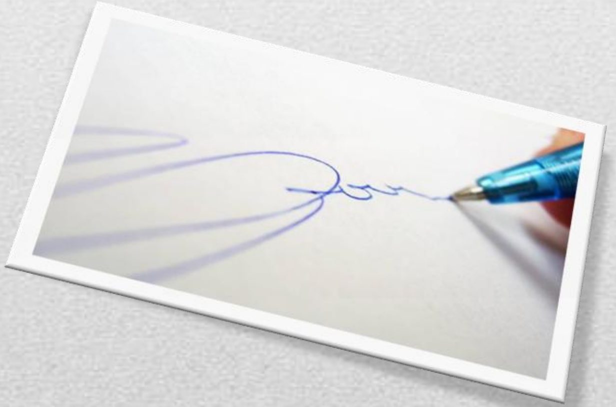
TRPA

Reconnaissance de signature



Caroline BEYNE - Thomas BOULANGER - Perrine BRUNEL - Pierre LE RENDU

Introduction

- **Plusieurs types de Reconnaissance de signatures**
  - Off-line ou statique
  - On-line ou dynamique

- Technique efficace pour reconnaitre les contrefaçons

Plan

- Prétraitements
- Distances
- Comparaison et classification
- Score et prise de décision

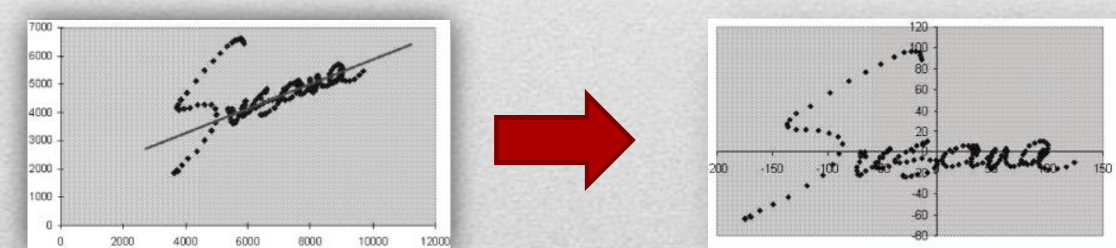
1. Prétraitements

- **Pourquoi ?**
  - Eliminer le bruit
  - Récupérer des données facilement comparables
  - Accélérer le traitement : Réduire le nombre de données
- **Les étapes**
  - Normalisation
  - Réduction du nombre de points

1.1. Normalisation

- **But ?**
  - Mettre toutes les signatures sur une même Norme
  - Plus faciles à comparer

- **Exemples?**
  - Rotation
  - Homothétie
  - Translation



1.2. Réduction des données

- **But?**
  - Supprimer les points inutiles

- **Différentes méthodes**
  - Algorithme génétique
  - Méthode de Brault
  - Sélection des points où la vitesse est minimum
  - Approximation polygonale

2. Distances

- Euclidienne

- Mahalanobis

- Temporelle

2. 1. Euclidienne

- Distance la plus simple et la plus utilisée
- Souvent utilisée pour le DTW
- **Définition**

$$d(X,Y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}$$

2. 3. Mahalanobis

- **Paramètres**
  - X et Y deux vecteurs à comparer
  - σ: matrice de covariance
- **Formule**

$$d(X,Y) = \sqrt{\sum_{i=1}^{p} \frac{(x_i - y_i)^2}{\sigma_i^2}}$$

- **Avantages**
  - Prend en compte la corrélation
  - Accorde un poids moins important aux composantes bruités

2. 4. Temporelle

P1: premier point de la signature 1

P1': premier point de la signature 2

- **Définition**
  - Formule: $d(P1, P2) = \left| t(P1) - t(P2) \right|$
  - Avec t(P) l'instant où P a été fait: t(P1) == t(P1')
- **Principe**
  - Permet de calculer le temps pour faire une signature

3. Comparaison et Classification

- HMM
- Réseau de neurones
- DTW
- **Notre proposition**

3.1. Notre Proposition

- Référence
- Algorithme
- Paramètres choisis
- Principe
- Distance

3.1.1. Référence

Pattern Recognition 42 (2009) 183 – 193

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

ELSEVIER

An efficient low cost approach for on-line signature recognition based on length normalization and fractional distances

Carlos Vivaracho-Pascual[a,*], Marcos Faundez-Zanuy[b], Juan M. Pascual[c]

[a]Departmento de Informática, Universidad de Valladolid, E.T.S.I. Informática, Campus Miguel Delibes, 47011 Valladolid, Spain
[b]Departament of de Telecomunicacions i Arquitectura de Computadors, Escola Universitària Politècnica de Mataró, Spain
[c]Universidad Europea Miguel de Cervantes, Spain

ARTICLE INFO

Article history:
Received 30 October 2007
Received in revised form 14 July 2008
Accepted 17 July 2008

Keywords:
On-line signature recognition
Signature normalization
Fractional distances
Dynamic time warping

ABSTRACT

This work presents a new proposal for an efficient on-line signature recognition system with very low computational load and storage requirements, suitable to be used in resource-limited systems like smart-cards. The novelty of the proposal is in both the feature extraction and classification stages, since it is based on the use of size normalized signatures, which allows for similarity estimation, usually based on dynamic time warping (DTW) or hidden Markov models (HMMs), to be performed by an easy distance calculation between vectors, which is computed using fractional distance, instead of the more typical Euclidean one, so as to overcome the concentration phenomenon that appears when data are high dimensional. Verification and identification tasks have been carried out using the MCYT database, achieving an EER (common threshold) of 6.6% and 1.8% with skilled and random forgeries, respectively, in the first task and 3.6% of error in the second. The proposed system outperforms DTW-based and HMM-based ones, even though these have proved to be very efficient in on-line signature recognition, with storage requirements between 9 and 90 times lesser and a processing speed between 181 and 713 times greater than the DTW-based systems.

1. Introduction

Automatic user identity recognition based on biometrics has became a focus of interest both for research and commercial purposes in recent years. Among the biometrics used, signature presents some advantages [1], such as for example, that it is widely accepted and commonly used in legal and commercial transactions as an authentication method. In addition, it is the second most important [2] of the behavioral biometrics [3], i.e., those biometrics based on measurements and data derived from an action performed by the user.

User recognition by means of his/her signature can be split into: (i) *static or off-line*, where the signature written on paper is digitized, (ii) *dynamic or on-line*, where users write their signature in a digitizing tablet, stylus-operated PDA, tablet-PC [4] or similar; the information acquired depends on the input device. Taking into account the highest security levels which can be achieved by dynamic systems, most of the efforts of the international scientific community are focused on this group [5,6]. Static systems are restricted to use in legal cases [7]. Our work is concerned with dynamic (on-line) signature recognition, so what follows will be mainly devoted to this category.

The different approaches that can be seen in the literature to extract relevant information can be broadly divided into [8]:

- Feature-based approaches, in which a holistic vector, consisting of *global* features (e.g., signature duration, standard deviation of x, etc.), is derived from the acquired signature trajectories.
- Function-based approaches, in which time sequences describing *local* properties (e.g., position trajectory, pressure, azimuth, etc.) are used.

The work shown in Ref. [1] will be used as reference one, since the performances of the main classifiers used in signature recognition are compared. Then, the same feature extraction, very simple *local* information, and the same experimental environment are followed. In the same way, identification and verification tasks have been approached, testing with skilled and random forgeries (imitated and not imitated impostor signatures) for the second task.

Focusing on the classification stage, different proposals can be found in the literature to measure the similarity between the claimed identity model and the input features. In the Signature Verification Competition 2004 (SVC04), dynamic time warping (DTW) [9]

* Corresponding author. Tel.: +34 983 423000x5618.
E-mail address: cevp@infor.uva.es (C. Vivaracho-Pascual).

0031-3203/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.07.008

« An efficient low cost approach for on-line signature recognition based on length normalization and fractional distances »

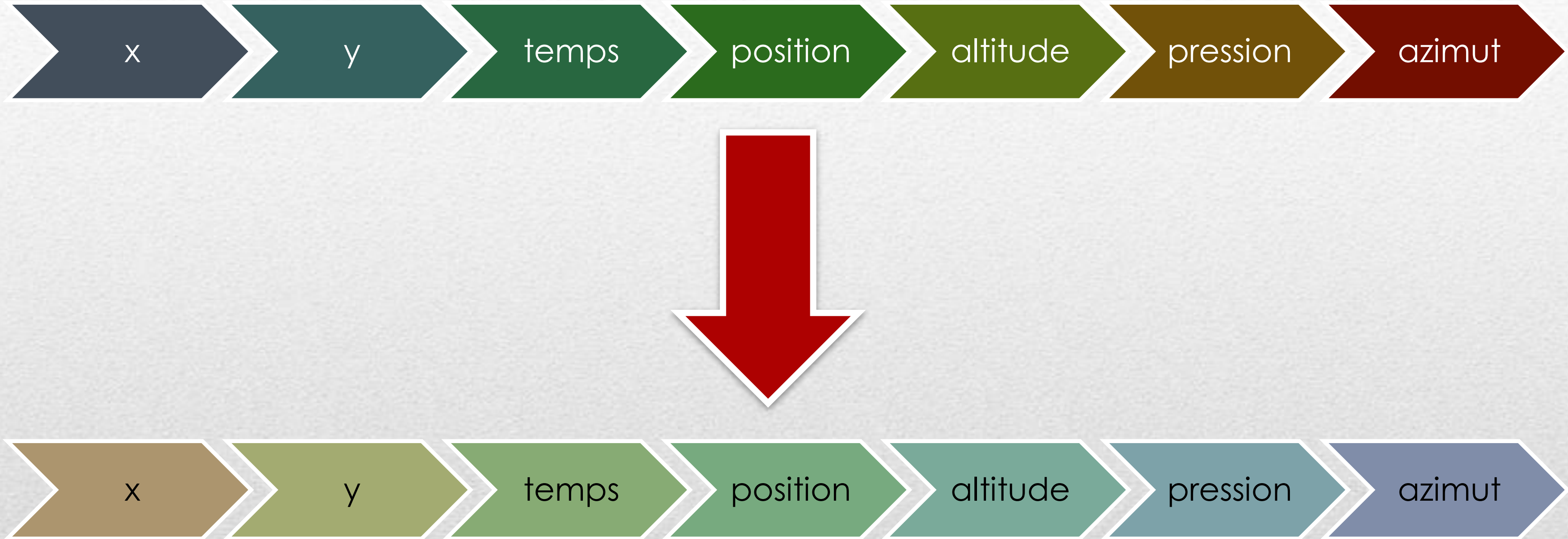
3.1.2. Algorithmes

- Ver_Sys :
  - Meilleurs résultats
  - « Skilled Forgeries » 5.2%
  - « Random » 1.8%
  - 9x moins d'espace mémoire
  - 328x + rapide

- Ver_Sys_Soft
  - Meilleures performances
  - « Skilled Forgeries » 5.4%
  - « Random » 1.6%
  - 90x moins d'espace mémoire
  - 1246x + rapide

- DTW
  - « Skilled Forgeries » 8.9%
  - « Random » 2.4%

3.1.3. Les paramètres choisis



- Pas de restriction
- Meilleure normalisation
- Meilleur modèle

3.1.4. Principe

4 étapes:

- La taille de la signature uniformisée
- La normalisation (par moyenne)
- La création du modèle
- La mesure de la distance

3.1.5. Taille et 3.1.6. Normalisation

- **Ver_Sys_Soft**
  - Taille : 30

$$S'_{jk} = \frac{\sum_{l=0}^{f_j - 1} S_{(i+l)k}}{f_j}$$

$S = \{S_0, \dots, S_{o-1}\}$

Vecteur origine

$\{I_0, \dots, I_{m-1}\}$

Intervalles /nb vecteurs

f_j Vecteur de I_j

k Composante k

$S' = \{S'_0, \dots, S'_{m-1}\}$

Vecteur normalisé

3.1.7. Modèle

- Modèle : **AT**
  - ***A**verage **T**emplate*
- Base d'entrainement
  - Fusion $\{tr_1,\dots,tr_n\} \rightarrow tr_i = \{x_1^i,\dots,x_m^i\} \rightarrow x_j^i = (x_{j1}^i,\dots,x_{jt}^i)$
  - $AT = \{\bar{x}_1,\dots,\bar{x}_m\} \rightarrow \bar{x}_j = \left(\frac{\sum_{i=1}^n x_{j1}^i}{n},\dots,\frac{\sum_{i=1}^n x_{jt}^i}{n}\right)$
- Exemplaire de test : ***te***
- Distance: $\boxed{d(AT,te)}$

3.1.8. Distance

- Espace **L^p**
- p-norme fractionnelle
- 0 < p < 1
  - Meilleurs résultats pour 0.2 < p < 0.4

$$\|Z\|_p = \left(\sum_i |z_i|^p \right)^{1/p}$$

- **Ver_Sys_Soft**
  - p = 0.2

$$d(y,z) = \left(\sum_{j=1}^{m} \sum_{l=1}^{t} |y_{jl} - z_{jl}|^p \right)^{1/p}$$

3.1.9. Récapitulatif Ver_Sys_Soft

- La taille normalisée à **30** vecteurs
- La normalisation (**par moyenne**) de chaque vecteur

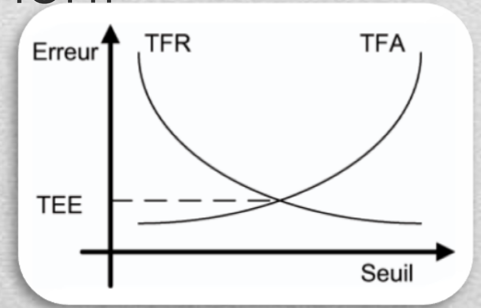
$$s'_{jk} = \frac{\sum_{l=0}^{f_j - 1} s_{(i+l)k}}{f_j}$$

- Le modèle **AT** : fusion des signatures d'apprentissage
- La mesure de la distance

$$d(AT, te) = \left(\sum_{j=1}^{m} \sum_{l=1}^{t} \left| AT_{jl} - te_{jl} \right|^{p} \right)^{1/p}$$

4. Score et prise de décision

- **Calcul du score**
  - score = 100 * (-seuil - log(distance))
- **Prise de décision**
  - log pour $[0;+\infty[\Rightarrow]-\infty;+\infty[$
  - Seuil déterminé expérimentalement
  - -seuil pour recentrer sur 0
    - « VRAI / TRUE » > 0
    - « FAUX / FALSE » < 0



Conclusion

- **Ver_Sys_Soft**
  - Rapide
  - Efficace
- **Score et prise de décision**
  - Seuil: -34
- **Résultats:**
  - Elimination des contrefaçons
  - Reconnaissance des authentiques

Questions ?

?