

a 146457

REVUE
D'HISTOIRE DES TEXTES

TOME CINQUIÈME

1975



ÉDITIONS DU
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

15, QUAI ANATOLE-FRANÇIS — 75700 PARIS
1977

TRAITEMENTS OPTIQUES D'ÉCRITURES

Au moyen de chaînes de traitement optiques, des éléments caractéristiques de l'écriture sont extraits des données essentiellement visuelles que constituent les morphologies de lettres manuscrites ou typographiées. Le degré de ressemblance entre deux écritures et la régularité de chacune d'elles sont mesurés; les traitements analogiques conduisent à la réalisation de lettres-moyennes représentatives d'une écriture aussi bien qu'à la détection d'éléments propres au scribe qui a transcrit un texte.

Outre la description des principes mis en œuvre et l'analyse des moyens expérimentaux utilisés, des résultats interprétables par le paléographe illustrent les travaux consignés dans cet article.

Nous nous proposons de décrire ici les études effectuées sur des écritures manuscrites au cours des cinq dernières années au laboratoire de Physique générale et Optique de Besançon, études suscitées par l'Institut de Recherche et d'Histoire des Textes (Paris).

En suivant l'ordre chronologique nous précisons à la fois les propos et les aboutissants de chaque étape de recherche en insistant sur les limites des techniques mises en œuvre. Le champ d'application d'une méthode peut se définir soit à partir d'un point de vue paléographique — en considérant les types d'écriture susceptibles d'être traités — soit en regardant le côté physique de l'analyse réalisée.

Ceci justifie *a priori* les passages apparemment brutaux de la paléographie à l'optique et vice-versa; la connexion constante des deux aspects étant naturelle du fait que les techniques de traitement analogique sont ici orientées essentiellement en vue de leur application à l'étude et à l'expertise d'écritures.

Dans tous les travaux dont nous rendons compte ci-après, on regarde l'écriture comme une « image » optique. Le but des traitements est alors de mettre en évidence les paramètres caractéristiques du graphisme des lettres. Les données extraites doivent permettre à la fois d'accroître la connaissance sur le mécanisme de l'écriture et de visualiser sous forme de résultats interprétables ce qui compose la morphologie d'une écriture et ce qui la rapproche ou la différencie d'une autre.

Tout en faisant ressortir les expériences effectuées par l'auteur, on ne saurait les scinder d'autres travaux assurant aux études sur les écritures menées par le groupe de traitement de l'information du LOBE¹,

1. LOBE : Laboratoire de Physique générale et Optique, Université de Besançon, 25030 Besançon Cedex.

le tout constituant un ensemble qui traduit l'évolution des recherches, leurs infléchissements et leur état actuel.

N'était une connaissance des concepts physiques mis en œuvre, il serait difficile, voire dangereux, d'interpréter directement les réponses fournies par une chaîne de traitement. Aussi est-il nécessaire d'introduire quelques notions sur les mécanismes décrivant le transfert d'information et le traitement analogique de données dans un système optique.

Le classement des activités se présente suivant deux volets :

— les procédés d'étude d'écritures non liées présentant des formes de lettres régulières et « stables » (mesure d'un degré de ressemblance entre deux écritures, calcul du graphisme moyen d'une lettre d'une écriture donnée) ;

— les méthodes plus générales d'expertise d'écriture destinées à faire ressortir les caractéristiques d'un scripteur (mise en évidence d'un facteur de forme).

Des indications sur les travaux en cours et l'orientation actuelle des recherches sont données en ce qui concerne chaque direction d'étude.

I. — REGARDS SUR UNE LETTRE. DOMAINE GÉOMÉTRIQUE ET DOMAINE SPECTRAL

Le dessin d'une lettre quelconque peut se décrire comme un être mathématique représentant la répartition d'encre sur le support d'écriture (papier, peau, etc...). Soit $f(x, y)$ une telle fonction. Plutôt que de la décomposer en de très nombreux éléments simples (largeur et inclinaison des hastes, dimensions des boucles, largeur des panses, taille des empattements, etc..), comme le ferait un analyste, un système optique transpose l'information contenue dans $f(x, y)$ — pour en donner, par exemple, une image $f'(x, y)$ — en faisant intervenir un seul type de variable géométrique, à savoir les fréquences spatiales.

En traversant une pupille transparente, une onde lumineuse monochromatique se déforme et ses modifications de structure portent une information caractéristique de la pupille. Ainsi le dessin d'une lettre — ou de plusieurs — mise sous forme de transparence sur un négatif photographique sert de fonction d'entrée dans un système optique éclairé par un laser ; à sa traversée, le faisceau lumineux est diffracté suivant des directions déterminées par la forme de la lettre, notamment par son contour.

Collectés par un objectif, les rayons diffractés s'organisent dans le plan focal de celui-ci pour décrire une *figure de diffraction à l'infini* — encore appelée *selvte*, dont les termes constitutifs correspondent à des fréquences spatiales, c'est-à-dire à des directions de diffraction (au niveau de l'objet). Ainsi, plus la diffraction suivant une direction est intense, plus la brillance du terme spectral correspondant à cette direction est importante.

L'information géométrique afférente au dessin d'une lettre (ou de toute autre forme, notamment d'un groupe de lettres tel que mots, textes) est alors transposée en une information spectrale; l'opération mathématique représentant le passage du domaine géométrique au domaine spectral s'appelle une « transformation de Fourier », d'où le nom de « plan de Fourier » communément donné au plan focal de l'objectif qui effectue ce calcul de façon analogique. Afin de donner une idée de ce qu'est l'analyse de Fourier, on peut comparer la description d'un objet en fréquences spatiales à la décomposition d'un son en fréquences temporelles : tout électrophone courant reconstitue grossièrement un morceau de musique enregistré sur un disque alors qu'une chaîne haute-fidélité restitue mieux le « volume » et le timbre de chaque instrument ; dans les deux cas, les basses fréquences temporelles sont véhiculées par l'appareil, mais les hautes fréquences déterminant les éléments « fins » du son ne passent que dans la chaîne Hi-Fi. Ceci illustre l'influence des fréquences lors du transfert d'une information. De même, dans le domaine spatial, les basses fréquences traduisent la continuité, c'est-à-dire l'uniformité des objets, alors que les détails fins (contours, points de discontinuité), caractéristiques de certains écarts entre les formes, correspondent à des fréquences élevées qui s'étalent sur les « bords » des spectres.

Il faut enfin noter que la transformation de Fourier est une opération réversible : le passage du domaine des fréquences spatiales au domaine de la variable géométrique se fait de la même façon (au moyen d'un objectif) que la transposition de l'objet en son spectre. L'espace géométrique et l'espace de Fourier sont des domaines réciproques l'un de l'autre.

La planche III illustre ces notions : les spectres de la lettre A dessinée suivant différents types d'écriture présentent d'autant plus de termes élevés en fréquences que la lettre contient de détails. En particulier, la lettre tracée peut être considérée comme le dessin de la lettre échantillonnée par un grand nombre de points équidistants, ce qui explique que son spectre comporte des termes à la fréquence spatiale élevée correspondant à la fréquence de présence d'un point dans la trame (nombre de points par millimètre). Il convient de noter que plus un détail (ou une lettre) est petit, plus son spectre est étendu.

Pour faire ressortir ce qui caractérise et ce qui différencie le graphisme des lettres, le tracé de tous les spectres des signes de l'alphabet hébraïque a été enregistré, ceci pour les types d'écriture de deux époques différentes (pl. IV). Les résultats permettent la comparaison visuelle entre l'évolution de l'écriture telle qu'on la voit sur le dessin des lettres et sur leur spectre où l'influence des détails semble accrue.

Cette comparaison visuelle demeure subjective.

Une suite naturelle à cette recherche consiste à établir un critère de détermination de la ressemblance entre deux formes et à mesurer celle-ci ; ce qui fut fait.

II. — DEGRÉ DE RESSEMBLANCE ENTRE DEUX FORMES VOISINES. DÉFINITION ET MESURE

La détermination des similarités et des dissemblances présentées par plusieurs objets peut s'effectuer par la comparaison des bandes de fréquences caractéristiques de leurs spectres.

D'autre part, la mise en mémoire holographique d'une forme quelconque permet de restituer l'image de celle-ci après lui avoir fait subir un filtrage de fréquences spatiales. En effet, en enregistrant sur une émulsion photographique un hologramme de Fourier — superposition d'un spectre et d'une onde plane de référence — on peut modifier à loisir l'amplitude de l'une des ondes qui interfèrent et par-là même conférer une pondération à chaque bande de fréquences. Lors de sa lecture ou de sa superposition avec un spectre, l'hologramme joue le rôle d'un filtre et permet de restituer une image dans laquelle on peut accroître à volonté l'importance des discontinuités ou faire ressortir des directions.

Remarque : Dans le cas des lettres, par exemple, on sait modifier l'importance relative des contours. Sans avoir recours à l'holographie, nous avons plusieurs fois utilisé une méthode simple de filtrage des fréquences spatiales pour améliorer la qualité d'une image, en particulier pour rehausser la lisibilité d'un texte.

Enfin, l'utilisation de tels hologrammes dans un montage optique permet d'effectuer des opérations compliquées, notamment d'obtenir la fonction de corrélation entre deux formes; celle-ci représente ce qui résulte du balayage d'une fonction par une autre. Ainsi l'inter-corrélation entre deux lettres présente un maximum correspondant à la quantité de lumière qui peut traverser les deux lettres lorsque celles-ci sont superposées au mieux et le profil de cette fonction d'inter-corrélation traduit les similarités dans les contours des lettres.

Ces constatations montrent que les techniques de l'optique cohérente semblent particulièrement bien appropriées pour réaliser des comparaisons quantitatives entre deux formes.

Nous avons proposé une définition du degré de ressemblance d'une forme géométrique indiquée 2 par rapport à une autre, voisine, indiquée 1 par le rapport :

$$R_{21} = \frac{\rho_{21}}{\rho_{11}} = \frac{\text{Max} [p_2(x, y) * p_1(x, y)]}{\text{Max} [p_1(x, y) * p_1(x, y)]}$$

où le signe * dénote l'opération de corrélation, et où Max [] signifie : valeur maximum de la fonction entre crochets.

Si $r_2(x, y)$ et $r_1(x, y)$ décrivent le dessin des lettres indicées 2 et 1, p_2 et p_1 représentent les images de ces lettres obtenues après un certain filtrage de fréquences spatiales.

La corrélation $p_2(x, y) * p_1(x, y)$ diffère donc de $r_2(x, y) * r_1(x, y)$ en donnant une prépondérance aux contours et en diminuant l'influence de la région centrale dans le dessin des lettres. Ainsi défini le degré de ressemblance d'une lettre par rapport à une autre est égal au niveau maximum d'inter-corrélation des deux lettres « pondérées » divisé par le niveau maximum d'auto-corrélation de l'une d'elles. Pour que cette définition ait un sens et rende compte des évidences du sens commun, on impose de plus une condition de réversibilité en exigeant des résultats tels que le degré de ressemblance d'une lettre (2) par rapport à une lettre (1) soit égal au degré de ressemblance de la lettre (1) par rapport à la lettre (2). Nous avons réalisé des filtres-hologrammes en respectant ces impératifs et nous avons pu ainsi mesurer des degrés de ressemblance. Il faut noter que cette méthode de mesure ne fournit des résultats justes que si elle est appliquée à des lettres de formes voisines et de mêmes dimensions. Elle perd tout son sens dès que l'on s'écarte de ces conditions d'utilisation.

EXAMEN DE L'ÉCRITURE D'UN MANUSCRIT

Ayant défini un critère de mesure de la ressemblance entre deux formes, nous avons réalisé un corrélateur optique fonctionnant avec un laser hélium-néon. Les réponses sont analysées par une caméra de télévision reliée à un oscilloscope sur lequel on mesure le niveau maximum des fonctions de corrélation. Ce système opto-électronique a été mis en œuvre pour étudier l'écriture d'un manuscrit hébraïque préalablement expertisé par un groupe de paléographes et répertorié H. 316 à la Bibliothèque nationale (Paris).

L'examen de toutes les lettres d'un manuscrit d'environ 400 folios représente, dans l'état actuel de l'appareillage, un travail extrêmement long; non seulement les mesures elles-mêmes ne se font pas rapidement, mais surtout la fabrication d'hologrammes adaptés aux problèmes de la ressemblance et l'enregistrement sur films spéciaux¹ de tout le texte de l'ouvrage nécessitent un temps très grand. Heureusement, les techniques d'échantillonnage permettent de réduire considérablement la quantité de travail tout en conduisant à des résultats statistiques interprétables. En premier lieu, nous avons cherché à sélectionner les lettres susceptibles

1. Pour des raisons techniques, les lettres devaient être réduites à des dimensions d'environ 1×1 mm. et photographiées sur des émulsions à bon contraste et à résolution élevée. Nous utilisons des plaques « Collodium orthochromatique » (marque Guilleminot). Le fabricant ayant cessé de produire ce matériau, nous l'avons depuis avantageusement remplacé par du film lith (0.81, Agfa).

d'être les plus significatives parmi les vingt-sept signes de l'alphabet hébraïque. Pour ce faire, nous avons tiré dix lettres qui comportent dans leur morphologie tous les éléments qui servent à constituer les dessins des vingt-sept signes (hastes, empattements, « barres » horizontales et verticales, boucles). Ces dix lettres ont été toutes corrélées entre elles dans les mêmes conditions, ceci en construisant et en mettant en œuvre un filtre-hologramme à adaptations multiples. Dans le tableau de résultats obtenus, on constate que parmi les cent fonctions de corrélation (10×10), une est particulièrement importante : il s'agit de l'auto-corrélation de la lettre aleph. Elle atteint une intensité qui, d'une part, est plus élevée que celles des autres fonctions, et, d'autre part, est importante devant celle des inter-corrélations de la lettre aleph avec les neuf autres lettres mises en jeu.

À la vue de ce résultat, nous avons estimé que, dans un premier temps, la seule lettre aleph pouvait servir d'étalon pour fournir des éléments de datation d'un texte hébraïque. Le problème est alors réduit à la comparaison de toutes les lettres aleph du texte entre elles.

Dans une seconde étape, nous avons diminué le nombre de corrélations à effectuer en sélectionnant à l'aide d'une table de nombres aléatoires une quantité restreinte de folios dans le manuscrit, ceci en se fixant un taux de sondage de l'ordre de 8 % de façon à ce que les résultats reposent sur des données utilisables en statistique.

Une série d'expériences a été conduite sur trois types d'écritures du manuscrit II. 316. Chaque lettre échantillonnée était « corrélée » avec cinq lettres-types de morphologies différentes, et une moyenne calculée sur l'ensemble des résultats permet de chiffrer le degré de ressemblance entre deux écritures. Les degrés de ressemblance moyens ainsi mesurés confirment la présence — préalablement découverte par les paléographes — de trois scribes différents dans le tracé des textes étudiés.

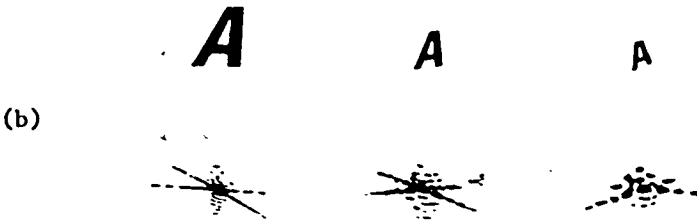
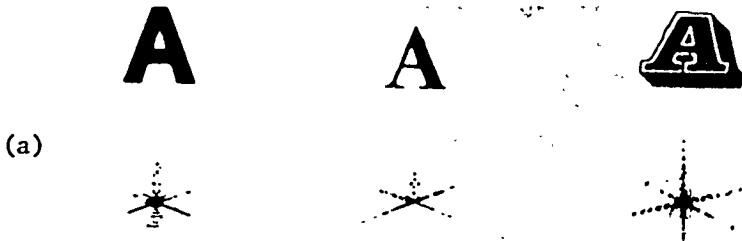
La planche V montre des portions des écritures étudiées avec en regard l'indication des degrés de ressemblance calculés.

Il ressort des travaux réalisés à cet effet qu'une lecture automatique d'une haute fidélité est possible sur une écriture hébraïque manuscrite, sous condition d'adapter l'hologramme au problème posé.

La planche VI traduit avec quel poids nous pouvons affirmer ceci.

LIMITES D'UTILISATION DE LA MESURE DE RESEMBLANCES

Les résultats dont il est fait état précédemment ont souvent donné naissance à un optimisme démesuré quant aux développements qu'ils devraient susciter dans l'étude des écritures. Tout en soulignant les espérances légitimes qu'ils apportent, principalement pour un accroissement de l'usage des techniques optiques en lecture automatique de caractères typographiques et manuscrits, il convient de préciser les conditions qu'il est nécessaire de respecter pour calculer la ressemblance entre diverses



(a) Différents graphismes de la lettre A et leurs spectres de Fourier. (b) Lettres de dimensions et d'orientations différentes et leur spectre. (c) Lettre traniée et son spectre.



א	ב	ט	ד	ה	ו	ז	ח	ט	י
א	ב	ט	ד	ה	ו	ז	ח	ט	י
ת	ר	ה	ק	ל	מ	נ	ס	ע	פ
ת	ר	ה	ק	ל	מ	נ	ס	ע	פ
ס	ע	ד	כ	ם	ה	ז	ו	ג	ה
ס	ע	ד	כ	ם	ה	ז	ו	ג	ה

Lettres hébraïques «modernes» et «anciennes» et leurs spectres.

SCRIBE n° 4
(réfèrece)

il a écrit tous les c_i

i = 1, 2, 3, 4, 5

c_j ∈ {c_i}

j = 1, 2, 3, 4, 5

$\rho(\bar{c}_i, \bar{c}_j) = 0.52$

היה אברהם אומר אתה אבינו אהב
נרין ולא זכר לבי אנו במהרה
איהו וצפיהו כל יוביה וגה דהה אכנה
איהב פטרנה ינהה לאגב
אני בטב ה רבני אנ וייב לדינ
איהב לא ריטה הוא ולא ביע אלנא
פטר ינה וכומו באינכ הינהה חיי
יורא החלה בפטישה סיב כאינכ פטר
טעיה הכלא ואגמא קטלה רנה בטניה

SCRIBE n° 2

il a écrit tous les c_i

i = 1, 2, ..., n'

$\rho(\bar{c}_i, \bar{c}_j) = 0.26$

$\rho' / \rho = 0.50$

יהא ביהא אנים היא
ביתבד הילכד ביכה ב
קיעה פטיה דלא ינהה
הם זכני הצור כיוו דתורה הוואב ר
הנה ליה לטיני אי איהי
לית הווקיה וכיוו דהא עבר הני
א דפטיט חייב תרע דהא רכניו
יהה דה שניעון בתורה דהא יהיב
צטיר יהנא אינהה תפיהו קיה י
יעיהא לפטור ררר דה הנ

SCRIBE n° 5

il a écrit tous les c_i

i = 1, 2, ..., n'

$\rho(\bar{c}_i, \bar{c}_j) = 0.21$

$\rho' / \rho = 0.40$

פיר וניה ואכיקטן קמו נהו ברטוה
א הט דלילה איטי הווא אבא
בט לטויה דנ פרוק דופוא אמא
מנרא חנה לקמיה דרב חסדא אמא לה
יכני כד תקט מטויה בליקחין
ה שרבי יגמי מטויה כהרענה ארבע
הה חנה לקמיה דרב עמי אמא להו היי
י רב נהואי כה אידי אמא נהוואי ינה...

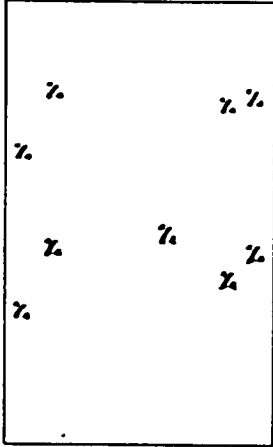
Fragments de trois types d'écritures hébraïques (XIV^e siècle) et degrés de ressemblances mesurés entre ces écritures.

(a) 

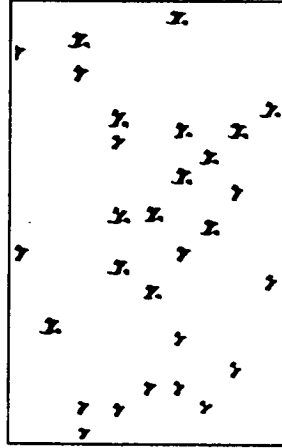
(b) 

(c) 

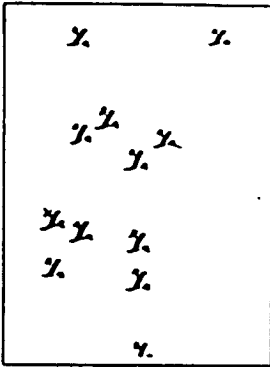
(a) Ligne d'écriture manuscrite. (b) «Points» obtenus par corrélation optique entre la ligne (a) et une lettre aleph de même écriture; *détection* des alephs de la ligne (a).
(c) Coupe photométrique de (b).



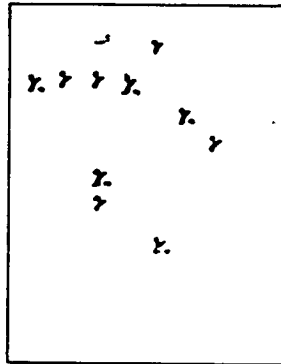
folio 101 r



folio 269 r



folio 72 r

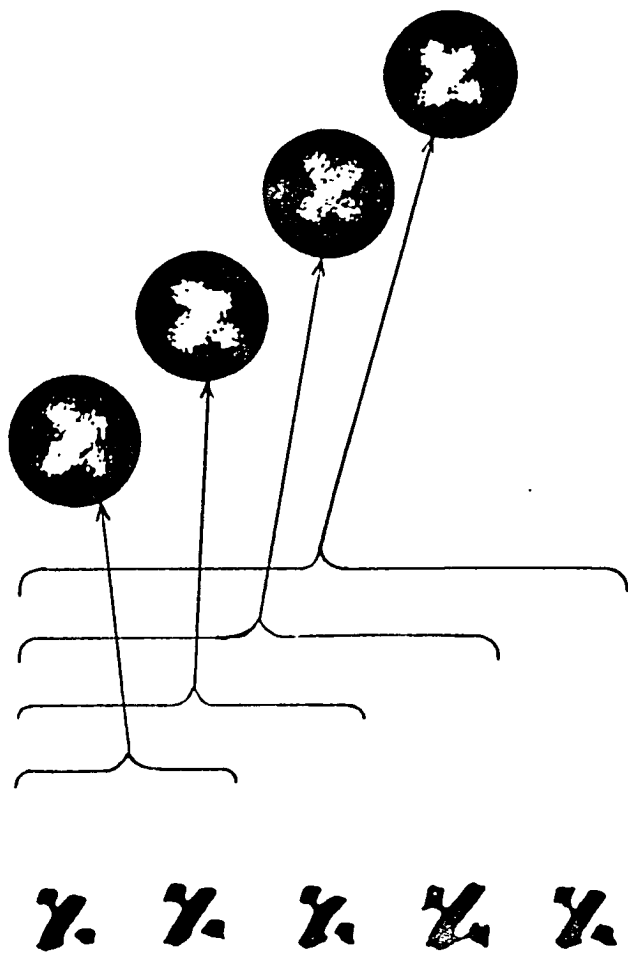


folio 14 r

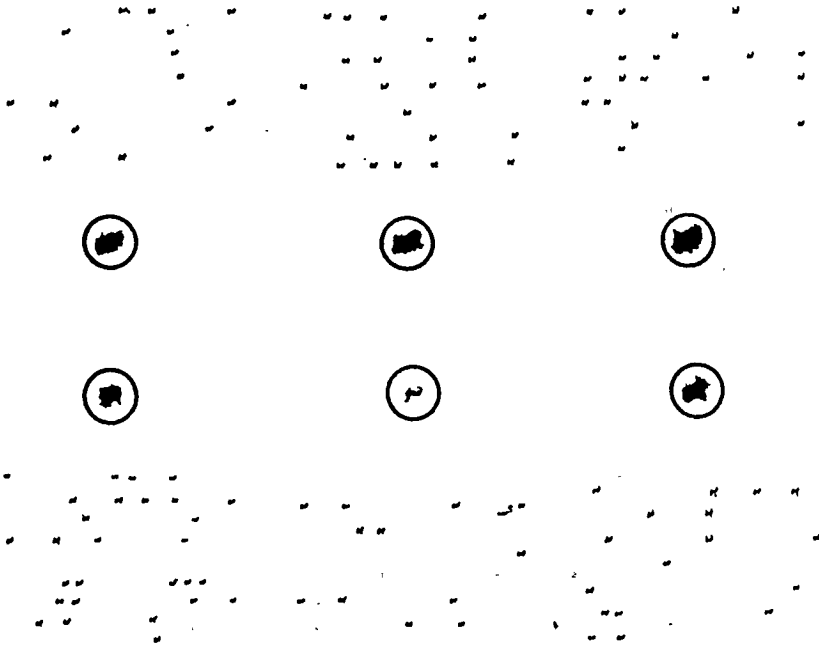
scribe
no 3

scribe
no 7

Lettres alephs telles qu'elles se présentent après effaçage de toutes les autres lettres d'un texte (divers fragments de 4 folios différents écrits par 2 scribes).



Superposition de 2, 3, 4, 5 lettres écrites par un même scribe.



Lettres alephs isolées dans 6 morceaux de textes écrits par un même scribe et lettres-moyennes construites à partir de ces 6 cartes de données.

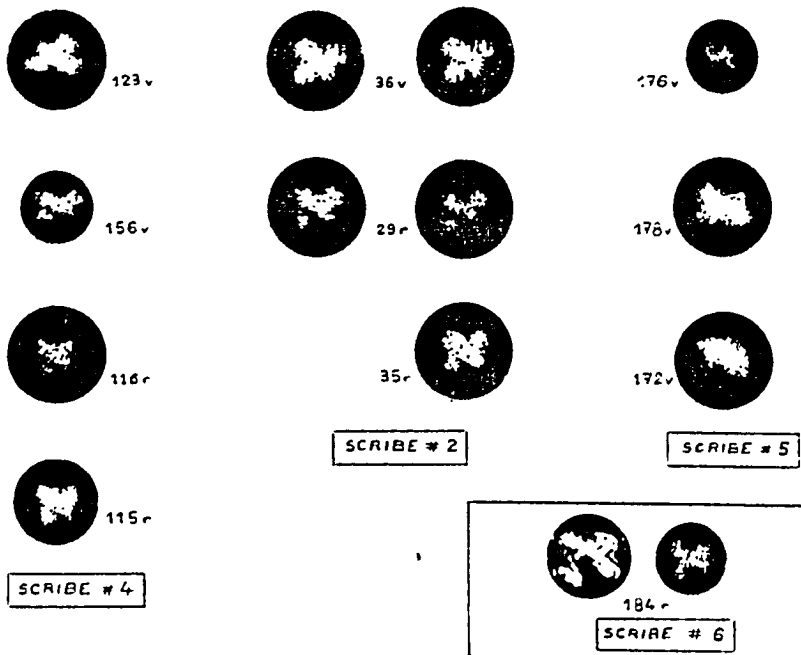


Tableau comparatif de lettres moyennes issues de 4 écritures différentes.

formes. Nous avons signalé précédemment que la détermination d'une ressemblance n'avait de sens que dans la comparaison de formes « voisines ». Ce vocable volontairement ambigu traduit la difficulté de définir clairement une frontière entre deux morphologies qui « se ressemblent » ou « ne se ressemblent pas ». On pourrait toujours fabriquer un hologramme qui conduirait à une corrélation non négligeable entre deux formes d'aspect très éloigné. Les « mesures » que sa mise en œuvre fournirait ne reposeraient plus sur les critères que l'expérience et la réflexion nous imposent et n'auraient alors aucune valeur.

D'autre part, il semble possible de mesurer une ressemblance sur certaines écritures cursives (sous condition que celles-ci soient assez régulières pour présenter des lettres de dessins voisins), mais ceci soulève de nouvelles difficultés dues au fait que la « machine optique » doit distinguer le début et la fin de chaque lettre.

Ce problème de la séparation des lettres dans une écriture cursive n'a pas encore été abordé sous son aspect expérimental.

Indiquons enfin que les dimensions des lettres ne constituent pas un obstacle important à l'utilisation d'un corrélateur. Une étude complète a abouti à la réalisation d'hologrammes susceptibles d'effectuer des opérations sur des objets de taille variable; on conçoit aisément qu'un dispositif optique soit capable de modifier à loisir les dimensions d'une lettre.

En conclusion, rappelons que la mesure des similarités au moyen de la comparaison des domaines de fréquences caractéristiques des spectres conduit à la détermination d'un degré de ressemblance entre deux formes.

La réalisation de filtres à adaptations multiples a conduit à une prise de décision quant au choix d'une lettre susceptible de servir d'élément de base dans l'étude des écritures d'une langue (aleph pour l'hébreu); elle a également facilité l'étude des variations du tracé d'une lettre par la comparaison simultanée à cinq étalons de toute lettre mise à l'entrée du dispositif.

Des mesures de ressemblance calculées sur divers folios d'un manuscrit hébraïque ont permis de reconnaître aussi bien deux types d'écriture (ashkenaze et sefarad) que deux scribes différents usant de la même écriture (ashkenaze). Parallèlement, les hologrammes construits pour ces expériences ont montré leur efficacité en tant que composants de lecture automatique de caractères, fussent-ils manuscrits, au moins dans le cadre d'écritures régulières et non liées.

III. — PRINCIPES ET RÉALISATION DE LETTRES MOYENNES

Depuis que la paléographie existe, les érudits ont toujours cherché à décrire les écritures et à présenter des aspects typiques des formes de lettres. Au-delà des lettres-types fabriquées artificiellement ou extraites

des manuscrits, plusieurs auteurs ont dessiné des « lettres-moyennes » en effectuant une statistique sur les éléments constitutifs des lettres. Outre leur intérêt intrinsèque (on rend visible quelque chose qui représente un ensemble de lettres sorties d'un texte écrit), de telles lettres-moyennes comportent en elles-mêmes un grand nombre de facteurs qui déterminent une écriture. Par leur calcul, on peut donc espérer caractériser complètement aussi bien un type d'écriture, qu'un scribe, qu'une écriture utilisée dans un scriptorium à telle ou telle époque.

Leur connaissance permet, à la limite, de ramener l'étude d'une écriture à celle des lettres moyennes qui la représentent : leurs comparaisons, la connaissance de leurs profils, la détermination de leurs évolutions fournissent des renseignements aussi complets que ceux obtenus en effectuant les mêmes opérations de façon fastidieuse, lettre par lettre. Les résultats sont alors directement utilisables par les paléographes.

Bien que ceci ne constitue qu'une approche du problème, mentionnons qu'une esquisse de forme-moyenne a été décrite dans le chapitre précédent : sans toutefois donner accès à la lettre-moyenne elle-même, les filtres hologrammes à adaptations multiples conduisent à une moyenne sur les mesures d'inter-corrélations entre les lettres mises en mémoire dans l'hologramme et toute lettre introduite à l'entrée du corrélateur.

Avant de développer plusieurs méthodes optiques d'obtention de formes moyennes, il faut remarquer que la notion de superposition existe implicitement dans le processus de formation d'un spectre. En effet, aucune limite restrictive n'est imposée quant aux formes contenues dans une fonction-objet lors du calcul du spectre de celle-ci par un objectif. Dans un tel objet peuvent être réparties plusieurs lettres de dessin quelconque, constituant une fonction dont la transformée de Fourier est la somme des transformées de Fourier des fonctions (lettres) composantes. Ainsi, le spectre d'un texte est la superposition exacte des spectres de toutes les lettres qui le constituent.

Un objectif conduit d'un ensemble de dessins de la même lettre (aleph par exemple) au spectre-moyen correspondant, où se superposent les fréquences spatiales constitutives des formes traitées. Parmi les bandes de fréquences représentatives de l'écriture analysée s'affichent entre autres celles qui sont propres au scribe : elles sont la transposition en termes de Fourier de la « fonction d'appareil » qu'est la main du scribe.

PRINCIPE DU CALCUL OPTIQUE D'UNE MOYENNE

« Extraction » des données à traiter

Par quelque méthode qu'elle soit faite, la réalisation d'une lettre moyenne nécessite une étape où les lettres que l'on veut « additionner » sont séparées des autres lettres du texte.

Un procédé consiste à extraire d'un fragment d'écriture les caractéristiques

tères à traiter (et à utiliser ensuite une technique de mixage pour réaliser leur moyenne). Ces prélèvements, qu'ils se fassent empiriquement par découpage ou très rapidement par un capteur élaboré (capteur optique par exemple) introduisent toujours un mouvement des données — c'est-à-dire des lettres — ce qui les entache de défauts d'orientation.

Un autre procédé consiste à supprimer d'un fragment d'écriture tous les caractères que l'on ne veut pas étudier. Ne restent alors que les données à traiter, sans que celles-ci aient subi aucun déplacement ni aucune modification. C'est ce dernier type de séparation des données que nous utilisons. Au stade du laboratoire, nous éliminons les lettres non retenues pour l'étude au moyen d'une peinture blanche appliquée sur des photographies de textes. Bien que fastidieux, cet artifice qui consiste à ne laisser dans un fragment d'écriture qu'un seul type de lettres (les a par exemple) conduit à des résultats qui intéressent le paléographe ; ils permettent de voir comment le tracé d'une lettre évolue dans un texte en dehors de toutes les autres lettres parasites. En témoigne la planche VII, où se côtoient toutes les lettres aleph écrites par deux scribes à quatre endroits différents d'un même ouvrage.

« Mixage » des données

Calculer une moyenne à partir d'un document où ne figurent que des dessins d'une même lettre consiste avant tout à « empiler » ces dessins les uns sur les autres. La morphologie globale obtenue contient les propriétés de l'écriture manuscrite de la lettre avec laquelle est faite l'expérience. Dans ce qui suit, nous allons décrire plusieurs méthodes permettant d'obtenir des formes moyennes ; dans tous les cas, on joue sur le trajet des rayons lumineux pour que ceux-ci, après traversée des formes transparentes à traiter (figées sur négatif photographique) et des divers composants optiques, fabriquent au moins une image où toutes les formes se superposent. Suivant les procédés utilisés, on peut faire subir un filtrage plus ou moins élaboré aux fréquences spatiales qui décrivent cette image.

Afin de superposer des formes, nous recourons au processus suivant :

- dans un premier temps, on détermine le barycentre d'énergie — disons plus grossièrement le « centre » — de chaque forme ;
- à la répartition des formes à traiter, on annexe la répartition de leurs centres ;

— on effectue, par un moyen quelconque, l'auto-corrélation de l'ensemble ou, au moins, l'inter-corrélation des deux répartitions.

Pour comprendre comment ceci mène à une forme moyenne, considérons deux cartes, l'une constituée par une distribution de lettres, l'autre — située dans le même plan que celle-ci, à une distance fixée d — étant un ensemble de points identique à la distribution des centres des lettres

à superposer. Si l'on déplaçait en bloc la carte des points sur la carte des lettres, on trouverait de nombreuses positions où un point chevauche une lettre, mais il en existerait une et une seule où tous les points recouvriraient le centre des lettres à traiter : celle-ci correspond à un déplacement égal à d de la répartition de points par rapport à sa localisation initiale. Dans le voisinage de cette position les points balaient, tous simultanément, la lettre avec laquelle ils sont associés. Ceci se traduit, sur la figure d'auto-corrélation de l'ensemble (lettres-points), par une superposition, située à la distance d du centre, de toutes les inter-corrélations entre chaque lettre et le point qui lui est associé. Or, la corrélation d'une forme quelconque avec un « point » n'est autre que la forme elle-même ; on obtient donc ainsi un empilement de lettres : les parties communes sont renforcées par addition et les éléments différents contribuent à former un fond plus ou moins uniforme. C'est cet aboutissement que, dans tout ce qui suit, nous appellerons « forme-moyenne ».

Cette description du processus appelle deux considérations quant à sa mise en œuvre :

— En expliquant le phénomène, nous avons supposé les « points » infiniment fins. En pratique, leur dimension maximale délimite la meilleure résolution géométrique que l'on puisse obtenir sur une forme moyenne. Par exemple, on ne saurait espérer une précision supérieure à 0,1 mm. sur la définition (au sens optique du mot) d'une lettre moyenne quand celle-ci est construite avec des « points » de 0,1 mm. de diamètre. De plus, la limitation de la qualité d'image obtenue dépend essentiellement du dispositif mis en œuvre lors de l'auto-corrélation.

— En principe, il est nécessaire de faire intervenir toutes les lettres d'une écriture pour fabriquer la lettre-moyenne de cette écriture.

Mais si un nombre restreint de lettres suffit à cela, la rapidité d'exécution est considérablement accrue. En fait, quand une écriture contient n fois la même lettre, il existe un écart entre la superposition de toutes les n lettres et la superposition de seulement p d'entre elles (p étant inférieur à n). On sait déterminer par le calcul le nombre minimal p qu'il suffit d'atteindre pour que cet écart ne dépasse pas une valeur fixée (que l'on peut rendre arbitrairement aussi petite que l'on veut).

Le nombre p est déterminé par la stabilité de l'écriture. En adoptant un modèle gaussien pour représenter la dispersion (c'est-à-dire les fluctuations) de l'écriture, la stabilité est décrite par l'écart quadratique moyen σ . Pour indiquer un ordre de grandeur, suivant ce modèle, pour un écart quadratique moyen de 0,2, il suffit de neuf lettres pour approcher la forme-moyenne à 10 % près alors qu'il en faut trente-trois pour l'atteindre avec une erreur inférieure à 5 %.

Connaissant la marge d'erreur inhérente au dispositif utilisé, on peut ainsi calculer le nombre minimal de lettres nécessaires à la représentation d'une écriture par ses lettres moyennes. Toutefois, on ne saurait mettre ceci à profit que lorsque l'écriture étudiée est préalablement connue.

PROCÉDÉS EXPÉRIMENTAUX DE RÉALISATION DE FORMES-MOYENNES¹

Construite par le biais des inter-corrélations entre divers graphismes et des points à eux appariés, une forme moyenne est obtenue par la réalisation de l'auto-corrélation d'un ensemble de données (lettres + points). Il en découle que, toutes les méthodes susceptibles de mener à une auto-corrélation peuvent être utilisées pour fabriquer des formes moyennes. De nombreux procédés analogiques effectuent cette opération; c'est le cas en particulier de tous les dispositifs de mesure du module d'une fonction de transfert optique. Parmi ces derniers, nous avons systématiquement éliminé l'emploi de techniques fournissant des résultats point par point pour éviter de trop longs temps d'exécution, ce qui explique pourquoi elles ne figurent pas dans les six types de méthodes décrites ci-après.

1. MISE EN ŒUVRE D'UN CORRÉLATEUR OPTIQUE FONCTIONNANT EN ÉCLAIRAGE COHÉRENT.

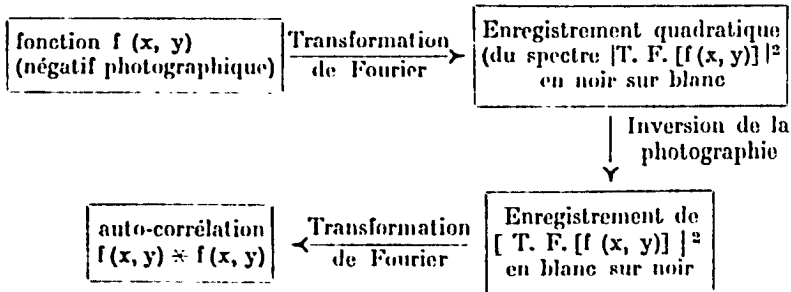
Il est possible de construire le filtre-hologramme d'une fonction quelconque et de corrélérer celle-ci avec elle-même au moyen d'un système analogue à celui utilisé pour les mesures de ressemblance. Toutefois, nous n'avons pas utilisé cette méthode pour deux raisons :

— la construction d'un hologramme adapté nécessite beaucoup de temps;

— les « points » figurant dans la fonction dont on veut faire l'auto-corrélation peuvent jouer eux-mêmes le rôle de référence dans la fabrication d'un hologramme, ce dont nous avons tiré parti comme en témoigne ce qui suit.

2. DIFFRACTION A L'INFINI DU MODULE CARRÉ DU SPECTRE D'UNE FONCTION.

Cette méthode consiste à enregistrer sur un récepteur photosensible une répartition proportionnelle au carré du module du spectre d'une fonction et à observer la figure de diffraction à l'infini de cette répartition. Le schéma synoptique suivant indique les étapes du processus :



1. Ce paragraphe fournit des renseignements techniques sur les diverses méthodes

Dans ce processus les « points » agissent comme des sources de référence. La photographie du spectre joue le rôle d'un hologramme de Fourier enregistré en source multiple. Chaque couple (lettre-point de référence) étant séparé par la même distance, et toutes les sources étant identiques, dans l'hologramme les spectres sont tous modulés par la même fréquence : ceci conduit à une restitution unique des images des lettres, autrement dit à la lettre-moyenne. Notons qu'une inversion photographique du spectre améliore considérablement la qualité du résultat (sans ce traitement la lettre-moyenne est noyée dans un bruit important). Pour conserver la linéarité que requiert par définition la corrélation, il est nécessaire d'enregistrer le spectre et son inverse suivant des conditions de développement convenablement appropriées. Bien qu'elle permette, au niveau spectral, de donner plus ou moins d'importance à certaines bandes de fréquences spatiales, l'utilisation de cette méthode s'accompagne de difficultés techniques, notamment en ce qui concerne le traitement des plaques photographiques ; aux conditions draconiennes imposées par le choix *a priori* du facteur de contraste s'ajoute l'impossibilité d'enregistrer fidèlement sur une émulsion ainsi développée un spectre dont la gamme d'éclairissement est très étendue. Aussi chaque expérience nécessite-t-elle de nombreux essais avant de fournir un résultat acceptable.

3. ENREGISTREMENT PUIS DIFFRACTION A L'INFINI DU « SPECTRE » D'UNE FONCTION MODULÉE PAR UN DIFFUSEUR.

Grossièrement, le principe de cette méthode est analogue à celui de la précédente. Un diffuseur de transparence $t(x, y)$ est placé devant la fonction $f(x, y)$ dont on veut construire l'auto-corrélation. L'ensemble produit une figure de diffraction d'aspect granuleux que l'on photographie dans un plan éloigné de l'objet (environ à 1 mètre de lui). L'information ainsi enregistrée est suffisamment proche de la figure de diffraction à l'infini du produit $t(x, y) \times f(x, y)$ pour se comporter comme sa transformée de Fourier. Il suffit alors de faire diffracter cette répartition pour observer, dans le plan focal d'une lentille, l'auto-corrélation de la fonction $f(x, y)$ multipliée par une fonction correspondant au diffuseur. La lettre moyenne recherchée est « couverte » par une structure de grains dont les dimensions dépendent du diffuseur utilisé. Toutes les fréquences interviennent avec le même poids.

Ce procédé est plus économique et plus rapide que le précédent puisqu'il ne requiert pas l'emploi d'émulsions photographiques spéciales (un film ordinaire développé sans précautions particulières suffit) et qu'il ne nécessite pas l'inversion de l'information enregistrée ; celle-ci se compose, en effet, d'autant de grains « noirs » que de grains « blancs » étalés

envisagées et utilisées et indique, pour chacune d'elles, son principe, ses avantages essentiels, ses inconvénients. Le lecteur pour qui la technologie et les procédés de fabrication ne présentent qu'un intérêt mineur peut négliger ces explications et passer de suite au paragraphe suivant : « RÉSULTATS EXPÉRIMENTAUX. »

sur toute la plaque et dans ce cas le théorème des écrans complémentaires s'applique parfaitement : l'image obtenue par diffraction est la même que le film soit inversé ou non.

4. CORRÉLATION EN TEMPS RÉEL FONCTIONNANT AVEC UN MILIEU NON-LINÉAIRE OUVERT PAR UN ÉCLAIREMENT TRÈS FORT.

Un laser de grande puissance fournit un pinceau lumineux qui, après passage dans plusieurs systèmes optiques, est divisé en trois faisceaux de rayons parallèles : deux intenses, inclinés l'un par rapport à l'autre, et un faible, de direction parallèle à celle de l'un des deux trajets précédents. Les deux faisceaux de même direction traversent dans un même plan les deux informations à corrélérer, c'est-à-dire deux copies identiques de la fonction-objet dont une seule lentille construit les deux spectres, parfaitement superposés. Dans le plan focal de la lentille passe également l'autre faisceau de forte puissance, qui interfère avec chaque spectre.

Les franges d'interférence formées par les deux ondes intenses sont suffisamment lumineuses pour exciter un milieu non-linéaire qui devient transparent et dans lequel se forme alors un hologramme de Fourier. On réalise ainsi, simultanément et dans le même plan, le produit d'un spectre (celui d'énergie la plus faible) et d'un hologramme (l'autre spectre, modulé par l'onde intense qui joue le rôle de référence). Une lentille transforme ce produit en la corrélation des fonctions dont les transformées de Fourier constituent le spectre et l'hologramme.

Un tel système fonctionne expérimentalement au LOBE avec un laser à rubis fournissant cent mégawatts par centimètre carré durant des impulsions de dix nanosecondes. Il présente l'avantage de donner en temps réel l'auto-corrélation d'une fonction dès que celle-ci est placée sous forme de deux transparences identiques dans les trajets des deux faisceaux parallèles. Parmi ses inconvénients, il faut mentionner la nécessité d'enregistrer les résultats (ceux-ci étant obtenus pendant des durées très brèves, on ne peut pas les regarder à loisir tant qu'ils ne sont pas mis en mémoire, sur une photographie par exemple), la complexité technique de l'appareillage, et surtout le prix très élevé du laser nécessaire pour faire fonctionner le système.

5. UTILISATION D'UN INTERFÉROMÈTRE QUI FOURNIT LA TRANSFORMÉE DE FOURIER D'UNE RÉPARTITION D'ÉCLAIREMENT INCOHÉRENTE.

La pupille étudiée, éclairée en onde plane monochromatique est placée devant une lentille. Dans le plan focal de celle-ci tourne un diffuseur qui transforme la répartition d'amplitude cohérente — c'est-à-dire le spectre de la fonction — en une répartition d'éclairement incohérente spatialement. Un interféromètre fournit la transformée de Fourier, modulée par une distribution de franges, de cette répartition d'énergie, autrement dit l'auto-corrélation de la fonction initiale. On peut resserrer à loisir l'espacement des franges, et notamment le rendre peu inférieur à la limite de résolution du récepteur.

Cette méthode possède avant tout l'avantage de fonctionner en temps réel. Elle nécessite l'emploi d'une source monochromatique, mais un laser n'est pas indispensable ; une lampe à arc ou une source de lumière blanche de haute luminance accompagnée d'un filtre à bande passante étroite peuvent convenir. Un prototype d'interféromètre construit au LOBE permet de réaliser la transformée de Fourier d'une répartition d'éclairage incohérente. Cependant, une importante difficulté demeure : pour obtenir la moyenne d'un nombre même restreint de lettres et observer les résultats à un grandissement suffisant, un éclairage très puissant est nécessaire à l'entrée du dispositif.

6. CORRÉLATION EN ÉCLAIRAGE INCOHÉRENT.

Pour calculer la corrélation entre deux fonctions bi-dimensionnelles représentées par des pupilles, il suffit d'éclairer l'une d'elles par un faisceau de rayons lumineux parallèles, puis de déplacer la seconde pupille dans un plan parallèle à la première. En collectant les rayons qui ont traversé les deux pupilles, on obtient une répartition d'énergie analogue à la corrélation entre les deux fonctions. Le déplacement mécanique effectuant le balayage peut être simulé en utilisant des rayons obliques dont les trajets couvrent toutes les directions. De nombreuses variantes de la méthode exposée ci-dessus et reposant sur le même principe ont été développées depuis plusieurs décades. Ces techniques conduisent à la corrélation de deux fonctions en éclairage spatialement et temporellement incohérent.

Un système simple réalisé récemment fournit des résultats de bonne qualité en utilisant des pupilles de taille assez importante. Il comprend une source de lumière blanche d'intensité réglable, un système collimateur, un diviseur d'amplitude (lame semi-transparente), un miroir et une lentille de mise au point. Il réalise l'auto-corrélation d'une fonction à partir d'un seul négatif photographique traversé deux fois par les mêmes rayons lumineux. Bien qu'une importante quantité de lumière soit nécessaire à l'affichage des résultats, un tel dispositif, peu coûteux et d'emploi très simple, apporte un gain de temps considérable dans le calcul d'auto-corrélation puisqu'il fonctionne en temps réel. Son utilisation pour la réalisation de formes-moyennes reste liée à des difficultés dues aux dimensions des lettres et des « points » imposées par le système optique.

Nous avons développé un autre système adapté à la superposition de formes sans faire intervenir un mécanisme d'auto-corrélation. En éclairage incohérent, il semble, en effet, plus simple et plus logique de placer, en suivant le sens de la lumière, la carte de lettres puis celle (semblable) de « points ». Ce montage fonctionne à l'inverse de celui utilisé pour multiplier une image dans les appareils photographiques à sténopés multiples. Une interprétation holographique du phénomène mis en jeu a guidé l'étude expérimentale du dispositif qui s'avère d'un emploi simple et d'un coût très faible.

A la fois pédagogique de par sa conception et utilisable par toute

personne ayant quelque connaissance élémentaire en photographie, ce système corrélateur fonctionne de façon rudimentaire, en éclairage incohérent, c'est-à-dire au moyen d'une source lumineuse aussi peu coûteuse que possible. Même s'il fournit des images d'une résolution assez faible, il devrait attirer nombre d'utilisateurs.

Dans l'état actuel des recherches, il est difficile de dire quelle technique sera la plus propice pour construire opérationnellement des lettres-moyennes. Le choix de l'appareillage dépend essentiellement de ce que requiert l'utilisateur et de ses impératifs : coût, rapidité d'exécution, fabrication des lettres moyennes par un spécialiste ou par le paléographe, qualité des images à la sortie du système. Outre les avantages et les inconvénients inhérents à chacune des méthodes décrites précédemment, il faut tenir compte de la chaîne d'entrée des données, autrement dit savoir quel système de prétraitement est le plus approprié pour isoler d'un document les lettres à étudier, construire la carte de leurs « centres », associer celle-ci à la distribution des lettres, introduire le tout à l'entrée d'un dispositif auto-corrélateur.

RÉSULTATS EXPÉRIMENTAUX

Compte tenu du fait que tous les procédés expérimentaux exposés précédemment sont tous très longs à mettre en œuvre en laboratoire, un choix a dû être fait avant de réaliser quelques lettres-moyennes. Nous avons utilisé les méthodes numérotées (2) et (3) (diffraction à l'infini du module carré du spectre d'une fonction et enregistrement puis diffraction à l'infini du « spectre » d'une fonction modulée par un diffuseur) pour fabriquer des formes-moyennes sur trois écritures du manuscrit H. 316 (Bibliothèque nationale, Paris), préalablement connues par des études paléographiques classiques et par des mesures analogiques optiques. Ces traitements ont conduit à la fois à une nouvelle connaissance sur l'écriture et à une comparaison, par le biais des résultats obtenus, entre les deux techniques employées.

L'évolution de la lettre-moyenne obtenue par superposition de deux, puis trois, puis quatre, puis cinq lettres d'une même écriture est représentée sur la planche VIII.

La planche IX, où sont placés côte à côte six fragments de texte où ne restent que les alephs écrits par un même scribe et les six lettres-moyennes construites à partir de ces données, illustre l'aspect obtenu en superposant des formes très fluctuantes issues d'une écriture de faible stabilité.

Sur la planche X, des échantillons de lettres-moyennes faites sur des folios différents couvrant les écritures de quatre scribes montrent l'évolution de l'écriture d'une même main le long d'un manuscrit et permettent également de comparer, par leur moyenne, les lettres alephs de diverses écritures. L'encart en bas à droite de la planche X contient

la même lettre moyenne fabriquée par deux méthodes différentes (numérotées 2 et 3 dans cet article); tout en traduisant l'imperfection actuelle des appareillages, ceci montre l'influence de la technique mise en œuvre et indique combien son choix doit être déterminé en fonction de l'exploitation des résultats.

PERSPECTIVES

Une comparaison entre les multiples procédés conduisant de façon analogique à des formes-moyennes paraît nécessaire. En dehors de ces aspects techniques, un développement tourné vers la paléographie et dont nous donnons ci-après l'idée maîtresse paraît d'un intérêt certain.

Au vu des formes obtenues en superposant plusieurs lettres, il est possible de décrire des écritures par des qualités telles que leur stabilité, leur régularité, leur évolution au cours de la fabrication d'un ouvrage ou en passant d'un monument à un autre.

Mais des mesures et une connaissance physique des profils des lettres-moyennes peuvent mener beaucoup plus loin. Nous envisageons de représenter les courbes décrivant ces profils; ceci devrait permettre non seulement d'aboutir à chiffrer la dispersion d'une écriture — autrement dit sa régularité — mais aussi de comparer celle-ci à des modèles théoriques connus, déterminés par le calcul à partir de lois statistiques telles que des distributions de Gauss, de Poisson, etc...

De telles expériences devraient apporter beaucoup à la connaissance de l'écriture et permettraient peut-être d'avancer dans le domaine de l'expertise; peut-on caractériser une main, dans un style d'écriture donné, par sa dispersion?

IV. — ÉTUDE GÉNÉRALE ET EXPERTISE D'UNE ÉCRITURE

Les travaux dont il est fait état précédemment n'ont de sens qu'appliqués à des écritures où les lettres possèdent une morphologie assez rigide; ils ne peuvent être que d'un apport très réduit pour les écritures irrégulières, cursives notamment. Aussi convenait-il de développer une méthode générale d'analyse susceptible de guider le paléographe vers la réponse à l'une des questions fondamentales qui se posent souvent à lui, à savoir: un ou plusieurs documents sont-ils ou non écrits par la même personne? Peut-on mettre en évidence quelque chose qui caractérise la main d'un scribe?

Pour résoudre ce problème, il s'avèrait nécessaire de poser une hypothèse décrivant l'acte « écrire », de vérifier son bien-fondé, puis de l'utiliser dans un système de traitement. D'où la démarche suivie dans les expériences rapportées ci-dessous.

1. MODÈLE DESCRIPTIF DE L'ACTE « ÉCRIRE ».

Notons d'abord que des langues différentes usant du même alphabet peuvent se distinguer, au niveau d'un texte, par la probabilité de présence de chaque lettre. Ainsi, on a 176 chances sur 1000 de rencontrer la lettre e en français alors qu'on a seulement 131 chances sur 1000 de la trouver dans un texte anglais. Il en va de même, avec des probabilités différentes, pour tous les caractères de l'alphabet latin.

A cette fréquence de présence de chaque lettre dans un texte s'ajoute, pour les manuscrits, le changement des stéréotypes des lettres (pour une écriture donnée) quand on passe d'une langue à une autre. En effet, le dessin de chaque caractère se trouve influencé par la langue du scribe, ne serait-ce que parce que la probabilité de voisinage entre deux lettres (e et t par exemple) varie d'une langue à l'autre.

Parce qu'il caractérise une langue, nous appelons *facteur linguistique* l'ensemble des données précédemment décrites (distribution statistique σ_i et forme des stéréotypes l_i composant l'alphabet). Une écriture résulte de la composition de ce facteur linguistique avec un *facteur de forme*, ou fonction d'appareil du scripteur. Celle-ci représente, pour un texte imprimé ou frappé, ce qui est dû à la machine (la forme des caractères) et à tout ce qui l'entoure (l'encrage, le support d'écriture, etc...).

Pour un manuscrit, ce facteur de forme est déterminé par le scribe, son instrument et la façon dont il le tient, les formes d'écriture qu'il choisit et son aptitude à les transcrire.

La localisation de chaque lettre — ou signal s_i — dans un texte étant repérée par l'indice j et représentable par une impulsion $\delta(x_{ij}, y_{ij})$, tout morceau de texte peut être décrit par la corrélation de chaque lettre avec sa fonction de position :

$$\sum_i \sum_j s_i(x, y) * \delta(x_{ij}, y_{ij})$$

Chaque lettre s_i est elle-même vue comme la corrélation d'un facteur linguistique σ_i , $l_i(x, y)$ et d'un facteur de forme $f(x, y)$ rendant compte de l'altération subie par chaque symbole :

$$s_i(x, y) = \sigma_i l_i(x, y) * f(x, y)$$

Une page d'écriture $e(x, y)$ est donc représentée par l'expression :

$$e(x, y) = \sum_i \sum_j [\sigma_i l_i(x, y) * f(x, y)] * \delta(x_{ij}, y_{ij})$$

Si l'on sait éliminer la fonction de position, tous les textes que l'on se propose de mieux connaître sont alors décrits par une équation à deux inconnues :

$$e(x, y) = l(x, y) * f(x, y)$$

On ne peut alors étudier que l'un des facteurs en fixant l'autre :

— pour des langues différentes, en utilisant les mêmes caractères typographiques, on reconnaît une langue à travers son facteur linguis-

lique. On pourrait aussi envisager l'étude de l'évolution de la distribution statistique des caractères de l'alphabet d'une langue ;

— pour une langue déterminée, par comparaison des facteurs de forme on détecte l'écriture d'un individu.

2. DESCRIPTION DE LA CHAÎNE OPTIQUE DE TRAITEMENT.

Toutes les expériences rapportées ci-après se font au moyen de dispositifs sphérocyllindriques dont il importe d'abord d'expliquer le fonctionnement.

Un doublet sphérocyllindrique se compose de l'association d'une lentille cylindrique et d'une lentille sphérique. Ce système donne, de toute information introduite devant lui et éclairée par une onde monochromatique plane, deux transformées de Fourier suivant deux directions orthogonales situées le long de lignes focales séparées dans l'espace. En plaçant le cliché d'un texte de façon telle que les lignes d'écriture soient parallèles aux génératrices de la lentille cylindrique, on obtient un « spectre » qui, suivant la deuxième ligne focale, est la transformée de Fourier suivant la direction perpendiculaire à la direction d'écriture ; l'information est conservée sans modification de structure le long de la direction d'écriture. En plaçant un récepteur quadratique — par exemple une plaque photosensible — dans ce plan spectral, on enregistre le carré du « spectre » et on perd ainsi l'information correspondant à la position des lettres, du moins en ce qui concerne la direction suivant laquelle s'effectue la transformée de Fourier. Après inversion de l'enregistrement du spectre, on obtient une fonction proportionnelle au carré de la transformée de Fourier réalisée qui, après une nouvelle diffraction par le système sphérocyllindrique, fournit l'auto-corrélation du texte suivant la direction (x) d'analyse ; l'information est toujours conservée telle quelle le long de la direction (y) d'écriture. Cette auto-corrélation contient :

— un terme central : la superposition, pour chaque y donné, de toutes les auto-corrélations des lettres suivant x ;

— des termes latéraux : les superpositions, pour chaque y donné, des inter-corrélations entre les lettres du texte suivant x. Ces termes latéraux sont espacés par des distances égales (au grandissement du dispositif près) aux intervalles entre les lignes d'écriture.

En masquant ces termes d'inter-corrélation, on peut ne laisser passer que les auto-corrélations (suivant x) entre les lettres et obtenir après diffraction dans un autre système sphérocyllindrique (couplé au précédent) une information spectrale qui représente :

— l'information initiale sans changement, suivant l'axe (y) d'écriture ;

— le carré de la transformée de Fourier du texte suivant l'axe (x) perpendiculaire à l'écriture, débarrassée des termes dus aux interactions entre les diverses lignes d'écriture.

En résumé, on met à profit le fait que l'écriture soit exécutée le long

caractères
différents (I.B.M.)

même langue



"dual gothic"
(1er texte)

"dual gothic"
(2ème texte)

"scribe"
(1er texte)

langues
différentes

mêmes caractères
("dual gothic")

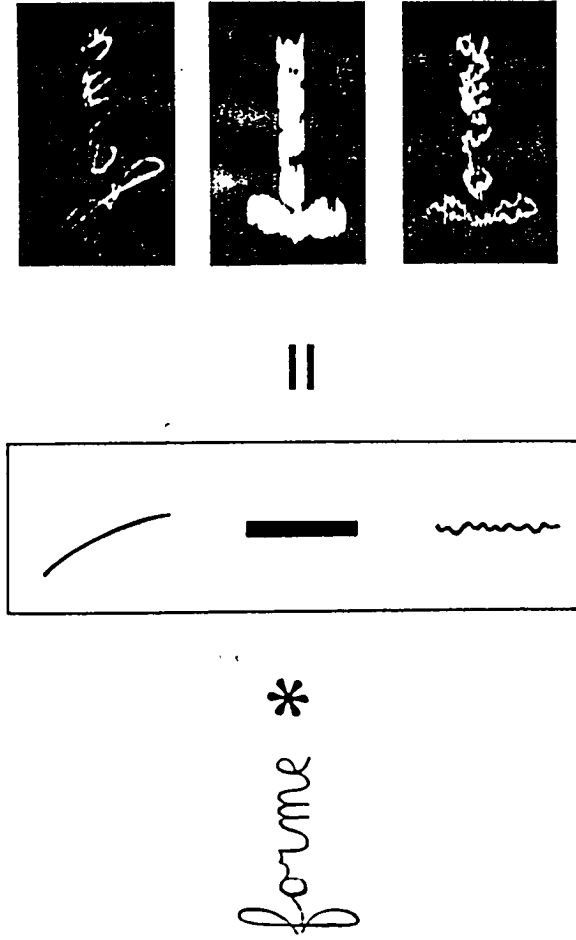


français
(1er texte)

français
(2ème texte)

anglais
(3ème texte)

Spectres unidimensionnels mettant en évidence les facteurs de formes et les facteurs linguistiques de diverses pages d'écriture typographique.



Corrélation unidimensionnelle d'un dessin du mot «forme» avec 3 facteurs de formes distincts.

(a)

direction d'analyse $\rightarrow x$

$Y \leftarrow$

texte

notation musicale

texte + notation musicale

DOUAI
Texte
+ notation musicale

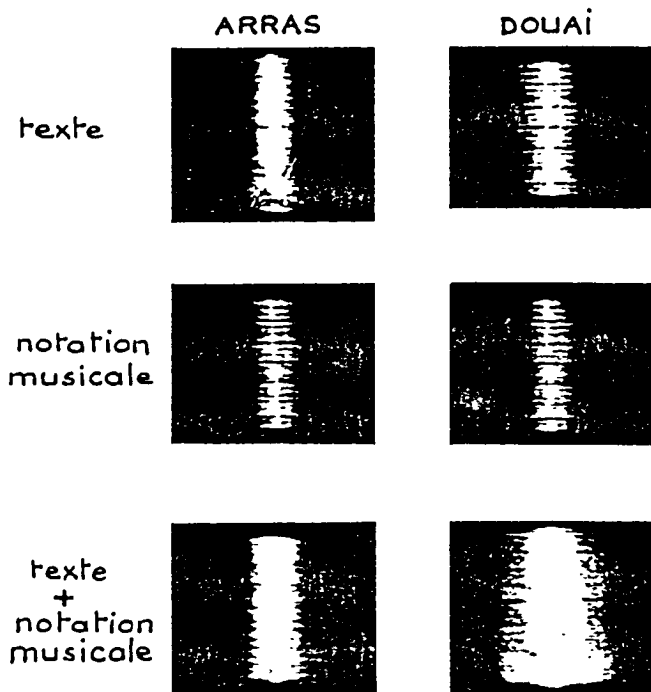
ia super omnes caelos no mirabilis est nomen domine hiesu qui beatum uirum dicitur gloria & honore coronasti S Dne da nra Confessio in domino xpi sacerdos uelutis multitudinis miraculoru inter semet & larus emicuit S In dno confido

(b)

ARRAS
Texte
+ notation musicale

U^o de... A^o...

(a) Vue partielle des données à traiter. (b) Fragments des partitions musicales provenant d'Arras et de Douai.



Résultats expérimentaux : mise en évidence des facteurs de forme des textes et des notations musicales traités. Noter en particulier les similarités des bandes de fréquences communes existant au niveau de l'ensemble des neumes.

de lignes parallèles et espacées pour éliminer au niveau du spectre les interférences dues aux diverses lignes d'écriture. Ceci nécessite l'emploi d'un système sphéro-cylindrique à double diffraction et ne permet d'effectuer une analyse du document que suivant une direction. D'autre part, un enregistrement photographique permet de supprimer la fonction de position des lettres dans le texte.

En bout de chaîne, après une diffraction, un enregistrement, une inversion, puis deux diffractions successives, on obtient le carré du produit des spectres des deux informations qui étaient corrélées dans l'écriture étudiée, c'est-à-dire le facteur linguistique spectral $\sum_i \alpha_i l_i$, multiplié par le facteur de forme spectral F .

$$\text{résultat} = \left(\sum_i \alpha_i l_i \right)^2 \cdot F^2$$

La somme des probabilités de présence de toutes les lettres étant égale à 1, si on utilise un document contenant suffisamment de lettres, on obtient finalement le carré du produit $L \cdot F$ du facteur linguistique global et du facteur de forme, ceci en termes spectraux suivant la direction d'analyse (perpendiculaire à celle d'écriture).

Ce type d'analyse comporte l'inconvénient de ne traiter l'information que suivant la direction verticale d'un texte. On pourrait craindre de ne tirer que peu d'éléments intéressants par une étude effectuée le long de cette direction, hormis le cas fréquent où l'écriture est « penchée ». De fait, les expériences réalisées même sur des textes produits à la machine et d'écriture « droite » confirment que l'analyse effectuée suffit à fournir les résultats escomptés.

3. VÉRIFICATION DE LA VALIDITÉ DU MODÈLE.

En disposant de données connues, nous avons vérifié *a posteriori* la qualité du fondement du modèle choisi pour décrire l'écriture. Pour ceci, nous avons utilisé des textes tapés à la machine, en anglais et en français, avec des caractères de types différents (lettres I. B. M.).

Chaque document utilisé pour l'expérience couvrait deux pages de format 21 × 29,7 tapé en simple interligne.

Les résultats consignés sur la planche XI montrent :

— pour une même langue : des facteurs de forme spectraux identiques correspondant à deux textes différents et des facteurs de formes spectraux différents pour deux textes, fussent-ils identiques dans leur contenu, frappés avec deux boules différentes (caractères I. B. M. « dual gothic » et « scribe ») ;

— pour une même police de caractères (ce qui fixe le facteur de forme), des facteurs linguistiques spectraux identiques pour deux textes différents d'une même langue (française) et des facteurs linguistiques spectraux différents pour deux textes écrits dans des langues différentes (français et anglais).

Au vu de ces résultats, on comprend immédiatement comment une écriture ou une langue est caractérisée, au niveau spectral, par des bandes de fréquences propres aux formes des lettres ou à la fréquence de présence des lettres.

4. AUTRE VÉRIFICATION DE LA VALIDITÉ DE L'HYPOTHÈSE.

A partir d'un document écrit, il n'est pas possible, dans l'état actuel, de séparer les deux termes (facteur linguistique et facteur de forme) qui, corrélés l'un à l'autre, constituent l'écriture telle que nous la voyons ; nous devons nous limiter à des comparaisons entre les déterminants de même nature de deux ou plusieurs écritures. Au mieux, on pourrait élaborer des modèles de facteur linguistique grâce à des données statistiques et en utilisant les archétypes de certaines écritures couramment utilisées et connues (l'onciale par exemple). A partir de là, il deviendrait possible d'isoler le facteur de forme d'une écriture manuscrite. En pratique, on ne peut atteindre un facteur de forme que sous l'aspect des fréquences qui se trouvent dans son spectre et nous ne savons pas, au stade actuel, passer du domaine spectral où l'on possède F^2 au domaine géométrique où l'on voudrait f , autrement dit, nous sommes incapables de calculer le dessin du facteur de forme dans l'espace habituel.

Cependant, la démarche inverse est possible, comme l'indique la planche XII, qui résume l'expérience suivante. Le mot « forme » écrit de façon quelconque sert de stéréotype arbitraire. Au moyen d'hologrammes adaptés, on réalise, dans un corrélateur optique sphérocyindrique, la corrélation entre ce stéréotype et un facteur de forme lui aussi dessiné arbitrairement dans l'espace géométrique. On modifie ainsi l'allure du mot « forme » qui devient courbé, tremblé, ou pâteux, suivant le facteur de forme utilisé pour construire l'hologramme. Les variations de morphologie d'une écriture peuvent, de cette manière, s'étendre à loisir.

Il est également possible de changer continûment le dessin d'une lettre, d'un mot ou d'un texte en déplaçant sur son spectre un hologramme convenablement adapté.

5. EXPERTISE DE DOCUMENTS ANCIENS.

Les considérations développées précédemment à propos des écritures s'étendent à l'étude d'autres domaines et peuvent apporter des connaissances sur la morphologie de toutes sortes de signes, voire sur le dessin en général.

Un exemple va montrer comment on peut aussi tirer des renseignements sur le graphisme de la notation musicale et peut-être relier celle-ci à l'écriture d'un texte.

Deux documents datés du XII^e siècle et conservés dans des bibliothèques, l'un à Arras, l'autre à Douai, portent le même chant liturgique. A première vue, l'écriture du texte est de facture différente dans les deux

documents, mais il est difficile de constater quelque chose sur la notation musicale. A cette époque, on n'écrivait pas encore les neumes sur des portées et la transcription du chant n'en était qu'à ses balbutiements. Des vues partielles des deux documents sont présentées sur la planche XIII, où se trouve également mentionnée la direction d'analyse suivie lors du traitement optique. Nous avons séparé la notation musicale de celle du texte en effaçant sur des photographies soit le texte, soit la musique, suivant ce que l'on voulait conserver. Ensuite, ces données ont subi un traitement adéquat dans le but de mettre en évidence les facteurs de forme. Comme les deux documents sont identiques tant dans leur texte (à certaines abréviations près) que dans la notation de leur chant, il n'est pas question ici de rencontrer des difficultés relatives au facteur « linguistique » ; toute différence entre les deux documents ne saurait être attribuée qu'aux formes des lettres et des neumes qu'ils contiennent.

Les résultats de l'expérience, consignés sur la planche XIV, montrent des similarités et des dissemblances, attribuables aux seuls facteurs de forme, entre les « spectres » sphérocyindriques des informations traitées, épurés des interactions ligne à ligne. Simplement au vu de ces réponses, on pressent qu'il existe une très forte probabilité pour que l'écriture des deux textes soit de mains différentes alors que la notation musicale a très vraisemblablement été transcrite par le même scribe dans les deux cas. Des mesures faites par corrélation optique entre les réponses ont confirmé la validité de cette impression visuelle.

Il semble, d'autre part, que le texte du manuscrit de Douai soit de la même facture que les neumes des deux documents, comme en témoignent les bandes de fréquences constitutives des spectres des facteurs de forme. Cependant, il ne faudrait pas considérer cette dernière interprétation comme une affirmation absolue, compte tenu du fait qu'il est dangereux de comparer des éléments qui ne sont pas de même nature.

6. EXPERTISE DE DOCUMENTS DE L'AFFAIRE DREYFUS.

Un travail sur les dossiers judiciaires de l'affaire Dreyfus a été entrepris à la fois pour tester cette méthode d'analyse optique des écritures et pour la faire connaître. Dans l'état actuel de cette recherche n'ont été traités que quelques-uns des écrits susceptibles d'être attribués soit à Dreyfus soit à Esterhazy, à savoir des lettres d'Alfred Dreyfus adressées, pendant sa détention, au ministre de la Guerre (14 février 1895) et à Monsieur le Président de la République (5 octobre 1895), des missives d'Esterhazy, l'une destinée au ministre de la Guerre (20 novembre 1897), une autre au même (14 septembre 1898), une à un baron (29 juin 1894), et enfin le « bordereau », sur lequel nous reviendrons. D'autres pièces du dossier, ayant eu à l'époque une célébrité certaine, tels la « lettre du buvard », le « petit bleu », le « faux Henry », seront analysées ultérieurement.

Sur la planche XV, les spectres contenant les facteurs de forme imputables à Dreyfus, à Esterhazy et à l'auteur du bordereau — pièce maîtresse de l'accusation et point de départ de « l'affaire » — côtoient des

fragments des écrits d'où ils sont extraits. Il s'avère commode de comparer simultanément toutes les lignes horizontales qui composent les facteurs de forme spectraux, d'où les courbes de la planche XVI, directement exploitables, et qui représentent l'analyse globale des spectres représentés planche XV. Ces graphes sont obtenus au moyen d'un balayage de télévision, d'un filtrage du signal vidéo, suivi de son affichage sur un écran d'oscilloscope.

Malgré le nombre assez faible de cas où un expert dispose de données en quantité suffisantes pour recevoir l'appui d'une analyse statistique, ces expériences traduisent sur un fait précis ce qu'un traitement optique peut apporter comme soutien à l'expertise en écriture. En effet, au vu des résultats présentés planche XV et planche XVI, on constate une forte proximité entre les facteurs de forme du scripteur du bordereau et d'une missive d'Esterhazy, ce qui ne saurait prouver indubitablement que ce dernier ait écrit le bordereau. Mais l'écart entre les spectres épurés du bordereau d'une part, d'écrits de Dreyfus d'autre part, est suffisamment important pour prouver que celui-là n'a pas été écrit par la main d'Alfred Dreyfus. Ce constat est en accord avec les rapports des experts en écriture actuels qui ont travaillé sur ces documents.

Notons que cette méthode de mise en évidence de facteurs de forme repose sur une seule hypothèse, dont nous nous sommes efforcés de montrer la validité. Elle ne peut s'appliquer qu'à des documents conséquents par leur taille puisqu'elle nécessite l'existence au sein de l'information traitée d'un grand nombre de lettres ou de signes. Si cette limite apparaît comme un inconvénient certain, il ne faut pas oublier l'importance de la capacité de traitement d'un système optique puisque celui-ci ne traite pas un texte de façon séquentielle lettre par lettre, mais au contraire travaille simultanément sur toute la page d'écriture que l'on place à son entrée. Les calculs optiques ainsi effectués ne requièrent aucune limite quant à la « qualité » de l'écriture. Fusse-t-elle illisible, celle-ci n'en demeure pas moins analysable quant à l'extraction des données spectrales de son facteur de forme.

Sans insister sur l'aspect linguistique mentionné précédemment, rappelons que par cette technique qui, certes, demeure longue, difficile à mettre en œuvre, assez coûteuse, nous avons cependant réussi à retrouver la trace d'un même scribe dans deux monuments anciens conservés en des lieux différents.

En délaissant pour le moment la paléographie musicale, nous envisageons de tester plus à fond la méthode pour mieux connaître ses possibilités et ses limites. Ceci sera fait en étudiant optiquement l'évolution au cours du temps (sur une période de plusieurs décades) de l'écriture d'un scribe préalablement très bien connue par les paléographes, et aussi en traitant les nombreux documents sélectionnés dans les dossiers de l'affaire Dreyfus.

Alent; sans barreaux; des
 sans formes d'écriture, sans de l'écrit
 de plus en plus que un d'elles par
 Comptes, je n'ai pu déceler, et je
 de leur nature que je n'ai pu en dire
 être, pour un compte, que je n'ai
 pour un compte, que je n'ai

Le 22 octobre 1892 (11 novembre 1892)
 Le d'écrit de l'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit

Pour d'écrit, sans barreaux, et d'écrit
 d'écrit, et d'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit
 d'écrit, et d'écrit, et d'écrit



DREYFUS



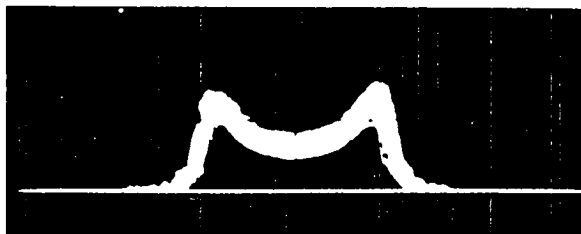
le bordereau



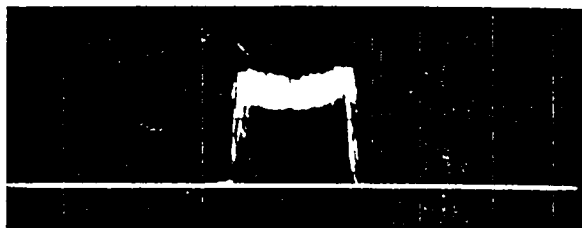
ESTERHAZY

Fragments de 3 écritures et facteurs de formes spectraux de celles-ci.

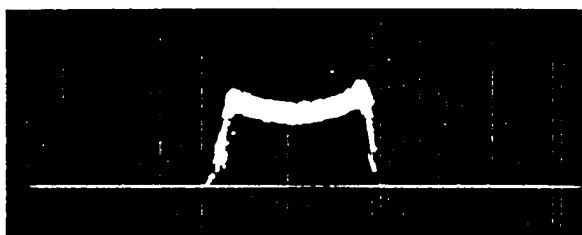
DREYFUS



Le bordereau



ESTERHAZY



Composantes spectrales caractéristiques de 3 écritures.

CONCLUSION

Plusieurs moyens de traitement analogique de la morphologie de lettres et d'écritures ont été étudiés et comparés entre eux à travers les réponses qu'ils fournissent au paléographe.

Lorsque le dessin des lettres est très régulier, un calcul de ressemblance de formes conduit à la mesure de l'évolution d'une écriture aussi bien qu'à celle de la distance qui sépare les écritures de deux scribes. Pour ce même type de formes, nous avons développé une méthode de construction de forme-moyenne qui, par diverses techniques qu'il sied de rendre opérationnelles, fournit une représentation visuelle de la stabilité d'une écriture à travers l'allure globale des graphismes d'un même signe de l'alphabet.

Une méthode générale d'analyse, applicable à tout type d'écriture, permet de détecter la présence d'une main dans un acte écrit, ou plutôt d'écartier la possibilité qu'un individu ait écrit un texte, pourvu que l'on dispose d'autres documents transcrits par le même scribe. Ces mesures comparatives s'effectuant par le traitement simultané d'un grand nombre de lettres portent les espoirs d'une méthode statistique d'expertise fine des écritures.

Ces résultats, associés à ceux obtenus par d'autres méthodes d'exploration utilisées dans les Sciences de l'Histoire, fournissent des éléments qui peuvent conduire les érudits vers un savoir plus précis concernant les monuments paléographiques.

Jean-Marc FOURNIER.

Besançon, mars 1975.

Laboratoire de Physique générale et Optique
de l'Université de Besançon.

Institut de Recherche et d'Histoire des Textes,
Paris.

BIBLIOGRAPHIE

Il nous est apparu très difficile de tenter le regroupement des éléments nécessaires à la réalisation d'une bibliographie exhaustive. Et surtout particulièrement délicat. D'une part, les méthodes analogiques, notamment optiques, n'en sont qu'à leurs débuts quant à leurs applications, et il ne paraît pas que d'éventuels résultats aient été publiés actuellement par les quelques groupes qui, à travers le monde, seraient susceptibles de traiter ces problèmes. Par contre, les informaticiens portent beaucoup d'intérêt à ces questions : leur domaine est cependant assez éloigné du nôtre et nous commençons seulement d'utiliser leurs résultats.

Enfin, et ceci est plus important, nous avons éludé la question d'une bibliographie des auteurs des Sciences humaines : en pratique, il eut été impossible, pour des physiciens, de faire un choix parmi les très nombreux textes et ouvrages que nous avons pu lire ou consulter et qui, nous en sommes conscients, ne représentent encore qu'une faible partie de ce que nous devrions sans doute connaître dès l'abord d'une étude des écritures.

Toutefois, nous donnons, à titre indicatif, une liste des conférences et pu-

blications ayant trait directement aux travaux dont il est question dans notre article.

- J.-Ch. VIÉNOT, *Optical Information Processing applied to the study of hebraic manuscripts (morphological resemblance in ancient handwritten texts)*, in *Proc. 5th World Congress of Jewish Studies (1969)*, ed. by Dr PINCHES PELI, World Union of Jewish Studies, Jerusalem, 1972; *Traitement optique de l'information appliqué à l'étude des manuscrits hébraïques (ressemblances morphologiques dans les textes manuscrits anciens)*, in *Revue d'Histoire des Textes*, t. 1, 1971, pp. 239-244.
- J.-M. FOURNIER, *Comparaison spectrale et mesure du degré de ressemblance de formes voisines*. Diplôme d'Études approfondies, Besançon, 1969.
- J.-Ch. VIÉNOT, J. BULABOIS, J. DUVERNOY, J.-M. FOURNIER, L.-R. GUY et G. TRIBILLON, *Progrès récents dans le traitement optique de l'information*, Société française de Physique, Section de Bourgogne/Franche-Comté, Besançon, novembre 1969.
- J.-M. FOURNIER et J.-Ch. VIÉNOT, *Essai de classification au moyen de techniques analogiques optiques*, Colloque sur l'Analyse des Données, Faculté des Lettres, Besançon, 1970.
- J.-M. FOURNIER et J.-Ch. VIÉNOT, *Fourier Transform holograms used as matched filters in hebraic Paleography*, Symposium on Engineering Applications of Lasers, Tel-Aviv, juin 1970, in *Israel Journ. of Tech.*, 9, n° 3, 1971, pp. 281-287.
- J.-M. FOURNIER, *Traitement optique de l'information appliqué à la mesure de la ressemblance et à la classification de formes voisines*, Thèse de Doctorat de 3^e cycle, Besançon, novembre 1970.
- J.-Ch. VIÉNOT et C. FROEHLY, *Traitement ultra-rapide d'informations optiques*, Contrat D. R. M. E., n° 72/34/112, rapport annuel, juillet 1973.
- A. LACOURT, C. FROEHLY et J. NEIRAS, *Détermination de fonctions de transfert optique par transformation de Fourier en lumière spatialement incohérente*, in *Optica Acta*, 19, n° 6, pp. 485-497.
- J.-M. FOURNIER et J.-Ch. VIÉNOT, *Mesures sur des tracés de lettres au moyen de techniques holographiques*, in *Les techniques de laboratoire dans l'étude des manuscrits et La paléographie hébraïque médiévale* (Colloques internationaux du Centre national de la recherche scientifique, n° 548, Paris, 13-15 septembre 1972), Paris, 1974, pp. 41-71.
- J. DUVERNOY, *Extraction des invariants morphologiques d'une information par corrélation optique et filtrage spatial*, in *Optics. Comm.*, 7, n° 2, pp. 142-145, 1973.
- J. DUVERNOY, *Estimation du nombre d'échantillons et du domaine spectral nécessaires pour déterminer une forme moyenne par une technique de superposition*, in *Optics. Comm.*, 7, n° 3, pp. 286-288, 1973.
- J. DUVERNOY, *Traitement d'informations en optique sphéro-cylindrique. Application à l'extraction de données géophysiques et morphologiques*, Thèse d'État, Besançon, mai 1973.
- F. MEMBREY et J. DUVERNOY, *Reconnaissance d'une forme et détermination de ses dimensions au moyen d'un filtre-hologramme à adaptations multiples*, in *Nouvelle revue d'Optique appliquée*, 4, n° 2, pp. 83-86, 1973.
- J.-Ch. VIÉNOT, J. DUVERNOY, G. TRIBILLON et J.-L. TRIBILLON, *Three methods of information assessment for optical data processing*, in *Applied Optics*, 12, n° 5, pp. 950-960, mai 1973.

- J.-M. FOURNIER, *Optical process in hebraic Paleography*, 6th World Congress of Jewish Studies, Jerusalem, août 1973.
- J. DUVERNOY, *Modèle de synthèse non-linéaire pour le traitement optique des écritures*, in *Optics Comm.*, 11, n° 4, pp. 373-378, 1974.
- J. DUVERNOY, *Codage de l'information dans un processus aléatoire. Influence des paramètres statistiques sur la densité spectrale d'énergie*, in *Optics Comm.*, 13, n° 1, pp. 48-53, 1974.
- B. CRETIN, *Corrélations en éclairage incohérent; application à la superposition d'images et à la réalisation de formes moyennes*, Diplôme d'Études approfondies, Besançon, 1974.
- J.-M. FOURNIER, *Approche analogique d'une expertise en écriture: un exemple concernant l'affaire Dreyfus*, soumis au « Courrier du C. N. R. S. », 1975.

FILM :

- J.-M. FOURNIER et J.-Ch. VIÉNOT, *Étude de documents, holographie optique*, séquence extraite du magazine *Synthèse* n° 2/1973, éd. par le Ministère des Affaires Étrangères (Direction Générale des Relations Culturelles, Scientifiques et Techniques).