Total Whitening for Online Signature Verification Based on Deep Representation

Communication Science Laboratories, NTT Corporation, Japan
Xiaomeng Wu, Akisato Kimura, and Kunio Kashino

Kyushu University, Japan
Seiichi Uchida
Introduction

• Deep Metric Learning for Online Signature Verification
  – Learn feature representations (via CNNs or RNNs) that can be easily separated using simple distance functions (such as the $l_2$-distance)

• Correlation between Feature Dimensions
  – Easily enlarged through highly nonlinear neural networks
  – Lead to suboptimal feature matching effectiveness

• Whitening
  – Assume that the features come from a Gaussian distribution
  – Learn a feature transformation that converts this Gaussian into a standard normal distribution

Feature Transformation (Whitening)
Three Types of Gaussian Distributions

- **Metric Gaussian** [K.M., ICCV07][F.R., ECCV16][F.R., TPAMI19]
  - Distribution of features in all signatures of all subjects

- **Class Gaussian** [R.M., ICPR96][F.A.-F., ICIP09][Y.K., ACPR13]
  - Distribution of features in all signatures of each particular subject

- **Instance Gaussian** [Y.L., NeurIPS17]
  - Distribution of features in each particular signature
Idea

• **Total Whitening**
  – *A novel method that embraces multiple whitening* derived from metric, class, and instance Gaussians

• **Assumption**
  – The local features of a signature are drawn from a total Gaussian (a combination of the multiple Gaussians)
  – The total Gaussian is somewhat consistent for signatures of the same subject, but different for signatures of different subjects
  – Transforming the total Gaussian to a standard normal distribution may reduce intra-class variation and enhance inter-class variation
Deep Online Signature Verification

Training Data

subject_1

sig_{1,1}  sig_{1,2}  \ldots  sig_{1,n}

subject_2

sig_{2,1}  sig_{2,2}  \ldots  sig_{2,n}

\ldots

subject_m

sig_{m,1}  sig_{m,2}  \ldots  sig_{m,n}

Matching Pairs (Same Subjects)

sig_{1,1}  sig_{1,1}  \ldots  sig_{1,1}  sig_{1,2}  \ldots

sig_{1,2}  sig_{1,3}  \ldots  sig_{1,n}  sig_{1,3}  \ldots

Non-matching Pairs (Different Subjects)

sig_{1,1}  sig_{1,1}  \ldots  sig_{1,n}  sig_{2,1}  \ldots

sig_{2,1}  sig_{2,2}  \ldots  sig_{m,n}  sig_{3,1}  \ldots

Training Signature

Siamese Network

Local Features

Contrastive Loss

Training Signature

Siamese Network

Local Features
Deep Online Signature Verification

Reference Data

subject_2

sig_{2,1}  sig_{2,2}  ...  sig_{2,n}

Test Data

test signature

Matching

e.g., mean distance & subject-independent threshold classifier

Reference Signature

Siamese Network

Local Features

Distance

Test Signature

Siamese Network

Local Features
The two signatures are aligned in time with DTW to remove temporal distortion.

Each pair of local features at the same timestamp can be regarded as a correspondence.
Metric Gaussian

Training Data

subject_1

<table>
<thead>
<tr>
<th>[ \sigma_{1,1} ]</th>
<th>[ \sigma_{1,2} ]</th>
<th>\ldots</th>
<th>[ \sigma_{1,n} ]</th>
</tr>
</thead>
</table>

subject_2

<table>
<thead>
<tr>
<th>[ \sigma_{2,1} ]</th>
<th>[ \sigma_{2,2} ]</th>
<th>\ldots</th>
<th>[ \sigma_{2,n} ]</th>
</tr>
</thead>
</table>

...  

subject_m

<table>
<thead>
<tr>
<th>[ \sigma_{m,1} ]</th>
<th>[ \sigma_{m,2} ]</th>
<th>\ldots</th>
<th>[ \sigma_{m,n} ]</th>
</tr>
</thead>
</table>

Matching Pairs (Same Subjects)

<table>
<thead>
<tr>
<th>[ \sigma_{1,1} ]</th>
<th>[ \sigma_{1,1} ]</th>
<th>\ldots</th>
<th>[ \sigma_{1,1} ]</th>
<th>[ \sigma_{1,2} ]</th>
<th>\ldots</th>
<th>[ \sigma_{1,3} ]</th>
<th>\ldots</th>
<th>[ \sigma_{1,n} ]</th>
</tr>
</thead>
</table>

Set of Local Feature Pairs \( (S_1) \)

Metric Mean & Metric Covariance

\[
\mu_1 = \frac{1}{2n_1} \sum_{S_1} (x + y)
\]

\[
\Sigma_1 = \frac{1}{n_1} \sum_{S_1} (x - y)(x - y)^T
\]
Class Mean & Class Covariance

\[
\mu_2 = \frac{1}{2n_2} \sum_{S_2} (x + y)
\]

\[
\Sigma_2 = \frac{1}{n_2} \sum_{S_2} (x - y)(x - y)^T
\]

One dataset has only one metric Gaussian, but has multiple class Gaussians, one for each subject.
Instance Gaussian

Reference Data / Test Data ($S_3$)

Instance Mean & Instance Covariance

$$\mu_3 = \frac{1}{n_3} \sum_{S_3} x$$

$$\Sigma_3 = \frac{1}{n_3} \sum_{S_3} (x - \mu_3)(x - \mu_3)^\top$$

One dataset has multiple instance Gaussians, one for each reference signature / test signature.
Total Whitening

Total Mean & Total Covariance

\[
\mu = \sum_{i=1}^{3} \omega_i \mu_i
\]

\[
\Sigma = \sum_{i=1}^{3} \omega_i \left[ \Sigma_i + (\mu_i - \mu)(\mu_i - \mu)^T \right]
\]

Center of multiple Gaussians

Relative positional relationship between multiple Gaussians

Total Gaussian

Total Whitening
Experimental Protocol

• **MCYT-100 (90/80/70/60/50%)**
  – 100 subjects: each has 25 genuine signatures and 25 skilled forgeries
  – The first 90/80/70/60/50% of subjects were used for training; the remaining subjects for testing.

• **FULL**
  – We combined three datasets (MCYT-100, Biosecurid SONOF, and SUSIG): the first 90% of subjects in each dataset were combined for training; all the remaining subjects were used for testing
  – The training set contains 375 subjects and 11,944 signatures; the testing set contains 41 subjects and 1,312 signatures

• **Evaluation Metric**
  – For each subject in the testing set, the first five genuine signatures were used as reference signatures
  – All the remaining signatures were used as test signatures
  – Equal Error Rate (EER)
### Results

<table>
<thead>
<tr>
<th>Method</th>
<th>MCYT (90%)</th>
<th>MCYT (80%)</th>
<th>MCYT (70%)</th>
<th>MCYT (60%)</th>
<th>MCYT (50%)</th>
<th>FULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Whitening [X.W., ICASSP19]</td>
<td>0.50</td>
<td>2.50</td>
<td>2.40</td>
<td>2.50</td>
<td>4.50</td>
<td>3.75</td>
</tr>
<tr>
<td>Metric Whitening</td>
<td>0.50</td>
<td>1.50</td>
<td>2.40</td>
<td>2.00</td>
<td>2.72</td>
<td>1.54</td>
</tr>
<tr>
<td>Class Whitening</td>
<td>0.50</td>
<td>1.50</td>
<td>4.00</td>
<td>4.00</td>
<td>4.30</td>
<td>2.11</td>
</tr>
<tr>
<td>Instance Whitening</td>
<td>0.50</td>
<td>1.00</td>
<td>2.17</td>
<td>2.50</td>
<td>2.90</td>
<td>2.11</td>
</tr>
<tr>
<td>Total (Metric &amp; Class)</td>
<td>0.50</td>
<td>1.20</td>
<td>3.47</td>
<td>3.20</td>
<td>3.30</td>
<td>1.73</td>
</tr>
<tr>
<td>Total (Metric &amp; Instance)</td>
<td>0.50</td>
<td>1.00</td>
<td>1.67</td>
<td>1.80</td>
<td>2.20</td>
<td>1.73</td>
</tr>
<tr>
<td>Total (Class &amp; Instance)</td>
<td>0.50</td>
<td>1.00</td>
<td>1.67</td>
<td>1.60</td>
<td>2.00</td>
<td>1.34</td>
</tr>
<tr>
<td>Total (Metric, Class &amp; Instance)</td>
<td>0.50</td>
<td>1.00</td>
<td>1.67</td>
<td>1.75</td>
<td>2.20</td>
<td>1.54</td>
</tr>
</tbody>
</table>

As long as the instance Gaussian was incorporated, total whitening was superior to all the baselines in almost all cases.

The class whitening was not very useful on its own, but it was effective when combined with the instance Gaussian.
Examples (1/2)

Two local features (top-white⁴ and bottom-black⁵ circles) in different signatures of different subjects

No Whitening¹

Class Whitening¹

Total Whitening¹

No Whitening²

Class Whitening²

Total Whitening²
Examples (2/2)

Two local features (top-white \(^1\) and bottom-black \(^2\) circles) in different signatures of different subjects

No Whitening \(^1\)

Instance Whitening \(^1\)

Total Whitening \(^1\)

No Whitening \(^2\)

Instance Whitening \(^2\)

Total Whitening \(^2\)
## EERs Published on MCYT

<table>
<thead>
<tr>
<th>Method</th>
<th>#Subject</th>
<th>#Reference</th>
<th>EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fierrez-Aguilar (2007)</td>
<td>145</td>
<td>10</td>
<td>3.36</td>
</tr>
<tr>
<td>Argones-Rua (2012)</td>
<td>100</td>
<td>10</td>
<td>2.85</td>
</tr>
<tr>
<td>Yanikoglu (2009)</td>
<td>100</td>
<td>5</td>
<td>7.80</td>
</tr>
<tr>
<td>Vivaracho-Pascual (2009)</td>
<td>280</td>
<td>5</td>
<td>6.60</td>
</tr>
<tr>
<td>Faundez-Zanuy (2007)</td>
<td>280</td>
<td>5</td>
<td>5.42</td>
</tr>
<tr>
<td>Nanni (2008)</td>
<td>100</td>
<td>5</td>
<td>5.20</td>
</tr>
<tr>
<td>Cpalka (2016)</td>
<td>100</td>
<td>5</td>
<td>4.88</td>
</tr>
<tr>
<td>Sae-Bae (2014)</td>
<td>100</td>
<td>5</td>
<td>4.02</td>
</tr>
<tr>
<td>Tang (2018)</td>
<td>100</td>
<td>5</td>
<td>3.16</td>
</tr>
<tr>
<td>Wu (2019)</td>
<td>50</td>
<td>5</td>
<td>4.50</td>
</tr>
<tr>
<td>* Total Whitening (Class &amp; Instance)</td>
<td>50</td>
<td>5</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Note: We need to use part of MCYT dataset to train our network, so our EER was obtained with only the testing set containing fewer subjects.