Noctis

A Statistical Signature Recognition System

Andrea Sergas
a.sergas@studenti.uniba.it

Donato Meoli
d.meoli1@studenti.uniba.it

June 12, 2017

University of Bari "Aldo Moro"
Department of Computer Science

Table of Contents

Overview

Signature as Biometric Authentication

Biometric authentication is gaining popularity as a more trustable alternative to password-based security systems. Signature is a behavioral biometric: it is not based on the physical properties, such as fingerprint or face, of the individual, but behavioral ones. As such, one's signature may change over time and it is not nearly as unique or difficult to forge as iris patterns or fingerprints, however signature's widespread acceptance by the public make it more suitable for certain lower-security authentication needs.

Online & Offline Signature Verification

Signature verification is split in two approaches, according to the input type of data:

- **offline** (static) signature verification takes as input the image of a signature; it is useful in automatic verification of signatures found on bank checks and documents;
- **online** (dynamic) signature verification uses signatures that are captured by pressure-sensitive tablets that extract dynamic properties of a signature, in addition to its shape.

Online Signature Verification

Dynamic features include the overall speed of the signature, the pen pressure at each point and other similar data. They make the signature more unique and difficult to forge. As a result, online signature verification is more reliable than offline signature verification.

In an online signature verification system, users are first enrolled by providing signature samples, called reference signatures. Then, when a user presents a signature claiming to be a particular individual, the test signature is compared with the reference signatures for that individual. If the similarity is under a certain threshold, the user is rejected, otherwise authenticated.

Introduction

SUSIG Dataset

For the train and test phases of the system it has been used an on-line handwritten signature database called SUSIG, which collects signatures for 94 person classified in two classes:

- 20 **genuine** signatures for user divided in two sessions:
  - *session 1* for the train phase;
  - *session 2* for the test phase.

  and vice versa in the 2-fold cross-validation approach used.
- 10 **forgery** signatures for user divided in two classes:
  - *skilled* forgery is signed by a person who has had access to a genuine signature for practice;
  - *highly skilled* forgery is signed is signed by a signatures' forgery professionist who has had access to a genuine signature for practice.

Noctis

Noctis - *Statistical Signature Recognition System* - is a recognition system for online handwritten signature verification based on a statistical approach.

Why ”Statistical”?

Features Extraction I

For each signature of the user, the basic data which are provided by SUSIG dataset are:

- x coordinate: scaled cursor position along the x-axis
- y coordinate: scaled cursor position along the y-axis
- timestamp: system time at which the event was posted
- pressure: adjusted state of the normal pressure in milliseconds

while, starting with these, the following features are extracted:

- Displacement:

$$d_i = \begin{cases} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} & 1 \leq i \leq n-1 \\ d_n - d_{n-1} & i = n \end{cases}$$

Features Extraction II

- Velocity:

$$v_i = \begin{cases} \frac{d_i}{t_{i+1} - t_i} & 1 \leq i \leq n-1 \\ v_n = v_{n-1} & i = n \end{cases}$$

- Acceleration:

$$a_i = \begin{cases} \frac{v_i}{t_{i+1} - t_i} & 1 \leq i \leq n-1 \\ v_n = v_{n-1} & i = n \end{cases}$$

- Angle between two consecutive points:

$$an_i = \begin{cases} \arctan(y_{i+1} - y_i, x_{i+1} - x_i) & 1 \leq i \leq n-1 \\ an_n = an_{n-1} & i = n \end{cases}$$

- Angle between a point and the origin:

$$an_{0_i} = \arctan(y_i - 0, x_i - 0) \quad 1 \leq i \leq n$$

Features Extraction III

- Pressure variations between two consecutive points:

$$\delta p_i = p_{i+1} - p_i$$

- Time variations between two consecutive points:

$$\delta t_i = t_{1+1} - t_i$$

- Total time:

$$t = \sum_{i=1}^{n} t_i$$

- Total path length:

$$l = \sum_{i=1}^{n} d_i$$

where n is the number of the samples for each data in the user's signatures representation.

Train Phase

Signature Representation

For each signature the features extracted are represented as a vector of local and global statistical features:

$$S_i = \langle \mu(d_i),\ \mu(v_i),\ \mu(a_i),\ \mu(\delta p_i),\ \mu(\delta t_i),\ \mu(an_i),\ \mu(an_{0_i}),$$

$$\sigma(d_i),\ \sigma(v_i),\ \sigma(a_i),\ \sigma(\delta p_i),\ \sigma(\delta t_i),\ \sigma(an_i),\ \sigma(an_{0_i}),\ t,\ l \rangle$$

where μ is the **mean** - *centrality* index - and σ is the **standard deviation** - *dispersion* index.
Only mean and standard deviation are considered to make a drastic reduction of redundant data to pass to the classifier by reducing each feature to only one significant and discriminating value.

Test Phase

User's SVM

For the users' signatures modeling phase it has been used one Support Vector Machine for each user. The train set for each SVM has been mostly auto generated: from the 10 genuine signatures' features vectors other 80 features vectors have been created by adding small noises, in order to maintain consistency.
No forgery signatures' features have been used in the train set.

User's Threshold

In order to set a specified threshold for each user right after the train phase 90 features vectors have been created by multiplying each feature used in the train set by the variance of that signature's features, excluding total time and total path.
The user's threshold is then obtained by getting the minimum score predicted from the trained SVM given the 90 auto-generated features vectors. A small constant is then added: the obtained threshold represents an average forgery signature's score.

Results & Conclusion

Evaluation

In evaluating the performances of a signature verification system, there are two important factors:

- **False Acceptance Rate** of forgery signatures which is defined as:

$$FAR = \frac{Number\ of\ accepted\ forgery\ attempts}{Total\ number\ of\ forgery\ attempts}$$

- **False Rejection Rate** of genuine signatures which is defined as:

$$FRR = \frac{Number\ of\ rejected\ legitimate\ attempts}{Total\ number\ of\ legitimate\ attempts}$$

Feature Development

This recognition system can be further improved by following two approaches, independent one from the other:

- using technologically advanced devices that provide **more features** about signed signatures such as:
  - *azimuth*: clockwise rotation of cursor about the z-axis;
  - *altitude*: angle upward toward the positive z-axis;
- using **hyper-parameterization algorithms** which prevent the phenomenon of overfitting through the use of the learning rate in a preventive manner: the system, at each step of optimizing its loss function, declines the learning rate.

License

Get the source of this software from:

github.com/DonatoMeoli/SSRS

github.com/MrNobody1992/Noctis

The End