Generating Handwritten Character Clones from an Incomplete Seed Character Set using Collaborative Filtering

Kazuaki Nakamura, Eiji Miyazaki, Naoko Nitta, and Noboru Babaguchi

Osaka University, Japan

on 6th Aug. 2018, at ICFHR2018
Research Background

• Handwriting generation
  • generate synthetic (or clone) images of handwritten characters resembling a target user’s actual handwriting.

Target user

Training dataset
- handwriting character images
- pen stroke (sequence of 2D pen-tip locations)

Generator (e.g. auto-encoders, GANs)

Automatically generated HCCs
- are applicable to communication tools (especially for hand-impaired people).
- serve a large scale dataset for handwriting recognition, faked signature detection, etc.
Requirements in Practice

• **Incomplete seed character set**

  • *Seed characters*: characters whose image(s) is in the training dataset
  
  • Seed characters are usually limited because it is difficult to collect a lot of images from the target user.
    • It is not rare that at most one or zero image is available per character, especially in the case of Asian languages.

• **Within-person variety**

  • Images of humans’ actual handwriting differ from each other even if the same writer writes the same character.

  ![Examples of handwriting](image)

  All of them have the similar characteristics but are slightly different from each other.
Goal

- **Goal:** To propose a HCC generation method
  - that can achieve the *within-person variety*
    Not a single HCC but its distribution should be created.
  - based on the *incomplete seed character set*
    At most one or zero image is available per character
    as a training data.
Related Work

• (Conventional) HCC generation [1, 2]

**INPUT**
A lot of images for each character

\[
\begin{array}{ccc}
\alpha & \alpha & \cdots \\
\beta & \beta & \cdots \\
\vdots & & \\
\gamma & \gamma & \cdots \\
\end{array}
\]

**OUTPUT**
HCC distribution for each character

\[
\begin{array}{c}
p(HCC|\alpha) \\
p(HCC|\beta) \\
\vdots \\
p(HCC|\gamma) \\
\end{array}
\]

- Incomplete seed character set
- within-person variety

• Font generation [3, 4, 5]

**INPUT**
A few images of several seed characters written in a certain style

\[
\begin{array}{ccc}
\alpha & n & c \\
\beta & e & x \\
\gamma & & \\
\end{array}
\]

**OUTPUT**
Images of the other characters that seem to be written in the same style

\[
\begin{array}{ccc}
\beta & d & x \\
\vdots & & \\
\gamma & x & \\
\end{array}
\]

- Incomplete seed character set
- within-person variety

---

Overview

- Character-wise HCC generation
- Offline (using only images)

**Training dataset**
- offered by the target user \(u\)
- incomplete seed character set

**Side dataset**
- collected from many other writers \(w_1, \ldots, w_J\)
- Each writer offers only a few images (e.g. an image per character)

**Encoder**
- feature extractor

**HCC**

**Decoder**
- Sample a new feature \(f \sim p(f | \theta)\)

**Parameter selection**
- select the parameter that is best-fit to \(\{f_u\}\)

**Parameter pool**

\[\theta_1, \theta_2, \theta_3, \ldots, \theta_K\]
Parameter Pool Construction

Hypothesis
- It is not rare that the shapes of two writer’s handwriting are very similar for some characters. IOW, there are a lot of writer-pairs whose handwriting shapes are similar for some characters. Not only the average shape but also the shape distribution of their handwriting would be similar.

Separately perform the following procedure for each character \( c \):
1. Extract a feature for each handwriting image in the side dataset
2. Cluster a set of the extracted features
3. Compute the mean \( t_k^c \) and the covariance \( \Sigma_k^c \) for each cluster \( k \) \( \Rightarrow \theta_k = (t_k^c, \Sigma_k^c) \)
For a seed character $c$,
- use the target user’s actual handwriting image $I_u^c$.
- select the parameter that is best-fit to $I_u^c$. 

BestFit strategy

$\widehat{\theta}^c = \theta_3^c$
For a non-seed character $c'$,
- there are no images of the target user’s actual handwriting.
- \textbf{BestFit} strategy cannot be used.
Parameter Selection for Non-seeds

• For a non-seed character $c'$, employ collaborative filtering (CF).
• To perform CF, first construct a writer-character matrix $\Phi$.
  • Estimate the best-fit parameters for not only the target user but also the other writers.
  • $\phi_{jm} \in \{1,2,\ldots,K\}$: ID of the best-fit parameter of $j$-th writer’s feature distribution for $m$-th character

\[
\Phi = \begin{bmatrix}
    \phi_{11} & \phi_{12} & \cdots & \phi_{1m} & \cdots & \phi_{1M} \\
    \phi_{21} & \phi_{22} & \cdots & \phi_{2m} & \cdots & \phi_{2M} \\
    \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
    \phi_{J1} & \phi_{J2} & \cdots & \phi_{Jm} & \cdots & \phi_{JM} \\
    \phi_{u1} & \phi_{u2} & \cdots & ? & \cdots & \phi_{uM} \\
\end{bmatrix}
\]

$\phi_{u1}, \phi_{u2}, \phi_{uM}$: known (estimated by Best-Fit strategy)
$\phi_{u,m}$: unknown  →  try to estimate it by collaborative filtering!
Collaborative Filtering

• User-based collaborative Filtering (UserCF)

Hypothesis

If the feature distributions of two writers are similar with each other for some characters, their distributions for another character also tend to be similar.

<table>
<thead>
<tr>
<th>similar writers</th>
<th>(c_1)</th>
<th>(c_2)</th>
<th>(c_3)</th>
<th>(c_4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(w_1)</td>
<td>(\phi_{11})</td>
<td>(\phi_{12})</td>
<td>(k)</td>
<td>(\phi_{14})</td>
</tr>
<tr>
<td>(w_2)</td>
<td>(\phi_{21})</td>
<td>(\phi_{22})</td>
<td>(k)</td>
<td>(\phi_{24})</td>
</tr>
<tr>
<td>(w_3)</td>
<td>(\phi_{31})</td>
<td>(\phi_{32})</td>
<td>(\phi_{33})</td>
<td>(\phi_{34})</td>
</tr>
<tr>
<td>(u)</td>
<td>(\phi_{u,1})</td>
<td>(\phi_{u,2})</td>
<td>(k)</td>
<td>(\phi_{u,4})</td>
</tr>
</tbody>
</table>

[Majority voting] For each \(w_j\), vote the similarity score \(\text{sim}(u, w_j)\) for \(\phi_{j3}\)-th parameter.
Experiment

• Dataset
  • ETL4: a set of Japanese Hiragana Characters
    • 48 characters, 120 writers, 48*120=5760 images
  • ETL5: a set of Japanese Katakana Characters
    • 48 characters, 208 writers, 48*204=9984 images

• Setting
  • Randomly select 3 writers as “target user”, i.e., $u$, and regard the remaining writers as “other writers”, i.e., $\{w_j\}$.
  • Generate the following 5 characters, regarding the other 43 characters as seed.
    • Hiragana: あ (a), し (shi), た (ta), は (ha), れ (re)
    • Katakana: ア (a), シ (shi), タ (ta), ハ (ha), レ (re)
  • Encoder & Decoder: Variational Autoencoder

• Compared methods
  • BestFit: using all of the 48 characters as seed (complete seed character set)
  • UserCF
  • ItemCF: item-based collaborative filtering
  • HybrCF: the method combining UserCF and ItemCF
  • Random: randomly selecting a parameter from the pool
### Result *(Hiragana in ETL4, \(K=40\))*

<table>
<thead>
<tr>
<th>Original</th>
<th>あしたはれ</th>
<th>あしたはれ</th>
<th>あしたはれ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BestFit</strong></td>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td><strong>UserCF</strong></td>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td><strong>HybrCF</strong></td>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td><strong>ItemCF</strong></td>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td><strong>Random</strong></td>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
</tbody>
</table>

- **BestFit** can generate HCCs quite similar with *Original*.
- **UserCF** and **HybrCF** can also generate good HCCs.
- The performance of **ItemCF** is almost same with that of *Random*.
  - Co-occurrence probability \(L\) becomes statistically unreliable with large \(K\).
**Result (Katakana in ETL5, K=40)**

<table>
<thead>
<tr>
<th>Original</th>
<th>あらたちヘレ</th>
<th>あらたちヘレ</th>
<th>あらたちヘレ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BestFit</strong></td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
</tr>
<tr>
<td><strong>UserCF</strong></td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
</tr>
<tr>
<td><strong>HybrCF</strong></td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
</tr>
<tr>
<td><strong>ItemCF</strong></td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
</tr>
<tr>
<td><strong>Random</strong></td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
<td>あらたちヘレ</td>
</tr>
</tbody>
</table>

- Similar result was obtained.
  - HCCs generated by *BestFit* are quite similar with Original.
  - *UserCF* and *HybrCF* also generate good HCCs
  - *ItemCF* did not work well.

- *UserCF* is more suitable to the HCC generation task.

**Average distance between feature of original image and that of generated HCCs**

- $K$: num. of clusters (size of parameter pool)
### Several Examples of HCC (ETL4)

<table>
<thead>
<tr>
<th>Original</th>
<th>BestFit</th>
<th>UserCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
<tr>
<td>あしたはれ</td>
<td>あしたはれ</td>
<td>あしたはれ</td>
</tr>
</tbody>
</table>

- HCCs generated by *BestFit* slightly differ from each other while keeping the similar shape with *original*.
- This is also the case with *UserCF*. 

> within-person variety
Several Examples of HCC (ETL5)

| Original | ア シ タ ハ レ | ア シ タ ハ レ | UserCF
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BestFit</td>
<td>ア シ ク ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ア シ タ ハ レ</td>
<td>ア シ タ ハ レ</td>
<td></td>
</tr>
</tbody>
</table>

- HCCs generated by *BestFit* slightly differ from each other while keeping the similar shape with *original*.
- This is also the case with *UserCF*.
Conclusion

• **Proposal:** A method for generating HCCs from a limited size of training data
  • The target writer only offers **at most one or zero image per character,** i.e., a **incomplete seed character set.**
  • To achieve **within-person variety,** the feature distribution of the target user’s handwriting is estimated for each character.

• **Idea:**
  • For seed characters: **BestFit strategy**
  • For non-seed characters: Collaborative Filtering (**UserCF**)  

• **Result:**
  • Examined the proposed method with a dataset of Japanese characters
  • **UserCF** can generate good HCCs with a certain level of within-person variety