

P. CHAUMEREUIL

J. P. VILLARD

Un exemple de discrimination de figures par l'analyse des correspondances

Les cahiers de l'analyse des données, tome 6, n° 1 (1981),
p. 108-114

http://www.numdam.org/item?id=CAD_1981__6_1_108_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/>

UN EXEMPLE DE DISCRIMINATION DE FIGURES
PAR L'ANALYSE DES CORRESPONDANCES
[DISCR. LETTRES]

par P. Chaumereuil et J. P. Villard

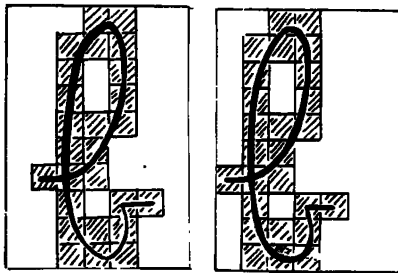
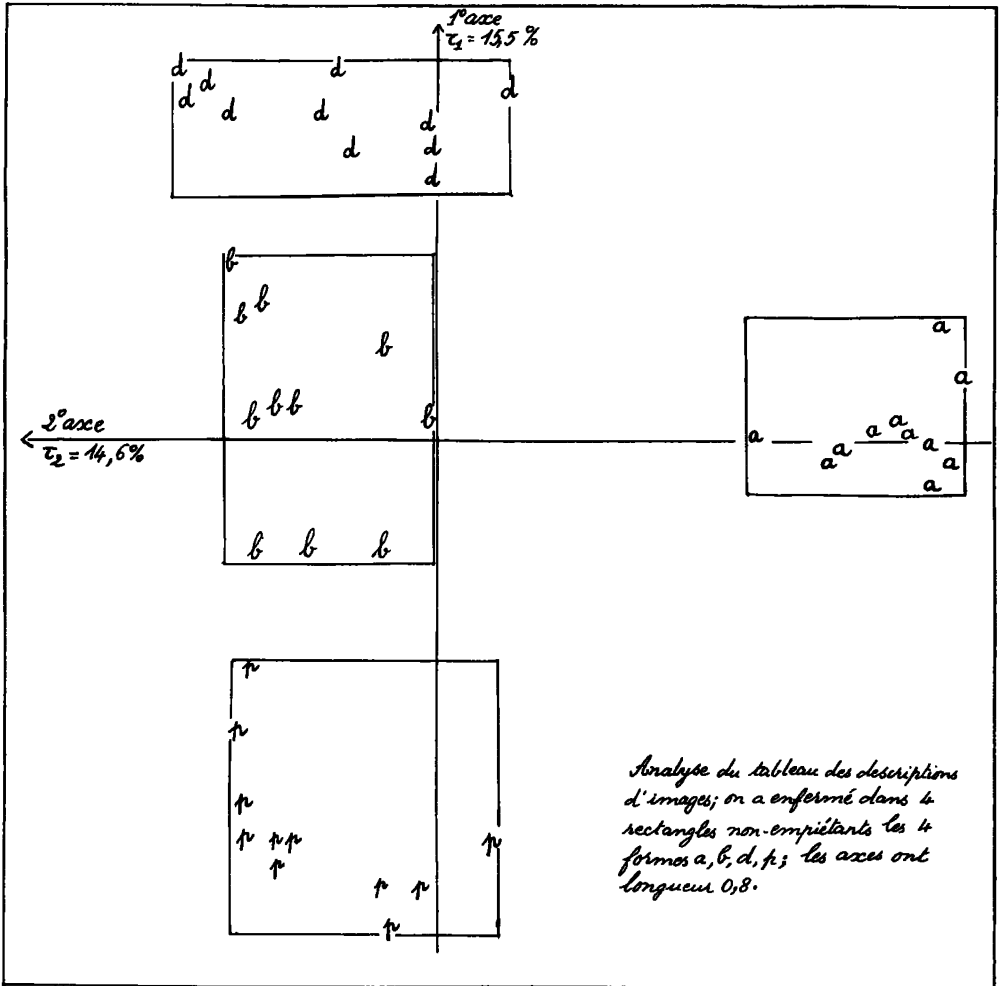
1 Les données et leur codage

Soit un cadre subdivisé en un ensemble C de cellules : toute image particulière i inscrite dans le cadre peut être approximativement décrite par un vecteur $\{k(i,c) | c \in C\}$, où $k(i,c)$ représente l'intensité du noir dans la cellule c . Dans l'exemple traité ici, le cadre est un rectangle 50×35 mm, subdivisé en 70 cellules carrées 5×5 mm ; les images sont des lettres minuscules cursives, a, b, d , ou p , inscrites dans ce cadre ; on pose $k(i,c) = 1$ si en traçant la lettre numérotée i on a rencontré le carré c , et zéro sinon. Pour illustration, sont figurées deux lettres b dans leur cadre : les cellules rencontrées par le trait sont hachurées. On considère un ensemble I de 44 figures inscrites (ou lettres) : chacune des quatre minuscules a, b, d, p figure onze fois dans notre échantillon ; pour chacune de ces lettres deux des figures individuelles les représentant sont égales à une translation près (cf. figure pour b). En somme, on a un tableau de correspondance $k_{I \times C}$, 44×70 , de 44 descriptions d'images (lettres).

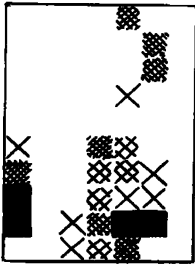
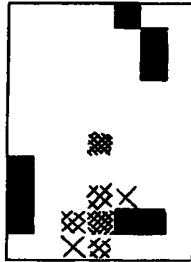
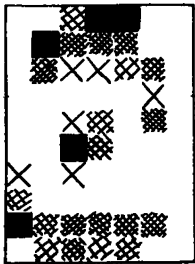
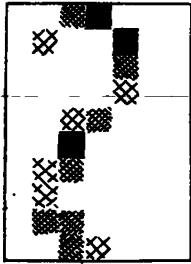
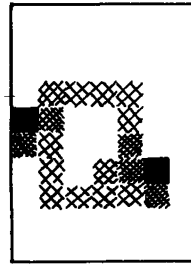
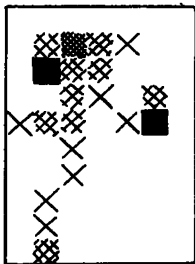
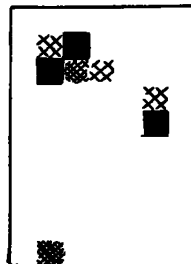
2 L'analyse des données

Le problème se pose de séparer les uns des autres les quatre sous-ensembles (relevant chacun d'une lettre) de 11 vecteurs à 70 composantes. On sait que l'analyse discriminante s'applique à placer des cloisons dans R^{70} entre les sous-ensembles, assimilés, par exemple, chacun à un nuage ellipsoïdal. D'un point de vue plus proche de l'origine graphique des données, certains recherchent par tâtonnement des grilles ou ensemble de traits tracés dans le même cadre que les images : la forme de chaque lettre devant être caractérisée par le fait de rencontrer certaines grilles et non d'autres ; ou encore on attribue à chaque image une note, combinaison linéaire, aux coefficients judicieusement choisis, des intensités du noir dans les diverses cellules. Dans le présent cas, l'analyse factorielle du tableau de correspondance $k_{I \times C}$ suffit.

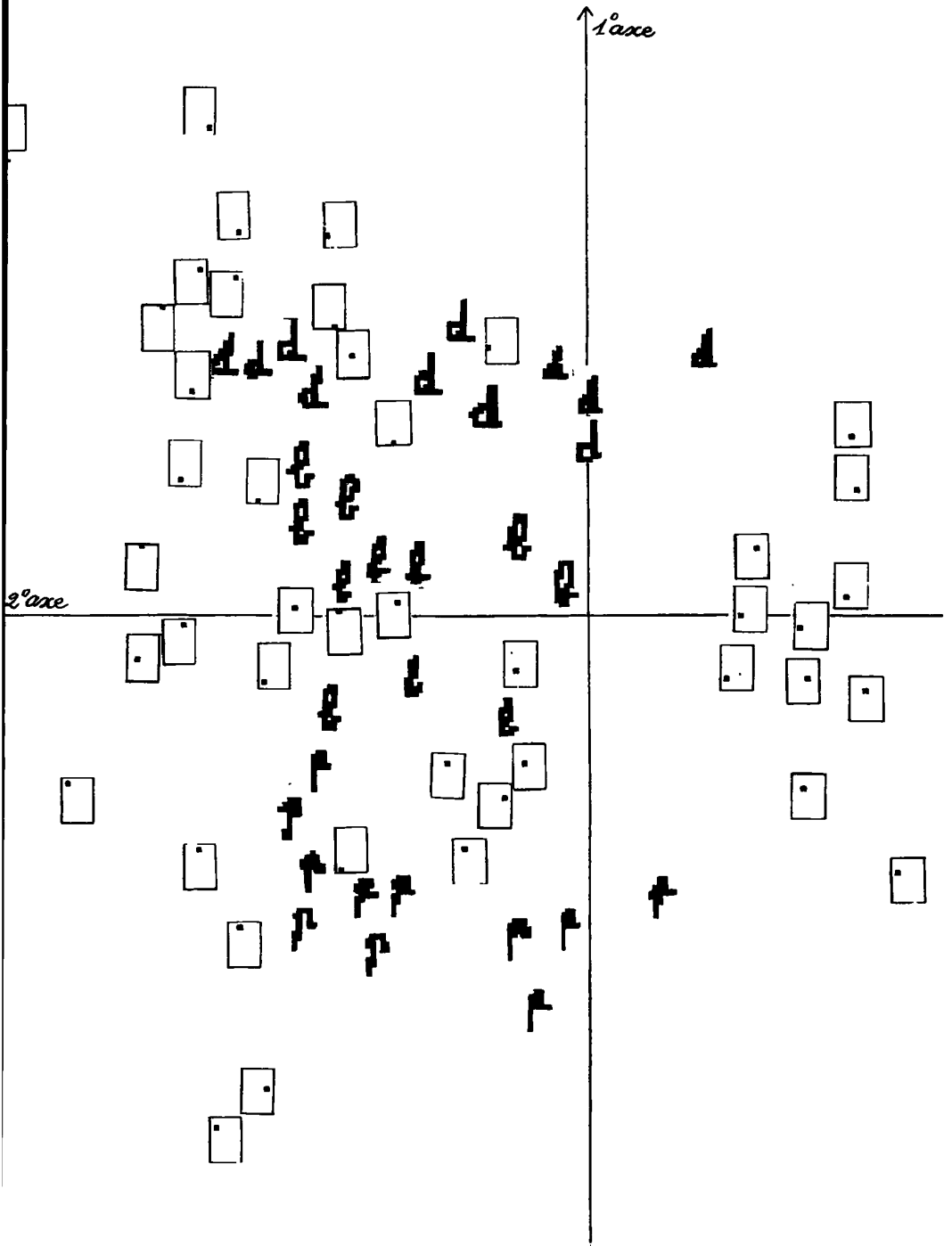
Considérons d'abord la représentation des ensembles I et C dans le plan des axes 1×2 . Chaque image i est exactement figurée à la place que lui assigne l'analyse factorielle, par la mosaïque de carrés noirs et blancs qui a servi à en coder le tracé manuel. Chaque cellule c est de même inscrite comme un carré noir de 1 mm^2 dans un cadre de 10×7 . (On notera toutefois que 19 des cellules de la périphérie du cadre ne sont pas représentées : c'est qu'elles ne sont jamais, ou presque jamais rencontrées par le trait : d'où dans le tableau $k_{I \times C}$ une

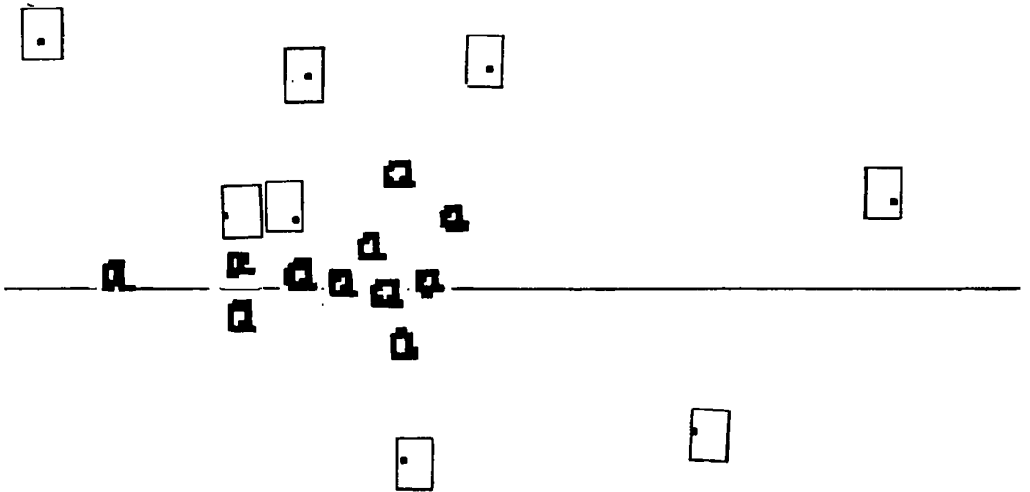


Exemple de codage d'une lettre et de sa translattée.

Grille $G_1 > 0$.Grille d: $G_2 > 0, G_4 > 0$ Grille $G_2 > 0$ Grille b: $G_2 > 0, |G_4|$ faibleGrille a: $G_2 < 0$ Grille $G_4 < 0$ Grille f: $G_2 > 0, G_4 < 0$

Analyse du tableau des descriptions
d'images: grilles des facteurs 1 et 2,
et caractéristiques des formes a, b, d, f.





*Analyse du tableau des descriptions
d'images: on a figuré dans leur
cadre les images comme les cellules*

colonne nulle ou quasi nulle). On voit que les quatre sous-ensembles a, b, d, p (de 11 images chacun) peuvent dans le plan des axes 1×2 être enfermés respectivement dans quatre rectangles non empiétant.

Or qu'est la valeur $F_\alpha(i)$ du α -ème facteur pour l'image i ? C'est d'après la formule de transition de l'analyse des correspondances :

$$F_\alpha(i) = \lambda_\alpha^{-N^2} \sum \{k(i,c) G_\alpha(c) / k(i) | c \in C\}$$

(où on note comme d'usage : λ_α valeur propre, $k(i)$ total de la ligne i , $G_\alpha(c)$ facteur pour la cellule c). Il est donc facile de passer des rectangles tracés dans le plan 1×2 et définis par des inégalités en F_1 et F_2 à un prisme (ou hyperprisme) de R^{70} défini par des inégalités portant sur des combinaisons linéaires des $k(i,c)$ (dont les coefficients sont les $G_\alpha(c)$). Ainsi se trouve résolu le problème de la discrimination par des cloisons.

De plus, ces cloisons sont définies par les $G_\alpha(c)$ qui, étant des fonctions sur C , peuvent être considérées comme des grilles. On a figuré séparément les parties positives et négatives de G_1 et G_2 ; par exemple dans la grille $G_2 < 0$, les carrés où $G_2 \geq 0$ sont laissés blancs ; ceux où $G_2 < 0$ sont noircis avec une intensité qui croît comme la valeur absolue de G_2 . On notera que cette grille $G_2 < 0$ dessine un a , lettre dont la forme a été caractérisée par $F_2(i) < 0$: dans le plan 1×2 issu de l'analyse de correspondance les i qui représentent des a sont, en effet nettement groupés avec autour d'eux la plupart des cellules c dont on se sert pour écrire un a . On a de même tenté de faire des grilles pour b, d, p ; en noircissant le plus les carrés qui tombent dans les mêmes zones que ces formes et dégradant progressivement la teinte. On rapprochera respectivement : d de $G_2 > 0$, b de $G_1 > 0$, p de $G_2 < 0$. Ces grilles sont moins clairement interprétables que celle de a .

On voit que le dessin des facteurs G_1 et G_2 résulte de la forme particulière des images de notre échantillon I : si on avait pris un échantillon I d'images très diverses, les premières associations entre cellules révélées par l'analyse factorielle eussent été, au contraire, les proximités géométriques dans le cadre rectangulaire : on aurait eu vraisemblablement un facteur G_1 positif en haut, négatif en bas ; puis un facteur G_2 positif à droite, négatif à gauche ; autrement dit les facteurs G_1 et G_2 eussent été l'ordonnée et l'abscisse des cellules.

3 Essai de reconnaissance de formes

Il se trouve que l'analyse des correspondances sans qu'on lui fournisse d'information relative à l'appartenance des images à une forme (ces onze-ci sont une même lettre ; ces onze-là en sont une autre etc.) a regroupé les images suivant leurs formes. Toutefois pour achever la reconnaissance des formes, il faut posséder une règle universelle valable non seulement pour 44 images données, mais pour l'infinité des tracés de lettres possibles.

Avons-nous atteint cette perfection ? En tout cas il est possible d'après la formule de transition rappelée ci-dessus de calculer

les facteurs $F_{\alpha}(i')$ pour une image supplémentaire i' . Cela a été fait pour 20 tracés de lettres : les b, d, p se sont tous placés dans les zones du plan 1×2 qui leur étaient assignées ; les a sont plus instables : cela tient sans doute à l'importance de la taille de la lettre.

4 Suggestions de recherches

Il serait instructif de faire une analyse avec plus de 4 classes: essayer par exemple de séparer les 10 chiffres. On améliorerait peut-être la séparation en introduisant dans l'analyse, avec chaque image, toutes les translatées contenues dans le cadre. Il faudrait sans doute utiliser plus de deux facteurs.

On sait que dans la perception d'une forme la direction du trait joue un rôle essentiel : une image est perçue plutôt comme un ensemble de traits que comme un ensemble de points. Evidemment, les traits ne sont que des combinaisons de points : mais le codage par traits joue un rôle intermédiaire essentiel. Ici, on pourrait appeler trait t une paire de cellules (c, c') de C ayant au moins un point commun. Dans la grille 10×7 , il y a 60 traits horizontaux 63 traits verticaux, 54 traits obliques suivant chaque direction à 45° sur les axes ; soit au total un ensemble T de 231 traits, dont quelques dizaines voisins des marges sont sans doute rares. On dira qu'une image i possède le trait $t = (c, c')$ si : $k(i, c) = k(i, c') = 1$; on posera alors $k(i, t) = 1$; et autrement (si $k(i, c) \times k(i, c') = 0$) on posera $k(i, t) = 0$. D'où un tableau k_{IT} à analyser.

Sur une page écrite à la main sans contrainte, les lettres ne sont pas enfermées dans des cadres rectangulaires déterminés ; et la segmentation de la chaîne manuscrite en lettres est sans doute l'étape première, et non la moins ardue, de la lecture automatique. A supposer cette étape accomplie, on pourrait cadrer chaque lettre dans un rectangle ; et, seulement alors subdiviser le rectangle en cellules, puis procéder un peu comme dans la présente note. Un tel cadrage pourrait même être utile pour les lettres isolées que nous traitons ici.