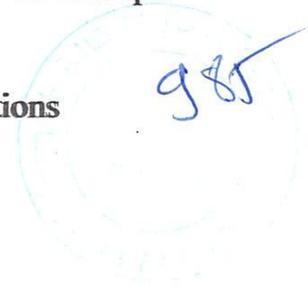


République Algérienne Démocratique et Populaire  
Ministère de l'enseignement supérieur et de la recherche scientifique  
Université 8Mai 1945 – Guelma  
Faculté des sciences et de la Technologie  
Département d'Electronique et Télécommunications



**Mémoire de fin d'étude  
pour l'obtention du diplôme de Master  
Académique**

**Domaine : Sciences et Technologie  
Filière : Electronique  
Spécialité : Instrumentation**

---

**TECHNIQUES DE RECONNAISSANCE DES FORMES :  
APPLICATION A L'AUTHENTIFICATION DE SIGNATURE  
MANUSCRITE EN LINGE**

---

Présenté par :  
**MEHIRA  
MOHAMMED RIDHA**

Sous la direction du :  
Professeur .BENNACER LAYACHI

**Juin 2018**



## Résumé

Notre travail s'inscrit dans le domaine de la reconnaissance des formes et est destiné à réaliser un système de vérification de signature manuscrite en ligne. La complexité de mise en œuvre d'un tel système croît avec le nombre de formes à reconnaître d'où l'impossibilité d'établir et de répertorier toutes les déformations issues de la signature d'un signataire quelconque ainsi que les autres formes de dégradation du tracé. L'objectif que nous nous sommes fixé pour contourner ce problème est de proposer une approche pour réaliser un système de vérification de signature. L'approche retenue est basée sur la transformée de Fourier. Le système est évalué sur deux bases de données : la base MCYT et la base SVC2004. Les résultats obtenus sont intéressants et avoisinent les 82% sur une population de 100 usagers

## Abstract

Our work is done in the field of pattern recognition and consists of realizing an on line handwritten signature verification system. The complexity to implement this system increases with the number of shapes to recognize, So it's impossible to establish and to index all the deformations resulting from the signature of all signatory and to define other types of shape degradation. The aim of this work is to achieve a system able de get round the writing shape deformation. Thus we propose an approach to realize a signature verification system based on the Fourier transform. The system is evaluated on two databases: MCYT and SVC2004. The results obtained are interesting and are close to 82% in a population of 100 users

## ملخص

تم عملنا في مجال التعرف على الأنماط ويتكون من تحقيق نظام التحقق من التوقيع بخط اليد. يزداد تعقيد تنفيذ هذا النظام مع عدد الأشكال التي يجب التعرف عليها، لذلك من المستحيل إنشاء وفهرسة جميع التشوهات الناتجة عن توقيع جميع الموقعين وتحديد أنواع أخرى من تدهور الشكل. والهدف من هذا العمل هو تحقيق نظام قادر على الحصول على شكل تشوه الكتابة. وبالتالي نقترح نهجاً لتحقيق نظام التحقق من التوقيع على أساس تحويل فورييه. يتم تقييم النظام على قاعدتي بيانات: MCYT و SVC2004. النتائج التي تم الحصول عليها هي مثيرة للاهتمام وتقترب من 82% من 100 مستخدم

# Remerciements

Je tiens tout d'abord à exprimer mes sincères remerciements à Monsieur le Professeur Bennacer Layachi qui a accepté d'encadrer et diriger le présent travail, qui m'a encouragé et soutenu le long de ce semestre. Je le remercie également pour tous ses conseils et orientations qu'il m'a apportés

Je remercie et salue humblement tous les enseignants qui ont accepté de juger ce modeste travail

# Table des matières

INTRODUCTION GENERALE :.....	1
Chapitre I ORGANISATION GENERALE D'UN SYSTEME BIOMETRIQUE ET ETAT DE L'ART .....	3
I.1 INTRODUCTION:.....	4
I.2 PROBLEMATIQUE : .....	4
I.3 ARCHITECTURE D'UN SYSTEME BIOMETRIQUE : .....	5
I.3.1 Cas d'un système de vérification de la signature en ligne : .....	5
I.3.2 Acquisition de la signature en ligne :.....	6
I.3.3 Les prétraitements : .....	7
I.3.3.1 Filtrage : .....	7
I.3.3.2 Ré-échantillonnage :.....	7
I.3.3.3 La normalisation :.....	8
I.3.4 Extraction de paramètres :.....	8
I.3.5 Comparaison : .....	9
I.3.6 Décision : .....	9
I.3.7 Entraînement du système : .....	10
I.4 EVALUATION DES SYSTEMES DE VERIFICATION DE SIGNATURES EN LIGNE :.....	11
CONCLUSION :.....	13
Chapitre II TECHNIQUES DE MODELISATION DES SIGNATURES MANUSCRITES.....	14
II.1 INTRODUCTION :.....	15
II.2 MODELE DE FRANCK LECLERC, GUY LORETTE & PLAMONNDON : .....	15
II.3 MODELISATION BASEE SUR L'OPERATEUR DE TEAGER-KAISER $\Psi_{\mathbb{R}}$ :.....	18
II.3.1 Description du modèle :.....	19
II.3.2 Applications et limites de l'opérateur $\Psi_{\mathbb{R}}$ :.....	22
II.4 MODELE D'HOLLERBACH :.....	22
II.5 MODELISATION PAR LA TRANSFORME DE FOURIER : (METHODE PROPOSEE PAR YANIKOGLU) : .....	23
II.6 MODELISATION BASEE SUR L'APPROCHE DE LA DECOMPOSITION EMPIRIQUE MODALE ET DE LA TRANSFORMEE DE HUANG HILBERT (THH) : ...	24
II.6.1 Décomposition modale empirique EMD : .....	24
II.6.1.1 Principe de la décomposition : .....	24
II.6.2 Transformée d'HILBERT : .....	25

II.6.2.1 Signal multi-composantes et mono-composante : .....	26
II.6.2.2 Estimation de la fréquence instantanée : .....	27
II.6.2.2 .1 Définition : .....	27
II.6.3 Transformation de Huang-Hilbert (THH) : .....	28
II.7 Modèle retenu.....	30
CONCLUSION .....	30
Chapitre III MECANISMES DE FONCTIONNEMENT DU SYSTEME ADOPTE .....	31
III.1 INTRODUCTION .....	32
III.2 LA TRANSFORMEE EN COSINUS DISCRETE .....	33
III.3 REDUCTION DE PARAMETRES.....	36
III.4 METRIQUES DE DIS-SIMILARITE UTILISEES .....	37
CONCLUSION .....	38
Chapitre IV RESULTATS EXPERIMENTAUX : Résultats des Performances du système évalués sur l'approche Comparaison.....	39
IV.1 SIGNAUX D'APPRENTISSAGE : .....	41
IV.2 PRE-TRAITEMENT : .....	41
IV.2.1 Extraction des segments de signature : .....	41
IV.2.2 Réduction des segments et obtention d'une signature mono-segment : .....	42
IV.2.3 Translation de la signature : .....	44
IV. 2.4 Normalisation en taille de la signature : .....	45
IV.2.5 Re-échantillonnage : .....	47
IV.3 OBTENTION DU SIGNAL DE REFERENCE : .....	48
IV.4 VERIFICATION : .....	49
IV.5 ALGORITHME D'OBTENTION DES SEUILS DE DECISION : .....	49
CONCLUSION : .....	49
CONCLUSION ET PERSPECTIVES : .....	59
Bibliographie : .....	60
Annexe : .....	63

# Liste des figures

Figure I.1 : Schéma bloc d'un système de reconnaissance biométrique.....	5
Figure I.2 : Phase d'entraînement d'un système biométrique à base de signature en ligne.....	6
Figure I.3 : Phase de vérification dans un système biométrique à base de signature en ligne.....	6
Figure I.4 : Dispositifs d'acquisition de la signature en ligne.....	6
Figure I-5 : Angles d'inclinaison du stylo.....	8
Figure II.1 : Processus de génération de tracés manuscrit.....	16
Figure II.2 : Modélisation du processus de génération.....	17
Figure II.3 : Schéma global du modèle.....	17
Figure II.4 : Schéma général du modèle utilisé.....	18
Figure II.5 : Vitesse curviligne d'un tracé manuscrit.....	18
Figure II.6 : L'opérateur d'énergie de Teager-Kaiser <i>dépiste</i> l'énergie fournie par la source sinusoïdale (normalisée par unité de masse-1) qui engendre le signal $x(t)$ .....	19
Figure II.7 : Représentation visuelle des équations d'Hollerbach, son excentricité ainsi que son orientation dépendent de $a$ , $b$ et $\phi = \phi_x - \phi_y$ dans le cas où les oscillateurs ont la même fréquence. Image empruntée à Hollerbach.....	23
Figure II.8 : Principe de décomposition de l'EMD.....	25
Figure II.9 : Principe d'estimation de la FI via le signal analytique $z_x(t)$ de $x(t)$ .....	28
Figure II.10 : Organigramme de la THH.....	29
Figure III.1: Exemple des signaux d'une signature en ligne ( $x(t)$ , $y(t)$ , $z(t)$ ).....	33
Figure III 2 : Transformée de Fourier des signaux de la signature montrée dans la figure (III-1).....	33
Figure III.3 : Transformée en cosinus des signaux de la signature montrée dans la figure III.1 ..	35
Figure III.4 : Coefficients paramètres DFT.....	36
Figure III.5 : Coefficients paramètres DCT.....	37
Figure IV.1 : Organisation générale du système étudié.....	40
Figure IV.2 : Exemple de signaux utilisés en phase d'apprentissage.....	41
Figure IV.3 : exemple de segment de tracer d'une signature.....	42
Figure IV.4 : exemple de suppression d'un segment ayant une longueur inférieure à un seuil égal 4.....	42
Figure IV.5 : algorithme d'extraction des segments d'une signature.....	43
Figure IV.6 : segmentation d'une signature afin d'obtenir une signature mono-segment.....	44
Figure IV.8 : normalisation en taille de la signature.....	46
Figure IV.9 : algorithme de la normalisation en taille de la signature.....	46
Figure IV.10 : Signaux d'apprentissage après normalisation.....	47
Figure IV.11: Signaux d'apprentissage après normalisation et filtrage.....	47
Figure IV.12 : Principe d'obtention du signal de référence à partir des signaux tests.....	48
Figure IV.13 : Courbes de performance pour la coordonnée TF[X].....	56
Figure IV.14 : Courbes de Performance pour la coordonnée TF[Y].....	57
Figure IV.15 : Courbes de Performance pour TF de la Pression.....	57

# Liste des tableaux

Tableau I.1 : Caractéristiques des bases de données de signatures en ligne. ....	12
Tableau IV.1: Déviation standard utilisée comme indice de ressemblance entre signatures. ....	51
Tableau IV.2 : détermination du taux de faux Rejet pour TH = 0.2. ....	52
Tableau IV.3 : Extraction du nombre de faux rejet. ....	53
Tableau IV.4 : Taux de faux rejet (FRR) obtenu. ....	53
Tableau IV.5 : détermination du taux de fausse acceptation pour TH = 0.2. ....	54
Tableau IV.6 : nombre de fausse acceptation (FAR) obtenu. ....	55
Tableau IV.7 : Taux de fausse acceptation (FAR) obtenu. ....	56
Tableau IV.8 : tableau des seuils estimés. ....	57
Tableau IV.9 : Tableau récapitulatif faisant intervenir la combinaison des paramètres avec le FAR, FRR, EER qui leurs est associés. ....	58

# Abréviations et Symboles

PDA	Personal Digital Assistant, Assistant Numérique Personnel.
DTW	Dynamic Time Warping, Déformation temporelle dynamique.
HMM	Hidden Markov Model, Modèle De Markov Cache
SVM	Support Vector Machine, Machine à vecteurs de support.
GMM	Gaussian Mixture Model, Modèle à mélange de gaussiennes.
EER	Equal Error Rate, Taux d'erreurs égales.
TKEO	Teager Kaiser Energy Operator, L'opérateur d'énergie de Teager-Kaiser.
FT	Fourier Transform, TF La transformée de Fourier.
DFT	Discrete Fourier Transform, TFD La Transformée de Fourier discrète.
DCT	Discrete Cosine Transform, TCD La transformée en cosinus discrète.
IDCT	Inverse Discrete Cosine Transform.
FFT	Fast Fourier Transform, La transformation de Fourier rapide.
EMD	Empirical Mode Decomposition.
IMF	Intrinsic Mode Function.
THH	Transformation de Huang-Hilbert.
FRR	False Rejection Rate, Taux de faux rejets.
FAR	False Acceptance Rate, Taux de fausses acceptations.
LDA	Linear Discriminant Analysis

## INTRODUCTION GENERALE

Malgré tous les progrès technologiques, la signature reste le moyen le plus utilisé pour authentifier un document, valider un contrat ou une transaction financière. La signature peut être également utilisée comme un mot de passe pour consulter des documents confidentiels, ou tout simplement pour accéder à son bureau. Elle est aussi fiable que l'identification vocale ou rétinienne. Les systèmes de vérification de signature sont des systèmes traitant le manuscrit ; ces derniers sont subdivisés en deux catégories les systèmes « *en ligne* » et les systèmes « *hors ligne* ». Un système hors ligne traite la signature à partir d'une image, tel que l'émargement sur un chèque ou sur un document. Un simple scanner suffit pour capturer la signature. Un prétraitement est indispensable à l'extraction de la signature et on ne dispose que des données statiques. Dans le cas d'un système en ligne, la signature est acquise directement sur une tablette à l'aide un stylo électronique, on peut donc relever des caractéristiques dynamiques telles que le temps de signature, la pression ou l'inclinaison du stylo.

Pour un système de vérification en ligne l'interface d'acquisition influe sur la conception du système, ainsi une tablette simple ne peut servir à des méthodes de vérifications nécessitant des paramètres tels que la pression ou l'inclinaison du stylo, pourtant des méthodes utilisant ces paramètres sont proposées par Y.K.T Ohishi et col [OHI00]. Une tablette classique suffira pour des algorithmes utilisant des paramètres tels que la vitesse et l'accélération [pla94], [WU98]. De plus en plus de méthodes comparent à la fois les paramètres et la trajectoire de la signature.

L'apparition des PDAs et des tablettes PCs relance la recherche sur la reconnaissance du manuscrit; ces ordinateurs personnels s'exploitent exclusivement via un écran tactile assimilable à une tablette. Là encore la signature manuscrite peut être utilisée en remplacement aux classiques mots de passe utilisés aussi bien pour la protection des documents qu'aux autorisations d'accès à certains répertoires confidentiels.

### **But et objectifs visés :**

Dans ce manuscrit nous proposons un système de vérification en ligne fondé sur une comparaison des paramètres dynamiques ainsi qu'une mesure de la ressemblance des formes des signatures. La signature capturée est mise à l'échelle, centrée et superposée à la signature spécimen. On calcule alors la distance entre le spécimen et la signature à tester grâce à une mise en correspondance entre vecteurs caractéristiques. La distance retenue dans ce travail est la

distance de la déviation standard. L'utilisation de cette distance pour l'évaluation des performances du système étudié est traitée.

L'objectif que nous nous sommes fixé est de proposer une nouvelle approche pour la vérification des signatures manuscrites en ligne. L'approche est basée sur la transformée de Fourier.

L'étude qui suit est organisée autour de quatre chapitres. Nous commencerons par présenter au chapitre I, l'organisation générale d'un système biométrique et l'état de l'art dans le domaine de l'identification des signatures manuscrites. Nous aborderons ensuite au chapitre II les différentes méthodes de modélisation des signatures manuscrites, où nous proposerons la méthode retenue pour l'identification des signatures. Les détails de cette démarche seront décrits au chapitre III. Enfin l'évaluation des performances de notre système sera étudiée au chapitre IV. Nous terminerons ce travail par une conclusion et les perspectives qui sous-tendent notre travail.

## **Chapitre I**

# **ORGANISATION GENERALE D'UN SYSTEME BIOMETRIQUE ET ETAT DE L'ART**

## **I.1 INTRODUCTION:**

La biométrie est la reconnaissance d'un individu à base de ses propres caractéristiques qui permettent de le différencier des autres gens. Donc, la reconnaissance biométrique est basée sur ce qui est un individu. Nous allons expliciter les fondements de cette reconnaissance biométrique en se basant sur les travaux effectués récemment. On distingue trois catégories des caractéristiques biométriques :

- Biologiques: comme le sang, la salive, l'urine, l'odeur ou encore l'ADN. Ces méthodes sont difficiles à mettre en œuvre pour une utilisation courante.
- Comportementales: comme la signature, les frappes clavier, la démarche (le mouvement des hanches, des bras et des épaules).
- Morphologiques: comme les empreintes digitales, le visage, l'iris, la rétine ou la forme de la main.

De nombreux systèmes biométriques sont envisagés dans les futurs systèmes en relation avec la sécurité. Particulièrement, un déploiement à large échelle des passeports et des cartes d'identités biométriques pour le contrôle de l'identité est prévu pour les années à venir au niveau mondial. Des délais sont alors imposés par les gouvernements pour la mise en place de ces systèmes.

## **I.2 PROBLEMATIQUE :**

La problématique de la reconnaissance des signatures manuscrites n'est pas nouvelle. De nombreuses études en témoignent, mais elles prouvent aussi que le problème n'est pas totalement résolu. Plusieurs défis rendent la reconnaissance par signature moins performante, que quelques modalités biométriques. Un problème important qui cause la dégradation des performances des systèmes de vérification par signature est l'intra-variabilité de la signature. La signature d'une personne change d'une fois à l'autre et deux signatures ne peuvent pratiquement jamais être identiques à cent pour cent. En plus, une personne peut changer complètement sa signature. Un autre problème qui entraîne des performances médiocres est la facilité d'imitation d'une signature en l'analysant et le pratiquant plusieurs fois. Un imposteur (imitateur) aura ainsi la possibilité de produire une signature très proche d'une signature authentique.

### I.3 ARCHITECTURE D'UN SYSTEME BIOMETRIQUE :

Un système biométrique est essentiellement un système de reconnaissance de formes [JRO08]. Il consiste principalement en deux phases : apprentissage (entraînement) et comparaison. Dans la première phase, plusieurs échantillons (généralement de 3 à 5) de la modalité biométrique d'une personne sont acquis par un dispositif dédié. Afin d'améliorer la qualité de la modalité, des prétraitements sont généralement appliqués sur les données acquises. Ensuite, des paramètres sont extraits de chaque échantillon capturé et sont stockés dans une base de données. Dans l'étape de comparaison (reconnaissance), la même modalité de la personne, que l'on veut soit authentifier (vérifier), soit identifier, est capturée et les mêmes paramètres sont extraits. Pour l'identification, les paramètres sont comparés à tous les enregistrements de la base afin de déterminer s'il y en a un qui correspond. En cas de vérification, les paramètres sont comparés uniquement à ceux de la personne proclamée et la décision d'accepter ou de refuser l'authentification est annoncée par le système. La figure 1 montre le schéma bloc d'un système biométrique.

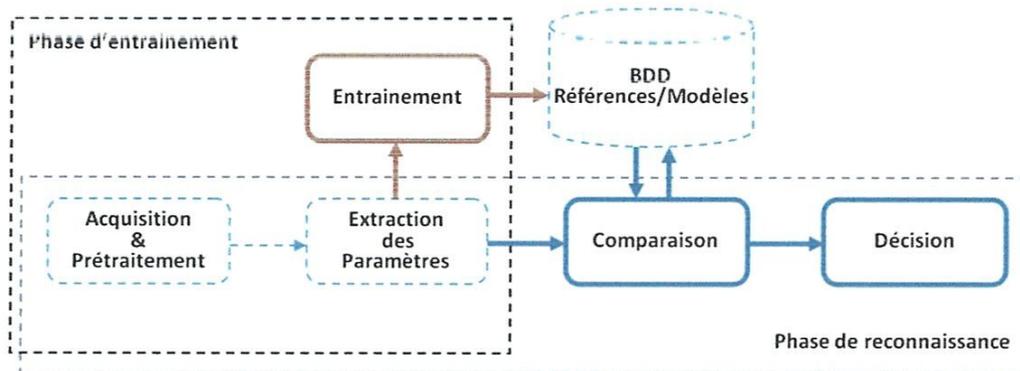


Figure I.1 : Schéma bloc d'un système de reconnaissance biométrique.

#### I.3.1 Cas d'un système de vérification de la signature en ligne :

Le schéma bloc d'un système de vérification à base de signature est le même que celui d'un système biométrique présenté dans la figure I.1. Pour illustrer le fonctionnement du système, nous avons séparé les étapes d'apprentissage (entraînement) schématisées dans la figure I.2 et celle de vérification dans la figure I.3. Dans la suite, nous allons explorer les principaux axes de recherche traités dans la littérature des systèmes de vérification de la signature en ligne.

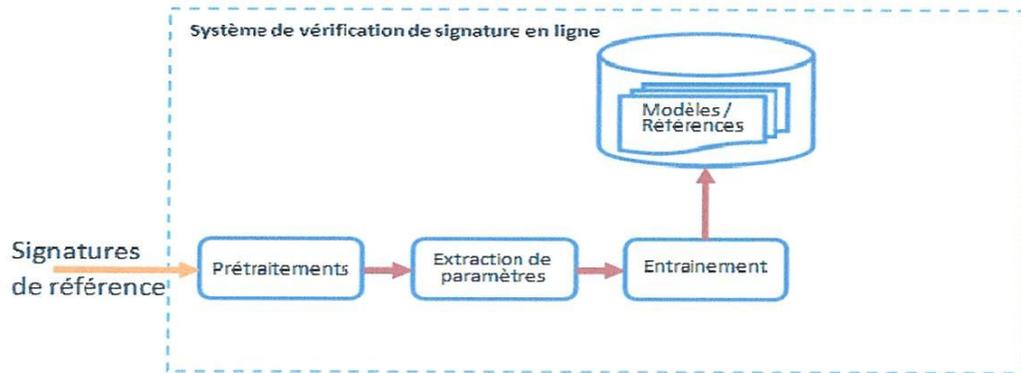


Figure I.2 : Phase d'entraînement d'un système biométrique à base de signature en ligne.

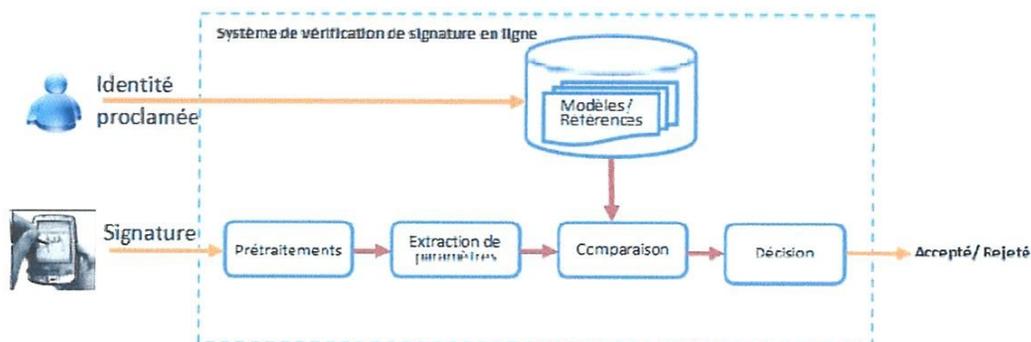


Figure I.3 : Phase de vérification dans un système biométrique à base de signature en ligne.

### I.3.2 Acquisition de la signature en ligne :

Différents dispositifs, disponibles sur le marché, rendent possible l'acquisition de la signature en ligne. Les tablettes graphiques, les tablettes PC, les PDA (Personal Digital Assistant) et les écrans tactiles (figure I.4) sont des exemples de ces dispositifs.



Figure I.4 : Dispositifs d'acquisition de la signature en ligne.

La signature est échantillonnée par un capteur à une fréquence fixe. A chaque point de l'échantillonnage, différentes informations sont enregistrés. Ces informations dépendent du type du dispositif d'acquisition de la signature. Pour les PDA et les écrans tactiles, seules les coordonnées x et y, tout au long de la trajectoire du stylo, et le moment d'acquisition sont fournies. Les tablettes PC et les tablettes graphiques permettent d'avoir, en plus de la forme de la signature, la pression à chaque point de la signature ainsi que les angles d'inclinaison du stylo (l'altitude et l'azimut du stylo) à chaque point de la signature. Ces signaux représentent la signature d'une personne d'une manière numérique, et permettent ainsi de réaliser différents traitements sur les signatures.

### **I.3.3 Les prétraitements :**

Le prétraitement est mené dans l'objectif d'éliminer les informations inutiles et d'améliorer la qualité du signal représentant la signature, ce qui améliore les performances du système de vérification. D'après [VAN05], les prétraitements sont indispensables car ils rendent les signaux indépendants du processus d'acquisition et ne préservent que les informations intéressantes pour les analyses ultérieures. Dans le cas de la signature dynamique, plusieurs prétraitements ont été appliqués dans différents travaux de recherche. On trouve le filtrage (transformée de Fourier), l'élimination du bruit (morphologie mathématique), lissage (filtre gaussien) et la normalisation.

#### **I.3.3.1 Filtrage :**

Le processus d'acquisition introduit du bruit sur les signaux de la signature. Afin d'éliminer ce bruit, des filtres sont appliqués aux signaux acquis. Des expériences dans [PLO89] affirment que les basses fréquences sont les plus informantes. En effet, ces expériences montrent que l'application d'une fréquence de coupure de 10Hz conserve plus de 90% de la puissance du signal. L'application d'un filtre permet aussi d'éliminer les oscillations involontaires dans la signature.

#### **I.3.3.2 Ré-échantillonnage :**

La vitesse de l'échantillonnage, souvent 100 points par seconde, introduit la redondance dans la trajectoire de la signature, spécialement quand la vitesse du stylo est lente. Le ré-échantillonnage des signatures est appliqué dans différents systèmes de vérification de signatures en ligne [FKO05, JGC02] dans le but d'éliminer les points redondants, d'avoir un espace équidistant entre les points et de fixer la taille des signatures. En outre, le ré-échantillonnage

permet d'améliorer la représentation de la forme de la signature. Cependant, Le ré-échantillonnage de la signature cause la perte de quelques informations temporelles selon la vitesse de la signature. Un autre problème s'ajoutant au précédent, est que certains systèmes reposent sur des points critiques (comme les points de début et de fin de la signature, les points de lever du stylo, les points d'inclinaison, etc.) afin de différencier deux signatures. Ces points peuvent être supprimés par le ré-échantillonnage.

Pour éviter ces problèmes, les paramètres temporels (comme la vitesse) et les points critiques de la signature sont extraits avant de ré-échantillonner la signature.

### I.3.3 La normalisation :

Consiste à standardiser la signature dans un domaine comme la taille, la position ou la durée. Certaines méthodes de normalisation en position réalisent une transformation sur les signatures de telle manière que le point de départ de la signature soit le même. D'autres alignent le centre des signatures. La taille d'une signature dépend de l'espace offert aux signataires. La normalisation de l'orientation de la signature consiste à faire une rotation des signatures selon un axe de telle sorte qu'elles soient alignées.

### I.3.4 Extraction de paramètres :

La signature en ligne est représentée par un ensemble de signaux numériques en fonction du temps. Comme nous l'avons mentionné, les paramètres obtenus dépendent du type de dispositif d'acquisition de la signature. Les paramètres basiques obtenus directement du dispositif d'acquisition sont : la forme de la signature, représentée par les coordonnées (abscisse  $x$  et ordonnée  $y$ ) du stylo à chaque point d'acquisition, la pression à chaque point, et les angles d'inclinaison du stylo à chaque point (composé de l'altitude et l'azimut, figure I.5).

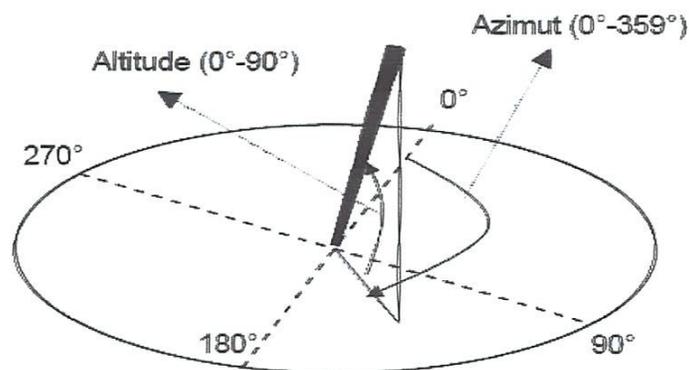


Figure I-5 : Angles d'inclinaison du stylo.

A partir de ces paramètres de base, plusieurs autres paramètres, représentant des caractéristiques dynamiques de la signature qui augmentent la discrimination entre les signatures, peuvent être extraits. Dans la littérature [IPI08], un grand nombre de paramètres sont utilisés dans différents systèmes de vérification de la signature en ligne. On distingue principalement deux grands types de paramètres. Premièrement, les paramètres globaux qui concernent la totalité de la signature comme la taille, la durée, les descripteurs de Fourier, le nombre de levers de stylo, etc. Les auteurs de [FNL05] recensent cent paramètres globaux et les classe selon leur pouvoir de discrimination. Le deuxième type sont les paramètres locaux qui sont extraits à chaque point de la signature comme la vitesse, l'accélération, la tangente, etc. Afin d'exploiter les caractéristiques de chaque type, certains systèmes de vérification de la signature en ligne [FKO05, FNL05] combinent les deux types.

### **I.3.5 Comparaison :**

Le composant comparaison consiste en un classificateur qui compare la signature à vérifier avec la référence de l'utilisateur et donne un score mesurant la similarité ou dissimilarité selon le classificateur utilisé. Plusieurs classificateurs ont été utilisés dans les systèmes de vérification de signature. Il existe deux catégories de classificateurs : à base de distance et à base de modèle. Dans le type à base de distance, la distance entre les vecteurs paramètres de deux signatures est calculée par une fonction. On trouve, entre autres, la distance Euclidienne, la distance Mahalanobis, et la Déformation temporelle dynamique (Dynamic Time Warping, DTW), etc. Dans le type à base de modèle, un modèle statistique est généré à partir des signatures référence pour chaque utilisateur. Les modèles les plus utilisés sont : le Modèle de Markov Caché (Hidden Markov Model, HMM), les réseaux de neurone (Neural Networks), machine à vecteurs de support (Support Vector Machine, SVM), le Modèle de Mixture de Gausse (Gaussian Mixture Model, GMM), etc. Les deux classificateurs qui ont prouvé leur puissance sont le Dynamique Time Warping (DTW) [KYA05] et le Hidden Markov Model (HMM) [CSH07]. Ces deux systèmes sont classés premier et deuxième respectivement dans la première compétition de vérification de signature en ligne SVC04 [YCX04].

### **I.3.6 Décision :**

Ce composant attribue une signature à vérifier selon le score de comparaison à l'une des deux classes client ou imposteur. La décision est prise selon la comparaison du score de la vérification au seuil de décision, fixé dans l'étape d'entraînement du système (voir section suivante). Un traitement peut être effectué sur les scores avant de prendre la décision. Des

exemples de ce traitement incluent : la fusion avec d'autres modalités ou d'autres systèmes de vérification de signature, l'intégration de la qualité des données acquise, la normalisation des scores, la prise en compte des caractéristiques spécifiques à l'utilisateur, etc.

### **I.3.7 Entraînement du système :**

Avant de mettre un système biométrique en fonction, on doit introduire les utilisateurs au système et régler ses paramètres de telle sorte à répondre à l'exigence de l'application n visées. Cette première étape dite entraînement (apprentissage), commence par l'acquisition de plusieurs signatures, qu'on appelle signatures références, pour chaque utilisateur du système. Comme nous l'avons signalé précédemment, la signature présente une intra-classe variabilité : la signature d'une personne change avec le temps (la personne peut changer complètement sa signature ou une modification mineure peut survenir sur la signature avec le temps) et selon le comportement du signataire (l'état moral influence la signature : on ne signe pas de la même façon quand on est à l'aise et quand on est dérangé) , l'environnement (environnement non familier, position du signataire, type du stylo et de la tablette, la taille de la surface de la tablette, etc.). Il faut donc choisir comme échantillon les signatures qui prennent en compte le maximum d'informations possibles et qui représentent le signataire de la meilleure façon. Pour ce faire, les signatures sont recueillies dans des sessions différentes, séparées par un intervalle de temps. En cas de disponibilité de plusieurs signatures pour une personne, des travaux [LWA08, LGD99, SKR97] proposent des techniques de sélection automatique des signatures références les plus appropriées. [LWA08], par exemple, propose deux méthodes de sélection de signatures références : la première est basée sur une classification hiérarchique des signatures selon leur similarité ; tandis que dans la deuxième, les signatures sont triées selon les distances deux à deux, puis les K premières seront choisies comme références. Notons que d'après les résultats obtenus dans les papiers référencés, toutes les méthodes donnent de meilleures performances par rapport à la sélection aléatoire des signatures de référence.

Une fois les signatures références de chaque utilisateur déterminées, elles seront utilisées pour entraîner le système pour qu'il puisse reconnaître l'utilisateur à l'étape de vérification. L'entraînement consiste à générer un modèle statistique ou un objet représentant les signatures de chaque utilisateur. Certaines méthodes gardent toutes les signatures de référence et effectuent une comparaison à chacune d'elle alors que d'autres combinent les signatures références pour avoir une seule représentante avec laquelle la comparaison sera effectuée.

L'entraînement du système permet de fixer les performances selon les exigences de l'application visée par le système. Si l'application demande un niveau de sécurité élevé, on doit minimiser le taux de fausses acceptations. Par contre, si l'application requiert une souplesse d'utilisation, on doit réduire le taux de faux rejets. Le seuil de décision est le paramètre qui contrôle les performances du système de vérification. Afin de déterminer les taux d'erreurs voulus, on fait varier le seuil de décision et on calcule à chaque fois les deux taux d'erreurs, FAR et FRR. Ensuite, le seuil correspondant aux performances adéquates est choisi pour le fonctionnement du système. Le seuil de décision peut être le même pour tous les utilisateurs, ce qui est le cas en général. On peut aussi choisir un seuil de décision par utilisateur, [JGC02] par exemple, adopte cette méthode.

#### **I.4 EVALUATION DES SYSTEMES DE VERIFICATION DE SIGNATURES EN LIGNE :**

Un système de vérification peut commettre deux types d'erreurs. Le système peut accepter une personne dont la signature est falsifiée (fausse acceptation), comme il peut rejeter une personne avec une vraie signature (taux rejet). Les tentatives des imposteurs peuvent être soit qualifiées, dans le cas où l'imposteur fait un effort pour imiter la signature après l'avoir étudiée, soit non qualifiées (à zéro effort) si l'imposteur ne fait pas d'effort pour simuler la signature. Les taux de fausses acceptations et faux rejets sont les indicateurs principaux des performances d'un système biométrique. Les performances des systèmes de vérification sont généralement exprimées en termes d'EER, qui correspond à la valeur où les taux de fausses acceptations et faux rejets sont égaux.

Afin d'évaluer les performances d'un système de vérification de signatures en ligne, les chercheurs ont rendu disponible plusieurs bases de données de signatures en ligne ainsi que des protocoles de tests associés. Aujourd'hui, plusieurs bases sont utilisées tant par les institutions académiques que par les entreprises commerciales pour évaluer leurs systèmes de vérification. Le tableau I.1 récapitule les caractéristiques des bases de données les plus utilisées. Les paramètres acquis diffèrent d'une base à une autre selon le dispositif d'acquisition utilisé. La tablette WACOM est la plus utilisée. Pour une signature, elle permet d'acquérir 100 points par seconde et les paramètres sont les coordonnées horizontale  $x$  et verticale  $y$ , la pression  $p$ , les angles d'inclinaison du stylo : azimut ( $az$ ) et altitude ( $al$ ) et le temps d'échantillonnage ( $t$ ). Dans la base SVC04, le paramètre contact qui indique si le stylo est posé sur la tablette (0) ou lever (1) est aussi enregistré.

La base Bio-Secure comporte le plus grand nombre d'utilisateurs. Sa deuxième partie DS3 est la seule base disponible pour les dispositifs mobiles. Cependant, la Base MCYT reste la plus répandue et la plus utilisée dans l'évaluation des systèmes de signatures en ligne. Le nombre de signatures authentiques et imitations par utilisateur diffère d'une base à l'autre, et même dans la même base pour certaines. Le protocole d'acquisition des imitations est aussi différent d'une base à l'autre. Ainsi on trouve des imitations de différents niveaux. Dans la base PHILIPS, trois types d'imitations :

- (1) l'imitateur observe et analyse le processus dynamique de signatures en cours d'écriture,
- (2) l'imitateur analyse l'image statique de la signature et la pratique plusieurs fois et
- (3) l'imitateur est un expert professionnel dans l'analyse de l'écriture.

La base SUSIG Visual contient des imitations hautement qualifiées, grâce à une simulation visuelle de la signature qui facilite à l'imitateur l'analyse des propriétés dynamiques de la signature. Afin de collecter le maximum d'intra-variabilité d'une signature, l'acquisition est souvent réalisée en plusieurs sessions séparées par un intervalle de temps plus ou moins long.

La dernière colonne du tableau I.1 indique les disponibilités des bases. Certaines sont téléchargeables gratuitement alors que d'autres sont payantes, avec différents prix suivant l'utilisation académique, recherche ou commerciale.

Base de données	Réf.	Paramètres	Utilisateurs	authentique	imitations	sessions	dispositif	Distribution
PHILIPS	[DAV98]	x,y,p, angles xOz et yOz	51	30	3 types variables	3-5	PHILIPS tablette	/
BIOMET	[GBC03]	x,y,p,az,al	84	15	12	2 (5 mois)	WACOM Intuos2	Payante
SVC2004	[YCX04]	x,y,p,az,al,t, contacte	40	20	20	2 (1 semaine)	WACOM Intuos	Gratuite
MCYT	[OFS03]	x,y,p,az,al	330	25	25	5	WACOM Intuos A6	Gratuite
BioSecure DS2	[OFA10]	x,y,p,az,al	600	30	20	2	WACOM Intuos3 A6.	Payante
BioSecure DS3	[OFA10]	x,y,t	700	30	20	2	PDA HP iPAQ hx2,790	Payante
SUSIG Visual	[KYA09]	x,y,p,t	100	20	10	2 (1 semaine)	Interlink ElectronicsePad-ink	Gratuite
SUSIG Blind	[KYA09]	x,y,p,t	100	8/10 (30/70)	10	1	Wacom's Graphire2	Gratuite
NISDCC	[BVF09]	x,y,p,t	43	12*5	31*12*5-15	1	Wacom Intuos2	Participants

Tableau I.1 : Caractéristiques des bases de données de signatures en ligne.

L'évaluation des différents systèmes sur la même base de données et avec le même protocole de test permet de comparer les uns avec les autres. Dans ce but, la première compétition de vérification de signatures SVC'2004 [YCX04] est organisée. SVC'2004 se composait de deux tâches distinctes de vérification de signatures: dans la première tâche, seules les coordonnées du stylo sont utilisées; dans la deuxième tâche, en plus des coordonnées, la pression et l'orientation (azimut et altitude) ont été inclus. Différents systèmes de l'industrie et de la recherche ont participé à la compétition (15 systèmes pour la première tâche, et 12 pour la seconde). Pour les deux tâches et lors des tests sur les imitations qualifiées, le système [KYA05] à base de la déformation temporelle dynamique DTW donne la plus faible valeur EER (EER = 2,84% dans la tâche 1 et EER = 2,89% dans la tâche 2). Le système à base de HMM [CSH07] suit avec un EER de 6% dans la tâche 1 et 5% dans la tâche 2. Pour les tests avec des imitations aléatoires, un système à base de HMM prend la première place avec EER = 2,12% dans la tâche 1 et EER = 1,70% dans la tâche 2.

Après le succès de SVC'2004, d'autres compétitions similaires, comme ICDAR'09[BVF09], ont été organisées.

Le projet Bio-Secure1 [GHL09] fournit un Framework pour l'évaluation des systèmes de la signature manuscrite en ligne. Il est composé de deux systèmes de référence open source. Il inclut les deux bases de données BIOMET et MCYT-100 et les protocoles de tests associés. Ce système facilite la comparaison d'un système de vérification de signatures en ligne. En effet, il suffit de suivre la documentation pour intégrer le nouveau système dans le Framework et le tester.

## **CONCLUSION :**

Après avoir évoqué, les principales étapes d'un système de reconnaissance de formes, dédiées à la reconnaissance de signatures manuscrites en ligne, l'état de l'art, et les difficultés rencontrées lors de la réalisation de ce système, nous exposons dans la suite de ce travail, les principales méthodes de modélisation des signatures manuscrites en ligne rencontrées dans la littérature, et le modèle retenu pour notre travail. Le chapitre II de ce mémoire décrit l'essentiel de ces méthodes.

## **Chapitre II**

# **TECHNIQUES DE MODELISATION DES SIGNATURES MANUSCRITES**

## II.1 INTRODUCTION :

L'étude de tracés manuscrits, signature et écriture, demande de bien connaître le processus qui les a engendrés .Ce processus complexe englobe :

- la représentation ou codage, c'est-à-dire la manière dont l'information de génération de tracés manuscrits est mémorisée dans le cerveau
- la génération, qui peut se diviser en une commande d'excitation (la commande nerveuse) et l'exécution de cette commande (la contraction des muscles).

Les systèmes testés sont tous des systèmes linéaires d'ordre supérieur ou égal à 2. Différents articles proposent un modèle basé sur la génération de profils de vitesse gaussiens.

Après une présentation de ce modèle et des concepts qui ont permis d'en établir les fondements, la manière dont les tracés manuscrits ont été traités est décrite, ainsi que les différentes étapes du traitement. Cette section permet de résumer succinctement certains modèles proposés dans la littérature

## II.2 MODELE DE FRANCK LECLERC, GUY LORETTE & PLAMONNDON [FRG92] :

Dans le modèle de génération de tracé proposé par ces auteurs, l'information permettant de générer un tracé manuscrit, écriture ou signature, est supposée être contenue dans le cerveau. Cette information de haut niveau peut être représentée sous la forme d'un programme moteur, décrit dans certains articles, qui sont transmis par les nerfs sous la forme d'ondes dépolarisantes.

Après transfert synaptique de la commande nerveuse par l'intermédiaire de la plaque motrice, les muscles sont activés pour produire le déplacement du crayon. La figure 1 résume le processus de génération de tracés manuscrits. Pour la modélisation du système générateur de tracés manuscrits, le programme moteur est pris comme entrée, et le déplacement du crayon comme sortie du système à modéliser. La figure 2 met en évidence ce modèle. Les paramètres du modèle seront calculés à partir des déplacements du crayon (signature ou écriture) enregistrés à partir d'une tablette à digitaliser.

Les figures II.1 et II.2 montrent que le système possède une seule entrée : le programme moteur, ce qui veut dire qu'aucun contrôle extérieur n'est possible. Le modèle suppose donc un mouvement balistique rapide, programmé d'avance.

Il n'y a aucune boucle de contre-réaction psycho-visuelle prenant en compte des informations de la sortie du système pour modifier son entrée. Pour des tracés manuscrits tels que les signatures, l'hypothèse d'un mouvement balistique rapide, programmé d'avance, semble tout à fait justifiée. Chaque personne signe de façon quasi automatique. La signature est un mouvement entièrement appris et stéréotypé qui se répète à volonté.

Avec une représentation des tracés manuscrits dans un repère cartésien, il n'est pas possible d'obtenir une invariance du codage des signaux correspondant après une rotation ou une translation. Il convient donc de trouver un codage insensible aux différentes transformations. La géométrie différentielle permet d'obtenir cette invariance. L'équation d'une courbe reste la même quelle que soit la transformation subie (translation et/ou rotation), seules changent les conditions initiales : point de début du tracé et direction du premier déplacement. Une courbe plane peut être entièrement caractérisée par son équation intrinsèque qui exprime sa courbure (variation angulaire) le long de l'abscisse curviligne.

Ceci s'exprime par l'équation (1)

$$c(\sigma) = \frac{d\theta(\sigma)}{d\sigma} \quad (II.1)$$

$\theta(\sigma)$  : angle le long de l'abscisse curviligne

$c(\sigma)$  : trajectoire courbe

$\sigma$  : abscisse curviligne

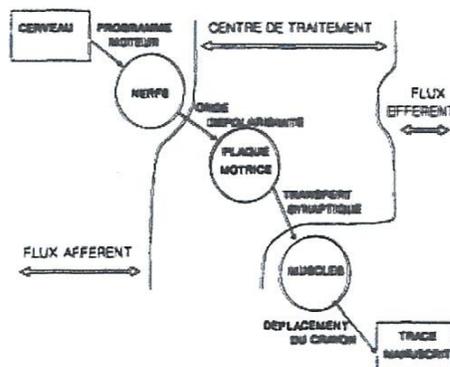


Figure II.1 : Processus de génération de tracés manuscrit.

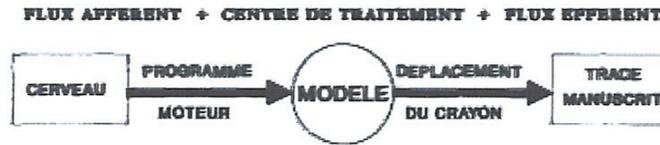


Figure II.2 : Modélisation du processus de génération.

Ainsi, pour réaliser un tracé manuscrit, le système doit contrôler le déplacement par rapport à un point d'origine, traduit mathématiquement par l'abscisse curviligne, et la direction dans laquelle s'effectue le mouvement caractérisée par la courbure le long de l'abscisse curviligne. Pour cela, le modèle doit posséder deux parties distinctes : un générateur de mouvements curvilignes et un générateur de mouvements angulaires, chacun d'eux étant commandé par un programme moteur. Les entrées des deux générateurs de mouvements sont constituées par des créneaux rectangulaires d'amplitude et de durée variables. La figure II.3 résume les différents aspects du modèle utilisé.

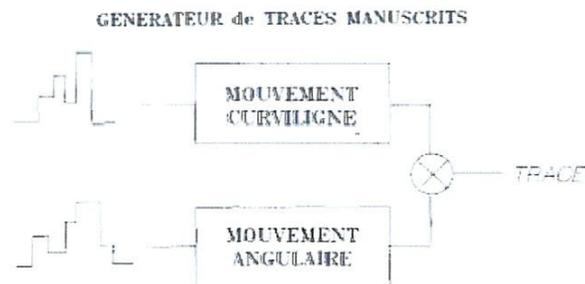


Figure II.3 : Schéma global du modèle.

En géométrie différentielle, à partir des coordonnées obtenues dans un repère orthonormé, la vitesse curviligne s'exprime par l'équation (II.2).

$$v(t) = \sqrt{v_x(t)^2 + v_y(t)^2} \quad (II.2)$$

avec :

$$v_x(t) = \frac{dx(t)}{dt} \quad v_y(t) = \frac{dy(t)}{dt}$$

$x(t)$  : abscisse du tracé manuscrit.

$y(t)$  : ordonnée du tracé manuscrit.

et la vitesse angulaire s'exprime par l'équation (II.3) :

$$\Omega(t) = \frac{d\theta(t)}{dt} \quad (II.3)$$

avec :

$$\theta(t) = \arctan\left(\frac{dy(t)}{dx(t)}\right)$$

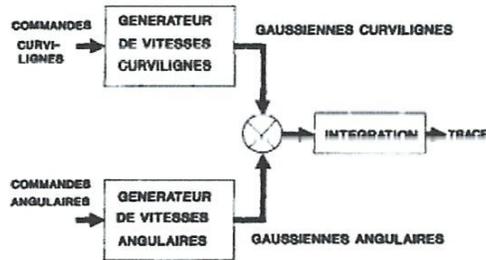


Figure II.4 : Schéma général du modèle utilisé.

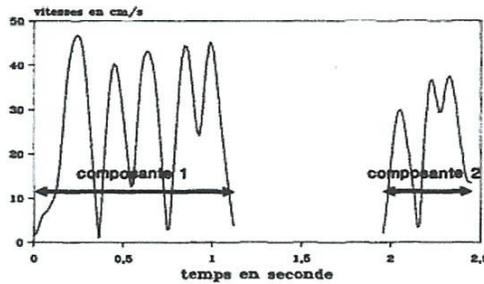


Figure II.5 : Vitesse curviligne d'un tracé manuscrit.

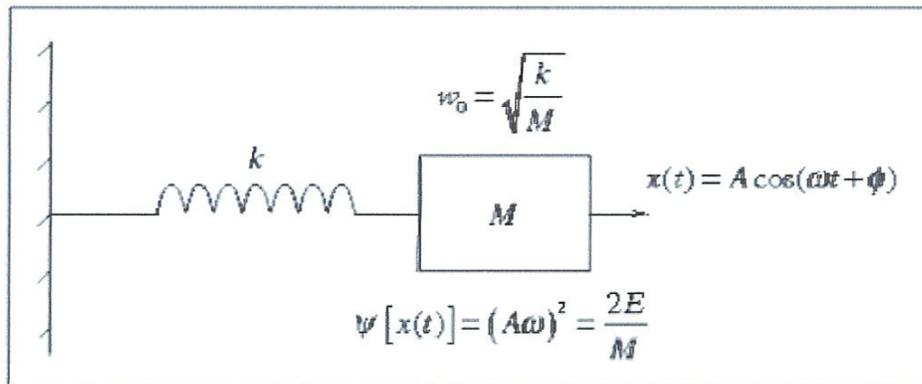
La vitesse angulaire  $\Omega(t)$  peut s'exprimer par l'équation (II.4) :

$$\Omega(t) = \frac{\frac{d^2y(t)}{dt^2} \cdot \frac{dx(t)}{dt} - \frac{d^2x(t)}{dt^2} \cdot \frac{dy(t)}{dt}}{v^2(t)} \quad (II.4)$$

### II.3 MODELISATION BASEE SUR L'OPERATEUR DE TEAGER-KAISER

$\Psi_{\mathbb{R}}$  [JCC05] :

Le modèle considéré est basé sur un modèle de représentation non linéaire. Au lieu de considérer le mouvement du tracé manuscrit décrit précédemment par les équations représentées précédemment (II.1- II.4), Teager-Kaiser ont associé à ces équations leur énergie en se basant sur le modèle harmonique (masse-ressort) (figure II.6). Le modèle considéré est décrit par les équations suivantes :



**Figure II.6 :** L'opérateur d'énergie de Teager Kaiser dépiste l'énergie fournie par la source sinusoïdale (normalisée par unité de masse-1) qui engendre le signal  $x(t)$ .

#### II.3.1 Description du modèle :

En appliquant le principe fondamental de la dynamique au système formé d'une masse  $m$  et d'un ressort de raideur  $k$  (figure II.6), nous obtenons l'équation différentielle suivante :

$$\frac{d^2x(t)}{dt^2} + \frac{k}{m}x(t) = 0 \quad (II.5)$$

La solution de cette équation est une oscillation périodique donnée par :

$$x(t) = A \cos(\Omega t + \phi) \quad (II.6)$$

Où  $A$  est l'amplitude des oscillations,  $\Omega_0 = \sqrt{\frac{k}{m}}$  la pulsation propre d'oscillation (fréquence angulaire), et  $\phi$  la phase initiale.

L'énergie totale  $E_x$  du système masse-ressort (oscillateur harmonique) est la somme de l'énergie potentielle du ressort et de l'énergie cinétique de la masse :

$$E_x = \frac{1}{2}kx^2 + \frac{1}{2}mv^2 \quad (II.7)$$

En remplaçant la vitesse  $v = dx/dt$  et la position  $x = A \cos(\Omega t + \phi)$  dans l'équation (II.7) nous obtenons :

$$E_x = \frac{1}{2}m\Omega^2 A^2 \quad (II.8)$$

$$E_x \simeq A^2 \Omega^2 \quad (II.9)$$

Ainsi, l'énergie de l'oscillateur est proportionnelle au carré de l'amplitude et au carré de la pulsation des oscillations. La valeur de  $E_x$  représente l'énergie du système masse-ressort nécessaire pour générer le signal sinusoïdal  $x(t)$ .

La forme suivante :

$$x_n = A \cos(\Omega n + \phi) \quad (II.10)$$

Où  $\Omega = 2\pi f/Fe$  est la pulsation numérique,  $f$  la fréquence et  $Fe$  la fréquence d'échantillonnage. Nous supposons que l'échantillonnage est régulier. Notons que l'équation II.10 a trois paramètres ou inconnues ( $A, \Omega, \phi$ ) qui peuvent être calculées à partir de trois échantillons du signal. Pour la résolution nous utilisons les échantillons  $x_{n-1}, x_n$ , et  $x_{n+1}$  :

$$x_n = A \cos(\Omega n + \phi) \quad (II.11)$$

$$x_{n+1} = A \cos((n+1)\Omega + \phi) \quad (II.12)$$

$$x_{n-1} = A \cos((n-1)\Omega + \phi) \quad (II.13)$$

En utilisant les relations trigonométriques :

$$\cos(\alpha + \beta) \cos(\alpha - \beta) = \frac{1}{2} [\cos(2\alpha) + \cos(2\beta)] \quad (II.14)$$

$$\cos(2\alpha) = 1 - 2 \sin^2 \alpha \quad (II.15)$$

Nous obtenons :

$$x_{n+1}x_{n-1} = A^2 \cos[(n+1)\Omega + \phi] \cos[(n-1)\Omega + \phi] \quad (II.16)$$

$$= \frac{A^2}{2} [2 \cos^2(\Omega n + \phi) + 1 - 2 \sin^2 \Omega] \quad (II.17)$$

$$= A^2 \cos^2(\Omega n + \phi) - A^2 \sin^2 \Omega \quad (II.18)$$

$$= x_n^2 - A^2 \sin^2 \Omega \quad (II.19)$$

$\Rightarrow$

$$A^2 \sin^2(\Omega) = x_n^2 - x_{n+1}x_{n-1} \quad (II.20)$$

Pour des faibles valeurs de  $\Omega$  nous avons  $\sin(\Omega) \approx \Omega$ . Si nous limitons les valeurs de  $\Omega$  à  $\Omega < \pi/4 = 0.7854$ , c'est à dire  $f/Fe < 1/8$ , l'erreur commise en approximant  $\sin(\Omega)$  par  $\Omega$  est de  $\Omega / \sin \Omega = \frac{\pi/4}{\sqrt{2}/2} \approx 11\%$  et ce qui permet d'écrire l'expression II.20 comme suit :

$$A^2 \Omega^2 \approx x_n^2 - x_{n+1}x_{n-1} \quad (II.21)$$

L'équation (II.21) donne une mesure de l'énergie du signal  $x(t)$  dont la fréquence d'échantillonnage est telle que  $Fe > 8f$  (c'est à dire au moins deux points dans chaque quart du cycle de l'oscillation sinusoïdale). L'algorithme nécessite alors une fréquence d'échantillonnage à peu près de 8 fois supérieure de la composante la plus élevée du signal à traiter. Ainsi, la relation (II.21) est vue comme un algorithme de calcul de l'énergie d'un signal mono-composante :

$$E_n = x_n^2 - x_{n+1}x_{n-1} = A^2 \sin^2(\Omega) \approx A^2 \Omega^2 \quad (II.22)$$

Où  $E_n$  est la sortie de l'algorithme et  $x_n$  est le signal à traiter. Notons que l'équation (II.22) est indépendante de la phase initiale de l'oscillation. Notons également que cet algorithme est symétrique dans le sens où le changement de  $x_n$  en  $x_{-n}$  ou de  $n$  en  $-n$  (en inversant le signal en temps) ne change pas la valeur de l'énergie du signal. L'algorithme est facile à mettre en œuvre et rapide car ne nécessite que trois échantillons.

Bien qu'historiquement, la version discrète de l'opérateur de Teager-Kaiser ait été définie avant la version continue, il est pourtant plus naturel de débiter par la définition de l'opérateur continu en fonction de  $t$ . L'opérateur d'énergie de Teager-Kaiser (TKEO)  $\Psi_{\mathbb{R}}[x(t)]$  d'un signal  $x(t)$  à valeurs réelles est défini dans le domaine continu par :

$$\Psi_{\mathbb{R}}[x(t)] = \left(\frac{dx(t)}{dt}\right)^2 - x(t) \frac{d^2x(t)}{dt^2} \quad (II.23)$$

ou encore :

$$\Psi_{\mathbb{R}}(x) = \dot{x}^2 - x\ddot{x} \quad (II.24)$$

avec  $x \equiv x(t)$ ,  $\dot{x} \equiv dx(t)/dt$  et  $\ddot{x} \equiv d^2x(t)/dt^2$  et en considérant la variable indépendante  $t$  comme implicite.

Dans le cas discret il vaut :

$$\Psi_{\mathbb{R}d} = \frac{x^2(n-1) + x(n)x(n-2)}{T_e^2} \quad (II.25)$$

En ignorant le décalage d'un échantillon et la normalisation par le facteur  $T_e^2$ , nous obtenons :

$$\Psi_{\mathbb{R}}[x(n)] = x^2(n) - x(n-1)x(n+1) \quad (II.26)$$

### II.3.2 Applications et limites de l'opérateur $\Psi_{\mathbb{R}}$ :

Actuellement, l'opérateur TKEO est un outil familier au sein de la communauté du traitement du signal et permet de définir une grandeur locale ayant la dimension d'une énergie pour les signaux oscillants. il peut être utilisé pour décrire les tracés manuscrit.

### II.4 MODELE D'HOLLERBACH [JMH81]:

Selon Hollerbach, l'écriture peut être vue comme la combinaison de deux oscillateurs dans le plan, l'un oscillant selon la direction verticale, l'autre selon la direction horizontale. À cela s'ajoute un mouvement constant :

$$\dot{x} = a \sin(\omega_x t + \phi_x) + c \quad (II.27)$$

$$\dot{y} = b \sin(\omega_y t + \phi_y) \quad (II.28)$$

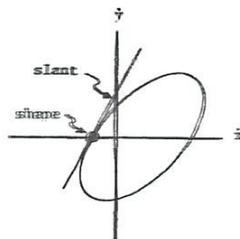
où  $a$  (resp.  $b$ ) est l'amplitude de la vitesse horizontale (resp. verticale),  $\omega_x$ ,  $\omega_y$ ,  $\phi_x$  et  $\phi_y$  sont respectivement les fréquences et les phases associées à ces directions.  $t$  représente le temps et  $c$  l'amplitude du déplacement constant vers la droite ( $c > 0$ ).

Afin de mieux appréhender les différents paramètres du modèle, il peut être intéressant de regarder le tracé de ces équations dans l'espace des vitesses. La figure II.7 montre le résultat des équations (II.27), (II.28), à paramètres constants dans l'espace des vitesses. Elles prennent la forme d'une ellipse dont l'excentricité et l'inclinaison dépendent des paramètres  $a$ ,  $b$  et

$\phi = \phi_x - \phi_y$ .  $\phi$  correspond à la phase relative entre les deux oscillateurs ; elle fait sens si  $\omega_x$  et  $\omega_y$  sont très proches. Le sens de parcours de l'ellipse est déterminé par le signe de  $\phi$  et la vitesse de parcours par  $\omega$  (en supposant  $\omega = \omega_x = \omega_y$ ).

Les paramètres du modèle (i.e.  $a$ ,  $b$ ,  $\omega_x$ ,  $\omega_y$ ,  $\phi_x$  et  $\phi_y$ ) sont constants par morceaux : il changent au moment où la vitesse verticale s'annule. Hollerbach a fait ce choix car il a remarqué que la forme des lettres semblait déterminée par la valeur de la vitesse horizontale au point d'annulation de la vitesse verticale, exprimée par :

$$\Psi = \dot{x}(t_{y_0}) = c - a \sin \phi \quad (II.29)$$



**Figure II.7 :** Représentation visuelle des équations d'Hollerbach, son excentricité ainsi que son orientation dépendent de  $a$ ,  $b$  et  $\phi = \phi_x - \phi_y$  dans le cas où les oscillateurs ont la même fréquence.

Image empruntée à Hollerbach

## II.5 MODELISATION PAR LA TRANSFORMEE DE FOURIER : (METHODE PROPOSEE PAR YANIKOGLU) [YKH09] :

La transformée de Fourier (TF) est une opération qui transforme une fonction d'un domaine à un autre. Dans le traitement du signal, le domaine de la fonction d'origine est généralement le temps et le domaine de la nouvelle fonction est généralement le domaine des fréquences, et la nouvelle fonction est une représentation dans le domaine fréquentiel de la fonction d'origine. La TF permet de montrer les fréquences composant un signal. En effet, la

transformée de Fourier décompose le signal en plusieurs signaux de différentes fréquences. La transformée de Fourier est une transformation séparable et symétrique. Elle montre également de bonnes caractéristiques de dé-corrélation d'énergie et de concentration. Cependant, la TF est une transformation complexe (le noyau est une exponentielle qui génère des coefficients complexes) et nécessite donc un traitement pour les paramètres phase et amplitude. En outre, la périodicité implicite de la TF provoque les discontinuités de frontière qui ont comme conséquence l'apparition des coupures et donc elle ne permet pas un codage précis.

## **II.6 MODELISATION BASEE SUR L'APPROCHE DE LA DECOMPOSITION EMPIRIQUE MODALE ET DE LA TRANSFORMEE DE HUANG HILBERT (THH) :**

### **II.6.1 Décomposition modale empirique EMD [HAF16] :**

La décomposition en mode empirique ou EMD (Empirical Mode Decomposition) est une technique flexible de décomposition du signal introduite par Huang et al [NEH71]

L'idée de base est de décomposer un signal en une somme de composantes noté IMF (Intrinsic Mode Function) plus un résidu. L'addition de toutes les IMFs avec le résidu reconstruit le signal original sans perte d'information (équation II.31). L'extraction des IMFs est non-linéaire, mais leur recombinaison pour la reconstruction du signal est linéaire.

L'EMD décompose un signal d'une façon auto-adaptative en une somme de composantes oscillantes. Le principe de l'EMD est défini par un algorithme. Dans la littérature les études menées sur l'EMD se font dans le cadre de simulations numériques avec des conditions contrôlées. Malgré le manque de formalisme mathématique pour décrire l'EMD, elle a montré des capacités d'analyse de signaux très intéressantes. Cette méthode s'applique dans des domaines tels que l'océanographie, la sismologie [NEH71], la biologie [NEH04] et plus généralement à des signaux non-stationnaires.

#### **II.6.1.1 Principe de la décomposition :**

Le principe de l'EMD est de décomposer un signal en une somme restreinte de composantes, décrivant localement le signal comme une succession de contributions d'oscillations rapides (hautes fréquences) sur des oscillations plus lentes (basses fréquences)

(équation II.30). Nous illustrons dans la (figure II.8) un exemple de décomposition d'un signal  $x(t)$  composé d'une oscillation rapide et d'oscillation lente.

$$x(t) = m(t) + d(t) \quad (II.30)$$

Signal = oscillation(a) + oscillation(b)

La reconstruction du signal  $x(t)$  peut être réalisée en sommant les deux courbes oscillations rapides et oscillation lente (figure II.8).

Dans le but d'obtenir une décomposition du signal en plusieurs composantes, il suffit de considérer le signal à l'échelle de ses oscillations locales, de soustraire l'oscillation la plus rapide du signal et de réitérer le processus sur le résidu de manière à écrire le signal  $x(t)$  comme une combinaison finie d'oscillations

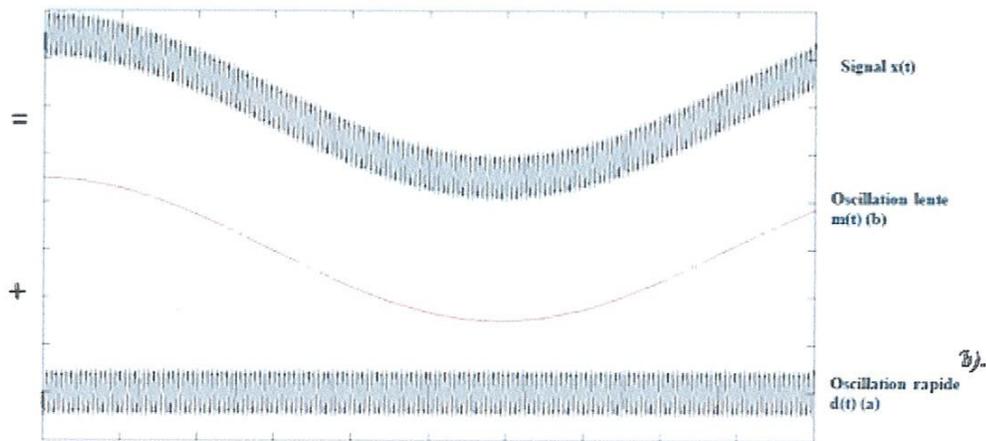


Figure II.8 : Principe de décomposition de l'EMD

$$x(t) = \sum IMF(t) + r(t) \quad (II.31)$$

ou  $IMF(t)$  est la  $k^{ième}$  oscillation  $r(t)$  le résidu de la décomposition et  $N$  le nombre d'IMFs (ou modes). Le résidu  $r(t)$  représente la tendance globale du signal original  $x(t)$   $r(t)$  n'est pas une quantité négligeable.

La décomposition EMD est entièrement pilotée par les données du signal (ou data driven) et adaptée à celui-ci. C'est pourquoi le nom de fonction modale intrinsèque (IMF) a été retenu. Le signal original peut être reconstruit en additionnant toutes les IMFs et le résidu. La

reconstruction est donc réalisée d'une manière linéaire et sans perte ou déformation de l'information initiale [N.E.H04].

## II.6.2 Transformée d'HILBERT :

L'analyse d'un phénomène physique passe souvent par l'acquisition d'un signal monodimensionnel fonction du temps. Même si la représentation temporelle d'un signal est la plus naturelle, elle ne permet pas toujours une analyse complète. La représentation fréquentielle obtenue par transformation de Fourier (TF) est la base de l'analyse spectrale qui fournit des informations auxquelles nous n'avons pas accès temporellement [BBO92], [PFL93].

Cependant, ces deux analyses, prises séparément, sont insuffisantes pour une large classe de signaux (signaux non-stationnaire, signaux multi-composantes).

Bien que toute l'information utile à l'analyse du signal soit contenue dans chacune de ces deux descriptions (temporelle et fréquentielle), celles-ci ne présentent que partiellement les caractéristiques d'un signal. On imagine alors qu'une approche mixte combinant simultanément les deux aspects temporel et fréquentiel devrait permettre une meilleure compréhension des différentes composantes présentes dans le signal.

### II.6.2.1 Signal multi-composantes et mono-composante :

Les notions des signaux mono/multi composantes passent par la définition des signaux dits à modulation d'amplitude et/ou modulation de fréquence (AM-FM). Un signal AMFM multi-composantes  $x(t)$  peut être représenté par le modèle suivant:

$$x(t) = \sum_{k=1}^N x_k(t) + n(t) \quad (II.32)$$

avec  $n(t)$  un bruit additif (représentant toutes les composantes indésirables), et  $N$  le nombre de composantes. Dans cette représentation,  $x_k(t)$  est un signal mono-composante caractérisé par une amplitude instantanée (AI)  $a_k(t)$  et une fréquence instantanée (FI)  $V_{ik}(\tau)$  telle que la forme complexe  $Z_k(t)$  associée à  $x_k(t)$  peut s'écrire :

$$Z_k(t) = a_k(t)e^{j\varphi_k(t)} \quad (II.33)$$

ou

$$\varphi_k(t) = \int_{-\infty}^t V_{ik}(\tau) d\tau \quad (II.34)$$

Dans ce modèle, si  $N=1$ , le signal est dit mono-composante et si  $N \geq 2$ , le signal est dit multi-composantes. Dans le cas le plus général, un signal  $x(t)$  peut être considéré comme un signal AM-FM multi-composantes, ou du moins comme la contribution de plusieurs signaux AM-FM. Pour interpréter et caractériser le signal, il est nécessaire de pouvoir estimer ces grandeurs FI et AI.

### II.6.2.2 Estimation de la fréquence instantanée :

Nous avons illustré dans le paragraphe précédent le concept de la FI mais pas les méthodes permettant son estimation. Pour cela, il est nécessaire d'associer à un signal réel  $x(t)$  un signal complexe  $z(t)$ , opération que l'on note  $z(t) = S [x(t)]$ .

#### II.6.2.2.1 Définition :

Le signal analytique SA  $Z_x(t)$  complexe associé au signal réel  $x(t)$  s'exprime sous la forme :

$$Z_x(t) = x(t) + jH\{x(t)\} \quad (II.35)$$

Avec

$$H\{x(t)\} = \frac{1}{\pi} vp \int_{-\infty}^{+\infty} \frac{x(s)}{t-s} ds$$

où H désigne la TH (Transformée d'Hilbert) de  $x(t)$  et 'vp' la valeur principale de l'intégrale au sens de Cauchy.

On note que la partie réelle du SA  $Z_x(t)$  est  $x(t)$  lui-même et que la partie imaginaire est obtenue à partir d'un filtre linéaire H [BBO92].

Les notions d'amplitude et de phase d'un signal réel  $x(t)$  non-stationnaire sont associées à un unique signal complexe  $Z_x(t)$  dont le module et l'argument définissent respectivement les grandeurs AI,  $|Z_x(t)|$  et la phase instantanée,  $\arg Z_x(t) = \varphi(t)$  du signal  $x(t)$ . La FI s'exprime comme la dérivée de la phase instantanée. Nous pouvons alors écrire que l'AI,  $a_x(t)$ , et que la FI,  $V_x(t)$  d'un signal  $x(t)$  sont de la forme :

$$a_x(t) = |Z_x(t)| \quad (II.36)$$

$$V_x(t) = \frac{1}{2\pi} \frac{d}{dt} \arg Z_x(t) \quad (II.37)$$

L'estimation de AI et FI d'un signal par la TH est schématisée dans la figure ci-dessus.

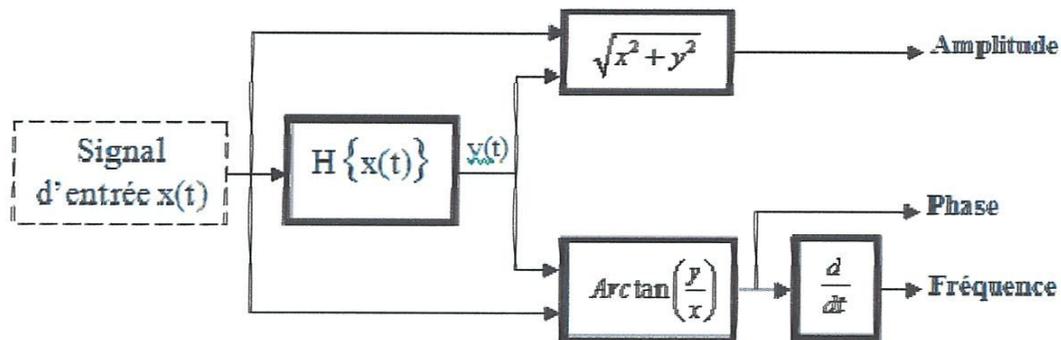


Figure II.9 : Principe d'estimation de la FI via le signal analytique  $z(t)$  de  $x(t)$ .

### II.6.3 Transformation de Huang-Hilbert (THH) :

La THH (Transformation de Huang-Hilbert) est une méthode d'analyse des signaux non stationnaires introduite par Huang et al [NEH71]. Cette méthode est la combinaison de l'EMD et de la transformation de Hilbert pour l'estimation des AI et FI. L'association de ces deux méthodes est appelée la Transformation de Huang-Hilbert (THH).

Nous avons parlé dans la partie précédente de l'intérêt de disposer de représentations combinant simultanément le temps et la fréquence afin d'analyser des signaux non stationnaires.

Cela conduit à représenter ces signaux comme des signaux multi-composant, combinaison linéaire des signaux mono-composante de type AM-FM. La description d'un tel signal (mono-composante) passe alors par l'estimation des deux grandeurs AI et FI.

Partant des limitations énumérées ci-dessus, Huang et al [NEH71] ont proposé une technique qui aborde sous un autre angle la problématique de l'analyse des signaux non-stationnaires avec l'approche de décomposition modale empirique (EMD). Une des motivations du développement de l'EMD est l'estimation de la FI du signal. Et comme nous avons présenté dans le chapitre 2, l'EMD repose sur une décomposition adaptée en décrivant localement le signal comme une succession de contributions d'oscillations rapides (hautes fréquences) sur des oscillations plus lentes (basses fréquences). En utilisant la décomposition EMD, un signal  $x(t)$  peut alors s'écrire :

$$x(t) = \sum_{j=1}^N IMF_j(t) + r(t), \quad n \in N^* \quad (II.38)$$

Et comme mentionné dans le chapitre précédant chaque IMF doit respecter les critères suivants :

a) de moyenne nulle,

- b) dont les nombres d'extrema et de passages à zéro diffèrent au plus de un (en d'autres termes, cela signifie qu'entre un minimum et un maximum successif, une IMF passe par zéro),
- c) qui suit une loi de modulation en amplitude et en fréquence (comportement oscillant) naturellement de type mono-composante.

Le point (a) est lié à la contrainte d'estimation de la FI d'un signal mono-composante avec la TH. Le point (b) repose sur la définition bande étroite d'un signal mono-composante. La dernière remarque (c) est plus une conséquence due au processus d'estimation de l'IMF. Ainsi, même si la notion d'IMF repose essentiellement sur un concept, celle-ci présente des caractéristiques très strictes et peut être formalisée par le modèle suivant :

$$IMF_i(t) = a_i(t)e^{i\varphi_i(t)} \tag{II.39}$$

avec  $a_i(t)$  l'amplitude de l'enveloppe de la  $i^{eme}$  IMF et  $\varphi_i(t)$  sa phase.

Pour résumer, l'analyse par THH d'un signal multi-composantes comporte essentiellement deux étapes : la décomposition temporelle (EMD) et l'estimation des FI et AI par la Transformée de Hilbert (TH) (voir figure ci-dessous) L'EMD combinée avec la transformée de Hilbert permet d'obtenir une nouvelle représentation temps-fréquence des signaux non stationnaires et multi-composantes.

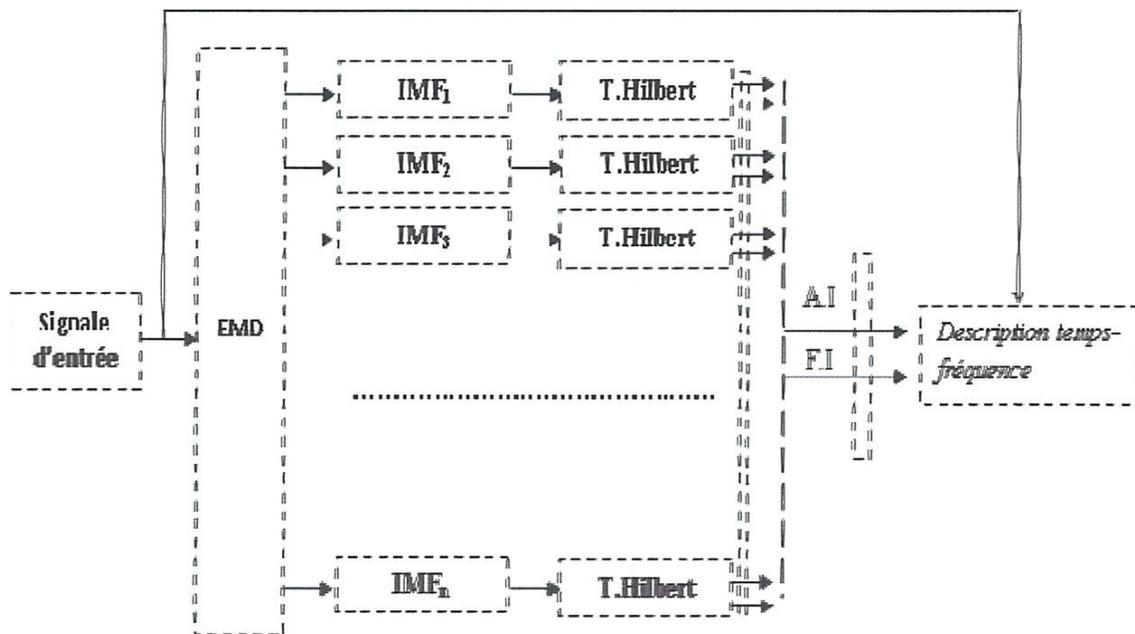


Figure II.10 : Organigramme de la THH.

### **II.7 Modèle retenu :**

Dans la suite de notre travail, nous avons retenu le modèle basé sur la transformée de Fourier. La description de ce modèle sera décrite dans le chapitre III.

### **CONCLUSION**

D'après l'étude proposée pour la modélisation des signatures manuscrites, il ressort, deux techniques de modélisation : la première basée sur l'information temporelle, la seconde sur l'information fréquentielle. Nous avons retenu, pour notre cas, les modèles basés sur les méthodes de transformations fréquentielles en particulier la transformée de Fourier. Les détails de cette méthode seront décrits dans le chapitre III qui suit.

## **Chapitre III**

# **MECANISMES DE FONCTIONNEMENT DU SYSTEME ADOPTE**

### III.1 INTRODUCTION :

Dans le cas d'un signal discret, on parle de Transformée de Fourier discrète (TFD). Il existe des algorithmes très rapides permettant de calculer la TFD. Dans notre cas (signature en ligne figure : III.1), afin d'avoir les paramètres fréquentiels d'une signature donnée, nous appliquons la TF à chaque signal discret  $f \in \{x, y, p, az, al\}$ . Les coefficients de Fourier d'un signal discret  $f$  de taille  $N$  sont obtenus selon les formules suivantes :

$$C_k = \frac{1}{N} \sum_{t=0}^{N-1} f(t) e^{-i2\pi kt/N} \text{ pour } k = 0, \dots, N-1 \quad (III.1)$$

$$\text{Ou } e^{-i2\pi kt/N} = \cos(2\pi kt/N) + i \sin(2\pi kt/N)$$

Les coefficients  $C_k$  sont des nombres complexes. Le traitement sera restreint aux modules de ces coefficients. Les modules  $|C_k|$  de chaque signal sont alors calculés :

$$|C_k| = \sqrt{\text{réel}(C_k)^2 + \text{Imag}(C_k)^2}, \text{ pour } k = 0, \dots, N-1 \quad (III.2)$$

Où  $\text{réel}(C_k)$  est la partie réelle de  $C_k$  et  $\text{Imag}(C_k)$  sa partie imaginaire. Notons qu'il est possible de revenir au signal d'origine à partir de sa transformée de Fourier, par l'application de la transformée inverse de la manière suivante :

$$f(t) = \sum_{k=0}^{N-1} c_k e^{i2\pi kt/N} \text{ pour } t = 0, \dots, N-1 \quad (III.3)$$

La figure (III.2) montre les modules de la transformée de Fourier des signaux de la signature de la figure (III.1). Nous pouvons bien remarquer la symétrie de la transformée ainsi que la concentration de l'énergie dans les fréquences basses. Seuls les modules des coefficients sont utilisés dans la suite. La phase est ainsi ignorée.

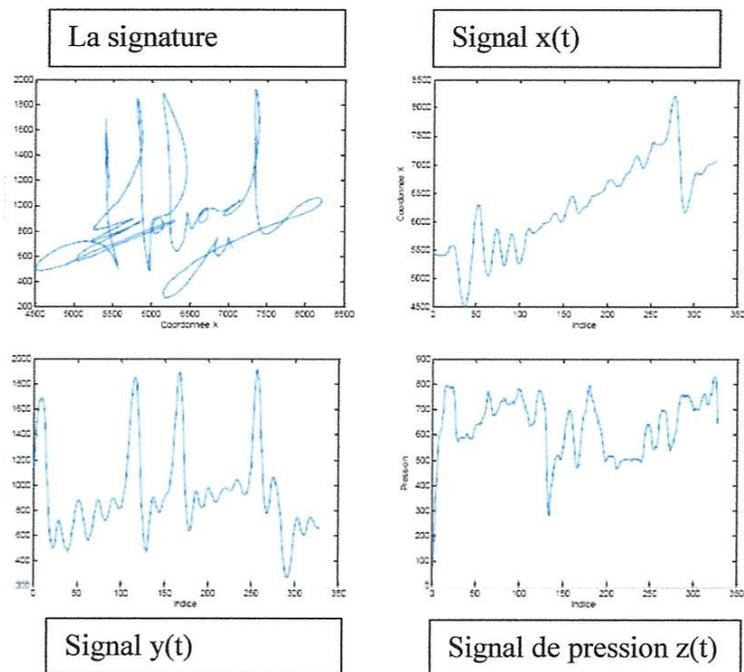


Figure III.1: Exemple des signaux d'une signature en ligne (x(t), y(t), z(t))

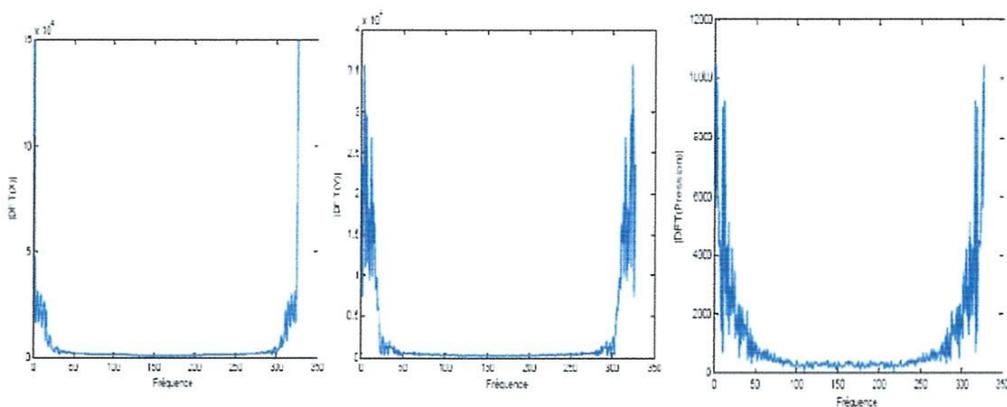


Figure III.2 : Transformée de Fourier des signaux de la signature montrée dans la figure (III-1)

### III.2 LA TRANSFORMEE EN COSINUS DISCRETE :

La transformée en cosinus discrète ou TCD (de l'anglais : DCT ou *Discrete Cosine Transform*) est une transformation proche de la transformée de Fourier discrète (DFT). Le noyau de projection est un cosinus et crée donc des coefficients réels, contrairement à la DFT, dont le noyau est une exponentielle complexe qui crée donc des coefficients complexes. La DCT est très utilisée en traitement du signal et de l'image, spécialement en compression. La DCT possède en effet une excellente propriété de regroupement de l'énergie. L'information est essentiellement portée par les coefficients de basses fréquences. Seuls un petit nombre de coefficients sont non-

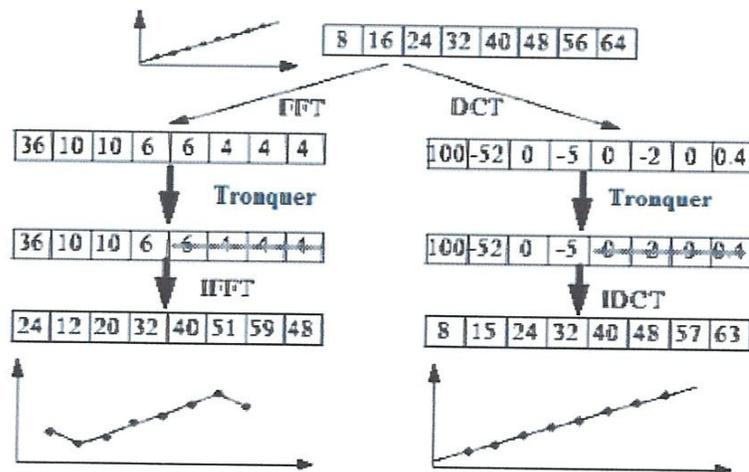
nuls, et peuvent être utilisés pour reconstruire le signal par la transformée inverse (IDCT).

Dans le cas de la compression, le gain vient de la suppression des coefficients nuls ou proches de zéro. La DCT est utilisée dans les standards de compression de l'image et du son (JPEG, MPEG, MP3, etc.).

La DCT est aussi employée pour la résolution de systèmes d'équations différentielles par des méthodes spectrales.

En fait, il existe une relation entre DCT et FFT. La DCT n'est qu'une version allégée de la FFT: seule la partie réelle de la FFT. La FFT est souvent utilisée pour calculer la DCT grâce à la rapidité de l'algorithme de calcul de la FFT.

La DCT est beaucoup plus couramment utilisés par rapport à la FFT. La transformée DCT fournit un bon compromis entre la capacité de regroupement d'informations et la complexité de calcul. Un autre avantage de la DCT, c'est que la plupart des composantes DCT sont généralement de très faible ampleur car la plupart des informations importantes existent dans les coefficients de basses fréquences [NLU08]. Pour montrer l'avantage de la DCT par rapport à la FFT, nous menons l'expérience suivante (voir schéma ci-après). On applique les transformées DCT et FFT sur le même vecteur de données. Ensuite, on élimine quelques coefficients des deux vecteurs de transformées. Enfin, on revient au signal d'origine par les transformées inverses.



On remarque bien que le signal obtenu par la DCT inverse est très proche du signal original alors que le signal obtenu par la FFT est déformé. Cette expérience montre l'efficacité de compression

de donnée apportée par la DCT. Toutes les raisons précédentes nous ont conduits à utiliser de la DCT, pour l'extraction de paramètres comme alternative à la DFT dans les travaux menés sur les systèmes de vérification de signatures manuscrite en ligne. A notre connaissance, la DCT n'a jamais été utilisée directement pour l'extraction des paramètres dans les systèmes de vérification de signatures en ligne. Le seul cas où nous avons rencontré la DCT dans de tels systèmes est le travail de Nanni *et al* [NLU08] où la DCT a été appliquée sur les paramètres ondelettes (Wavelette) pour des fins de réduction de dimensionnalité. Il existe quelques variantes standards pour calculer la DCT. La variante la plus courante de la transformée en cosinus discrète est la suivante :

$$F_k = \alpha(k) \sum_{t=0}^{N-1} f_t \cos \left[ \frac{\pi}{N} \left( t + \frac{1}{2} \right) k \right], k = 0, \dots, N - 1 \quad (III.4)$$

$\alpha(k)$  est défini comme suit :

$$\alpha(k) = \begin{cases} \sqrt{1/N}, \text{ pour } k = 0 \\ \sqrt{2/N}, \text{ pour } k \neq 0 \end{cases} \quad (III.5)$$

La transformée inverse est donnée par :

$$f_t = \sum_{k=0}^{N-1} \alpha(k) F_k \cos \left[ \frac{\pi}{N} k \left( t + \frac{1}{2} \right) \right], t = 0, \dots, N - 1 \quad (III.6)$$

Dans notre cas, comme nous avons fait avec la DFT, la DCT est appliquée séparément sur chaque signal de la signature. La figure suivante illustre la transformée en DCT des signaux de la signature précédente (montrée dans la figure III.1) :

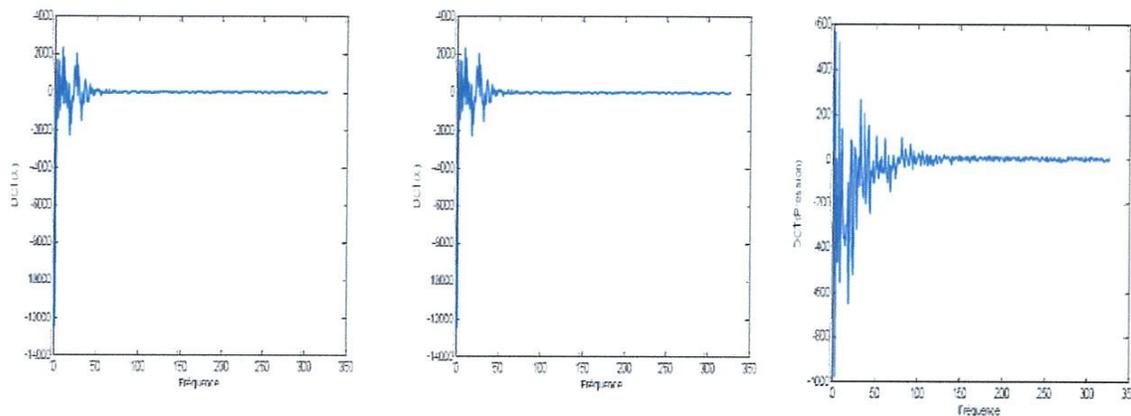


Figure III.3 : Transformée en cosinus des signaux de la signature montrée dans la figure III.1

### III.3 REDUCTION DE PARAMETRES :

Généralement, le processus d'échantillonnage introduit du bruit sur les signaux de la signature échantillonnée par la tablette à digitaliser. En effet, il est acquis que dans le signal de la signature en ligne, les basses fréquences sont les plus informantes, les hautes fréquences, dans la plupart du temps, représentent les oscillations involontaires sur le tracé de la signature [VAN05]. D'après [PLO89], des observations sur des exemples de signature ont conduit à un résultat relativement constant: un filtre passe-bas de fréquence de coupure 10Hz conserve plus de 90% de la puissance du signal.

Nous avons étudié plusieurs signatures de la base de données SUSIG Visual et nous avons remarqué que seul un nombre restreint de coefficients de basses fréquences ont une valeur significative. Les autres coefficients sont tous nuls ou presque nuls. Le travail dans le domaine fréquentiel simplifie l'élimination des composés de hautes fréquences. En effet, il suffit d'ignorer les derniers coefficients nuls ou proches de zéro de chaque vecteur de paramètres. Pour les paramètres de Fourier, comme la TF est symétrique, on ne considère que la moitié des coefficients de chaque signal. La réduction de paramètres permet d'accélérer davantage les calculs (seuls les premiers coefficients significatifs sont calculés) et de faire la comparaison entre les signatures sans dégrader les performances du système. Les figures suivantes montrent les paramètres DFT et DCT de la signature après réduction à 50 coefficients.

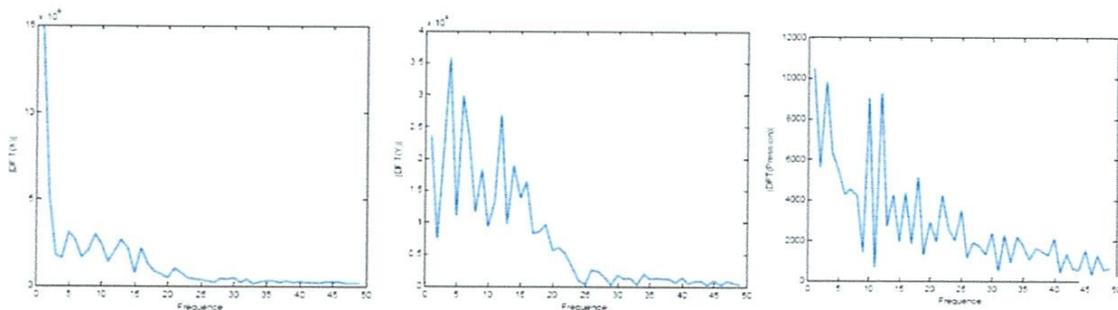


Figure III.4 : Coefficients paramètres DFT

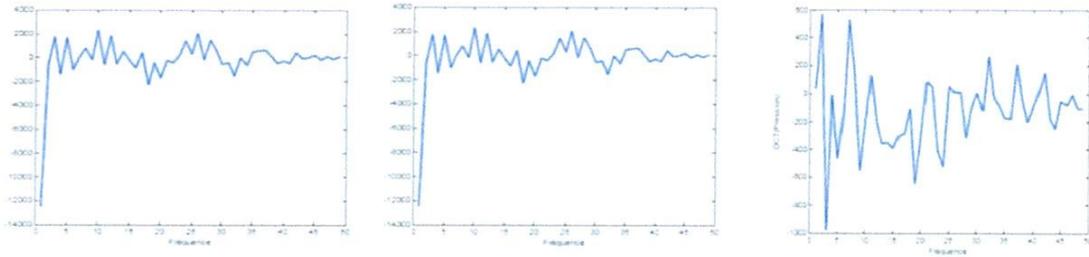


Figure III.5 : Coefficients paramètres DCT

### III.4 METRIQUE DE DIS-SIMILARITE UTILISEE :

Dans la phase d'entraînement du système, chaque utilisateur  $j$  fournit un ensemble de signatures références  $R_j$  utilisés pour accepter ou refuser une requête d'authentification. Au moment de la vérification, une signature est soumise au système avec l'identité proclamée. La dis-similarité entre les paramètres de cette signature par rapport à ceux de la signature de référence de la personne proclamée sera calculée. Dans notre cas, la dis-similarité d'une requête  $q$  est calculée comme la distance entre ses paramètres  $F_q$  et les paramètres de chacune des signatures références  $F_{R_j}$

$$D_j(F_q, F_{R_j}) = \|F_q - F_{R_j}\| \quad (III.7)$$

Le score de comparaison est considéré en combinant toute les distances de la requête par rapport aux différentes signatures références. On rencontre dans la littérature plusieurs types de distance Soient  $X(x_1, x_2, \dots, x_N)$  et  $Y(y_1, y_2, \dots, y_N)$  deux vecteurs de dimension  $N$ . On définit :

- La distance L1 :

$$D_{L1}(X, Y) = \sum_{i=1}^N |x_i - y_i| \quad (III.8)$$

- La distance L2 (Distance Euclidienne) :

$$D_{L2}(X, Y) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \quad (III.9)$$

- La distance Mahalanobis : la distance Mahalanobis d'un vecteur  $X$  par rapport à un ensemble de donnée de moyenne  $\mu$  et matrice de covariance  $\Sigma$  est donnée par :

$$D_{L2}(X) = (X - \mu)^T \Sigma^{-1} (X - \mu) \quad (III.10)$$

- l'évaluation de la déviation standard entre deux signatures (distance moyenne). Si  $S_{ref}$  et  $S_T$  deux signatures données, leur déviation standard est définie par :

$$DS = \sqrt{\frac{|C_t - C_{ref}|^2}{I_{ref}}} \quad (III.11)$$

- DS est la déviation standard.
- $C_t$  est une composante du signal de test.
- $C_{ref}$  est une composante du signal de référence.
- $I_{ref}$  est la longueur du signal de référence.

Nous avons retenu cette distance dans notre travail.

### **CONCLUSION :**

La méthode proposée est importante dans l'étape de para-métrisation des signatures, car elle permet une réduction de données au niveau des signatures sans perte significative de l'information de départ. On peut ainsi reconstruire les signatures d'origine sans difficulté. Le critère de comparaison retenu peut-être modifié en prenant d'autres types de distances (distance de Mahalanobis, euclidienne, distance entre gaussienne, ...etc.). La mise en œuvre pratique de cette méthode est examinée au chapitre IV suivant.

## **Chapitre IV**

### **RESULTATS EXPERIMENTAUX : Résultats des Performances du système évalués sur l'approche Comparaison**

Seules les coordonnées  $x(t)$ ,  $y(t)$ ,  $p(t)$  et l'état du stylo sont retenues pour l'évaluation des performances du système. Les résultats qui suivent sont évalués sur les coordonnées  $x(t)$  seulement. Il en sera de même pour les autres paramètres ( $y(t)$  et  $p(t)$ ). L'organisation générale du système étudié est présentée en figure IV.1

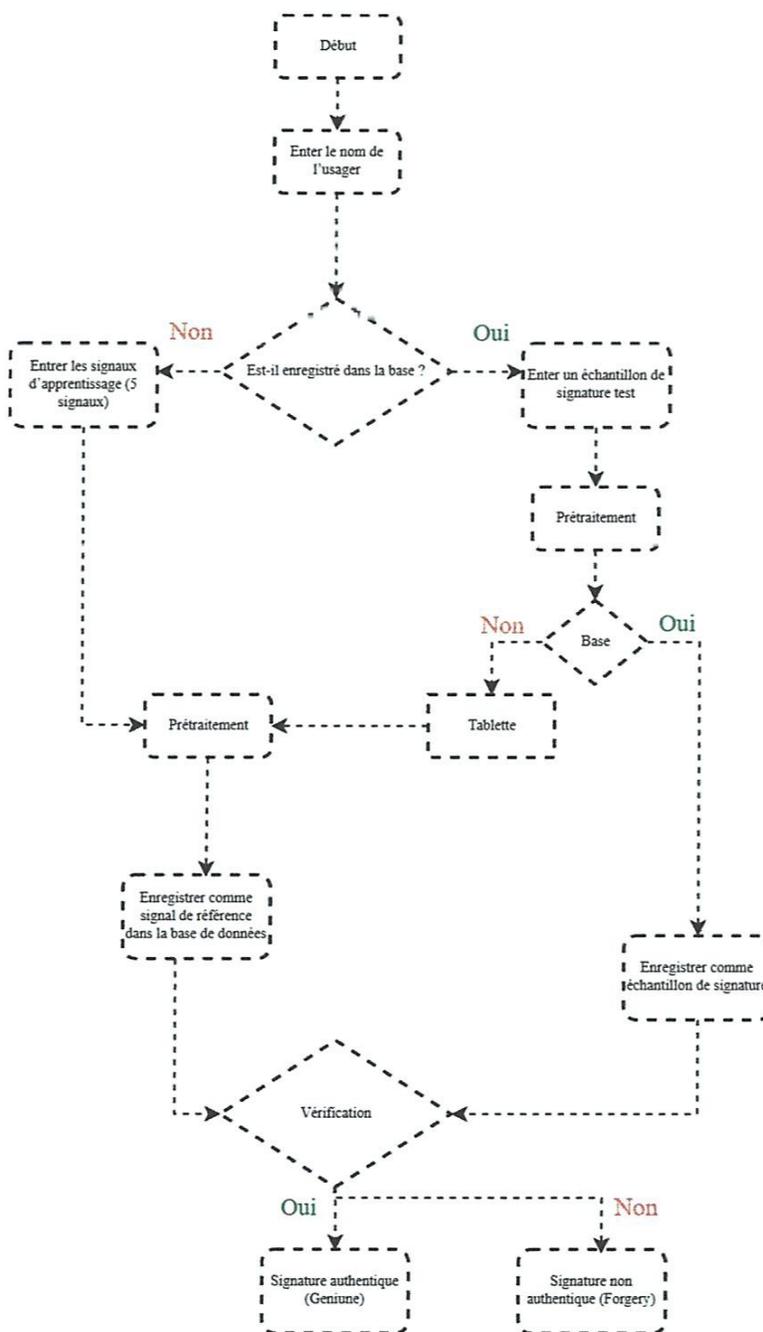


Figure IV.1 : Organisation générale du système étudié

Nous décrivons dans ce qui suit les principales opérations à effectuer pour l'évaluation des performances de notre système.

#### IV.1 SIGNAUX D'APPRENTISSAGE :

Dans le système de vérification de signatures en ligne proposé, les utilisateurs sont d'abord introduits dans le système par l'enregistrement de quelques échantillons de leurs signatures qui servent de références (phase d'enrôlement). Ensuite, quand un utilisateur, qui prétend être un client particulier du système, présente sa signature pour la vérification, elle est comparée avec les signatures de référence de l'individu proclamé. La figure IV.2 illustre l'allure des signaux servant à la création du signal de référence (dans notre cas 5 signaux  $x(t)$ ).

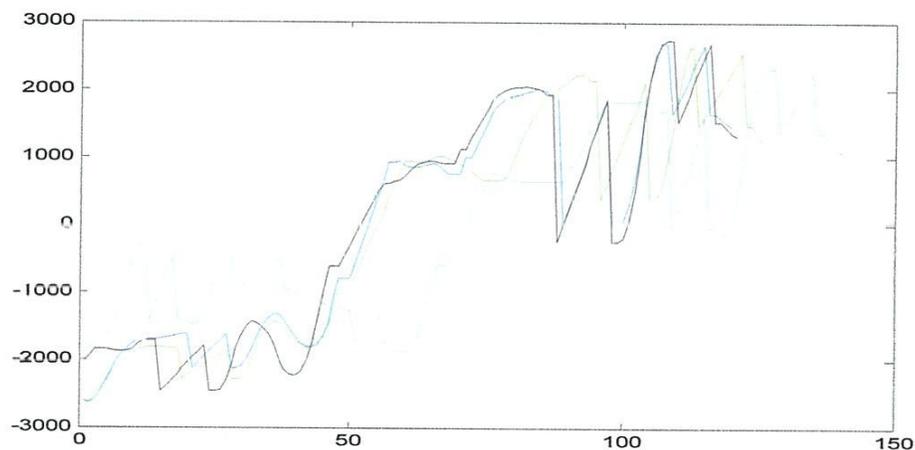


Figure IV.2 : Exemple de signaux utilisés en phase d'apprentissage.

#### IV.2 PRETRAITEMENT :

##### IV.2.1 Extraction des segments de signature :

Cette procédure consiste à faire ressortir le nombre de segments présents dans une signature donnée. Un segment est défini comme une suite de coordonnées codée 1 (stylet appuyé sur la tablette) séparé par. Deux coordonnées codées 0, définissant chacune, le début (DEBUTSEG) et la fin du segment (FINSEG). La figure suivante, montre le codage correspondant à un segment de tracé dans une signature manuscrite.



appartenant à l'équation de la droite reliant ces deux points est introduit. L'opération permettant l'ajout de ces nombre est évaluée en prenant en considération le temps de levé de plume ( $\Delta t$ ) calculé comme suit :

$$\Delta t = T_{aqseg2} - T_{aqseg1} \quad (IV.1)$$

Où  $T_{aqseg2}$  est le temps de début d'acquisition du segment 2 et  $T_{aqseg1}$  est le temps d'acquisition du segment 1.

Le nombre des coordonnées à ajouter est calculé suivant cette formule :

$$\text{Nbr de point} = \Delta t / T_e \quad (IV.2)$$

Où  $T_e$  est la période d'échantillonnage.

L'algorithme suivant illustre le principe d'obtention de ce segment :

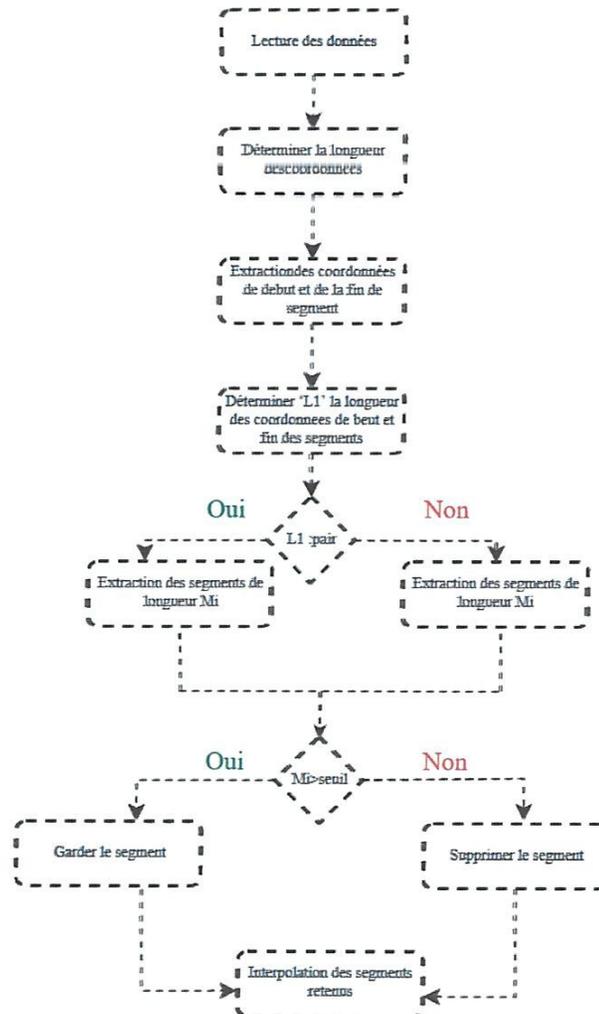


Figure IV.5 : algorithme d'extraction des segments d'une signature.

Les figures suivantes (montrent) illustre le principe d'obtention des segments, leurs interpolations pour l'obtention d'une signature mono-segment

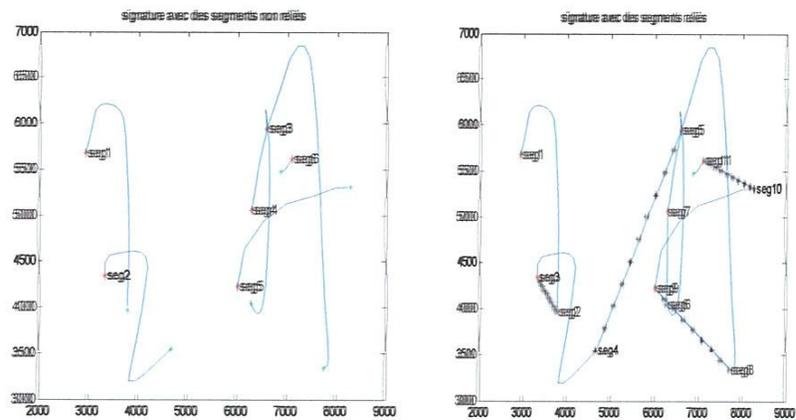


Figure IV.6 : segmentation d'une signature afin d'obtenir une signature mono-segment.

### IV.2.3 Translation de la signature :

En réalité, les coordonnées du stylo dépendent beaucoup de la position de la signature sur la tablette: les coordonnées sont petites lorsqu'on signe près de l'origine et deviennent grandes lorsqu'on signe loin de cette origine. Une translation de la signature pour normaliser les coordonnées, est donc indispensable si nous voulons les utiliser directement. Certains auteurs tradent la signature, parfois afin que les coordonnées du premier point de la signature soient à l'origine, parfois afin que les valeurs minimales des coordonnées soient nulles. Cependant, ces translations accentuent l'importance d'un point spécifique de la signature, et ne tiennent pas compte du reste de la signature. Ceci génère de graves problèmes lorsque le point de repère pour la normalisation n'est pas stable. Etant donné la variabilité de la signature manuscrite, ce problème est assez fréquent. En principe, une bonne translation doit tenir compte de l'information de toute la signature afin de minimiser l'effet lié à l'instabilité de la signature. Dans nos travaux, nous appliquons une translation de la signature telle que le centre de gravité de la signature soit à l'origine. Pour cela, nous devons calculer tout d'abord les coordonnées du centre de gravité par les deux équations ci-dessous:

$$x_g = \frac{1}{T} \sum_{i=1}^T x_i \quad (IV.3)$$

$$y_g = \frac{1}{T} \sum_{i=1}^T y_i \quad (IV.4)$$

où  $T$  est le nombre total de points de la signature.

Comme le centre de gravité est calculé à partir de tous les points de la signature, il tient compte de toute l'information de cette dernière, et est ainsi plus stable qu'un point spécifique de la signature. Les coordonnées sont ensuite normalisées comme suit:

$$x_i = x_i - x_g, \quad 1 \leq i \leq T \quad (IV.5)$$

$$y_i = y_i - y_g, \quad 1 \leq i \leq T \quad (IV.6)$$

La translation de la signature par rapport à son centre de gravité est aussi proposée par Igarza. En effet, l'auteur compare la performance du système de vérification de signatures en utilisant les deux repères pour la translation (par rapport au premier point et au centre de gravité) et prouve par expérimentation que la translation par rapport au centre de gravité donne le meilleur résultat.

#### IV. 2.4 Normalisation en taille de la signature :

On a appliqué dans notre démarche la normalisation en taille de la signature basée sur la méthode de la variance est utilisée. Celle-ci considère que la taille d'une signature dépend de l'espace offert au signataire. La normalisation des signatures en taille consiste à transformer les signatures de telles sortes qu'elles soient englobées dans un rectangle de taille donnée. Une normalisation basée sur la variance utilise les formules suivantes :

$$x_n = x * \frac{C_x}{V_x} \quad (IV.7)$$

$$y_n = y * \frac{C_y}{V_y} \quad (IV.8)$$

Où  $x_n$  et  $y_n$  sont les coordonnées après normalisation et  $x$  et  $y$  sont les coordonnées initiales de la signature.

$C_x$  et  $C_y$  sont des constantes données. Le choix de  $C_x$  et  $C_y$  se fait à travers les formules suivantes:

$$C_x = V_x * \frac{X}{x} \quad (IV.9)$$

$$C_y = V_y * \frac{Y}{y} \quad (IV.10)$$

Où  $X$  et  $Y$  sont les coordonnées du rectangle ou sont englobées les signatures

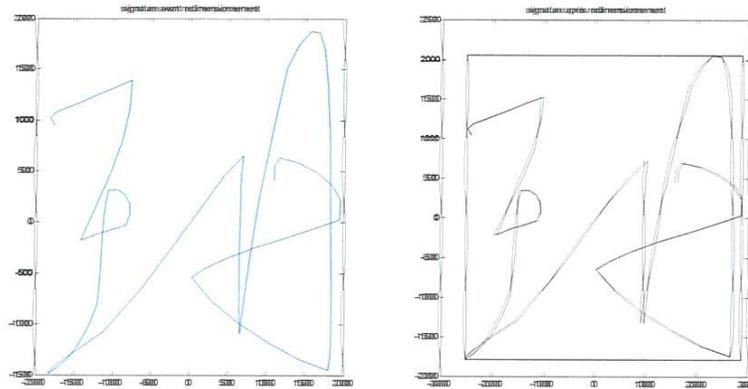


Figure IV.8 : normalisation en taille de la signature.

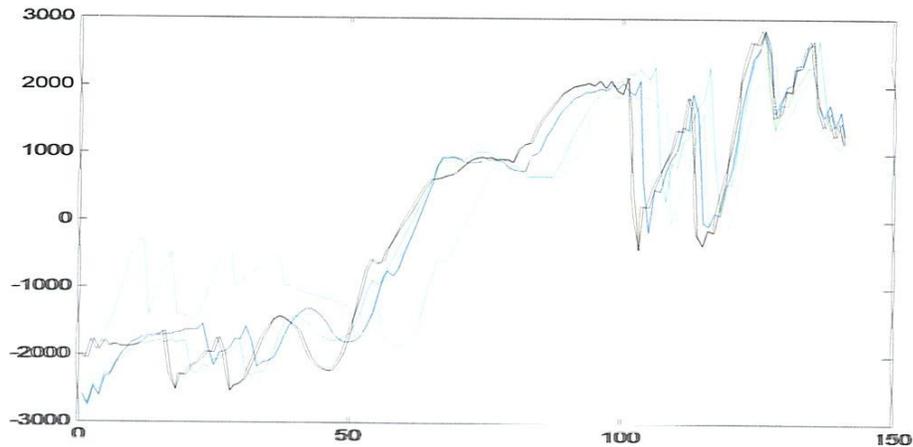
L'algorithme de normalisation est le suivant :



Figure IV.9 : algorithme de la normalisation en taille de la signature.

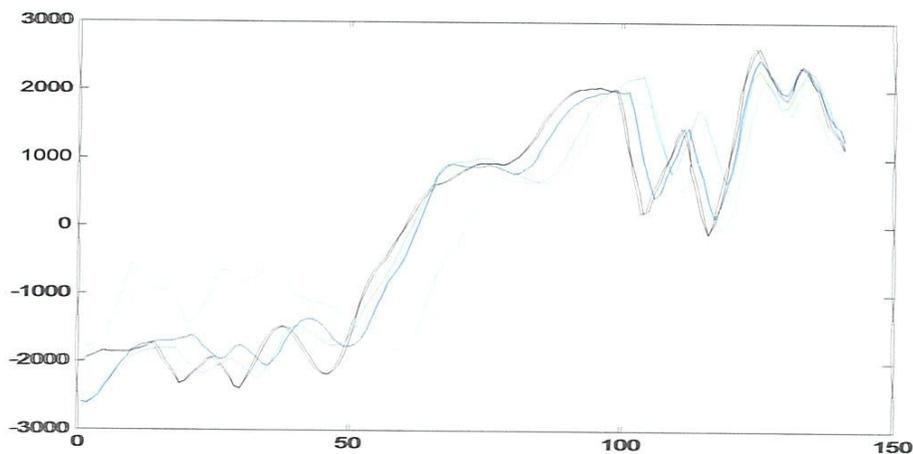
### IV.2.5 Re-échantillonnage :

Les signaux une fois normalisés sont ré-échantillonnés de manière à conserver la même taille c.-à-d. la même longueur durant tout le processus (figure IV.10).



**Figure IV.10 :** Signaux d'apprentissage après normalisation.

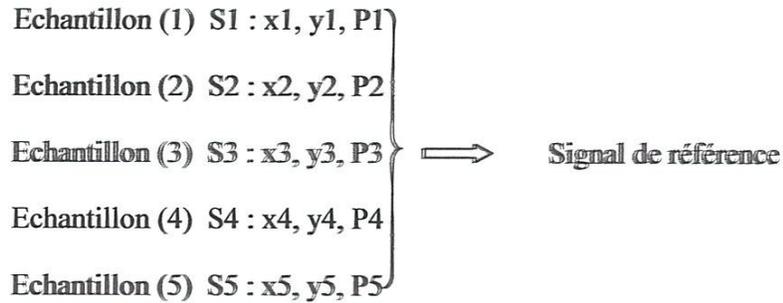
La normalisation des signaux introduit du bruit (faux maxima et de faux minima au niveaux des maxima et minima locaux), une phase de filtrage (smoothing) est introduite pour supprimer ce bruit (Figure IV.11).



**Figure IV.11:** Signaux d'apprentissage après normalisation et filtrage.

### IV.3 OBTENTION DU SIGNAL DE REFERENCE :

Pour obtenir le signal de référence, l'utilisateur entre dans le système 5 échantillons de signature. Si la suite des 5 premiers échantillons de l'utilisateur est décrite par :



Et

Echantillon (T) ST : xT, yT, pT : échantillon de test quelconque

Le signal de référence est obtenu après normalisation par :

$$S_{ref} = S1 + S2 + S3 + S4 + S5 / 5 \tag{IV.11}$$

La figure IV.12 illustre le principe d'obtention du signal de référence

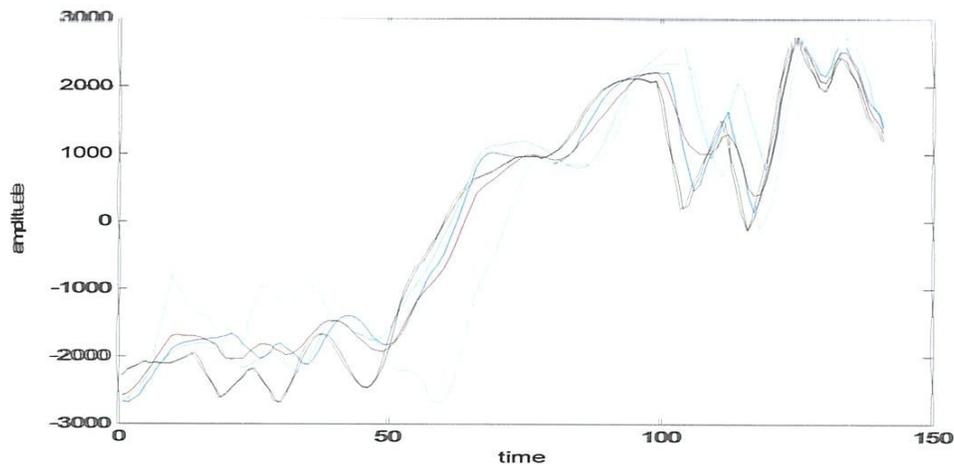


Figure IV.12 : Principe d'obtention du signal de référence à partir des signaux tests.

#### IV.4 VERIFICATION :

Après obtention de la transformé de Fourier du signal de référence et de son prétraitement, le système passe à la phase de vérification. La première approche pour l'évaluation de notre système repose sur l'approche par comparaison dont le principe a été décrit (voir paragraphe). Dans cette approche le composant comparaison consiste en un classificateur qui compare la transformé de la signature à vérifier avec la transformée du signal de référence de l'utilisateur et donne un score mesurant la similarité ou dis-similarité selon le classificateur utilisé. Le classificateur retenu dans notre cas est celui basé sur l'évaluation de la déviation standard entre deux signatures.

Si  $S_{ref}$ , et  $S_T$  deux signatures données, leur déviation standard est définie par :

$$DS = \sqrt{\frac{|c_t - c_{ref}|^2}{I_{ref}}} \quad (IV.12)$$

DS est la déviation standard.

$C_t$  est une composante du signal de test.

$C_{ref}$  est une composante du signal de référence.

$I_{ref}$  est la longueur du signal de référence.

#### IV.5 ALGORITHME D'OBTENTION DES SEUILS DE DECISION :

Nous proposons dans ce qui suit une démarche qui permet d'estimer les taux de fausse acceptation (FAR) et faux rejet (FFR) pour les signatures paramétrées par TF[x(t), y(t), p(t)]. Nous rappelons que seule l'étude portée sur TF[x(t)] est retenue dans ce projet. La même démarche sera appliquée pour TF[y(t)] et TF[p(t)].

Pour l'estimation du taux FFR, nous procédons comme suit :

- 1- Calcul de la déviation standard entre les TF des signatures de références obtenues des usagers inscrits dans la base, et les TF de leurs signatures authentiques (voir tableau 1). Notons que le calcul de la déviation s'effectue entre les coordonnées fréquentielles des deux signatures TF[X(t)] et TF[Y(t)] définies par :

$$TF[X(t)] = TF[(x_1(t), y_1(t), p_1(t))], \text{ et } TF[Y(t)] = TF[(x_2(t), y_2(t), p_2(t))].$$

A partir des résultats obtenus en tableau IV.1, il faut déterminer les seuils pour les signatures qualifiées de Faux rejet. Pour ce faire, la démarche suivante est adoptée

- a) Faire ressortir du tableau IV.1, les déviations standard maximales (max) et les déviations minimales (min). Ce qui introduit un intervalle de variation associé à la variabilité des signatures authentiques introduites dans la base.
- b) Découper l'intervalle ainsi obtenu en 10 sections. Chaque section est pondérée par un poids TH ( $0 < TH < 1$ ). Une pondération pour une section quelconque est définie par:

$$\text{SECTION (i)} = (\text{max} - \text{min}) \text{ TH (i)}.$$

- c) Détermination du seuil de décision pour TF[x(t)], TF[y(t)], TF[p(t)]. Par exemple pour TH = 0.2, nous pouvons calculer le seuil TH\_x pour la coordonnée TF[x(t)] par :

$$\text{TH}_x = \text{max} - (\text{max}-\text{min}) \text{ TH}.$$

- d) Evaluation des performances du système en fonction des seuils adoptés (on comptabilise le nombre de faux rejet pour chaque TH considéré),

Les résultats expérimentaux portés par les tableaux (IV.1, IV.2, IV.3, IV.4, et IV.5) témoignent de l'efficacité de la démarche adoptée.

TF [Sig ref du Signataire]	TF [Ech1]	TF [Ech2]	TF [Ech3]	TF [Ech4]	TF [Ech5]	TF [Ech6]	TF [Ech7]	TF [Ech8]	TF [Ech9]	TF [Ech10]
1	0,427731	0,753071	2,045180	4,689628	3,852012	7,494711	5,750783	6,309137	8,055018	14,25512
2	13,64794	9,829835	6,333199	1,662462	7,591366	13,43454	4,444029	5,959379	3,776343	0,886633
3	0,833684	11,20562	5,249825	2,091101	3,551452	13,68398	11,22172	19,98382	11,91162	8,548886
4	3,408885	8,297970	1,735132	1,217178	2,338589	4,559224	5,799321	0,493492	0,400978	1,959067
5	3,097295	6,872498	1,777278	2,247216	2,551847	12,61899	7,806666	5,697192	5,129710	0,864526
6	0,854220	2,070535	1,153737	7,089047	7,955768	10,78677	2,800584	1,321047	2,155463	2,173053
7	0,952830	4,088588	4,261496	0,904201	3,734279	2,441053	1,450659	1,618568	10,16391	4,125937
8	3,570084	7,753858	15,85450	10,38474	11,50928	12,62941	10,88507	8,021034	4,747857	10,53146
9	11,60386	4,538592	5,500050	9,204532	7,167577	1,526443	0,281502	3,011485	7,472803	2,798155
10	9,345324	1,332220	15,03091	15,94296	2,174904	5,443063	4,372841	4,344080	4,853131	4,162314
11	4,636002	0,845899	0,842223	1,577777	2,613162	3,405130	2,371199	5,672134	7,579721	0,208328
12	0,164531	6,568030	0,866379	1,039632	4,150562	4,620902	5,155998	2,792040	1,841940	5,508093
13	3,818568	1,123193	2,090615	0,076724	1,271536	7,455445	0,341197	1,573058	1,503444	2,544989
14	2,000639	2,971518	0,829040	1,226056	3,101643	4,917290	5,356105	0,744146	4,251525	0,435941
15	3,477338	3,152469	4,202218	3,458161	3,908806	0,122208	0,803414	5,452816	7,848943	10,14847
16	2,468873	4,039969	1,379034	3,021542	1,813854	1,003376	2,481452	2,973082	0,618485	0,734258
17	2,925391	1,160383	3,640600	6,565938	1,457170	1,211915	2,462841	1,921489	11,16092	0,865878
18	2,923554	1,422216	8,383265	1,982513	1,368606	8,579684	8,050645	0,789037	5,531107	3,937064
19	2,483052	1,572410	5,158652	1,863012	2,149907	2,756857	1,502941	5,825890	5,129240	1,772295
20	3,782421	2,936864	4,365527	1,575488	0,732387	4,517463	0,683138	4,404164	0,646490	2,549702

**Tableau IV.1:** Déviation standard utilisée comme indice de ressemblance entre signatures.

TF [Sig ref du Signataire]	TF [Ech1]	TF [Ech2]	TF [Ech3]	TF [Ech4]	TF [Ech5]	TF [Ech6]	TF [Ech7]	TF [Ech8]	TF [Ech9]	TF [Ech10]
1	0,427731	0,753071	2,045180	4,689628	3,852012	7,494711	5,750783	6,309137	8,055018	14,25512
2	13,64794	9,829835	6,333199	1,662462	7,591366	13,43454	4,444029	5,959379	3,776343	0,886633
3	0,833684	11,20562	5,249825	2,091101	3,551452	13,68398	11,22172	19,98382	11,91162	8,548886
4	3,408885	8,797970	1,735137	1,217178	2,338589	4,559774	5,799371	0,493497	0,400978	1,959067
5	3,097295	6,872498	1,777278	2,247216	2,551847	12,61899	7,806666	5,697192	5,129710	0,864526
6	0,854220	2,070535	1,153732	7,089047	7,955768	10,78672	2,800584	1,321047	2,155463	2,173053
7	0,957830	4,088588	4,261496	0,904201	3,734279	2,441053	1,450659	1,618568	10,16391	4,125937
8	3,570084	7,753858	15,85450	10,38474	11,50928	12,62941	10,88507	8,021034	4,747857	10,53146
9	11,60386	4,538592	5,500050	9,204532	7,167577	1,526443	0,281502	3,011485	7,472803	2,798155
10	9,345324	1,332220	15,03091	15,94296	2,174904	5,443063	4,372841	4,344080	4,853131	4,162314
11	4,636002	0,845899	0,842223	1,577777	2,613162	3,405130	2,371199	5,672134	7,579721	0,208328
12	0,164531	6,568030	0,866379	1,039632	4,150562	4,620902	5,155998	2,792040	1,841940	5,508093
13	3,818568	1,123193	2,090615	0,076724	1,271536	7,455445	0,341197	1,573058	1,503444	2,544989
14	2,000639	2,971518	0,829040	1,226056	3,101643	4,917290	5,356105	0,744146	4,251525	0,435941
15	3,477338	3,152469	4,202218	3,458161	3,908806	0,122208	0,803414	5,452816	7,848943	10,14847
16	2,468873	4,039969	1,379034	3,021542	1,813854	1,003376	2,481452	2,973082	0,618485	0,734258
17	2,925391	1,160383	3,640600	6,565938	1,457170	1,211915	2,462841	1,921489	11,16092	0,865878
18	2,923554	1,422216	8,383265	1,982513	1,368606	8,579684	8,050645	0,789037	5,531107	3,937064
19	2,483052	1,572410	5,158652	1,863012	2,149907	2,756857	1,502941	5,825890	5,129240	1,772295
20	3,782421	2,936864	4,365527	1,575488	0,732387	4,517463	0,683138	4,404164	0,646490	2,549702

Tableau IV.2 : détermination du taux de faux Rejet pour TH = 0.2.

TH	X(f)	Y(f)	P(f)
0.0	0	0	0
0.1	1	2	1
0.2	1	2	9
0.3	5	5	14
0.4	10	6	20
0.5	22	11	25
0.6	31	20	34
0.7	48	35	47
0.8	86	58	71
0.9	134	122	125
1	199	199	199

**Tableau IV.3 :** Extraction du nombre de faux rejet.

TH	X(f)	Y(f)	P(f)
0.0	0	0	0
0.1	0.5	1	0.5
0.2	0.5	1	4.5
0.3	2.5	2.5	7
0.4	5	3	10
0.5	11	5.5	12.5
0.6	15.5	10	17
0.7	24	17.5	23.5
0.8	43	29	37
0.9	66.5	61	62.5
1	99.5	99.5	99.5

**Tableau IV.4 :** Taux de faux rejet (FRR) obtenu.

TF [Sig ref du Signatai re]	TF [Ech1]	TF [Ech2]	TF [Ech3]	TF [Ech4]	TF [Ech5]	TF [Ech6]	TF [Ech7]	TF [Ech8]	TF [Ech9]	TF [Ech1 0]
1	16,15793	6,634197	13,75236	14,19111	2,406620	0,360556	0,089619	10,01178	3,267129	13,13136
2	16,39458	12,07454	0,404586	5,023028	3,828483	3,895740	2,637995	5,783809	6,945252	1,202514
3	3,820409	1,493654	2,286178	0,955412	4,886718	8,580600	13,20701	1,116741	22,02022	12,65632
4	3,828073	1,657347	12,04944	4,177688	6,631320	0,442392	3,038972	5,660009	2,001367	10,39574
5	13,16905	11,53013	14,23095	11,55281	5,867763	9,891402	2,874431	3,835762	0,697313	13,35296
6	10,36165	13,56634	13,06778	1,255828	5,323689	2,632113	0,595298	6,497062	18,70273	13,00552
7	2,498502	4,406067	3,885637	1,181856	1,751534	12,18931	5,122517	1,769993	2,243819	0,972667
8	2,408378	6,118439	3,364571	8,302328	6,729677	1,673238	3,245358	5,063303	13,32560	16,80931
9	13,12914	9,801011	5,819640	7,801369	12,191046	12,117760	8,484115	3,970097	16,00999	12,65667
10	9,732118	22,20134	5,843507	19,01503	15,85605	12,62391	15,34719	8,728894	22,21599	20,03600
11	7,616256	14,44760	7,803727	6,469580	0,749967	0,199536	2,510534	2,638166	1,055341	4,659655
12	0,744934	4,653306	4,344165	1,982751	2,121639	10,78916	1,992842	2,814269	7,728734	6,462407
13	2,120439	1,353986	0,448056	1,315053	11,15434	7,641929	10,52776	9,536298	12,75906	12,39768
14	0,772850	3,682562	2,854043	0,647388	5,716073	1,389746	2,905427	1,372750	7,273122	12,24351
15	7,983794	3,870250	0,780534	0,885897	0,298189	3,729077	0,391242	0,946306	10,49996	5,386229
16	3,938932	3,145986	3,297639	4,757858	3,466801	0,919344	1,809880	3,304640	1,695951	4,179679
17	4,018458	3,870916	2,361362	4,100940	3,729865	0,350146	3,325453	0,483836	3,491377	10,35819
18	2,717410	3,240901	1,314905	2,097312	1,175415	3,678444	4,257927	1,949493	8,225754	6,850167
19	4,904647	3,644775	8,232609	2,452365	8,073498	5,157165	6,090873	1,362380	6,037280	8,179332
20	4,386846	1,899343	2,833026	2,632214	3,414242	5,938041	14,12695	10,18784	4,974372	6,152520

Tableau IV.5 : détermination du taux de fausse acceptation pour TH = 0.2

Le FRR est calculé par la formule suivante :

$$FRR = \frac{No\_of\_false\_rejection\_sample}{No\_of\_geniune\_signature} \times 100\% \quad (IV.13)$$

Pour le calcul du taux de fausse acceptation(FAR), celui-ci est obtenu en effectuant une comparaison entre les signatures de références des 10 usagers et l'ensemble des signatures de la base (Geniune + imposteurs) ceci dans un souci de tester la sensibilité du système considéré. Même les signatures authentiques dans ce cas précis peuvent être considérées comme des signatures d'attaques au système. Par un raisonnement analogue aux précédant, nous dressons tous les tableaux nécessitants la détermination des seuils pour les signatures qualifiées de fausse acceptation ainsi que le nombre qui leur est associé, suivant le TH considéré. Les tableaux (IV.6, IV.7, IV.8, IV.9), illustrent les résultats obtenus pour l'obtention des seuils du FAR, nécessaires à l'étude de la performance du système étudié.

Le FAR est calculé par la formule suivante :

$$FAR = \frac{No\_of\_false\_acceptance\_sample}{No\_of\_forgery\_signature} \times 100\% \quad (IV.14)$$

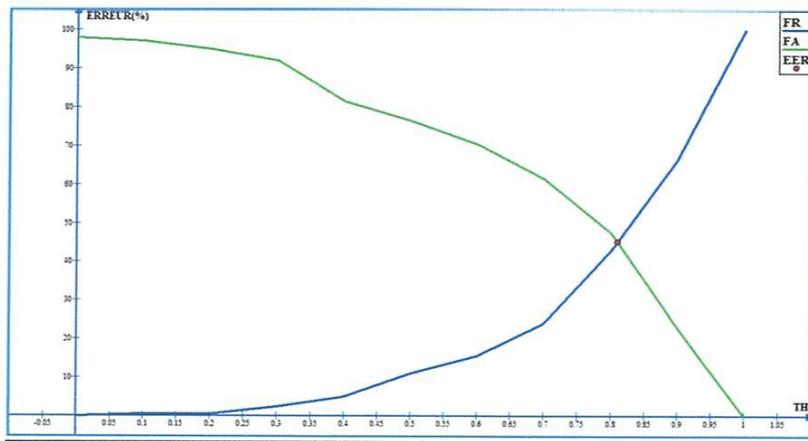
TH	X(f)	Y(f)	P(f)
0.0	196	200	149
0.1	194	198	138
0.2	190	197	129
0.3	184	194	121
0.4	163	190	117
0.5	153	181	104
0.6	141	171	86
0.7	123	151	60
0.8	95	117	36
0.9	46	71	16
1	0	0	1

**Tableau IV.6 :** nombre de fausse acceptation (FAR) obtenu.

TH	X(f)	Y(f)	P(f)
0.0	98	100	74.5
0.1	97	99	69
0.2	95	98.5	64.5
0.3	92	97	60.5
0.4	81.5	95	58.5
0.5	76.5	90.5	52
0.6	70.5	85.5	43
0.7	61.5	75.5	30
0.8	47.5	58.5	18
0.9	23	35.5	8
1	0	0	0.5

**Tableau IV.7 :** Taux de fausse acceptation (FAR) obtenu.

Les performances du système sont évaluées à travers les courbes représentatives suivantes :



**Figure IV.13 :** Courbes de performance pour la coordonnée TF[X].

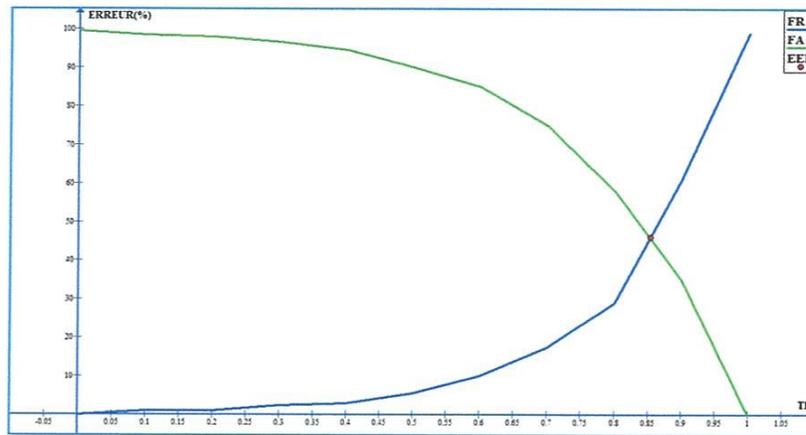


Figure IV.14 : Courbes de Performance pour la coordonnée TF[Y].

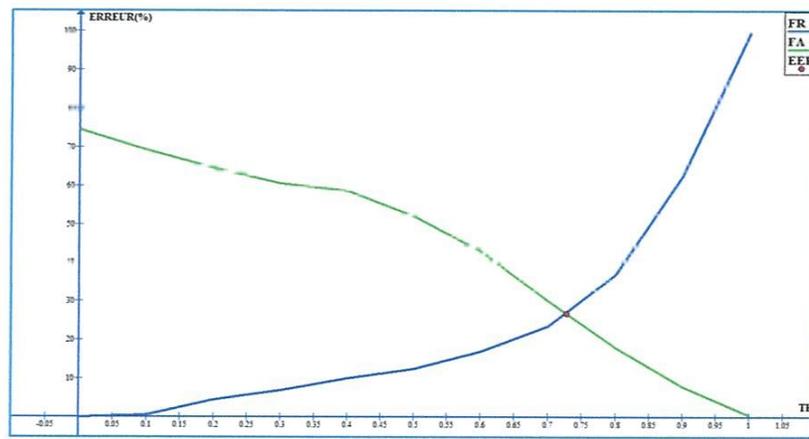


Figure IV.15 : Courbes de Performance pour la TF de la Pression.

Tableau pour estimation des seuils pour les paramètres X, Y, P accompagnés de l'ERR

Paramètres	Seuil estimé	EER (%)
TF [X]	0.8094	45.2031
TF [Y]	0.8527	45.8727
TF [P]	0.7255	26.9412

Tableau IV.8 : tableau des seuils estimés.

Paramètres	FRR (%)	FAR (%)	TER(%)
TF [X]	44.5	45	44.75
TF [Y]	42.5	46.5	44.5
TF [P]	27	27.5	27.25
TF [X], TF [Y]	20	24.5	22.25
TF [X], TF [P]	14	14.5	14.25
TF [Y], TF [P]	11.5	14.5	13
TF [X], TF [Y], TF [P]	7	8.5	7.75

**Tableau IV.9 :** Tableau récapitulatif faisant intervenir la combinaison des paramètres avec le FAR, FRR, EER qui leurs est associés.

Selon les résultats obtenus dans cette approche on remarque le meilleur taux obtenu est 92.25% lorsque nous avons employé tous les paramètres, on constate aussi que les paramètres géométrique X et P sont facile à falsifier par contre un imposteur ne peut pas falsifier le paramètre dynamique de la pression qu'il présente le meilleur paramètre pour révéler un imposteur.

## CONCLUSION :

Dans ce chapitre, nous avons présenté quelques tests effectués dans le but de fixer les paramètres de notre système et de dégager les meilleures performances de celui-ci. Pour ce faire nous avons utilisé une méthode fréquentielle basée sur l'utilisation de la DCT(FFT) et un critère de décision basé sur la déviation standard. Notre système de vérification de signature manuscrite a donné de bons résultats. Le meilleur résultat obtenu avec les paramètres X, Y et P est de 6.7% avec une erreur EER de 3,7057. Notons que nos tests ont été effectués sur deux bases de données MCVT et SVC 2004. Les tests ont été effectués sur un nombre réduit de signataires (100)

## CONCLUSION GENERALE ET PERSPECTIVES :

La biométrie est une technologie en plein essor. Elle est de plus en plus utilisée dans les applications en lien avec la sécurité vu les avantages qu'elle offre contrairement aux anciennes méthodes.

Dans ce mémoire, nous avons étudié un système biométrique de vérification de l'identité à base de la signature manuscrite en ligne. Après une présentation du contexte général de notre étude, les concepts fondamentaux des systèmes biométriques sont étudiés au premier chapitre. Nous avons présenté les systèmes de vérification de signature manuscrite en ligne, ses problématiques traités dans la littérature et les solutions proposées pour pallier aux difficultés rencontrées.

Dans le deuxième chapitre, nous avons évoqué les différentes méthodes de modélisation des signatures manuscrites rencontrées dans la littérature dans le but est de retenir une méthode de modélisation servant comme modèle retenue pour notre travail.

Les détails de l'approche retenue sont développés au troisième chapitre. Dans ce chapitre, deux contributions ont été proposées pour la réalisation de notre système de vérification de signature manuscrite en ligne. La première, consistait à utiliser les paramètres DCT qui ont donné de bons résultats. Ces paramètres mis en place ont un pouvoir discriminant supérieur à celui des paramètres FFT, en plus de leur gain en temps d'exécution. La deuxième, était la proposition d'une métrique pour l'évaluation des scores. Celle-ci est basée la déviation standard. Les résultats obtenus au quatrième chapitre sont satisfaisants. En effet, nous avons pu avoir des performances en termes d'EER qui avoisinent 3.7% sur des bases très utilisées dans la recherche. Afin de donner plus de valeur à notre travail, nous envisageons de tester notre système sur un grand nombre de signataires, au niveau des bases MCYT et SVC04. Cela nous permettra d'étudier la consistance de la discrimination des paramètres DCT par rapport au changement des données. En plus, l'exploitation des autres informations de la signature, azimuth et altitude, fournies dans ces bases, pourront apporter davantage d'amélioration à notre système. La classification des signatures est actuellement faite simplement par la déviation standard. L'emploi d'autres classificateurs comme le SVM (Support Vector Machine), la LDA (Linear Discriminant Analysis) ou les réseaux de neurones ont de forte possibilité d'améliorer notre classification.

# Bibliographies

- [BBO92] B. Boashash, *Time-Frequency Signal Analysis-Methods and Applications*, 1992.
- [BOUT11] Système biométrique de vérification de signatures manuscrites en ligne, thèse de magistère, Oued-Smar Alger.
- [BVF09] V. L. Blankers, C.E. Van den Heuvel, K.Y. Franke, L.G. Vuurpijl, The ICDAR 2009 Signature Verification Competition, 10th International Conference on Document Analysis and Recognition, IEEE, 2009.
- [CSH07] W. D. Chang, J. Shin, Modified Dynamic Time Warping for Stroke-based On-line Signature Verification, Proc. of the 9th International Conference on Document Analysis and Recognition, vol. 2, pp 724–728, Brazil, 2007.
- [FKO05] J. Fierrez-Aguilar, S. Krawczyk, Fusion of Local and Regional Approaches for On-Line Signature Verification, Springer, 2005.
- [FNL05] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Peñalba, J. Ortega-Garcia, D. Maltoni, An On-Line Signature Verification System Based on Fusion of Local and Global Information, in: T.Kanade, A. Jain, and N.K. Ratha, AVBPA 2005, LNCS 3546, pp. 523–532, Springer, 2005.
- [FRG92] Des gaussiennes pour la modélisation des signatures et la segmentation de tracés manuscrits, revue traitement du Signal, volume 9 n°4, P347-358, France, 1992.
- [GHL09] S. Garcia-Salicetti, N. Houmani, B. Ly-Van, et al, Online Handwritten Signature Verification, In: D. Petrovska-Delacrétaz et al., Guide to Biometric Reference Systems and Performance Evaluation, Springer, 2009
- [HAF16] Hafs.T & al. Empirical mode decomposition for online handwritten signature verification. IET Biometrics 5(3): 190-199 (2016)
- [IPI08] D. Impedovo, G. Pirlo, Automatic Signature Verification: The State of the art, IEEE transactions on systems, man, and cybernetics, part c: applications and reviews, vol. 38, no. 5, pp 609-635, September 2008.
- [ISI10] <http://www.isical.ac.in/~icfhr2010/CallforParticipation4NSigComp2010.html>
- [JCC05] Analyse des signaux non-stationnaires par transformation de Huang, Operateur de Teager-Kaiser, et Transformation de Huang-Teager (THT).thèse de doctorat Université Rennes 1, 2005.

- [JGC02] J. Ortega-Garcia, A. K. Jain, A. K. Jain, F. D. Griess, S. D. Connell, On-line Signature Verification, *Pattern Recognition* Vol. 35, pp. 2963 – 2972, 2002.
- [JMH81] JM.Hollerbach, An Oscillation Theory of Handwriting. *Biological Cybernetics*, 156:139–156, 1981.
- [JRO08] A. K. Jain, A. Ross, Introduction to Biometrics, in: *handbook of biometrics*, Springer, 2008.
- [KYA05] A. Kholmatov, B. A. Yanikoglu, Identity Authentication Using Improved Online
- [LGD99] V. D. Lecce, A. Guerriero, G. Dimauro, et al. Selection of Reference Signatures for Automatic Signature Verification, *Proc. of International Conference on Document Analysis and Recognition*, pp.597 –600, 1999.
- [LWA08] N. Liu, Y. Wang, Template Selection for On-line Signature Verification, *IEEE*, 2008
- [NEH04] N. E. H. a. Z.Wu, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society of London*, vol. 460, pp. 1597–1611, 2004.
- [NEH71] N. E. Huang, "The Empirical Mode Decomposition and the Hilbert spectrum for non- Linear and non-stationary time series analysis "Proceedings - Royal Society". *Mathematical, physical and engineering sciences* vol. 454, pp.903-995, 1971.
- [NLU08] L. Nanni, A. Lumini, A Novel Local On-line Signature Verification System, *Pattern Recognition Letters*, vol. 29, no. 5, pp. 559–568, 2008
- [OHI00] Ohishi. Y.K.T et Matsumoto.T, "On line Signature Verification Using Pen Position, Pen Pressure and Pen Inclination trajectories", *ICPR, International Conference on Pattern Recognition*, 2000, p. 547-550.
- [PFL93] P. Flandrin, "Temps-Fréquence," *Traité des Nouvelles Technologies, série Traitement du Signal*. Hermès, 1993.
- [Pla94] Plamondon.R et Parizeau.MA, "Signature verification from position, velocity and acceleration signals: a comparative study", *Progress in automatic signature verification*, New York, Editions World Scientific, 1994.
- [PLO89] R. Plamondon, G. Lorette, Automatic Signature Verification and Writer Identification: The state of the art, *Pattern Recognition*, vol. 22, no. 2, pp. 107-131, 1989.

**[SKR97]** C. Schmidt, K.F. Kraiss, Establishment of Personalized Templates for Automatic Signature Verification, proc. of International Conference on Document Analysis and Recognition, vol.1, pp. 263–267, 1997

**[VAN05]** B. L. VAN, Réalisation d'un Système de Vérification de Signature Manuscrite En ligne

**[WU98]** Wu. Q.Z, Lee. S.Y et Jou.I.C, “On Line Signature Verification Based ON Logarithmic Spectrum“, Pattern Recognition Letters, 1998, Vol 31, p. 1871-1998.

**[YCX04]** D. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, G. Rigoll. Svc2004: First international signature verification competition. Proc. of International Conference on Biometric Authentication (ICBA), pp. 16 – 22, Hong Kong, China, 2004.

**[YKH09]** B. Yanikoglu, A. Kholmatov, Fourier Descriptors for On-Line Signature Verification, Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing, Volume

# Annexe

## Exemples de signaux issus à partir de la base de données SVC2004 :

X	Y	P	Z	az	al	t
2288	7111	75748770	0	1310	680	244
2252	7058	75748780	1	1320	650	247
2207	6996	75748790	1	1300	630	247
2185	6857	75748800	1	1310	610	245
2178	6662	75748810	1	1290	600	287
2173	6429	75748820	1	1280	580	344
2173	6121	75748830	1	1280	570	425
2173	5747	75748840	1	1260	560	485
2178	5328	75748850	1	1250	550	515
2190	4901	75748860	1	1240	540	537
2206	4574	75748870	1	1220	540	554
2227	4314	75748880	1	1200	540	562
2252	4106	75748890	1	1170	550	568
2276	3984	75748900	1	1160	550	562
2290	3920	75748910	1	1130	560	565
2296	3920	75748920	1	1120	560	565
2286	3974	75748930	1	1100	560	548
2259	4070	75748940	1	1070	570	522
2220	4302	75748950	1	1060	570	504
2185	4615	75748960	1	1040	570	496
2156	4981	75748970	1	1030	570	490
2156	5417	75748980	1	1030	570	490
2177	5814	75748990	1	1030	570	490
2229	6186	75749000	1	1020	560	486
2307	6514	75749010	1	1020	560	449
2401	6762	75749020	1	1040	560	438
2507	6957	75749030	1	1050	560	440
2616	7082	75749040	1	1080	560	441
2727	7146	75749050	1	1110	550	441
2826	7170	75749060	1	1130	560	448
2927	7133	75749070	1	1150	550	448
3031	7019	75749080	1	1170	550	464
3134	6853	75749090	1	1190	540	496
3203	6662	75749100	1	1220	540	515
3251	6456	75749110	1	1230	530	542
3251	6241	75749120	1	1240	530	580
3204	6041	75749130	1	1250	530	586
3130	5854	75749141	1	1250	530	589
3025	5676	75749151	1	1250	530	592
2910	5545	75749161	1	1250	530	594

Figure : codage pour USER1\_1

X	Y	P	Z	az	al	t
1530	6223	75766286	0	1020	580	100
1518	6443	75766296	1	1030	580	369
1489	6498	75766306	1	1050	580	382
1462	6534	75766316	1	1080	580	369
1462	6534	75766326	1	1090	570	369
1488	6501	75766336	1	1100	560	381
1515	6393	75766346	1	1120	560	388
1542	6194	75766356	1	1130	560	438
1570	5934	75766366	1	1150	550	517
1597	5576	75766376	1	1160	550	566
1631	5167	75766386	1	1180	560	586
1669	4725	75766396	1	1200	550	585
1714	4319	75766406	1	1200	550	603
1764	3989	75766416	1	1210	560	617
1704	3710	75766426	1	1210	560	627
1799	3566	75766436	1	1200	570	631
1810	3504	75766446	1	1190	580	622
1810	4498	75766456	1	1160	580	613
1795	3533	75766466	1	1150	600	592
1765	3699	75766476	1	1130	600	572
1722	3953	75766486	1	1110	610	560
1680	4313	75766496	1	1110	610	570
1642	4706	75766506	1	1080	630	570
1642	5122	75766516	1	1080	630	564
1662	5522	75766526	1	1070	640	567
1715	5843	75766536	1	1050	640	567
1793	6112	75766546	1	1050	640	564
1888	6311	75766556	1	1050	640	563
1997	6464	75766566	1	1070	640	560
2118	6586	75766576	1	1090	640	570
2248	6638	75766586	1	1130	630	560
2368	6619	75766596	1	1150	630	578
2489	6553	75766606	1	1180	620	595
2609	6393	75766616	1	1190	610	611
2707	6185	75766626	1	1220	600	645
2790	5945	75766636	1	1240	600	669
2835	5689	75766646	1	1250	590	699
2843	5459	75766656	1	1270	590	721
2826	5246	75766666	1	1280	580	722
2776	5032	75766676	1	1270	570	721

Figure : codage pour USER1\_2

**Exemples de signaux issus à partir de la base de données SUSing :**

X	Y	P	Z	t
3235	2154	208	1	1
3244	2146	242	1	1
3254	2139	275	1	1
3265	2132	305	1	1
3276	2127	328	1	1
3288	2122	346	1	1
3288	2122	359	1	1
3304	2117	367	1	1
3304	2117	374	1	1
3316	2114	379	1	1
3316	2114	383	1	1
3316	2114	387	1	1
3316	2114	392	1	1
3316	2114	399	1	1
3317	2113	406	1	1
3316	2113	414	1	1
3295	2113	422	1	1
3289	2114	427	1	1
3237	2117	432	1	1
3200	2122	436	1	1
3157	2129	440	1	1
3109	2137	442	1	1
3056	2147	446	1	1
3000	2159	449	1	1
2941	2174	452	1	1
2883	2190	455	1	1
2829	2208	456	1	1
2779	2226	457	1	1
2735	2246	458	1	1
2695	2265	460	1	1

**Figure :** codage pour la base de données SUSing

