



**HAL**  
open science

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

Christian Kahindo, Sonia Garcia-Salicetti, Mounim El Yacoubi, Victoria Cristacho-Lacroix, Hélène Kerherve, Anne-Sophie Rigaud

### ► To cite this version:

Christian Kahindo, Sonia Garcia-Salicetti, Mounim El Yacoubi, Victoria Cristacho-Lacroix, Hélène Kerherve, et al. Caractérisation de la Maladie d'Alzheimer à travers l'Écriture manuscrite acquise sur Tablette graphique. Journées d'Étude sur la TélÉSANTé, 6ème édition, Pôle Capteurs, Université d'Orléans, May 2017, Bourges, France. hal-01564669

**HAL Id: hal-01564669**

**<https://hal.science/hal-01564669v1>**

Submitted on 19 Jul 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Caractérisation 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

Dans cette étude, nous analysons l'écriture manuscrite acquise sur tablette graphique de patients touchés par la maladie d'Alzheimer (MA) à un stade précoce, et de Troubles Cognitifs Légers (TCL), suivis à l'hôpital Broca. L'objectif est de caractériser par des paramètres spatio-temporels de l'écriture la maladie d'Alzheimer à ses débuts. Nous proposons une méthode de sélection automatique de paramètres extraits globalement sur un texte imposé. Cette sélection repose sur un critère d'information mutuelle normalisée entre un ensemble de paramètres et les différents profils cognitifs (MA, TCL et sujets « contrôle »). Cette méthode intègre un regroupement automatique par Classification Hiérarchique Ascendante (CAH).

L'étude a été réalisée auprès de 182 participants dont 29 Alzheimer et 86 Troubles Cognitifs Légers. Nous montrons que certains paramètres cinématiques de l'écriture manuscrite permettent de caractériser les MA par rapport aux sujets « contrôle », via un Clustering. Aussi, il apparaît que les paramètres cinématiques sélectionnés par notre méthode, révèlent deux profils différents pour les Troubles Cognitifs Légers: un davantage proche de l'écriture des MA et l'autre davantage proche de celle des sujets « contrôle ».

**Mots-clés:** Ecriture manuscrite, tablette graphique, Alzheimer, Clustering, information mutuelle normalisée.

## 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, peut être 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. En effet, plusieurs travaux dans la littérature portent sur l'analyse de l'écriture et la caractérisation de diverses pathologies ou états du sujet: la maladie de Parkinson [10,11,12,13], d'Alzheimer [3,4,5,6] et des états émotionnels (stress, dépression, anxiété) [23]. La caractérisation de la maladie d'Alzheimer (MA) et des Troubles Cognitifs Légers « Mild Cognitive Impairment » (MCI) 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 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 sur 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.

Plus récemment, une étude faite par Kawa et al [21] sur des patients souffrant de TCL et des personnes contrôle, a montré que

sur un exercice de dictée d'un texte, l'écriture des TCL prend plus de place en termes de nombre de lignes et taille des lettres, que celle des personnes contrôle; aussi, la vitesse d'écriture des TCL s'avère faible comparativement à celle des personnes contrôle. Enfin, Yu et al. [22] ont analysé la motricité chez les MA, TCL et personnes contrôle. Sur un exercice qui consiste à faire une série de « l » cursifs, ils concluent par des tests statistiques qu'il n'y a pas de différence entre les différents profils cognitifs, sur la vitesse d'exécution, la hauteur et la largeur des « l ». En revanche, sur les exercices de précision qui consistent à faire des allers-retours entre deux cibles, la vitesse diffère de façon significative entre les différents profils cognitifs.

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 suffisant. Notre démarche proposée dans la suite du travail vise à pallier ces limites.

### **2.2 Acquisition**

L'acquisition des données se déroule au sein de l'Hôpital Broca. Nous ciblons principalement des patients touchés par l'Alzheimer à un stade débutant (MMSE>20) 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.

Les critères d'inclusion suivants sont retenus :

- Etre âgé de plus de 60 ans ;
- Parler couramment le français ;
- 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 cognitif) 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) ;

### **Protocole d'acquisition**

Dans notre protocole, chaque participant doit effectuer d'office un bilan neuropsychologique établissant son profil cognitif. Quant aux patients Alzheimer et aux Troubles Cognitifs Légers en plus

du bilan neuropsychologique, d'autres examens (IRM, ponction lombaire etc.) sont effectués afin de s'assurer qu'ils sont bien Alzheimer ou TCL.

L'acquisition complète de l'écriture manuscrite pour un participant donnée dure en moyenne de 45 minutes.

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. Ces données nous permettront de faire l'étude longitudinale.

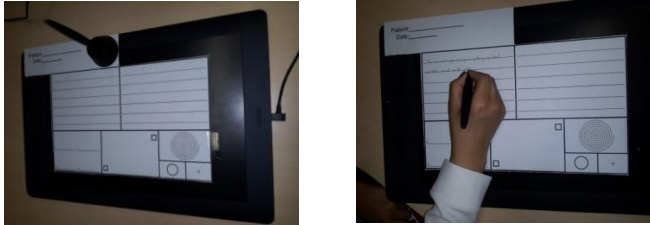


Figure 1:Tablette Wacom avec feuille lors d'une acquisition

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). Nous posons une feuille sur la tablette afin de permettre un retour visuel du tracé effectué; 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.

Nous demandons aux participants de réaliser une série d'exercices. Le premier consiste à recopier un texte imposé (voir Figure 5). 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.

### 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 (HCC), 6 personnes souffrant de démence inconnue et 28 personnes dont on n'a pas encore de diagnostic (Non-d).

La répartition des données en fonction de la latéralité et du sexe est résumée dans le Tableau 1 ci-dessous.

Tableau 1: 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).

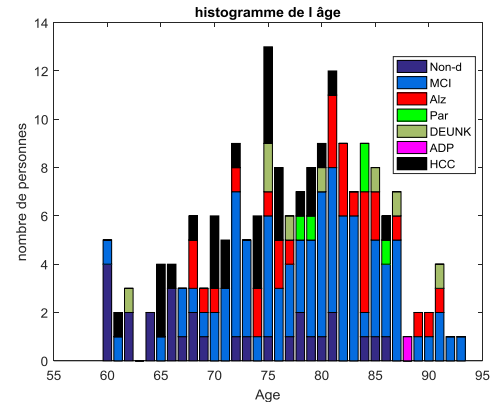


Figure 2: Répartition de l'âge des participants

Etant donné que nous ne travaillons que sur la maladie d'Alzheimer et les troubles Cognitifs Légers, nous avons gardé 142 participants parmi tous ceux qui ont participé; qui sont soit Alzheimer, Troubles Cognitifs Légers ou Contrôle. Ces données sont présentées sur la Figure 4.

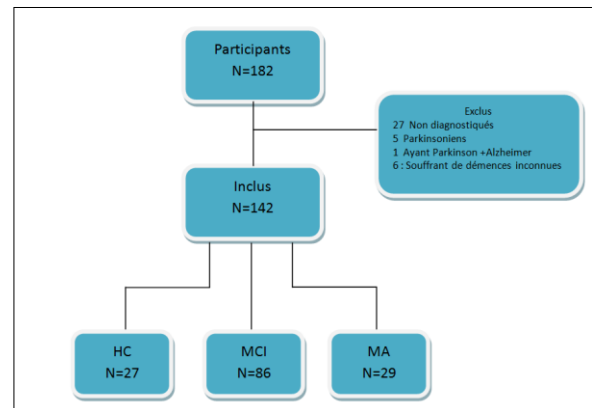


Figure 3: Organigramme des participants retenus et exclus

### 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 » ou dérivée de l'accélération (en  $x$ ,  $y$ , et le module), ...

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 ( $n+1$ ) et ( $n-1$ ) comme décrit dans la littérature [1].

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

$$\cos\theta(n)=\Delta x(n)/\Delta s(n)$$

$$\text{où } \Delta s(n)=\sqrt{\Delta x(n)^2 + \Delta y(n)^2}$$

$$\text{et } \Phi(n)=\theta(n+1)-\theta(n-1).$$

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 un 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], les 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 les valeurs des vitesses en Y ( $V_y$ ) par un code couleur.

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

### Procédure de quantification

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  $V_y$ .

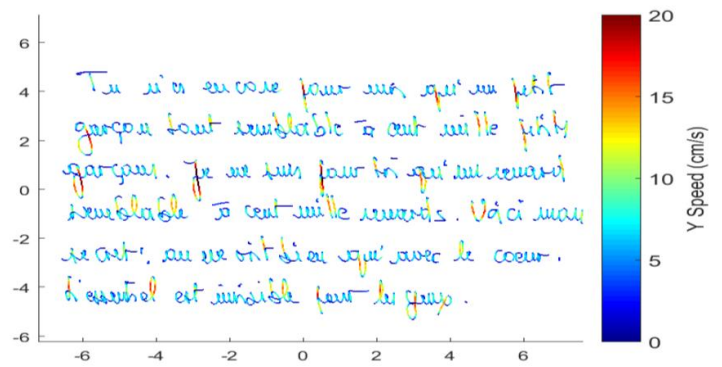


Figure 4: Visualisation des vitesses en Y ( $V_y$ ) sur le texte imposé

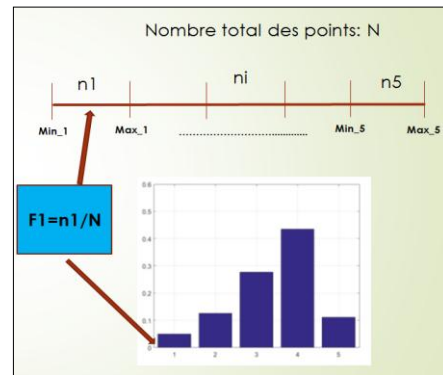


Figure 5: Illustration de la procédure de quantification vectorielle avec k-moyennes

## 2. 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 :

$$D(C1,C2)=\frac{P1P2}{P1+P2}d^2(G1,G2) \quad (1)$$

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

Pour sélectionner les paramètres les plus discriminants et déterminer le nombre optimal de clusters, nous utilisons une procédure itérative de sélection des paramètres pénalisant l'augmentation du nombre de clusters. Cette procédure est basée sur l'information mutuelle normalisée définie par l'expression (2). A chaque itération  $j$ , on rajoute le  $j$ ème paramètre qui avec ceux précédemment sélectionnés permet de maximiser l'information mutuelle normalisée :

$$IMN(C,A)=\frac{IM(C,A)}{(H(C)+H(A))/2} \quad (2)$$

où  $C$  fait référence aux clusters,  $A$  correspond aux classes et  $IM$  est l'information mutuelle définie par :

$$IM(C, A) = \sum_{k=1}^{N_C} \sum_{i=1}^{N_A} p(C_k \cap A_i) \log_2 \left( \frac{p(C_k \cap A_i)}{p(C_k) \times p(A_i)} \right) \quad (3)$$

avec  $N_A$  et  $N_C$  correspondant respectivement aux nombres de classes et de clusters, et  $H(C)$  à l'entropie sur les clusters [9]:

$$H(C) = - \sum_{k=1}^{N_C} p(C_k) \log_2(p(C_k)) \quad (4)$$

où

$N_C$  et  $C_k$  : sont respectivement le nombre de clusters et le cluster  $k$ .

et  $H(A)$  l'entropie sur les classes :

$$H(A) = - \sum_{i=1}^{N_A} p(A_i) \log_2(p(A_i)) \quad (5)$$

avec

$N_A$  et  $A_i$  correspondant respectivement au nombre de classes et à la classe  $i$  (MA, TCL, HCC).

Plus l'information mutuelle normalisée se rapproche de 1, plus les profils cognitifs sont homogènes dans les clusters ; plus elle est petite (proche de 0), plus les profils cognitifs sont hétérogènes dans les clusters. Ainsi, plus un paramètre permet d'accroître l'homogénéité des profils cognitifs dans l'ensemble des clusters, plus il sera de bonne qualité. La meilleure combinaison de paramètres est ainsi celle permettant d'obtenir la valeur maximale de l'information mutuelle normalisée.

Nous avons appliqué cette méthode dans deux cas différents.

Dans le premier cas, nous avons effectué le Clustering à l'aide de la CAH sur 2 classes seulement, les HCC et les MA. Les meilleurs paramètres obtenus pour ce cas sont : Le jerk vertical sur la trajectoire en l'air, le Jerk normalisé par rapport à la longueur et à la durée sur la trajectoire en l'air, le module de la vitesse sur la trajectoire sur tablette et le nombre de stokes (portion du tracé telle que la vitesse verticale  $V_y$  est nulle sur la trajectoire en l'air). Les résultats obtenus sont donnés dans la Table 3.

Table 3 : Clusters obtenus pour le cas HCC vs MA

	HCC	MA	total
Cluster_1	24	6	30
Cluster_2	4	23	27
Nbr Sujets	28	29	57

Dans ce tableau, on constate que dans le cluster 1, nous avons une majorité des personnes contrôle (HCC) et dans le cluster 2, une majorité des malades d'Alzheimer (MA).

Dans le second cas, nous effectuons un Clustering en prenant les trois profils cognitifs à savoir les HCC, TCL et les MA. Avec la même méthode, l'ensemble des paramètres les plus discriminants est constitué des paramètres suivants : le nombre de points à accélération verticale nulle sur la trajectoire en l'air (caractérisant les extrema locaux de la vitesse en  $y$  (verticale)), le temps inter-mot, et le jerk vertical ( $J_y$ ) sur la trajectoire sur tablette.

Les clusters obtenus sont présentés dans la Table 4.

Table 4 : Clusters obtenus pour les 3 profils HCC, TCL et MA

	HCC	TCL	MA	total
Cluster_1	2	2	5	9
Cluster_2	2	38	21	61
Cluster_3	24	47	3	74
Nbr Sujets	28	87	29	144

On constate que le cluster 2 qui contient la plupart des MA et une partie des TCL, et que le cluster 3 contient presque tous les HCC et une partie des TCL. On a donc deux groupes de TCL qui se dégagent : un qui est proche des MA et un autre qui est proche des HCC.

Dans les deux cas, on constate que le jerk (variation de l'accélération) fait partie des meilleurs paramètres. 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.

On remarque également que certains paramètres sur la trajectoire en l'air font partie des meilleurs paramètres. 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 avec un autre marqueur, le jerk vertical, qui caractérise les variations de l'accélération verticale.

### 3 CONCLUSION

Dans cette étude, nous proposons une méthode de sélection automatique des paramètres cinématiques de l'écriture, basée sur un critère d'information mutuelle entre ces paramètres et les profils cognitifs. Cette sélection est faite itérativement, en utilisant un clustering hiérarchique. Nous avons ainsi dégagé l'ensemble de paramètres optimaux pour caractériser d'abord les malades Alzheimer à un stade précoce par rapport aux sujets sains : sur la trajectoire en l'air, le jerk vertical, le jerk normalisé, et le nombre de stokes, puis, sur la trajectoire sur tablette, le module de la vitesse.

En un deuxième temps, nous avons effectué la sélection des paramètres en considérant aussi les Troubles Cognitifs Légers. Les paramètres optimaux sont les suivants : sur la trajectoire en l'air, le nombre de points à accélération verticale nulle et le temps inter-mot ; sur la trajectoire sur tablette, le jerk vertical.

Ces résultats montrent que lorsque la personne écrit sur la tablette, sa vitesse d'écriture et son jerk dans la direction verticale sont porteurs d'information discriminante de la classe des Alzheimer par rapport à celle des sujets sains. Aussi, dans ce cas (MA vs. HCC), nous constatons que lorsque la personne a le stylo levé lors de la tâche d'écriture (entre deux mots, ou quand elle change de ligne, ou pour une autre raison), les variations de l'accélération (jerk) sont porteuses d'information ainsi que le temps écoulé entre mots successifs. Dans la littérature, sur le même type de tâche (recopie d'un texte imposé), Werner et al. ont conclu par un test statistique que la vitesse moyenne permet de différencier les Alzheimer des sujets sains ; en revanche, ils concluent que c'est

surtout le temps en l'air et la pression qui permettent de différencier les deux classes, ce que nous ne trouvons pas. Nos résultats montrent que sur la trajectoire en l'air, c'est le jerk et le nombre de strokes qui s'avèrent informants.

En perspective, nous comptons approfondir cette étude, notamment via la recherche d'autres paramètres discriminants. Aussi, nous étudierons d'autres tâches réalisées par les patients lors de l'acquisition des données manuscrites. 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., Smekal, 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.
- [21] Kawa, Jacek, Bednorz, Adam, Stepień, Paula, et al. *Spatial and Dynamical Handwriting Analysis in Mild Cognitive Impairment*. *Computers in Biology and Medicine*, 2017.
- [22] YU, Nan-Ying et CHANG, Shao-Hsia. Kinematic Analyses of Graphomotor Functions in Individuals with Alzheimer's Disease and Amnesic Mild Cognitive Impairment. *Journal of Medical and Biological Engineering*, 2016, vol. 36, no 3, p. 334-343.
- [23] L. Likforman-Sulem, A. Eposito, M. Faundez-Zanuy, S. Cléménçon and G. Cordasco, "EMOTHAW: A Novel Database for Emotional State Recognition From Handwriting and Drawing". *IEEE Transactions on Human-Machine Systems*, 2017, vol. 47, no 2, pp. 273-284.