Generating Handwriting via Decoupled Style Descriptors


Atsunobu Kotani

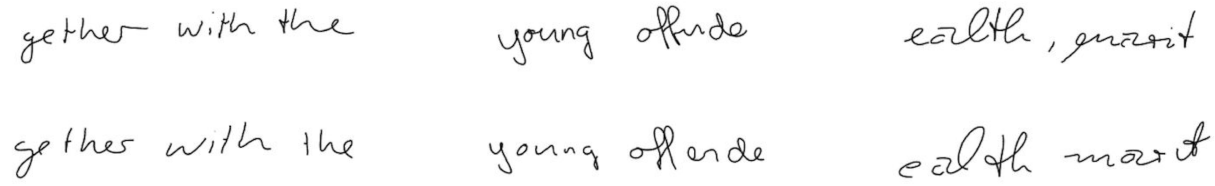

Stefanie Tellex


James Tompkin

BROWN
Computer Science

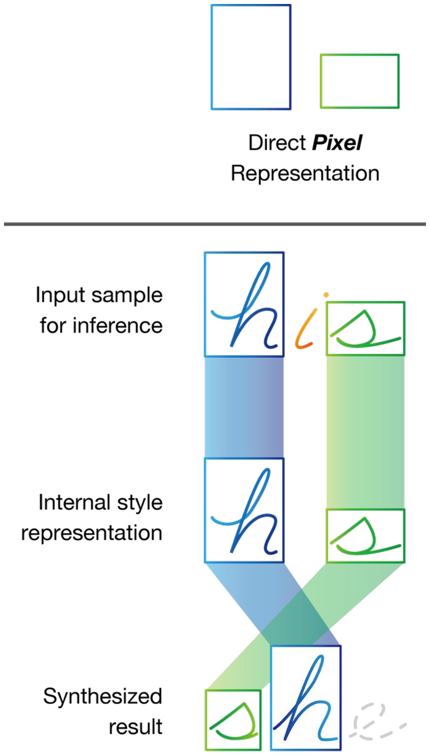

Target

Ours


gether with the
young offrde
ealth, guarit
gether with the
young offende
ealth marit

Handwriting generation:

- Pixel representation [2]

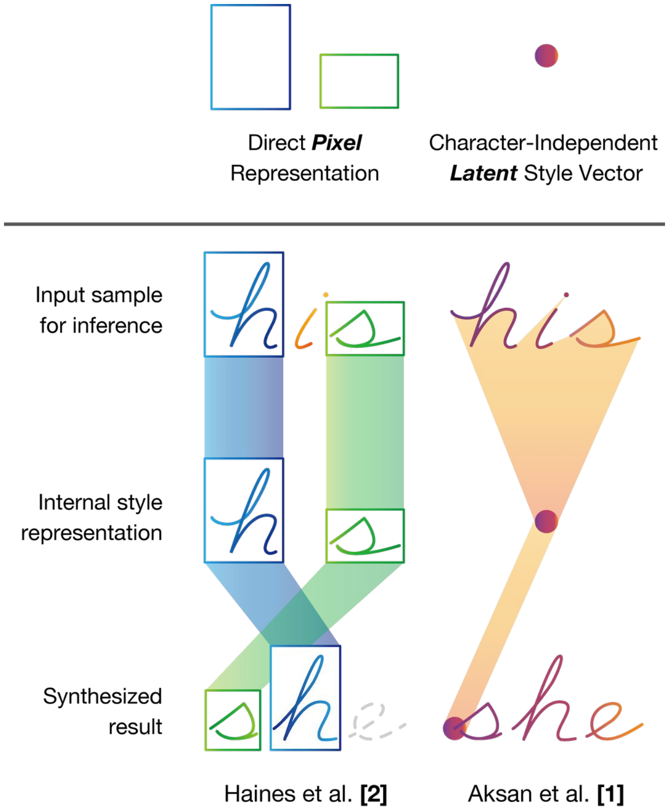


Haines et al. **[2]**

[1] Aksan, E., Pece, F., Hilliges, O.: DeepWriting: Making Digital Ink Editable via Deep Generative Modeling. SIGCHI (2018).
[2] Haines, T.S.F., Mac Aodha, O., Brostow, G.J.: My text in your handwriting. ACM Trans. Graph. 35(3) (May 2016).

Handwriting generation:

- Pixel representation [2]
- Learned by neural networks [1]



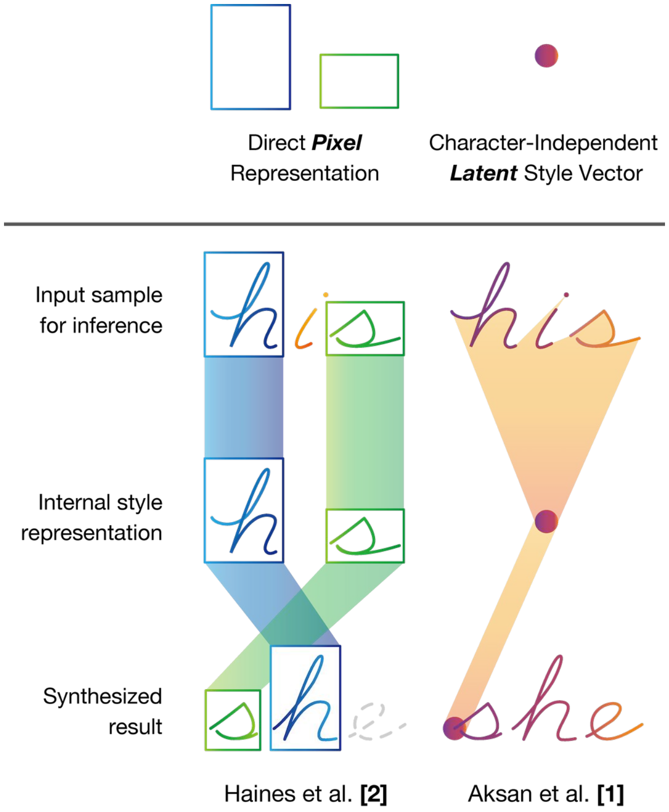
[1] Aksan, E., Pece, F., Hilliges, O.: DeepWriting: Making Digital Ink Editable via Deep Generative Modeling. SIGCHI (2018).
[2] Haines, T.S.F., Mac Aodha, O., Brostow, G.J.: My text in your handwriting. ACM Trans. Graph. 35(3) (May 2016).

Handwriting generation:

- Pixel representation [2]
- Learned by neural networks [1]

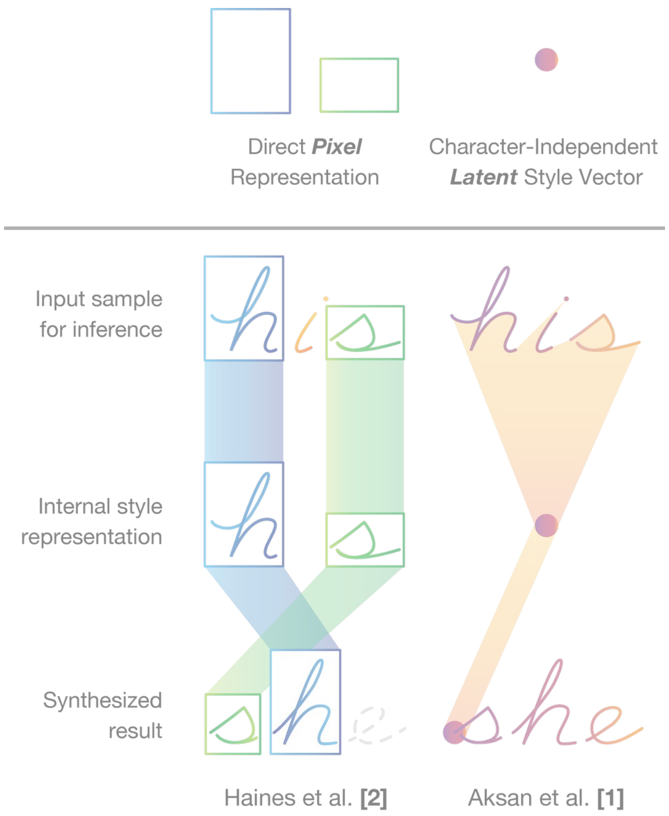
Limitations:

- Generating missing characters
- Generating fine details



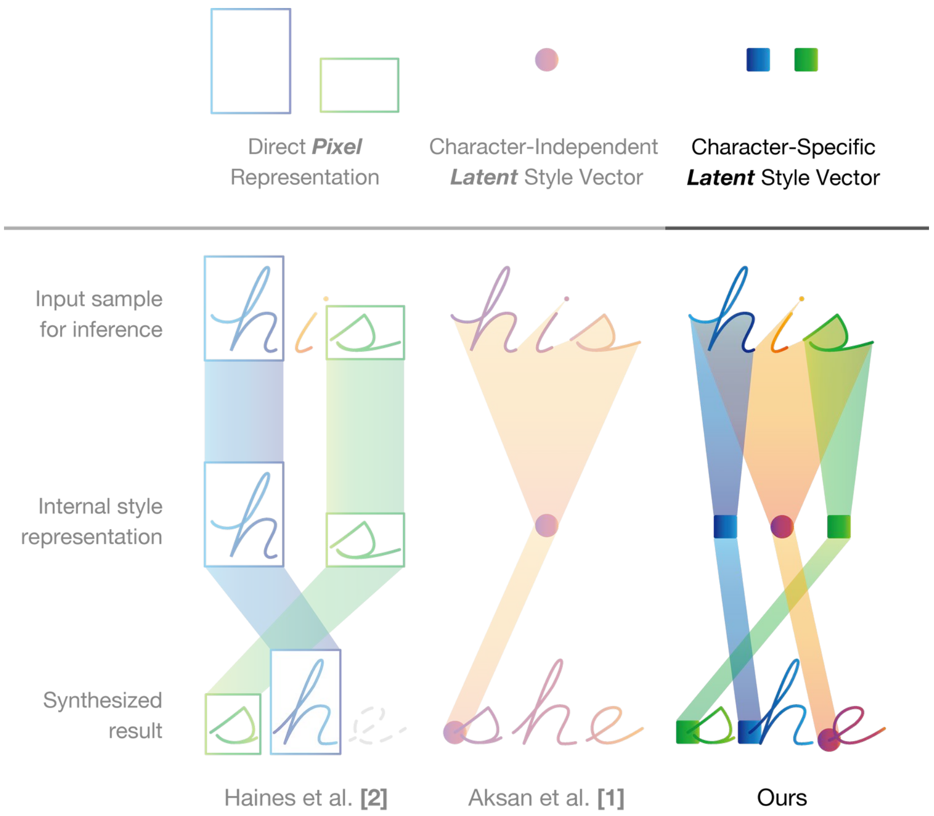
[1] Aksan, E., Pece, F., Hilliges, O.: DeepWriting: Making Digital Ink Editable via Deep Generative Modeling. SIGCHI (2018).
[2] Haines, T.S.F., Mac Aodha, O., Brostow, G.J.: My text in your handwriting. ACM Trans. Graph. 35(3) (May 2016).

Underlying problem:
No explicit separation of writer style from character style.



[1] Aksan, E., Pece, F., Hilliges, O.: DeepWriting: Making Digital Ink Editable via Deep Generative Modeling. SIGCHI (2018).
[2] Haines, T.S.F., Mac Aodha, O., Brostow, G.J.: My text in your handwriting. ACM Trans. Graph. 35(3) (May 2016).

We learn to decouple writer and character style into specific vectors.

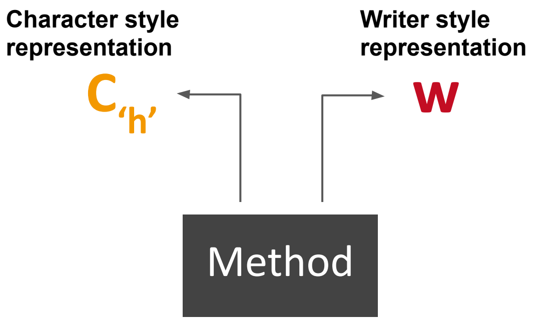


[1] Aksan, E., Pece, F., Hilliges, O.: DeepWriting: Making Digital Ink Editable via Deep Generative Modeling. SIGCHI (2018).
[2] Haines, T.S.F., Mac Aodha, O., Brostow, G.J.: My text in your handwriting. ACM Trans. Graph. 35(3) (May 2016).

Problem statement

Desired output data:

- *Writer-independent* character style representation $c_{\text{'h'}}$
- *Character-independent* writer style representation $\mathbf{w}$



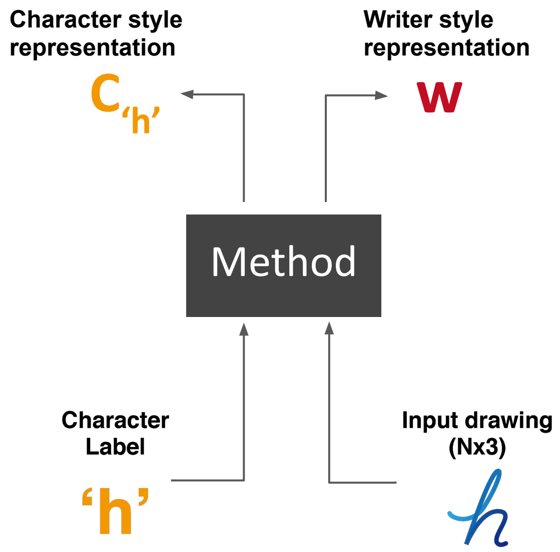
Problem statement

Desired output data:

- *Writer-independent* character style representation $C_{\text{‘h’}}$
- *Character-independent* writer style representation $\mathbf{w}$

Input data:

- Strokes as point sequences (x, y, t)
- Character labels as one-hot vectors



Decoupled Style Descriptors (DSD)

We learn a ***linear*** relationship between

- Writer-DSD **w** and

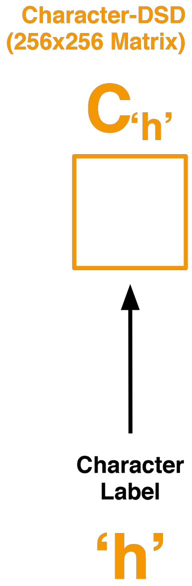
- Character-DSD $\mathbf{C}_{\text{‘h’}}$

Decoupled Style Descriptors (DSD)

We learn a ***linear*** relationship between

- Writer-DSD **w** and

- Character-DSD $\mathbf{C}_{\text{‘h’}}$

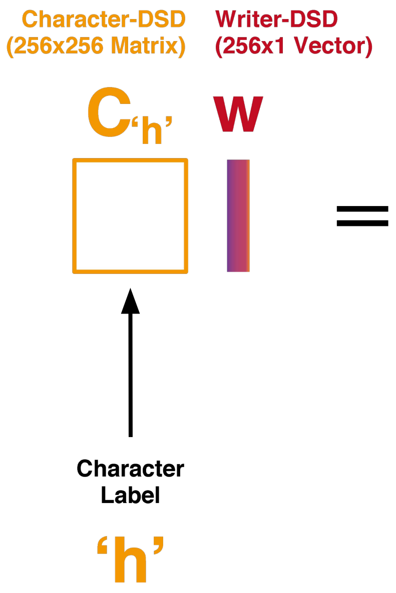


Decoupled Style Descriptors (DSD)

We learn a ***linear*** relationship between

- Writer-DSD **w** and

- Character-DSD $\mathbf{C}_{\text{'h'}}$

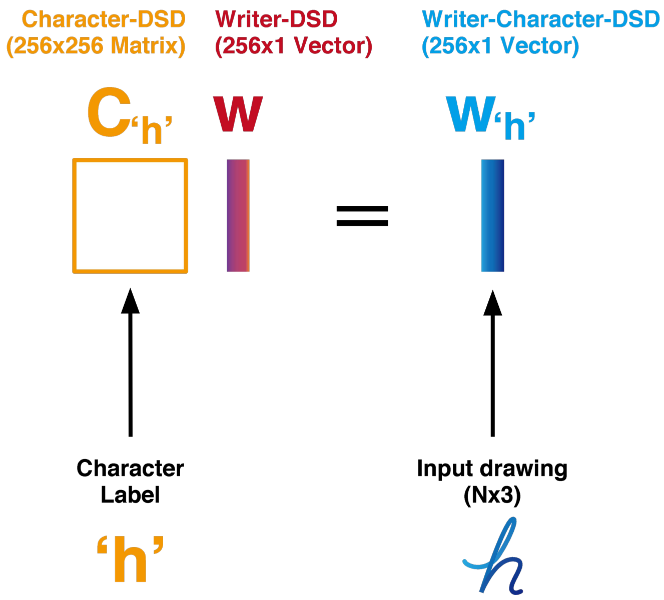


Decoupled Style Descriptors (DSD)

We learn a ***linear*** relationship between

- Writer-DSD **w** and

- Character-DSD $\mathbf{C}_{\text{‘h’}}$



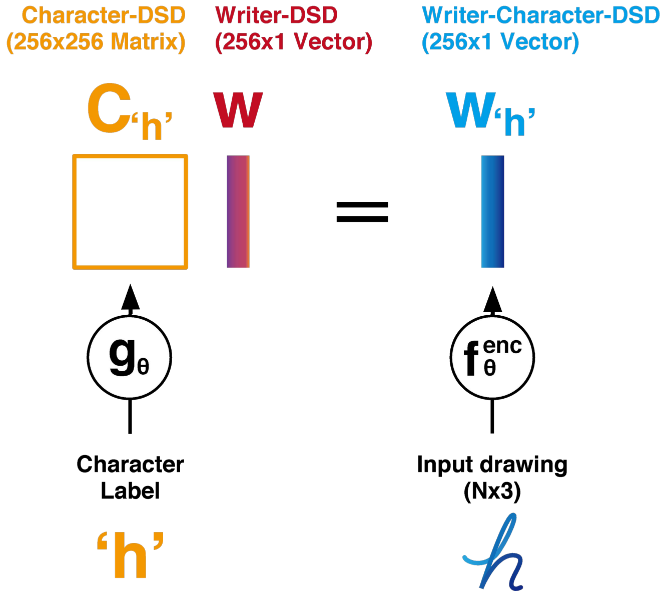
Decoupled Style Descriptors (DSD)

We learn a ***linear*** relationship between

- Writer-DSD **w** and

- Character-DSD $\mathbf{C}_{\text{'h'}}$

through learned LSTM encoders $\mathbf{g}_\theta$ and $\mathbf{f}_\theta$.



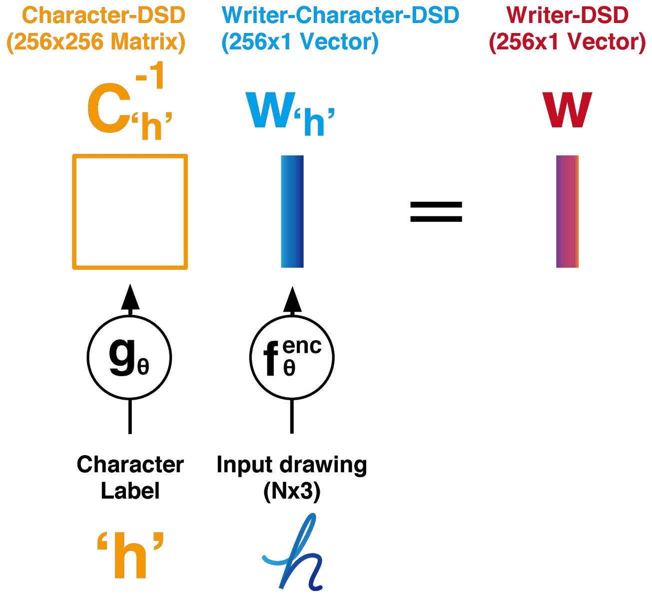
Decoupled Style Descriptors (DSD)

We learn a ***linear*** relationship between

- Writer-DSD **w** and
- Character-DSD $\mathbf{C}_{\text{'h'}}$

through learned LSTM encoders $\mathbf{g}_\theta$ and $\mathbf{f}_\theta$.

Simply invert $\mathbf{C}_{\text{'h'}}$ to recover **w** from $\mathbf{w}_{\text{'h'}}$.



Decoupled Style Descriptors (DSD)

We learn a ***linear*** relationship between

- Writer-DSD **w** and

- Character-DSD $\mathbf{C}_{\text{‘h’}}$

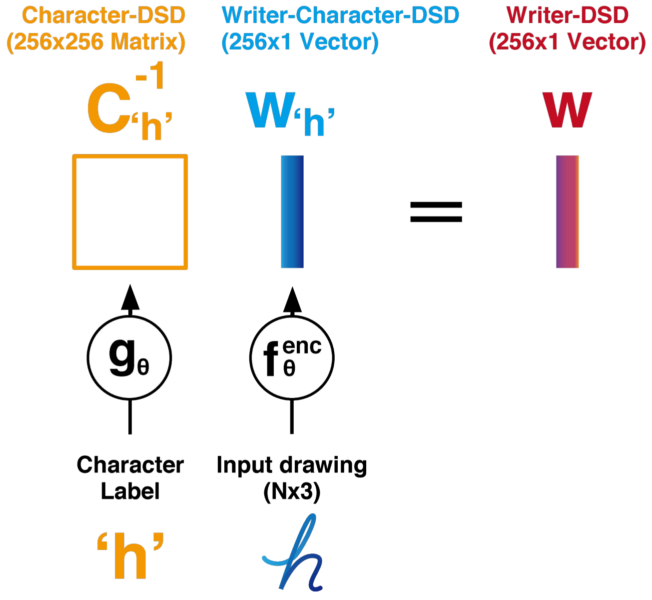
through learned LSTM encoders $\mathbf{g}_\theta$ and $\mathbf{f}_\theta$.

Simply invert $\mathbf{C}_{\text{‘h’}}$ to recover **w** from $\mathbf{w}_{\text{‘h’}}$.

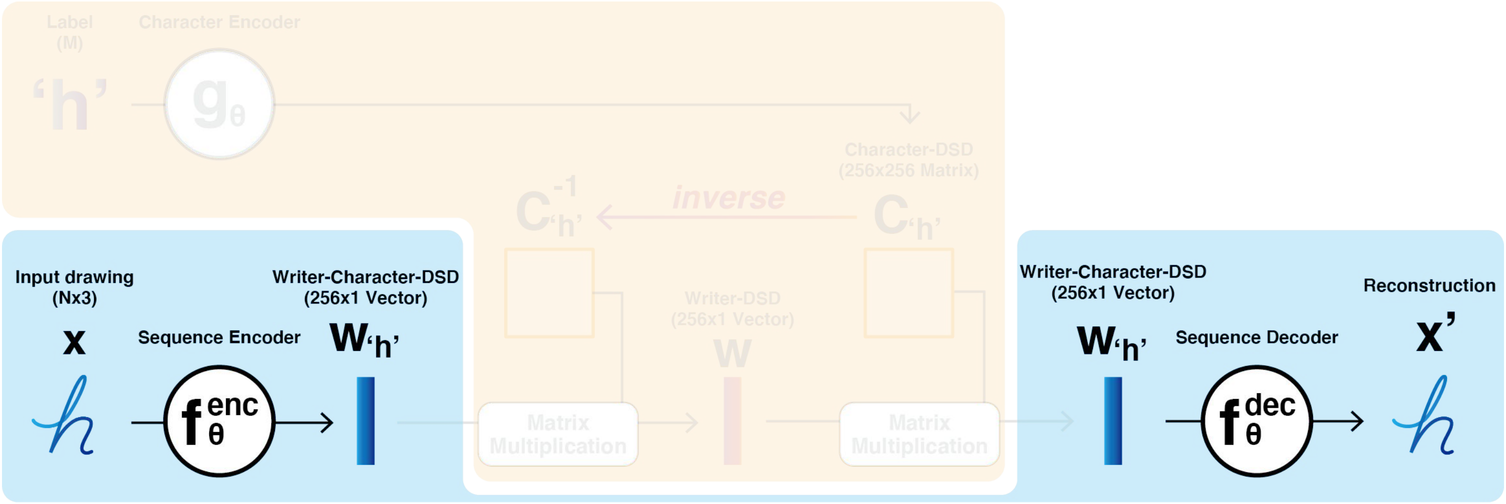
Retains ***more fine detail***.

Allows ***few-shot learning*** for new characters,

and ***writer identification***.



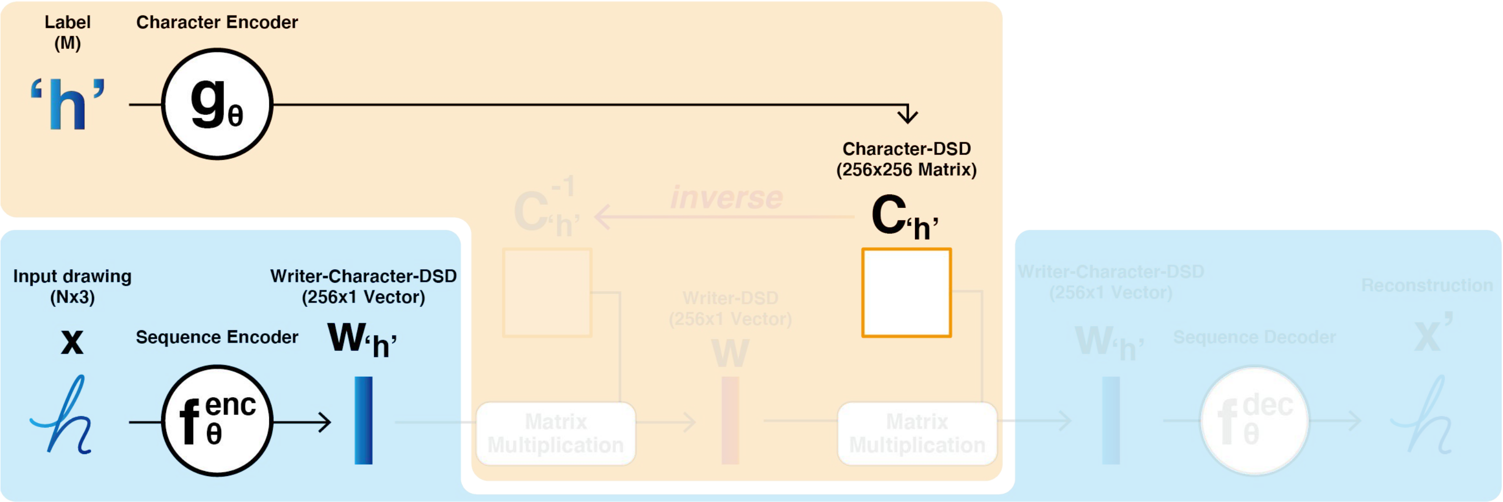
Decoupled Style Descriptors (DSD)



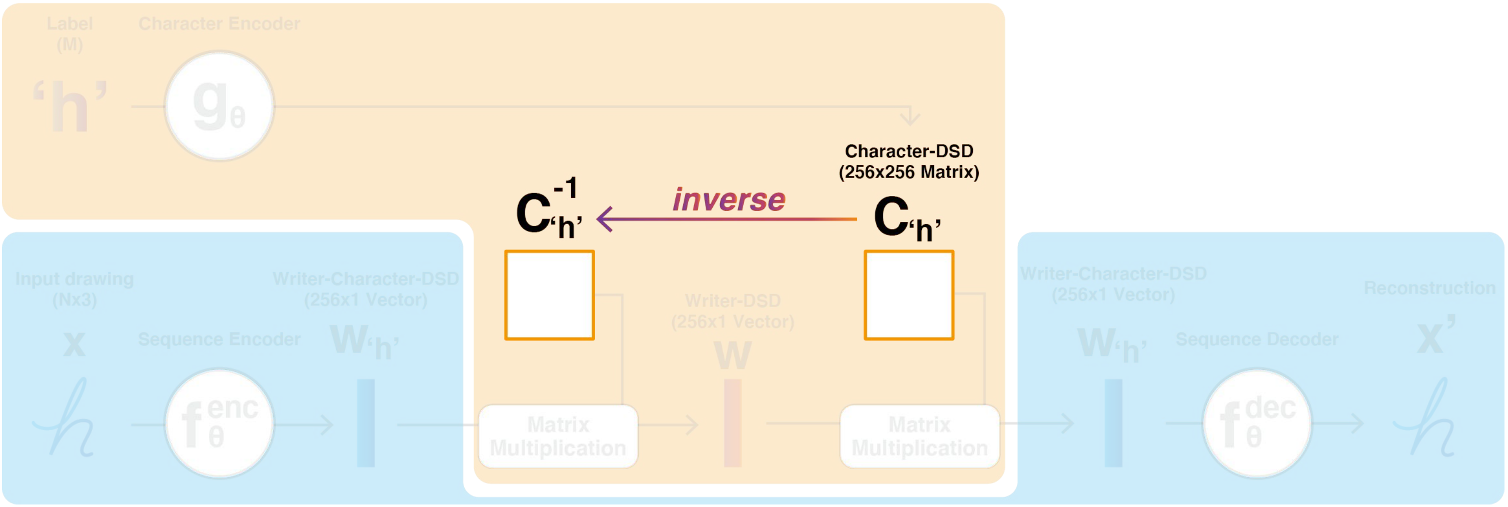
At the core, we have LSTM-based autoencoder, similar to the work by Graves [3].

[3] Graves, A.: Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850 (2013)

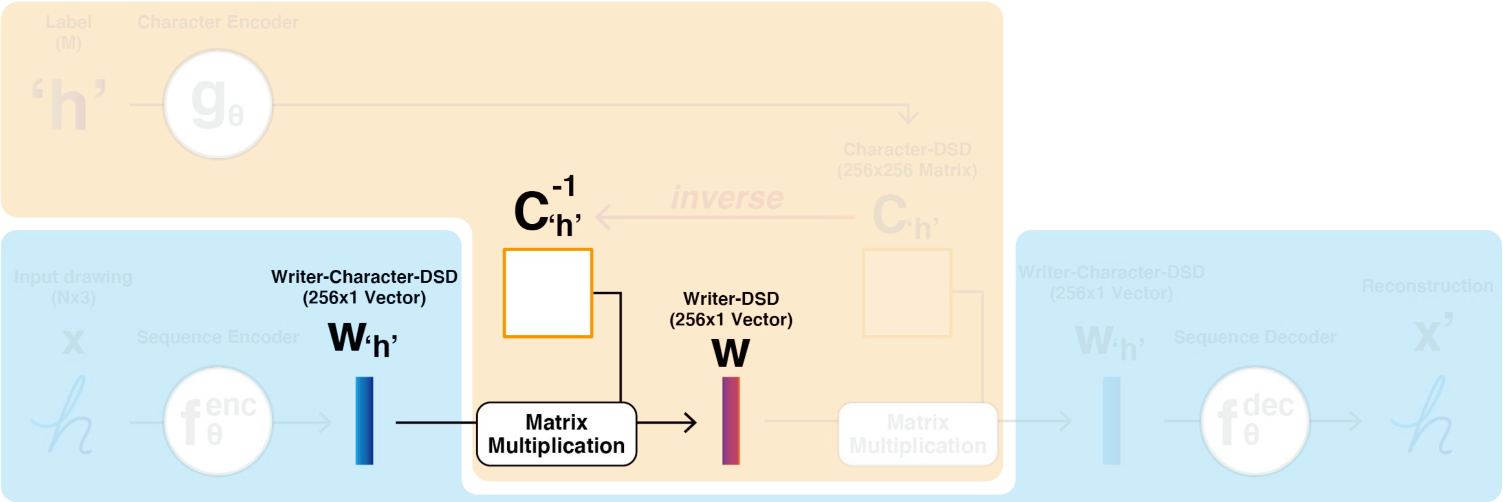
Decoupled Style Descriptors (DSD)



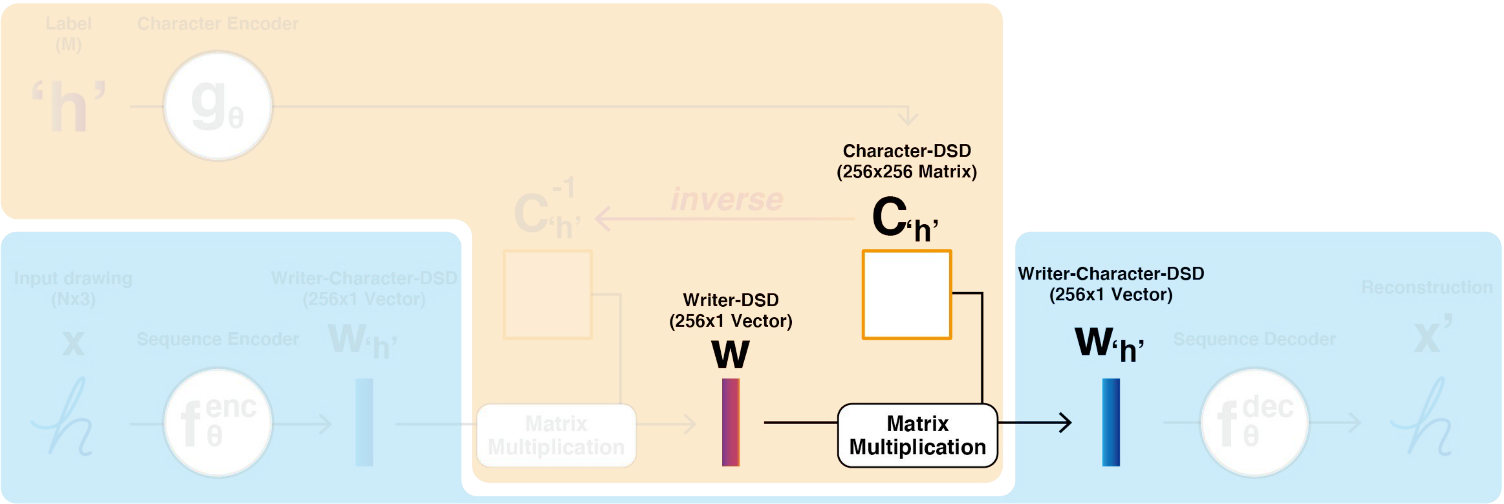
Decoupled Style Descriptors (DSD)



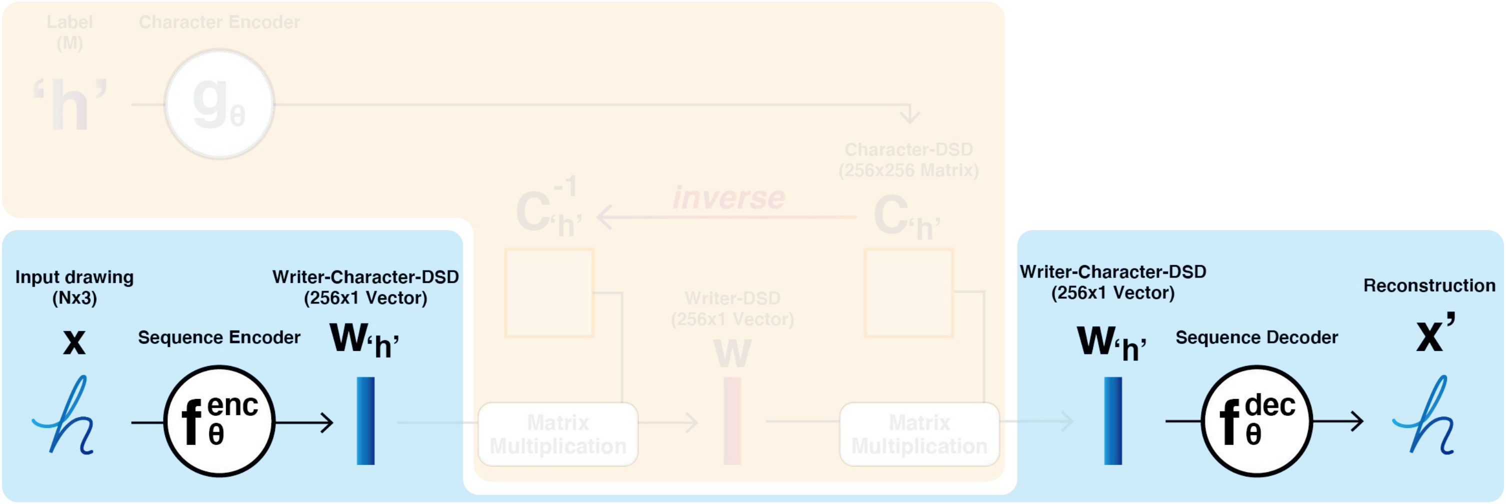
Decoupled Style Descriptors (DSD)



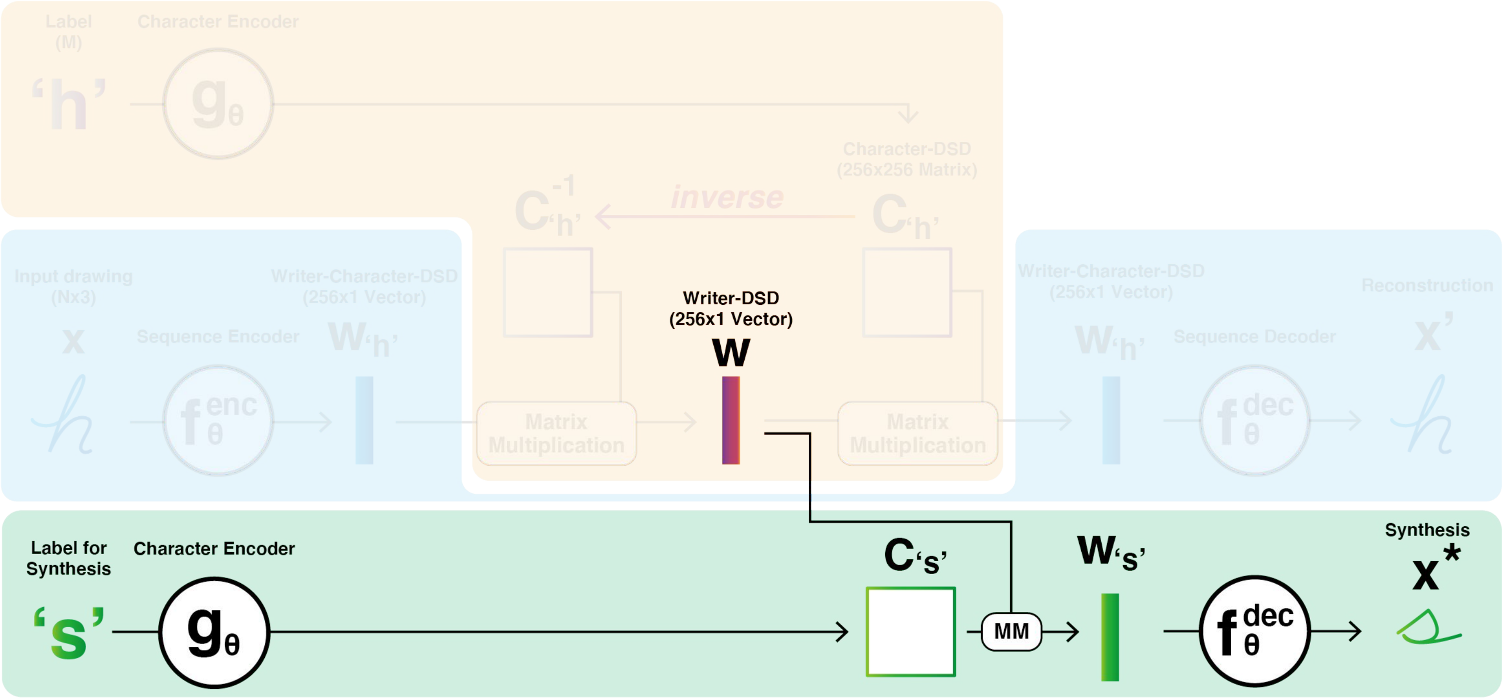
Decoupled Style Descriptors (DSD)



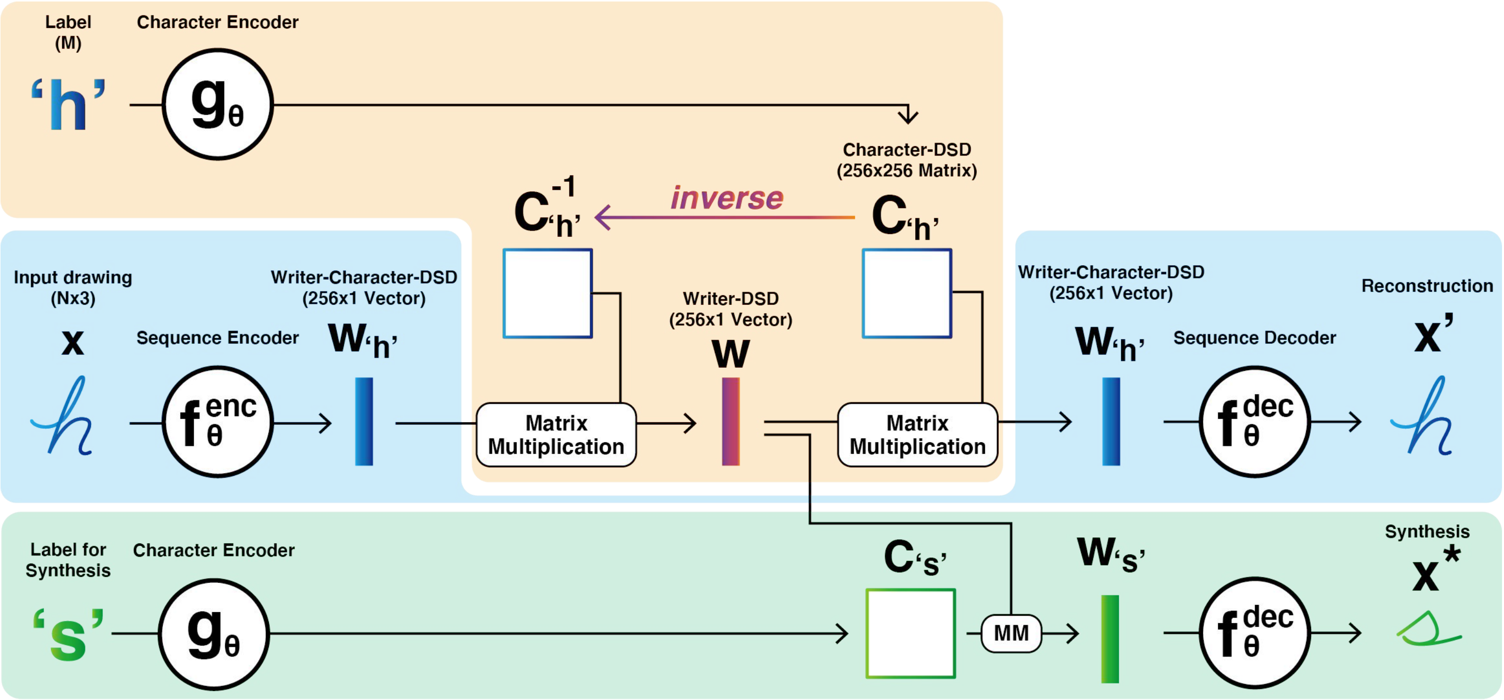
Decoupled Style Descriptors (DSD)



Decoupled Style Descriptors (DSD)



Decoupled Style Descriptors (DSD)



Subsequences rather than single characters

Writing is complex:

- Cursive
- Character pairs (ligatures)
- Delayed strokes (f, t, i, j)

Subsequences rather than single characters

Writing is complex:

- Cursive
- Character pairs (ligatures)
- Delayed strokes (f, t, i, j)

Our approach actually represents latent space of *all subsequences of characters.*

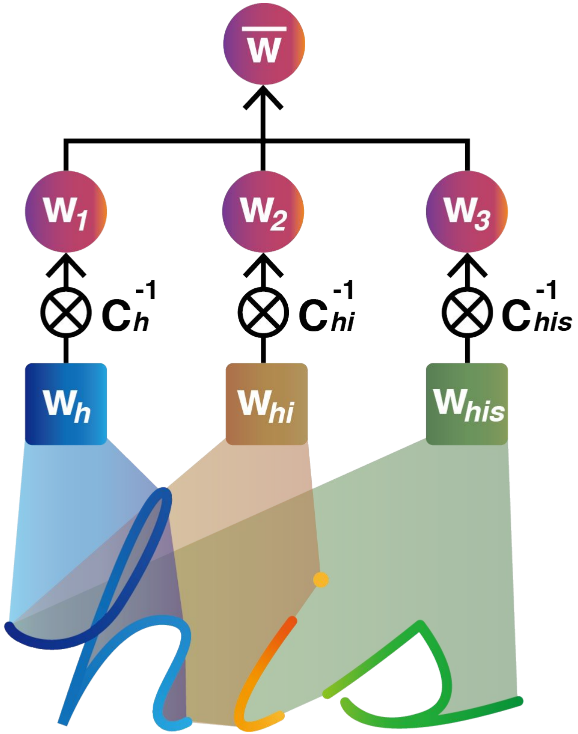
Word 'his' has representations $\mathbf{C}_{\text{'h'}}$, $\mathbf{C}_{\text{'hi'}}$, and $\mathbf{C}_{\text{'his'}}$.

and $\mathbf{w}_{\text{'h'}}$, $\mathbf{w}_{\text{'hi'}}$, and $\mathbf{w}_{\text{'his'}}$.

Recovering Writer-DSD **w** from handwriting samples

Take mean **w** over subsequences:

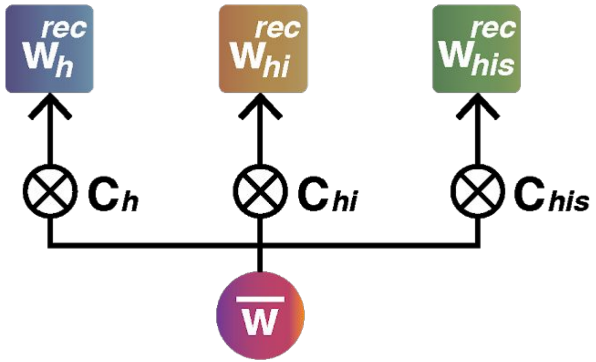
$$\overline{\mathbf{w}} = \frac{1}{M}\sum_{t=1}^{M}\mathbf{C}_{c_t}^{-1}\mathbf{w}_{c_t}$$



Generating handwriting using a *global* Writer-DSD **w**

Given a target word 'his':

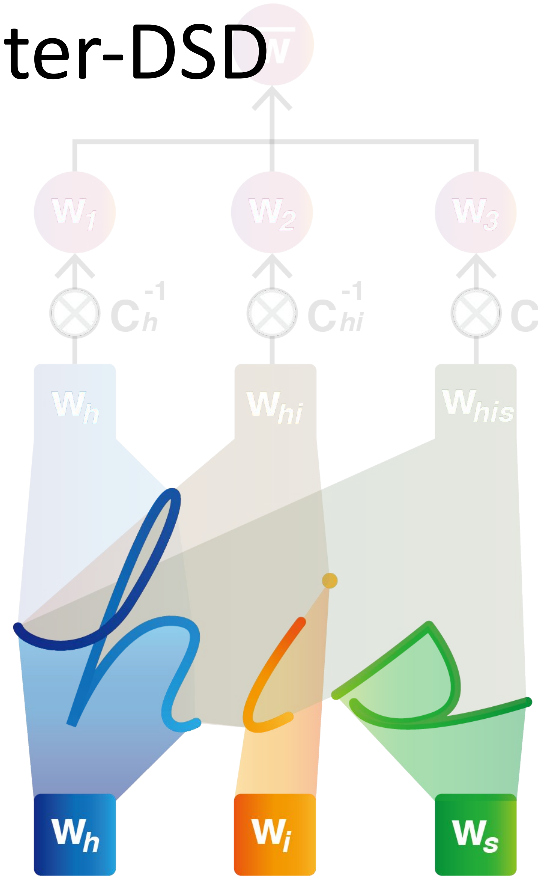
- Predict $\mathbf{C}_{\text{'h'}}$, $\mathbf{C}_{\text{'hi'}}$, and $\mathbf{C}_{\text{'his'}}$
- Multiply by **w** to create $\mathbf{w}_{\text{'h'}}$, $\mathbf{w}_{\text{'hi'}}$, and $\mathbf{w}_{\text{'his'}}$
- Decode ($\mathbf{w}_{\text{'h'}}, \mathbf{w}_{\text{'hi'}}, \mathbf{w}_{\text{'his'}}$) into stroke sequence



Single-character Writer-Character-DSD

There are relatively few single-character $\mathbf{w}_{\text{'h'}}$.

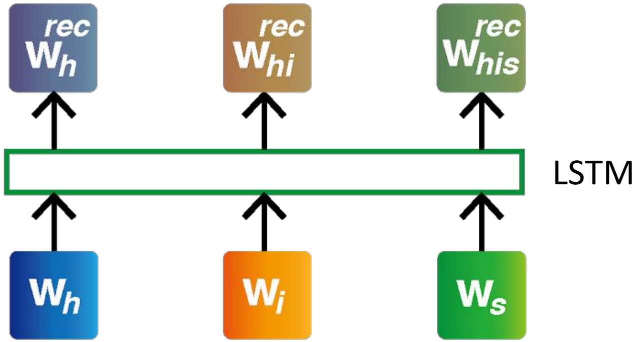
If we extract them from a writing sample, we can save them in a **database** and **sample** them during generation.



Generating handwriting using ***sampled*** Writer-Character-DSD

Retrieve relevant single-characters $\mathbf{w}_{\text{'h'}}$, $\mathbf{w}_{\text{'i'}}$, $\mathbf{w}_{\text{'s'}}$.

Restore temporal dependencies via LSTM.



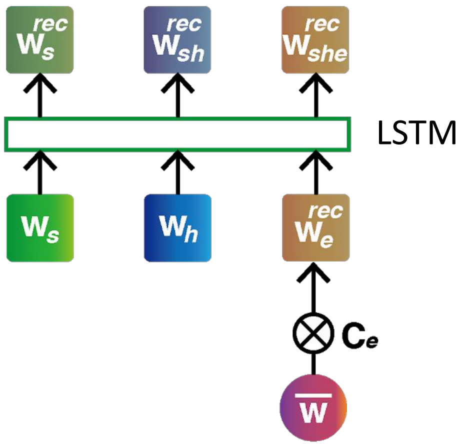
Cannot cope with any missing characters in reference handwriting samples.

Combined method with ***sampling***

Reference sample: ‘his’

Generation target: ‘she’

- Compute mean **w** from all substrings
- Predict $\mathbf{w}_{‘e’}$ with $\mathbf{C}_{‘e’}$ and the mean **w**
- Extract single-character Writer-Character-DSDs
- *Restore temporal dependencies with LSTM.*

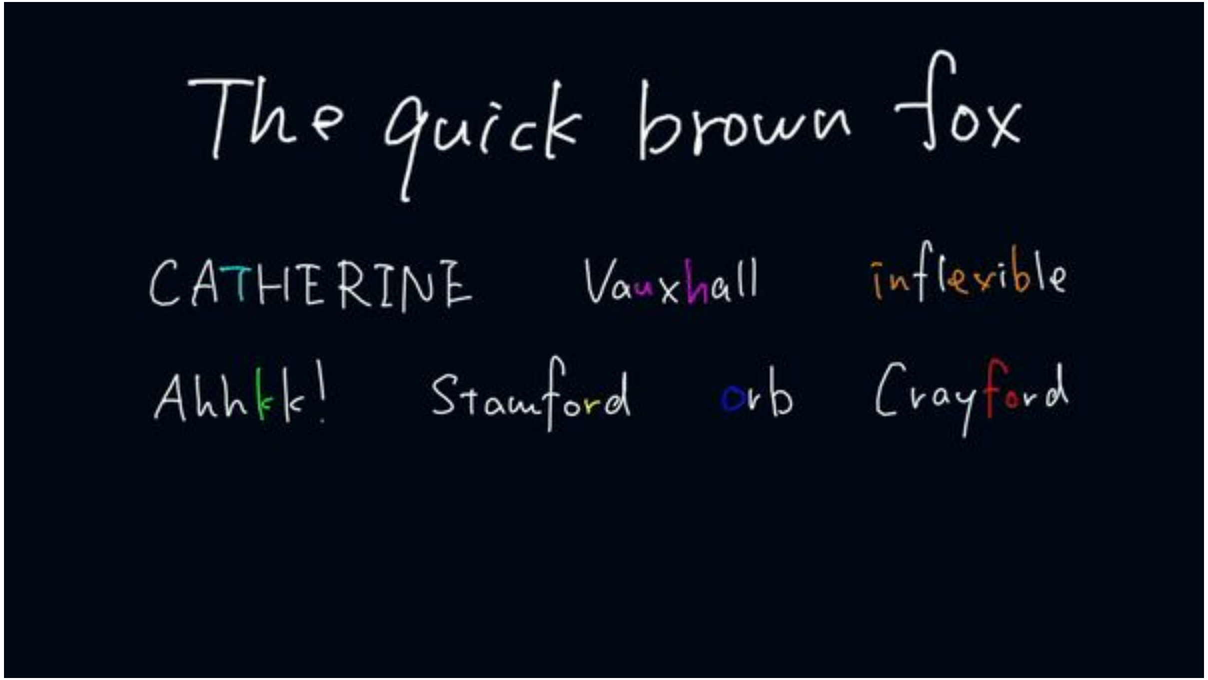


Generated Results

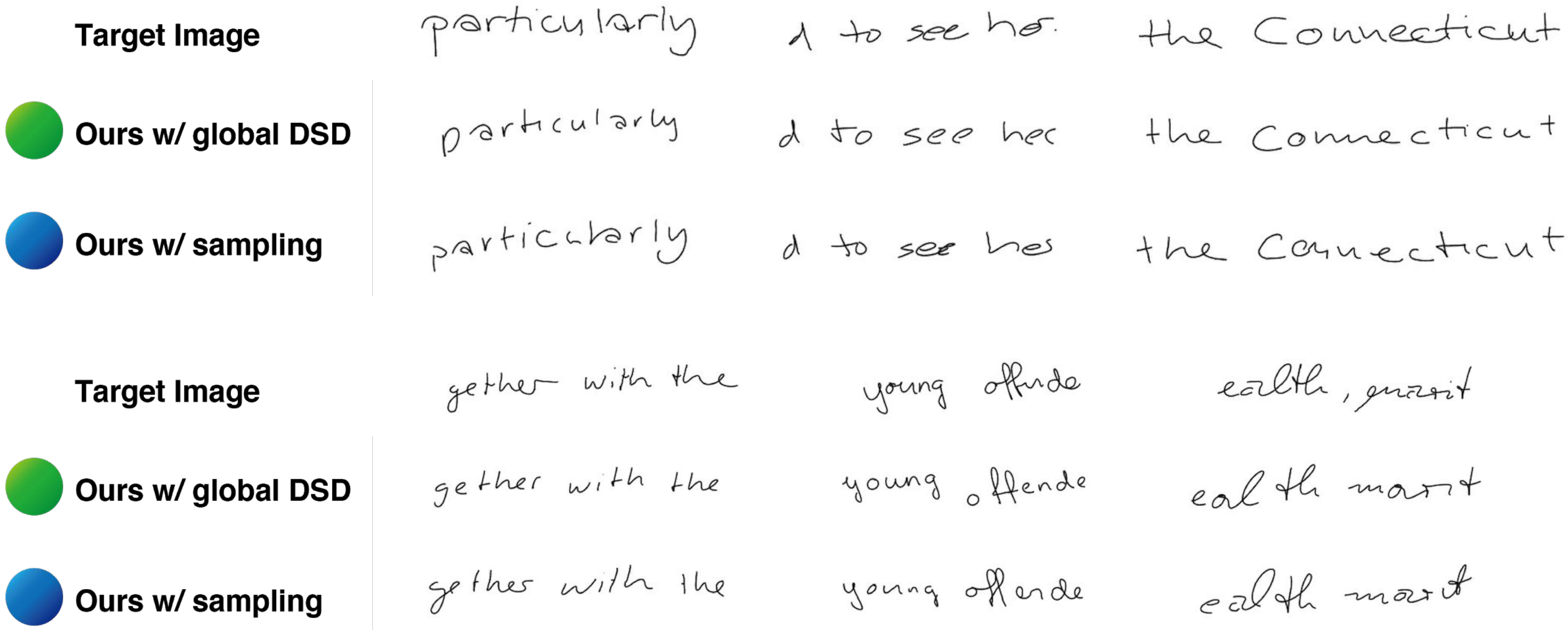
Colored characters match between provided writing samples and desired output.

These are generated from retrieved $\mathbf{w}_{\text{'h'}}$

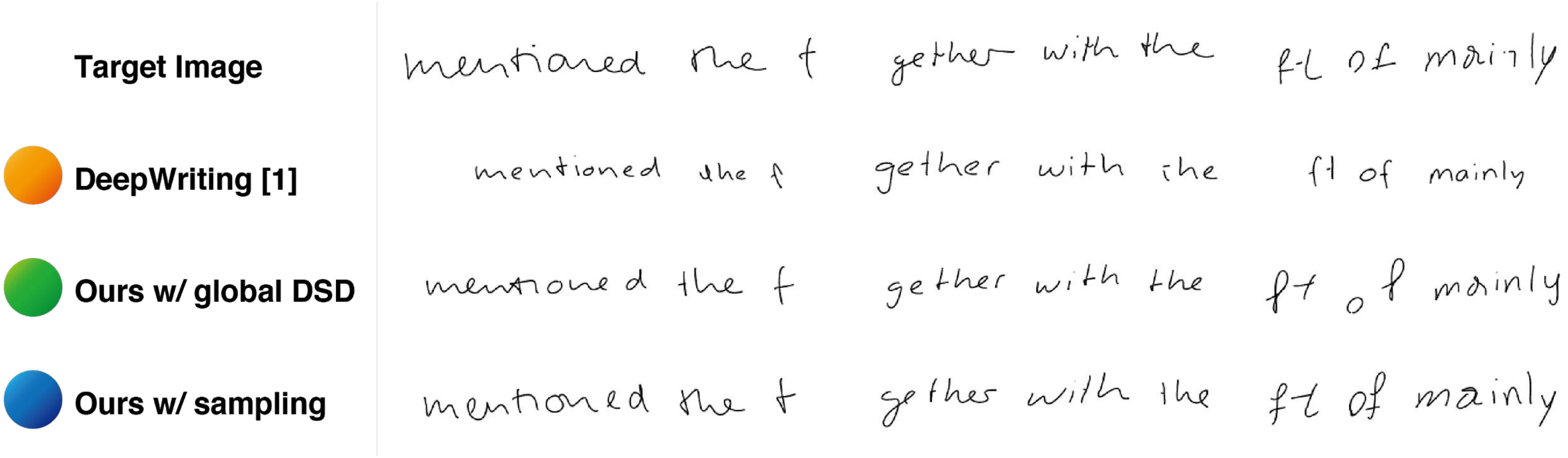
Missing characters are generated from global $\mathbf{w}$



Generated Results



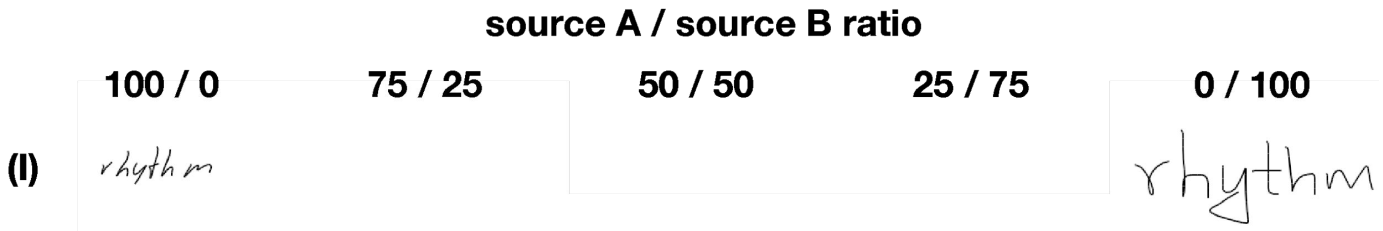
Generated Results - Comparison



[1] Aksan, E., Pece, F., Hilliges, O.: DeepWriting: Making Digital Ink Editable via Deep Generative Modeling. SIGCHI (2018).

Interpolation

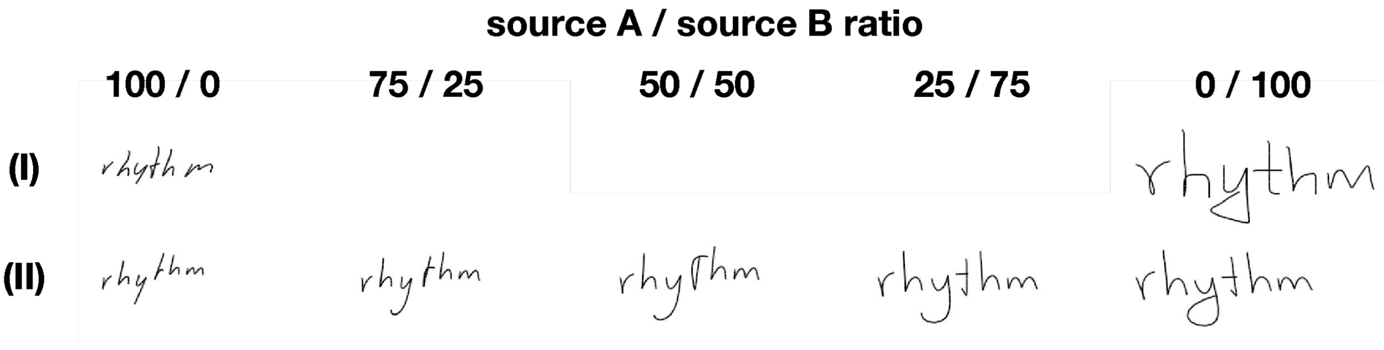
True samples



Interpolation

True samples

Writer-DSD level

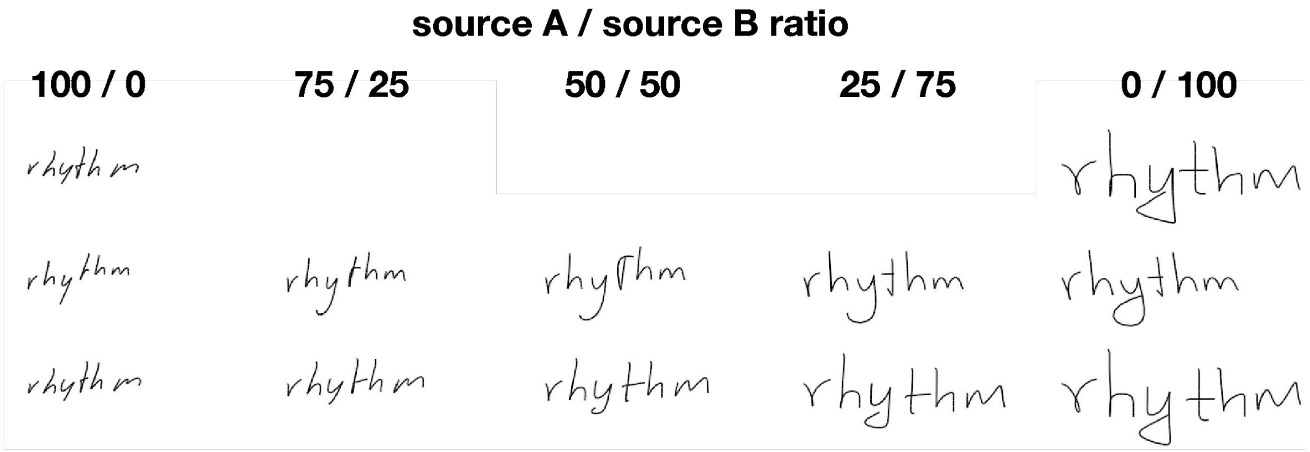


Interpolation

True samples (I)

Writer-DSD level (II)

Writer-Character-DSD level (III)



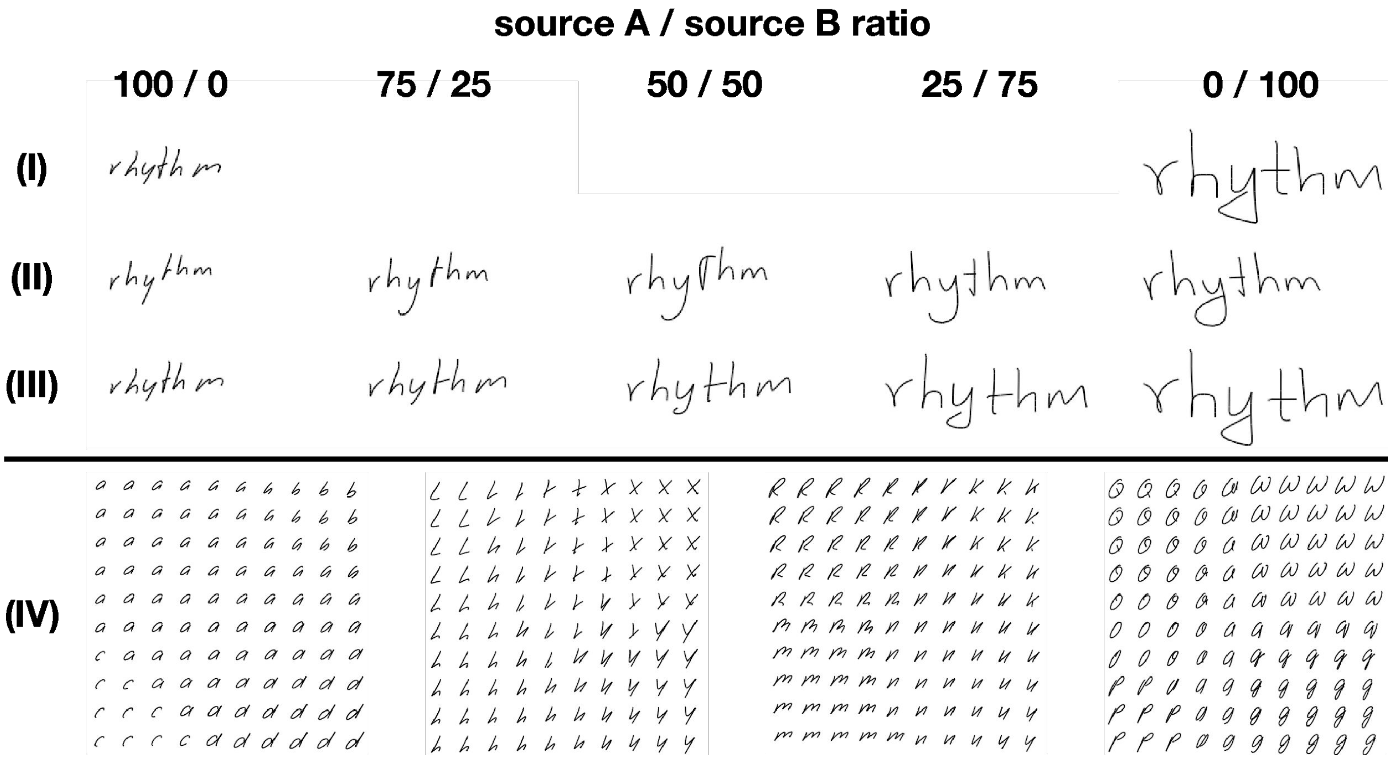
Interpolation

True samples

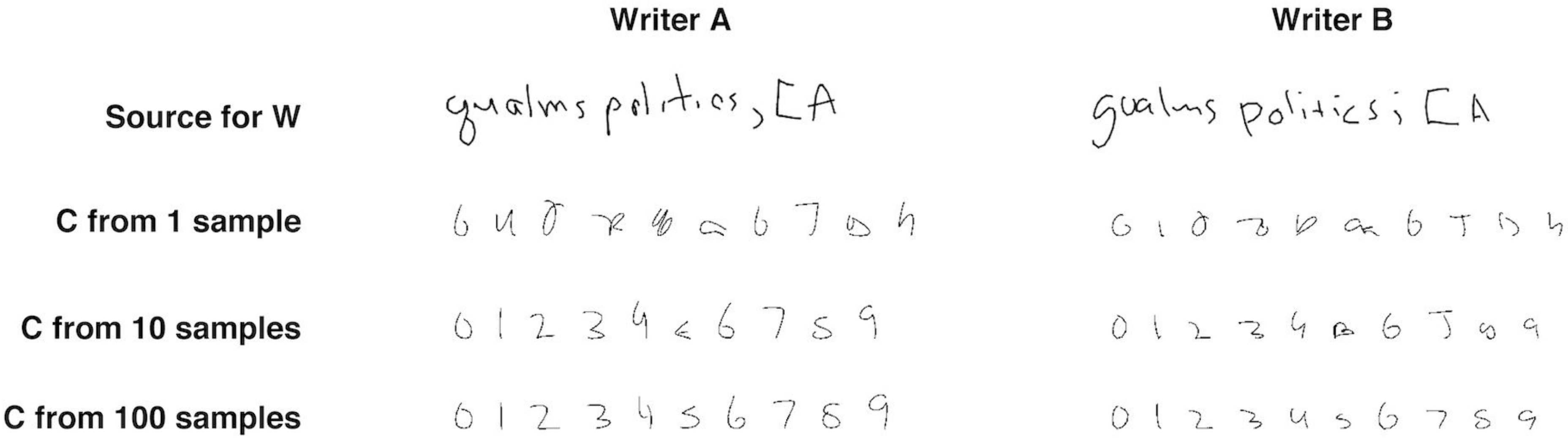
Writer-DSD level

Writer-Character-DSD level

Character-DSD level



Few-shot learning of New Characters



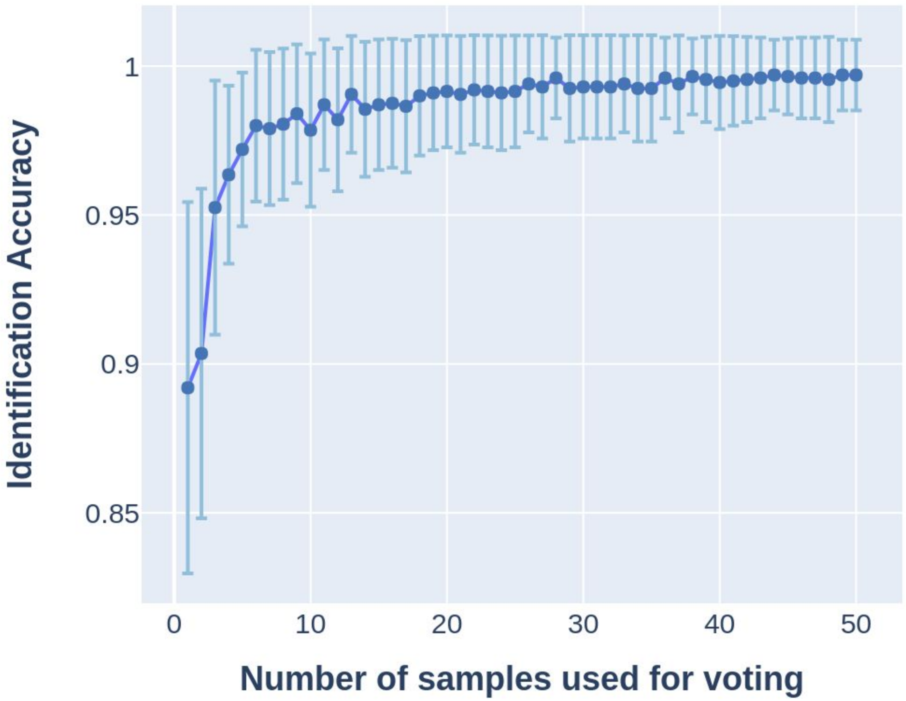
Writer Identification

In a 20-writer identification task,

our model achieved:

89.20% ± 6.23 for **1** word sample

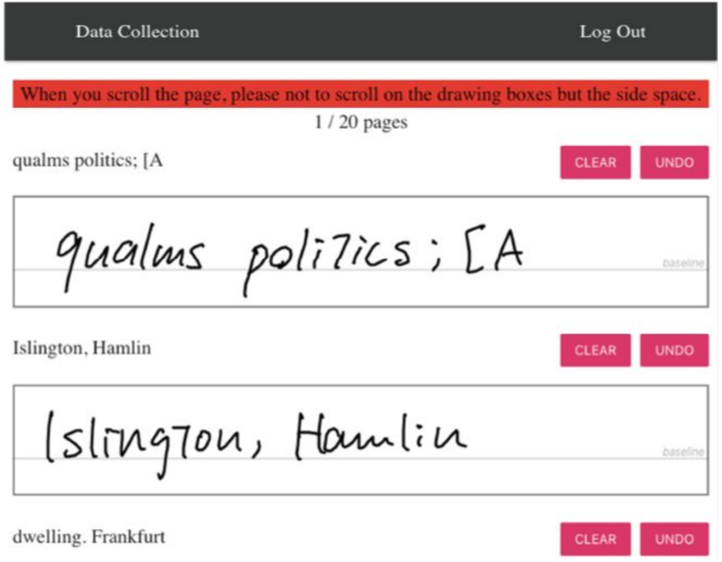
99.70% ± 1.18 for **50** word samples



BRUSH dataset

BRUSH dataset contains handwriting of:

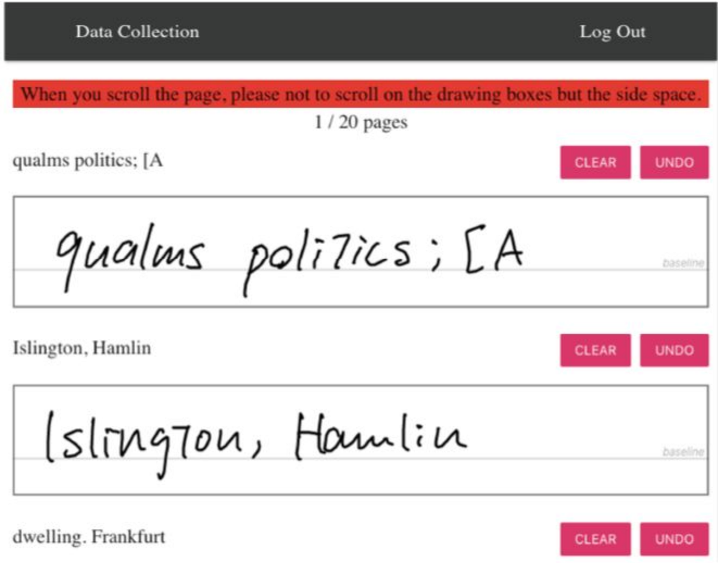
- 170 writers, 86 Latin alphabet characters
- 488 common words written by all writers
  - 99.5% coverage of two-character letter space
- + 3668 rarer words written across writers
  - 99.9% coverage in total



BRUSH dataset

BRUSH dataset contains handwriting of:

- 170 writers, 86 Latin alphabet characters
- 488 common words written by all writers
  - 99.5% coverage of two-character letter space
- + 3668 rarer words written across writers
  - 99.9% coverage in total



Common word samples & character-level labels

are not present in IAM dataset [4].

[4] Marti, U.V., Bunke, H.: The iam-database: an english sentence database for offline handwriting recognition. IJDAR (2002).

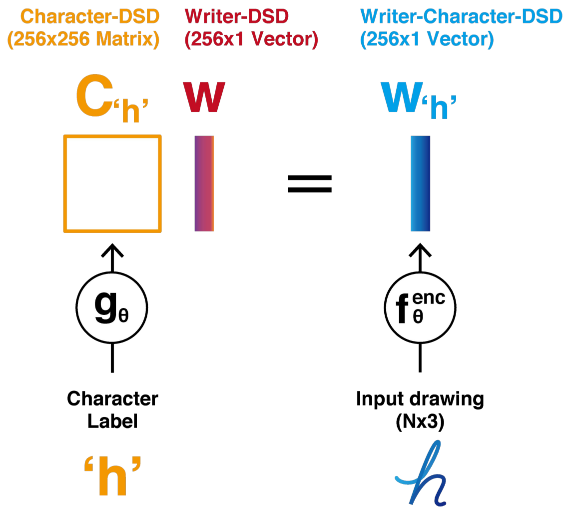
Generating Handwriting via Decoupled Style Descriptors

Atsunobu Kotani, Stefanie Tellex, James Tompkin

BROWN Computer Science

Contributions:

- DSDs decouple character style from writer style.
- Allow flexible handwriting generation.
- BRUSH: new dataset for online handwriting.



Code & dataset available at

http://dsd.cs.brown.edu

Generating Handwriting via Decoupled Style Descriptors

Atsunobu Kotani, Stefanie Tellex, James Tompkin

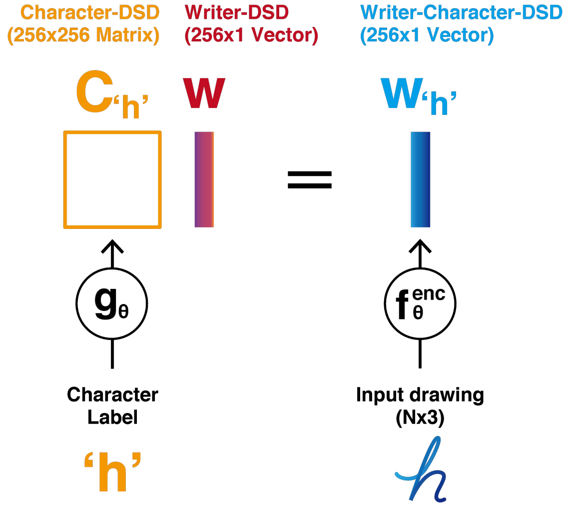
BROWN Computer Science

Contributions:

- DSDs decouple character style from writer style.
- Allow flexible handwriting generation.
- BRUSH: new dataset for online handwriting.

Potential Extension:

- Explicit decoupling may work better than implicit decoupling / disentanglement.
- DSDs could be applied to other sequential data (e.g. speech, motion capture data)



Code & dataset available at http://dsd.cs.brown.edu