

# Identification de Marqueurs de la Maladie d'Alzheimer à travers l'Écriture manuscrite acquise sur Tablette graphique

Christian Kahindo<sup>1\*</sup>, Mounîm A. El-Yacoubi<sup>1</sup>, Sonia Garcia-Salicetti<sup>1</sup>, Victoria Cristancho-Lacroix<sup>2,3</sup>,  
Hélène Kerhervé<sup>2,3</sup>, Anne-Sophie Rigaud<sup>2,3</sup>

<sup>1</sup>SAMOVAR, Telecom SudParis, CNRS, University Paris Saclay, France

<sup>2</sup>AP-HP, Groupe Hospitalier Cochin Paris Centre, Hôpital Broca, Pôle Gériatrie, Paris, France

<sup>3</sup>Université Paris Descartes, EA 4468, Paris, France

## RESUME

Ce travail porte sur l'analyse de l'écriture manuscrite acquise sur tablette numérique graphique à des fins de santé. Ces tablettes permettent l'acquisition du tracé au cours du temps, tracé dit « en-ligne », riche en information cinématique du scripteur. Notre but est la caractérisation de la maladie d'Alzheimer (MA) et des Troubles Cognitifs Légers via l'exploitation des paramètres spatiotemporels de l'écriture manuscrite. L'étude a un volet transversal, objet de cet article, et un longitudinal. L'acquisition des données manuscrites, réalisée à l'Hôpital Broca, est présentée. A ce jour, 182 personnes ont participé à l'étude dont 29 Alzheimer et 86 Troubles Cognitifs Légers. Une étude sur les paramètres extraits du tracé manuscrit est ici effectuée. Elle aboutit à une classification de ces derniers basée sur une mesure de qualité. Nos résultats préliminaires sont confrontés à la littérature et donnent notamment lieu à de nouvelles caractéristiques prometteuses pour caractériser les profils cognitifs.

**Mots-clés:** Ecriture manuscrite, tablette graphique, Alzheimer, Clustering, entropie.

## 1 INTRODUCTION

L'analyse de l'écriture manuscrite a été pendant longtemps concentrée sur la problématique de la reconnaissance automatique du signal manuscrit, notamment pour des applications bancaires et postales. Plus récemment, l'analyse de l'écriture manuscrite a évolué vers le champ de la santé. Le contexte actuel qui voit un développement spectaculaire des tablettes numériques graphiques accentue les possibilités d'innover dans ce domaine. Ces tablettes permettent l'acquisition du tracé au cours du temps, tracé dit « en-ligne », riche en information cinématique du scripteur.

Ecrire est une faculté de haut niveau qui requiert un contrôle moteur fin et une coordination neuromusculaire spécifique. Une fois qu'un sujet a appris à écrire, l'échange (feedback) entre perception visuelle et contrôle musculaire requis pour écrire, devient automatique, d'où des mouvements extrêmement rapides. En effet, un programme de contrôle moteur s'est à ce stade développé au niveau cérébral. Ainsi, la détérioration de l'écriture, quelle que soit son importance, est le signe d'une détérioration ou d'un dysfonctionnement de ce programme moteur fin.

Dans le champ de la santé, on distingue dans la littérature deux axes qui ont en commun l'analyse du manuscrit : le premier est celui des modèles de génération du manuscrit [16,17,18] ; le deuxième concerne l'analyse du manuscrit « en-ligne » pour le diagnostic de pathologies ou pour évaluer quantitativement les

effets d'une thérapie médicamenteuse. Dans le premier axe, on trouve les modèles moteurs de la production du manuscrit, qui font intervenir la cinématique du mouvement et l'activité neuromusculaire [16,17,18]. Les travaux de Schomaker [17] ont apporté un éclairage sur le processus de production de l'écriture, qu'il a modélisé par des oscillateurs, en décomposant les mouvements horizontaux et verticaux. D'autres travaux ont utilisé le paradigme oscillatoire, reposant sur une décomposition en séries de Fourier au niveau des portions du tracé situées entre deux levés de stylo [19]. Teulings [18] a ensuite étudié le contrôle moteur fin et son développement, en se basant sur des paramètres dynamiques acquis sur tablette graphique, dans le cadre de différents troubles : maladie de Parkinson et certains troubles psychiatriques. Ces travaux ont montré d'une part que le manuscrit est un marqueur comportemental qui permet de quantifier les effets de traitements médicamenteux, et que cela est possible grâce aux paramètres cinématiques qu'on peut extraire des signaux acquis sur tablette graphique. D'autre part, ses études sur le contrôle moteur fin révèlent la contribution du signal manuscrit dans la détection et le suivi médical de troubles aussi divers que la maladie de Parkinson, les dysgraphies, la schizophrénie, les dyspraxies (troubles de la coordination motrice), la coordination œil-main, le stress, parmi d'autres troubles [18]. Dans le deuxième axe, on trouve l'analyse quantitative des tremblements et plus généralement de troubles du contrôle moteur. Cette direction a été en particulier investie grâce à l'apparition des tablettes graphiques. Les premières études sur l'utilisation de l'écriture manuscrite en-ligne pour caractériser les troubles du contrôle moteur portent sur la maladie de Parkinson [10,11,12,13]. Ce n'est qu'à la fin des années 1990 que sont apparues des études portant sur l'analyse cinématique de l'écriture chez les personnes atteintes de la maladie d'Alzheimer (MA) et des Troubles Cognitifs Légers (TCL) ou « Mild Cognitive Impairment » (MCI) [3,4,5,6]. La caractérisation de la maladie d'Alzheimer (MA) et des Troubles Cognitifs Légers via l'exploitation des paramètres cinématiques de l'écriture manuscrite reste toutefois relativement peu abordée dans la littérature. Or détecter précocement la maladie est aujourd'hui un véritable enjeu sociétal au vu de son importance : elle est la principale cause de démence, et représente au moins les deux tiers des cas diagnostiqués en France. D'autre part, le nombre de patients atteints y est en croissance constante : 900000 cas environ aujourd'hui, avec des prévisions de 1,3 million en 2020 et 2,1 millions en 2040. Détecter précocement la maladie permet de mettre en place des prises en charge basées sur la stimulation cognitive et d'infléchir sa progression.

De multiples efforts ont été réalisés ces dernières années par la communauté scientifique pour établir des critères de diagnostic de la maladie à la fois précoces, valides et discriminants. Cependant, le début de la maladie étant insidieux, il reste difficilement repérable dès les premiers stades. Ainsi, on estime qu'elle débute plusieurs années avant qu'elle ne soit vraiment diagnostiquée. Le moment exact de transition entre la phase asymptomatique et la phase symptomatique pré-démontielle, puis celle de cette dernière

---

\*christian.kahindo@telecom-sudparis.eu

vers la démence débutante sont très difficiles à établir avec certitude. Le but de notre étude est de caractériser l'écriture de personnes touchées par ces troubles spécifiques en vue de détecter quels sont les paramètres spatio-temporels de l'écriture qui seraient prédictifs de la maladie. Pour cela, nous réalisons une étude statistique transversale et longitudinale sur l'analyse automatique de l'écriture auprès de patients ayant un Alzheimer débutant ou un Trouble Cognitif Léger, suivis au service Gériatrie de l'Hôpital Broca.

La suite de cet article sera organisée comme suit :

- ✓ La première section présentera l'état de l'art dans l'analyse de l'écriture et la maladie d'Alzheimer et des troubles cognitifs légers, et motivera nos choix pour l'acquisition des données et l'extraction des paramètres de l'écriture des patients; aussi, elle présentera les données collectées à ce jour ;
- ✓ La deuxième partie présentera les résultats préliminaires ;
- ✓ La troisième partie présentera nos conclusions et les perspectives du présent travail.

## **2 ETAT DE L'ART ET MOTIVATION DE NOS CHOIX POUR L'ACQUISITION DES DONNEES**

### **2.1 Analyse de l'écriture et maladie d'Alzheimer**

Plusieurs auteurs ont étudié si certains paramètres permettraient de discriminer les patients ayant la MA, les patients souffrant de TCL, et les personnes contrôle, à savoir des personnes à profil cognitif normal. Schröter et al. [5] ont fait une étude comparative auprès de 4 groupes (MA, TCL, patients avec dépression, sujets sains) où les participants devaient réaliser une tâche graphique sur tablette, apparentée à l'écriture manuscrite, en dessinant des cercles concentriques superposés pendant une durée de quelques secondes. L'analyse cinématique du tracé révèle des différences significatives entre les groupes par le biais de tests statistiques classiques (ANOVA, Kruskal-Wallis) [5]. En effet, les auteurs constatent que les variations de la vitesse (écart-type) sont plus importantes chez les patients atteints de MA que chez les TCL et chez les déprimés ; de même, ces variations sont plus importantes chez les patients TCL et déprimés que chez les sujets contrôle. Etant donnée la co-morbidité décrite dans la littérature entre TCL et dépression, les auteurs ont suggéré la réalisation d'études longitudinales auprès de patients TCL et avec dépression, afin de déterminer la valeur prédictive de l'analyse cinématique du tracé dans ces populations.

D'autres études sur l'écriture manuscrite montrent que, chez les personnes atteintes de MA, l'écriture est beaucoup moins fluide, ce qui se traduit par une hausse du « jerk », la dérivée de l'accélération en chaque point du tracé [4]. Aussi, les personnes atteintes de TCL et de MA passent davantage de temps avec le stylo en l'air que les personnes contrôle [3]. Enfin, les malades d'Alzheimer ont une longueur de « strokes » (portions du tracé entre deux levés de stylo) significativement variable par rapport à celle des personnes contrôle [6].

Ces travaux affirment que les indicateurs les plus influents dans la caractérisation du vieillissement pathologique (MA, TCL, dépression) s'avèrent être la vitesse et sa variation (écart-type), la vitesse relative (normalisée par l'écart-type), la fréquence d'écriture en nombre de « strokes » par seconde, les pics de vitesse des « strokes », la longueur des « strokes », la durée du mouvement, la fluidité de l'écriture (mesurée par les variations de l'accélération (« jerk »)), la pression du stylo sur la tablette.

Globalement, la caractérisation de la maladie d'Alzheimer (MA) et des Troubles Cognitifs Légers via des paramètres cinématiques de l'écriture manuscrite reste relativement peu abordée dans la littérature. Notamment, on constate l'absence à ce jour d'une étude longitudinale. C'est un des objectifs de notre étude. De plus, les études de la littérature ci-dessus sont fondées sur des tests statistiques déterminant les différences statistiques significatives entre les différents profils cognitifs des sujets (MA, TCL, personnes contrôle) à partir des paramètres cités précédemment. Cependant, ces tests statistiques se révèlent parfois peu concluants voire contradictoires en raison de la non-homogénéité d'un groupe de même profil cognitif ; la maladie d'Alzheimer, par exemple, a une évolution non homogène chez les patients. En outre, ces études ont été faites sur peu d'échantillons d'où la nécessité de mener l'étude avec un nombre des participants beaucoup plus grand. Notre démarche proposée dans la suite du travail vise à pallier ces limites.

### **2.2 Acquisition**

L'acquisition des données est en cours depuis un an au sein du Service Gériatrie de l'Hôpital Broca. Nous ciblons principalement des patients touchés par l'Alzheimer à un stade débutant ainsi que ceux ayant des Troubles Cognitifs Légers. Concernant les personnes contrôle, nous collectons leurs données manuscrites auprès de participants provenant d'associations de seniors, principalement OLD-UP et Génération 13.

#### **2.2.1 Critères d'inclusion et de non inclusion**

Afin de contrôler les données, nous avons établi certains critères d'inclusion et de non inclusion pour la sélection des participants. Ces critères sont les suivants :

##### **Critères d'inclusion**

- Etre âgé de plus de 60 ans ;
- Accepter de participer librement à l'étude (signer le formulaire de consentement) ;
- Etre un patient suivi en consultation gériatrique à l'hôpital de jour à l'hôpital Broca ;
- Critères d'inclusion par groupe, sur la base du diagnostic établi par l'équipe médicale :
  - Groupe 1 : diagnostic MA débutant, critères DSM-IV, MMSE (Mini Mental State Examination : échelle mesurant le niveau de dégradation cognitive) entre 20 et 25[14] ;
  - Groupe 2 : diagnostic TCL (Troubles Cognitifs Légers), critères de Petersen [15] ;
  - Groupe 3 : patients « contrôle » (ayant un vieillissement cognitif normal) ;

##### **Critères de non-inclusion:**

- Patients ne comprenant pas le français ;
- Patients ne sachant ni lire, ni écrire ;
- Patients avec des troubles visuels ou auditifs leur empêchant de réaliser les tâches ;
- Sujets privés de liberté par décision judiciaire ou administrative ;
- Sujets sous régime de protection curatelle ou tutelle.

#### **2.2.2 Protocole d'acquisition**

Nous avons défini un protocole d'acquisition et un ensemble d'instructions pour effectuer une acquisition auprès d'un participant. Ces instructions ont été établies de telle manière qu'il

n'y ait pas de variations dans l'acquisition dues au simple fait qu'on change d'évaluateur.

L'évaluation complète d'un patient dure une heure et trente minutes : trente minutes pour acquérir l'écriture et une heure pour effectuer un bilan neuropsychologique. Quant aux personnes « contrôle », on leur fait passer un bilan afin de connaître leur profil cognitif pour s'assurer, entre autres, qu'ils ont un profil cognitif normal.

A la fin de la campagne d'acquisition, tous les participants seront revus deux fois. La première visite a lieu au mois M0 et la seconde au mois M12 (12 mois plus tard) pour réacquérir leurs données avec le même protocole et refaire un bilan neuropsychologique. Ces données nous permettront de faire l'étude longitudinale.

L'acquisition de l'écriture manuscrite se fait sur une tablette graphique « Wacom » (Figure 1). La tablette graphique permet d'enregistrer au cours du temps les coordonnées du stylo  $(x(t), y(t))$ , la pression exercée sur la tablette, l'inclinaison du stylo ainsi que les coordonnées du stylo lorsqu'il est en l'air (jusqu'à 1.5 cm au dessus de la tablette). Cette tablette ne permet pas un retour visuel du tracé effectué mais permet de capturer les signaux à travers une feuille de papier posé sur elle ; avec un stylo spécial « Inking Pen » de Wacom (stylo à encre), la personne peut avoir un retour visuel du tracé qu'elle effectue sur le papier. Le tracé du stylo est échantillonné à intervalles de temps réguliers (entre 100 et 200 Hz habituellement). Ainsi la tablette permet de récupérer, pendant la tâche, plusieurs paramètres au cours du temps : les coordonnées du stylo, la pression exercée par le stylo sur la tablette (pression axiale) et enfin l'inclinaison du stylo dans l'espace à 3 dimensions lorsque la personne écrit.

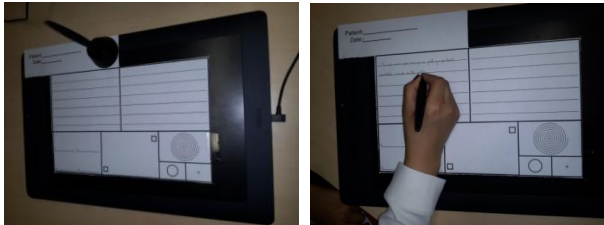


Figure 1: Tablette Wacom avec feuille lors d'une acquisition

Nous demandons aux participants de réaliser une série d'exercices. Le premier consiste à recopier un texte imposé (voir Figure 2). L'analyse présentée dans cet article ne concerne que cette tâche du protocole complet qui comprend aussi un texte libre, des boucles, des cercles, une spirale, parmi d'autres tâches.

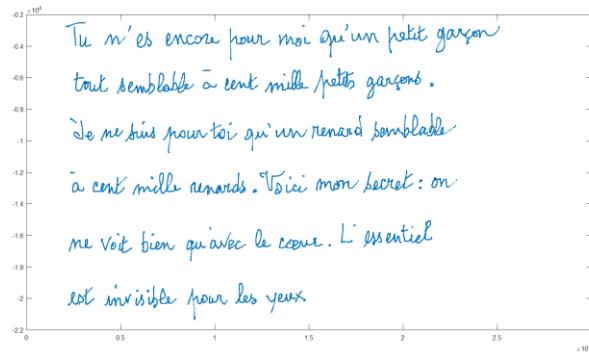


Figure 2: Texte imposé à recopier

### 2.3 Présentation des données acquises

A ce jour, nous avons pu acquérir les données auprès de 182 personnes parmi lesquelles nous avons : 29 patients Alzheimer (MA), 86 personnes souffrant des troubles cognitifs légers (TCL) dont (41 TCL exécutif, 8 TCL amnésique et 37 TCL multi-domaines), 1 personne souffrant à la fois de l'Alzheimer et du Parkinson (ADP), 5 parkinsoniens (Par), 27 personnes contrôles, 6 personnes souffrant de démence inconnue et 28 personnes dont on n'a pas encore de diagnostic (Non-d). La répartition de ces différentes données selon le profil cognitif des sujets est présentée dans le Tableau 1.

Tableau 1: Répartition des différents profils cognitifs

Diagnostic	Nombre personnes
MA	29
ADP	1
HCC	28
TCL	86
Non-d	27
Par	5
DEUNK	6

Nos données sont constituées de 128 femmes et 54 hommes. Concernant la latéralité des participants, nous avons 165 droitiers, 3 gauchers, 13 gauchers contrariés et 1 ambidextre. Ces données sont présentées dans le Tableau 2 ci-dessous.

Tableau 2: répartition des données en fonction du sexe et de la latéralité

Sexe		Latéralité		
Hommes	Femmes	Droitiers	Gauchers	Ambidextre
53	129	165	4	13

L'âge des participants varie majoritairement entre 70 et 85 ans, comme on peut l'observer sur la Figure 3, où chaque couleur représente un profil cognitif (voir légende).

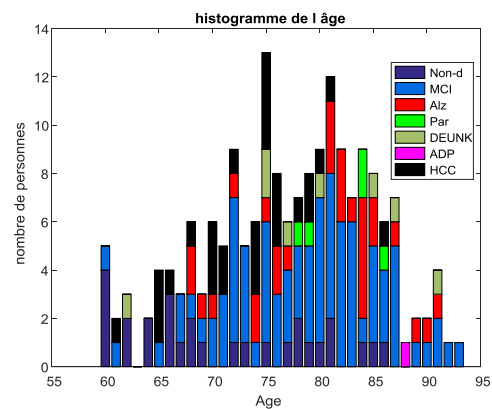


Figure 3: Répartition de l'âge des participants

Parmi toutes les données collectées, nous en avons gardé 142 ; en effet, nous avons exclu toutes les personnes pour qui nous n'avons pas encore eu les bilans cognitifs, les parkinsoniens, et les personnes souffrant de démence inconnue. Cette description est donnée sur la Figure 4.

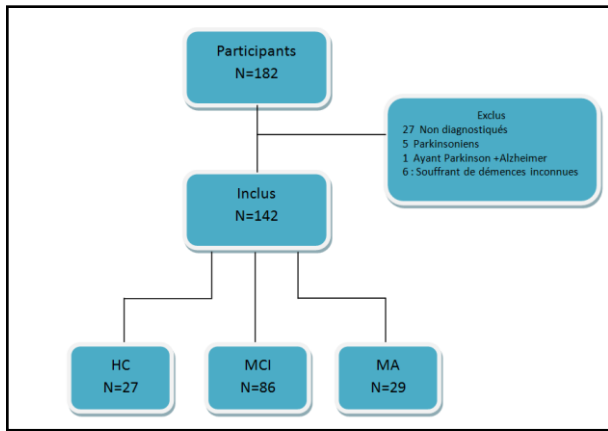


Figure 4: Organigramme des participants retenus et exclus

## 2.4 Extraction des paramètres

La fréquence d'échantillonnage du tracé est telle que nous avons un point tous les 8 millisecondes (125 Hz). En chaque point  $n$ , nous avons le triplet  $(x(n), y(n), P(n))$  représentant respectivement l'abscisse, l'ordonnée et la pression du stylo.

Les paramètres que nous avons extraits peuvent être classés en 3 catégories :

Cinématiques : vitesse (en  $x$ ,  $y$ , et le module), accélération, « jerk » (en  $x$ ,  $y$ , et le module) ou dérivée de l'accélération, ...

Spatiaux: direction, courbure, longueur des espaces intra-mots, longueur des espaces inter-mots, hauteur des mots, complexité de la trajectoire du stylo (au moment où le stylo est en l'air)...

Dynamiques : pression, variation de pression.

Dans la catégorie de paramètres cinématiques, les vitesses horizontale et verticale sont estimées en chaque point  $n$  de la façon suivante :

$V_x(n) = \Delta x(n) / \Delta t(n)$  et  $V_y(n) = \Delta y(n) / \Delta t(n)$  où  $\Delta x(n) = x(n+1) - x(n-1)$ ,  $\Delta y(n) = y(n+1) - y(n-1)$  et  $\Delta t(n) = t(n+1) - t(n-1)$ ; en effet, du fait que la fréquence d'échantillonnage est très élevée, on peut estimer la dérivée en un point  $n$  en considérant deux points voisins ( $n+1$ ) and ( $n-1$ ) comme décrit dans la littérature [1].

Dans la catégorie de paramètres spatiaux, la direction  $\theta$  du déplacement du stylo et la courbure  $\Phi[1]$ , sont estimées localement sur le tracé comme en [1] :

$$\cos \theta(n) = \frac{\Delta x(n) / \Delta s(n)}{\sqrt{\Delta x(n)^2 + \Delta y(n)^2}}$$

où  $\Delta s(n) = \sqrt{\Delta x(n)^2 + \Delta y(n)^2}$   
 Et  $\Phi(n) = \theta(n+1) - \theta(n-1)$ .

Aussi, nous mesurons la complexité du tracé en l'air du stylo par le nombre d'intersections sur la trajectoire en l'air induite.

L'ensemble complet des paramètres est donné dans le Tableau 3. Au total 46 paramètres sont extraits dont 22 correspondant au moment où le stylo est en contact avec la tablette (pen-down) et 24 au moment où il y a levé de stylo (pen-up). Il est à noter que l'extraction des paramètres sur la trajectoire du stylo en l'air, a été exploitée par Peter et.al [20], dans le cadre de la maladie du parkinson.

Dans la littérature [3,4,5,6], des moyennes et écart-types des différents paramètres sont considérés, ce qui constitue une perte importante d'information locale du tracé. Nous proposons de coder cette information via une quantification automatique des paramètres décrite ci-dessous. Ainsi, nous gardons une version compacte de la distribution des valeurs locales des paramètres sur l'ensemble du tracé.

Afin d'illustrer ces informations locales, nous présentons en Figure 5 un texte écrit par un participant montrant toutes les valeurs des vitesses en Y ( $V_y$ ) par un code couleur.

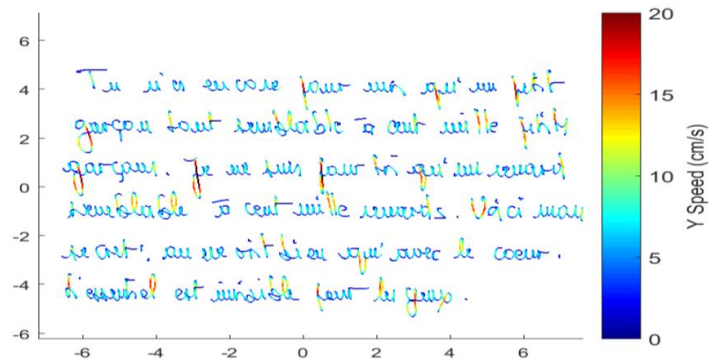


Figure 5: Visualisation des vitesses en Y ( $V_y$ )

La couleur bleu représente les vitesses faibles et la couleur rouge représente les vitesses fortes sur des plages allant de 0 à 20 cm/s.

### Quantification vectorielle

La quantification est effectuée à l'aide de l'algorithme des k-moyennes [2] car nous avons préféré extraire les frontières des « bins » de façon naturelle. Pour chaque paramètre, on considère ses valeurs pour toutes les personnes, et à l'aide de l'algorithme des k-moyennes, on fait un clustering sur 5 « bins ». Pour chaque cluster, on récupère le minimum et le maximum qui serviront de frontières des « bins » obtenus.

Pour une personne donnée, nous calculons les fréquences relatives des points dans chacun des « bins » obtenus, ce qui permet d'obtenir une description sur 5 dimensions du paramètre donné.

Cette procédure est illustrée sur la Figure 6, pour une personne donnée et pour le paramètre  $V_y$ .

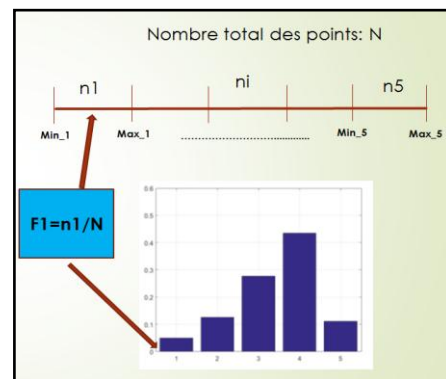


Figure 6: Illustration de la procédure de quantification vectorielle avec k-moyennes

## 3. Analyse préliminaire

A ce stade de notre travail, nous procédons à une analyse paramètre par paramètre afin de bien pouvoir analyser l'apport de chacun dans la caractérisation des profils cognitifs. Pour chaque paramètre codé sur 5 « bins », nous effectuons une classification hiérarchique ascendante (CAH) [7] avec le critère de Ward [8]

comme distance entre éléments, qui est moins sensible aux valeurs aberrantes. La distance entre deux clusters C1 et C2, est donnée par :

$$D(C1,C2)=\frac{P1P2}{PA+PB}d^2(G1,G2) \quad (1)$$

où P1, P2 sont les poids des deux clusters et G1 et G2 sont leurs respectifs.

Pour mesurer la qualité de chaque paramètre sur sa capacité à regrouper les sujets selon leurs profils cognitifs, on mesure la qualité du clustering fondée sur ce paramètre à l'aide de l'entropie normalisée [9] dont l'expression est la suivante :

$$\eta(c_k) = - \sum_{i=1}^{N_A} \frac{p(A_i|C_k) \log_2(p(A_i|C_k))}{\log_2(N_A)} \quad (2)$$

où  $N_A$  représente le nombre des profils cognitifs, à savoir le profil personne contrôle, le profil Alzheimer et le profil Troubles Cognitifs Légers,  $C_k$  le nombre total d'individus situés dans le cluster  $k$  et  $A_i$  le nombre d'individus ayant un profil cognitif  $i$ .

L'entropie totale pour un paramètre est donnée par :

$$E[\eta] = \sum_{k=1}^{N_C} \frac{|C_k|}{|U_{j=1}^{N_C} C_j|} \eta(c_k) \quad (3)$$

Plus cette entropie est faible (proche de 0), plus les profils cognitifs sont homogènes dans les clusters ; plus elle est élevée (proche de 1), plus les profils cognitifs sont hétérogènes dans les clusters. Plus un paramètre permet d'accroître l'homogénéité des profils cognitifs dans l'ensemble des clusters, plus il sera de bonne qualité. Nous avons ainsi pu classer les paramètres en fonction de leur qualité, par ordre croissant d'entropie. L'entropie varie de 0.72 à 0.77 pour nos 46 paramètres comme le montre le Tableau 3. La différence observée entre les valeurs de l'entropie de différents paramètres n'est pas très grande et ces entropies sont relativement élevées par rapport à 0, qui est le cas parfait où chaque cluster contient un seul profil cognitif ; Il faut rappeler cependant que ces expériences sont réalisées paramètre par paramètre

Nos résultats préliminaires montrent que la complexité du tracé en l'air et le nombre de points en l'air où l'accélération verticale est nulle (le nombre de fois où le scripteur accélère puis décélère ou inversement dans la trajectoire en l'air) sont les paramètres ayant une entropie qui est relativement faible par rapport aux autres paramètres. Il est à noter qu'ils correspondent à la trajectoire « inconsciente » réalisée par la personne lors de la tâche d'écriture du texte imposé. Les hésitations et le déficit de planification dans la succession des gestes à effectuer pour recopier le texte ont en effet tendance à augmenter la complexité de la trajectoire en l'air. D'autre part, le nombre de points en l'air à accélération verticale nulle indique le nombre de fois où le scripteur accélère puis décélère ou inversement. Dans [3], les auteurs ont remarqué que les trajectoires en l'air étaient plus lentes et plus longues pour les malades d'Alzheimer comparativement aux personnes contrôle lorsqu'un texte de 107 caractères en hébreu est recopié. Notre résultat sur la trajectoire en l'air va dans le sens de cette étude mais apporte d'autres marqueurs : complexité et variations de l'accélération sur la verticale.

Enfin, le troisième paramètre donnant un bon résultat correspond au module du « jerk » quand la personne écrit sur la tablette. Ce paramètre caractérise la fluidité du geste d'écriture. Il apparaît dans la littérature comme étant important pour caractériser la maladie, ce que nous retrouvons bien.

Contrairement à Werner et al. [3], nous ne trouvons pas que la pression est un bon paramètre pour caractériser les profils cognitifs.

Tableau 3: Les 46 paramètres spatiotemporels

Paramètre	Entropie
Complexité du tracé en l'air	0.71
Nombre de points en l'air à accélération verticale nulle	0.71
Module du Jerk sur tracé sur tablette	0.72
Module de l'accélération sur tablette	0.73
Temps en l'air inter-mot	0.73
Module de la vitesse sur tablette	0.73
«Jerk» vertical normalisé	0.73
Vitesse en y sur tablette	0.73
« Jerk » en x en l'air	0.74
Vitesse en y en l'air	0.74
Vitesse en x en l'air	0.74
« Jerk » en y en l'air	0.74
Longueur des portions t.q. Vy=0	0.74
Nombre de points en l'air à Jerk vertical nul	0.74
Accélération en y en l'air	0.74
Accélération en x en l'air	0.74
Module de la vitesse en l'air	0.74
Longueur des portions t.q. Vx=0 en l'air	0.74
Accélération en y sur tablette	0.75
Jerk en x normalisé en l'air	0.75
Nombre de points en l'air à jerk horizontal nul	0.75
Angle mot-ligne de base	0.75
Temps en l'air intra-mot	0.75
Module du Jerk en l'air	0.75
Variation de pression	0.75
Jerk vertical sur tablette	0.75
Direction en l'air	0.75
Nombre de points en l'air à Ax null	0.75
Nombre des portions t.q. Vy=0 en l'air	0.75
temps des portions t.q. Vy=0 sur tablette	0.75
Module de l'accélération en l'air	0.75
Nombre de points sur tablette à Jerk en y null	0.75
Jerk normalisé en y sur tablette	0.76
temps des portions t.q. Vx=0 en l'air	0.76
Longueur des portions t.q. Vy=0 sur tablette	0.76
Temps des portions t.q. Vy=0 en l'air	0.76
Espace intra-mots	0.76
Nb_Ay_null Pen-down	0.76
Nombre des portions t.q. Vx=0 en l'air	0.76
Courbure déplacement en l'air	0.76
Ecrasement horizontal	0.77
Pression	0.77
Courbure déplacement en l'air	0.77
Espace inter-mots	0.77
Direction du déplacement sur tablette	0.77
Taille du corps du mot	0.77

### 3 CONCLUSION

Lors de cette phase préliminaire de nos travaux, nous avons procédé à une analyse des paramètres un à un, en effectuant un clustering par classification hiérarchique ascendante, puis en mesurant la qualité de chacun par rapport à sa capacité à séparer les profils cognitifs dans les clusters obtenus. Une mesure d'entropie a été utilisée à ces fins. Ces valeurs d'entropie nous ont permis de classer les paramètres selon leur apport. Nous notons

néanmoins que ces valeurs d'entropie sont moyennement élevées car il s'agit d'une mesure effectuée à la suite d'un clustering en prenant chacun des paramètres séparément. Nous avons ainsi dégagé la complexité du tracé en l'air, le nombre de fois où le scripteur accélère puis décélère ou inversement dans la trajectoire en l'air, et le module du « jerk » sur la tablette comme étant les trois meilleurs paramètres. Ce résultat préliminaire montre que la trajectoire en l'air qui est inconsciente est porteuse d'information comme l'avait déjà montré Peter et al dans le cas de la maladie du Parkinson [20]. Concernant l'Alzheimer, la littérature le mentionne mais avec un autre paramètre [3] (la durée la trajectoire en l'air) qui n'a pas donné de bons résultats comparativement à la complexité de la trajectoire en l'air dans notre étude. Ainsi, le paramètre de complexité est nouveau par rapport à l'état de l'art et celui de l'accélération n'y apparaît pas pour la trajectoire en l'air.

En perspective, nous comptons sélectionner les meilleurs paramètres pour procéder à un regroupement automatique des données, pour dégager différents profils sans connaissance a priori des profils cognitifs. Pour cela, nous comparerons notre démarche à des tests statistiques (ANOVA, Kruskal-Wallis) utilisés dans la littérature.

Cette étude transversale permettra de conclure quels paramètres caractérisent l'écriture des patients Alzheimer, comparativement aux TCL et aux personnes contrôles.

Une fois cette analyse transversale terminée, nous ferons l'étude longitudinale, afin de déterminer quels paramètres ont une valeur prédictive de la maladie d'Alzheimer.

## REFERENCES

- [1] I. Guyon, P. Albrecht, Y. Le Cun, J. Denker, and W. Hubbard, "Design of a neural network character recognizer for a touch terminal". *Pattern Recognition*. Vol. 24 Issue 2, 1991, pp. 105-119.
- [2] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100-108.
- [3] Werner P., Rosenblum S., Bar-On G., Heinik J. and Korczyn A., Handwriting Process Variables Discriminating Mild Alzheimer's Disease and Mild Cognitive Impairment, in *Journal of Gerontology: Psychological Sciences*, Volume 61B (4), pages 228-236, 2006.
- [4] Yan J. H., Rountree S., Massman P., Smith Doody R. and Li H., Alzheimer's disease and mild cognitive impairment deteriorate fine movement control, in *Journal of Psychiatric Research*, volume 42(14), pages 1203-1212, 2008.
- [5] Schröter A., Mergl R., Bürger K., Hampel H., Möller H.-J. and Hegerl U., Kinematic Analysis of Handwriting Movements in Patients with Alzheimer's Disease, Mild Cognitive Impairment, Depression and Healthy Subjects, in *Dementia and Geriatric Cognitive Disorders*, Volume 15(3), pages 132-142, 2003.
- [6] Slavin M. J., Phillips J. G., Bradshaw J. L., Hall K. A. and Presnell I., Consistency of handwriting movements in dementia of the Alzheimer's type: A comparison with Huntington's and Parkinson's diseases, in *Journal of the International Neuropsychological Society*, Volume 5(1), pages 20-25, 1999.
- [7] Rokach, L., & Maimon, O. (2005). Clustering methods. In *Data mining and knowledge discovery handbook* (pp. 321-352). Springer US.
- [8] Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301), 236-244.
- [9] Shannon, C. E. (1948). A mathematical theory of communication, *bell System technical Journal* 27: 379-423 and 623-656. *Mathematical Reviews (MathSciNet): MR10, 133e*.
- [10] Teulings, H. L., Contreras-Vidal, J. L., Stelmach, G. E., & Adler, C. H. (1997). Parkinsonism reduces coordination of fingers, wrist, and arm in fine motor control. *Experimental neurology*, 146(1), 159-170.
- [11] van Gemmert, A. W., Teulings, H. L., & Stelmach, G. E. (1998). The influence of mental and motor load on handwriting movements in Parkinsonian patients. *Acta psychologica*, 100(1), 161-175.
- [12] Drotár P., Mekyska J., Rektorová I., Masarová L., Smekal Z and Faundez-Zanuy M., A New Modality for Quantitative Evaluation of Parkinson's Disease: In-Air Movement, *IEEE 13th Intl. Conf. on Bioinformatics and Bioengineering (BIBE)*, 2013.
- [13] Rosenblum S, Samuel M., Zlotnik S, Erikh I, Schlesinger I., Handwriting as an objective tool for Parkinson's disease diagnosis, *Journal of Neurology*, 2013; DOI: 10.1007/s00415-013-6996-x.
- [14] American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- [15] Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcomes. *Archives of Neurology*, 56, 303-308.
- [16] C. O'Reilly and R. Plamondon, "Design of a neuromuscular disorders diagnostic system using human movement analysis", in *Proc. of the 11<sup>th</sup> International Conference on Information Sciences, Signal Processing and their Applications*, pp.787-792, 2012.
- [17] L. R. B. Schomaker, "Simulation and recognition of handwriting movements: a vertical approach to modeling human motor behavior", Ph.D. Thesis, Nijmegen University, Netherlands, 1991.
- [18] A.W.A. Van Gemmert and H; L. Teulings, "Advances in graphonomics: studies on fine motor control, its development and disorders". *Human Movement Science*, vol. 25, pp. 447-453, 2006.
- [19] John M. Hollerbach, "An Oscillation Theory of Handwriting", *Biol. Cybernetics* 39, 139-156, 1981.
- [20] Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M. (2014). Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease. *Computer methods and programs in biomedicine*, 117(3), 405-411.