TRAINING ON-LINE HANDWRITING RECOGNIZERS USING SYNTHETICALLY GENERATED TEXT

Daniel Martín-Albo[†], Réjean Plamondon[*] and Enrique Vidal[†]

[†] PRHLT Research Center
Universitat Politècnica de València

[*] Laboratoire Scribens
École Polytechnique de Montréal

ICFHR — CRETE, SEPTEMBER 2014

Motivation

- On-line handwriting recognition systems are very common nowadays.
- Good results can be achieved in writer-independent systems, but performance degrades if the user writing style is different from training.
- Possible solution: writer adaptation. Tune writer-independent models using writer-specific data.
- Trade-off: Amount of new training data requested versus performance.
- Proposal: Generate synthetic human-like handwritten words from real samples using the Kinematic Theory of rapid human movements.

Kinematic Theory (Σ-Lognormal)

- The Kinematic Theory can be used to analyze rapid human movements.
- Relies on lognormals for describing the velocity response of a neuromuscular network.
- The goal is to decompose a complex movement into a sequence of simpler strokes:

$$\vec{v}(t) = \sum_{i=1}^{N} D_i \begin{bmatrix} \cos\left(\phi_i(t)\right) \\ \sin\left(\phi_i(t)\right) \end{bmatrix} \Lambda_i(t; t_{0_i}, \mu_i, \sigma_i^2)$$

- Each stroke is described by six parameters:

$$p_i = (D_i,\, t_{0_i},\, \mu_i,\, \sigma_i,\, \theta_{s_i},\, \theta_{e_i})$$

Generation of Synthetic Samples

- Relationship between the fluctuations of the Σ-lognormal parameters and the handwritten variability.

- The synthetic generation algorithm presents three different stages:

1. Extraction of Σ-lognormal parameters.
2. Add noise (n) to the Σ-lognormal parameters.

$$n = [0, 0, n_\mu, n_\sigma, 0, 0] \qquad \begin{aligned} n_\mu &\in \mathcal{U}(-0.15\,\mu, 0.15\,\mu) \\ n_\sigma &\in \mathcal{U}(-0.15\,\sigma, 0.15\,\sigma) \end{aligned}$$

3. Recalculate the velocity profile and recover $(x,\ y)$.

$$x(t) = \int_0^t v_x(\tau)\, d\tau \qquad y(t) = \int_0^t v_y(\tau)\, d\tau$$

Σ-Lognormal Parameter Extraction

- We used the Robust Xzero [1] based algorithm to estimate the Σ-lognormal parameters.

- The quality of the calculated parameters is expressed using the signal-to-noise ratio (SNR).

$$\mathrm{SNR} = 10\ \log\left(\frac{\int v_{x_n}^2 + v_{y_n}^2\, dt}{\int (v_{x_n} - v_{x_\Sigma})^2 + (v_{y_n} - v_{y_\Sigma})^2 dt}\right)$$

[1] C. O'Reilly and R. Plamondon, "Development of a Sigma-lognormal representation for on-line signatures", Pattern Recognition, vol. 42, no. 12, pp. 3324 – 3337, 2009, new Frontiers in Handwriting Recognition.

Some Synthetic Samples

Solid: Original Sample
Dashed: Synthetic Sample

Some Synthetic Samples (II)

Solid: Original Sample
Dashed: Synthetic Sample

Experiments

1. Assess the quality of the reconstruction using the Robust Xzero algorithm.

2. Study the impact of increasing the training dataset with

   - a varying number of real writer-specific samples

   - and a varying number of synthetic samples.

Experimental Setup

- Unipen-ICROW-03 database has been used in the experiments.
  - Composed of 13,119 isolated words from 72 people.
  - 10,496 words from 56 writers were selected as training set (**trn**).
    - From the remaining writers, 70% of their words were randomly chosen as test set (**tst**) leaving 30% as adaptation set (**adp**).
- 7 features representing pen positions, writing-speeds and curvature.
- Closed 1-gram word language model consisting of 884 unique lexical words.
- We evaluate the results by means of the Recognition Error Rate (% RER).

Validation of the Reconstruction

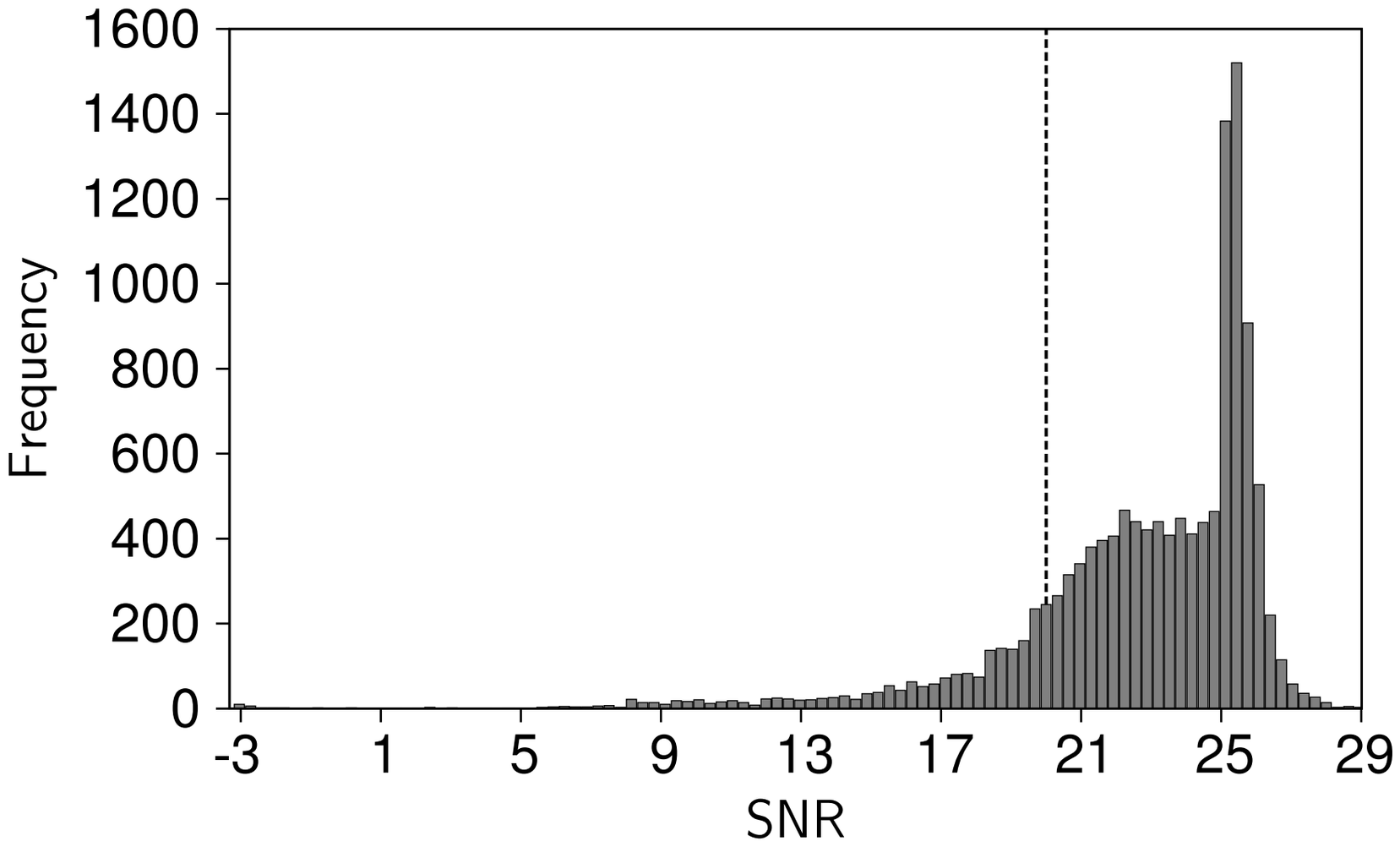
- Determine whether the technique provides a correct reconstruction for handwritten recognition purposes.

- Recognize the words contained in **tst** using **trn** as training

  a) employing the original samples. (10.3% RER)

  b) replacing each word by its reconstruction. (12.6% RER)

Validation of the Reconstruction (II)



SNR histogram for the reconstructed words of **tst** & **trn**

Impact of Synthetic Samples

- For each writer, recognize his words on **tst** using **trn** as training and

  - a varying number of real words (w) picked from his words in **adp**
  - and a varying number of synthesized (or replicated) words (s) created from the previously chosen.

$w \in \{20, 35, 50\}$ $s \in \{10, 20, 50, 150, 200, 250\}$

Impact of Synthetic Samples (II)

# writer-specific samples (w)	# replicated or synthetic samples (s)					
	10	20	50	100	150	200
20	11.7	11.4	10.2	9.5	9.5	8.9
35	10.8	10.2	9.4	8.2	8.2	8.6
50	10.4	9.8	8.7	8.1	8.5	8.5
20	11.0	10.4	9.5	8.9	8.5	8.4
35	9.9	9.1	7.9	7.0	6.6	6.4
50	9.6	8.8	7.6	7.0	6.9	6.9

Test set Recognition Error Rate (%)

Top: Using only real replicated adaptation samples.
Bottom: Using both real and synthetic adaptation samples.

Summary

- A method for the generation of synthetic on-line samples has been evaluated on handwritten words.

- The distortion of the Σ-lognormal parameters introduces realistic variability.

- Only a small number of real writer-specific samples are needed.

- The use of synthetic samples allows to improve the performance of a writer-specific system.

Future Work

- Compare with different writer adaptation techniques.
- Test with more corpora.
- ...