s'écrire :

$$k(i, J) = \{k(i, j) \mid j \in J\},$$

où $k(i, j)$ désigne l'intensité (de blanc... ou de noir...) de l'image i sur la cellule j. Superposer les images, c'est simplement calculer ce que selon le langage familier de l'analyse des correspondances, on appelle la loi marginale sur J :

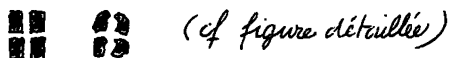
$$k(j) = \Sigma\{k(i, j) \mid i \in I\};$$

il s'agit bien là d'une image rapportée à la trame J ; avec, là où toutes les images i s'accordent, une forte intensité ; et ailleurs, un gris estompé.

Reste que le tableau kIJ se prête à une analyse factorielle, propre à révéler les oppositions majeures qui subsistent entre les images après que celles-ci ont été alignées. A une telle analyse se sont livrés Frank et coll. sur la suggestion de J.P. Bretauière ; en utilisant un programme installé par ce dernier en collaboration avec J. B. Kazmierczak.

Nous ne répéterons pas ici tout ce que l'analyse des correspondances peut apporter au traitement des images : un rapport déjà ancien de P. Chaumereuil et J.P. Villard, publié dans le présent *cahier* avait ouvert la voie : il reste à traiter de grands ensembles de données réelles ; comme le permettent les moyens informatiques aujourd'hui disponibles. Avec Frank et coll. cette étape semble abordée. Voyons donc leurs premiers résultats.

3 Premiers résultats d'analyse : Les objets considérés sont des molécules d'hémocyanine (pigment respiratoire) d'un arthropode la limule (*Limalus polyphenus*) dont le nom anglais de *horseshoe crab* évoque une sorte de crabe en fer à cheval (bien qu'il ne s'agisse pas en toute rigueur d'un crustacé). Ces molécules (de masse moléculaire $3,3 \cdot 10^6$...) se composent de huit blocs chacun formé de six unités (monomères) identiques. Sur la micrographie électronique analysée figurent des demi-molécules, c'est-à-dire des assemblages de 4 hexamères:



en gros ces demi-m. se présentent comme un rectangle divisé en 4 : avec pour séparer les quarts, deux sillons dont l'un est plus large que l'autre ; en fait ces quarts ne sont guère réguliers, et sur les images individuelles ils semblent parfois se tordre, empiéter mutuellement...

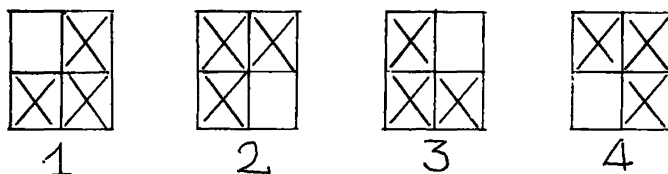
Après analyse factorielle, on trouve dans le plan 1×2 les objets répartis en quatre groupes. Pour Frank et coll. cette répartition résulte de deux oppositions fondamentales :

1) Les objets présentent une certaine torsion ; ils ont un sens ; la disposition des blocs suggérant, dirons-nous, un tourbillon. Ce que les auteurs interprètent en disant que les objets ont deux faces ; et sont vus suivant l'une ou l'autre de celles-ci selon les cas. (Frank et coll. disent : "Flip Flop" ; "pile ou face" ; cf. figure : Flop = A ou B ; Flip = C ou D). *A posteriori*, cette interprétation appelle la suggestion présentée (sous le n° 2) au § 1 : essayer systématiquement de retourner les objets en les alignant.

2) Selon F. et coll. les objets (demi-molécules) ne sont pas plans : les 4 hexamères (les quatre blocs de notre schéma) seraient plutôt au sommet d'un tétraèdre. Dans l'observation en microscopie électronique, les objets se posent nécessairement de travers (les auteurs parlent de "rocking" ; basculement) ; par le fait la coloration (plus exactement l'opacification) des objets se fait de manière inégale ; comme si l'un des quatre coins restait en l'air. Ici encore il est facile de corriger l'alignement : une rotation de π y suffit (cf. *supra* § 1 ; sous le n° 1).

Sur la figure dessinée par décalque d'après celle de F. et coll. et complétée, on voit que A et B se correspondent le mieux par déplacement direct (rotation de π), (et de même pour C et D) ; tandis que l'on passe de {A,B} à {C,D} (i.e. de flop à flip selon leur terminologie) par un déplacement inverse. On notera que si l'alignement avait été fait selon nos suggestions, il n'y aurait plus qu'une seule classe de figures. Les auteurs déclarent avoir décelé à l'oeil nu, l'opposition entre flip et flop ; tandis que les subdivisions de ces classes en deux (flop en A et B ; flip en C et D) n'ont été découvertes qu'à l'analyse factorielle.

Quoiqu'il en soit de ces interprétations, les auteurs remarquent à juste titre que par la représentation simultanée de l'ensemble I des images et de l'ensemble J des cellules, l'analyse donne en terme de traits pertinents l'explication des groupements d'images observés. Frank et coll. montrent sur la grille les cellules situées dans un même quartier du plan 1×2 ; ainsi se dessinent des sortes de masques ; qu'on peut colorer en gris d'inégale intensité d'après les coordonnées sur les axes factoriels. De façon précise, ainsi que l'expose la note citée de P. Chaumereuil et J.P. Villard, la formule de transition permet de calculer exactement les valeurs des facteurs par produit scalaire d'une image avec de tels masques.



P.S. : Sans prétendre contester l'interprétation donnée par Frank et coll. qui ont de leur problème une connaissance bien plus intime que la notre, il nous paraît bon de proposer une autre vue sur ce qu'ils appellent l'effet de torsion ou de pile ou face.

Imaginons une image divisée en quatre secteurs dont l'un est moins intensément coloré que les autres ; particularité liée à ce que les auteurs appellent basculement de l'objet non plan. Il y aura quatre apparences typiques que nous avons figurées avec les n°s 1, 2, 3, 4 . 1 est relié à 2 par une rotation d'angle π ; et de même 3 à 4 . Mais pour passer de {1,2} à {3,4}, il faut changer le signe d'une coordonnée (miroir).

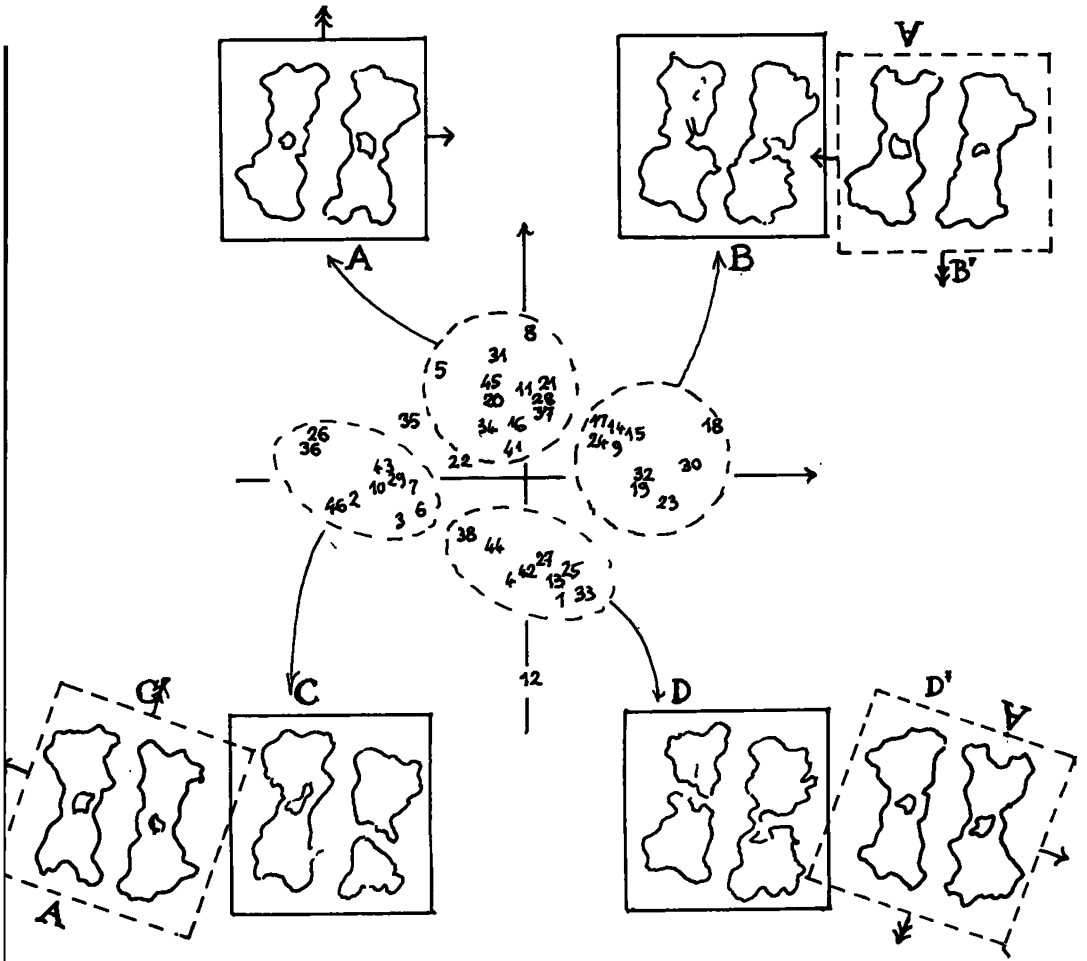
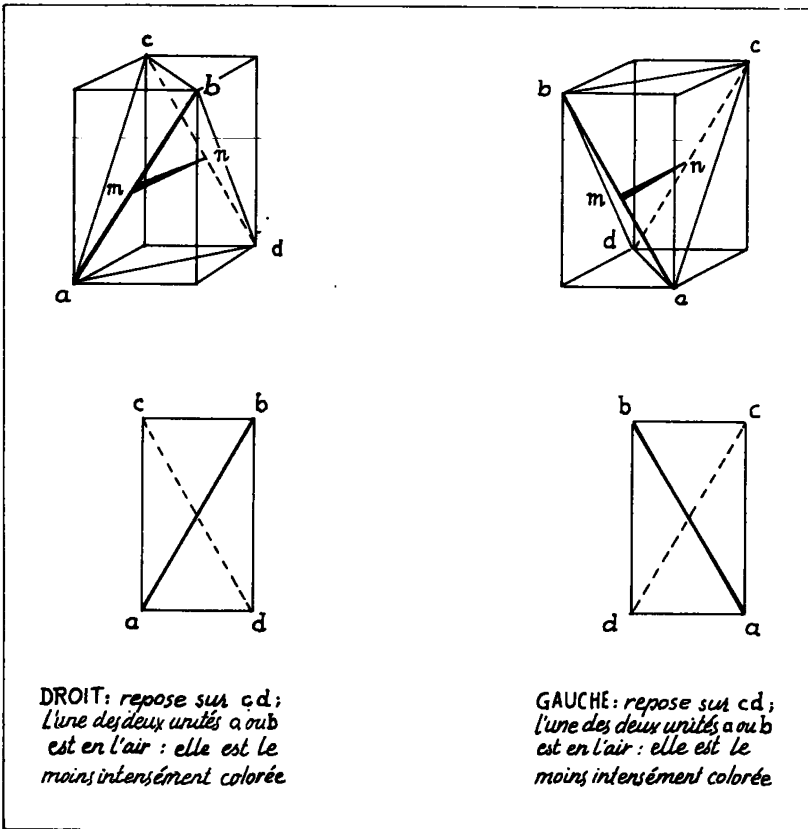


Figure d'après Frank et van IJckel, modifiée : les images des molécules sont supprimées par de simples contours ; à côté des images moyennes des classes B, C, D, on a placé des images B', C', D' obtenues à partir de la classe A : B' résulte de A par simple rotation d'angle π (déplacement direct) ; C' et D' résultent de A par des déplacements inverses (figures décalquées après retournement) ; les flèches simples ou doubles permettent de repérer les axes

Quant à la représentation spatiale on peut reprendre comme suit la proposition des auteurs. L'objet se compose de quatre unités (a,b,c,d) dont les centres sont disposés aux sommets d'un tétraèdre ayant deux arêtes opposées ab, cd (plus longues que les autres), telles que la ligne mn (très courte) joignant leurs milieux est la perpendiculaire commune à ab et cd. En projection horizontale ab et cd sont vues suivant les diagonales d'un rectangle ; l'une de ces arêtes repose sur le plan du porte-objet ; la deuxième arête ne touche ce plan qu'à une de ses extrémités ; l'autre extrémité est en l'air et détermine le secteur le moins intensément coloré du rectangle.

Comme le montre la figure ci-jointe, il existe selon ce modèle des objets de deux types : droit et gauche, qui diffèrent quelle que soit la face sous laquelle on les regarde.



Poser les objets sur l'autre arête ne change rien aux apparences : L'opposition entre DROIT et GAUCHE n'est pas une opposition entre faces (flip-flop), mais une opposition entre objets inversement égaux (de-mi-molécules droites et demi-molécules gauches).

Bibliographie : N. B. Seuls les deux premiers titres concernent l'analyse factorielle ; les trois derniers étant exclusivement consacrés à l'élaboration et l'alignement des images.

Marin van Heel et Joachim Frank : Classification of particles in noisy electron micrographs using correspondence analysis ; in *Pattern Recognition in Practice*, E.S. Gelsema et L.N. Kanal (ed.) ; N. Holl (1980).

eorumdem : Intelligent averaging of single molecules using computer alignment and correspondence analysis: I the basic method ; II Localisation of image features. in *Electron Microscopy*, Vol 2 . (1980).

M. Kessel, J. Frank, et W. Goldfarb : Low dose electron microscopy of individual biological macromolecules ; in *Electron Microscopy at Molecular Dimension*, W. Baumeister et W. Vogell (ed.) ; Springer(1980).

J. Frank et W. Goldfarb : Methods for averaging of single molecules and lattice-fragments ; *ibid.*

N.O. Saxton et J. Frank : Motif detection in quantum-noise limited electron micrographs by cross-correlation ; in *Ultramicroscopy* ; N. Holl (1977).